

Lineær Algebra

Lars Hesselholt og Nathalie Wahl

3. udgave, august 2023

Forord

Denne bog er beregnet til et første kursus i lineær algebra, men vi har lagt vægt på at fremstille dette materiale på en sådan måde, at det implicit fremstår i sin naturlige generalitet. Vi har også lagt vægt på ikke-kommutativitet, og vi benytter således ikke, at ab og ba er ens, medmindre det er nødvendigt.

Vores hovedkilde som inspiration har været den uovertrufne præsentation af lineær algebra i N. Bourbaki, Algebra I og Algebra II. Vores gennemgang af determinanten er også stærkt inspireret af H. A. Nielsen, Lineær Algebra, Aarhus Universitet fra 1988. Endvidere er bogens layout og visse figurer overført fra Niels Vigand Pedersen, Lineær Algebra, Københavns Universitet, 2. udgave fra 2009, revideret af Morten Risager.

Vi takker Anna Schøtt-Edholm og Katja Thorseth for nøje gennemlæsning bogen.

Vi takker endvidere Fabien Pazuki for mange gode bidrag til denne nye udgave, samt Louise Bøtchier Meyer og Søren Ladegaard Kristensen for nyttige feedback.

København, august 2023,

Lars Hesselholt og Nathalie Wahl

Hvordan man læser denne bog

Matematik adskiller sig fra andre naturvidenskaber ved, at matematiske begreber har præcise definitioner, og at matematiske udsagn har præcise beviser. Alle matematiske begreber er opbygget fra det grundlæggende begreb “mængde”, der som det eneste er udefineret, og alle matematiske beviser fremkommer i princippet ved at kombinere de grundlæggende udsagn om mængder, der er beskrevet i Zermelo-Fraenkel aksiomerne.¹

Det formelle grundlag betyder dog ikke, at ideer, der udtrykkes i matematik, er løsrevet fra resten af verden. Snarere er matematik i denne sammenhæng ligesom musik i den forstand, at man skal lære at spille et instrument, før man kan udtrykke sine ideer. Det er også afgørende, at man lærer definitioner helt rigtigt og ikke kun nogenlunde rigtigt, ligesom det i musik er vigtigt, at man rammer de rigtige toner.

At læse en matematisk tekst som denne er derfor en aktiv proces, hvor man typisk ikke forstår nye begreber fuldt ud, den første gang man læser dem, men må gå tilbage og genlæse definitioner og sætninger samtidigt med, at man læser eksempler og laver opgaver eller browser i Wikipedia.² Det kan også tit være nyttigt at tænke på ekstreme tilfælde, såsom “Hvad siger denne definition, hvis A er den tomme matrix?”

Skal man også læse og forstå beviser? Med tiden, ja. Beviser er matematikkens svar på eksperimenter i den forstand, at beviser benyttes til at afgøre om et udsagn gælder. Det er ikke så let at sige præcist, hvad et formelt bevis er,³ men uformelt består et bevis i at angive, hvordan et udsagn, som man ønsker at bevise, logisk følger fra en række udsagn, som antages at gælde, ved hjælp af logiske operationer. At finde et bevis for et udsagn er dog ikke nogen formel proces og kræver typisk en dyb forståelse af det udsagn, man ønsker at vise. Omvendt forstår man typisk også et udsagn langt bedre, hvis man forstår dets bevis (eller beviser),⁴ og da beviser er en matematikers eneste værktøjer, er det vigtigt at lære sig disse.

Bogen kan læses i flere omgange og på flere niveauer. De fleste begreber i lineær algebra har således mening i meget stor generalitet, ligesom de fleste sætninger gælder meget generelt. Vi tillader derfor skalarer, der er mere generelle end reelle tal og komplekse tal, da de første fem kapitler ikke anvender de mere specielle (arkimediske) egenskaber ved de reelle og komplekse tal. I første omgang er læseren dog mere end velkommen til at forestille sig, at skalarer enten er reelle tal eller komplekse tal.

¹ For eksempel lyder det første aksiom “ $\forall x \forall y (\forall z (z \in x \Leftrightarrow z \in y) \Rightarrow x = y)$ ”, hvilket betyder, at hvis to mængder x og y har de samme elementer, så er $x = y$. Det andet aksiom lyder “ $\exists x \forall y (\neg y \in x)$ ”, hvilket betyder, at der findes en mængde, der ingen elementer har, og ifølge det første aksiom findes der højst én sådan mængde. Vi kalder denne mængde for den tomme mængde og betegner den med “ $\emptyset$ ”.

² Den engelske version af Wikipedia er ofte en fremragende kilde, især hvad eksempler angår.

³ Definitionen af et formelt bevis kan findes i Kenneth Kunen, *The foundations of mathematics, Studies in Logic* (London), 19, Mathematical Logic and Foundations. College Publications, London, 2009.

⁴ Sætning 4.4.14 er eksempel herpå.

Indhold

0	Skalarer	1
0.1	Afbildninger	1
0.2	Legemer	3
0.3	Højre og venstre multiplikation	6
0.4	Opgaver	7
1	Lineære ligningssystemer	11
1.1	Lineære ligningssystemer	11
1.2	Rækkeoperationer og Gauss-elimination	13
1.3	Rang og løsningsmængde	31
1.4	Opgaver	34
2	Matricer og lineære afbildninger	39
2.1	Matricer	39
2.2	Vektorrummet $\mathbb{F}^m$	47
2.3	Lineære afbildninger	52
2.4	Invertible matricer	61
2.5	Operationsmatricer	68
2.6	Hermitiske former	80
2.7	Opgaver	87
3	Determinant	97
3.1	Determinant af 2×2 -matrix	97
3.2	Determinant for $n \times n$ -matricer	101
3.3	Triangulære matricer og beregning af determinanter	116
3.4	Determinant og invers matrix	120
3.5	Polynomier	122
3.6	Opgaver	129
4	Vektorrum	135
4.1	Vektorrum	135
4.2	Lineære afbildninger	141
4.3	Basis for et vektorrum	145
4.4	Matrixrepræsentation af lineære afbildninger	160
4.5	Kerne og billede	172
4.6	Affine underrum og kvotientrum	178

Indhold

4.7	Opgaver	184
5	Egenverdier og egenrum	199
5.1	Egenverdier og egenrum for kvadratiske matricer	199
5.2	Diagonaliserbare matricer	207
5.3	Egenverdier og egenrum for lineære endomorfier	216
5.4	Opgaver	221
6	Vektorrum med indre produkt	227
6.1	Indre produkt	227
6.2	Ortogonalitet	232
6.3	Lineære isometrier	242
6.4	Spektralsætningen	251
6.5	Klassifikation af hermitiske former	260
6.6	Opgaver	263
	Indeks	271

0 Skalarer

Hele vejen igennem bogen vil vi arbejde med en grundlæggende mængde af “tal”, som vi betegner $\mathbb{F}$. Vi siger, at elementer $a \in \mathbb{F}$ er *skalarer*. De to hovedeksempler, vi har i tankerne, er de reelle tal $\mathbb{F} = \mathbb{R}$ og de komplekse tal $\mathbb{F} = \mathbb{C}$. Vi specificerer nedenfor de aritmetiske strukturer “+” og “·” på mængden af skalarer, som vi skal gøre brug af. Disse aritmetiske strukturer er alle velkendte fra de reelle og de komplekse tal.

0.1 Afbildninger

En *afbildning* er en triple bestående af to mængder X og Y og en regel¹ f , der til ethvert element $x \in X$ tilordner et element $f(x) \in Y$. Vi skriver denne triple som

$$f: X \longrightarrow Y \quad \text{eller} \quad X \xrightarrow{f} Y$$

og kalder mængderne X og Y for afbildningens *domæne* og *codomæne*.

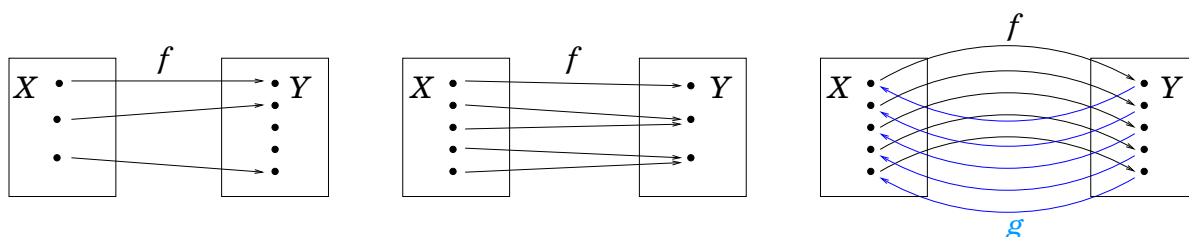
Eksempel 0.1.1 Afbildningerne $f: \mathbb{R} \rightarrow \mathbb{R}$ og $g: \mathbb{R} \rightarrow [0, \infty)$, hvor $f(x) = x^2$ og $g(x) = x^2$, er to forskellige afbildninger, da deres codomæner er to forskellige mængder.

Vi minder om, at en afbildning $f: X \rightarrow Y$ er *injektiv*, hvis der for alle $y \in Y$ findes højst ét $x \in X$, sådan at $f(x) = y$; at den er *surjektiv*, hvis der for alle $y \in Y$ findes mindst ét $x \in X$, sådan at $f(x) = y$; og den er *bijektiv*, hvis der for alle $y \in Y$ findes præcis ét $x \in X$, sådan at $f(x) = y$. Vi bemærker, at en afbildning er bijektiv, hvis og kun hvis den både er injektiv og surjektiv.

Eksempel 0.1.2 Afbildningen $f: \mathbb{R} \rightarrow \mathbb{R}$ fra eksempel 0.1.1 er hverken injektiv eller surjektiv, mens afbildningen $g: \mathbb{R} \rightarrow [0, \infty)$ er surjektiv, men ikke injektiv. For hvis $y \in [0, \infty)$, da er $g(x) = y$ for $x = \pm\sqrt{y}$, så $g: \mathbb{R} \rightarrow [0, \infty)$ er surjektiv, men da det for $y > 0$ gælder, at $+\sqrt{y} \neq -\sqrt{y}$, så er $g: \mathbb{R} \rightarrow [0, \infty)$ ikke injektiv.

¹Formelt definerer vi en *regel* til at være en delmængde $f \subset X \times Y$, sådan at det for alle $x \in X$ gælder, at der præcis findes ét element $y \in Y$, sådan at $(x, y) \in f$. Vi skriver da $y = f(x)$ i stedet for $(x, y) \in f$.

0 Skalarer



Figur 0.1: Injektive, surjektive og bijektive afbildninger

Givet afbildninger $g: Y \rightarrow Z$ og $f: X \rightarrow Y$, sådan at domænet af den første afbildning er lig med codomænet af den anden, da definerer vi den *sammensatte afbildning*

$$X \xrightarrow{g \circ f} Z$$

ved $(g \circ f)(x) = g(f(x))$. Vi siger også, at diagrammet

$$\begin{array}{ccc} & Y & \\ f \nearrow & & \searrow g \\ X & \xrightarrow{h} & Z \end{array}$$

kommuterer, hvis $h = g \circ f$. Vi definerer *identitetsafbildningen* af en mængde X til at være afbildningen $\text{id}_X: X \rightarrow X$ givet ved $\text{id}_X(x) = x$.

Lemma 0.1.3 *En afbildning $f: X \rightarrow Y$ er bijektiv, hvis og kun hvis der eksisterer en afbildning $g: Y \rightarrow X$, sådan at $f \circ g = \text{id}_Y$ og $g \circ f = \text{id}_X$.*

Bevis Vi antager først, at der findes en afbildning $g: Y \rightarrow X$, sådan at $f \circ g = \text{id}_Y$ og $g \circ f = \text{id}_X$, og viser, at $f: X \rightarrow Y$ er bijektiv. Givet $y \in Y$, da opfylder $x = g(y) \in X$, at

$$f(x) = f(g(y)) = (f \circ g)(y) = \text{id}_Y(y) = y,$$

så $f: X \rightarrow Y$ er surjektiv. Hvis også $x' \in X$ opfylder, at $f(x') = y$, da er

$$x' = \text{id}_X(x') = (g \circ f)(x') = g(f(x')) = g(y) = g(f(x)) = (g \circ f)(x) = \text{id}_X(x) = x,$$

så $f: X \rightarrow Y$ er også injektiv og dermed bijektiv. Omvendt, hvis $f: X \rightarrow Y$ er bijektiv, så lader vi $g: Y \rightarrow X$ være afbildningen, der til $y \in Y$ tilordner det entydigt bestemte element $x \in X$, sådan at $f(x) = y$. Hvis $x \in X$ og $y \in Y$, da er udsagnene “ $f(x) = y$ ” og “ $x = g(y)$ ” derfor ensbetydende. Specielt er udsagnene “ $f(g(y)) = y$ ” og “ $g(y) = g(y)$ ” ensbetydende, hvilket viser, at $f \circ g = \text{id}_Y$. Tilsvarende er udsagnene “ $g(f(x)) = x$ ” og “ $f(x) = f(x)$ ” ensbetydende, hvilket viser, at også $g \circ f = \text{id}_X$. $\square$

Hvis $f: X \rightarrow Y$ er bijektiv, så er afbildningen $g: Y \rightarrow X$ med $f \circ g = \text{id}_Y$ og $g \circ f = \text{id}_X$, entydigt bestemt. For hvis også $g': Y \rightarrow X$ opfylder $f \circ g' = \text{id}_Y$ og $g' \circ f = \text{id}_X$, da er

$$g = g \circ \text{id}_Y = g \circ (f \circ g') = (g \circ f) \circ g' = \text{id}_X \circ g' = g'.$$

Vi kalder afbildningen $g: Y \rightarrow X$ for den *inverse afbildning* af $f: X \rightarrow Y$. Den betegnes sommetider med $f^{-1}: Y \rightarrow X$. Vi bemærker, at det ifølge lemma 0.1.3 kun er bijektive afbildninger, der har en invers afbildning.

Eksempel 0.1.4 Vi lader $a \in \mathbb{R}$ og betragter afbildningen $f: \mathbb{R} \rightarrow \mathbb{R}$, der er defineret ved $f(x) = ax$. Hvis $a \neq 0$, da er $f: \mathbb{R} \rightarrow \mathbb{R}$ bijektiv, og dens inverse afbildning $g: \mathbb{R} \rightarrow \mathbb{R}$ er givet ved $g(y) = a^{-1}y$. Hvis $a = 0$, da er $f(x) = 0$ for alle $x \in \mathbb{R}$, så $f: \mathbb{R} \rightarrow \mathbb{R}$ er hverken injektiv eller surjektiv, og har derfor ikke nogen invers afbildning.

0.2 Legemer

Vi lader $\mathbb{F}$ være et vilkårligt legeme, hvilket vil sige en mængde af skalarer udstyret med sum og produkt, der opfylder de sædvanlige aritmetiske regler. Den præcise definition er som følger.

Definition 0.2.1 Et *legeme* er en triple $(\mathbb{F}, +, \cdot)$, der består af en mængde $\mathbb{F}$ samt to afbildninger $+: \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F}$ og $\cdot: \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F}$, sådan at det gælder:

- (A1) For alle $a, b, c \in \mathbb{F}$ er $(a + b) + c = a + (b + c)$.
- (A2) Der findes et element $0 \in \mathbb{F}$, sådan at $a + 0 = a = 0 + a$ for alle $a \in \mathbb{F}$.
- (A3) For alle $a \in \mathbb{F}$, findes der et $b \in \mathbb{F}$, sådan at $a + b = 0 = b + a$.
- (A4) For alle $a, b \in \mathbb{F}$ er $a + b = b + a$.
- (P1) For alle $a, b, c \in \mathbb{F}$ er $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- (P2) Der findes et element $1 \in \mathbb{F}$, sådan at $a \cdot 1 = a = 1 \cdot a$ for alle $a \in \mathbb{F}$.
- (P3) For alle $0 \neq a \in \mathbb{F}$, findes $b \in \mathbb{F}$, sådan at $a \cdot b = 1 = b \cdot a$. Endvidere er $0 \neq 1$.
- (P4) For alle $a, b \in \mathbb{F}$ er $a \cdot b = b \cdot a$.
- (D1) For alle $a, b, c \in \mathbb{F}$ er $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.
- (D2) For alle $a, b, c \in \mathbb{F}$ er $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$.

Vi forkorter normalt og skriver ab i stedet for $a \cdot b$. Vi forkorter også normalt og skriver blot $\mathbb{F}$ for legemet $(\mathbb{F}, +, \cdot)$. Dette er et eksempel på misbrug af notation, idet vi benytter symbolet $\mathbb{F}$ til at betegne to forskellige objekter.

Eksempel 0.2.2 Vi giver følgende mere eller mindre velkendte eksempler på legemer.

- ($\mathbb{R}$) Det bedst kendte eksempel på et legeme er nok legemet af reelle tal $(\mathbb{R}, +, \cdot)$, hvor $\mathbb{R}$ er mængden af reelle tal, og hvor “+” og “ $\cdot$ ” er henholdsvis den sædvanlige sum og det sædvanlige produkt af reelle tal.
- ($\mathbb{Q}$) Et andet velkendt eksempel er legemet af rationale tal $(\mathbb{Q}, +, \cdot)$, hvor $\mathbb{Q} \subset \mathbb{R}$ er delmængden af rationale tal, og hvor igen “+” og “ $\cdot$ ” er de sædvanlige sum og produkt operationer.
- ($\mathbb{C}$) Et meget vigtigt eksempel er legemet af komplekse tal $(\mathbb{C}, +, \cdot)$, som består af mængden af komplekse tal

$$\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$$

med “+” og “ $\cdot$ ” defineret ved henholdsvis

$$\begin{aligned}(a + ib) + (c + id) &= (a + c) + i(b + d) \\ (a + ib) \cdot (c + id) &= (ac - bd) + i(ad + bc).\end{aligned}$$

Det modsatte element af $a + ib$ er $-a + i(-b)$, som vi også skriver $-a - ib$, og, hvis $a + ib \neq 0$, da er det multiplikativt inverse element af $a + ib$ givet ved

$$\frac{1}{a + ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a - ib}{a^2 + b^2} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2}.$$

- ($\mathbb{F}_2$) Det mindste legeme er legemet $(\mathbb{F}_2, +, \cdot)$, hvor $\mathbb{F}_2 = \{0, 1\}$, og hvor $1 + 1$ er defineret til at være lig med 0.

Der findes mange andre legemer, endelige som uendelige, og lineær algebra fungerer over et vilkårligt legeme. I dette kursus vil vi dog kun bruge $\mathbb{R}$ og $\mathbb{C}$ som eksempler.

De grundlæggende egenskaber (A1)–(A4), (P1)–(P4) og (D1)–(D2), der per definition gælder for sum og produkt i et legeme, har flere konsekvenser, som er velkendte for de reelle og de komplekse tal, men som altid gælder så snart de grundlæggende betingelser er opfyldt. Vi nævner her nogle af de vigtigste:

- (1) Elementet $0 \in \mathbb{F}$, der opfylder (A2), er entydigt bestemt. For hvis 0 og $0'$ begge opfylder (A2), da gælder

$$0 = 0 + 0' = 0'.$$

Vi kalder $0 \in \mathbb{F}$ for *nul-elementet* i $\mathbb{F}$.

- (2) Givet $a \in \mathbb{F}$, da er elementet $b \in \mathbb{F}$, der opfylder (A3) ligeledes entydigt bestemt. For hvis både b og b' opfylder (A3), da er

$$b = b + 0 = b + (a + b') = (b + a) + b' = 0 + b' = b'.$$

Vi skriver $-a$ for dette element b og kalder det for det *modsatte element* af a .

- (3) Man viser tilsvarende, at elementet $1 \in \mathbb{F}$, der opfylder (P2), er entydigt bestemt; vi kalder dette element for *et-elementet* i $\mathbb{F}$. Hvis $a \in \mathbb{F}$, og der findes $b \in \mathbb{F}$, sådan at $ab = 1 = ba$, så ser vi ligeledes, at elementet $b \in \mathbb{F}$ er entydigt bestemt af $a \in \mathbb{F}$. Vi siger da, at a er *invertibel*, og at vi skriver a^{-1} for det entydigt bestemte $b \in \mathbb{F}$, sådan at $ab = 1 = ba$, og kalder det for det *multiplikativt inverse* element af a . Så ifølge (P3) gælder det i et legeme $\mathbb{F}$, at ethvert element $a \neq 0$ er invertibelt.

- (4) Hvis $a + c = b + c$, så er $a = b$, idet

$$\begin{aligned} a &\stackrel{(A2)}{=} a + 0 \stackrel{(A3)}{=} a + (c + (-c)) \stackrel{(A1)}{=} (a + c) + (-c) \\ &= (b + c) + (-c) \stackrel{(A1)}{=} b + (c + (-c)) \stackrel{(A3)}{=} b + 0 \stackrel{(A2)}{=} b. \end{aligned}$$

Det vil sige, at vi kan trække et vilkårligt $c \in \mathbb{F}$ fra på begge sider af et lighedstegn. Tilsvarende gælder det for alle $c \in \mathbb{F}$ med $c \neq 0$, at hvis $a \cdot c = b \cdot c$, så er $a = b$, fordi

$$\begin{aligned} a &\stackrel{(P2)}{=} a \cdot 1 \stackrel{(P3)}{=} a \cdot (c \cdot c^{-1}) \stackrel{(P1)}{=} (a \cdot c) \cdot c^{-1} \\ &= (b \cdot c) \cdot c^{-1} \stackrel{(P1)}{=} b \cdot (c \cdot c^{-1}) \stackrel{(P3)}{=} b \cdot 1 \stackrel{(P2)}{=} b. \end{aligned}$$

Det vil sige, at vi kan dividere med $c \in \mathbb{F}$ på begge sider af et lighedstegn, forudsat at c er invertibel. Vi kan derimod ikke dividere med 0, fordi 0 ikke er invertibel.

Vi nævner derudover de følgende to konsekvenser af Definition 0.2.1:

Sætning 0.2.3 *Lad $\mathbb{F}$ være et legeme. For alle $a \in \mathbb{F}$ gælder følgende:*

(E1) $a \cdot 0 = 0 = 0 \cdot a.$

(E2) $a \cdot (-1) = -a = (-1) \cdot a.$

Bevis Vi beviser først (E1). Givet $a \in \mathbb{F}$, da er

$$a \cdot 0 \stackrel{(A2)}{=} a \cdot (0 + 0) \stackrel{(D1)}{=} (a \cdot 0) + (a \cdot 0),$$

og ved at trække $a \cdot 0$ fra venstre og højre side får vi da, at $0 = (a \cdot 0)$ som ønsket.

0 Skalarer

Vi viser dernæst (E2). Ifølge bemærkning (2) ovenfor, er $-a \in \mathbb{F}$ det entydigt bestemte element, der opfylder $a + (-a) = 0 = (-a) + a$. Men udregningerne

$$\begin{aligned} a + (-1) \cdot a &\stackrel{(P2)}{=} 1 \cdot a + (-1) \cdot a \stackrel{(D2)}{=} (1 + (-1)) \cdot a \stackrel{(A3)}{=} 0 \cdot a \stackrel{(E1)}{=} 0 \\ (-1) \cdot a + a &\stackrel{(P2)}{=} (-1) \cdot a + 1 \cdot a \stackrel{(D1)}{=} a \cdot (1 + (-1)) \stackrel{(A3)}{=} a \cdot 0 \stackrel{(E1)}{=} 0 \end{aligned}$$

viser, at elementet $(-1) \cdot a$ også har denne egenskab, hvilket viser (E2). $\square$

Bemærkning 0.2.4 Triplen $(\mathbb{Z}, +, \cdot)$, hvor $\mathbb{Z}$ er mængden af hele tal, og hvor “+” og “ $\cdot$ ” er de sædvanlige sum og produkt operationer, opfylder alle aksiomerne i Definition 0.2.1 med undtagelse af (P3). For +1 og -1 er de eneste elementer i $\mathbb{Z}$, der er invertible, så ikke alle $0 \neq x \in \mathbb{Z}$ er invertible. Dermed er $(\mathbb{Z}, +, \cdot)$ altså ikke noget legeme. En triple af denne art kaldes for en *kommutativ ring*. Disse forekommer i enorm variation.²

0.3 Højre og venstre multiplikation

Hvis $(\mathbb{F}, +, \cdot)$ er et legeme, så gælder det ifølge aksiom (P4), at $a \cdot b = b \cdot a$ for alle $a, b \in \mathbb{F}$. Med andre ord, så gør det ingen forskel, om vi ganger a med b fra højre eller fra venstre. Der forekommer dog naturligt også tripler $(\mathbb{F}, +, \cdot)$, der opfylder alle aksiomerne for at være et legeme med undtagelse af (P4), og vi siger da, at $(\mathbb{F}, +, \cdot)$ er et *skævlegeme*.

Et klassisk eksempel på et skævlegeme er Hamilton’s kvaternioner, som vi beskriver i eksempel 0.3.1 nedenfor. Enhedskvaternioner bestemmer rotationer af 3-dimensionalt rum, og “ $a \cdot b$ ” svarer da til først at udføre rotationen bestemt af b og dernæst rotationen bestemt af a , mens “ $b \cdot a$ ” svarer til at udføre disse rotationer i modsat rækkefølge, hvilket typisk ikke er det samme.

I lineær algebra spiller multiplikation fra højre og venstre forskellige roller, og vi gør det derfor til en regel altid at være forsigtige med, hvilken side vi ganger fra. Den kommutative lov (P4) gælder således *ikke* for multiplikation af matricer, som vi indfører i kapitel 2. For eksempel kan rotationer af 3-dimensionalt rum udtrykkes ved matricer, og sammensætning af rotationer svarer da til matrixmultiplikation. Endelig er (P4) slet ikke nødvendig i de fleste af de efterfølgende kapitler, og mange beviser bliver faktisk lettere, når man ikke tillader sig at bruge (P4).

Eksempel 0.3.1 Hamilton’s *kvaternioner* er skævlegemet $(\mathbb{H}, +, \cdot)$, hvor

$$\mathbb{H} = \{a + ib + jc + kd \mid a, b, c, d \in \mathbb{R}\}$$

er mængden af kvaternioner, og hvor “+” er defineret ved

$$\begin{aligned} (a_1 + ib_1 + jc_1 + kd_1) + (a_2 + ib_2 + jc_2 + kd_2) \\ = (a_1 + a_2) + i(b_1 + b_2) + j(c_1 + c_2) + k(d_1 + d_2), \end{aligned}$$

² Det mere end 5000 sider lange Stacks Project (stacks.math.columbia.edu) omhandler kommutative ringe og deres egenskaber.

mens “ $\cdot$ ” er defineret ved

$$\begin{aligned} & (a_1 + ib_1 + jc_1 + kd_1) \cdot (a_2 + ib_2 + jc_2 + kd_2) \\ &= (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) + i(a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2) \\ & \quad + j(a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2) + k(a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2). \end{aligned}$$

For eksempel er $i \cdot j = k$, mens $j \cdot i = -k$, så multiplikation er ikke kommutativ.

0.4 Opgaver

0.1 Vis, at afbildningen $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defineret ved

$$f \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + v \\ u - v \end{pmatrix}$$

er bijektiv, og angiv den inverse afbildning $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$.

0.2 Afgør i hvert af følgende tilfælde, om den angivne afbildning $f: X \rightarrow Y$ er injektiv, surjektiv og/eller bijektiv, og angiv billedet af $f: X \rightarrow Y$. Hvis $f: X \rightarrow Y$ er bijektiv, angiv da også den inverse afbildning $g: Y \rightarrow X$.

a) $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$.

b) $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^3$.

c) $f: \mathbb{R}^2 \rightarrow \mathbb{R}, f \begin{pmatrix} x \\ y \end{pmatrix} = xy$.

d) $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x - y \end{pmatrix}$.

e) $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, f(x) = \frac{1}{x}$.

f) $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} \setminus \{0\}, f(x) = \frac{1}{x}$.

g) $f: [0, 1] \rightarrow [0, 1], f(x) = x^2$.

0.3 Afgør, om hver af de følgende afbildninger er injektiv, surjektiv og/eller bijektiv. Hvis afbildningen er bijektiv, angiv da også den inverse afbildning.

a) Afbildningen $f: \mathbb{C} \rightarrow \mathbb{R}^2$ defineret ved

$$f(z) = \begin{pmatrix} \operatorname{Re}(z) \\ \operatorname{Im}(z) \end{pmatrix}.$$

b) Afbildningen $g: \mathbb{C}^2 \rightarrow \mathbb{R}^2$ defineret ved

$$g \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} \operatorname{Re}(z) \\ \operatorname{Im}(z) \end{pmatrix}.$$

0 Skalarer

c) Afbildningen $h: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ defineret ved

$$h \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} \bar{z} \\ w \end{pmatrix}.$$

d) Afbildningen $l: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ defineret ved

$$l \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} i \cdot z - w \\ z \end{pmatrix}.$$

0.4 Vi betragter mængderne

$$\begin{array}{ll} A = \{a, b, c\} & B = \{2, 3, 5, 7, 11\} \\ D = \{\alpha, \omega\} & E = \{I, V, X, L, C, D, M\}. \end{array}$$

Angiv for hver af følgende afbildninger, om den er injektiv og/eller surjektiv. Angiv den inverse afbildning i de tilfælde, hvor afbildningen er bijektiv.

a) Afbildningen $f: A \rightarrow B$ defineret ved

$$f(a) = 2, \quad f(b) = 11, \quad f(c) = 2.$$

b) Afbildningen $g: D \rightarrow A$ defineret ved

$$g(\alpha) = b, \quad g(\omega) = a.$$

c) Afbildningen $h: E \rightarrow A$ defineret ved

$$\begin{array}{llll} h(I) = c, & h(V) = a, & h(X) = a, & h(L) = b, \\ h(C) = c, & h(D) = b, & h(M) = a. & \end{array}$$

d) Afbildningen $i: A \cup D \rightarrow B$ defineret ved

$$i(a) = 7, \quad i(b) = 3, \quad i(c) = 11, \quad i(\alpha) = 2, \quad i(\omega) = 5.$$

0.5 Afbildningerne $f, g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ er givet ved

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + x_3 \end{pmatrix}, \quad g \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 + 2x_3 \\ x_2 + x_3 \\ x_1 + x_3 \end{pmatrix}.$$

a) Gør rede for, at $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ er bijektiv, mens $g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ikke er bijektiv.

b) Angiv den inverse afbildning til $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

0.6 Lad X , Y og Z være mængder, og lad $f: X \rightarrow Y$ og $g: Y \rightarrow Z$ være afbildninger.

- a) Vis, at hvis $f: X \rightarrow Y$ og $g: Y \rightarrow Z$ begge er injektive, da er $g \circ f: X \rightarrow Z$ også injektiv.
- b) Vis, at hvis $f: X \rightarrow Y$ og $g: Y \rightarrow Z$ begge er surjektive, da er $g \circ f: X \rightarrow Z$ også surjektiv.
- c) Vis, at hvis $f: X \rightarrow Y$ og $g: Y \rightarrow Z$ begge er bijektive, da er $g \circ f: X \rightarrow Z$ bijektiv, og at det endvidere gælder, at $(g \circ f)^{-1} = f^{-1} \circ g^{-1}: Z \rightarrow X$.
- d) Vis, at hvis $g \circ f: X \rightarrow Z$ er injektiv, så er også $f: X \rightarrow Y$ injektiv.
- e) Vis, at hvis $g \circ f: X \rightarrow Z$ er surjektiv, så er også $g: Y \rightarrow Z$ surjektiv.

0.7 (★) Angiv en bijektiv afbildning $f: [0, 1] \rightarrow [0, 1)$.

[Vink: Afbildningen $f: [0, 1] \rightarrow [0, 1)$ kan ikke være kontinuert.]

1 Lineære ligningssystemer

I dette kapitel indfører vi lineære ligningssystemer og beskriver en algoritme til at finde deres løsninger. Denne algoritme eller metode, der allerede er beskrevet i den kinesiske tekst “Ni kapitler om den matematiske kunst” skrevet omkring begyndelsen af vor tidsregning, blev langt senere genfundet af først Newton og senere Gauss, og går nu, noget uretfærdigt, under navnet *Gauss-elimination*. Denne algoritme er utroligt effektiv og er grundlaget for stort alle beregninger i lineær algebra.

Algoritmen består i at omforme ligningssystemer ved hjælp af *rækkeoperationer*, der ikke ændrer løsningsmængden, men gør det muligt umiddelbart at angive denne. Vi samler al information om et lineært ligningssystem i en *matrix*, og vi indfører en speciel form for matricer, som vi kalder matricer på *reduceret echelonform*, hvor det tilhørende ligningssystem umiddelbart kan løses. Gauss-elimination er da en algoritme, der ved hjælp af rækkeoperationer omdanner en given matrix til en entydigt bestemt matrix på reduceret echelonform.

Løsning af lineære af ligningssystem ved Gauss-elimination er beskrevet i detaljer i sætning 1.2.17 nedenfor. I grove træk afgør algoritmen, om der er løsninger eller ej, og hvis der er løsninger, så bliver de alle fundet af algoritmen. Specielt afgør algoritmen, om der er nul, én eller flere løsninger.

Vi minder om fra kapitel 0, at vi arbejder med skalarer i et legeme $\mathbb{F}$, som for os er enten de reelle eller de komplekse tal. De eneste egenskaber ved de reelle og komplekse tal, som vi anvender, er, at de kan adderes og multipliceres, og at disse operationer opfylder legemesaksiomerne beskrevet i definition 0.2.1.

1.1 Lineære ligningssystemer

Det lineære ligningssystem

$$4x_1 + x_2 = 9$$

$$x_1 - x_2 = 1$$

består af to ligninger “ $4x_1 + x_2 = 9$ ” og “ $x_1 - x_2 = 1$ ” med to ubekendte x_1 og x_2 . Det er ikke svært at se, at $x_1 = 2$ og $x_2 = 1$ tilfredsstiller begge ligninger, og derfor er en løsning til ligningssystemet, samt at dette er den eneste løsning. Vi angiver denne løsning som

$$\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

1 Lineære ligningssystemer

Generelt siger vi, at et system af ligninger på formen

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\&\vdots \\a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m\end{aligned}$$

med $a_{ij}, b_i \in \mathbb{F}$ er et *lineært ligningssystem* med m ligninger i n ubekendte over $\mathbb{F}$, og vi kalder skalarerne a_{ij} og b_i for ligningssystemets *koefficienter* og *konstanter*. Ved en *løsning* til ligningssystemet forstår vi en familie af n skalarer

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

der samtidigt tilfredstiller de m ligninger i systemet. Den mængde, der består af alle løsninger til ligningssystemet, kaldes for ligningssystemets *løsningsmængde*.

Hvis konstanterne b_i alle er lig med 0, siger vi, at ligningssystemet er *homogent*, og ellers siger vi, at det er *inhomogent*. Ethvert homogent ligningssystem har løsningen

$$\mathbf{x} = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

mens et inhomogent ligningssystem ikke behøver at have nogen løsning.

Eksempel 1.1.1 (Homogent og inhomogent ligningssystem) Ligningssystemet

$$\begin{aligned}2x_1 - 3x_2 + x_3 &= -1 \\-x_1 + x_2 - x_3 &= 2 \\3x_1 + 2x_2 + 2x_3 &= 3\end{aligned}$$

med $m = 3$ ligninger i $n = 3$ ubekendte er inhomogent; og ligningssystemet

$$\begin{aligned}2x_1 - 3x_2 + x_3 &= 0 \\-x_1 + x_2 - x_3 &= 0 \\3x_1 + 2x_2 + 2x_3 &= 0\end{aligned}$$

er det tilhørende homogene ligningssystem.

1.2 Rækkeoperationer og Gauss-elimination

Det er nyttigt ikke at skulle skrive de variable x_j hele tiden, og derfor samler vi ligningssystemets koefficienter a_{ij} i en $m \times n$ -matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

som vi kalder ligningssystemets *koefficientmatrix*. Vi siger, at matricen har m *rækker* og n *søjler*. Hvis vi tilføjer den ekstra søjle

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

der består af ligningssystemets konstanter, så får vi en $m \times (n + 1)$ -matrix

$$(A | \mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right),$$

som vi kalder ligningssystemets *totalmatrix*. Vi har her, som det er sædvane, inkluderet en lodret linje, der adskiller koefficienter og konstanter. Linjen er ikke af matematisk betydning og har udelukkende til formål at huske os på, hvilket ligningssystem denne matrix repræsenterer.

Eksempel 1.1.2 (Totalmatrix) Ligningssystemet i eksempel 1.1.1 har totalmatrix

$$\left(\begin{array}{ccc|c} 2 & -3 & 1 & -1 \\ -1 & 1 & -1 & 2 \\ 3 & 2 & 2 & 3 \end{array} \right).$$

Den har 3 rækker og 4 søjler, og er dermed en 3×4 -matrix.

1.2 Rækkeoperationer og Gauss-elimination

I dette afsnit indfører vi rækkeoperationer og viser, hvordan de anvendes til at løse lineære ligningssystemer. Vi illustrerer først dette med det simple eksempel

$$\begin{aligned} 4x_1 + x_2 &= 9 \\ x_1 - x_2 &= 1, \end{aligned}$$

1 Lineære ligningssystemer

hvor vi nu beskriver hver skridt i udregningen. Vi begynder med at addere den anden ligning til den første, og får da ligningssystemet

$$\begin{aligned}5x_1 &= 10 \\ x_1 - x_2 &= 1.\end{aligned}$$

Vi ganger dernæst den første ligning med $\frac{1}{5}$ og den anden med -1 , hvilket giver

$$\begin{aligned}x_1 &= 2 \\ -x_1 + x_2 &= -1.\end{aligned}$$

Endelig adderer vi den første ligning til den anden og får derved

$$\begin{aligned}x_1 &= 2 \\ x_2 &= 1.\end{aligned}$$

De fire ligningssystemer har alle samme løsningsmængde, da de operationer, vi har udført, kan inverteres og derfor bevarer løsningsmængden. Vi konkluderer dermed, at den fælles løsningsmængde består af den entydige løsning

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Vi skal nu vise, at denne løsningsmetode virker for et generelt lineært ligningssystem. Så vi betragter lineære ligningssystemer

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m\end{aligned}$$

bestående af m ligninger i n ubekendte og indfører de følgende tre typer af operationer, der omformer et sådant lineært ligningssystem til et nyt lineært ligningssystem med samme løsningsmængde.

Type M: Multiplikation fra venstre af en ligning med en skalar $c \neq 0$.

Type S: Addition af et venstre multiplum af en ligning til en anden ligning.

Type T: Ombytning af to ligninger.

Her indikerer “M”, “S” og “T” henholdsvis *multiplikation*, *sum* og *transposition*, og som vi forklarede i afsnit 0.3, så skelner vi mellem multiplikation med skalarer fra venstre og fra højre.

1.2 Rækkeoperationer og Gauss-elimination

De tre typer af operationer bevarer både antallet m af ligninger og antallet n af ubekendte. Endvidere gælder det, at enhver løsning til et givet ligningssystem igen er en løsning til det nye ligningssystem, der fremkommer ved at udføre en af de tre typer operationer.

Desuden har hver af de tre typer operationer en *invers* operation: For hvis vi ganger ligning i med $c \neq 0$ fra venstre, så kan vi gange den nye ligning i med $c^{-1} \neq 0$ fra venstre og derved få den oprindelige ligning i tilbage. Og hvis vi adderer c gange ligning j til ligning $i \neq j$, så kan vi addere $-c$ gange ligning j til den nye ligning i og derved få den oprindelige ligning i tilbage. Endelig, hvis vi ombytter ligning i og j , så kan vi ombytte dem igen, og derved komme tilbage til, der hvor vi startede. Derfor er enhver løsning til det nye ligningssystem også en løsning til det oprindelige ligningssystem. Med andre ord har de to ligningssystemer den samme løsningsmængde.

Vi udtrykker nu de tre typer operationer ved ligningssystemernes totalmatricer

$$(A | \mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right),$$

hvor de svarer til de følgende tre typer *rækkeoperationer* på matricer:

Type M: Multiplikation fra venstre af en række med en skalar $c \neq 0$.

Type S: Addition af et venstre multiplum af en række til en anden række.

Type T: Ombytning af to rækker.

Vi bemærker, at disse rækkeoperationer bevarer antallet af rækker og antallet af søjler.

Rækkeoperationer forstås lettest ved gennemgang af nogle eksempler. Disse eksempler illustrerer samtidigt, hvordan vi angiver rækkeoperationer.

Eksempel 1.2.1 Vi betragter igen ligningssystemet

$$\begin{aligned} 4x_1 + x_2 &= 9 \\ x_1 - x_2 &= 1, \end{aligned}$$

som vi løste i starten af afsnittet. Vi udfører nu de rækkeoperationer på ligningssyste-

1 Lineære ligningssystemer

met totalmatrix $(A | \mathbf{b})$, der svarer til de rækkeoperationer, vi anvendte tidligere.

$$\begin{aligned}(A | \mathbf{b}) &= \left(\begin{array}{cc|c} 4 & \color{red}{1} & 9 \\ 1 & -1 & 1 \end{array} \right) +R_2 \\ &\quad \left(\begin{array}{cc|c} \color{red}{5} & 0 & 10 \\ 1 & -1 & 1 \end{array} \right) \frac{1}{5} \cdot R_1 \\ &\quad \left(\begin{array}{cc|c} 1 & 0 & 2 \\ \color{red}{1} & -1 & 1 \end{array} \right) (-1) \cdot R_2 \\ &\quad \left(\begin{array}{cc|c} 1 & 0 & 2 \\ \color{red}{-1} & 1 & -1 \end{array} \right) +R_1 \\ (A' | \mathbf{b}') &= \left(\begin{array}{cc|c} \color{blue}{1} & 0 & 2 \\ 0 & \color{blue}{1} & 1 \end{array} \right)\end{aligned}$$

Ligningssystemet hørende til $(A' | \mathbf{b}')$ er

$$\begin{aligned}x_1 &= 2 \\ x_2 &= 1,\end{aligned}$$

hvilket er trivielt at løse.

Bemærkning 1.2.2 Vi bemærker, at de fem matricer i eksempel 1.2.1 *ikke* er lig med hinanden, og at det derfor er forkert at skrive lighedstegn mellem dem.

Vi viser nu, at rækkeoperationer er invertible i den forstand, at hvis en given rækkeoperation omformer matricen A til matricen A' , så omformer den inverse rækkeoperation matricen A' til matricen A .

Sætning 1.2.3 Rækkeoperationerne af type M , S og T er invertible, og de inverse rækkeoperationer er givet følger:

- (1) Den inverse af rækkeoperationen af type M , der består i at gange række i med $c \neq 0$ fra venstre, er rækkeoperationen af type M , der består i at gange række i med c^{-1} fra venstre.
- (2) Den inverse af rækkeoperationen af type S , der består i at addere c gange række j til række $i \neq j$, er rækkeoperationen af type S , der består i at addere $-c$ gange række j til række i .
- (3) Den inverse af rækkeoperationen af type T , der består i at ombytte række i og j , er den samme rækkeoperation af type T .

1.2 Rækkeoperationer og Gauss-elimination

Bevis Vi beviser (2); (1) og (3) vises tilsvarende. Så lad $A = (a_{kl})$ være en $m \times n$ -matrix, og lad $A' = (a'_{kl})$ være $m \times n$ -matricen, der fremkommer fra A ved rækkeoperationen af type S, der adderer c gange den i 'te række til den j 'te række. Da gælder det, at

$$a'_{kl} = \begin{cases} a_{jl} + c \cdot a_{il} & \text{hvis } k = j, \\ a_{kl} & \text{hvis } k \neq j. \end{cases}$$

Lad $A'' = (a''_{kl})$ være $m \times n$ -matricen, der fremkommer fra A' ved rækkeoperationen af type S, der adderer $-c$ gange den i 'te række til den j 'te række. Da gælder det, at

$$a''_{kl} = \begin{cases} a'_{jl} + (-c) \cdot a'_{il} & \text{hvis } k = j, \\ a'_{kl} & \text{hvis } k \neq j, \end{cases}$$

og vi skal vise, at $A'' = A$. Hvis $k \neq j$, da er $a''_{kl} = a'_{kl} = a_{kl}$, og hvis $k = j$, da er

$$\begin{aligned} a''_{jl} &= a'_{jl} + (-c) \cdot a'_{il} = (a_{jl} + c \cdot a_{il}) + (-c) \cdot a_{il} = a_{jl} + (c \cdot a_{il} + (-c) \cdot a_{il}) \\ &= a_{jl} + (c + (-c)) \cdot a_{il} = a_{jl} + 0 \cdot a_{il} = a_{jl} + 0 = a_{jl}. \end{aligned}$$

Dette viser sætningen. □

Eksempel 1.2.4 Vi betragter igen matricerne $(A \mid \mathbf{b})$ og $(A' \mid \mathbf{b}')$ fra eksempel 1.2.1. Vi udfører nu de inverse rækkeoperationer af de rækkeoperationer, vi udførte for at omdanne $(A \mid \mathbf{b})$ til $(A' \mid \mathbf{b}')$, og vi udfører dem endvidere i den omvendte rækkefølge.

$$\begin{aligned} (A' \mid \mathbf{b}') &= \left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array} \right) + (-1) \cdot R_1 \\ &= \left(\begin{array}{cc|c} 1 & 0 & 2 \\ -1 & 1 & -1 \end{array} \right) + (-1) \cdot R_2 \\ &= \left(\begin{array}{cc|c} 1 & 0 & 2 \\ 1 & -1 & 1 \end{array} \right) + 5 \cdot R_1 \\ &= \left(\begin{array}{cc|c} 5 & 0 & 10 \\ 1 & -1 & 1 \end{array} \right) + (-1) \cdot R_2 \\ (A \mid \mathbf{b}) &= \left(\begin{array}{cc|c} 4 & 1 & 9 \\ 1 & -1 & 1 \end{array} \right) \end{aligned}$$

De inverse rækkeoperationer, anvendt i omvendt rækkefølge, omdanner altså matricen $(A' \mid \mathbf{b}')$ til den oprindelige matrix $(A \mid \mathbf{b})$.

1 Lineære ligningssystemer

Sætning 1.2.5 Hvis totalmatricen $(A \mid \mathbf{b})$ for et lineært ligningssystem kan omdannes til totalmatricen $(A' \mid \mathbf{b}')$ for et andet lineært ligningssystem ved at udføre et endeligt antal rækkeoperationer, så har de to ligningssystemer den samme løsningsmængde.

Bevis Vi beviser sætningen ved induktion på antallet $N \geq 0$ af rækkeoperationer, der benyttes til at omdanne $(A \mid \mathbf{b})$ til $(A' \mid \mathbf{b}')$. Hvis $N = 0$, har man udført ingen rækkeoperationer, så er $(A \mid \mathbf{b}) = (A' \mid \mathbf{b}')$, og der er derfor ikke noget at vise. Vi antager dernæst, at påstanden er bevist for $N = r - 1$, og beviser den for $N = r$.

Så lad $(A \mid \mathbf{b})$ være totalmatricen for et lineært ligningssystem, og lad $(A' \mid \mathbf{b}')$ være en matrix, der er fremkommet fra matricen $(A \mid \mathbf{b})$ ved at udføre r rækkeoperationer. Vi skal vise, at ligningssystemerne hørende til $(A \mid \mathbf{b})$ og $(A' \mid \mathbf{b}')$ har den samme løsningsmængde. Lad $(B \mid \mathbf{c})$ være den matrix, der fremkommer fra $(A \mid \mathbf{b})$ ved at udføre de $r - 1$ første af de r rækkeoperationer. Per induktion ved vi da, at ligningssystemerne hørende til $(A \mid \mathbf{b})$ og $(B \mid \mathbf{c})$ har samme løsningsmængde, og det er derfor tilstrækkeligt at vise, at ligningssystemerne hørende til $(B \mid \mathbf{c})$ og $(A' \mid \mathbf{b}')$ har samme løsningsmængde. Men $(A' \mid \mathbf{b}')$ fremkommer fra $(B \mid \mathbf{c})$ ved at udføre en enkelt rækkeoperation, og som vi allerede har bemærket har de tilhørende ligningssystemer derfor den samme løsningsmængde. Dette viser induktionsskridtet og dermed sætningen. $\square$

Eksempel 1.2.6 Hvis man er *meget* forsigtig, så kan to eller flere operationer udføres i samme skridt forudsat, at de er *ombyttelige*. De fire operationer

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 7 & 3 & 1 \\ 0 & 4 & 6 & -6 \end{pmatrix} \begin{array}{l} R_1 \leftrightarrow R_2 \\ \frac{1}{2} \cdot R_3 \\ \begin{pmatrix} 1 & 7 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 2 & 3 & -3 \end{pmatrix} \begin{array}{l} +(-1) \cdot R_2 \\ +3 \cdot R_2 \end{array} \end{array} \\ A' &= \begin{pmatrix} 1 & 6 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 5 & 9 & 0 \end{pmatrix} \end{aligned}$$

omdanner A til A' . De første to operationer er ombyttelige, da de ikke involverer de samme rækker, og det samme gælder for de sidste to operationer.

Vi vil anvende rækkeoperationer til at omdanne totalmatricerne af ligningssystemer, til matricer på en særlig form, som vi kalder *reduceret echelonform*. Definitionen, som vi nu giver, forstås lettest ved samtidigt at betragte eksempel 1.2.8 nedenfor.

Definition 1.2.7 En $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$,

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

siges at være på *reduceret echelonform*, hvis der findes $r \geq 0$ og $1 \leq j_1 < \cdots < j_r \leq n$, sådan at følgende gælder:

(1) For alle $1 \leq s \leq r$ gælder det, at

$$a_{ij_s} = \begin{cases} 1 & \text{hvis } i = s, \\ 0 & \text{hvis } i \neq s. \end{cases}$$

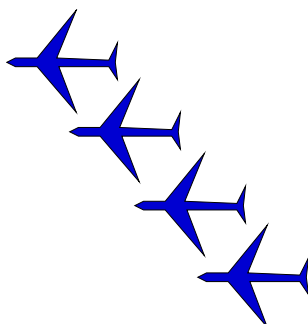
(2) For alle $1 \leq s \leq r$ og $1 \leq j < j_s$ gælder det, at $a_{sj} = 0$.

(3) For alle $r < i \leq m$ og $1 \leq j \leq n$ gælder det, at $a_{ij} = 0$.

Eksempel 1.2.8 Den følgende 5×9 -matrix er på reduceret echelonform.

$$\begin{pmatrix} 0 & \textcolor{blue}{1} & 0 & 7 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & \textcolor{blue}{1} & 3 & 0 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & \textcolor{blue}{1} & 6 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \textcolor{blue}{1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

For denne matrix er $r = 4$, mens søjlerne $j_1 = 2$, $j_2 = 3$, $j_3 = 5$ og $j_4 = 8$ er dens specielle søjler. Betingelsen (1) udtrykker, at i søjle nummer j_s er indgangen a_{sj_s} lig med 1, mens alle øvrige indgange er lig med 0. Indgangene $a_{sj_s} = 1$, som vi har markeret med blåt, kaldes for matrixens ledende indgange. Betingelsen (2) udtrykker, at i en række, der indeholder en ledende indgang, er alle indgange til venstre for den ledende indgang lig med 0. Endelig udtrykker betingelsen (3), at i de sidste $m - r$ rækker er alle indgange lig med 0. Vi bemærker, at disse $m - r$ rækker er netop de rækker, der ikke indeholder en ledende indgang.



Figur 1.1: Betegnelsen “echelonform” har sin oprindelse i den militære betegnelse “echelonformation”. Figuren viser fly, der flyver i echelonformation.

Definition 1.2.9 I en matrix A , der er på reduceret echelonform, kaldes antallet r af ikke-nul rækker for matrixens *rang*, mens den første ikke-nul indgang fra venstre i en række kaldes for rækkens *ledende indgang*.

I en matrix på reduceret echelonform er de ledende indgange altså alle lig med 1, og antallet af ledende indgange i matrixen er lig med rangen r .

Vi siger endvidere, at en $m \times n$ -matrix $A = (a_{ij})$ er på *echelonform*, hvis den opfylder betingelserne (2) og (3) i definition 1.2.7 samt den svagere betingelse

(1') For alle $1 \leq s \leq r$ gælder det, at $a_{s,j_s} \neq 0$.

Hvis en matrix A er på echelonform, så kalder vi igen antallet r af ikke-nul rækker for matrixens rang, og vi kalder den første ikke-nul indgang fra venstre i en række for rækkens ledende indgang. Rangen r er således igen lig med antallet af ledende indgange i matrixen.

Eksempel 1.2.10 De to følgende 4×8 -matricer, hvor “*” indikerer en vilkårlig skalar, er henholdsvis på reduceret echelonform og på echelonform.

$$\begin{pmatrix} 1 & * & * & 0 & * & 0 & * & * \\ 0 & 0 & 0 & 1 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 2 & * & * & * & * & * & * \\ 0 & 0 & 4 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 5 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & * \end{pmatrix}$$

Deres rang er henholdsvis 3 og 4, og de ledende indgange er markeret med blåt. Vi bemærker, at i matrixen, der er på reduceret echelonform, er de ledende indgange alle lig med 1, og at i de søjler, der indeholder en ledende indgang, er alle øvrige indgange lig med 0.

1.2 Rækkeoperationer og Gauss-elimination

Matricer på reduceret echelonform svarer til totalmatricer for ligningssystemer som umiddelbart kan løses. Det simpleste tilfælde er det tilfælde, hvor totalmatricen har en ledende indgang i alle søjler undtagen konstantssøjlen som i følgende eksempel.

Eksempel 1.2.11 Matricen

$$(A | \mathbf{b}) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array} \right)$$

er på reduceret echelonform af rang $r = 3$. Den er totalmatricen for ligningssystemet

$$\begin{array}{rcl} x_1 & = & 8 \\ x_2 & = & 1 \\ x_3 & = & 3 \end{array}$$

som umiddelbart ses at have præcis den ene løsning

$$\mathbf{x} = \begin{pmatrix} 8 \\ 1 \\ 3 \end{pmatrix}.$$

Som allerede nævnt består Gauss-elimination i at anvende rækkeoperationer til at omdanne totalmatricen hørende til et lineært ligningssystem til en matrix på reduceret echelonform. Det tilhørende lineære ligningssystemer har da ifølge sætning 1.2.5 samme løsningsmængde, men kan nu umiddelbart løses. Den følgende sætning viser, at denne løsningsmetode altid virker.

Sætning 1.2.12 *Enhver matrix A kan omdannes til en matrix A' på reduceret echelonform ved at udføre et endeligt antal rækkeoperationer.*

Vi viser i sætning 2.5.11, at matricen A' er entydigt bestemt af matricen A , hvilket kræver lidt flere værktøjer, end vi lige nu har til rådighed. Beviset for sætning 1.2.12 består i at angive den algoritme, der omformer A til A' . Algoritmen er beskrevet ved induktion på antallet af søjler i A , og for at forstå den generelle algoritme i beviset er det en god idé sideløbende at læse eksempel 1.2.13-1.2.16, hvor algoritmen er anvendt på konkrete matricer.

Bevis Vi lader A være en $m \times n$ -matrix og viser ved induktion på antallet $n \geq 1$ af søjler, at A ved brug af rækkeoperationer kan omdannes til en matrix A' på reduceret

1 Lineære ligningssystemer

echelonform. Vi betragter først det basale tilfælde $n = 1$, hvor

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

er en enkelt søjle. Hvis $a_{i1} = 0$ for alle i , så er $A' = A$ allerede på reduceret echelonform af rang $r = 0$, og i modsat fald findes $1 \leq i \leq m$ med $a_{i1} \neq 0$. Vi anvender nu først en rækkeoperation af type M til at gange den i 'te række med a_{i1}^{-1} , hvorefter vi anvender en rækkeoperation af type T til at ombytte den i 'te række med den første række. Herved får vi en matrix B på formen

$$B = \begin{pmatrix} 1 \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix}.$$

Endelig anvender vi for alle $2 \leq i \leq m$ en rækkeoperation af type S til at addere $-b_{i1}$ gange den første række til den i 'te række, hvorved vi får matricen

$$A' = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

som er på reduceret echelonform af rang $r = 1$. Dette viser påstanden for $n = 1$.

Så vi antager induktivt, at påstanden er vist for $n = p - 1$ og beviser den for $n = p$. Hvis alle indgange i den første søjle er lig med 0, så er

$$A = \left(\begin{array}{c|c} 0 & B \end{array} \right),$$

hvor B er en $m \times (p - 1)$ -matrix. Den induktive antagelse viser, at der findes en følge af rækkeoperationer, som omdanner B til en matrix B' på reduceret echelonform, og den samme følge af rækkeoperationer omdanner da matricen A til matricen

$$A' = \left(\begin{array}{c|c} 0 & B' \end{array} \right),$$

der som ønsket er på reduceret echelonform. Hvis ikke alle indgange i den første søjle er lig med 0, så findes $1 \leq i \leq m$ med $a_{i1} \neq 0$, og vi kan da som i tilfældet $n = 1$ først anvende

1.2 Rækkeoperationer og Gauss-elimination

en rækkeoperation af type M til at gange den i 'te række med a_{i1}^{-1} og derefter anvende en rækkeoperation af type T til at ombytte den i 'te række med den første række. Endelig anvender vi for alle $2 \leq i \leq m$ en rækkeoperation af type S til at addere $-b_{i1}$ gange den første række til den i 'te række, hvorved vi får matricen

$$B = \left(\begin{array}{c|ccc} 1 & b_{12} & \cdots & b_{1p} \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array} \right),$$

hvor C er en $(m-1) \times (p-1)$ -matrix. Den induktive antagelse viser nu, at der findes en følge af rækkeoperationer, der omdanner C til en matrix C' på reduceret echelonform, og da alle indgange i den første søjle i B med undtagelse af $b_{11} = 1$ er lig med 0, så omdanner de tilsvarende rækkeoperationer $m \times p$ -matricen B til matricen

$$D = \left(\begin{array}{c|ccc} 1 & b_{12} & \cdots & b_{1p} \\ 0 & & & \\ \vdots & & C' & \\ 0 & & & \end{array} \right).$$

Endelig lader vi $1 \leq j_2 < \cdots < j_r \leq p$ nummerere de søjler i D , der indeholder de ledende indgange i C' , og anvender for alle $2 \leq s \leq r$ rækkeoperationen af type S , der adderer $-b_{1j_s}$ gange den s 'te række i D til den første række i D . Matricen A' , der fremkommer herved, er da på reduceret echelonform af rang r . Dette viser induktionsskridtet og dermed sætningen. $\square$

I de følgende eksempler anvender vi algoritmen, der er beskrevet i beviset for sætning 1.2.12, på totalmatricen $(A | \mathbf{b})$ hørende til et givet lineært ligningssystem. Herved får vi en matrix $(A' | \mathbf{b}')$ på reduceret echelonform, og vi kan da umiddelbart angive løsningsmængden for det tilhørende ligningssystem. Denne løsningsmængde er ifølge sætning 1.2.5 den samme som løsningsmængden for det oprindelige ligningssystem.

Eksempel 1.2.13 (Entydig løsning) Vi anvender Gauss-elimination til at bestemme løsningsmængden til det lineære ligningssystem

$$\begin{array}{rrcr} 2x_1 & -3x_2 & +x_3 & = & -1 \\ -x_1 & +x_2 & -x_3 & = & 2 \\ 3x_1 & +2x_2 & +2x_3 & = & 3. \end{array}$$

Vi opskriver ligningssystemets totalmatrix $(A | \mathbf{b})$ og anvender rækkeoperationer til at

1 Lineære ligningssystemer

omdanne denne til en matrix $(A' | \mathbf{b}')$ på reduceret echelonform.

$$\begin{aligned}
 (A | \mathbf{b}) &= \left(\begin{array}{ccc|c} 2 & -3 & 1 & -1 \\ -1 & 1 & -1 & 2 \\ 3 & 2 & 2 & 3 \end{array} \right) R_1 \leftrightarrow R_2 \\
 &\left(\begin{array}{ccc|c} -1 & 1 & -1 & 2 \\ \color{red}{2} & -3 & 1 & -1 \\ \color{red}{3} & 2 & 2 & 3 \end{array} \right) \begin{array}{l} \\ +2R_1 \\ +3R_1 \end{array} \\
 &\left(\begin{array}{ccc|c} -1 & \color{red}{1} & -1 & 2 \\ 0 & -1 & -1 & 3 \\ 0 & \color{red}{5} & -1 & 9 \end{array} \right) \begin{array}{l} +R_2 \\ \\ +5R_2 \end{array} \\
 &\left(\begin{array}{ccc|c} \color{red}{-1} & 0 & -2 & 5 \\ 0 & \color{red}{-1} & -1 & 3 \\ 0 & 0 & \color{red}{-6} & 24 \end{array} \right) \begin{array}{l} (-1) \cdot R_1 \\ (-1) \cdot R_2 \\ (-\frac{1}{6}) \cdot R_3 \end{array} \\
 &\left(\begin{array}{ccc|c} 1 & 0 & \color{red}{2} & -5 \\ 0 & 1 & \color{red}{1} & -3 \\ 0 & 0 & 1 & -4 \end{array} \right) \begin{array}{l} +(-2)R_3 \\ +(-1)R_3 \\ \end{array} \\
 (A' | \mathbf{b}') &= \left(\begin{array}{ccc|c} \color{blue}{1} & 0 & 0 & 3 \\ 0 & \color{blue}{1} & 0 & 1 \\ 0 & 0 & \color{blue}{1} & -4 \end{array} \right)
 \end{aligned}$$

Her har vi markeret de indgange, vi har ønsket at ændre, med rødt; og i den endelige matrix $(A' | \mathbf{b}')$ har vi markeret de tre ledende indgange med blå. Ligningssystemet, der har $(A' | \mathbf{b}')$ som totalmatrix, er nu

$$\begin{array}{rcl}
 x_1 & = & 3 \\
 x_2 & = & 1 \\
 x_3 & = & -4.
 \end{array}$$

Dette ligningssystem har tydeligvis præcis den ene løsning

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix}.$$

Ifølge sætning 1.2.5 er denne løsning også den entydige løsning til det oprindelige ligningssystem.

Eksempel 1.2.14 (Løsningsmængde med én parameter) Vi anvender Gauss-elimination til at bestemme løsningsmængden til det lineære ligningssystem

$$\begin{array}{rrcr} -x_1 & -2x_2 & -5x_3 & = & -3 \\ 2x_1 & +3x_2 & +8x_3 & = & 4 \\ 2x_1 & +6x_2 & +14x_3 & = & 10. \end{array}$$

Vi opskriver igen totalmatricen for ligningssystemet og omdanner den til en matrix på reduceret echelonform ved hjælp af rækkeoperationer.

$$\begin{aligned} (A | \mathbf{b}) &= \left(\begin{array}{ccc|c} -1 & -2 & -5 & -3 \\ \color{red}{2} & 3 & 8 & 4 \\ \color{red}{2} & 6 & 14 & 10 \end{array} \right) \begin{array}{l} \\ +2R_1 \\ +2R_1 \end{array} \\ &\left(\begin{array}{ccc|c} -1 & \color{red}{-2} & -5 & -3 \\ 0 & -1 & -2 & -2 \\ 0 & \color{red}{2} & 4 & 4 \end{array} \right) \begin{array}{l} \\ +(-2)R_2 \\ +2R_2 \end{array} \\ &\left(\begin{array}{ccc|c} \color{red}{-1} & 0 & -1 & 1 \\ 0 & \color{red}{-1} & -2 & -2 \\ 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{l} (-1) \cdot R_1 \\ (-1) \cdot R_2 \\ \end{array} \\ (A' | \mathbf{b}') &= \left(\begin{array}{ccc|c} \color{blue}{1} & 0 & 1 & -1 \\ 0 & \color{blue}{1} & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right). \end{aligned}$$

Vi har igen markeret de ledende indgange med blå og de indgange, vi ønsker at ændre, med rødt. Ligningssystemet, der har $(A' | \mathbf{b}')$ som totalmatrix, er nu

$$\begin{array}{rrcr} x_1 & & + & x_3 & = & -1 \\ & x_2 & + & 2x_3 & = & 2 \\ & & & 0 & = & 0. \end{array}$$

Vi ser heraf, at der for hver værdi t af x_3 , findes præcis en værdi af x_1 og af x_2 , der giver en løsning til ligningssystemet, nemlig

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1-t \\ 2-2t \\ t \end{pmatrix}.$$

Med andre ord er den fælles løsningsmængde for de to ligningssystemer

$$\left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1-t \\ 2-2t \\ t \end{pmatrix} \mid t \in \mathbb{F} \right\}.$$

Løsningsmængden er altså *parametriseret* ved parameteren $t \in \mathbb{F}$. Vi siger, at variabelen x_3 er en *fri variabel*, da den kan tilordnes en vilkårlig værdi, og bemærker, at den

1 Lineære ligningssystemer

svarer til den søjle i A' , der *ikke* indeholder en ledende indgang. Vi siger endvidere, at variablene x_1 og x_2 , der svarer til de søjler i A' , der indeholder en ledende indgang, er *ledende variable*.

Generelt kan de ledende variable udtrykkes entydigt ved de frie variable.

Eksempel 1.2.15 (Tom løsningsmængde) Vi anvender igen Gauss-elimination til at bestemme løsningsmængden til det lineære ligningssystem

$$\begin{array}{rrcr} -x_1 & -2x_2 & -5x_3 & = & -3 \\ 2x_1 & +3x_2 & +8x_3 & = & 4 \\ 2x_1 & +6x_2 & +14x_3 & = & 5. \end{array}$$

Dette ligningssystem har samme koefficientmatrix som ligningssystemet i eksempel 1.2.14, men konstant søjlen er en anden. Vi opskriver totalmatricen for ligningssystemet og omdanner denne til en matrix på reduceret echelonform.

$$\begin{aligned} (A | \mathbf{b}) &= \left(\begin{array}{ccc|c} -1 & -2 & -5 & -3 \\ \color{red}{2} & 3 & 8 & 4 \\ \color{red}{2} & 6 & 14 & 5 \end{array} \right) \begin{array}{l} \\ +2R_1 \\ +2R_1 \end{array} \\ &\left(\begin{array}{ccc|c} -1 & \color{red}{-2} & -5 & -3 \\ 0 & -1 & -2 & -2 \\ 0 & \color{red}{2} & 4 & -1 \end{array} \right) \begin{array}{l} \\ +(-2)R_2 \\ +2R_2 \end{array} \\ &\left(\begin{array}{ccc|c} \color{red}{-1} & 0 & -1 & 1 \\ 0 & \color{red}{-1} & -2 & -2 \\ 0 & 0 & 0 & \color{red}{-1} \end{array} \right) \begin{array}{l} (-1) \cdot R_1 \\ (-1) \cdot R_2 \\ (-1) \cdot R_3 \end{array} \\ &\left(\begin{array}{ccc|c} 1 & 0 & 1 & \color{red}{-1} \\ 0 & 1 & 2 & \color{red}{2} \\ 0 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} +R_3 \\ +(-2)R_3 \\ \end{array} \\ (A' | \mathbf{b}') &= \left(\begin{array}{ccc|c} \color{blue}{1} & 0 & 1 & 0 \\ 0 & \color{blue}{1} & 2 & 0 \\ 0 & 0 & 0 & \color{blue}{1} \end{array} \right). \end{aligned}$$

Vi har igen markeret de indgange, vi ønsker at ændre, med rødt, og vi har markeret de tre ledende indgange med blå. Ligningssystemet med $(A' | \mathbf{b}')$ som totalmatrix er nu

$$\begin{array}{rrcr} x_1 & & + & x_3 & = & 0 \\ & x_2 & + & 2x_3 & = & 0 \\ & & & 0 & = & 1. \end{array}$$

Dette ligningssystem har *ingen* løsninger, da den sidste ligning “ $0 = 1$ ” ikke er opfyldt uanset hvilke værdier, vi tilskriver de variable. Ifølge sætning 1.2.5 har det oprindelige ligningssystem derfor heller ingen løsninger.

1.2 Rækkeoperationer og Gauss-elimination

Generelt har et ligningssystem med totalmatrix $(A \mid \mathbf{b})$ ingen løsninger, hvis og kun hvis konstantsøjlen i matrixen $(A' \mid \mathbf{b}')$ på reduceret echelonform, der fremkommer fra $(A \mid \mathbf{b})$ ved rækkeoperationer, indeholder en ledende indgang.

Eksempel 1.2.16 (Løsningsmængde med to parametre) Vi anvender Gauss-elimination til at bestemme løsningsmængden til det lineære ligningssystem

$$\begin{array}{rrrrr} 2x_3 & -x_4 & +8x_5 & = & -13 \\ x_1 & -2x_2 & +3x_3 & +2x_4 & +x_5 & = & 10 \\ 3x_1 & -6x_2 & +10x_3 & +6x_4 & +5x_5 & = & 27 \end{array}$$

Så vi opskriver totalmatrixen for ligningssystemet og omdanner denne til en matrix på reduceret echelonform ved hjælp af rækkeoperationer.

$$\begin{aligned} (A \mid \mathbf{b}) &= \left(\begin{array}{ccccc|c} 0 & 0 & 2 & -1 & 8 & -13 \\ 1 & -2 & 3 & 2 & 1 & 10 \\ 3 & -6 & 10 & 6 & 5 & 27 \end{array} \right) R_1 \leftrightarrow R_2 \\ &\left(\begin{array}{ccccc|c} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 2 & -1 & 8 & -13 \\ \textcolor{red}{3} & -6 & 10 & 6 & 5 & 27 \end{array} \right) +(-3)R_1 \\ &\left(\begin{array}{ccccc|c} 1 & -2 & \textcolor{red}{3} & 2 & 1 & 10 \\ 0 & 0 & \textcolor{red}{2} & -1 & 8 & -13 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array} \right) \begin{array}{l} +(-3)R_3 \\ +(-2)R_3 \end{array} \\ &\left(\begin{array}{ccccc|c} 1 & -2 & 0 & \textcolor{red}{2} & -5 & 19 \\ 0 & 0 & 0 & -1 & 4 & -7 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array} \right) +2R_2 \\ &\left(\begin{array}{ccccc|c} 1 & -2 & 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & \textcolor{red}{-1} & 4 & -7 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array} \right) (-1) \cdot R_2 \\ &\left(\begin{array}{ccccc|c} 1 & -2 & 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 1 & -4 & 7 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array} \right) R_2 \leftrightarrow R_3 \\ (A' \mid \mathbf{b}') &= \left(\begin{array}{ccccc|c} \textcolor{blue}{1} & -2 & 0 & 0 & 3 & 5 \\ 0 & 0 & \textcolor{blue}{1} & 0 & 2 & -3 \\ 0 & 0 & 0 & \textcolor{blue}{1} & -4 & 7 \end{array} \right) \end{aligned}$$

Vi har markeret de indgange, vi ønsker at ændre, med rødt; og vi har markeret de tre ledende indgange med blåt. Ligningssystemet med $(A' \mid \mathbf{b}')$ som totalmatrix er nu

$$\begin{array}{rrrr} x_1 & -2x_2 & & +3x_5 & = & 5 \\ & & x_3 & & +2x_5 & = & -3 \\ & & & x_4 & -4x_5 & = & 7. \end{array}$$

1 Lineære ligningssystemer

Så for hver værdi $x_2 = t_1$ og $x_5 = t_2$ af de frie variable x_2 og x_5 er

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 5 & +2t_1 & -3t_2 \\ & t_1 & \\ -3 & & -2t_2 \\ 7 & & +4t_2 \\ & & t_2 \end{pmatrix}$$

en løsning. Den fælles løsningsmængde til de to ligningssystemer er parametriseret ved de to parametre $t_1 \in \mathbb{F}$ og $t_2 \in \mathbb{F}$.

Sætning 1.2.17 (Løsning af lineære ligningssystemer ved Gauss-elimination)

Løsningsmængden for det lineære ligningssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

bestående af m ligninger i n ubekendte over $\mathbb{F}$ bestemmes som følger:

(1) Ligningssystemets totalmatrix

$$(A \mid \mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right)$$

omdannes ved en følge af rækkeoperationer til en matrix $(A' \mid \mathbf{b}')$ på reduceret echelonform af rang $0 \leq r \leq \min\{m, n\}$.

- (2) Hvis en af de ledende indgange i $(A' \mid \mathbf{b}')$ er indeholdt i $\mathbf{b}'$, så har ligningssystemet ingen løsninger.
- (3) Hvis ingen af de ledende indgange i $(A' \mid \mathbf{b}')$ er indeholdt i $\mathbf{b}'$, og hvis $r = n$, så har ligningssystemet præcis én løsning.
- (4) Hvis ingen af de ledende indgange i $(A' \mid \mathbf{b}')$ er indeholdt i $\mathbf{b}'$, og hvis $r < n$, så parametriseres løsningsmængden ved at (i) opdele de n variable i de r ledende variable $x_{j_1}, \dots, x_{j_r}$ svarende til de søjler i A' , der indeholder en ledende indgang, og de resterende $p = n - r$ frie variable; (ii) tilordne de frie variable vilkårlige værdier $t_1, \dots, t_p \in \mathbb{F}$; og (iii) løse ligningssystemet med totalmatrix $(A' \mid \mathbf{b}')$ for at bestemme værdierne af de ledende variable.

1.2 Rækkeoperationer og Gauss-elimination

Bevis Vi har allerede vist i sætning 1.2.12, at der findes en følge af rækkeoperationer, der omdanner totalmatricen $(A \mid \mathbf{b})$ til en matrix $(A' \mid \mathbf{b}')$ på reduceret echelonform. Med andre ord kan skridtet (1) altid udføres. Desuden har vi vist i sætning 1.2.5, at det givne ligningssystem og ligningssystemet med totalmatrix $(A' \mid \mathbf{b}')$ har den samme løsningsmængde. I resten af beviset vil vi derfor udelukkende betragte det sidstnævnte ligningssystem.

Vi viser nu (2), så vi antager, at $\mathbf{b}'$ indeholder en af de ledende indgange i matricen $(A' \mid \mathbf{b}')$. Da denne $m \times (n+1)$ -matrix er på reduceret echelonform med $\mathbf{b}'$ som sidste søjle, gælder det nødvendigvis, at $j_r = n+1$, og derfor er $a'_{r,j} = 0$ for alle $1 \leq j \leq n$, mens $b'_r = 1$. Den r 'te ligning i det tilhørende ligningssystem er derfor

$$0 \cdot x_1 + \cdots + 0 \cdot x_n = 1,$$

hvilket som ønsket viser, at ligningssystemet ingen løsninger har.

Vi viser dernæst (3). Da $m \times (n+1)$ -matricen $(A' \mid \mathbf{b}')$ er på reduceret echelonform og har $r = n$ ledende indgange, og da den sidste søjle $\mathbf{b}'$ ikke indeholder nogen ledende indgang, må alle søjler i A' nødvendigvis indeholde en ledende indgang. De første $r = n$ ligninger i det tilhørende ligningssystem er da henholdsvis

$$\begin{array}{rcl} x_1 & & = b'_1 \\ & \ddots & \vdots \\ & & x_n = b'_n, \end{array}$$

mens de sidste $m - r$ ligninger alle er lig den trivielle ligning " $0 = 0$ ". Dette viser som ønsket, at ligningssystemet har en entydig løsning.

Vi viser endelig (4). Da $m \times (n+1)$ -matricen $(A' \mid \mathbf{b}')$ er på reduceret echelonform og har $r < n$ ledende indgange, og da den sidste søjle $\mathbf{b}'$ ikke indeholder nogen ledende indgang, konkluderer vi derfor, at r af de n søjler i A' indeholder en ledende indgang, mens de resterende $p = n - r$ søjler ikke indeholder en ledende indgang. Vi indicerer nu med $1 \leq j_1 < \cdots < j_r \leq n$ og $1 \leq k_1 < \cdots < k_p \leq n$ de søjler i A' , der henholdsvis indeholder og ikke indeholder en ledende indgang. De første r ligninger i ligningssystemet giver da

$$\begin{array}{rcl} x_{j_1} & = & b'_1 - (a'_{1,k_1}x_{k_1} + \cdots + a'_{1,k_p}x_{k_p}) \\ & \vdots & \\ x_{j_r} & = & b'_r - (a'_{r,k_1}x_{k_1} + \cdots + a'_{r,k_p}x_{k_p}), \end{array}$$

idet vi flytter de led, der indeholder en af de fri variable $x_{k_1}, \dots, x_{k_p}$, til højre side af lighedstegnet. De resterende $m - r$ ligninger er igen alle lig den trivielle ligning " $0 = 0$ ". Dette viser som ønsket, at for hver værdi $x_{k_1} = t_1, \dots, x_{k_p} = t_p$ af de frie variable, har ligningssystemet præcis én løsning. $\square$

Som allerede nævnt så viser vi i sætning 2.5.11, at matricen $(A' \mid \mathbf{b}')$ på reduceret echelonform og dens rang r er entydigt bestemt af matricen $(A \mid \mathbf{b})$ og ikke afhænger af valget af rækkeoperationer, der omdanner $(A \mid \mathbf{b})$ til $(A' \mid \mathbf{b}')$.

1 Lineære ligningssystemer

Eksempel 1.2.18 (Ligningssystem med komplekse tal) Vi anvender Gauss-elimination til at løse det lineære ligningssystem

$$\begin{array}{rrcr} ix_1 + (-1+i)x_2 & -x_3 & = & 4-i \\ -x_1 & -2ix_2 + (2-i)x_3 & = & -2+7i \end{array}$$

med to ligninger i tre variable over de komplekse tal $\mathbb{C}$. Vi opskriver ligningssystemets totalmatrix $(A \mid \mathbf{b})$ og omdanner den til en matrix $(A' \mid \mathbf{b}')$ på reduceret echelonform ved hjælp af rækkeoperationer.

$$\begin{aligned} (A \mid \mathbf{b}) &= \left(\begin{array}{ccc|c} i & -1+i & -1 & 4-i \\ -1 & -2i & 2-i & -2+7i \end{array} \right) +i \cdot R_2 \\ &\left(\begin{array}{ccc|c} 0 & 1+i & 2i & -3-3i \\ -1 & -2i & 2-i & -2+7i \end{array} \right) \frac{1}{2}(1-i) \cdot R_1 \\ &\left(\begin{array}{ccc|c} 0 & 1 & 1+i & -3 \\ -1 & -2i & 2-i & -2+7i \end{array} \right) +2i \cdot R_1 \\ &\left(\begin{array}{ccc|c} 0 & 1 & 1+i & -3 \\ -1 & 0 & i & -2+i \end{array} \right) (-1) \cdot R_2 \\ &\left(\begin{array}{ccc|c} 0 & 1 & 1+i & -3 \\ 1 & 0 & -i & 2-i \end{array} \right) R_1 \leftrightarrow R_2 \\ (A' \mid \mathbf{b}') &= \left(\begin{array}{ccc|c} 1 & 0 & -i & 2-i \\ 0 & 1 & 1+i & -3 \end{array} \right) \end{aligned}$$

Vi har som tidligere markeret de indgange, vi ønsker at ændre, med rødt og de ledende indgange i matricen $(A' \mid \mathbf{b}')$ med blå. Da $\mathbf{b}'$ ikke indeholder ledende indgange, har ligningssystemet løsninger; og da rangen $r = 2$ af $(A' \mid \mathbf{b}')$ er mindre end antallet $n = 3$ af variable, så er løsningsmængden uendelig og parametriseret af $p = n - r = 1$ parameter. Endelig aflæser vi umiddelbart fra matricen $(A' \mid \mathbf{b}')$, at for hver værdi $t \in \mathbb{C}$ af den frie variabel x_3 er der præcis den ene løsning

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} (2-i) + i \cdot t \\ -3 - (1+i) \cdot t \\ t \end{pmatrix}$$

til ligningssystemet. Vi understreger, at $t \in \mathbb{C}$ er en kompleks parameter.

1.3 Rang og løsningsmængde

Vi har hovedsageligt anvendt rækkeoperationer på totalmatricen $(A | \mathbf{b})$ hørende til et lineært ligningssystem, men vi kan også anvende rækkeoperationer til at omdanne ligningssystemets koefficientmatrix A til en matrix A' på reduceret echelonform. Vi afslutter dette kapitel med den følgende sætning om betydningen af rangen af matricen A' for løsningsmængden. Sætningen siger blandt andet, at hvis A er en kvadratisk matrix og A' har maksimal rang, så har ligningssystemet med totalmatrix $(A | \mathbf{b})$ altid præcis én løsning, uanset hvad konstant søjlen $\mathbf{b}$ er. Sætningen får stor betydning for vores forståelse af lineære afbildninger i kapitel 2.

Sætning 1.3.1 *Lad A være en $m \times n$ -matrix, der ved hjælp af rækkeoperationer kan omdannes til en $m \times n$ -matrix A' på reduceret echelonform af rang $0 \leq r \leq \min\{m, n\}$. Da gælder:*

- (1) *Ligningssystemet med totalmatrix $(A | \mathbf{b})$ har mindst én løsning for alle valg af konstant søjle $\mathbf{b}$, hvis og kun hvis $r = m$.*
- (2) *Ligningssystemet med totalmatrix $(A | \mathbf{b})$ har højst én løsning for alle valg af konstant søjle $\mathbf{b}$, hvis og kun hvis $r = n$.*
- (3) *Ligningssystemet med totalmatrix $(A | \mathbf{b})$ har netop én løsning for alle valg af konstant søjle $\mathbf{b}$, hvis og kun hvis $r = m = n$.*

Bevis Vi vælger en gang for alle en følge af rækkeoperationer, der omdanner A til A' .

Vi viser først (1). Hvis $r = m$, da indeholder hver række i A' en ledende indgang. Derfor gælder det for alle $\mathbf{b}$, at matricen $(A' | \mathbf{b}')$, der fremkommer fra $(A | \mathbf{b})$ ved at anvende den valgte følge af rækkeoperationer, er på reduceret echelonform, og at de ledende indgange i $(A' | \mathbf{b}')$ alle er indholdt i A' . Da $\mathbf{b}'$ altså ikke indeholder en ledende indgang, viser sætning 1.2.17 (3)–(4), at ligningssystemet har en løsning. Hvis $r < m$, da er de sidste $m - r$ rækker i A' alle lig med nul-rækken. Lad nu $\mathbf{b}' = (b'_i)$, hvor

$$b'_i = \begin{cases} 1 & \text{hvis } i = r + 1, \\ 0 & \text{hvis } i \neq r + 1. \end{cases}$$

Da er matricen $(A' | \mathbf{b}')$ på reduceret echelonform af rang $r + 1$, og da $\mathbf{b}'$ indeholder en ledende indgang, viser sætning 1.2.17 (2), at det tilhørende ligningssystem ingen løsninger har. Ifølge sætning 1.2.5 er det samme derfor tilfældet for ligningssystemet, hvis totalmatrix er den matrix $(A | \mathbf{b})$, der fås fra $(A' | \mathbf{b}')$ ved i omvendt rækkefølge at udføre de inverse rækkeoperationer svarende til rækkeoperationerne i den valgte følge af rækkeoperationer. Dette viser (1).

1 Lineære ligningssystemer

Vi viser dernæst (2). Hvis $r = n$, indeholder hver søjle i A' en ledende indgang. Lad nu $\mathbf{b}$ være et valg af konstantsøjle. Den valgte følge af rækkeoperationer omdanner matricen $(A | \mathbf{b})$ til en matrix $(A' | \mathbf{c})$, og vi kan da omdanne $(A' | \mathbf{c})$ til en matrix $(A' | \mathbf{b}')$ på reduceret echelonform ved om nødvendigt at anvende yderligere rækkeoperationer. Da hver søjle i A' indeholder en ledende indgang, er rangen r' af $(A' | \mathbf{b}')$ enten $r' = r$ eller $r' = r + 1$. Hvis $r' = r + 1$, da indeholder $\mathbf{b}'$ en ledende indgang, hvorfor sætning 1.2.17 (2) viser, at ligningssystemet med totalmatrix $(A | \mathbf{b})$ ikke har nogen løsninger. Og hvis $r' = r$, da indeholder $\mathbf{b}'$ ikke en ledende indgang, hvorfor sætning 1.2.17 (3) viser, at ligningssystemet med totalmatrix $(A | \mathbf{b})$ præcis har én løsning. Med andre ord gælder det for alle valg af $\mathbf{b}$, at ligningssystemet med totalmatrix $(A | \mathbf{b})$ højst har én løsning. Omvendt, hvis $r < n$, da gælder det for $\mathbf{b} = \mathbf{0}$, at ligningssystemet med totalmatrix $(A | \mathbf{b})$ har uendeligt mange løsninger. For dette følger fra sætning 1.2.17 (4), idet den valgte følge af rækkeoperationer omdanner $(A | \mathbf{0})$ til matricen $(A' | \mathbf{0})$, der er på reduceret echelonform af rang $r < n$ og ikke har nogen ledende indgange i konstantsøjlen. Dette viser (2), og endelig følger (3) ved at kombinere (1) og (2). $\square$

Lad A være en matrix, og lad A' være den matrix på reduceret echelonform, der fremkommer fra A ved hjælp af rækkeoperationer. Det følgende resultat viser, at hvis vi i stedet omdanner A til en matrix B på echelonform, hvilket normalt kræver mindre arbejde, da har A' og B samme rang. Hvis vi kun ønsker at bestemme rangen af A' , så er det derfor nok at udregne B . Vi bemærker, at modsat matricen A' så er matricen B ikke entydigt bestemt af A .

Lemma 1.3.2 *Enhver matrix B på echelonform af rang r kan ved endeligt mange rækkeoperationer omdannes til en matrix B' på reduceret echelonform af samme rang r .*

Bevis Lad B være en $m \times n$ -matrix på echelonform, og lad $1 \leq j_1 < \dots < j_r \leq n$ indicere de søjler i B , der indeholder en ledende indgang. For alle $1 \leq s \leq r$ anvender vi først en rækkeoperation af type M til at gange den s 'te række med $b_{sj_s}^{-1}$. Herved opnår vi en ny matrix C på echelonform af rang r med ledende indgange

$$c_{sj_s} = b_{sj_s}^{-1} \cdot b_{sj_s} = 1,$$

hvor $1 \leq s \leq r$. Dernæst anvender vi for alle $2 \leq s \leq r$ og $1 \leq i \leq s - 1$ den rækkeoperation af type S, der adderer $-c_{ij_s}$ gange den s 'te række i C til den i 'te række i C . Herved opnår vi en matrix B' på reduceret echelonform af rang r med ledende indgange

$$b'_{sj_s} = c_{sj_s} = 1,$$

hvor $1 \leq s \leq r$, hvilket viser lemmaet. $\square$

Sætning 1.2.12 viser, at totalmatricen $(A | \mathbf{b})$ for et givet lineært ligningssystem ved rækkeoperationer kan omdannes til en matrix $(A' | \mathbf{b}')$ på reduceret echelonform, og sætning 1.2.17 giver en fuldstændig beskrivelse af ligningssystemets løsningmængde

1.3 Rang og løsningsmængde

udtrykt ved matricen $(A' | \mathbf{b}')$. Hvis vi imidlertid kun ønsker at bestemme kvalitative egenskaber ved løsningsmængden såsom at bestemme, hvorvidt den er tom eller ej, eller bestemme antallet $p = n - r$ af parametre for løsningsmængden, så kan vi ifølge lemma 1.3.2 nøjes med at omdanne $(A | \mathbf{b})$ til en matrix $(B | \mathbf{c})$ på echelonform. Da dette normalt kræver færre rækkeoperationer end at omdanne $(A | \mathbf{b})$ til en matrix på reduceret echelonform, så kan man herved spare sig lidt arbejde.

Eksempel 1.3.3 (Echelon/reduceret echelonform) Betragt ligningssystemet

$$\begin{array}{rrcr} 2x_1 + 3x_2 & -4x_3 & -x_4 & = & 12 \\ -2x_1 & & +10x_3 + 3x_4 & = & 4 \\ 2x_1 + 6x_2 & +2x_3 & & = & 2. \end{array}$$

Vi omdanner først dets totalmatrix $(A | \mathbf{b})$ til en matrix $(B | \mathbf{c})$ på echelonform.

$$\begin{aligned} (A | \mathbf{b}) &= \left(\begin{array}{cccc|c} 2 & 3 & -4 & -1 & 12 \\ -2 & 0 & 10 & 3 & 4 \\ 2 & 6 & 2 & 0 & 2 \end{array} \right) \begin{array}{l} \\ +R_1 \\ +(-1) \cdot R_1 \end{array} \\ &= \left(\begin{array}{cccc|c} 2 & 3 & -4 & -1 & 12 \\ 0 & 3 & 6 & 2 & 16 \\ 0 & 3 & 6 & 1 & -10 \end{array} \right) +(-1) \cdot R_2 \\ (B | \mathbf{c}) &= \left(\begin{array}{cccc|c} 2 & 3 & -4 & -1 & 12 \\ 0 & 3 & 6 & 2 & 16 \\ 0 & 0 & 0 & -1 & -26 \end{array} \right) \end{aligned}$$

Vi har her som tidligere markeret de ledende indgange med blåt. Hvis vi yderligere omdanner $(B | \mathbf{c})$ til en matrix $(A' | \mathbf{b}')$ på reduceret echelonform, så vil positionen af de ledende indgange ikke ændres. Derfor kan løsningsmængden parametriseres ved

$$p = n - r = 4 - 3 = 1$$

parameter. Hvis vi ønsker at bestemme løsningsmængden mere præcist, så må vi først bestemme matricen $(A' | \mathbf{b}')$, hvilket vi i dette tilfælde overlader til læseren.

Eksempel 1.3.4 Vi betragter 3×3 -matricen

$$A = \begin{pmatrix} -1 & -2 & -5 \\ 2 & 3 & 8 \\ 2 & 6 & 14 \end{pmatrix}$$

og undersøger først om, der findes en konstantsøjle $\mathbf{b}$ sådan, at ligningssystemet med totalmatrix $(A | \mathbf{b})$ ikke har nogen løsninger. Vi anvender derfor rækkeoperationer til

1 Lineære ligningssystemer

at omdanne A til en matrix B på echelonform.

$$A = \begin{pmatrix} -1 & -2 & -5 \\ 2 & 3 & 8 \\ 2 & 6 & 14 \end{pmatrix} \begin{matrix} \\ +2R_1 \\ +2R_1 \end{matrix}$$
$$\begin{pmatrix} -1 & -2 & -5 \\ 0 & -1 & -2 \\ 0 & 2 & 4 \end{pmatrix} +2R_2$$
$$B = \begin{pmatrix} -1 & -2 & -5 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

Da B er på echelonform af rang $r = 2 < 3 = m$, konkluderer vi fra sætning 1.3.1 (1), at en sådan konstantsøjle $\mathbf{b}$ findes. Endvidere fortæller beviset for sætningen os, hvordan vi bærer os ad med at finde et sådant $\mathbf{b}$:

$$(B | \mathbf{c}) = \left(\begin{array}{ccc|c} -1 & -2 & -5 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) +(-2)R_2$$
$$\left(\begin{array}{ccc|c} -1 & -2 & -5 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 2 & 4 & 1 \end{array} \right) \begin{matrix} \\ +(-2)R_1 \\ +(-2)R_1 \end{matrix}$$
$$(A | \mathbf{b}) = \left(\begin{array}{ccc|c} -1 & -2 & -5 & 0 \\ 2 & 3 & 8 & 0 \\ 2 & 6 & 14 & 1 \end{array} \right)$$

Ligningssystemet med totalmatrix $(B | \mathbf{c})$ har tydeligvis ingen løsninger, og det samme er derfor tilfældet for ligningssystemet med totalmatrix $(A | \mathbf{b})$. Her har vi udført de inverse af de rækkeoperationer, vi benyttede til at omdanne A til B .

1.4 Opgaver

1.1 Omform ved hjælp af rækkeoperationer matricen

$$A = \begin{pmatrix} 1 & -2 & 3 & 2 & 1 & 10 \\ 2 & -4 & 8 & 3 & 10 & 7 \\ 3 & -6 & 10 & 6 & 5 & 27 \end{pmatrix}$$

til en matrix A' på reduceret echelonform.

1.2 Omform ved hjælp af rækkeoperationer matricen

$$B = \begin{pmatrix} 1 & 2 & 1 & 4 \\ 3 & 8 & 7 & 20 \\ 2 & 7 & 9 & 23 \end{pmatrix}$$

til en matrix B' på reduceret echelonform.

1.3 Omform ved hjælp af rækkeoperationer matricen

$$C = \begin{pmatrix} 2+i & 0 & -4-i \\ 5 & 1-3i & -2+2i \\ 0 & 3+i & 3 \end{pmatrix}$$

til en matrix C' på reduceret echelonform.

1.4 Løs følgende tre reelle lineære ligningssystemer:

$$\text{a) } \begin{cases} 2x_1 - x_2 + x_3 = 3 \\ -x_1 + 2x_2 + 4x_3 = 6 \\ x_1 + x_2 + 5x_3 = 9 \end{cases}$$

$$\text{b) } \begin{cases} 2x_1 - x_2 + x_3 = 4 \\ -x_1 + 2x_2 + 4x_3 = 6 \\ x_1 + x_2 + 5x_3 = 9 \end{cases}$$

$$\text{c) } \begin{cases} 2x_1 - x_2 + 2x_3 = 4 \\ -x_1 + 2x_2 + 4x_3 = 6 \\ x_1 + x_2 + 5x_3 = 9 \end{cases}$$

1.5 Løs det reelle lineære ligningssystem

$$\begin{aligned} x + y + 2z &= 3 \\ 2x - y + 4z &= 0 \\ x + 3y - 2z &= 3 \\ -3x - 2y + z &= 0. \end{aligned}$$

1.6 Løs det reelle lineære ligningssystem

$$\begin{aligned} 2x_1 + 4x_2 - x_3 - 2x_4 + 2x_5 &= 6 \\ x_1 + 3x_2 + 2x_3 - 7x_4 + 3x_5 &= 9 \\ 5x_1 + 8x_2 - 7x_3 + 6x_4 + x_5 &= 4. \end{aligned}$$

1.7 Løs det komplekse ligningssystem

$$\begin{aligned} ix_1 + 2x_2 &= 1 \\ (1 + 2i)x_1 + (2 + 2i)x_2 &= 3i. \end{aligned}$$

1 Lineære ligningssystemer

1.8 Løs det komplekse ligningssystem

$$\begin{aligned}ix_1 + (-2 + i)x_2 - 2x_3 &= 1 \\2x_1 + 6ix_2 + (4 + 2i)x_3 &= 2i \\3ix_1 + (-5 + 2i)x_2 + (-7 + i)x_3 &= 3.\end{aligned}$$

1.9 Løs det komplekse ligningssystem

$$\begin{aligned}(1 - i)x_1 + ix_2 + 3x_3 &= 0 \\2ix_2 + 2x_3 &= 0 \\2x_1 + (1 - i)x_2 + (1 + i)x_3 &= 0.\end{aligned}$$

1.10 Vis først, at det reelle ligningssystem

$$\begin{aligned}x + y - z &= 2 \\2x + y + z &= a \\x + 2z &= 3\end{aligned}$$

ikke har nogen løsning (x, y, z) , hvis $a \neq 5$. Vis dernæst, at ligningssystemet har uendeligt mange løsninger, hvis $a = 5$, og angiv disse.

1.11 Vis, at det komplekse ligningssystem

$$\begin{aligned}x - 2y + (3 + i)z &= -2 + 3i \\y + 2z &= a \\ix + (1 - 2i)y + (1 + 3i)z &= -4 - i\end{aligned}$$

ikke har nogen løsning (x, y, z) , hvis $a \neq -1 + i$. Vis dernæst, at ligningssystemet har uendeligt mange løsninger, hvis $a = -1 + i$, og angiv disse.

1.12 (★) Find alle reelle løsninger til ligningen

$$7x_1 + 4x_2 - 13x_3 + x_4 + 6x_5 - x_6 = 30.$$

1.13 (★) Vis, at $(x, y) = (0, 0)$ er den eneste løsning til ligningssystemet

$$\begin{aligned}ax + by &= 0 \\cx + dy &= 0,\end{aligned}$$

hvis og kun hvis $ad - bc \neq 0$.

1.14 Betragt et lineært ligningssystem med 5 ligninger i 6 ubekendte.

- a. Er det muligt at ligningssystemet ingen løsning har?

- b. Har ligningssystemet nødvendigvis ingen løsning?
- c. Er det muligt at ligningssystemet har præcist én løsning?
- d. Er det muligt at ligningssystemet har uendeligt mange løsninger?
- e. Hvad bliver svarene til spørgsmål a–d hvis vi i stedet betragter et lineært ligningssystem med 6 ligninger i 5 ubekendte?

1.15 Afgør, hvilke af følgende afbildninger der er lineære, og angiv i givet fald matricen, der repræsenterer afbildningen med hensyn til standardbaserne.

- a. Afbildningen $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ defineret ved

$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 + 1 \\ x_1 \\ x_1 + x_2 \end{pmatrix}.$$

- b. Afbildningen $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ defineret ved

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 \\ x_3 - x_1 \end{pmatrix}.$$

- c. Afbildningen $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defineret ved

$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 + x_1 \\ \sin(x_2) \end{pmatrix}.$$

1.16 (a) Find samtlige løsninger til ligningssystemet

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + 6x_3 + x_4 &= 0 \\ x_1 - x_2 + 5x_3 + x_4 &= 0 \\ 2x_1 + x_2 + 7x_3 + x_4 &= 0. \end{aligned}$$

(b) Bestem et reelt tal a så at ligningssystemet

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 + 6x_3 + x_4 &= a \\ x_1 - x_2 + 5x_3 + x_4 &= a \\ 2x_1 + x_2 + 7x_3 + x_4 &= 2a \end{aligned}$$

har løsninger, og finde disse løsninger for den fundne værdi af a .

1 Lineære ligningssystemer

- 1.17 (★) Find et kubisk polynomium $p(x) = ax^3 + bx^2 + cx + d$, som opfylder, at $p(0) = -1$, $p'(0) = 5$, $p(1) = 3$, $p'(1) = 1$.
- 1.18 (★) Betragt afbildningen $R_\theta: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, som roterer omkring x -aksen med vinklen $\theta \in [0, 2\pi)$.
- Argumenter geometrisk for, at R_θ er en lineær afbildning.
 - Find en matrix A_θ , som opfylder at $R_\theta(\mathbf{x}) = A_\theta \mathbf{x}$ for alle $\mathbf{x} \in \mathbb{R}^3$.
 - Gør det samme for rotationen omkring y -aksen og z -aksen.

2 Matricer og lineære afbildninger

I dette kapitel indfører vi matricer og deres algebra, som vi allerede har brugt i kapitel 1. Vi indfører også vektorrummene $\mathbb{F}^m$ af søjlevektorer og endelig lineære afbildninger mellem disse. Disse begreber er alle tæt forbundne: For mens en lineær afbildning $f: \mathbb{F} \rightarrow \mathbb{F}$ er givet ved venstre multiplikation med en skalar a , så er en lineær afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ givet ved venstre multiplikation med en $m \times n$ -matrix A , som vi tænker på som “hældningskoefficienten” af den lineære afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$. Vi skal se, at Gauss-elimination er det centrale værktøj, hvad beregninger angår, samt fortolke nogle af resultaterne fra kapitel 1 som sætning om lineære afbildninger. Vi vil arbejde over et generelt legeme $\mathbb{F}$, men læseren er velkommen til at tænke på, at $\mathbb{F}$ er enten $\mathbb{R}$ eller $\mathbb{C}$.

2.1 Matricer

Vi har allerede brugt matricer, så det er nu på høje tid, at vi definerer dem ordentligt. Vi siger, at en afbildning $x: I \rightarrow X$ fra en mængde I til en mængde X er en *familie* af elementer i X indexeret ved I , og vi skriver $(x_i)_{i \in I}$, hvor $x_i = x(i)$.

Definition 2.1.1 Lad m og n være naturlige tal. En $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$ er en familie $A = (a_{ij})$ af elementer i $\mathbb{F}$ indexeret ved mængden af par (i, j) af naturlige tal $1 \leq i \leq m$ og $1 \leq j \leq n$. Vi siger, at elementet a_{ij} er den (i, j) 'te indgang i A . Mængden bestående af alle $m \times n$ -matricer med indgange i $\mathbb{F}$ betegnes $M_{m,n}(\mathbb{F})$.

Om nødvendigt skriver vi $a_{i,j}$ i stedet for a_{ij} , og vi skriver endvidere

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

for at anskueliggøre matricen $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$. Bemærk, at der ikke er kommaer mellem indgangene i matricen. Givet $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$, kalder vi $1 \times n$ -matricen

$$(a_{i1} \ a_{i2} \ \dots \ a_{in}) \in M_{1,n}(\mathbb{F})$$

2 Matricer og lineære afbildninger

for den i 'te række i A , mens vi kalder $m \times 1$ -matricen

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in M_{m,1}(\mathbb{F})$$

for den j 'te søjle i A . Hvis A er en $m \times n$ -matrix, så siger vi også, at A er en matrix af dimensioner $m \times n$, hvilket læses " m gange n ". Vi bemærker, at m og n er antallet af henholdsvis rækker og søjler i A . Hvis $m = n$, så siger vi, at matricen A er *kvadratisk* af orden n , og vi skriver $M_n(\mathbb{F})$ i stedet for $M_{n,n}(\mathbb{F})$ for mængden kvadratiske matricer af orden n med indgange i $\mathbb{F}$.

Eksempel 2.1.2 Matricerne

$$A = \begin{pmatrix} 2 & -1 & 0 & 3 \\ 3 & 2 & 0 & 0 \end{pmatrix} \in M_{2,4}(\mathbb{R}) \quad \text{og} \quad B = \begin{pmatrix} 1/2 + i/2 & 1/2 - i/2 \\ 1/2 - i/2 & 1/2 + i/2 \end{pmatrix} \in M_2(\mathbb{C})$$

er henholdsvis en 2×4 -matrix med indgange i $\mathbb{R}$ og en kvadratisk matrix af orden 2 med indgange i $\mathbb{C}$.

Eksempel 2.1.3 (1) Nulmatricen af dimensioner $m \times n$ er matricen

$$O_{m,n} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in M_{m,n}(\mathbb{F}),$$

hvis indgange alle er lig med 0. Vi benytter forkortelsen $O = O_{m,n}$, hvis m og n fremgår fra sammenhængen.

(2) Den *modsatte* af matricen $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$ er matricen $-A = (-a_{ij}) \in M_{m,n}(\mathbb{F})$ af samme dimensioner, hvis indgange er de modsatte elementer af indgangene i den oprindelige matrix.

(3) *Identitetsmatricen* $I_n \in M_n(\mathbb{F})$ er den kvadratiske matrix af orden n , hvis (i,j) 'te indgang er Kroneckers δ_{ij} , som er enten 1 eller 0, eftersom $i = j$ eller $i \neq j$.

$$I_1 = (1), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \dots \quad I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Vi forkorter ofte $I = I_n$, hvis n fremgår fra sammenhængen.

Vi indfører nu de fundamentale aritmetiske operationer på matricer.

Definition 2.1.4 (1) *Summen* af to matricer

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix},$$

der begge har dimensioner $m \times n$, er matricen

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix},$$

der ligeledes har dimensioner $m \times n$.

(2) *Produktet* af to matricer

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{pmatrix}$$

af dimensioner henholdsvis $m \times n$ og $n \times p$, er matricen

$$A \cdot B = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mp} \end{pmatrix}$$

hvis (i, k) 'te indgang er defineret ved

$$c_{ik} = \sum_{j=1}^n a_{ij}b_{jk} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk}.$$

Produktmatricen $A \cdot B$ har dimensioner $m \times p$.

Vi understreger, at summen af to matricer A og B kun er defineret, hvis A og B har de samme dimensioner, og at produktet af to matricer A og B kun er defineret, hvis antallet af søjler i A er lig antallet af rækker i B . Det har således ikke mening at tale

2 Matricer og lineære afbildninger

om $A + B$ eller $A \cdot B$, medmindre dimensionerne af A og B er som foreskrevet. Som for multiplikation af skalarer, forkorter vi normalt og skriver AB i stedet for $A \cdot B$.

Eksempel 2.1.5 Summen af matricerne

$$A = \begin{pmatrix} 2 & 0 & 4 \\ 1 & -3 & 6 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 1 & -7 & 4 \\ 1 & 2 & 0 \end{pmatrix}$$

findes, da de to matricer har samme dimensioner 2×3 , og er givet ved

$$\begin{aligned} A + B &= \begin{pmatrix} 2 & 0 & 4 \\ 1 & -3 & 6 \end{pmatrix} + \begin{pmatrix} 1 & -7 & 4 \\ 1 & 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2+1 & 0+(-7) & 4+4 \\ 1+1 & -3+2 & 6+0 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -7 & 8 \\ 2 & -1 & 6 \end{pmatrix}. \end{aligned}$$

Her har vi markeret indgangene i A og B med henholdsvis blå og sort.

Matrixprodukt kræver nok lidt mere tilvænning end matrixsum, så vi forklarer her, hvordan det kan anskueliggøres. Givet en $1 \times n$ -matrix og en $n \times 1$ -matrix

$$A = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n) \quad \text{og} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

da er matrixproduktet AB defineret, og det er lig med 1×1 -matricen

$$AB = (\mathbf{a}_1 b_1 + \mathbf{a}_2 b_2 + \dots + \mathbf{a}_n b_n).$$

Generelt, for en $m \times n$ -matrix A og en $n \times p$ -matrix B , kan vi betragte deres produkt som værende den $m \times p$ -matricen $AB = C$, hvis indgange er alle de mulige produkter af en række i A med en søjle i B :

$$\begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & & & \vdots \\ \boxed{a_{i1}} & \dots & \boxed{a_{ij}} & \dots & \boxed{a_{in}} \\ \vdots & & & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & \boxed{b_{1k}} & \dots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{j1} & & \boxed{b_{jk}} & & b_{jp} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \dots & \boxed{b_{nk}} & \dots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1k} & \dots & c_{1p} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \dots & \boxed{c_{ik}} & \dots & c_{ip} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \dots & c_{mk} & \dots & c_{mp} \end{pmatrix}$$

Da der er m rækker i A og p søjler i B , er der altså $m \times p$ indgange i $AB = C$.

Når man udregner et matrixprodukt AB , kan det hjælpe at tegne kasser omkring rækkerne i A og søjlerne i B , som vi har gjort ovenfor. Vi fortsætter denne praksis i de følgende eksempler.

Eksempel 2.1.6 Produktet af matricerne

$$A = \begin{pmatrix} 2 & 7 \\ -1 & 3 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 2 & -1 & 0 \\ 4 & 1 & 2 \end{pmatrix}$$

er defineret, da antallet søjler i A er lig med antallet af rækker i B , og er givet ved

$$\begin{aligned} A \cdot B &= \begin{pmatrix} \boxed{2} & \boxed{7} \\ \boxed{-1} & \boxed{3} \end{pmatrix} \begin{pmatrix} \boxed{2} & \boxed{-1} & \boxed{0} \\ \boxed{4} & \boxed{1} & \boxed{2} \end{pmatrix} = \begin{pmatrix} 2 \cdot 2 + 7 \cdot 4 & 2 \cdot (-1) + 7 \cdot 1 & 2 \cdot 0 + 7 \cdot 2 \\ (-1) \cdot 2 + 3 \cdot 4 & (-1) \cdot (-1) + 3 \cdot 1 & (-1) \cdot 0 + 3 \cdot 2 \end{pmatrix} \\ &= \begin{pmatrix} 32 & 5 & 14 \\ 10 & 4 & 6 \end{pmatrix} \end{aligned}$$

Her har vi markeret indgangene i den første række i A med rødt og indgangene i den anden række med blå. Vi bemærker, at produktet $B \cdot A$ ikke er defineret, da antallet af søjler i B er forskelligt fra antallet af rækker i A .

Eksempel 2.1.7 Produktet af matricerne

$$A = \begin{pmatrix} 1 & i+1 & 2i-1 \\ 4 & 3i & 8 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 2i \\ -3 \\ 0 \end{pmatrix}$$

er defineret, da antallet af søjler i A er lig med antallet af rækker i B , og er givet ved

$$\begin{aligned} A \cdot B &= \begin{pmatrix} \boxed{1} & \boxed{i+1} & \boxed{2i-1} \\ \boxed{4} & \boxed{3i} & \boxed{8} \end{pmatrix} \begin{pmatrix} \boxed{2i} \\ \boxed{-3} \\ \boxed{0} \end{pmatrix} = \begin{pmatrix} 1 \cdot 2i + (i+1) \cdot (-3) + (2i-1) \cdot 0 \\ 4 \cdot 2i + 3i \cdot (-3) + 8 \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 2i - 3i - 3 + 0 \\ 8i - 9i + 0 \end{pmatrix} = \begin{pmatrix} -3 - i \\ -i \end{pmatrix} \end{aligned}$$

hvor vi igen brugt rødt og blå for rækkerne i A .

Eksempel 2.1.8 Produktmatricen

$$AB = \begin{pmatrix} \boxed{1} & \boxed{0} & \boxed{3} \end{pmatrix} \begin{pmatrix} \boxed{2} \\ \boxed{4} \\ \boxed{1} \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 0 \cdot 4 + 3 \cdot 1 \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix}$$

2 Matricer og lineære afbildninger

er en 1×1 -matrix, da der er 1 række i A og 1 søjle i B , mens produktmatricen

$$BA = \begin{pmatrix} \boxed{2} \\ \boxed{4} \\ \boxed{1} \end{pmatrix} \left(\boxed{1} \mid \boxed{0} \mid \boxed{3} \right) = \begin{pmatrix} 2 \cdot 1 & 2 \cdot 0 & 2 \cdot 3 \\ 4 \cdot 1 & 4 \cdot 0 & 4 \cdot 3 \\ 1 \cdot 1 & 1 \cdot 0 & 1 \cdot 3 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 6 \\ 4 & 0 & 12 \\ 1 & 0 & 3 \end{pmatrix}$$

er en 3×3 -matrix, da der er 3 rækker i B og 3 søjler i A . Vi bemærker, at det ikke er muligt at sammenligne matricerne AB og BA , da deres dimensioner er forskellige. Så spørgsmålet om, hvorvidt $AB = BA$, er meningsløst.

Eksempel 2.1.9 I kapitel 1 har vi betragtet lineære ligningssystemer

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

med $a_{ij}, b_i \in \mathbb{F}$. Vi betragter nu matricerne

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{og} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

som har dimensioner henholdsvis $m \times n$, $n \times 1$ og $m \times 1$. Produktmatricen

$$A\mathbf{x} = \begin{pmatrix} \boxed{a_{11} \quad a_{12} \quad \cdots \quad a_{1n}} \\ \boxed{a_{21} \quad a_{22} \quad \cdots \quad a_{2n}} \\ \boxed{\vdots \quad \vdots \quad \ddots \quad \vdots} \\ \boxed{a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}} \end{pmatrix} \left(\begin{pmatrix} \boxed{x_1} \\ \boxed{x_2} \\ \boxed{\vdots} \\ \boxed{x_n} \end{pmatrix} \right) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

er derfor en $m \times 1$ -matrix ligesom $\mathbf{b}$. Det oprindelige lineære ligningssystem kan altså ækvivalent skrives som ligning af $m \times 1$ -matricer

$$A\mathbf{x} = \mathbf{b},$$

og vi vil i det følgende hovedsageligt benytte denne korte skrivemåde.

Sætning 2.1.10 For matricer af passende dimensioner med indgange i legeme $\mathbb{F}$ gælder følgende identiteter:

$$(A1) \quad (A + B) + C = A + (B + C)$$

$$(A2) \quad A + O = A = O + A$$

$$(A3) \quad A + (-A) = O = (-A) + A$$

$$(A4) \quad A + B = B + A$$

$$(P1) \quad (AB)C = A(BC)$$

$$(P2) \quad A I_n = A = I_m A$$

$$(D1) \quad A(B + C) = AB + AC$$

$$(D2) \quad (A + B)C = AC + BC$$

$$(E1) \quad A O_{n,n} = O_{m,n} = O_{m,m} A$$

$$(E2) \quad A(-I_n) = -A = (-I_m)A$$

Bevis Ifølge definition 0.2.1 og sætning 0.2.3 gælder disse identiteter for addition og multiplikation af elementer i $\mathbb{F}$, og vi benytter nu dette til at vise, at identiteterne også gælder for addition og multiplikation af matricer med indgange i $\mathbb{F}$.

Identiteterne (A1)–(A4) omhandler matricer af samme dimensioner og følger umiddelbart af de tilsvarende identiteter (A1)–(A4) for skalarer, idet addition for matricer per definition er givet ved addition i $\mathbb{F}$ af de respektive indgange.

Identiteten (P1) omhandler matricer A , B og C af dimensioner $m \times n$, $n \times p$ og $p \times q$. Produkt matricerne $(AB)C$ og $A(BC)$ er begge af dimensioner $m \times q$, og vi ønsker at vise, at deres (i, l) 'te indgange er identiske. På den ene side er den (i, l) 'te indgang i $(AB)C$ lig med $\sum_{k=1}^p e_{ik} c_{kl}$, hvor $e_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$ er den (i, k) 'te indgang i AB ; og på den anden side er den (i, l) 'te indgang i $A(BC)$ lig med $\sum_{j=1}^n a_{ij} d_{jl}$, hvor $d_{jl} = \sum_{k=1}^p b_{jk} c_{kl}$ er den (j, l) 'te indgang i BC . Vi skal derfor vise, at der gælder følgende identitet af skalarer:

$$\sum_{k=1}^p \left(\sum_{j=1}^n a_{ij} b_{jk} \right) c_{kl} = \sum_{j=1}^n a_{ij} \left(\sum_{k=1}^p b_{jk} c_{kl} \right).$$

For at være helt præcis, så definerer vi her de itererede summer til at være

$$\sum_{j=1}^n a_{ij} b_{jk} = (\dots((a_{i1} b_{1k} + a_{i2} b_{2k}) + a_{i3} b_{3k}) + \dots + a_{in} b_{nk}),$$

2 Matricer og lineære afbildninger

etc. Det følger dog fra (A1) for skalarer, at ethvert valg af summationsorden vil tilordne den samme værdi til den itererede sum. Identiteten mellem de (i, l) 'te indgange i $(AB)C$ og $A(BC)$ fås nu som følger:

$$\begin{aligned} \sum_{k=1}^p \left(\sum_{j=1}^n a_{ij} b_{jk} \right) c_{kl} &= \sum_{k=1}^p \sum_{j=1}^n (a_{ij} b_{jk}) c_{kl} = \sum_{k=1}^p \sum_{j=1}^n a_{ij} (b_{jk} c_{kl}) \\ &= \sum_{j=1}^n \sum_{k=1}^p a_{ij} (b_{jk} c_{kl}) = \sum_{j=1}^n a_{ij} \left(\sum_{k=1}^p b_{jk} c_{kl} \right) \end{aligned}$$

Her følger den første identitet fra (D2) for skalarer; den anden identitet fra (P1) for skalarer; og den tredje identitet fra (A1) og (A4) for skalarer; og den sidste identitet fra (D1) for skalarer. Dette beviser, at (P1) gælder for matricer.

Vi beviser dernæst identiteten $AI_n = A$ i (P2); beviset for identiteten $I_m A = A$ er helt tilsvarende. Identiteten $AI_n = A$ omhandler en matrix A af dimensioner $m \times n$ og identitetsmatricen I_n af dimensioner $n \times n$. Vi husker på fra eksempel 2.1.3(3), at den (j, k) 'te indgang δ_{jk} i I_n er lig 1, hvis $j = k$, og lig med 0, hvis $j \neq k$. På (i, k) 'te indgange svarer identiteten $AI_n = A$ nu til følgende identitet af skalarer:

$$\sum_{j=1}^n a_{ij} \delta_{jk} = a_{ik}.$$

så identiteten følger fra (A2), (P2) og (E1) for skalarer. Vi har hermed bevist, at (P2) gælder for matricer.

Vi beviser dernæst identiteten (D1); identiteten (D2) bevises helt tilsvarende. Denne identitet omhandler en $m \times n$ -matrix A og $n \times p$ -matricer B og C og svarer på (i, k) 'te indgange til følgende identitet af skalarer:

$$\sum_{j=1}^n a_{ij} (b_{jk} + c_{jk}) = \sum_{j=1}^n a_{ij} b_{jk} + \sum_{j=1}^n a_{ij} c_{jk}.$$

Denne identitet følger umiddelbart fra (A1), (A4) og (D1) for skalarer, hvilket beviser, at (D1) også gælder for matricer.

Endelig fås identiteterne (E1) og (E2) fra de øvrige identiteter ved at repetere beviset for sætning 0.2.3. Vi har hermed bevist sætningen. $\square$

Eksempel 2.1.11 Vi betragter 2×2 -matricerne

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

og udregner de to produkter AB og BA , der begge findes og er af dimensioner 2×2 .

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad BA = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Da matricerne AB og BA er af samme dimensioner, har det mening at spørge, om de er ens. Men udregningen ovenfor viser, at $AB \neq BA$.

2.2 Vektorrummet $\mathbb{F}^m$

Vi lader igen $\mathbb{F}$ være et legeme, for eksempel $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og definerer

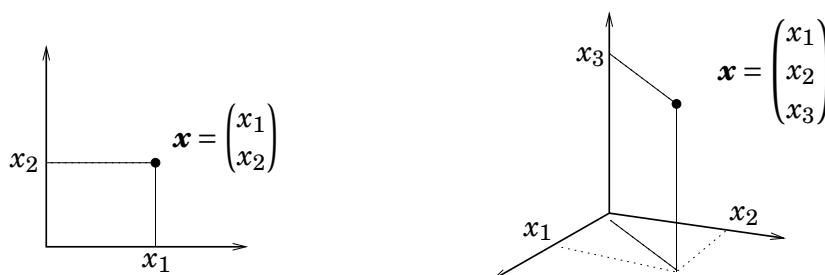
$$\mathbb{F}^m = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mid x_1, \dots, x_m \in \mathbb{F} \right\}$$

til at være mængden $M_{m,1}(\mathbb{F})$ af $m \times 1$ -matricer med indgange i $\mathbb{F}$, som vi også kalder for mængden af *søjlevektorer* i $\mathbb{F}$ af dimension m . Vi vil tilsvarende kalde mængden $M_{1,n}(\mathbb{F})$ af $1 \times n$ -matricer med indgange i $\mathbb{F}$ for mængden af *rækkevektorer* i $\mathbb{F}$ af dimension n , men vi vil ikke indføre nogen speciel notation for denne mængde.

Eksempel 2.2.1 Vi kan visualisere

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\} \quad \text{og} \quad \mathbb{R}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1, x_2, x_3 \in \mathbb{R} \right\}$$

som mængden af punkter i henholdsvis planen og rummet; se figur 2.1.



Figur 2.1: De reelle vektorrum $\mathbb{R}^2$ og $\mathbb{R}^3$.

Vi har allerede defineret sum og produkt af matricer, og disse regneoperationer kan specielt anvendes på søjlevektorer som følger. Givet to søjlevektorer

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \quad \text{og} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

i $\mathbb{F}$ af samme dimension m , da er deres sum givet ved

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_m + y_m \end{pmatrix},$$

2 Matricer og lineære afbildninger

hvilken igen er en søjlevektor i $\mathbb{F}$ af dimension m .

Vi husker dernæst på, at hvis vi ganger en $m \times 1$ -matrix med en 1×1 -matrix, så får vi igen en $m \times 1$ -matrix. Vi kalder denne operation for *skalarmultiplikation*. Som det er sædvanen, vil vi også misbruge notation og identificere en 1×1 -matrix $(a) \in M_{1,1}(\mathbb{F})$ med dens indgang $a \in \mathbb{F}$. Så det skalare multiplum af en søjlevektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{F}^m$$

med en skalar $a \in \mathbb{F}$ er søjlevektoren

$$\mathbf{x} \cdot a = \begin{pmatrix} x_1 a \\ \vdots \\ x_m a \end{pmatrix} \in \mathbb{F}^m.$$

Vi forkorter normalt og skriver $\mathbf{x}a$ i stedet for $\mathbf{x} \cdot a$. Vi understreger, at det omvendte produkt $a \cdot \mathbf{x}$ ikke har mening som et matrixprodukt, medmindre $m = 1$, da antallet af søjler i a ikke er lig med antallet af rækker i $\mathbf{x}$. Endvidere er den *modsatte vektor* af $\mathbf{x}$ og nulvektoren $\mathbf{0}$ givet ved henholdsvis

$$-\mathbf{x} = \begin{pmatrix} -x_1 \\ \vdots \\ -x_m \end{pmatrix} \quad \text{og} \quad \mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Vi henviser til figur 2.2 for en illustration af disse aritmetiske operationer.

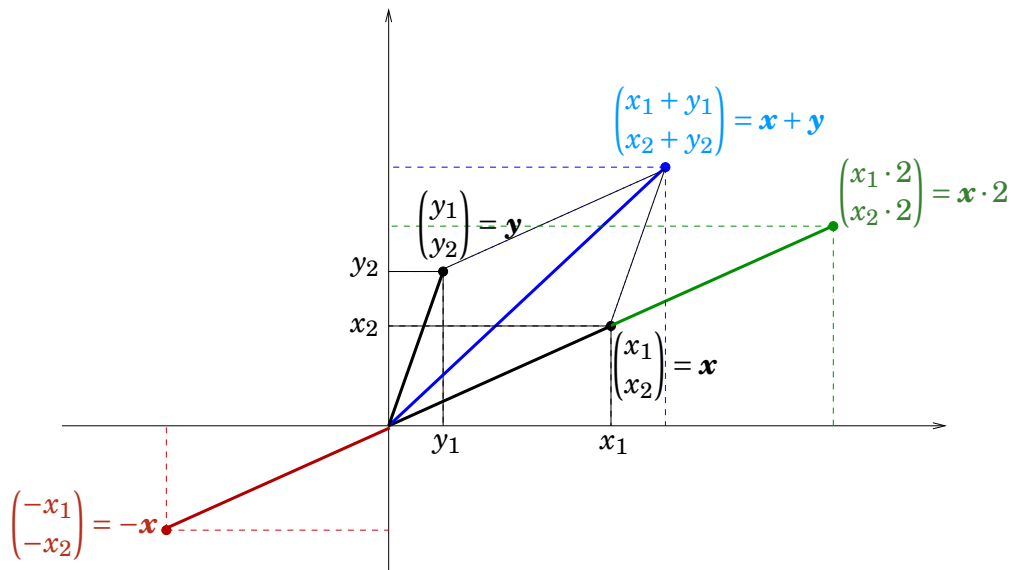
Eksempel 2.2.2 Vi betragter vektorerne $\mathbf{x}, \mathbf{y} \in \mathbb{C}^3$ givet ved

$$\mathbf{x} = \begin{pmatrix} 5 \\ i \\ 2i \end{pmatrix} \quad \text{og} \quad \mathbf{y} = \begin{pmatrix} i+1 \\ 3 \\ -i \end{pmatrix}.$$

For disse vektorer er

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} i+6 \\ i+3 \\ i \end{pmatrix}, \quad \mathbf{y} \cdot 3 = \begin{pmatrix} 3i+3 \\ 9 \\ -3i \end{pmatrix} \quad \text{og} \quad -\mathbf{y} = \begin{pmatrix} -i-1 \\ -3 \\ i \end{pmatrix}.$$

Sætning 2.1.10 ovenfor specialiserer nu til den følgende sætning, der udtrykker, at mængden $\mathbb{F}^m$ sammen med operationerne vektorsum og skalarmultiplikation udgør, hvad vi senere vil kalde et $\mathbb{F}$ -vektorrum; se definition 4.1.1.



Figur 2.2: Addition, skalarmultiplikation og modsat vektor i $\mathbb{R}^2$.

Sætning 2.2.3 For søjlevektorer $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}^m$ og skalarer $a, b \in \mathbb{F}$ gælder følgende:

(A1) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$

(A2) $\mathbf{x} + \mathbf{0} = \mathbf{x} = \mathbf{0} + \mathbf{x}$

(A3) $\mathbf{x} + (-\mathbf{x}) = \mathbf{0} = (-\mathbf{x}) + \mathbf{x}$

(A4) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$

(P1) $(\mathbf{x} \cdot a) \cdot b = \mathbf{x} \cdot (a \cdot b)$

(P2) $\mathbf{x} \cdot 1 = \mathbf{x}$

(D1) $\mathbf{x} \cdot (a + b) = \mathbf{x} \cdot a + \mathbf{x} \cdot b$

(D2) $(\mathbf{x} + \mathbf{y}) \cdot a = \mathbf{x} \cdot a + \mathbf{y} \cdot a$

(E1) $\mathbf{x} \cdot \mathbf{0} = \mathbf{0}$

(E2) $\mathbf{x} \cdot (-1) = -\mathbf{x}$

Vi siger, at en familie af elementer i en mængde X indexeret ved $I = \{1, 2, \dots, m\}$ er en m -tuple af elementer i X og skriver $(x_1, x_2, \dots, x_m)$ i stedet for $(x_i)_{i \in I}$.

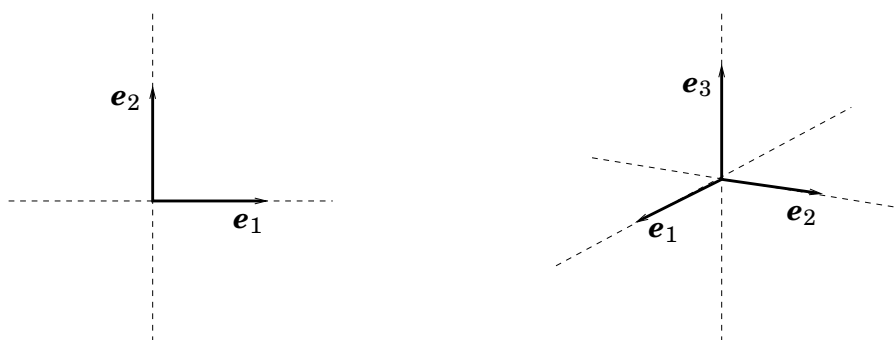
2 Matricer og lineære afbildninger

Definition 2.2.4 Hvis $\mathbb{F}$ er et legeme, og hvis m et naturligt tal, da kaldes m -tuplen $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ af vektorer i $\mathbb{F}^m$, hvor

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

for *standardbasen* for $\mathbb{F}^m$.

Vektorerne $\mathbf{e}_1, \dots, \mathbf{e}_m$ kaldes for *standardenhedsvektorerne* i $\mathbb{F}^m$. Den følgende figur illustrerer disse vektorer for $\mathbb{F} = \mathbb{R}$.



Figur 2.3: Standardenhedsvektorerne i $\mathbb{R}^2$ og $\mathbb{R}^3$.

Lad $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ være en r -tuple af vektorer i $\mathbb{F}^m$. Vi siger, at en vektor $\mathbf{u} \in \mathbb{F}^m$ er en *linearkombination* af $(\mathbf{u}_1, \dots, \mathbf{u}_r)$, hvis der findes skalarer $a_1, \dots, a_r \in \mathbb{F}$, sådan at

$$\mathbf{u} = \mathbf{u}_1 a_1 + \mathbf{u}_2 a_2 + \dots + \mathbf{u}_r a_r.$$

Specielt er nulvektoren $\mathbf{0}$ en linearkombination af 0-tuplen $()$, idet vi vedtager, at en tom sum er lig med $\mathbf{0}$.

Eksempel 2.2.5 Vi betragter de følgende fire vektorer i $\mathbb{R}^3$:

$$\mathbf{u} = \begin{pmatrix} 2 \\ 5 \\ 3 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Vi ser da, at $\mathbf{u}$ er en linearkombination af familien $(\mathbf{u}_1, \mathbf{u}_2)$, idet

$$\mathbf{u} = \begin{pmatrix} 2 \\ 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot 2 + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \cdot 3 = \mathbf{u}_1 \cdot 2 + \mathbf{u}_2 \cdot 3.$$

På den anden side kan $\mathbf{v}$ ikke skrives som linearkombination af $\mathbf{u}_1$ og $\mathbf{u}_2$. For antag, at der findes a_1, a_2 , sådan at $\mathbf{u}_1 a_1 + \mathbf{u}_2 a_2 = \mathbf{v}$. Da er

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot a_1 + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \cdot a_2 = \begin{pmatrix} a_1 \\ a_1 + a_2 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

hvilket ikke er muligt, idet ligningssystemet

$$\begin{aligned} a_1 &= 1 \\ a_1 + a_2 &= 1 \\ a_2 &= 1. \end{aligned}$$

ikke har nogen løsning.

Lemma 2.2.6 *Lad $\mathbb{F}$ være et legeme og m et naturligt tal. Enhver vektor $\mathbf{x} \in \mathbb{F}^m$ kan på entydig vis udtrykkes som en linearkombination*

$$\mathbf{x} = \mathbf{e}_1 a_1 + \mathbf{e}_2 a_2 + \cdots + \mathbf{e}_m a_m$$

af standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ for $\mathbb{F}^m$.

Bevis Givet en vilkårlig vektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{F}^m,$$

ønsker vi at undersøge, om der findes $a_1, a_2, \dots, a_m \in \mathbb{F}$, sådan at

$$\mathbf{x} = \mathbf{e}_1 a_1 + \mathbf{e}_2 a_2 + \cdots + \mathbf{e}_m a_m.$$

Vi udregner derfor højresiden, som er lig med

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} a_1 + \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} a_2 + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} a_m = \begin{pmatrix} a_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix},$$

hvoraf vi ser, at $a_1 = x_1, a_2 = x_2, \dots, a_m = x_m$ er den entydige løsning til dette problem. $\square$

2.3 Lineære afbildninger

Vi lader igen $\mathbb{F}$ være et legeme, for eksempel $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og betragter vektorrummet

$$\mathbb{F}^n = M_{n,1}(\mathbb{F}).$$

Hvis $A \in M_{m,n}(\mathbb{F})$ er en $m \times n$ -matrix med indgange i $\mathbb{F}$, så kan vi for alle $\mathbf{x} \in \mathbb{F}^n$ danne produktet $A\mathbf{x} \in M_{m,1}(\mathbb{F}) = \mathbb{F}^m$. På denne måde giver $m \times n$ -matricen $A \in M_{m,n}(\mathbb{F})$ dermed anledning til en afbildning

$$f: \mathbb{F}^n \rightarrow \mathbb{F}^m$$

defineret ved $f(\mathbf{x}) = A\mathbf{x}$.

Eksempel 2.3.1 Matricen

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 6 & -1 & 9 \end{pmatrix} \in M_{2,3}(\mathbb{R}).$$

giver anledning til afbildningen $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ givet ved

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 6 & -1 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 + 3x_3 \\ 6x_1 - x_2 + 9x_3 \end{pmatrix}.$$

En matrix $A \in M_{m,n}(\mathbb{F})$ bestemmer afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ givet ved $f(\mathbf{x}) = A\mathbf{x} \in \mathbb{F}^m$. Vi skal i dette afsnit vise, at de afbildninger, der fremkommer på denne måde, præcis er de *lineære* afbildninger, som vi nu definerer.

Definition 2.3.2 Lad $\mathbb{F}$ være et legeme. En afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er lineær, hvis den opfylder følgende betingelser (L1)–(L2) for alle $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ og $a \in \mathbb{F}$.

$$(L1) \quad f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}).$$

$$(L2) \quad f(\mathbf{x} \cdot a) = f(\mathbf{x}) \cdot a.$$

Enhver lineær afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ afbilder nødvendigvis $\mathbf{0} \in \mathbb{F}^n$ til $\mathbf{0} \in \mathbb{F}^m$. For

$$f(\mathbf{0}) = f(\mathbf{0} \cdot 0) = f(\mathbf{0}) \cdot 0 = \mathbf{0},$$

hvor den første og sidste identitet fås fra (E1) i sætning 2.2.3, mens den midterste identitet er (L2). Ved at anvende (E2) i sætning 2.2.3 samt (L2), ser vi tilsvarende, at

$$f(-\mathbf{x}) = f(\mathbf{x} \cdot (-1)) = f(\mathbf{x}) \cdot (-1) = -f(\mathbf{x}).$$

Eksempel 2.3.3 Afbildningen $f: \mathbb{R} \rightarrow \mathbb{R}$ givet ved $f(x) = 2 \cdot x$ er lineær, idet

$$\begin{aligned} f(x+y) &= 2 \cdot (x+y) \stackrel{(D1)}{=} 2 \cdot x + 2 \cdot y = f(x) + f(y), \\ f(x \cdot a) &= 2 \cdot (x \cdot a) \stackrel{(P1)}{=} (2 \cdot x) \cdot a = f(x) \cdot a. \end{aligned}$$

Derimod er afbildningen $g: \mathbb{R} \rightarrow \mathbb{R}$ givet ved $g(x) = x^2$ er ikke lineær, da den hverken opfylder (L1) eller (L2). For eksempel er

$$\begin{aligned} g(1+1) &= (1+1)^2 = 2^2 = 4, & \text{ mens } & g(1) + g(1) = 1^2 + 1^2 = 2, & \text{ og} \\ g(1 \cdot 2) &= (1 \cdot 2)^2 = 2^2 = 4, & \text{ mens } & g(1) \cdot 2 = 1^2 \cdot 2 = 2. \end{aligned}$$

Afbildningen $h: \mathbb{R} \rightarrow \mathbb{R}$ givet ved $h(x) = 2 \cdot x + 3$ er heller ikke lineær, da for eksempel $h(0) \neq 0$. Vi siger, at en afbildning som denne, der er en sum af en lineær afbildning og en konstant afbildning, er en *affin* afbildning. Vi bemærker, at i dele af litteraturen anvendes “lineær” i betydningen “affin”.

Sætning 2.3.4 Hvis $A \in M_{m,n}(\mathbb{F})$ er en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$, da er afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ defineret ved $f(\mathbf{x}) = A\mathbf{x}$ en lineær afbildning.

Bevis Det følger fra sætning 2.1.10, at $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er lineær. For (D1) viser, at

$$f(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = f(\mathbf{x}) + f(\mathbf{y}),$$

mens (P1) tilsvarende viser, at

$$f(\mathbf{x}a) = A(\mathbf{x}a) = (A\mathbf{x})a = f(\mathbf{x})a,$$

så afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ opfylder både (L1) og (L2). □

Eksempel 2.3.5 Multiplikation med Pauli matricen

$$A = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \in M_2(\mathbb{C})$$

giver anledning til afbildningen $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ defineret ved

$$f \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} -iz_2 \\ iz_1 \end{pmatrix},$$

som ifølge sætning 2.3.4 er lineær.

2 Matricer og lineære afbildninger

Eksempel 2.3.6 Vi viser direkte, at afbildningen $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ fra eksempel 2.3.1,

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 + 3x_3 \\ 6x_1 - x_2 + 9x_3 \end{pmatrix},$$

er lineær. Udregningen

$$\begin{aligned} f(\mathbf{x} + \mathbf{y}) &= f \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix} = \begin{pmatrix} 2(x_1 + y_1) + (x_2 + y_2) + 3(x_3 + y_3) \\ 6(x_1 + y_1) - (x_2 + y_2) + 9(x_3 + y_3) \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 + x_2 + 3x_3 \\ 6x_1 - x_2 + 9x_3 \end{pmatrix} + \begin{pmatrix} 2y_1 + y_2 + 3y_3 \\ 6y_1 - y_2 + 9y_3 \end{pmatrix} = f(\mathbf{x}) + f(\mathbf{y}) \end{aligned}$$

viser, at (L1) er opfyldt, og udregningen

$$\begin{aligned} f(\mathbf{x} \cdot a) &= f \begin{pmatrix} x_1 a \\ x_2 a \\ x_3 a \end{pmatrix} = \begin{pmatrix} 2(x_1 a) + (x_2 a) + 3(x_3 a) \\ 6(x_1 a) - (x_2 a) + 9(x_3 a) \end{pmatrix} \\ &= \begin{pmatrix} (2x_1 + x_2 + 3x_3)a \\ (6x_1 - x_2 + 9x_3)a \end{pmatrix} = f(\mathbf{x}) \cdot a \end{aligned}$$

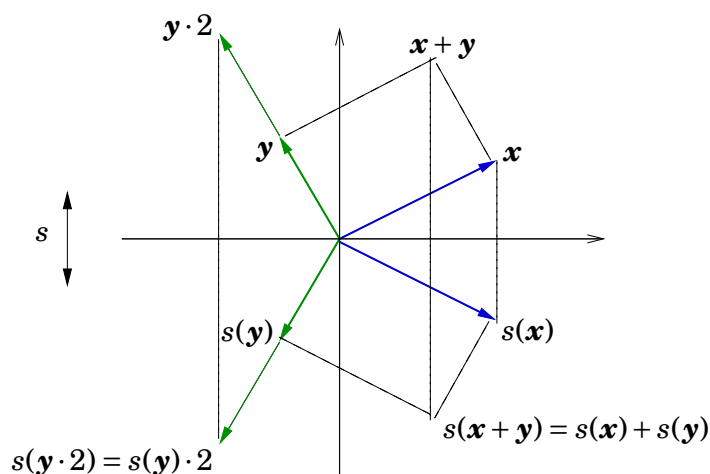
viser tilsvarende, at (L2) er opfyldt.

Eksempel 2.3.7 (Spejling) Vi definerer $s: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ til at være afbildningen, der er givet ved spejling i x -aksen i $\mathbb{R}^2$; se figur 2.4. I koordinater er denne givet ved

$$s \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix},$$

og den følgende udregning viser, at den opfylder (L1) og (L2).

$$\begin{aligned} s(\mathbf{x} + \mathbf{y}) &= s \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ -(x_2 + y_2) \end{pmatrix} \\ &= \begin{pmatrix} x_1 + y_1 \\ -x_2 - y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ -y_2 \end{pmatrix} = s(\mathbf{x}) + s(\mathbf{y}) \\ s(\mathbf{x} \cdot a) &= s \begin{pmatrix} x_1 \cdot a \\ x_2 \cdot a \end{pmatrix} = \begin{pmatrix} x_1 \cdot a \\ -(x_2 \cdot a) \end{pmatrix} = \begin{pmatrix} x_1 \cdot a \\ (-1) \cdot (x_2 \cdot a) \end{pmatrix} \\ &= \begin{pmatrix} x_1 \cdot a \\ ((-1) \cdot x_2) \cdot a \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} \cdot a = s(\mathbf{x}) \cdot a \end{aligned}$$



Figur 2.4: Spejling s i x -aksen i $\mathbb{R}^2$ er en lineær afbildning

Som forberedelse til at vise, at enhver lineær afbildning fremkommer ved multiplikation med en matrix, viser vi nu, at lineære afbildninger bevarer linearkombinationer.

Lemma 2.3.8 *Lad $\mathbb{F}$ være et legeme, og lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ være en lineær afbildning. Givet vektorer $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{F}^n$ og skalarer $a_1, \dots, a_k \in \mathbb{F}$, da gælder det, at*

$$f(\mathbf{x}_1 a_1 + \mathbf{x}_2 a_2 + \dots + \mathbf{x}_k a_k) = f(\mathbf{x}_1) a_1 + f(\mathbf{x}_2) a_2 + \dots + f(\mathbf{x}_k) a_k.$$

Bevis Vi beviser påstanden ved induktion på $k \geq 0$. I tilfældet $k = 0$, da er påstanden, at $f(\mathbf{0}) = \mathbf{0}$, hvilket vi allerede har vist ovenfor. For den tomme sum er per definition lig med nulvektoren. Så vi antager, at påstanden allerede er bevist for $k = r - 1$ og beviser den for $k = r$. Hertil udregner vi

$$\begin{aligned} f(\mathbf{x}_1 a_1 + \dots + \mathbf{x}_{r-1} a_{r-1} + \mathbf{x}_r a_r) &= f(\mathbf{x}_1 a_1 + \dots + \mathbf{x}_{r-1} a_{r-1}) + f(\mathbf{x}_r a_r) \\ &= f(\mathbf{x}_1 a_1 + \dots + \mathbf{x}_{r-1} a_{r-1}) + f(\mathbf{x}_r) a_r = f(\mathbf{x}_1) a_1 + \dots + f(\mathbf{x}_{r-1}) a_{r-1} + f(\mathbf{x}_r) a_r, \end{aligned}$$

hvor de tre identiteter fås fra henholdsvis (L1) med $\mathbf{x} = \mathbf{x}_1 a_1 + \dots + \mathbf{x}_{r-1} a_{r-1}$ og $\mathbf{y} = \mathbf{x}_r a_r$, fra (L2) med $\mathbf{x} = \mathbf{x}_r$ og $a = a_r$, og fra den induktive hypotese. Dette viser induktions-skridtet og dermed lemmaet. $\square$

Vi har vist i sætning 2.3.4, at multiplikation med en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$ definerer en lineær afbildning fra $\mathbb{F}^n$ til $\mathbb{F}^m$. Vi viser nu omvendt, at enhver lineær afbildning fra $\mathbb{F}^n$ til $\mathbb{F}^m$ fremkommer ved multiplikation med en $m \times n$ -matrix med indgange i $\mathbb{F}$. Vi betragter denne matrix som den generaliserede hældningskoefficient af den lineære afbildning.

2 Matricer og lineære afbildninger

Sætning 2.3.9 Lad $\mathbb{F}$ være et legeme, lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ være en lineær afbildning, og lad $A \in M_{m,n}(\mathbb{F})$ være matricen, hvis j 'te søjle er $f(\mathbf{e}_j) \in \mathbb{F}^m$. Da gælder det for alle $\mathbf{x} \in \mathbb{F}^n$, at

$$f(\mathbf{x}) = A\mathbf{x},$$

og matricen $A \in M_{m,n}(\mathbb{F})$ er entydigt bestemt med denne egenskab.

Bevis Entydighedsudsagnet i sætning følger umiddelbart af, at der for enhver matrix $A \in M_{m,n}(\mathbb{F})$ gælder, at dens j 'te søjle er lig med $A\mathbf{e}_j$. Så vi lader nu

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M_{m,n}(\mathbb{F})$$

være matricen, hvis j 'te søjle er lig med $f(\mathbf{e}_j)$, og viser, at $f(\mathbf{x}) = A\mathbf{x}$ for alle $\mathbf{x} \in \mathbb{F}^n$. Ifølge lemma 2.2.6 kan vi skrive $\mathbf{x} \in \mathbb{F}^n$ entydigt som

$$\mathbf{x} = \mathbf{e}_1x_1 + \mathbf{e}_2x_2 + \cdots + \mathbf{e}_nx_n = \sum_{j=1}^n \mathbf{e}_jx_j,$$

og da $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er lineær, så viser lemma 2.3.8, at

$$f(\mathbf{x}) = f\left(\sum_{j=1}^n \mathbf{e}_jx_j\right) = \sum_{j=1}^n f(\mathbf{e}_j)x_j.$$

Per definition af matricen A gælder det endvidere, at

$$f(\mathbf{e}_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = \mathbf{e}_1a_{1j} + \mathbf{e}_2a_{2j} + \cdots + \mathbf{e}_ma_{mj} = \sum_{i=1}^m \mathbf{e}_ia_{ij},$$

hvilket vi substituerer i formelen for $f(\mathbf{x})$ overfor. Herved får vi

$$f(\mathbf{x}) = \sum_{j=1}^n \left(\sum_{i=1}^m \mathbf{e}_ia_{ij} \right) x_j = \sum_{j=1}^n \sum_{i=1}^m (\mathbf{e}_ia_{ij}x_j) = \sum_{j=1}^n \sum_{i=1}^m (\mathbf{e}_ia_{ij}x_j) = \sum_{i=1}^m \mathbf{e}_i \left(\sum_{j=1}^n a_{ij}x_j \right),$$

hvilket viser, at

$$f(\mathbf{x}) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A\mathbf{x}$$

som ønsket. □

Definition 2.3.10 Lad $\mathbb{F}$ være et legeme og lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ være en lineær afbildning. Den entydigt bestemte matrix $A \in M_{m,n}(\mathbb{F})$, sådan at der for alle $\mathbf{x} \in \mathbb{F}^n$ gælder det, at

$$f(\mathbf{x}) = A\mathbf{x},$$

kaldes for *matricen, der repræsenterer $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ med hensyn til standardbaserne* for domænet $\mathbb{F}^n$ og codomænet $\mathbb{F}^m$.

Vi skal senere definere og betragte andre baser for $\mathbb{F}^n$ og $\mathbb{F}^m$ end standardbaserne, og vi er derfor forsigtige med at sige, at det her er standardbaserne, vi benytter.

Eksempel 2.3.11 Vi så i eksempel 2.3.7, at afbildningen $s: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, der er givet ved spejling i x -aksen, er lineær. Matricen, der repræsenterer denne afbildning med hensyn til standardbaserne for domænet og codomænet, er

$$A = (s(\mathbf{e}_1) \quad s(\mathbf{e}_2)) = \left(s \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad s \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

hvilket udregningen

$$A\mathbf{x} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} = s(\mathbf{x})$$

bekræfter.

Eksempel 2.3.12 Vi efterviste i eksempel 2.3.6, at afbildningen $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ givet ved

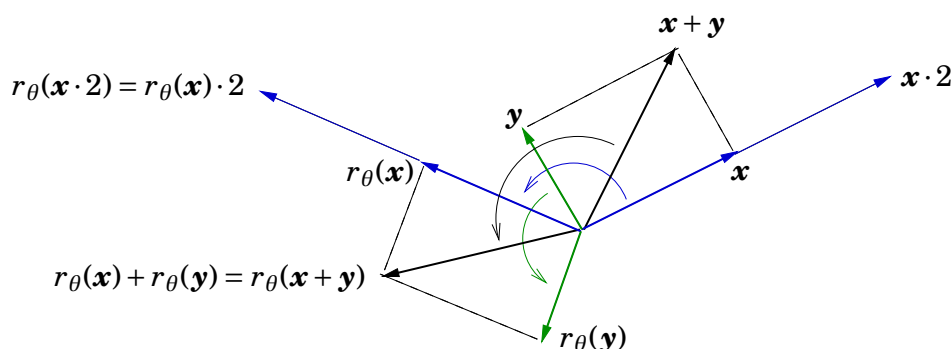
$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 + 3x_3 \\ 6x_1 - x_2 + 9x_3 \end{pmatrix}$$

er lineær. Matricen, der repræsenterer $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ med hensyn til standardbaserne for domænet og codomænet, er

$$A = (f(\mathbf{e}_1) \quad f(\mathbf{e}_2) \quad f(\mathbf{e}_3)) = \begin{pmatrix} 2 & 1 & 3 \\ 6 & -1 & 9 \end{pmatrix},$$

hvilket vi selvfølgelig allerede vidste fra eksempel 2.3.1.

2 Matricer og lineære afbildninger



Figur 2.5: Rotation r_θ med vinkel θ omkring $\mathbf{0}$ i $\mathbb{R}^2$ er en lineær afbildning

Eksempel 2.3.13 Vi betragter den afbildning $r_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, der er givet ved en rotation på θ radianer imod urets retning omkring $\mathbf{0} \in \mathbb{R}^2$. Figur 2.5 illustrerer, at denne afbildning opfylder (L1)–(L2) og derfor er lineær. Matricen, der repræsenterer $r_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ med hensyn til standardbaserne for domænet og codomænet, er

$$B = (r_\theta(\mathbf{e}_1) \quad r_\theta(\mathbf{e}_2)) = \begin{pmatrix} \cos \theta & \cos(\theta + \frac{\pi}{2}) \\ \sin \theta & \sin(\theta + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

hvor vi har brugt, at

$$r_\theta \begin{pmatrix} 1 \\ 0 \end{pmatrix} = r_\theta \begin{pmatrix} \cos 0 \\ \sin 0 \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad \text{og} \quad r_\theta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = r_\theta \begin{pmatrix} \cos(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} \cos(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) \end{pmatrix}.$$

Eksempel 2.3.14 Identitetsafbildningen $\text{id}: \mathbb{F}^n \rightarrow \mathbb{F}^n$ er defineret ved

$$\text{id}(\mathbf{x}) = \mathbf{x}$$

er tydeligvis lineær. Matricen, der repræsenterer denne afbildning med hensyn til standardbaserne for *både* domænet og codomænet, er identitetsmatricen

$$A = (\text{id}(\mathbf{e}_1) \quad \text{id}(\mathbf{e}_2) \quad \cdots \quad \text{id}(\mathbf{e}_n)) = (\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n.$$

I forbindelse med koordinatskift skal vi senere betragte matricen, der repræsenterer identitetsafbildningen med hensyn til *forskellige* baser for domænet og codomænet, og denne matrix er da *ikke* identitetsafbildningen.

Vi viser nu, at sammensætning af lineære afbildninger svarer til multiplikation af de matricer, der repræsenterer dem, i *samme* rækkefølge.

Sætning 2.3.15 *Lad $\mathbb{F}$ være et legeme, lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ og $g: \mathbb{F}^p \rightarrow \mathbb{F}^n$ være to lineære afbildninger og lad $A \in M_{m,n}(\mathbb{F})$ og $B \in M_{n,p}(\mathbb{F})$ være de matricer, der repræsenterer henholdsvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ og $g: \mathbb{F}^p \rightarrow \mathbb{F}^n$ med hensyn til de respektive standardbaser. Den sammensatte afbildning $f \circ g: \mathbb{F}^p \rightarrow \mathbb{F}^m$ er da lineær, og matricen $C \in M_{m,p}(\mathbb{F})$, der repræsenterer $f \circ g: \mathbb{F}^p \rightarrow \mathbb{F}^m$ med hensyn til de respektive standardbaser, er*

$$C = AB.$$

Bevis For alle $\mathbf{x} \in \mathbb{F}^p$ gælder det, at

$$(f \circ g)(\mathbf{x}) = f(g(\mathbf{x})) = A(B\mathbf{x}) = (AB)\mathbf{x}.$$

Her er den første lighed definitionen af den sammensatte afbildning, den anden lighed er definitionen af matricerne, der repræsenterer de lineære afbildninger med hensyn til de respektive standardbaser, mens den sidste lighed er sætning 2.1.10 (P1). Det følger derfor fra sætning 2.3.4, at afbildningen $f \circ g: \mathbb{F}^p \rightarrow \mathbb{F}^m$ er lineær, og da matricen C , der repræsenterer denne afbildning med hensyn til standardbaserne for $\mathbb{F}^p$ og $\mathbb{F}^m$, er entydigt bestemt, konkluderer vi endvidere, at $C = AB$ som påstået. $\square$

Eksempel 2.3.16 Lad $s: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ være den lineære afbildning fra eksempel 2.3.7, der er givet ved spejling i x -aksen, og lad $r = r_{\pi/2}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ være den lineære afbildning fra eksempel 2.3.13, der er givet ved rotation gennem $\pi/2$ radianer imod urets retning omkring $\mathbf{0} \in \mathbb{R}^2$. Vi ved fra eksempel 2.3.11 og 2.3.13, at $r: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ og $s: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ er repræsenteret med hensyn til de respektive standardbaser ved henholdsvis

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

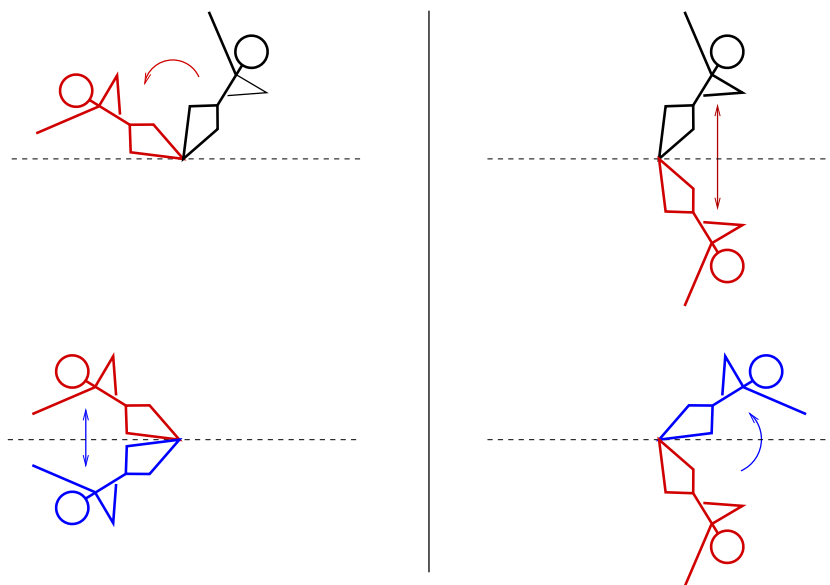
Ifølge sætning 2.3.15 repræsenterer produktmatricerne BA og AB derfor henholdsvis de sammensatte afbildninger $s \circ r: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ og $r \circ s: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, hvoraf den første er “først drej, så spejl”, mens den anden er “først spejl, så drej”. Vi udregner

$$BA = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

$$AB = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Vi bemærker, at de to matricer er forskellige, og konkluderer derfor, at de sammensatte afbildninger $s \circ r$ og $r \circ s$ ligeledes er forskellige. De to sammensatte afbildninger er illustreret i figur 2.6.

2 Matricer og lineære afbildninger



Figur 2.6: På venstre side afbildes den sorte figur ved r i den røde figur, som derefter ved s afbildes i den blå figur. På højre side afbildes samme sorte figur ved s i den røde figur, som derefter ved r afbildes i den blå figur.

Lineære afbildninger optræder i mange sammenhænge. Vi afslutter dette afsnit med et eksempel fra operationsanalyse, der illustrerer en lineær sammenhæng mellem varer og råvarer i en produktion. Hvis produktionsbegrænsninger endvidere er udtrykt ved lineære uligheder, da kan den optimale produktion bestemmes effektivt ved hjælp af Dantzig's simpleksalgoritme¹. Dette område kaldes for lineær programmering.

Eksempel 2.3.17 En fabrik fremstiller to varer X_1 og X_2 og anvender dertil tre råvarer Y_1 , Y_2 og Y_3 . Hvis der dagligt fremstilles x_1 enheder af X_1 og x_2 enheder af X_2 , så siger vi, at fabrikkens *produktion* er

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Hvis fabrikken dagligt forbruger y_1 enheder af Y_1 , y_2 enheder af Y_2 og y_3 enheder af Y_3 , så siger vi tilsvarende, at fabrikkens *forbrug* er

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Til produktion af én enhed af X_1 kræves tre enheder af Y_1 , to enheder af Y_2 og én enhed af Y_3 ; og til produktion én enhed af X_2 kræves fem enheder af Y_1 , fem enheder

¹Se en.wikipedia.org/wiki/Simplex_algorithm

af Y_2 og tre enheder af Y_3 . Sammenhængen mellem forbrug og produktion er altså givet ved

$$\begin{aligned}y_1 &= 3x_1 + 5x_2 \\y_2 &= 2x_1 + 5x_2 \\y_3 &= x_1 + 3x_2.\end{aligned}$$

Så hvis vi lader $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ være den afbildning, der til en ønsket produktion $\mathbf{x}$ af varer tilordner det tilsvarende forbrug $\mathbf{y}$ af råvarer, da gælder det, at

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 3x_1 + 5x_2 \\ 2x_1 + 5x_2 \\ x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Dermed er $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ altså en lineær afbildning.

2.4 Invertible matricer

I dette afsnit undersøger vi de matricer A , for hvilke ligningssystemet $A\mathbf{x} = \mathbf{b}$ præcis har én løsning for alle $\mathbf{b}$. Disse matricer, som vi kalder invertible, repræsenterer således bijektive lineære afbildninger, og vi viser, at de nødvendigvis er kvadratiske.

Lemma 2.4.1 *Lad $\mathbb{F}$ være et legeme, lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ være en lineær afbildning, og lad $A \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ med hensyn til de respektive standardbaser. Da er følgende udsagn ækvivalente:*

- (1) *Afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er bijektiv.*
- (2) *For alle $\mathbf{b} \in \mathbb{F}^m$ har det lineære ligningssystem $A\mathbf{x} = \mathbf{b}$ præcis én løsning $\mathbf{x} \in \mathbb{F}^n$.*

Bevis At $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er bijektiv, betyder per definition, at der for alle $\mathbf{b} \in \mathbb{F}^m$ præcis findes ét $\mathbf{x} \in \mathbb{F}^n$, sådan at $f(\mathbf{x}) = \mathbf{b}$. Da $f(\mathbf{x}) = A\mathbf{x}$ for alle $\mathbf{x} \in \mathbb{F}^n$, følger lemmaet. $\square$

Eksempel 2.4.2 Vi undersøger, om den lineære afbildning $f: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ givet ved

$$f(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 \end{pmatrix}$$

er bijektiv, hvilket ifølge lemma 2.4.1 er ækvivalent til at undersøge, hvorvidt det for alle $\mathbf{b} \in \mathbb{F}^2$ gælder, at ligningssystemet $A\mathbf{x} = \mathbf{b}$ præcis har én løsning. Vi anvender

2 Matricer og lineære afbildninger

derfor rækkeoperationer til at omdanne ligningssystemets totalmatrix $(A \mid \mathbf{b})$ til en matrix på $(A' \mid \mathbf{b}')$ på reduceret echelonform.

$$\begin{aligned} (A \mid \mathbf{b}) &= \left(\begin{array}{cc|c} 1 & 2 & b_1 \\ 2 & 1 & b_2 \end{array} \right) + (-2)R_1 \\ &= \left(\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & -3 & -2b_1 + b_2 \end{array} \right) \left(-\frac{1}{3} \right) \cdot R_2 \\ &= \left(\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & 1 & \frac{2b_1 - b_2}{3} \end{array} \right) + (-2)R_2 \\ (A' \mid \mathbf{b}') &= \left(\begin{array}{cc|c} 1 & 0 & \frac{2b_2 - b_1}{3} \\ 0 & 1 & \frac{2b_1 - b_2}{3} \end{array} \right) + (-2)R_2 \end{aligned}$$

Vi ser, at for ethvert $\mathbf{b} \in \mathbb{F}^2$ har ligningssystemet $A\mathbf{x} = \mathbf{b}$ netop den ene løsning

$$\mathbf{x} = \begin{pmatrix} 2b_2 - b_1 \\ 2b_1 - b_2 \end{pmatrix} \cdot \frac{1}{3},$$

og derfor er $f: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ altså bijektiv.

Vi viser nu, at en lineær afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ ikke kan være bijektiv, medmindre matricen $A \in M_{m,n}(\mathbb{F})$, der repræsenterer $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ med hensyn til standardbaserne, er en kvadratisk matrix.

Sætning 2.4.3 Hvis $\mathbb{F}$ er et legeme, og hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er en bijektiv lineær afbildning, så gælder det nødvendigvis, at $m = n$.

Bevis Lad $A \in M_{m,n}(\mathbb{F})$ være den matrix, der repræsenterer $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ med hensyn til de respektive standardbaser, og lad A' være matricen på reduceret echelonform af rang $0 \leq r \leq \min\{m, n\}$, der fremkommer fra A ved at udføre en følge af rækkeoperationer. Ifølge lemma 2.4.1 er $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ bijektiv, hvis og kun hvis det lineære ligningssystem $A\mathbf{x} = \mathbf{b}$ præcis én løsning for alle $\mathbf{b} \in \mathbb{F}^m$, og ifølge sætning 1.3.1 er dette tilfældet, hvis og kun hvis $r = m = n$. Så hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er bijektiv, da er $m = n$ som påstået. $\square$

Bemærkning 2.4.4 Givet vilkårlige positive heltal m og n , så kan man vise, at der altid findes en bijektiv afbildning $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Sætning 2.4.3 viser imidlertid, at en sådan bijektiv afbildning kun kan være lineær, hvis $m = n$.

Som beviset for sætning 2.4.3 afslører, er en lineær afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ bijektiv, hvis og kun hvis matricen $A \in M_n(\mathbb{F})$, der repræsenterer $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ med hensyn til standardbaserne, kan omdannes ved rækkeoperationer til identitetsmatricen.

Eksempel 2.4.5 Vi undersøger om den lineære afbildning $g: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ givet ved

$$g(\mathbf{z}) = B\mathbf{z} = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_1 + iz_2 \\ iz_1 - z_2 \end{pmatrix}$$

er bijektiv. Så vi omdanner totalmatricen $(B \mid \mathbf{b})$ for ligningssystem $B\mathbf{z} = \mathbf{b}$ til en matrix $(B' \mid \mathbf{b}')$ på reduceret echelonform.

$$(B \mid \mathbf{b}) = \left(\begin{array}{cc|c} 1 & i & b_1 \\ i & -1 & b_2 \end{array} \right) + (-i)R_1$$

$$(B' \mid \mathbf{b}') = \left(\begin{array}{cc|c} 1 & i & b_1 \\ 0 & 0 & -ib_1 + b_2 \end{array} \right)$$

Vi ser, at $g: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ ikke er bijektiv.

Vi husker fra lemma 0.1.3, at en afbildning $f: X \rightarrow Y$ er bijektiv, hvis og kun hvis der findes en afbildning $g: Y \rightarrow X$, sådan at $f \circ g = \text{id}_Y$ og $g \circ f = \text{id}_X$. Afbildningen $g: Y \rightarrow X$ er i givet fald entydigt bestemt og kaldes den inverse afbildning af $f: X \rightarrow Y$.

Sætning 2.4.6 Hvis $\mathbb{F}$ er et legeme, og hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ er en bijektiv lineær afbildning, så er den inverse afbildning $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ også lineær.

Bevis Vi skal vise, at den inverse afbildning $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ opfylder (L1) og (L2). Vi så i beviset for lemma 0.1.3, at “ $\mathbf{u} = g(\mathbf{v})$ ” og “ $f(\mathbf{u}) = \mathbf{v}$ ” er ækvivalente udsagn. Specielt er udsagnene “ $g(\mathbf{x}) + g(\mathbf{y}) = g(\mathbf{x} + \mathbf{y})$ ” og “ $f(g(\mathbf{x}) + g(\mathbf{y})) = \mathbf{x} + \mathbf{y}$ ” derfor ækvivalente, og det sidste udsagn gælder, fordi $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ opfylder (L1):

$$f(g(\mathbf{x}) + g(\mathbf{y})) = f(g(\mathbf{x})) + f(g(\mathbf{y})) = \mathbf{x} + \mathbf{y}.$$

Dette viser, at $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ opfylder (L1). Tilsvarende er “ $g(\mathbf{x})a = g(\mathbf{x}a)$ ” og “ $f(g(\mathbf{x})a) = \mathbf{x}a$ ” ækvivalente, og det sidste gælder, idet $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ opfylder (L2):

$$f(g(\mathbf{x})a) = f(g(\mathbf{x}))a = \mathbf{x}a.$$

Dette viser, at $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ også opfylder (L2). Altså er $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ lineær. $\square$

Eksempel 2.4.7 Den lineære afbildning $f: \mathbb{R} \rightarrow \mathbb{R}$ givet ved $f(x) = cx$, hvor $c \in \mathbb{R}$, er bijektiv, hvis og kun hvis $c \neq 0$. I givet fald er den inverse afbildning $g: \mathbb{R} \rightarrow \mathbb{R}$ givet ved $g(x) = c^{-1}x$, og den er derfor også lineær.

2 Matricer og lineære afbildninger

Lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ være en bijektiv lineær afbildning, og lad $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ være den inverse afbildning, der ifølge sætning 2.4.6 også er lineær. Lad $A \in M_n(\mathbb{F})$ og $B \in M_n(\mathbb{F})$ være de kvadratiske matricer, der repræsenterer henholdsvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ og $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ med hensyn til standardbaserne. Ifølge sætning 2.3.15 og eksempel 2.3.14 er identiteterne $f \circ g = \text{id} = g \circ f$ og matrixidentiteterne

$$AB = I = BA$$

da ækvivalente. Dette motiverer den følgende definition.

Definition 2.4.8 En kvadratisk matrix A er *invertibel*, hvis der findes en kvadratisk matrix B , sådan at $AB = I = BA$, hvor I er identitetsmatricen.

Hvis A er en kvadratisk matrix, og hvis de to kvadratiske matricer B og B' opfylder, at $AB = I = BA$ og $AB' = I = B'A$, da er $B = B'$. For ifølge sætning 2.1.10 er

$$B = BI = B(AB') = (BA)B' = IB' = B'.$$

Vi kalder matricen B for den *inverse matrix* af matricen A og betegner den A^{-1} . Den følgende sætning opsummerer ovenstående.

Sætning 2.4.9 Lad $\mathbb{F}$ være et legeme, lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ være en lineær afbildning, og lad $A \in M_n(\mathbb{F})$ være den kvadratiske matrix, der repræsenterer $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ med hensyn til standardbaserne. Da er følgende udsagn (1)–(4) ækvivalente.

- (1) Afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ er bijektiv.
- (2) Der findes en lineær afbildning $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$, sådan at $f \circ g = \text{id}_{\mathbb{F}^n} = g \circ f$.
- (3) Den kvadratiske matrix A er invertibel.
- (4) Det lineære ligningssystem $A\mathbf{x} = \mathbf{b}$ har præcis én løsning for alle $\mathbf{b} \in \mathbb{F}^n$.

Bevis Vi viste i lemma 2.4.1, at (1) og (4) er ækvivalente, og i sætning 2.4.6, at (1) og (2) er ækvivalente, og vi viser nu, at også (2) og (3) er ækvivalente. Hvis (2) holder, da opfylder matricen $B \in M_n(\mathbb{F})$, der repræsenterer $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ med hensyn til standardbasen, ifølge sætning 2.3.15 og eksempel 2.3.14, at $AB = I = BA$, så (3) holder. Omvendt, hvis (3) holder, da lader vi B være den inverse matrix af A og definerer $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$ til at være den lineære afbildning givet ved $g(\mathbf{y}) = B\mathbf{y}$. Sætning 2.3.15 og eksempel 2.3.14 viser da, at $f \circ g = \text{id}_{\mathbb{F}^n} = g \circ f$, så (2) holder. $\square$

Hvis en kvadratisk matrix A er invertibel, og hvis dens inverse matrix B er kendt, da er den entydige løsning til ligningssystemet $A\mathbf{x} = \mathbf{b}$ givet ved $\mathbf{x} = B\mathbf{b}$. For

$$A(B\mathbf{b}) = (AB)\mathbf{b} = I\mathbf{b} = \mathbf{b}.$$

Specielt er den j 'te søjle $\mathbf{b}_j = B\mathbf{e}_j$ i den inverse matrix lig med den entydige løsning til ligningssystemet $A\mathbf{b}_j = \mathbf{e}_j$. Vi kan derfor finde den inverse matrix B ved hjælp af rækkeoperationer. Vi vender tilbage til dette i næste afsnit.

Eksempel 2.4.10 Vi betragter matricerne

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 3/5 & -2/5 \\ 1/5 & 1/5 \end{pmatrix}$$

og udregner

$$AB = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 3/5 & -2/5 \\ 1/5 & 1/5 \end{pmatrix} = \begin{pmatrix} 3/5 + 2/5 & -2/5 + 2/5 \\ -3/5 + 3/5 & 2/5 + 3/5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$$BA = \begin{pmatrix} 3/5 & -2/5 \\ 1/5 & 1/5 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 3/5 + 2/5 & 6/5 - 6/5 \\ 1/5 - 1/5 & 2/5 + 3/5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

hvilket viser, at $B = A^{-1}$ er den inverse matrix af A .

Hvis en kvadratisk matrix med heltalsindgange er invertibel, da vil dens inverse matrix som i ovenstående eksempel typisk have indgange, der er brøker med samme nævner. Forklaringen herpå finder vi i kapitel 3.

Vi afslutter dette afsnit med at undersøge invertibilitet af diagonalmatricer, som vi nu definerer.

Definition 2.4.11 En kvadratisk matrix på formen

$$\text{diag}(a_1, \dots, a_n) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

kaldes for en *diagonalmatrix*.

Produktet af to diagonal matricer er igen en diagonalmatrix, idet

$$\text{diag}(a_1, \dots, a_n) \cdot \text{diag}(b_1, \dots, b_n) = \text{diag}(a_1 b_1, \dots, a_n b_n).$$

2 Matricer og lineære afbildninger

For eksempel er

$$\begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 \cdot 5 & 0 \\ 0 & -3 \cdot 1 \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ 0 & -3 \end{pmatrix}.$$

Vi bemærker også, at identitetsmatricen en diagonalmatrix.

Sætning 2.4.12 *Lad $\mathbb{F}$ være et legeme. Diagonalmatricen $\text{diag}(a_1, \dots, a_n)$ er invertibel, hvis og kun hvis diagonalindgangene $a_1, \dots, a_n \in \mathbb{F}$ alle er invertible, og i givet fald er*

$$\text{diag}(a_1, \dots, a_n)^{-1} = \text{diag}(a_1^{-1}, \dots, a_n^{-1}).$$

Bevis Hvis $a_1, \dots, a_n$ alle er invertible i $\mathbb{F}$, så er

$$\text{diag}(a_1, \dots, a_n) \text{diag}(a_1^{-1}, \dots, a_n^{-1}) = \text{diag}(1, \dots, 1) = I,$$

$$\text{diag}(a_1^{-1}, \dots, a_n^{-1}) \text{diag}(a_1, \dots, a_n) = \text{diag}(1, \dots, 1) = I,$$

hvilket viser, at $\text{diag}(a_1, \dots, a_n)$ er invertibel, og at dens inverse matrix er som angivet. Omvendt, hvis $A = \text{diag}(a_1, \dots, a_n)$ er invertibel med invers matrix B , så viser

$$AB = \begin{pmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} = \begin{pmatrix} a_1 b_{11} & \dots & a_1 b_{1n} \\ \vdots & \ddots & \vdots \\ a_n b_{n1} & \dots & a_n b_{nn} \end{pmatrix} = I,$$

$$BA = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \begin{pmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_n \end{pmatrix} = \begin{pmatrix} b_{11} a_1 & \dots & b_{1n} a_n \\ \vdots & \ddots & \vdots \\ b_{n1} a_1 & \dots & b_{nn} a_n \end{pmatrix} = I,$$

at a_i er invertibel med invers $a_i^{-1} = b_{ii}$ for alle $1 \leq i \leq n$, som ønsket. $\square$

Eksempel 2.4.13 Vi beskriver tre lineære afbildninger, der alle er repræsenteret med hensyn til standardbaserne af diagonalmatricer. Matricen

$$A = \text{diag}(2, 2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

repræsenterer den lineære afbildning $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, der skalerer alle vektorer med en faktor 2. Den er invertibel, og den inverse matrix $A^{-1} = \text{diag}(1/2, 1/2)$ repræsenterer den lineære afbildning $f^{-1}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, der skalerer alle vektorer med en faktor $1/2$. Og vi har allerede set i eksempel 2.3.7, at matricen

$$B = \text{diag}(1, -1) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

repræsenterer den lineære afbildning $s: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, der er givet ved spejling i x -aksen. Den er invertibel og er sin egen inverse matrix. Endelig repræsenterer matricen

$$C = \text{diag}(1, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

den lineære afbildning $p: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, der projicerer på x -aksen langs y -aksen. Den er ikke invertibel, men opfylder i stedet at $p \circ p = p$. Vi siger generelt, at en lineær afbildning $p: \mathbb{F}^n \rightarrow \mathbb{F}^n$ er en *projektion*, hvis denne formel gælder.

Sætning 2.4.14 Hvis A og B er invertible matricer af samme dimensioner, da er AB også en invertibel matrix, og dens inverse matrix er givet ved

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Bevis Vi sætter $C = B^{-1}A^{-1}$ og udregner

$$(AB)C = (AB)(B^{-1}A^{-1}) = ((AB)B^{-1})A^{-1} = (A(BB^{-1}))A^{-1} = (AI)A^{-1} = AA^{-1} = I,$$

$$C(AB) = (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}(AB)) = B^{-1}((A^{-1}A)B) = B^{-1}(IB) = B^{-1}B = I.$$

Per entydighed af den inverse matrix konkluderer vi derfor, at $C = (AB)^{-1}$. □

Bemærkning 2.4.15 Hvis $\mathbb{F}$ er et legeme, så skriver vi

$$GL_n(\mathbb{F}) \subset M_n(\mathbb{F})$$

for delmængden bestående af de invertible $n \times n$ -matricer. Det følger fra sætning 2.4.14, at matrixproduktet definerer en produktafbildning

$$GL_n(\mathbb{F}) \times GL_n(\mathbb{F}) \xrightarrow{“\cdot”} GL_n(\mathbb{F}),$$

der opfylder gruppeaksiomerne (G1)–(G3) nedenfor. Generelt defineres en *gruppe* til at være et par $(G, \cdot)$ af en mængde G og en afbildning $\cdot: G \times G \rightarrow G$, der opfylder:

(G1) For alle $g, h, k \in G$ er $(g \cdot h) \cdot k = g \cdot (h \cdot k)$.

(G2) Der findes et element $e \in G$, sådan at $g \cdot e = g = e \cdot g$ for alle $g \in G$.

(G3) For alle $g \in G$ findes $h \in G$, sådan at $g \cdot h = e = h \cdot g$.

Elementet $e \in G$, der opfylder (G2), er entydigt bestemt og kaldes for *et-elementet* i G . Man skriver ofte “1” i stedet for “ e ” for dette element. Givet $g \in G$ er elementet $h \in G$, der opfylder (G3), ligeledes entydigt bestemt. Det kaldes for det *inverse element* af $g \in G$ og betegnes $g^{-1} \in G$. Grupper optræder mange steder i både matematik og fysik, fordi de beskriver symmetrier. Den generelle lineære gruppe $GL_n(\mathbb{F})$ er således den største gruppe af symmetrier af vektorrummet $\mathbb{F}^n$.

2 Matricer og lineære afbildninger

$$M_i(c) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & c & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \quad S_{ij}(c) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \\ & & & & c \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \quad T_{ij} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 0 & \\ & & & \ddots \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

Figur 2.7: Operationsmatricerne $M_i(c)$, $S_{ij}(c)$ og T_{ij} adskiller sig fra I_m som anført.

2.5 Operationsmatricer

I dette afsnit introducerer vi *operationsmatricer*, som er matricer med den egenskab, at de ved venstre multiplikation effektuerer de rækkeoperationer, vi definerede i kapitel 1. Vi beviser også flere sætninger fra kapitel 1 i en mere præcis form. Specielt viser vi, at hvis en matrix A ved rækkeoperationer omdannes til en matrix A' på reduceret echelonform, da afhænger A' kun af A og ikke af den valgte følge af rækkeoperationer. Vi anvender endvidere Gauss-elimination til at afgøre, om en matrix er invertibel, og i givet fald finde den inverse matrix.

Vi definerer nu operationsmatricerne, der er kvadratiske matricer og opdelt i tre typer svarende til de respektive rækkeoperationer i kapitel 1.

Definition 2.5.1 Lad $\mathbb{F}$ være et legeme. Operationsmatricerne i $M_m(\mathbb{F})$ er de følgende tre typer af matricer.

Type M: For alle $1 \leq i \leq m$ og $0 \neq c \in \mathbb{F}$ er $M_i(c)$ den matrix, der fås fra I_m ved at multiplicere den i 'te række med c fra venstre.

Type S: For alle $1 \leq i \neq j \leq m$ og $c \in \mathbb{F}$ er $S_{ij}(c)$ den matrix, der fås fra I_m ved at addere c gange den j 'te række til den i 'te række.

Type T: For alle $1 \leq i < j \leq m$ er T_{ij} den matrix, der fås fra I_m ved ombytning af den i 'te række og den j 'te række.

Figur 2.7 illustrerer disse matricer.

Eksempel 2.5.2 For $m = 3$ har vi de følgende eksempler på operationsmatricer

$$M_2(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S_{23}(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, \quad T_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

hvor vi har markeret de indgange, der adskiller sig fra de tilsvarende indgange i identitetsmatricen I_3 , med blåt.

Vi viser nu som lovet, at operationsmatricerne effektuerer rækkeoperationerne fra kapitel 1. Beviset forstås lettest ved sideløbende at betragte eksempel 2.5.4 nedenfor.

Sætning 2.5.3 Lad $A \in M_{m,n}(\mathbb{F})$ være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$.

- (1) For alle $1 \leq i \leq m$ og $0 \neq c \in \mathbb{F}$ er $M_i(c)A$ den matrix, der fremkommer fra A ved at multiplicere den i 'te række med c fra venstre.
- (2) For alle $1 \leq i \neq j \leq m$ og $c \in \mathbb{F}$ er $S_{ij}(c)A$ den matrix, der fremkommer fra A ved at addere c gange den j 'te række til den i 'te række.
- (3) For alle $1 \leq i < j \leq m$ er $T_{ij}A$ den matrix, der fremkommer fra A ved at ombytte den i 'te række og den j 'te række.

Bevis Lad $\delta_r \in M_{1,m}(\mathbb{F})$ være $1 \times m$ -matricen med $(\delta_r)_{1j} = \delta_{rj}$. Matricen δ_r er altså den r 'te række i identitetsmatricen $I_m \in M_{m,m}(\mathbb{F})$, og hvis $B \in M_{m,n}(\mathbb{F})$ er en $m \times n$ -matrix, da er $1 \times n$ -matricen $\delta_r \cdot B \in M_{1,n}(\mathbb{F})$ den r 'te række i B . Dette bruger vi i resten af beviset.

Vi viser nu (1). Per definition af $M_i(c)$ er

$$\delta_r \cdot M_i(c) = \begin{cases} c \cdot \delta_r & \text{hvis } r = i, \\ \delta_r & \text{hvis } r \neq i. \end{cases}$$

Den r 'te række i $M_i(c) \cdot A$ er derfor lig med henholdsvis

$$\delta_r \cdot (M_i(c) \cdot A) = (\delta_r \cdot M_i(c)) \cdot A = (c \cdot \delta_r) \cdot A = c \cdot (\delta_r \cdot A),$$

hvis $r = i$, og

$$\delta_r \cdot (M_i(c) \cdot A) = (\delta_r \cdot M_i(c)) \cdot A = \delta_r \cdot A,$$

hvis $r \neq i$, hvilket netop er påstanden (1).

Vi viser dernæst (2). Operationsmatricen $S_{ij}(c)$ er per definition givet ved

$$\delta_r \cdot S_{ij}(c) = \begin{cases} \delta_r + c \cdot \delta_j & \text{hvis } r = i, \\ \delta_r & \text{hvis } r \neq i, \end{cases}$$

og den r 'te række i $S_{ij}(c) \cdot A$ er derfor lig med henholdsvis

$$\begin{aligned} \delta_r \cdot (S_{ij}(c) \cdot A) &= (\delta_r \cdot S_{ij}(c)) \cdot A = (\delta_r + c \cdot \delta_j) \cdot A \\ &= \delta_r \cdot A + (c \cdot \delta_j) \cdot A = \delta_r \cdot A + c \cdot (\delta_j \cdot A), \end{aligned}$$

hvis $r = i$, og

$$\delta_r \cdot (S_{ij}(c) \cdot A) = (\delta_r \cdot S_{ij}(c)) \cdot A = \delta_r \cdot A,$$

hvis $r \neq i$. Dette viser påstanden (2).

2 Matricer og lineære afbildninger

Endelig viser vi (3). Operationsmatricen T_{ij} er defineret ved

$$\delta_r \cdot T_{ij} = \begin{cases} \delta_j & \text{hvis } r = i, \\ \delta_i & \text{hvis } r = j, \\ \delta_r & \text{hvis } r \neq i \text{ og } r \neq j, \end{cases}$$

og derfor er den r 'te række i $T_{ij} \cdot A$ lig med

$$\delta_r \cdot (T_{ij} \cdot A) = (\delta_r \cdot T_{ij}) \cdot A = \begin{cases} \delta_j \cdot A & \text{hvis } r = i, \\ \delta_i \cdot A & \text{hvis } r = j, \\ \delta_r \cdot A & \text{hvis } r \neq i \text{ og } r \neq j, \end{cases}$$

hvilket viser (3). □

Eksempel 2.5.4 Vi multiplicerer 3×4 -matricen

$$A = \begin{pmatrix} 1 & 3 & 2 & 0 \\ 2 & 5 & 2 & 4 \\ 0 & 1 & 2 & 1 \end{pmatrix}$$

med operationsmatricerne fra eksempel 2.5.2:

$$\begin{aligned} M_2(c)A &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 & 0 \\ 2 & 5 & 2 & 4 \\ 0 & 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 0 \\ c \cdot 2 & c \cdot 5 & c \cdot 2 & c \cdot 4 \\ 0 & 1 & 2 & 1 \end{pmatrix} \\ S_{23}(c)A &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 & 0 \\ 2 & 5 & 2 & 4 \\ 0 & 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 0 \\ 2+c \cdot 0 & 5+c \cdot 1 & 2+c \cdot 2 & 4+c \cdot 1 \\ 0 & 1 & 2 & 1 \end{pmatrix} \\ T_{13}A &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 & 0 \\ 2 & 5 & 2 & 4 \\ 0 & 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 2 & 5 & 2 & 4 \\ 1 & 3 & 2 & 0 \end{pmatrix} \end{aligned}$$

Vi ser, at $M_2(c)A$ er matricen, der fremkommer fra A ved multiplikation af den anden række med c fra venstre; at matricen $S_{23}(c)A$ er matricen, der fremkommer fra A ved at addere c gange den tredje række til den anden række; og at $T_{13}A$ er den matrix, der fremkommer fra A ved ombytning af den første og tredje række.

I kapitel 1 indførte vi en invers rækkeoperation hørende til hver rækkeoperation. Oversat til operationsmatricer, giver det følgende resultat.

Sætning 2.5.5 Operationsmatricerne $M_i(c)$, $S_{ij}(c)$ og T_{ij} er alle invertible, og deres inverse matricer er givet ved henholdsvis

$$M_i(c)^{-1} = M_i(c^{-1}), \quad S_{ij}(c)^{-1} = S_{ij}(-c) \quad \text{og} \quad T_{ij}^{-1} = T_{ij}.$$

Bevis Hvis A er lig med $M_i(c)$, $S_{ij}(c)$ eller T_{ij} , så sætter vi B til at være lig med henholdsvis $M_i(c^{-1})$, $S_{ij}(-c)$ eller T_{ij} , og skal da vise, at $AB = I = BA$. Nu er

$$AB = (AB)I = A(BI),$$

og ifølge sætning 2.5.3 er $A(BI)$ den matrix, der fremkommer fra identitetsmatricen ved først at anvende den rækkeoperation, der svarer til B , og derefter den rækkeoperation, der svarer til A . I alle tre tilfælde, er disse rækkeoperationer præcis hinandens inverse rækkeoperationer. Dette viser, at $AB = I$, og $BA = I$ vises tilsvarende. Alternativt kan sætningen også let bevises ved direkte udregning af de relevante matrixprodukter. $\square$

Eksempel 2.5.6 Vi udregner, at

$$S_{23}(c)S_{23}(-c) = \begin{pmatrix} \boxed{1} & \boxed{0} & \boxed{0} \\ \boxed{0} & \boxed{1} & \boxed{c} \\ \boxed{0} & \boxed{0} & \boxed{1} \end{pmatrix} \begin{pmatrix} \boxed{1} & \boxed{0} & \boxed{0} \\ \boxed{0} & \boxed{1} & \boxed{-c} \\ \boxed{0} & \boxed{0} & \boxed{1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -c+c \\ 0 & 0 & 1 \end{pmatrix} = I,$$

og ved at erstatte c med $-c$ ser vi, at også $S_{23}(-c)S_{23}(c) = I$, hvilket viser, at de to matricer er hinandens inverse.

Definition 2.5.7 En ækvivalensrelation på en mængde X er en delmængde $R \subset X \times X$, der opfylder de følgende aksiomer (R1)–(R3).

(R1) For alle $x \in X$ er $(x, x) \in R$.

(R2) For alle $x, y \in X$ gælder det, at $(x, y) \in R$, hvis og kun hvis $(y, x) \in R$.

(R3) For alle $x, y, z \in X$ gælder det, at hvis $(x, y) \in R$ og $(y, z) \in R$, da er $(x, z) \in R$.

Vi tænker på “ $(x, y) \in R$ ” som “ x er R -ækvivalent med y ”. En ækvivalensrelation R på en mængde X giver anledning til en opdeling af X i disjunkte ækvivalensklasser, hvor to elementer $x, y \in X$ per definition tilhører den samme R -ækvivalensklasse, hvis og kun hvis $(x, y) \in R$. Vi anvender rækkeoperationer til at definere en ækvivalensrelation på mængden $M_{m,n}(\mathbb{F})$ af $m \times n$ -matricer med indgange i et legeme $\mathbb{F}$.

2 Matricer og lineære afbildninger

Definition 2.5.8 Hvis $A, B \in M_{m,n}(\mathbb{F})$ er $m \times n$ -matricer med indgange i et legeme $\mathbb{F}$, da siges matricen A at være *rækkeækvivalent* til matricen B , hvis der findes en endelig følge af operationsmatricer $P_1, \dots, P_k \in M_m(\mathbb{F})$, sådan at $B = P_k \cdots P_2 P_1 A$.

Ifølge sætning 2.5.3 er A altså rækkeækvivalent til B , hvis B fremkommer fra A ved at anvende endeligt mange rækkeoperationer. Vi viser nu, at rækkeækvivalens er en ækvivalensrelation på mængden $M_{m,n}(\mathbb{F})$ af $m \times n$ -matricer med indgange i $\mathbb{F}$.

Sætning 2.5.9 Lad $\mathbb{F}$ være et legeme, og lad m og n være naturlige tal.

- (1) For alle $A \in M_{m,n}(\mathbb{F})$ gælder det, at A er rækkeækvivalent til A .
- (2) For alle $A, B \in M_{m,n}(\mathbb{F})$ gælder det, at A er rækkeækvivalent til B , hvis og kun hvis B er rækkeækvivalent til A .
- (3) For alle $A, B, C \in M_{m,n}(\mathbb{F})$ gælder det, at hvis A er rækkeækvivalent til B , og hvis B er rækkeækvivalent til C , så er A rækkeækvivalent til C .

Bevis En matrix A fremkommer fra sig selv ved at anvende $k = 0$ rækkeoperationer, hvilket viser (1). Hvis $B = P_k \cdots P_1 A$, da er $A = P_1^{-1} \cdots P_k^{-1} B$, hvilket viser (2). Her har vi anvendt, at operationsmatricer ifølge sætning 2.5.3 er invertible. Hvis $B = P_k \cdots P_1 A$ og $C = Q_l \cdots Q_1 B$, da er $C = Q_l \cdots Q_1 P_k \cdots P_1 A$, hvilket viser (3) og dermed sætningen. $\square$

Vi viste i sætning 1.2.5, at hvis totalmatricen for et ligningssystem kan omdannes til totalmatricen for et andet ligningssystem ved hjælp af rækkeoperationer, da har de to ligningssystemer den samme løsningsmængde. Vi omformulerer nu denne sætning og gentager beviset ved hjælp af operationsmatricer.

Sætning 2.5.10 Hvis totalmatricerne $(A \mid \mathbf{b})$ og $(B \mid \mathbf{c})$ for ligningssystemerne $A\mathbf{x} = \mathbf{b}$ og $B\mathbf{x} = \mathbf{c}$ er rækkeækvivalente, så har de to ligningssystemer samme løsningsmængde.

Bevis Hvis $(B \mid \mathbf{c}) = P_k \cdots P_1 (A \mid \mathbf{b})$, da er specielt $B = P_k \cdots P_1 A$ og $\mathbf{c} = P_k \cdots P_1 \mathbf{b}$. Hvis derfor $A\mathbf{x} = \mathbf{b}$, da er også $B\mathbf{x} = P_k \cdots P_1 A\mathbf{x} = P_k \cdots P_1 \mathbf{b} = \mathbf{c}$. Omvendt, hvis $B\mathbf{x} = \mathbf{c}$, da er også $A\mathbf{x} = P_1^{-1} \cdots P_k^{-1} B\mathbf{x} = P_1^{-1} \cdots P_k^{-1} \mathbf{c} = \mathbf{b}$. $\square$

Vi vender nu tilbage til Gauss-elimination og beviser som lovet den følgende mere præcise version af sætning 1.2.12. Denne sætning er vores første hovedsætning.

Sætning 2.5.11 *Lad A være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$.*

(1) *Der findes en endelig følge af operationsmatricer $P_1, \dots, P_k \in M_m(\mathbb{F})$, sådan at*

$$A' = P_k \dots P_2 P_1 A$$

er på reduceret echelonform af rang $0 \leq r \leq \min\{m, n\}$.

(2) *Matricen A' og dens rang r er entydigt bestemte af matricen A og afhænger ikke af den anvendte følge af operationsmatricer.*

Bevis Udsagnet (1) er ifølge sætning 2.5.3 ækvivalent til sætning 1.2.12, så vi viser (2). Sætning 2.5.9 viser, at rækkeækvivalens er en ækvivalensrelation på $M_{m,n}(\mathbb{F})$, og det udsagn, vi skal vise, er, at hver rækkeækvivalensklasse præcis indeholder én matrix på reduceret echelonform. Med andre ord skal vi vise, at hvis A og B er rækkeækvivalente $m \times n$ -matricer, der begge er på reduceret echelonform, da er $A = B$, hvilket vi gør ved induktion på $n \geq 1$. For $n = 1$ er de eneste $m \times 1$ -matricer på reduceret echelonform

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

og for enhver operationsmatrix $P \in M_m(\mathbb{F})$ er $P\mathbf{0} = \mathbf{0}$, hvilket viser, at $\mathbf{0}$ og $\mathbf{e}_1$ ikke er rækkeækvivalente. Derfor er to rækkeækvivalente $m \times 1$ -matricer A og B nødvendigvis ens, hvilket viser påstanden for $n = 1$.

Vi antager induktivt, at påstanden er vist for $n = p - 1$ og beviser den for $n = p$. Hertil lader vi C og D være de $m \times (p - 1)$ -matricer, der fremkommer fra henholdsvis A og B ved at fjerne den sidste søjle, sådan at

$$A = \left(C \left| \begin{array}{c} a_{1p} \\ \vdots \\ a_{mp} \end{array} \right. \right) \quad \text{og} \quad B = \left(D \left| \begin{array}{c} b_{1p} \\ \vdots \\ b_{mp} \end{array} \right. \right).$$

Da A og B er på reduceret echelonform, så er også C og D på reduceret echelonform, og de er endvidere rækkeækvivalente. For hvis $P_1, \dots, P_k$ er operationsmatricer, sådan at $B = P_k \dots P_1 A$, da er også $C = P_k \dots P_1 D$. Den induktive antagelse viser derfor, at $m \times (p - 1)$ -matricerne C og D er ens, hvorefter vi konkluderer, at

$$A - B = \left(C - D \left| \begin{array}{c} a_{1p} - b_{1p} \\ \vdots \\ a_{mp} - b_{mp} \end{array} \right. \right) = \left(O \left| \begin{array}{c} a_{1p} - b_{1p} \\ \vdots \\ a_{mp} - b_{mp} \end{array} \right. \right).$$

2 Matricer og lineære afbildninger

Vi ønsker at vise, at også den sidste søjle er lig med $\mathbf{0}$, sådan at $A = B$, og antager derfor modsætningsvist, at der findes et $1 \leq j \leq p$, sådan at $a_{jp} \neq b_{jp}$. Da A og B er rækkeækvivalente, så har ligningssystemerne $A\mathbf{x} = \mathbf{0}$ og $B\mathbf{x} = \mathbf{0}$ ifølge sætning 2.5.10 samme løsningsmængde. Hvis $\mathbf{x}$ er en løsning til disse ligningssystemer, da er

$$(A - B)\mathbf{x} = A\mathbf{x} - B\mathbf{x} = \mathbf{0} - \mathbf{0} = \mathbf{0},$$

og $\mathbf{x}$ er derfor også en løsning til ligningssystemet $(A - B)\mathbf{x} = \mathbf{0}$. Den j 'te ligning i dette ligningssystem er " $(a_{jp} - b_{jp})x_p = 0$ ", og da vi har antaget, at $a_{jp} - b_{jp} \neq 0$, følger det, at $x_p = 0$. Så enhver løsning $\mathbf{x}$ til $A\mathbf{x} = \mathbf{0}$ og $B\mathbf{x} = \mathbf{0}$ har nødvendigvis $x_p = 0$, og derfor er x_p ikke en fri variabel i disse ligningssystemer. Da A og B er på reduceret echelonform medfører dette, at deres p 'te søjler begge indeholder en ledende indgang. Vi har allerede set, at $m \times (p - 1)$ -matricerne C og D , der er på reduceret echelonform, er identiske, og derfor har de specielt samme rang $r - 1$. Da $m \times p$ -matricerne A og B er på reduceret echelonform, og da deres sidste søjler indeholder en ledende indgang, er disse søjler derfor begge nødvendigvis lig med $\mathbf{e}_r$. Dermed er $A = B$, hvilket strider mod vores antagelse, at $A \neq B$, så denne antagelse var altså forkert. Så $A = B$, hvilket viser induktionsskridtet og dermed sætningen. $\square$

Bemærkning 2.5.12 Som nævnt i beviset, så siger sætning 2.5.11 ækvivalent, at hver rækkeækvivalensklasse af $m \times n$ -matricer præcis indeholder én $m \times n$ -matrix, der er på reduceret echelonform.

Vi har i definition 1.2.9 defineret rangen af en matrix på reduceret echelonform til at være antallet af ikke-nul rækker eller ækvivalent antallet af ledende indgange. Ifølge sætning 2.5.11 kan vi nu udvide denne definition til alle matricer, som følger.

Definition 2.5.13 Lad A være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$ og lad A' være den entydigt bestemte matrix på reduceret echelonform, der er rækkeækvivalent med A . Da er *rangen* af A defineret til at være rangen af A' og betegnes $\text{rank}(A)$.

Vi bemærker, at lemma 1.3.2 viser, at rangen af en matrix på echelonform også er lig med antallet af ikke-nul rækker eller ækvivalent antallet af ledende indgange.

Eksempel 2.5.14 (1) Nul-matricen $O_{m,n}$ har rang 0.

(2) Matricen

$$A = \begin{pmatrix} 8 & 1 & 3 & 3 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

er på echelonform og har tre ledende indgange, som vi har markeret med blåt. Så ifølge lemma 1.3.2 har den rang 3.

(3) Vi bestemmer rangen af matricen

$$B = \begin{pmatrix} 2 & 0 & 3 & 3 \\ 0 & 1 & 4 & 0 \\ 2 & 1 & 7 & 3 \end{pmatrix}$$

ved at omdanne denne til en rækkeækvivalent matrix C på echelonform.

$$\begin{aligned} B &= \begin{pmatrix} 2 & 0 & 3 & 3 \\ 0 & 1 & 4 & 0 \\ 2 & 1 & 7 & 3 \end{pmatrix} + (-1)R_1 \\ &= \begin{pmatrix} 2 & 0 & 3 & 3 \\ 0 & 1 & 4 & 0 \\ 0 & 1 & 4 & 0 \end{pmatrix} (-1)R_2 \\ C &= \begin{pmatrix} 2 & 0 & 3 & 3 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Da B og C er rækkeækvivalente, har de samme rang, og da C er på echelonform, er dens rang lig med antallet af ledende indgange. Så B har rang 2.

Vi skal senere vise, at rangen af $A \in M_{m,n}(\mathbb{F})$ er lig med dimensionen af billedet af den lineære afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ givet ved $f(\mathbf{x}) = A\mathbf{x}$, men det må vente til, vi har defineret vektorrum og deres dimension i kapitel 4. Vi relaterer herunder rangen af matricen A til injektivitet og surjektivitet af afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$.

Sætning 2.5.15 *Lad $\mathbb{F}$ være et legeme, lad A være en $m \times n$ -matrix med indgange i $\mathbb{F}$, lad $0 \leq r \leq \min\{m, n\}$ være rangen af A og lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ være den lineære afbildning defineret ved $f(\mathbf{x}) = A\mathbf{x}$. Følgende gælder:*

- (1) *Afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er surjektiv, hvis og kun hvis $r = m$.*
- (2) *Afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er injektiv, hvis og kun hvis $r = n$.*
- (3) *Afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er bijektiv, hvis og kun hvis $r = m = n$.*

Bevis Per definition er $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ injektiv (resp. surjektiv, resp. bijektiv), hvis og kun hvis ligningssystemet $A\mathbf{x} = \mathbf{b}$ har højst én løsning (resp. mindst én løsning, resp. præcis én løsning) for alle $\mathbf{b} \in \mathbb{F}^m$. Sætning følger derfor fra sætning 1.3.1 og definition 2.5.13. $\square$

Da rangen r af en matrix ikke kan være større end antallet af rækker eller antallet af søjler, medfører sætning 2.5.15 umiddelbart følgende resultat.

2 Matricer og lineære afbildninger

Korollar 2.5.16 Lad $\mathbb{F}$ være et legeme, og lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ være en lineær afbildning.

- (1) Hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er surjektiv, da er $n \geq m$.
- (2) Hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er injektiv, da er $n \leq m$.
- (3) Hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er bijektiv, da er $n = m$.

Eksempel 2.5.17 Vi betragter 3×4 -matricerne A og B fra eksempel 2.5.14. Matricen A har rang 3, så afbildningen $f: \mathbb{F}^4 \rightarrow \mathbb{F}^3$ givet ved $f(\mathbf{x}) = A\mathbf{x}$ er surjektiv. Derimod har matricen B kun rang 2, så afbildningen $g: \mathbb{F}^4 \rightarrow \mathbb{F}^3$ givet ved $g(\mathbf{x}) = B\mathbf{x}$ er ikke surjektiv. Det er da også let at se, at ligningssystemet $B\mathbf{x} = \mathbf{e}_3$ ikke har nogen løsninger.

Blandt kvadratiske matricer af orden n er identitetsmatricen den eneste matrix, der både er på reduceret echelonform og af rang $r = n$. Denne bemærkning har følgende nyttige konsekvens.

Sætning 2.5.18 Lad $\mathbb{F}$ være et legeme. En kvadratisk matrix $A \in M_n(\mathbb{F})$ af orden n er invertibel, hvis og kun hvis den har rang $r = n$. I givet fald findes en endelig følge af operationsmatricer $P_1, \dots, P_k \in M_n(\mathbb{F})$, sådan at $I = A' = P_k \cdots P_1 A$, og

$$A^{-1} = P_k \cdots P_1.$$

Bevis Ifølge sætning 2.4.9 er A invertibel hvis og kun hvis den lineære afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ givet ved $f(\mathbf{x}) = A\mathbf{x}$ er invertibel, og ifølge sætning 2.5.15 (3) gælder dette, hvis og kun hvis matricen A har rang $r = n$. I givet fald findes der operationsmatricer $P_1, \dots, P_k \in M_n(\mathbb{F})$, sådan at $I = A' = P_k \cdots P_1 A$, og vi konkluderer derfor, at

$$A^{-1} = I A^{-1} = (P_k \cdots P_1 A) A^{-1} = P_k \cdots P_1 (A A^{-1}) = P_k \cdots P_1 I = P_k \cdots P_1$$

som ønsket. □

Sætning 2.5.11 giver følgende effektive metode til at bestemme den inverse matrix af en invertibel matrix.

Korollar 2.5.19 En $n \times n$ -matrix $A \in M_n(\mathbb{F})$ med indgange i et legeme $\mathbb{F}$ er invertibel, hvis og kun hvis $n \times 2n$ -matricen $(A \mid I) \in M_{n,2n}(\mathbb{F})$ ved rækkeoperationer kan omdannes til en $n \times 2n$ -matrix på formen $(I \mid B) \in M_{n,2n}(\mathbb{F})$, og i givet fald er $A^{-1} = B$.

Bevis Ifølge sætning 2.5.18 er A invertibel, hvis og kun hvis den ved rækkeoperationer kan omdannes til identitetsmatricen, og dette er tilfældet, hvis og kun hvis $(A \mid I)$ ved rækkeoperationer kan omdannes til en $n \times 2n$ -matrix på formen $(I \mid B)$. I givet fald lader vi $P_1, \dots, P_k \in M_n(\mathbb{F})$ være en følge af operationsmatricer, sådan at $P_k \cdots P_1 A = I$. Sætning 2.5.18 viser da, at

$$P_k \cdots P_1(A \mid I) = (P_k \cdots P_1 A \mid P_k \cdots P_1 I) = (I \mid A^{-1}),$$

og derfor er $B = A^{-1}$ som ønsket. $\square$

Eksempel 2.5.20 Vi anvender korollar 2.5.19 til at undersøge om 3×3 -matricen

$$A = \begin{pmatrix} 8 & 5 & 3 \\ 2 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

er invertibel og i givet fald finde dens inverse matrix. Vi omdanner derfor 3×6 -matricen $(A \mid I)$ til en matrix på reduceret echelonform ved hjælp af rækkeoperationer:

$$\begin{aligned} (A \mid I) &= \left(\begin{array}{ccc|ccc} 8 & 5 & 3 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} +(-3)R_3 \\ +(-1)R_3 \end{array} \\ &= \left(\begin{array}{ccc|ccc} -1 & 2 & 0 & 1 & 0 & -3 \\ -1 & 1 & 0 & 0 & 1 & -1 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} +(-1)R_1 \\ +3R_1 \end{array} \\ &= \left(\begin{array}{ccc|ccc} -1 & 2 & 0 & 1 & 0 & -3 \\ 0 & -1 & 0 & -1 & 1 & 2 \\ 0 & 7 & 1 & 3 & 0 & -8 \end{array} \right) \begin{array}{l} +2R_2 \\ +7R_2 \end{array} \\ &= \left(\begin{array}{ccc|ccc} -1 & 0 & 0 & -1 & 2 & 1 \\ 0 & -1 & 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & -4 & 7 & 6 \end{array} \right) \begin{array}{l} (-1) \cdot R_1 \\ (-1) \cdot R_2 \end{array} \\ (I \mid B) &= \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & -1 \\ 0 & 1 & 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & -4 & 7 & 6 \end{array} \right) \end{aligned}$$

Her har vi som tidligere indikeret de indgange, vi ønsker at ændre med rødt, og de ledende indgange med blå. Vi konkluderer heraf, at A er invertibel, og at

$$A^{-1} = B = \begin{pmatrix} 1 & -2 & -1 \\ 1 & -1 & -2 \\ -4 & 7 & 6 \end{pmatrix}$$

er den inverse matrix.

Eksempel 2.5.21 Vi anvender korollar 2.5.19 til at undersøge om 3×3 -matricen

$$A = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -2 \\ 1 & 2 & 0 \end{pmatrix}$$

er invertibel og i givet fald finde dens inverse matrix. Vi omdanner derfor 3×6 -matricen $(A | I)$ til en matrix på reduceret echelonform ved hjælp af rækkeoperationer:

$$\begin{aligned} (A | I) &= \left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ -1 & 0 & -2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} +(-2)R_3 \\ +R_3 \end{array} \\ &= \left(\begin{array}{ccc|ccc} 0 & -1 & 1 & 1 & 0 & -2 \\ 0 & 2 & -2 & 0 & 1 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array} \right) R_1 \leftrightarrow R_3 \\ &= \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 2 & -2 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 & 0 & -2 \end{array} \right) +2R_3 \\ &= \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 & 2 \\ 0 & -1 & 1 & 1 & 0 & -2 \end{array} \right) R_2 \leftrightarrow R_3 \\ &= \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & -1 & 1 & 2 \end{array} \right) \end{aligned}$$

Her har vi som tidligere indikeret de indgange, vi ønsker at ændre med rødt, og de ledende indgange med blå. Matricen er på echelonform, og vi kan derfor aflæse, at A har rang $r = 2 < 3 = n$, og derfor er A ikke invertibel.

For en kvadratisk matrix A er de tre udsagn (1)–(3) i sætning 1.3.1 identiske, da antallet af rækker og søjler er ens. Dette har følgende bemærkelsesværdige konsekvens.

Sætning 2.5.22 Hvis A og B er $n \times n$ -matricer med indgange i et legeme $\mathbb{F}$, da er de følgende udsagn ækvivalente:

- (1) $AB = I$.
- (2) $BA = I$.
- (3) $AB = I$ og $BA = I$.

Bevis Det er klart, at (3) medfører både (1) og (2), så vi skal vise, at (1) medfører (3), og at (2) medfører (3). Vi viser nu, at (1) medfører (3), og ved at ombytte A og B viser det samme argument, at også (2) medfører (3). Ifølge sætning 2.5.18 er (3) ækvivalent til udsagnet, at $r = n$, så det er nok at vise, at (1) medfører, at $r = n$. Men $AB = I$ medfører, at ligningssystemet $A\mathbf{x} = \mathbf{b}$ har mindst en løsning, nemlig $\mathbf{x} = B\mathbf{b}$, og derfor viser sætning 1.3.1, at $r = n$ som ønsket. Dette beviser sætningen. $\square$

Bemærkning 2.5.23 Sætning 2.5.22 gælder kun for *kvadratiske* matricer, også selvom matrixprodukterne AB og BA begge findes. Eksemplet

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = I_1 \quad \text{og} \quad BA = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq I_2$$

viser, at udsagnene “ $AB = I$ ” og “ $BA = I$ ” generelt ikke er ækvivalente.

På samme måde specialiserer sætning 1.3.1 for $m = n$ til følgende resultat, der kun gælder for lineære afbildninger. For eksempel er $\exp: \mathbb{R} \rightarrow \mathbb{R}$ injektiv men ikke surjektiv.

Korollar 2.5.24 Hvis $\mathbb{F}$ er et legeme og $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ en lineær afbildning, da er de følgende udsagn ækvivalente:

- (1) Afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ er surjektiv.
- (2) Afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ er injektiv.
- (3) Afbildningen $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ er bijektiv.

Bemærkning 2.5.25 Vi skal senere i kapitel 3 også anvende søjleoperationer til at definere og udregne determinanten. Ligesom rækkeoperationer svarer til at gange med operationsmatricer fra venstre, så svarer søjleoperationer til at gange med disse fra højre. Hvis $A \in M_{m,n}(\mathbb{F})$ er en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$, da gælder:

- (1) For alle $n \geq j \geq 1$ og $0 \neq c \in \mathbb{F}$ er $AM_j(c)$ den matrix, der fremkommer fra A ved at gange den j 'te søjle med c fra højre.
- (2) For alle $n \geq k \neq j \geq 1$ og $c \in \mathbb{F}$ er $AS_{kj}(c)$ den matrix, der fremkommer fra A ved at adderer den k 'te søjle gange c til den j 'te søjle.
- (3) For alle $n \geq k > j \geq 1$ er AT_{kj} den matrix, der fremkommer fra A ved at ombytte den j 'te søjle og den k 'te søjle.

Her er alle operationsmatricerne kvadratiske matricer af orden n . Vi understreger dog, at søjleoperationer *ikke* kan bruges til at løse lineære ligningssystemer, da de modsat rækkeoperationer ændrer løsningsmængden.

2.6 Hermitiske former

På legemet af komplekse tal $\mathbb{C}$ findes der en meget interessant ekstra struktur, nemlig kompleks konjugering, som er afbildningen $(-)^*: \mathbb{C} \rightarrow \mathbb{C}$ der til $z = a + ib \in \mathbb{C}$ tilordner

$$z^* = \bar{z} = a - ib.$$

Kompleks konjugering opfylder aksiomerne, at for alle $z, w \in \mathbb{C}$ er

$$(z + w)^* = z^* + w^*, \quad (z \cdot w)^* = w^* \cdot z^* \quad \text{og} \quad (z^*)^* = z,$$

og den er vores hovedeksempel på en *skævinvolution*, som defineret nedenfor.

Givet et (skæv)legeme $\mathbb{F}$ med en skævinvolution $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$ såsom $\mathbb{C}$ med kompleks konjugering $(-)^*: \mathbb{C} \rightarrow \mathbb{C}$, definerer vi dette afsnit den adjungerede matrix A^* hørende til en matrix A med indgange i $\mathbb{F}$ og viser, at denne operation på matricer opfylder de samme aksiomer som kompleks konjugering.² Det er i denne forbindelse vigtigt at huske, at $(-)^*$ ombytter multiplikationsrækkefølgen.

Kvadratiske matricer A , som opfylder at $A^* = A$, kaldes for *hermitiske* matricer. Hvis $\mathbb{F} = \mathbb{C}$ og $(-)^*: \mathbb{C} \rightarrow \mathbb{C}$ er kompleks konjugering, da spiller de hermitiske matricer en speciel rolle, blandt andet fordi de, som vi skal se i kapitel 6, har reelle egenverdier. Hvis $\mathbb{F} = \mathbb{R}$ og $(-)^*: \mathbb{R} \rightarrow \mathbb{R}$ er identitetsafbildningen, da kaldes de hermitiske matricer for *symmetriske* matricer og spiller ligeledes en særlig rolle.

Generelt giver en hermitisk $n \times n$ -matrix A anledning til en *hermitisk form* på $\mathbb{F}^n$, som er afbildningen $\langle -, - \rangle: \mathbb{F}^n \times \mathbb{F}^n \rightarrow \mathbb{F}$ defineret ved $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* A \mathbf{y}$, og vores hovedformål i dette afsnit er at studere hermitiske former. Disse anvendes til at beskrive geometriske strukturer på $\mathbb{F}^n$ i bred forstand. For eksempel er Einstein-geometrien i speciel relativitetsteori beskrevet af Minkowskiformen

$$\langle \mathbf{x}, \mathbf{y} \rangle = -x_1^* y_1 + x_2^* y_2 + \cdots + x_n^* y_n = -x_1 y_1 + x_2 y_2 + \cdots + x_n y_n,$$

der svarer til den symmetriske reelle matrix $A = \text{diag}(-1, 1, \dots, 1)$.

Vi vil nedenfor arbejde over et generelt skævlegeme $\mathbb{F}$, hvor det ikke nødvendigvis gælder, at $ab = ba$, selvom vi stadigvæk har $\mathbb{F} = \mathbb{R}$ og $\mathbb{F} = \mathbb{C}$ som hovedeksempler.

Definition 2.6.1 En *skævinvolution* på et skævlegeme $\mathbb{F}$ er en afbildning $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$, sådan at følgende gælder for alle $a, b \in \mathbb{F}$:

$$(I1) \quad (a + b)^* = a^* + b^*$$

$$(I2) \quad (ab)^* = b^* a^*$$

$$(I3) \quad 1^* = 1$$

$$(I4) \quad (a^*)^* = a$$

²Fysiklitteraturen anvender betegnelsen $A^\dagger$ i stedet for A^* for den adjungerede matrix

Vi siger “skæv” og “involution” for at antyde henholdsvis (I2) og (I4).

Eksempel 2.6.2 (1) På $\mathbb{F} = \mathbb{R}$ er identitetsafbildningen en skævinvolution, og man kan vise, at der ikke findes andre.

(2) På $\mathbb{F} = \mathbb{C}$ er både identitetsafbildningen og kompleks konjugering skævinvolutioner, men vi vil udelukkende betragte kompleks konjugering.

(3) På kvaternionerne $\mathbb{F} = \mathbb{H}$ fra eksempel 0.3.1 er kvaternionisk konjugering, som er afbildningen $(-)^*: \mathbb{H} \rightarrow \mathbb{H}$, der til $q = a + ib + jc + kd$ tilordner

$$q^* = \bar{q} = a - ib - jc - kd,$$

en skævinvolution. Derimod er identitetsafbildningen ikke en skævinvolution, da (I2) ikke er opfyldt.

Vi bemærker, at det for $z \in \mathbb{C}$ gælder, at $\bar{\bar{z}} = z$, hvis og kun hvis $z \in \mathbb{R}$. Dette er en af grundene til, at vi altid anvender kompleks konjugering som skævinvolution på $\mathbb{C}$.

Definition 2.6.3 Lad $\mathbb{F}$ være et skævlegeme med skævinvolution $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$. Da er den *adjungerede matrix* af en $m \times n$ -matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

med indgange i $\mathbb{F}$ den $n \times m$ -matrix med indgange i $\mathbb{F}$, der er givet ved

$$A^* = \begin{pmatrix} a_{11}^* & a_{21}^* & \dots & a_{m1}^* \\ a_{12}^* & a_{22}^* & \dots & a_{m2}^* \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}^* & a_{2n}^* & \dots & a_{mn}^* \end{pmatrix}.$$

En kvadratisk matrix A med indgange i $\mathbb{F}$ er *hermitisk*, hvis $A^* = A$.

Vi bemærker, at antallet af rækker i A^* er lig med antallet af søjler i A og vice versa. Det giver altså kun mening at spørge om $A^* = A$, hvis A er en kvadratisk matrix.

2 Matricer og lineære afbildninger

Eksempel 2.6.4 Vi betragter de følgende komplekse matricer og deres adjungerede matricer, hvor vi anvender kompleks konjugering som skævinvolution.

$$\begin{aligned} A &= \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 0 & 3 \end{pmatrix}, & A^* &= \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 3 \end{pmatrix}. \\ B &= \begin{pmatrix} 2 & -1 & -3+2i \\ 0 & 1+i & 2 \\ 0 & 0 & -1 \end{pmatrix}, & B^* &= \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1-i & 0 \\ -3-2i & 2 & -1 \end{pmatrix}. \\ C &= \begin{pmatrix} -2 & i \\ -i & -1 \end{pmatrix}, & C^* &= \begin{pmatrix} -2 & i \\ -i & -1 \end{pmatrix}. \end{aligned}$$

Matricerne B og C er kvadratiske, og det er de adjungerede matricer B^* og C^* dermed også. Matricen C er endvidere hermitisk.

Bemærkning 2.6.5 Hvis $\mathbb{F}$ er et legeme, da er identitetsafbildningen $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$ en skævinvolution. Den adjungerede matrix A^* med hensyn til denne skævinvolution kaldes for den *transponerede matrix* og betegnes A^t , mens en kvadratisk matrix A , der opfylder $A^t = A$, kaldes for en *symmetrisk matrix*.

Eksempel 2.6.6 Vi betragter de følgende reelle matricer og deres transponerede matricer.

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 & 4 \\ 0 & 7 & 4 \\ 3 & 0 & 4 \end{pmatrix}, & A^* = A^t &= \begin{pmatrix} 1 & 0 & 3 \\ 2 & 7 & 0 \\ 4 & 4 & 4 \end{pmatrix} \\ B &= \begin{pmatrix} 1 & 2 & 4 \\ 2 & 7 & 4 \\ 4 & 4 & 4 \end{pmatrix}, & B^* = B^t &= \begin{pmatrix} 1 & 2 & 4 \\ 2 & 7 & 4 \\ 4 & 4 & 4 \end{pmatrix} \\ C &= \begin{pmatrix} 6 \\ 2 \\ -1 \end{pmatrix}, & C^* = C^t &= (6 \quad 2 \quad -1) \end{aligned}$$

hvor $B = B^t$ er symmetrisk. Bemærk også at den adjungerede matrix af en søjlevektor er en rækkevektor og vice versa.

Vi viser nu, at adjungering af matricer har præcis de samme egenskaber som en skævinvolution, hvilket er grunden til, at vi anvender den samme notation for begge.

Sætning 2.6.7 Lad $\mathbb{F}$ være et skævlegeme og lad $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$ være en skævinvolution. For matrixer af passende dimensioner med indgange i $\mathbb{F}$ gælder følgende:

$$(I1) \quad (A + B)^* = A^* + B^*$$

$$(I2) \quad (AB)^* = B^* A^*$$

$$(I3) \quad I^* = I$$

$$(I4) \quad (A^*)^* = A$$

Bevis For $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$ skriver vi $A^* = (a'_{ij}) \in M_{n,m}(\mathbb{F})$, sådan at $a'_{ij} = (a_{ji})^*$ per definition af den adjungerede matrix. For at vise (I1) lader vi $A, B \in M_{m,n}(\mathbb{F})$ og sætter $C = A + B \in M_{m,n}(\mathbb{F})$. Udregningen

$$c'_{ij} = c^*_{ji} = (a_{ji} + b_{ji})^* = a^*_{ji} + b^*_{ji} = a'_{ij} + b'_{ij}$$

viser da, at $C^* = (A + B)^*$ som ønsket. For at vise (I2) lader vi tilsvarende $A \in M_{m,n}(\mathbb{F})$ og $B \in M_{n,p}(\mathbb{F})$ og sætter $C = AB \in M_{m,p}(\mathbb{F})$, hvorefter udregningen

$$c'_{ik} = c^*_{ki} = \left(\sum_{j=1}^n a_{kj} b_{ji} \right)^* = \sum_{j=1}^n (a_{kj} b_{ji})^* = \sum_{j=1}^n b^*_{ji} a^*_{kj} = \sum_{j=1}^n b'_{ij} a'_{jk}$$

viser, at $C^* = B^* A^*$. Videre følger (I3) fra $\delta^*_{ji} = \delta_{ij} = \delta_{ij}$, idet $I = (\delta_{ij})$ og $I^* = (\delta^*_{ji})$, og for at vise (I4) lader vi $A \in M_{m,n}(\mathbb{F})$ og betragter $C = A^* \in M_{n,m}(\mathbb{F})$. Udregningen

$$c'_{ij} = c^*_{ji} = (a'_{ji})^* = (a^*_{ij})^* = a_{ij}$$

viser da, at $C^* = A$, hvilket beviser (I4) og dermed sætningen. $\square$

Definition 2.6.8 Lad $\mathbb{F}$ være et skævlegeme, lad $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$ være en skævinvolution, og lad n være et naturligt tal. En *hermitisk form* afbildning

$$\langle -, - \rangle: \mathbb{F}^n \times \mathbb{F}^n \rightarrow \mathbb{F},$$

for hvilken det gælder følgende: For alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}^n$ og $a \in \mathbb{F}$ er

$$(H1) \quad \langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle.$$

$$(H2) \quad \langle \mathbf{x}, \mathbf{y} \cdot a \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \cdot a.$$

$$(H3) \quad \langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle.$$

$$(H4) \quad \langle \mathbf{x} \cdot a, \mathbf{y} \rangle = a^* \cdot \langle \mathbf{x}, \mathbf{y} \rangle.$$

$$(H5) \quad \langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle^*.$$

2 Matricer og lineære afbildninger

Hermitiske former er opkaldt efter Charles Hermite, der også er kendt for at have bevist, at grundtallet e for den naturlige logaritme er transcendentalt.

Eksempel 2.6.9 Standard-indreproduktet på $\mathbb{C}^n$, som er givet ved

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y} = (\bar{x}_1 \quad \dots \quad \bar{x}_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \bar{x}_1 y_1 + \dots + \bar{x}_n y_n,$$

er en hermitiske form på $\mathbb{C}^n$, hvor vi igen anvender kompleks konjugering som skævinvolution på $\mathbb{C}$. Ligeledes er standard-indreproduktet på $\mathbb{R}^n$, som er givet ved

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y} = \mathbf{x}^t \mathbf{y} = (x_1 \quad \dots \quad x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \dots + x_n y_n,$$

en hermitiske form på $\mathbb{R}^n$.

Vi skal nu vise, at der er en kanonisk 1-1 korrespondance mellem hermitiske former på $\mathbb{F}^n$ og hermitiske matricer $A \in M_n(\mathbb{F})$. Vi bemærker hertil, at hvis $\mathbf{x} \in \mathbb{F}^n = M_{n,1}(\mathbb{F})$ er en søjlevektor af dimension n , så er $\mathbf{x}^* \in M_{1,n}(\mathbb{F})$ en rækkevektor af dimension n .

Sætning 2.6.10 Lad $\mathbb{F}$ være et skævlegeme, lad $(-)^* : \mathbb{F} \rightarrow \mathbb{F}$ være en skævinvolution, og lad n være et naturligt tal.

- (1) Hvis $A \in M_n(\mathbb{F})$ er en hermitisk matrix, så er afbildningen $\langle -, - \rangle : \mathbb{F}^n \times \mathbb{F}^n \rightarrow \mathbb{F}$ givet ved $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* A \mathbf{y}$ en hermitisk form.
- (2) Hvis $\langle -, - \rangle : \mathbb{F}^n \times \mathbb{F}^n \rightarrow \mathbb{F}$ er en hermitisk form, så er matricen $A = (a_{ij}) \in M_n(\mathbb{F})$ med indgange $a_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle$ en hermitisk matrix, og $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* A \mathbf{y}$.

Bevis For at bevise påstanden (1), lader vi $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}^n$ og $a \in \mathbb{F}$ og udregner

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle &= \mathbf{x}^* A(\mathbf{y} + \mathbf{z}) = \mathbf{x}^* A \mathbf{y} + \mathbf{x}^* A \mathbf{z} = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle \\ \langle \mathbf{x}, \mathbf{y} \cdot a \rangle &= \mathbf{x}^* A \mathbf{y} a = \langle \mathbf{x}, \mathbf{y} \rangle \cdot a \\ \langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle &= (\mathbf{x} + \mathbf{y})^* A \mathbf{z} = \mathbf{x}^* A \mathbf{z} + \mathbf{y}^* A \mathbf{z} = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle \\ \langle \mathbf{x} \cdot a, \mathbf{y} \rangle &= (\mathbf{x} \cdot a)^* A \mathbf{y} = a^* \cdot \mathbf{x}^* A \mathbf{y} = a^* \cdot \langle \mathbf{x}, \mathbf{y} \rangle \\ \langle \mathbf{y}, \mathbf{x} \rangle &= \mathbf{y}^* A \mathbf{x} = \mathbf{y}^* A^* (\mathbf{x}^*)^* = (\mathbf{x}^* A \mathbf{y})^* = \langle \mathbf{x}, \mathbf{y} \rangle^*, \end{aligned}$$

hvilket viser påstanden. Her har vi anvendt sætning 2.6.7.

2.6 Hermitiske former

Den første del af (2) følger fra (H5), idet $a_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \langle \mathbf{e}_j, \mathbf{e}_i \rangle^* = a_{ji}^*$, og udregningen

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle &= \left\langle \sum_{i=1}^n \mathbf{e}_i x_i, \sum_{j=1}^n \mathbf{e}_j y_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i^* \langle \mathbf{e}_i, \mathbf{e}_j \rangle y_j \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i^* a_{ij} y_j = \sum_{i=1}^n x_i^* A y = \mathbf{x}^* A \mathbf{y} \end{aligned}$$

viser den anden del af (2). □

Vi kalder matricen $A = (\langle \mathbf{e}_i, \mathbf{e}_j \rangle)$ i sætning 2.6.10 (1) for *matricen, der repræsenterer den hermitiske form $\langle -, - \rangle$ med hensyn til standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ for $\mathbb{F}^n$* , og omvendt kalder vi den hermitiske form $\langle -, - \rangle$ i sætning 2.6.10 (2) for *den hermitiske form, der hører til matricen A med hensyn til standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ for $\mathbb{F}^n$* .

Definition 2.6.11 Hvis $\mathbb{F}$ er et skævlegeme, og $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$ er en skævinvolution, så siges en hermitisk form $\langle -, - \rangle$ på $\mathbb{F}^n$ at være *ikke-singulær*, hvis den matrix $A \in M_n(\mathbb{F})$, der repræsenterer $\langle -, - \rangle$ med hensyn til standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ er invertibel.

Eksempel 2.6.12 (1) Standard-indreprodukt på $\mathbb{F}^n$ er den hermitiske form

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* I \mathbf{y} = \mathbf{x}^* \mathbf{y} = x_1^* y_1 + \dots + x_n^* y_n,$$

der hører til identitetsmatricen med hensyn til standardbasen; sammenlign med eksempel 2.6.9. Den er ikke-singulær.

(2) Minkowskis hermitiske form på $\mathbb{F}^n$ er den hermitiske form

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \text{diag}(-1, 1, \dots, 1) \mathbf{y} = -x_1^* y_1 + x_2^* y_2 + \dots + x_n^* y_n,$$

der hører til diagonalmatricen $\text{diag}(-1, 1, \dots, 1)$ med hensyn til standardbasen. Den er ikke-singulær, selvom der findes ikke-nul vektorer $\mathbf{x}$, sådan at $\langle \mathbf{x}, \mathbf{x} \rangle = 0$.

(3) Den hyperbolske form på $\mathbb{F}^2$ er den hermitiske form

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{y} = x_1^* y_2 + x_2^* y_1,$$

der hører til den indikerede hermitiske matrix $H \in M_2(\mathbb{F})$. Den er ikke-singulær, selvom der igen findes ikke-nul vektorer $\mathbf{x}$, sådan at $\langle \mathbf{x}, \mathbf{x} \rangle = 0$.

Givet hermitiske former på $\mathbb{F}^m$ og $\mathbb{F}^n$, da ønsker vi at betragte lineære afbildninger $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$, der er kompatible med de givne hermitiske former. Sådanne afbildninger

2 Matricer og lineære afbildninger

kaldes *lineære isometrier*, idet de bevarer de geometriske strukturer, som de hermitiske former beskriver. Den formelle definition er som følger.

Definition 2.6.13 Lad $\mathbb{F}$ være et skævlegeme, lad $(-)^* : \mathbb{F} \rightarrow \mathbb{F}$ være en skævinvolution, og lad $\langle -, - \rangle$ og $\langle -, - \rangle'$ være hermitiske former på henholdsvis $\mathbb{F}^n$ og $\mathbb{F}^m$. En lineær afbildning $f : \mathbb{F}^n \rightarrow \mathbb{F}^m$ er en *lineær isometri* med hensyn til $\langle -, - \rangle$ og $\langle -, - \rangle'$, hvis

$$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle' = \langle \mathbf{x}, \mathbf{y} \rangle$$

for alle $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$.

Eksempel 2.6.14 (1) En bijektiv lineær isometri $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ med hensyn til det standard indre produkt $\langle -, - \rangle$ kaldes for en euklidisk transformation. For eksempel er rotationer og spejlinger i $\mathbb{R}^2$ isometriske isomorfier.

(2) En bijektiv lineær isometri $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ med hensyn til Minkowskiformen $\langle -, - \rangle$ kaldes for en Lorentz-transformation. De skyldes Hendrik Lorentz, som viste, at Maxwells ligninger er invariante under disse transformationer. Hvis $-1 < v < 1$, da er afbildningen $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ repræsenteret ved matricen

$$P = \begin{pmatrix} 1/\sqrt{1-v^2} & -v/\sqrt{1-v^2} \\ -v/\sqrt{1-v^2} & 1/\sqrt{1-v^2} \end{pmatrix}$$

en Lorentz-transformation. Hvis $\mathbf{x}$ er rumtidskoordinaterne med hensyn til et givet inertialsystem, da er $f(\mathbf{x})$ rumtidskoordinaterne med hensyn til det inertialsystem, der bevæger sig i forhold til det første med konstant hastighed v langs x_2 -aksen. Vi anvender her de naturlige enheder, hvor lysets hastighed er $c = 1$.

Sætning 2.6.15 Lad $\mathbb{F}$ være et skævlegeme og lad $(-)^* : \mathbb{F} \rightarrow \mathbb{F}$ være en skævinvolution. Lad $\langle -, - \rangle$ og $\langle -, - \rangle'$ være hermitiske former på $\mathbb{F}^n$ og $\mathbb{F}^m$, og lad $A \in M_n(\mathbb{F})$ og $A' \in M_m(\mathbb{F})$ være de matricer, der repræsenterer $\langle -, - \rangle$ og $\langle -, - \rangle'$ med hensyn til standardbaserne. En lineær afbildning $f : \mathbb{F}^n \rightarrow \mathbb{F}^m$ er en lineær isometri med hensyn til $\langle -, - \rangle$ og $\langle -, - \rangle'$, hvis og kun hvis den matrix $P \in M_{m,n}(\mathbb{F})$, der repræsenterer denne afbildning med hensyn til de respektive standardbaser, tilfredsstiller, at

$$P^* A' P = A.$$

Bevis Hvis $P^* A' P = A$, da viser udregningen

$$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle' = (P\mathbf{x})^* A' P\mathbf{y} = \mathbf{x}^* P^* A' P\mathbf{y} = \mathbf{x}^* A\mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle,$$

hvor $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, at $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er en lineær isometri. Omvendt, hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er en lineær isometri, da viser udregningen

$$\mathbf{e}_i^* P^* A' P \mathbf{e}_j = (P \mathbf{e}_i)^* A' P \mathbf{e}_j = \langle f(\mathbf{e}_i), f(\mathbf{e}_j) \rangle' = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \mathbf{e}_i^* A \mathbf{e}_j,$$

at de (i, j) 'te indgange i matricerne $P^* A' P$ og A er identiske for alle $1 \leq i, j \leq n$, og derfor gælder det, at $P^* A' P = A$ som ønsket. $\square$

Eksempel 2.6.16 Ifølge sætning 2.6.15 viser udregningen

$$\begin{aligned} P^* A P &= \begin{pmatrix} 1/\sqrt{1-v^2} & -v/\sqrt{1-v^2} \\ -v/\sqrt{1-v^2} & 1/\sqrt{1-v^2} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{1-v^2} & -v/\sqrt{1-v^2} \\ -v/\sqrt{1-v^2} & 1/\sqrt{1-v^2} \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{1-v^2} & -v/\sqrt{1-v^2} \\ -v/\sqrt{1-v^2} & 1/\sqrt{1-v^2} \end{pmatrix} \begin{pmatrix} -1/\sqrt{1-v^2} & v/\sqrt{1-v^2} \\ -v/\sqrt{1-v^2} & 1/\sqrt{1-v^2} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = A, \end{aligned}$$

at afbildningen $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ i eksempel 2.6.14 (2) er en Lorentz-transformation. Ligeledes viser udregningen

$$Q^* A' Q = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = A,$$

at afbildningen $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ givet ved

$$g(\mathbf{x}) = Q\mathbf{x} = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \mathbf{x}$$

er en bijektiv lineær isometri fra $\mathbb{R}^2$ med Minkowskiformen til $\mathbb{R}^2$ med den hyperbolske form.

2.7 Opgaver

2.1 Udregn følgende matrixprodukter

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}.$$

2 Matricer og lineære afbildninger

2.2 Afgør, ved at overveje dimensioner, hvilke af matricerne

$$A_1 = \begin{pmatrix} 2 & 5 \\ 0 & 1 \\ 4 & -2 \end{pmatrix} \quad A_2 = \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 2+i \end{pmatrix} \quad A_3 = \begin{pmatrix} 4 & 4 & 5 \end{pmatrix}$$

$$A_4 = \begin{pmatrix} 2 \\ 8 \\ 0 \\ -4 \end{pmatrix} \quad A_5 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad A_6 = \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix},$$

det giver mening at multiplicere med hinanden, og bestem alle matrixprodukter, der er meningsfulde. (Der skal være 11 i alt.)

2.3 Denne øvelse viser, at multiplikation af skalarer har nogle egenskaber, som ikke deles af matrixmultiplikation. Kommenter i hvert tilfælde, hvilken egenskab ved multiplikation af skalarer, der er tale om. (Sammenlign med definition 0.2.1.)

a) Udregn AB og BA , hvor

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}.$$

b) Udregn AB , hvor

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}.$$

c) Udregn A^2 , hvor

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}.$$

d) Udregn A^2 , hvor

$$A = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

e) Udregn A^2 , hvor

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

2.4 Afgør om følgende er rigtigt eller forkert:

- a) Hvis matrixprodukterne AB og BA begge eksisterer, da er begge matricerne A og B nødvendigvis kvadratiske.
- b) $A(BC) = (AB)C$.
- c) Hvis A og B er kvadratiske matricer af orden n , da er $AB = BA$.

d) Hvis $AB = O$, da er $A = O$ eller $B = O$.

2.5 Udregn følgende to matrixprodukter, idet der redegøres for den anvendte strategi.
[Vink: Matrixproduktet er associativt.]

$$\text{a)} \quad \begin{pmatrix} 17 & 231 & 100 \\ 91 & 640 & 77 \\ -11 & 1003 & 1 \end{pmatrix} \begin{pmatrix} 23 & 546 & 0 \\ -19 & -34 & 1 \\ 22 & 1001 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$\text{b)} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 670 & 546 & 45 \\ 1 & 0 & 0 \\ 22 & 1001 & 99 \end{pmatrix} \begin{pmatrix} 17 & 231 & 100 \\ 91 & 640 & 77 \\ -11 & 1003 & 1 \end{pmatrix}.$$

2.6 Betragt matricerne

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{og} \quad C = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

Udregn følgende matrixprodukter:

a) A^2, A^3, A^2A^3, A^5 .

b) B^{10} .

c) C^3 .

2.7 Udregn matrixprodukterne AB, BA, AC, CA , hvor

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad B = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix} \quad \text{og} \quad C = \begin{pmatrix} p & q \\ r & s \end{pmatrix}.$$

2.8 Udregn matrixprodukterne AB, BA, CD, DC, AC og CA , hvor

$$A = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & d & f \\ 0 & 1 & e \\ 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ r & 1 & 0 \\ t & s & 1 \end{pmatrix} \quad \text{og} \quad D = \begin{pmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ w & v & 1 \end{pmatrix}.$$

2.9 (★) Hvis $\mathbb{F}$ er et legeme, da kaldes afbildningen $\text{tr} : M_n(\mathbb{F}) \rightarrow \mathbb{F}$ defineret ved

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}$$

for *sporet*. Sporet af en matrix er således summen af matrixens diagonalindgange.

a) Vis, at det for alle $A, B \in M_n(\mathbb{F})$ gælder, at

$$\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B).$$

2 Matricer og lineære afbildninger

- b) Vis, at det for $A, B \in M_n(\mathbb{F})$ generelt *ikke* gælder, at

$$\operatorname{tr}(AB) = \operatorname{tr}(A)\operatorname{tr}(B).$$

[Vink: Find et modeksempel med 2×2 -matricer.]

- c) Vis imidlertid, at det for alle $A \in M_{m,n}(\mathbb{F})$ og $B \in M_{n,m}(\mathbb{F})$ gælder, at

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

- d) Vis endvidere, at hvis $A \in M_{m,n}(\mathbb{F})$, $B \in M_{n,p}(\mathbb{F})$ og $C \in M_{p,m}(\mathbb{F})$, da er

$$\operatorname{tr}(ABC) = \operatorname{tr}(BCA).$$

Vi kalder denne egenskab ved sporet for *invarians under cyklisk permutation*.

- e) Vis, at hvis $A \in M_n(\mathbb{F})$, og hvis $P \in M_n(\mathbb{F})$ er invertibel, da er

$$\operatorname{tr}(P^{-1}AP) = \operatorname{tr}(A).$$

[Vink: Anvend invarians af sporet under cyklisk permutation.]

- f) Vis, at hvis $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$ er skævinvolution, da er

$$\operatorname{tr}(A^*) = \operatorname{tr}(A)^*$$

for alle $A \in M_n(\mathbb{F})$. Her er A^* den adjungerede matrix.

2.10 Om en lineær afbildning $f: \mathbb{R}^2 \rightarrow \mathbb{R}^4$ vides, at

$$f \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \\ 14 \\ 1 \end{pmatrix} \quad \text{og} \quad f \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} -4 \\ -3 \\ 10 \\ 5 \end{pmatrix}.$$

Bestem matricen $A \in M_{4,2}(\mathbb{R})$, der repræsenterer $f: \mathbb{R}^2 \rightarrow \mathbb{R}^4$ med hensyn til standardbaserne.

2.11 Lad $f, g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ være afbildningerne defineret ved

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + x_3 \end{pmatrix} \quad \text{og} \quad g \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 + 2x_3 \\ x_2 + x_3 \\ x_1 + x_3 \end{pmatrix}.$$

Vi har tidligere i opgave 0.5 vist, at $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ er bijektiv, samt at $g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ikke er bijektiv.

- a) Vis, at afbildningerne $f, g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ begge er lineære.

- b) Bestem matricerne, der repræsenterer henholdsvis $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $f^{-1}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ og $g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ med hensyn til standardbaserne.

2.12 Afgør, hvilke af følgende afbildninger $f, g, h: \mathbb{C}^2 \rightarrow \mathbb{C}^2$, der er lineære.

$$f\begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} w \\ \bar{z} \end{pmatrix}, \quad g\begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} iz \\ z-w \end{pmatrix}, \quad h\begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} z \\ \operatorname{Im}(w) \end{pmatrix}.$$

For de af afbildningerne, der er lineære, bestem den matrix, der repræsenterer afbildningen med hensyn til standardbaserne.

2.13 En fabrik producerer to varer X_1 og X_2 ud fra de tre råvarer Y_1 , Y_2 og Y_3 . Vi betegner tilhørende produktionssæt og forbrugssæt med henholdsvis (x_1, x_2) og (y_1, y_2, y_3) . Om den pågældende produktion gælder det, at

$$\text{produktion af en enhed af } X_1 \text{ kræver } \begin{cases} 2 \text{ enheder af } Y_1 \\ 1 \text{ enhed af } Y_2 \\ 5 \text{ enheder af } Y_3, \end{cases}$$

og

$$\text{produktion af en enhed af } X_2 \text{ kræver } \begin{cases} 1 \text{ enhed af } Y_1 \\ 6 \text{ enheder af } Y_2 \\ 1 \text{ enhed af } Y_3. \end{cases}$$

- a) Opskriv den afbildning $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, der fortæller, hvad sammenhængen er mellem forbrugssæt og produktionssæt.
- b) Vis, at afbildningen $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ er lineær, og angiv matricen, der repræsenterer denne med hensyn til standardbaserne.

I den virkelige verden, er der begrænsninger på råvarerne, som typisk er repræsenteret af lineære uligheder. For eksempel kan antallet af enheder af råvarer ikke være negativt. Lineær programmering er det fagområde inden for matematikken, der anvendes til at optimere produktionen i sådan en situation. I praksis finder Danzigs simpleksalgoritme hurtigt den optimale løsning, selvom det er muligt at konstruere problemer, hvor algoritmen ikke standser.

2.14 a) Vis, at matricen

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R}),$$

er en løsning til ligningen " $x^2 + 1 = 0$ ", det vil sige, at $J^2 + I = O$.

- b) Findes der andre matricer $A \in M_2(\mathbb{R})$, der er løsninger til denne ligning?
- c) (★) Find samtlige løsninger $A \in M_2(\mathbb{R})$ og $B \in M_2(\mathbb{C})$ til ligningen " $x^2 + 1 = 0$ ".

2 Matricer og lineære afbildninger

- 2.15 a) Vis, at hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ og $g: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er lineære afbildninger, da er også afbildningen $f + g: \mathbb{F}^n \rightarrow \mathbb{F}^m$ defineret ved $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ lineær.
- b) Vis endvidere, at hvis afbildningerne $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ og $g: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er repræsenteret af henholdsvis $A \in M_{m,n}(\mathbb{F})$ og $B \in M_{m,n}(\mathbb{F})$ med hensyn til standardbaserne, da er $f + g: \mathbb{F}^n \rightarrow \mathbb{F}^m$ repræsenteret af $A + B \in M_{m,n}(\mathbb{F})$ med hensyn til standardbaserne.
- c) Giv et eksempel, der viser at hvis $f, g: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er to afbildninger, sådan at $f + g: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er lineær, da er afbildningerne $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ og $g: \mathbb{F}^n \rightarrow \mathbb{F}^m$ ikke nødvendigvis lineære.

2.16 Vi betragter de følgende matricer:

$$\begin{aligned} A_1 &= \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} & A_2 &= \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix} & A_3 &= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \\ A_4 &= \begin{pmatrix} 0 & 1/4 & 1/8 \\ 0 & -1/4 & 3/8 \\ 1 & -1/4 & -1/8 \end{pmatrix} & A_5 &= \begin{pmatrix} 1 & 0 & 1 \\ 3 & -1 & 0 \\ 2 & 2 & 0 \end{pmatrix} & A_6 &= \begin{pmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ 2 & 2 & 0 \end{pmatrix} \\ A_7 &= \begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix} & A_8 &= \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

For hvilke værdier af i og j gælder det, at $A_i = A_j^{-1}$?

[Vink: Overvej dimensionen af matricerne, og undersøg om $A_i A_j = I = A_j A_i$.]

2.17 Lad $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ være afbildningen givet ved $f(\mathbf{x}) = A\mathbf{x}$, hvor

$$A = \begin{pmatrix} 0 & 0 & 1/4 \\ 0 & -2 & 0 \\ 3 & 0 & 0 \end{pmatrix}$$

Gør rede for, at $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ er bijektiv, og find den inverse afbildning $g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Bestem herved den inverse matrix A^{-1} .

2.18 (★) Lad $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ være afbildningen givet ved $f(\mathbf{x}) = A\mathbf{x}$, hvor

$$A = \begin{pmatrix} 0 & 0 & a \\ 0 & b & 0 \\ c & 0 & 0 \end{pmatrix}.$$

Undersøg, for hvilke værdier af $a, b, c \in \mathbb{F}$ afbildningen $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ er bijektiv, og bestem i givet fald den inverse afbildning $g: \mathbb{F}^3 \rightarrow \mathbb{F}^3$. Bestem i disse tilfælde den inverse matrix A^{-1} .

2.19 Vis, gerne ved brug af Maple, at det for matricen

$$C = \begin{pmatrix} -1 & -19 & 10 \\ 1 & 6 & -2 \\ 1 & 11 & -5 \end{pmatrix}$$

gælder, at $C^3 = I$. Vis, at dette medfører, at C er invertibel med $C^{-1} = C^2$.

2.20 Giv et eksempel på en lineær afbildning $f: \mathbb{R}^3 \rightarrow \mathbb{R}^4$, som er injektiv, og hvis billede $\text{im}(f) \subset \mathbb{R}^4$ indeholder vektorerne

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{og} \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}.$$

Er en sådan afbildning entydigt bestemt?

2.21 Vi betragter afbildningerne $g: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ og $h: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ givet ved

$$g \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u+v \\ u-v \\ u+2v \end{pmatrix} \quad \text{og} \quad h \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ y+z \end{pmatrix}.$$

- Gør rede for, at afbildningerne $g: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ og $h: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ er lineære, og bestem matricerne $B \in M_{3,2}(\mathbb{R})$ og $C \in M_{2,3}(\mathbb{R})$, der repræsenterer disse med hensyn til standardbaserne.
- Gør rede for, at $g: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ er injektiv, men ikke surjektiv.
[Vink: Overvej rangen af matricen B .]
- Gør rede for, at $h: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ er surjektiv, men ikke injektiv.
[Vink: Overvej rangen af matricen C .]
- Bestem matricerne, der repræsenterer $g \circ h: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ og $h \circ g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ med hensyn til standardbaserne.

2.22 a) Bestem rangen af matricerne

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

- Afgør, om afbildningerne $f, g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ givet ved $f(\mathbf{x}) = A\mathbf{x}$ og $g(\mathbf{x}) = B\mathbf{x}$ er injektive, surjektive og/eller bijektive.

2 Matricer og lineære afbildninger

2.23 Vis, at for alle $a, b, c \in \mathbb{F}$ er matricen

$$A = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$$

invertibel, og bestem den inverse matrix A^{-1} .

2.24 Bestem rangen af følgende matricer og forklar, hvad rangen siger om antallet af løsninger til ligningssystemet med totalmatrix $(A_i \mid \mathbf{b})$, hvor $\mathbf{b}$ er en arbitrær konstantsøjle.

$$\begin{aligned} A_1 &= \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} & A_2 &= \begin{pmatrix} -1 & 1 \\ 0 & 2 \\ -3 & 1 \end{pmatrix} & A_3 &= \begin{pmatrix} 2 & -2 & 4 \\ -1 & 1 & -2 \end{pmatrix} \\ A_4 &= \begin{pmatrix} -1 & 2 & i \\ 2i & -4i & 1 \end{pmatrix} & A_5 &= \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix} & A_6 &= \begin{pmatrix} 1 & 2 & -2 \\ 3 & -1 & 0 \\ 2 & -3 & 2 \end{pmatrix} \end{aligned}$$

2.25 Bestem rangen af den komplekse matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -i \\ i & 1-i & 1 \end{pmatrix},$$

og afgør derved, om den er invertibel eller ej.

2.26 Afgør, om matricerne

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 5 & 7 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix}$$

er invertible eller ej, og bestem i givet fald deres inverse matricer.

2.27 Vis, at den lineære afbildning $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ givet ved

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 + x_2 \\ x_1 + x_2 + x_3 \\ x_1 + x_2 + x_3 + x_4 \end{pmatrix}$$

er bijektiv, og find matricen, der repræsenterer den inverse afbildning $g: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ med hensyn til standardbaserne.

2.28 Vis, at den lineære afbildning $f: \mathbb{F}^4 \rightarrow \mathbb{F}^4$ givet ved

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 4x_1 + x_3 + x_4 \\ 3x_1 + x_2 + 3x_3 + x_4 \\ x_2 + 2x_3 \\ 3x_1 + 2x_2 + 4x_3 + x_4 \end{pmatrix}$$

er bijektiv, og find matricen, der repræsenterer den inverse afbildning $g: \mathbb{F}^4 \rightarrow \mathbb{F}^4$ med hensyn til standardbaserne.

2.29 a) Vis, at identitetsafbildning $\text{id}: \mathbb{R} \rightarrow \mathbb{R}$ er en skævinvolution på $\mathbb{R}$.

b) Vis, at kompleks konjugering $(-)^*: \mathbb{C} \rightarrow \mathbb{C}$, som er defineret ved

$$(a + ib)^* = a - ib,$$

er en skævinvolution på $\mathbb{C}$.

c) (★) Vis, at kvaternionisk konjugering $(-)^*: \mathbb{H} \rightarrow \mathbb{H}$, som er defineret ved

$$(a + ib + jc + kd)^* = a - ib - jc - kd,$$

er en skævinvolution på $\mathbb{H}$. Addition og multiplikation af kvaternioner er defineret i eksempel 0.3.1.

2.30 Afgør, hvilke af de følgende komplekse matricer, der er hermitiske:

$$A = \begin{pmatrix} 1 & 2i & 7 \\ 2i & 4 & 1+i \\ 7 & 1+i & i \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2i & 7 \\ -2i & 4 & 1-i \\ 7 & 1+i & 1 \end{pmatrix} \quad \text{og} \quad C = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}.$$

Her betragter vi som altid de komplekse tal $\mathbb{C}$ med kompleks konjugering som skævinvolution.

2.31 Udregn matrixprodukterne AA^t og A^tA , hvor A er den reelle matrix

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -2 & 3 & 1 \end{pmatrix}.$$

2.32 (★) Lad $\mathbb{F}$ være et legeme med skævinvolution $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$. Vis, at afbildningen

$$\langle -, - \rangle: \mathbb{F}^2 \times \mathbb{F}^2 \rightarrow \mathbb{F}$$

defineret ved $\langle \mathbf{x}, \mathbf{y} \rangle = (x_1 + x_2)^*(y_1 + y_2)$ er en hermitisk form. Angiv endvidere matricen $A \in M_2(\mathbb{F})$, der repræsenterer $\langle -, - \rangle$ med hensyn til standardbasen.

2.33 Lad $A_1, \dots, A_k \in M_n(\mathbb{F})$ være invertible matricer. Bevis, at

$$(A_1 A_2 \cdots A_k)^{-1} = (A_k)^{-1} \cdots (A_1)^{-1}.$$

3 Determinant

Givet n vektorer i $\mathbb{R}^n$, ønsker vi at definere og udregne det n -dimensionale volume af det parallellepipedum, de n vektorer udspænder. Dette problem fører naturligt til definitionen af determinanten af en $n \times n$ -matrix med indgange i $\mathbb{R}$, og denne definition viser sig da at være meningsfuld for en $n \times n$ -matrix med indgange i et vilkårligt legeme $\mathbb{F}$. Definitionen nødvendiggør, at multiplikation af skalarer opfylder den kommutative lov, så vi skal derfor i dette kapitel for første gang bruge, at det for skalarer $a, b \in \mathbb{F}$ gælder, at $ab = ba$. En vigtig teoretisk anvendelse af determinanten er, at den afgør, hvorvidt en kvadratisk matrix er invertibel eller ej, og deraf navnet *determinant*, som skyldes Gauss. Determinanten udregnes lettest ved at benytte dens definerende egenskaber. Denne udregningsmetode kan minde om Gauss-elimination med den forskel, at nogle operationer ændrer determinanten, omend på en kontrolleret måde. Man skal dog passe på med ikke at glemme denne forskel. Endelig findes der en lukket formel for determinanten af en $n \times n$ -matrix, som skyldes Leibniz, men denne formel indeholder $n!$ summander og er derfor mest af teoretisk interesse.

3.1 Determinant af 2×2 -matrix

Lad os først betragte det 1-dimensionale tilfælde. Det 1-dimensionale parallellepipedum udspændt af en vektor $\mathbf{a} = (a) \in \mathbb{R}^1$ defineres som intervallet

$$P = \{\mathbf{a}c \in \mathbb{R}^1 \mid 0 \leq c \leq 1\} \subset \mathbb{R}^1.$$

Vi definerer det 1-dimensionale volume eller længden af P til at være absolutværdien

$$\text{vol}(P) = |a|.$$

Tilsvarende er det 2-dimensionale parallellepipedum udspændt af de to vektorer

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

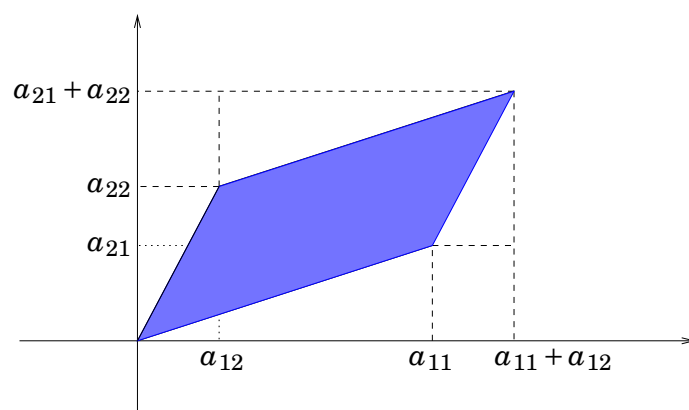
defineret til at være parallelogrammet

$$P = \{\mathbf{a}_1 c_1 + \mathbf{a}_2 c_2 \mid 0 \leq c_1, c_2 \leq 1\} \subset \mathbb{R}^2;$$

se Figur 3.1 for en illustration af denne delmængde. Som vores notation antyder, vil vi tænke på de to vektorer $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^2$ som søjlevektorerne i matricen

$$A = (\mathbf{a}_1 \quad \mathbf{a}_2) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in M_2(\mathbb{R}).$$

3 Determinant

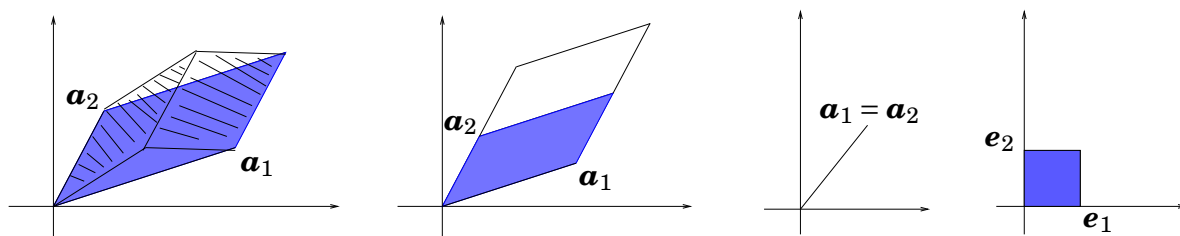


Figur 3.1: Parallelogrammet udspændt af to vektorer

Vi forestiller os, at det 2-dimensionale volumen eller areal af parallelogrammet P tilsvarende skal være givet som absolutværdien

$$\text{vol}(P) = |\det(A)|$$

af en skalar $\det(A)$, der afhænger af 2×2 -matricen A . Vi diskuterer nu, hvordan vi vil forvente, at afbildningen $\det: M_2(\mathbb{R}^2) \rightarrow \mathbb{R}$ skal opføre sig, hvilket figuren illustrerer.



Figur 3.2: Determinanten opfylder (1) $\det(\mathbf{b}_1 + \mathbf{c}_1 \ \mathbf{a}_2) = \det(\mathbf{b}_1 \ \mathbf{a}_2) + \det(\mathbf{c}_1 \ \mathbf{a}_2)$, (2) $\det(\mathbf{a}_1 \ \mathbf{a}_2 c) = \det(\mathbf{a}_1 \ \mathbf{a}_2) c$, (3) $\det(\mathbf{a} \ \mathbf{a}) = 0$ (4) $\det(\mathbf{e}_1 \ \mathbf{e}_2) = 1$

Hvis vektoren $\mathbf{a}_1 = \mathbf{b}_1 + \mathbf{c}_1$ er en sum af to vektorer, da forventer vi rimeligvis, at arealet af parallelogrammet udspændt af $\mathbf{a}_1$ og $\mathbf{a}_2$ er lig med summen af arealerne af parallelogrammerne udspændt af henholdsvis $\mathbf{b}_1$ og $\mathbf{a}_2$ og af $\mathbf{c}_1$ og $\mathbf{a}_2$. Tilsvarende, hvis $\mathbf{a}_2 = \mathbf{b}_2 + \mathbf{c}_2$ er en sum af to vektorer, da vil vi forvente, at arealet af parallelogrammet udspændt af $\mathbf{a}_1$ og $\mathbf{a}_2$ er lig med summen af arealerne af parallelogrammerne udspændt af henholdsvis $\mathbf{a}_1$ og $\mathbf{b}_2$ og af $\mathbf{a}_1$ og $\mathbf{c}_2$. Hvis vi erstatter $\mathbf{a}_1$ med $\mathbf{a}_1 \cdot c$, men bibeholder $\mathbf{a}_2$, da vil vi forvente, at arealet af de tilsvarende parallelogrammer ganges med $|c|$; tilsvarende, hvis vi erstatter $\mathbf{a}_2$ med $\mathbf{a}_2 \cdot c$, men bibeholder $\mathbf{a}_1$, da vil vi igen forvente, at arealet af de tilsvarende parallelogrammer ganges med $|c|$. Hvis $\mathbf{a}_1 = \mathbf{a}_2$, da forventer vi, at arealet af det parallelogram, de udspænder, er lig med 0. Endelig, vælger vi

3.1 Determinant af 2×2 -matrix

at normere vores arealmål, sådan at enhedskvadratet har areal 1. Vi ønsker altså, at afbildningen $\det: M_2(\mathbb{R}) \rightarrow \mathbb{R}$ har de følgende egenskaber:

(D1) Hvis $\mathbf{a}_1 = \mathbf{b}_1 + \mathbf{c}_1$ eller $\mathbf{a}_2 = \mathbf{b}_2 + \mathbf{c}_2$, da gælder det, at

$$\det(\mathbf{b}_1 \ \mathbf{a}_2) + \det(\mathbf{c}_1 \ \mathbf{a}_2) = \det(\mathbf{a}_1 \ \mathbf{a}_2) = \det(\mathbf{a}_1 \ \mathbf{b}_2) + \det(\mathbf{a}_1 \ \mathbf{c}_2).$$

(D2) Hvis $\mathbf{a}_1$ erstattes af $\mathbf{a}_1 \cdot c$ eller $\mathbf{a}_2$ erstattes af $\mathbf{a}_2 \cdot c$, da gælder det, at

$$\det(\mathbf{a}_1 \cdot c \ \mathbf{a}_2) = \det(\mathbf{a}_1 \ \mathbf{a}_2) \cdot c = \det(\mathbf{a}_1 \ \mathbf{a}_2 \cdot c).$$

(D3) Hvis $\mathbf{a}_1 = \mathbf{a}_2$, da er $\det(\mathbf{a}_1 \ \mathbf{a}_1) = 0$.

(D4) Hvis $\mathbf{a} = \mathbf{e}_1$ og $\mathbf{a}_2 = \mathbf{e}_2$ er standardenhedsvektorerne, da er $\det(\mathbf{e}_1 \ \mathbf{e}_2) = 1$.

Det viser sig, at der netop findes én sådan afbildning:

Sætning 3.1.1 *Lad $\mathbb{F}$ være et legeme. Afbildningen $\det: M_2(\mathbb{F}) \rightarrow \mathbb{F}$ defineret ved*

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

opfylder (D1)–(D4) og er entydigt bestemt herved.

Bemærkning 3.1.2 (1) I mere læsevenlig form siger formelen i sætning 3.1.1, at

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

(2) Hvis man begynder med arealet $(a_{11} + a_{12})(a_{21} + a_{22})$ af det store rektangel i figur 3.1 og derfra trækker arealet af de rektangler, der tilsammen udgør området udenfor parallelogrammet, så får man samme formel $a_{11}a_{22} - a_{12}a_{21}$.

Før vi beviser sætning 3.1.1, viser vi, at egenskaberne (D1)–(D3) medfører de følgende yderligere egenskaber:

(D5) Hvis vektorerne $\mathbf{a}_1$ og $\mathbf{a}_2$ ombyttes, da gælder det, at

$$\det(\mathbf{a}_1 \ \mathbf{a}_2) = -\det(\mathbf{a}_2 \ \mathbf{a}_1).$$

(D6) Hvis et multiplum af en af vektorerne $\mathbf{a}_1$ og $\mathbf{a}_2$ adderes til den anden, da gælder det, at

$$\det(\mathbf{a}_1 + \mathbf{a}_2 \cdot c \ \mathbf{a}_2) = \det(\mathbf{a}_1 \ \mathbf{a}_2) = \det(\mathbf{a}_1 \ \mathbf{a}_1 \cdot c + \mathbf{a}_2).$$

3 Determinant

Her følger (D5) fra udregningen

$$\begin{aligned}\det(\mathbf{a}_1 \ \mathbf{a}_2) + \det(\mathbf{a}_2 \ \mathbf{a}_1) &= \det(\mathbf{a}_1 \ \mathbf{a}_2) + \det(\mathbf{a}_1 \ \mathbf{a}_1) + \det(\mathbf{a}_2 \ \mathbf{a}_1) + \det(\mathbf{a}_2 \ \mathbf{a}_2) \\ &= \det(\mathbf{a}_1 \ \mathbf{a}_1 + \mathbf{a}_2) + \det(\mathbf{a}_2 \ \mathbf{a}_1 + \mathbf{a}_2) \\ &= \det(\mathbf{a}_1 + \mathbf{a}_2 \ \mathbf{a}_1 + \mathbf{a}_2) = 0,\end{aligned}$$

hvor de fire identiteter følger fra henholdsvis (D3), (D1), (D1) og (D3). På lignende vis følger den første identitet i (D6) fra udregningen

$$\begin{aligned}\det(\mathbf{a}_1 + \mathbf{a}_2 \cdot c \ \mathbf{a}_2) &= \det(\mathbf{a}_1 \ \mathbf{a}_2) + \det(\mathbf{a}_2 \cdot c \ \mathbf{a}_2) \\ &= \det(\mathbf{a}_1 \ \mathbf{a}_2) + \det(\mathbf{a}_2 \ \mathbf{a}_2) \cdot c \\ &= \det(\mathbf{a}_1 \ \mathbf{a}_2),\end{aligned}$$

hvor de tre identiteter følger fra (D1), (D2), og (D3). Den anden identitet i (D6) vises helt tilsvarende.

Bevis (for Sætning 3.1.1) Man viser let, at den givne afbildning opfylder (D1)–(D4), så vi nøjes med at vise, at disse egenskaber bestemmer afbildningen entydigt. Vi skriver derfor først $\mathbf{a}_1 = \mathbf{e}_1 a_{11} + \mathbf{e}_2 a_{21}$ og får da fra (D1) og (D2), at

$$\det(\mathbf{a}_1 \ \mathbf{a}_2) = \det(\mathbf{e}_1 a_{11} \ \mathbf{a}_2) + \det(\mathbf{e}_2 a_{21} \ \mathbf{a}_2) = \det(\mathbf{e}_1 \ \mathbf{a}_2) a_{11} + \det(\mathbf{e}_2 \ \mathbf{a}_2) a_{21}.$$

Dernæst skriver vi $\mathbf{a}_2 = \mathbf{e}_1 a_{12} + \mathbf{e}_2 a_{22}$ og får fra henholdsvis (D6), (D2) og (D4), at

$$\det(\mathbf{e}_1 \ \mathbf{a}_2) = \det(\mathbf{e}_1 \ \mathbf{e}_2 a_{22}) = \det(\mathbf{e}_1 \ \mathbf{e}_2) a_{22} = a_{22},$$

mens henholdsvis (D6), (D2), (D5) og (D4) viser, at

$$\det(\mathbf{e}_2 \ \mathbf{a}_2) = \det(\mathbf{e}_2 \ \mathbf{e}_1 a_{12}) = \det(\mathbf{e}_2 \ \mathbf{e}_1) a_{12} = -\det(\mathbf{e}_1 \ \mathbf{e}_2) a_{12} = -a_{12}.$$

Tilsammen viser disse udregninger altså, at (D1)–(D4) medfører, at

$$\det(\mathbf{a}_1 \ \mathbf{a}_2) = a_{22} a_{11} - a_{12} a_{21},$$

som ønsket. □

Eksempel 3.1.3 Vi illustrerer, hvordan (D1)–(D4) og de afledte egenskaber (D5)–(D6) kan anvendes til at udregne determinanten.

$$\det \begin{pmatrix} 5 & 3 \\ 7 & 4 \end{pmatrix} \stackrel{(D6)}{=} \det \begin{pmatrix} -1 & 3 \\ -1 & 4 \end{pmatrix} \stackrel{(D6)}{=} \det \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \stackrel{(D6)}{=} \det \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \stackrel{(D2)}{=} \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot (-1) \stackrel{(D4)}{=} -1.$$

I de første tre ligheder bruger vi (D6) på følgende måde: i den første erstatter vi den første søjle med den første søjle plus den anden søjle gange -2 ; i den anden erstatter vi den anden søjle med den anden søjle plus den første søjle gange 3 ; og i den tredje erstatter vi den første søjle med den første søjle plus den anden søjle. Den fjerde lighed benytter (D2) på den første søjle, og den sidste lighed følger fra (D4). Naturligvis giver formelen " $a_{11}a_{22} - a_{12}a_{21}$ " fra sætning 3.1.1 samme resultat.

3.2 Determinant for $n \times n$ -matricer

Vi definerer nu determinanten af en $n \times n$ -matrix med indgange i et legeme $\mathbb{F}$ ved at generalisere den definition for 2×2 -matricer, som vi gav i afsnit 3.1. For $\mathbb{F} = \mathbb{R}$ definerer vi endvidere det n -dimensionale volumen af det n -dimensionale parallellepipedum

$$P = \{\mathbf{a}_1 c_1 + \cdots + \mathbf{a}_n c_n \mid 0 \leq c_1, \dots, c_n \leq 1\} \subset \mathbb{R}^n$$

udspændt af $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^n$ til at være absolutværdien

$$\text{vol}(P) = |\det(\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n)|$$

af determinanten af den reelle $n \times n$ -matrix, der har $\mathbf{a}_1, \dots, \mathbf{a}_n$ som søjler. Den følgende sætning er vores anden hovedsætning.

Sætning 3.2.1 *Lad $\mathbb{F}$ være et legeme og lad n være et naturligt tal. Der findes da en entydigt bestemt afbildning $\det: M_n(\mathbb{F}) \rightarrow \mathbb{F}$ med følgende egenskaber:*

(D1) *For alle $1 \leq k \leq n$ og $\mathbf{a}_1, \dots, \mathbf{a}_{k-1}, \mathbf{b}_k, \mathbf{c}_k, \mathbf{a}_{k+1}, \dots, \mathbf{a}_n \in \mathbb{F}^n$ gælder det, at*

$$\det(\mathbf{a}_1 \ \dots \ \mathbf{b}_k + \mathbf{c}_k \ \dots \ \mathbf{a}_n) = \det(\mathbf{a}_1 \ \dots \ \mathbf{b}_k \ \dots \ \mathbf{a}_n) + \det(\mathbf{a}_1 \ \dots \ \mathbf{c}_k \ \dots \ \mathbf{a}_n).$$

(D2) *For alle $1 \leq k \leq n$, $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{F}^n$ og $c \in \mathbb{F}$ gælder det, at*

$$\det(\mathbf{a}_1 \ \dots \ \mathbf{a}_k \cdot c \ \dots \ \mathbf{a}_n) = \det(\mathbf{a}_1 \ \dots \ \mathbf{a}_k \ \dots \ \mathbf{a}_n) \cdot c.$$

(D3) *For alle $1 \leq k < l \leq n$ og $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{F}^n$ med $\mathbf{a}_k = \mathbf{a}_l$, gælder det, at*

$$\det(\mathbf{a}_1 \ \dots \ \mathbf{a}_k \ \dots \ \mathbf{a}_l \ \dots \ \mathbf{a}_n) = 0.$$

(D4) *For standardenhedsvektorerne $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \in \mathbb{F}^n$ gælder det, at*

$$\det(\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_n) = 1.$$

Bemærkning 3.2.2 Vi kan betragte determinanten som en afbildning

$$\det: \mathbb{F}^n \times \cdots \times \mathbb{F}^n \rightarrow \mathbb{F},$$

der til en n -tuple $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ af søjlevektorer tilordner skalaren $\det(\mathbf{a}_1 \ \dots \ \mathbf{a}_n)$, og da udtrykker (D1)–(D2), at determinanten opfylder betingelserne (L1)–(L2) for en lineær afbildning i hver faktor $\mathbb{F}^n$. Vi siger, at en sådan afbildning er *multilineær*. Endvidere siges en multilineær afbildning, der også opfylder (D3), at være *alternerende*. Således

3 Determinant

er determinanten den entydige alternerende multilineære afbildning, der på standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ tager værdien 1. Denne sidste egenskab (D4) er naturligvis temmelig arbitrær, og beviset for sætning 3.2.1 viser da også mere generelt, at der for ethvert $a \in \mathbb{F}$ findes en entydig multilinear alternerende afbildning, der på standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ antager værdien $a \in \mathbb{F}$.

I beviset for sætning 3.2.1 anvender vi undervejs, at de definerede egenskaber ved determinanten har følgende konsekvenser.

Sætning 3.2.3 *Lad $\mathbb{F}$ være et legeme, og lad n være et naturligt tal. Hvis en afbildning $\det: M_n(\mathbb{F}) \rightarrow \mathbb{F}$ tilfredsstiller (D1)–(D3), da gælder endvidere følgende udsagn:*

(D5) *Hvis B er matricen, der fremkommer fra A ved ombytning af to søjler, da er*

$$\det(B) = -\det(A).$$

(D6) *Hvis B er matricen, der fremkommer fra A ved at addere et multiplum af en søjle i i A til en anden søjle j i A , da er*

$$\det(B) = \det(A).$$

Bevis Vi viser først (D5). Så lad A være en $n \times n$ -matrix og lad B være matricen, der fremkommer fra A ved ombytning af den k 'te og l 'te søjle, hvor $1 \leq k < l \leq n$.

$$\begin{aligned} \det(A) + \det(B) &= \det(\dots \mathbf{a}_k \dots \mathbf{a}_l \dots) + \det(\dots \mathbf{a}_l \dots \mathbf{a}_k \dots) \\ &\stackrel{(D3)}{=} \det(\dots \mathbf{a}_k \dots \mathbf{a}_l \dots) + \det(\dots \mathbf{a}_l \dots \mathbf{a}_k \dots) \\ &\quad + \det(\dots \mathbf{a}_k \dots \mathbf{a}_k \dots) + \det(\dots \mathbf{a}_l \dots \mathbf{a}_l \dots) \\ &\stackrel{(D1)}{=} \det(\dots \mathbf{a}_k \dots \mathbf{a}_k + \mathbf{a}_l \dots) + \det(\dots \mathbf{a}_l \dots \mathbf{a}_k + \mathbf{a}_l \dots) \\ &\stackrel{(D1)}{=} \det(\dots \mathbf{a}_k + \mathbf{a}_l \dots \mathbf{a}_k + \mathbf{a}_l \dots) \stackrel{(D3)}{=} 0, \end{aligned}$$

hvilket beviser (D5). Tilsvarende, lad B være matricen der fremkommer fra A ved at addere multiplummet $\mathbf{a}_l \cdot c$ af den l 'te søjle $\mathbf{a}_l$ til den k 'te søjle $\mathbf{a}_k$. Hvis $k < l$, da er

$$\begin{aligned} \det(B) &= \det(\dots \mathbf{a}_k + \mathbf{a}_l \cdot c \dots \mathbf{a}_l \dots) \\ &\stackrel{(D1)}{=} \det(\dots \mathbf{a}_k \dots \mathbf{a}_l \dots) + \det(\dots \mathbf{a}_l \cdot c \dots \mathbf{a}_l \dots) \\ &\stackrel{(D2)}{=} \det(\dots \mathbf{a}_k \dots \mathbf{a}_l \dots) + \det(\dots \mathbf{a}_l \dots \mathbf{a}_l \dots) \cdot c \\ &\stackrel{(D3)}{=} \det(\dots \mathbf{a}_k \dots \mathbf{a}_l \dots) = \det(A), \end{aligned}$$

hvilket beviser (D6) for $k < l$, og beviset for $k > l$ er ganske tilsvarende. □

3.2 Determinant for $n \times n$ -matricer

For at vise at determinantaftbildningen findes, giver vi i beviset for sætning 3.2.1 en rekursiv definition af determinanten, som bruger følgende definition.

Definition 3.2.4 Lad A være en $m \times n$ -matrix. Givet $1 \leq i \leq m$ og $1 \leq j \leq n$, er den (i, j) 'te *undermatrix* af A den $(m-1) \times (n-1)$ -matrix A_{ij} , der fremkommer fra A ved at fjerne den i 'te række og den j 'te søjle:

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

Hvis A er en kvadratisk matrix af orden n , da er den (i, j) 'te undermatrix A_{ij} specielt en kvadratisk matrix af orden $n-1$.

Eksempel 3.2.5 Vi betragter som eksempel 3×3 -matricen

$$A = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix}.$$

Den $(1, 1)$ 'te undermatrix A_{11} og den $(2, 3)$ 'te undermatrix A_{23} er da henholdsvis

$$A_{11} = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 6 & 4 \end{pmatrix} \quad \text{og} \quad A_{23} = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ 2 & 6 \end{pmatrix},$$

som begge er 2×2 -matricer.

Bevis (for sætning 3.2.1) Vi ønsker at vise eksistens og entydighed af en afbildning $\det: M_n(\mathbb{F}) \rightarrow \mathbb{F}$, der opfylder (D1)–(D4), og vi begynder med eksistensen, som vi viser ved induktion på $n \geq 0$. For $n = 0$ har mængden $M_n(\mathbb{F})$ præcis ét element, nemlig den tomme matrix $\begin{pmatrix} \end{pmatrix}$, og vi definerer da $\det: M_0(\mathbb{F}) \rightarrow \mathbb{F}$ til at være afbildningen givet ved

$$\det\left(\begin{pmatrix} \end{pmatrix}\right) = 1.$$

Afbildningen opfylder trivielt (D1)–(D3), da der ikke findes $1 \leq j \leq n$ eller $1 \leq j < k \leq n$, mens (D4) er opfyldt per definition.

3 Determinant

Vi antager derfor induktivt, at determinant afbildningen med de ønskede egenskaber er defineret for $n = p - 1$, og definerer afbildningen $\det: M_p(\mathbb{F}) \rightarrow \mathbb{R}$ ved formelen

$$\det(A) = \sum_{j=1}^p (-1)^{1+j} a_{1j} \det(A_{1j}) \quad (3.2.6)$$

hvor A_{1j} er den $(1, j)$ 'te undermatrix af A , som er kvadratisk af orden $p - 1$. Vi skal vise at denne afbildningen opfylder (D1)–(D4). For at vise (D1), betragter vi kvadratiske matricer $A, B, C \in M_p(\mathbb{F})$, sådan at

$$a_{ij} = b_{ij} = c_{ij}$$

for alle $1 \leq i, j \leq p$ med $j \neq k$, mens

$$a_{ik} = b_{ik} + c_{ik},$$

for alle $1 \leq i \leq p$. Så for alle $1 \leq j \leq p$ med $j \neq k$ er den j 'te søjle i A lig med den j 'te søjle i både B og C , mens den k 'te søjle i A er lig med summen af den k 'te søjle i B og den k 'te søjle i C , og vi ønsker at vise, at $\det(A) = \det(B) + \det(C)$. Da (D1) per induktion gælder for kvadratiske matricer af orden $p - 1$, gælder det, at

$$\det(A_{1j}) = \det(B_{1j}) + \det(C_{1j})$$

for alle $1 \leq j \leq p$ med $j \neq k$, mens

$$\det(A_{1k}) = \det(B_{1k}) + \det(C_{1k}),$$

fordi $A_{1k} = B_{1k} + C_{1k}$. Derfor får vi nu ved at indsætte i (3.2.6), at

$$\begin{aligned} \det(A) &= \sum_{j=1}^p (-1)^{1+j} a_{1j} \det(A_{1j}) \\ &= \sum_{j=1}^{k-1} (-1)^{1+j} a_{1j} (\det(B_{1j}) + \det(C_{1j})) + (-1)^{1+k} (b_{1k} + c_{1k}) \det(A_{1k}) \\ &\quad + \sum_{j=k+1}^p (-1)^{1+j} a_{1j} (\det(B_{1j}) + \det(C_{1j})) \\ &= \sum_{j=1}^p (-1)^{1+j} b_{1j} \det(B_{1j}) + \sum_{j=1}^p (-1)^{1+j} c_{1j} \det(C_{1j}) = \det(B) + \det(C) \end{aligned}$$

som ønsket. Dette viser, at (D1) gælder for $n = p$, og beviset for (D2) er tilsvarende.

For at vise (D3), betragter vi først tilfældet, hvor $1 \leq k < p$ og den k 'te og $(k+1)$ 'te søjle i A er ens. Hvis $j \neq k$ og $j \neq k+1$, da har undermatricen A_{1j} igen to ens søjler, sådan at $\det(A_{1j}) = 0$ per induktion. Desuden er $A_{1k} = A_{1,k+1}$, sådan at

$$\begin{aligned} \det(A) &= \sum_{j=1}^p (-1)^{1+j} a_{1j} \det(A_{1j}) \\ &= (-1)^{1+k} a_{1k} \det(A_{1k}) + (-1)^{1+k+1} a_{1,k+1} \det(A_{1,k+1}) = 0, \end{aligned}$$

3.2 Determinant for $n \times n$ -matricer

hvilket viser, at (D3) gælder i tilfældet, hvor to nabosøjler er ens. Dette medfører ved et argument svarende til beviset for (D5) i sætning 3.2.3, at determinanten af en kvadratisk matrix af orden p skifter fortegn, hvis to nabosøjler ombyttes, hvilket vi nu bruger vi for at vise, at (D3) gælder i det generelle tilfælde. Så antag, at den j 'te og k 'te søjle i A er ens, hvor $1 \leq j < k \leq p$. Ved at lave $k - j - 1$ ombytninger af nabosøjler kan vi omdanne A til en matrix B vis j 'te og $j + 1$ 'st søjler er ens, og ved at bruge (D5) gentagende gange for nabosøjler, ser vi, at $\det(A) = (-1)^{k-j-1} \det(B)$, og vi har lige vist at $\det(B) = 0$ fordi den har to ens nabosøjler. Så $\det(A) = 0$ og (D3) følger.

Endelig følger det fra (3.2.6), at

$$\det(\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_m) = 1 \cdot \det(\mathbf{e}_1 \ \dots \ \mathbf{e}_{m-1}) + 0 + \dots + 0 = \det(\mathbf{e}_1 \ \dots \ \mathbf{e}_{m-1}),$$

og per induktion er den fælles værdi lig med 1, hvilket viser at også (D4) gælder. Vi har nu vist induktionsskridtet, og vi har således for alle $n \geq 0$ defineret en afbildning

$$\det: M_n(\mathbb{F}) \rightarrow \mathbb{F},$$

der opfylder (D1)–(D4).

Vi mangler at vise, at afbildningen $\det: M_n(\mathbb{F}) \rightarrow \mathbb{F}$ defineret ovenfor er den eneste afbildning, der opfylder (D1)–(D4), hvilket vi igen gør ved induktion på $n \geq 0$. Hvis $n = 0$, og $D: M_0(\mathbb{F}) \rightarrow \mathbb{F}$ opfylder (D1)–(D4), da er $D(\) = 1$ ifølge (D4), hvilket viser, at $D = \det$ som ønsket. Så vi antager, at $\det: M_{p-1}(\mathbb{F}) \rightarrow \mathbb{F}$ er den eneste afbildning, der opfylder (D1)–(D4), og lader $D: M_p(\mathbb{F}) \rightarrow \mathbb{F}$ være en afbildning, der opfylder (D1)–(D4), og viser, at $D = \det$. Hvis $A = (\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) \in M_p(\mathbb{F})$, så kan vi anvende lemma 2.2.6 til at skrive den første søjle $\mathbf{a}_1$ som en linearkombination

$$\mathbf{a}_1 = \mathbf{e}_1 a_{11} + \mathbf{e}_2 a_{21} + \dots + \mathbf{e}_p a_{p1}$$

af standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p)$ for $\mathbb{F}^p$ og konkluderer da, at

$$\begin{aligned} D(A) &= D(\mathbf{e}_1 a_{11} \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) + D(\mathbf{e}_2 a_{21} \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) + \dots + D(\mathbf{e}_p a_{p1} \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) \\ &= D(\mathbf{e}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) a_{11} + D(\mathbf{e}_2 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) a_{21} + \dots + D(\mathbf{e}_p \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) a_{p1}, \end{aligned}$$

hvor vi anvender henholdsvis (D1) og (D2) for afbildningen D . Vi påstår, at

$$D(\mathbf{e}_i \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) = (-1)^{i-1} \det(A_{i1})$$

for alle $1 \leq i \leq p$. Vi konkluderer heraf, at

$$D(A) = \sum_{i=1}^p (-1)^{i-1} \det(A_{i1}) a_{i1},$$

hvilket viser, at $D = \det$, idet den samme formel nødvendigvis gælder for $\det(A)$, da også $\det: M_p(\mathbb{F}) \rightarrow \mathbb{F}$ opfylder (D1)–(D4). For at bevise påstanden, lader vi $1 \leq i \leq p$ være fastholdt og anvender (D6), som er en konsekvens af (D1)–(D3), til at skrive

$$D(\mathbf{e}_i \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p) = D(\mathbf{e}_i \ \mathbf{a}_2 - \mathbf{e}_i a_{i2} \ \dots \ \mathbf{a}_p - \mathbf{e}_i a_{ip}).$$

3 Determinant

Vi bemærker, at matricen til højre kun afhænger af den kvadratiske matrix A_{i1} af orden $p-1$. Mere præcist, hvis $B \in M_{p-1}(\mathbb{F})$ er en vilkårlig kvadratisk matrix af orden $p-1$, da definerer vi $\tilde{B} \in M_p(\mathbb{F})$ til at være matricen

$$\tilde{B} = \begin{pmatrix} 0 & b_{1,1} & \dots & b_{1,p-1} \\ \vdots & \vdots & & \vdots \\ 0 & b_{i-1,1} & \dots & b_{i-1,p-1} \\ 1 & 0 & \dots & 0 \\ 0 & b_{i,1} & \dots & b_{i,p-1} \\ \vdots & \vdots & & \vdots \\ 0 & b_{p-1,1} & \dots & b_{p-1,p-1} \end{pmatrix},$$

der er kvadratisk af orden p , og vi har da, at

$$\begin{pmatrix} \mathbf{e}_i & \mathbf{a}_2 - \mathbf{e}_i a_{i2} & \dots & \mathbf{a}_p - \mathbf{e}_i a_{ip} \end{pmatrix} = \begin{pmatrix} 0 & a_{1,2} & \dots & a_{1,p} \\ \vdots & \vdots & & \vdots \\ 0 & a_{i-1,2} & \dots & a_{i-1,p} \\ 1 & 0 & \dots & 0 \\ 0 & a_{i+1,2} & \dots & a_{i+1,p} \\ \vdots & \vdots & & \vdots \\ 0 & a_{p,2} & \dots & a_{p,p} \end{pmatrix} = \widetilde{A_{i1}}.$$

Vi bemærker dernæst, at afbildningen $d: M_{p-1}(\mathbb{F}) \rightarrow \mathbb{F}$ defineret ved

$$d(B) = (-1)^{i-1} D(\tilde{B})$$

opfylder (D1)–(D4). For egenskaberne (D1)–(D3) følger fra de tilsvarende egenskaber for D , mens (D4) følger fra udregningen

$$d(\mathbf{e}_1 \dots \mathbf{e}_{p-1}) = (-1)^{i-1} D(\mathbf{e}_i \mathbf{e}_1 \dots \mathbf{e}_{i-1} \mathbf{e}_{i+1} \dots \mathbf{e}_p) = D(\mathbf{e}_1 \dots \mathbf{e}_p) = 1,$$

hvor den første identitet er definitionen af d , den anden følger fra (D5) for D , idet vi for at ændre $(\mathbf{e}_i \mathbf{e}_1 \dots \mathbf{e}_{i-1} \mathbf{e}_{i+1} \dots \mathbf{e}_p)$ til $(\mathbf{e}_1 \dots \mathbf{e}_{i-1} \mathbf{e}_i \mathbf{e}_{i+1} \dots \mathbf{e}_p)$ skal bytte om på to rækker $i-1$ gange, mens den tredje følger fra (D4) for D . Det følger derfor fra den induktive hypotese, at $d(B) = \det(B)$ for alle $B \in M_{p-1}(\mathbb{F})$. Især er

$$(-1)^{i-1} D(\widetilde{A_{i1}}) = d(\widetilde{A_{i1}}) = \det(A_{i1}),$$

hvilket viser den ønskede påstand, at

$$D(\mathbf{e}_i \mathbf{a}_2 \dots \mathbf{a}_p) = D(\widetilde{A_{i1}}) = (-1)^{i-1} \det(A_{i1}).$$

Dette beviser induktionsskridtet og dermed entydigheden af determinanten. $\square$

Eksempel 3.2.7 Vi anvender determinantens egenskaber til at udregne:

$$\begin{aligned} \det \begin{pmatrix} 2 & 0 & 0 \\ 2 & 3 & 0 \\ 2 & 3 & 4 \end{pmatrix} &\stackrel{(D2)}{=} \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 3 & 4 \end{pmatrix} \cdot 2 \stackrel{(D2)}{=} \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 4 \end{pmatrix} \cdot 3 \cdot 2 \stackrel{(D2)}{=} \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \cdot 4 \cdot 3 \cdot 2 \\ &\stackrel{(D6)}{=} \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot 4 \cdot 3 \cdot 2 \stackrel{(D6)}{=} \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot 4 \cdot 3 \cdot 2 \stackrel{(D4)}{=} 1 \cdot 4 \cdot 3 \cdot 2 = 24, \end{aligned}$$

hvor de første tre ligheder bruger (D2) på den første, anden, og tredje søjle i matricen, den fjerde erstatter den første søjle med den første søjle minus den anden søjle, og den femte lighed erstatter den anden søjle med den anden søjle minus den tredje søjle.

Det følger fra beviset for sætning 3.2.1, at der gælder følgende rekursive formler for determinanten. Disse formler kaldes for *Laplace-udvikling*, da de skyldes Laplace.

Sætning 3.2.8 (Laplace) Lad A være en $n \times n$ -matrix med indgange i et legeme $\mathbb{F}$.

(D7) Determinanten af A kan udregnes ved udvikling langs i 'te række:

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}).$$

(D8) Determinanten af A kan udregnes ved udvikling langs j 'te søjle:

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} \det(A_{ij}) a_{ij}.$$

Bevis Vi beviser først (D7). I eksistensdelen af beviset for sætning 3.2.1 beviste vi, at $\det(A)$ defineret ved udvikling langs første række opfylder (D1)–(D4). Valget af den første række spillede dog ingen rolle, og en mindre modifikation af beviset viser at udvikling af determinanten langs en vilkårlig række ligeledes opfylder (D1)–(D4). Resultatet følger derfor fra entydigheden af determinanten.

Vi beviser dernæst (D8). Ifølge beviset for entydighedsdelen af sætning 3.2.1 er

$$\det(A) = \sum_{i=1}^n (-1)^{i-1} \det(A_{i1}) a_{i1} = \sum_{i=1}^n (-1)^{i+1} \det(A_{i1}) a_{i1},$$

hvilket viser (D8) for $j = 1$. Vi lader nu B være den matrix, der fremkommer fra A ved

3 Determinant

at ombytte den første søjle og den j 'te søjle, og har da identiteten

$$\det(B) = \sum_{i=1}^n (-1)^{i+1} \det(B_{i1}) b_{i1},$$

som vi netop har bemærket. Det følger da fra (D5) er $\det(B) = -\det(A)$, og per definition af matricen B er $b_{i1} = a_{ij}$. Endvidere fremkommer B_{i1} fra $n \times (n-1)$ -matricen

$$B_1 = (\mathbf{b}_2 \cdots \mathbf{b}_{j-1} \mathbf{b}_j \mathbf{b}_{j+1} \cdots \mathbf{b}_n) = (\mathbf{a}_2 \cdots \mathbf{a}_{j-1} \mathbf{a}_1 \mathbf{a}_{j+1} \cdots \mathbf{a}_n)$$

ved at fjerne den i 'te række, mens A_{ij} fremkommer fra $n \times (n-1)$ -matricen

$$A_j = (\mathbf{a}_1 \mathbf{a}_2 \cdots \mathbf{a}_{j-1} \mathbf{a}_{j+1} \cdots \mathbf{a}_n)$$

ved at fjerne den i 'te række. Vi bemærker nu, at A_j kan omdannes til B_1 ved at foretage $j-2$ ombytninger af to søjler. Det følger heraf, at A_{ij} ligeledes kan omdannes til B_{i1} ved at foretage $j-2$ ombytninger af to søjler, og derfor viser (D5), at

$$\det(B_{i1}) = (-1)^{j-2} \det(A_{ij}).$$

Vi har således alt i alt vist, at

$$\begin{aligned} \det(A) &= -\det(B) = -\sum_{i=1}^n (-1)^{i+1} \det(B_{i1}) b_{i1} \\ &= -\sum_{i=1}^n (-1)^{i+1+j-2} \det(A_{ij}) a_{ij} = \sum_{i=1}^n (-1)^{i+j} \det(A_{ij}) a_{ij} \end{aligned}$$

som ønsket. Dette viser (D8) og dermed sætningen. □

Eksempel 3.2.9 Vi udregner determinanten af matricen

$$A = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix}$$

fra eksempel 3.2.5. Laplace-udvikling langs første søjle giver

$$\det(A) = \det(A_{11}) \cdot 5 - \det(A_{21}) \cdot 0 + \det(A_{31}) \cdot 2,$$

hvor

$$A_{11} = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 6 & 4 \end{pmatrix} \quad \text{og} \quad A_{31} = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & -3 \end{pmatrix}.$$

Vi behøver ikke at udregne A_{21} og dens determinant, da $a_{21} = 0$. Fra sætning 3.1.1 har vi nu, at $\det(A_{11}) = 2 \cdot 4 - (-3) \cdot 6 = 26$ og $\det(A_{31}) = (-1) \cdot (-3) - 1 \cdot 2 = 1$, så

$$\det(A) = 26 \cdot 5 + 1 \cdot 2 = 132.$$

Vi bemærker, at det altid er en fordel at udvikle en determinant langs en række eller søjle, der indeholder mange nuller, jo flere desto bedre. For dermed behøver man ikke udregne de tilsvarende summander i Laplace-udviklingen. I dette eksempel kunne vi lige så godt have valgt at udvikle determinanten langs den anden række.

Bemærkning 3.2.10 Det er naturligvis vigtigt at huske fortegnene i udviklingen af determinant efter en række eller en søjle. De minder om et skakbræt

$$(+), \quad \begin{pmatrix} + & - \\ - & + \end{pmatrix}, \quad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}, \quad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}, \quad \dots,$$

hvor man begynder med “+” i det øverste venstre hjørne.

Sætning 3.2.11 Lad $\mathbb{F}$ være et legeme og lad $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$ være en skævinvolution. Hvis A er en $n \times n$ -matrix med indgange i $\mathbb{F}$, og hvis A^* er den adjungerede matrix, da er

$$\det(A^*) = \det(A)^*.$$

Bevis Beviset er igen ved induktion på $n \geq 0$. For $n = 0$ er sætningen trivielt, så vi antager, at sætningen allerede er bevist for $n = p - 1$ og beviser den for $n = p$. Lad os skrive $a'_{ij} = a^*_{ji}$ for den (i, j) 'te indgang i A^* . Da endvidere $(A^*)_{ij} = (A_{ji})^*$, gælder det, at

$$\begin{aligned} \det(A^*) &= \sum_{i=1}^p (-1)^{i+1} \det((A^*)_{i1}) a'_{i1} = \sum_{j=1}^p (-1)^{1+j} \det((A_{1j})^*) a^*_{1j} \\ &= \sum_{j=1}^p (-1)^{1+j} \det(A_{1j})^* a^*_{1j} = \left(\sum_{j=1}^p (-1)^{1+j} a_{1j} \det(A_{1j}) \right)^* = \det(A)^*, \end{aligned}$$

hvor den første lighed følger fra (D8); den anden fra vores indledende bemærninger; den tredje fra den induktive hypotese, at sætningen gælder for $(p-1) \times (p-1)$ -matricer; den fjerde fra definitionen af en antiinvolution; og den femte fra (D7). Vi har nu vist induktionsskridtet og dermed sætningen. $\square$

3 Determinant

Korollar 3.2.12 Hvis A er en kvadratisk matrix med indgange i et legeme $\mathbb{F}$, og hvis A^t er den transponerede matrix, da gælder det, at $\det(A^t) = \det(A)$.

Bevis Dette er udsagnet i sætning 3.2.11 i det tilfælde, hvor $(-)^* = \text{id}: \mathbb{F} \rightarrow \mathbb{F}$. □

Eksempel 3.2.13 Korollar 3.2.12 og eksempel 3.2.7 viser, at

$$\det \begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{pmatrix} = \det \begin{pmatrix} 2 & 0 & 0 \\ 2 & 3 & 0 \\ 2 & 3 & 4 \end{pmatrix} = 24.$$

Desuden viser sætning 3.1.1, at

$$\det \begin{pmatrix} 2 & 1 \\ -i & 4 \end{pmatrix} = 2 \cdot 4 - 1 \cdot (-i) = 8 + i \quad \text{og} \quad \det \begin{pmatrix} 2 & i \\ 1 & 4 \end{pmatrix} = 2 \cdot 4 - i \cdot 1 = 8 - i.$$

Så de to komplekse matrixer, der er hinandens adjungerede matrix, har konjugerede determinanter, som de skal have ifølge sætning 3.2.11.

Korollar 3.2.14 Hvis A er en hermitisk kompleks matrix, da er $\det(A)$ et reelt tal.

Bevis Vi lader $\mathbb{F} = \mathbb{C}$ med kompleks konjugation som skævinvolution. Vi minder om, at en kvadratisk matrix A med indgange i $\mathbb{C}$ per definition er hermitisk, hvis $A = A^*$. Vi konkluderer derfor fra sætning 3.2.11, at

$$\det(A) = \det(A^*) = \det(A)^*,$$

hvilket viser, at $\det(A)$ er et reelt tal. For $z \in \mathbb{C}$ er reel, hvis og kun hvis $z = z^*$. □

Eksempel 3.2.15 Ifølge sætning 3.1.1 er

$$\det \begin{pmatrix} 2 & i \\ -i & 4 \end{pmatrix} = 2 \cdot 4 - (-i) \cdot i = 7.$$

Determinanten af denne komplekse hermitiske matrix er altså reel, hvilket den skal være ifølge korollar 3.2.14.

Den følgende sætning giver en version af (D1)–(D3) og (D5)–(D6) for rækkevektorer i stedet for søjlevektorer i matrixen.

Sætning 3.2.16 *Lad $\mathbb{F}$ være et legeme og n et naturligt tal. Følgende udsagn gælder:*

(D1') *For alle $1 \leq j \leq n$ og $\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{b}_j, \mathbf{c}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n \in M_{1,n}(\mathbb{F})$ er*

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{b}_j + \mathbf{c}_j \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{b}_j \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{c}_j \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

(D2') *For alle $1 \leq j \leq n$, $\mathbf{a}_1, \dots, \mathbf{a}_n \in M_{1,n}(\mathbb{F})$ og $c \in \mathbb{F}$ gælder det, at*

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ c \cdot \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = c \cdot \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

(D3') *Hvis $A \in M_{n,n}(\mathbb{F})$ har to ens rækker, da er $\det(A) = 0$.*

(D5') *Hvis B fremkommer ved ombytning af to rækker i A , da er $\det(B) = -\det(A)$.*

(D6') *Hvis B fremkommer fra A ved at addere et multiplum af en række i A til en anden række i A , da er $\det(B) = \det(A)$.*

Bevis Hvis $\mathbf{a}_1, \dots, \mathbf{a}_n \in M_{1,n}(\mathbb{F})$ er rækkevektorer, da er $\mathbf{a}_1^t, \dots, \mathbf{a}_n^t \in M_{n,1}(\mathbb{F})$ søjlevektorer, og ifølge korollar 3.2.14 har $n \times n$ -matricen

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \in M_n(\mathbb{F})$$

den samme determinant som dens transponerede $n \times n$ -matrix

$$A^t = \begin{pmatrix} \mathbf{a}_1^t \\ \vdots \\ \mathbf{a}_n^t \end{pmatrix} = (\mathbf{a}_1^t \cdots \mathbf{a}_n^t) \in M_n(\mathbb{F}).$$

De ønskede identiteter (D1')–(D3') og (D5')–(D6') angående rækkevektorer følger nu fra de tilsvarende identiteter (D1)–(D3) og (D5)–(D6) angående søjlevektorer. $\square$

3 Determinant

Vi illustrerer i detaljer i næste afsnit, hvordan rækkeoperationer og søjleoperationer anvendes til at udregne determinanter. Men først viser vi et par yderligere resultater om determinanten.

Sætning 3.2.17 Hvis A og B er $n \times n$ -matricer med indgange i et legeme $\mathbb{F}$, da er

$$\det(AB) = \det(A)\det(B).$$

Bevis Givet $a \in \mathbb{F}$, betragter vi afbildningen $\det_a: M_n(\mathbb{F}) \rightarrow \mathbb{F}$ defineret ved

$$\det_a(B) = a \cdot \det(B).$$

Ved en mindre modifikation af beviset for entydighedsdelen af sætning 3.2.1 kan man vise, at denne afbildning er entydigt bestemt ved, at den opfylder (D1)–(D3) i sætning 3.2.1 samt den nye betingelse (D4)_a, at $\det_a(I_n) = a$. Det er derfor tilstrækkeligt at vise, at afbildningen $d: M_n(\mathbb{F}) \rightarrow \mathbb{F}$ defineret ved

$$d(B) = \det(AB)$$

opfylder (D1)–(D3) samt (D4)_a, hvor $a = \det(A)$. For da giver entydighedsudsagnet, at $d(B) = \det_a(B)$, hvilket præcis giver $\det(AB) = \det(A)\det(B)$ som ønsket. Vi viser nu, at afbildningen $d: M_n(\mathbb{F}) \rightarrow \mathbb{F}$ opfylder (D1). Hvis $B = (\mathbf{b}_1 \ \cdots \ \mathbf{b}_k \ \cdots \ \mathbf{b}_n)$, da er

$$AB = \left(\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{array} \right) \left(\begin{array}{c|c|c|c|c} b_{11} & \cdots & b_{1k} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2k} & \cdots & b_{2n} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{nn} \end{array} \right) = (A\mathbf{b}_1 \ \cdots \ A\mathbf{b}_k \ \cdots \ A\mathbf{b}_n)$$

ifølge definitionen af matrix produktet. Hvis derfor $\mathbf{b}_k = \mathbf{c}_k + \mathbf{d}_k$, da er

$$\begin{aligned} d(\mathbf{b}_1 \ \cdots \ \mathbf{b}_k \ \cdots \ \mathbf{b}_n) &= \det(A\mathbf{b}_1 \ \cdots \ A(\mathbf{c}_k + \mathbf{d}_k) \ \cdots \ A\mathbf{b}_n) \\ &= \det(A\mathbf{b}_1 \ \cdots \ A\mathbf{c}_k + A\mathbf{d}_k \ \cdots \ A\mathbf{b}_n) \\ &= \det(A\mathbf{b}_1 \ \cdots \ A\mathbf{c}_k \ \cdots \ A\mathbf{b}_n) + \det(A\mathbf{b}_1 \ \cdots \ A\mathbf{d}_k \ \cdots \ A\mathbf{b}_n) \\ &= d(\mathbf{b}_1 \ \cdots \ \mathbf{c}_k \ \cdots \ \mathbf{b}_n) + d(\mathbf{b}_1 \ \cdots \ \mathbf{d}_k \ \cdots \ \mathbf{b}_n), \end{aligned}$$

hvilket viser, at $d: M_n(\mathbb{F}) \rightarrow \mathbb{F}$ opfylder (D1). Her følger den første lighed og fjerde lighed fra definitionen af matrix produktet, den anden lighed følger fra den distributive lov for matrixmultiplikation, mens den tredje lighed følger fra (D1) for determinanten. Beviset for (D2) og (D3) er tilsvarende, mens (D4)_a fås fra $d(I) = \det(AI) = \det(A) = a$. $\square$

Korollar 3.2.18 Hvis en kvadratisk matrix A er invertibel, så er $\det(A)$ invertibel, og

$$\det(A^{-1}) = \det(A)^{-1}.$$

Især er $\det(A) \neq 0$.

Bevis Da $AA^{-1} = I = A^{-1}A$, viser sætning 3.2.17, at

$$\det(A) \cdot \det(A^{-1}) = \det(AA^{-1}) = 1 = \det(A^{-1}A) = \det(A^{-1}) \cdot \det(A),$$

idet $\det(I) = 1$. Derfor er $\det(A)$ invertibel med invers $\det(A^{-1})$. $\square$

Eksempel 3.2.19 Vi har set i eksempel 2.4.10, at matricerne

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 3/5 & -2/5 \\ 1/5 & 1/5 \end{pmatrix} = A^{-1}$$

er hinandens inverse, og ifølge sætning 3.1.1 er deres determinanter

$$\det(A) = 1 \cdot 3 - (-1) \cdot 2 = 5 \quad \text{og} \quad \det(B) = \frac{3}{5} \cdot \frac{1}{5} - \frac{1}{5} \cdot \left(-\frac{2}{5}\right) = \frac{5}{25} = \frac{1}{5}.$$

De er dermed hinanden inverse, som de skal være ifølge korollar 3.2.18.

Bemærkning 3.2.20 Hvis A og B er $n \times n$ -matricer, da medfører den kommutative lov for multiplikation af skalarer samt sætning 3.2.17, at det altid gælder, at

$$\det(AB) = \det(A)\det(B) = \det(B)\det(A) = \det(BA),$$

også selvom matricerne AB og BA sædvanligvis ikke er ens. Vi bemærker også, at

$$\det(A+B) \neq \det(A) + \det(B),$$

undtaget i trivielle tilfælde.

Eksempel 3.2.21 Vi betragter 2×2 -matricerne

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{med} \quad AB = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{og} \quad BA = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}.$$

Selvom $AB \neq BA$, så er $\det(AB) = -1 = \det(BA)$.

3 Determinant

Vi giver endelig en lukket¹ formel for determinanten af en $n \times n$ -matrix. En bijektiv afbildning $\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ kaldes for en *permutation af n bogstaver*. Der er

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 1$$

forskellige permutationer af n bogstaver. En permutation σ af n bogstaver kaldes for en *transposition*, hvis der findes $1 \leq k < l \leq n$, sådan at

$$\sigma(i) = \begin{cases} l & \text{hvis } i = k, \\ k & \text{hvis } i = l, \\ i & \text{ellers.} \end{cases}$$

Man kan vise, at enhver permutation kan udtrykkes som en sammensætning af et antal transpositioner. En sådan opskrivning er dog generelt ikke entydig.

Eksempel 3.2.22 De $2! = 2$ permutationer af 2 bogstaver er identitetsafbildningen $\text{id}: \{1, 2\} \rightarrow \{1, 2\}$ og transpositionen $\sigma: \{1, 2\} \rightarrow \{1, 2\}$, der ombytter 1 og 2.

Til en permutation af n bogstaver $\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ tilordner vi matricen

$$P(\sigma) = (\mathbf{e}_{\sigma(1)} \quad \mathbf{e}_{\sigma(2)} \quad \dots \quad \mathbf{e}_{\sigma(n)}) \in M_n(\mathbb{F}),$$

som vi kalder for den tilhørende *permutationsmatrix*. Hvis σ og τ er to permutationer af n bogstaver, så er deres sammensætning $\sigma \circ \tau$ igen en permutation af n bogstaver, og

$$P(\sigma \circ \tau) = P(\sigma) \cdot P(\tau).$$

Vi definerer *fortegnet* af en permutationen σ af n bogstaver til at være determinanten

$$\text{sign}(\sigma) = \det(P(\sigma)) = \det(\mathbf{e}_{\sigma(1)} \quad \mathbf{e}_{\sigma(2)} \quad \dots \quad \mathbf{e}_{\sigma(n)}) \in \mathbb{F}.$$

Eksempel 3.2.23 For de to permutationer af to bogstaver i eksempel 3.2.22 er

$$P(\text{id}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{og} \quad P(\sigma) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

så deres fortegn er henholdsvis $\text{sign}(\text{id}) = 1$ og $\text{sign}(\sigma) = -1$.

¹ Dette betyder, at formlen i modsætning til Laplace-udvikling ikke er rekursiv.

3.2 Determinant for $n \times n$ -matricer

Det følger fra sætning 3.2.17, at hvis σ og τ er to permutationer af n bogstaver, da er

$$\text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \text{sign}(\tau),$$

og det følger endvidere fra (D5) at fortegnet af en transposition er -1 . Generelt gælder det altså, at hvis σ er en permutation af n bogstaver, og hvis σ er en sammensætning af r transpositioner, da er $\text{sign}(\sigma) = (-1)^r \in \mathbb{F}$.

Sætning 3.2.24 (Leibniz) Lad $A = (a_{ij})$ være en kvadratisk matrix af orden n med indgange i et legeme $\mathbb{F}$. Da er

$$\det(A) = \sum_{\sigma} \text{sign}(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n},$$

hvor summen løber over de $n!$ mulige permutationer af n bogstaver.

Bevis For alle $1 \leq j \leq n$ skriver vi den j 'te søjle $\mathbf{a}_j$ i A som en linearkombination

$$\mathbf{a}_j = \sum_{i=1}^n \mathbf{e}_i a_{ij}$$

af standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ for $\mathbb{F}^n$. Vi får da ved gentagen brug af (D1) og (D2), at

$$\begin{aligned} \det(A) &= \det(\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n) \\ &= \sum_{\sigma} \det(\mathbf{e}_{\sigma(1)} \quad \mathbf{e}_{\sigma(2)} \quad \cdots \quad \mathbf{e}_{\sigma(n)}) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}, \end{aligned}$$

hvor summen løber over de n^n mulige afbildninger $\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$. Hvis en sådan afbildning σ ikke er injektiv, da har matrixen $(\mathbf{e}_{\sigma(1)} \quad \mathbf{e}_{\sigma(2)} \quad \cdots \quad \mathbf{e}_{\sigma(n)})$ to eller flere søjler, der er ens, og derfor viser (D3), at dens determinant lig med nul. Det er således kun summander, der er indicerede af injektive, eller ækvivalent, bijektive afbildninger $\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$, som bidrager til summen. Dette viser sætningen. $\square$

Eksempel 3.2.25 For en 2×2 -matrix specialiserer sætning 3.2.24 til den formel

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12},$$

som vi gav i sætning 3.1.1; og for en 3×3 -matrix, får vi formlen

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

med $3! = 6$ summander. For en 4×4 -matrix har formlen $4! = 24$ summander, mens den for en 5×5 -matrix giver den $5! = 120$ summander.

3.3 Triangulære matricer og beregning af determinanter

Der er to hovedmetoder til beregning af determinanter: (1) at bruge Laplace-udvikling langs en række eller en søjle samt (2) at anvende rækkeoperationer og søjleoperationer til at omdanne matricen til en triangulær matrix, som vi nu definerer. Man skal dog her være forsigtig med at huske på, at en rækkeoperation af type $M_i(c)$ og en søjleoperation af type $M_j(c)$ begge ændrer determinantens værdi med faktoren c .

Definition 3.3.1 En $n \times n$ -matrix $B = (b_{ij})$ kaldes for *nedre triangulær*, hvis $b_{ij} = 0$ for alle $1 \leq i < j \leq n$; *øvre triangulær*, hvis $b_{ij} = 0$ for alle $1 \leq j < i \leq n$; og *triangulær*, hvis den er enten nedre triangulær eller øvre triangulær.

Specielt er en diagonal matrix en triangulær matrix. Generelt er diagonalindgangene i en triangulær matrix af særlig betydning.

Eksempel 3.3.2 Blandt 3×3 -matricerne

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 8 & 7 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 6 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 12 \end{pmatrix} \quad \text{og} \quad C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

er A nedre triangulær og B øvre triangulær, mens C er både øvre og nedre triangulær og derfor en diagonal matrix. Vi har markeret diagonalindgangene med blå.

Sætning 3.3.3 Determinanten af en triangulær $n \times n$ -matrix B er lig med produktet

$$\det(B) = b_{11}b_{22} \cdots b_{nn}$$

af diagonalindgangene.

Bevis Vi beviser udsagnet for B nedre triangulær; beviset for B øvre triangulær er ganske tilsvarende. Beviset er ved induktion på $n \geq 0$, og tilfældet $n = 0$ gælder trivielt. Vi antager derfor, at udsagnet er vist for $n = p - 1$ og viser det for $n = p$. Så vi lader

$$B = \begin{pmatrix} b_{11} & 0 & 0 & \cdots & 0 \\ b_{21} & b_{22} & 0 & \cdots & 0 \\ b_{31} & b_{32} & b_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & b_{p3} & \cdots & b_{pp} \end{pmatrix}$$

3.3 Triangulære matricer og beregning af determinanter

og får ved Laplace-udvikling langs første række, at

$$\det(B) = b_{11} \cdot \det(B_{11}) + 0 \cdot \det(B_{12}) + \cdots + 0 \cdot \det(B_{1p}) = b_{11} \cdot \det(B_{11}).$$

Endvidere er undermatricen B_{11} en nedre triangulær $(p-1) \times (p-1)$ -matrix med diagonalindgange $b_{22}, \dots, b_{pp}$, så ifølge den induktive antagelse er

$$\det(B_{11}) = b_{22} \cdots b_{pp}.$$

Dette viser induktionsskridtet og dermed sætningen. $\square$

Det er altså let at bestemme determinanten af en triangulær matrix. Generelt for en kvadratisk matrix A består den mest effektive strategi til bestemmelse af $\det(A)$ i at anvende række- og søjleoperationer til at omdanne A til en triangulær matrix B , hvorved $\det(A)$ udtrykkes ved $\det(B)$ ved hjælp af sætning 3.2.1, 3.2.3 og 3.2.16. Da $\det(B)$ er givet ved sætning 3.3.3, bestemmer dette $\det(A)$.

Eksempel 3.3.4 Vi betragter 4×4 -matricen

$$A = \begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 0 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix},$$

og anvender rækkeoperationer til at omdanne den til en triangulær matrix B , idet vi er omhyggelige med at notere, hvorledes disse operationer påvirker determinanten.

$$\begin{aligned} \det(A) &\stackrel{(D5')}{=} -\det \begin{pmatrix} 1 & 1 & 3 & 5 \\ \color{red}{1} & 2 & 1 & 3 \\ 0 & 0 & 2 & 3 \\ \color{red}{3} & 1 & 5 & 9 \end{pmatrix} \stackrel{(D1')}{=} -\det \begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 2 & 3 \\ 0 & \color{red}{-2} & -4 & -6 \end{pmatrix} \\ &\stackrel{(D1')}{=} -\det \begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & \color{red}{-8} & -10 \end{pmatrix} \stackrel{(D1')}{=} -\det \begin{pmatrix} \color{blue}{1} & 1 & 3 & 5 \\ 0 & \color{blue}{1} & -2 & -2 \\ 0 & 0 & \color{blue}{2} & 3 \\ 0 & 0 & 0 & \color{blue}{2} \end{pmatrix} = -4. \end{aligned}$$

Her har vi først anvendt (D5') til at ombytte den anden og tredje række, hvilket ændrer determinantens fortegn. Derefter har vi anvendt (D1') til at addere -1 gange 1. række til 2. række og -3 gange 1. række til 4. række. Vi har så igen anvendt (D1') til at addere 2 gange 2. række til 4. række, og endelig har vi anvendt (D1') til at addere 4 gange 3. række til 4. række. Sætning 3.3.3, at determinanten af den øvre triangulære matrix B er 4, og vi konkluderer derfor, at $\det(A) = -4$.

3 Determinant

Eksempel 3.3.5 Til sammenligning bestemmer vi determinanten af 4×4 -matricen A i eksempel 3.3.4 ved Laplace-udvikling langs 2. række.

$$\begin{aligned}
 \det \begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 0 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix} &= -0 + 0 - 2 \cdot \det \begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 0 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 0 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix} \\
 &= -2 \cdot \det \begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 9 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} \\
 &= -2 \cdot \left(\det \begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 9 \end{pmatrix} \cdot 5 - \det \begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 9 \end{pmatrix} \cdot 3 + \det \begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 9 \end{pmatrix} \cdot 9 \right) \\
 &\quad + 3 \cdot \left(3 \cdot \det \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} + 5 \cdot \det \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} \right) \\
 &= -2 \cdot \left(\det \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \cdot 5 - \det \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \cdot 3 + \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot 9 \right) \\
 &\quad + 3 \cdot \left(3 \cdot \det \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + 5 \cdot \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \right) \\
 &= -2 \cdot \left((-5) \cdot 5 - (-2) \cdot 3 + 1 \cdot 9 \right) + 3 \cdot \left(3 \cdot (-5) - 1 \cdot (-2) + 5 \cdot 1 \right) \\
 &= -2 \cdot (-10) + 3 \cdot (-8) = 20 - 24 = -4.
 \end{aligned}$$

Vi har valgt at udvikle determinanten langs 2. række, fordi den har mange nuller. Vi har derefter udviklet den første og anden 3×3 -matrix efter henholdsvis 3. søjle og 2. række, markeret med rødt. Og endelig har vi brugt formelen i sætning 3.1.1 til at udregne de seks 2×2 -determinanter.

Laplace-udvikling er mest anvendelig til bestemmelse af determinanten af matricer, hvor de fleste indgange er nul.² Det er dog lettere at lave fejl og sværere at opdage dem ved Laplace-udvikling end ved udregning ved hjælp af række- og søjleoperationer.

Eksempel 3.3.6 Vi udregner determinanten af matricen

$$A = \begin{pmatrix} 2 & 16 & 3 \\ 4 & 8 & -6 \\ 8 & 16 & 12 \end{pmatrix}$$

²En sådan matrix kaldes for en "sparse matrix" på engelsk.

3.3 Triangulære matricer og beregning af determinanter

ved at anvende søjleoperationer til at omdanne A til en triangulære matrix.

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} 2 & 16 & 3 \\ 4 & 8 & -6 \\ 8 & 16 & 12 \end{pmatrix} \stackrel{(D2)}{=} \det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & -2 \\ 4 & 2 & 4 \end{pmatrix} \cdot 2^4 \cdot 3 \stackrel{(D6)}{=} \det \begin{pmatrix} 1 & 0 & 0 \\ 2 & -3 & -4 \\ 4 & -6 & 0 \end{pmatrix} \cdot 2^4 \cdot 3 \\ &\stackrel{(D2)}{=} \det \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 1 \\ 4 & 2 & 0 \end{pmatrix} \cdot 2^6 \cdot 3^2 \stackrel{(D6)}{=} \det \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 2 & -2 \end{pmatrix} \cdot 2^6 \cdot 3^2 = -2^7 3^2 = -1152. \end{aligned}$$

Vi har først anvendt (D2) til at flytte den fælles faktor 2 fra 1. søjle, den fælles faktor 8 fra 2. søjle og den fælles faktor 3 fra 3. søjle udenfor determinanten; dernæst har vi anvendt (D6) og har adderet 1. søjle gange -2 til 2. søjle og 1. søjle gange -1 til 3. søjle; vi har så anvendt (D2) til at flytte den fælles faktor -3 i 2. søjle og den fælles faktor -4 i 3. søjle udenfor determinanten; og endelig har vi anvendt (D6) og har adderet 2. søjle gange -1 til 3. søjle. Sætning 3.3.3 giver determinanten af den triangulære matrix.

Eksempel 3.3.7 Man kan ofte med fordel anvende både (D1)–(D6) og de tilsvarende rækkevarianter (D1')–(D6') til at udregne determinanter. Som eksempel udregner vi følgende determinant ved at ombytte rækker og søjler, idet vi husker, at en ombytning af to rækker eller søjler skifter fortegn på determinanten.

$$\begin{aligned} \det \begin{pmatrix} 0 & 0 & 0 & 0 & 3 \\ 0 & 2 & 0 & 0 & 5 \\ 0 & 3 & -2 & 0 & -4 \\ 0 & -6 & 1 & 2 & 2 \\ 8 & 1 & 2 & 3 & 4 \end{pmatrix} &\stackrel{(D5)}{=} -\det \begin{pmatrix} 8 & 1 & 2 & 3 & 4 \\ 0 & 2 & 0 & 0 & 5 \\ 0 & 3 & -2 & 0 & -4 \\ 0 & -6 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \\ &\stackrel{(D5)}{=} +\det \begin{pmatrix} 8 & 1 & 2 & 3 & 4 \\ 0 & -6 & 1 & 2 & 2 \\ 0 & 3 & -2 & 0 & -4 \\ 0 & 2 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \stackrel{(D5')}{=} -\det \begin{pmatrix} 8 & 3 & 2 & 1 & 4 \\ 0 & 2 & 1 & -6 & 2 \\ 0 & 0 & -2 & 3 & -4 \\ 0 & 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} = 2^6 \cdot 3 = 192. \end{aligned}$$

Her har vi markeret med rødt de rækker og søjler, vi ombytter, og i den øvre triangulære matrix til sidst har vi markeret diagonalindgangene med blåt. Til sidst har vi benyttet sætning 3.3.3 til at udregne determinanten af den triangulære matrix.

3.4 Determinant og invers matrix

Vi har set i korollar 3.2.18, at determinanten af en invertibel matrix er invertibel, og vi skal nu vise det omvendte udsagn. Mere præcist giver vi en formel for den inverse matrix udtrykt ved determinanten. Vi husker på, at et element i et legeme er invertibelt, hvis og kun hvis det er forskellig fra nul.

Sætning 3.4.1 *Lad $\mathbb{F}$ være et legeme, lad $A \in M_n(\mathbb{F})$, og antag, at $\det(A) \in \mathbb{F}$ er invertibel. Da er matricen A invertibel og matricen $C \in M_n(\mathbb{F})$, hvis (i, j) 'te indgang er*

$$c_{ij} = (-1)^{i+j} \det(A_{ji}) \det(A)^{-1},$$

er den inverse matrix af A . Her er A_{ji} den (j, i) 'te undermatrix af A .

Bevis Vi udregner den (i, k) 'te indgang i produkt matricen $D = AC$,

$$d_{ik} = \sum_{j=1}^n a_{ij} c_{jk} = \sum_{j=1}^n (-1)^{j+k} a_{ij} \det(A_{kj}) \det(A)^{-1},$$

og betragter $i = k$ og $i \neq k$ særskilt. Hvis $i = k$, da er

$$d_{ii} = \left(\sum_{j=1}^n (-1)^{j+i} a_{ij} \det(A_{ij}) \right) \det(A)^{-1} = \det(A) \det(A)^{-1} = 1,$$

hvor den midderste identitet er Laplace-udvikling af $\det(A)$ langs i 'te række. Hvis $i \neq k$, da betragter vi matricen B , der fremkommer fra A ved at erstatte den k 'te række med den i 'te række. På den ene side viser (D3'), at $\det(B) = 0$, da den i 'te og k 'te række i B er ens. Og på den anden side giver Laplace-udvikling langs den k 'te række, at

$$\det(B) = \sum_{j=1}^n (-1)^{k+j} b_{kj} \det(B_{kj}) = \sum_{j=1}^n (-1)^{k+j} a_{ij} \det(A_{kj}),$$

hvor den anden lighed følger fra definitionen af B . Derfor er

$$d_{ik} = (-1)^{i+k} \det(B) \det(A)^{-1} = 0.$$

Dermed har vi vist at $AC = D = I$. Vi ser ligeledes, at $CA = I$, ved enten at foretage en lignende udregning eller ved at bruge sætning 2.5.22, så $C = A^{-1}$. $\square$

Korollar 3.4.2 *Lad $\mathbb{F}$ være et legeme, lad $A \in M_n(\mathbb{F})$. Matricen A er invertibel i $M_n(\mathbb{F})$, hvis og kun hvis $\det(A)$ er invertibel i $\mathbb{F}$. I givet fald er $\det(A^{-1}) = \det(A)^{-1}$.*

Bevis Det følger af korollar 3.2.18 og sætning 3.4.1. $\square$

Bemærkning 3.4.3 Vi lader $A \in M_n(\mathbb{R})$ være en invertibel matrix og antager, at alle indgange i A er heltal. Da er $\det(A)$ og $\det(A_{ji})$ også heltal, og sætning 3.4.1 viser derfor, at indgangene i $C = A^{-1}$ er rationale tal med samme nævner $\det(A)$. Dette forklarer at matricen A^{-1} i eksempel 2.4.10 havde indgangene af form $m/5$ med $m \in \mathbb{Z}$, idet $\det(A) = 5$ i dette eksempel.

Eksempel 3.4.4 For 2×2 -matricer viser sætning 3.4.1, at

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \begin{pmatrix} a_{22}d^{-1} & -a_{12}d^{-1} \\ -a_{21}d^{-1} & a_{11}d^{-1} \end{pmatrix},$$

forudsat at determinanten $d = a_{11}a_{22} - a_{21}a_{12}$ er invertibel.

Bemærkning 3.4.5 Hvis $\mathbb{F}$ er et legeme, så skriver vi

$$SL_n(\mathbb{F}) \subset GL_n(\mathbb{F})$$

for delmængden af de invertible $n \times n$ -matricer, der består af de $n \times n$ -matricer, der har determinant lig med 1. At sådanne matricer er invertible følger fra sætning 3.4.1, idet $1 \in \mathbb{F}$ er invertibel. Det følger endvidere fra korollar 3.2.17, at hvis $A, B \in SL_n(\mathbb{F})$, da er også $AB \in SL_n(\mathbb{F})$. Dermed er $(SL_n(\mathbb{F}), \cdot)$ en gruppe, som defineret i bemærkning 2.4.15. Vi kalder denne gruppe for den specielle lineære gruppe. Gruppen $(SL_n(\mathbb{R}), \cdot)$ består således af de lineære isomorfier af $\mathbb{R}^n$, der bevarer volumen.

Hvis A er invertibel, så har ligningssystemet $A\mathbf{x} = \mathbf{b}$ den entydige løsning $\mathbf{x} = A^{-1}\mathbf{b}$. Denne løsning kan også udtrykkes ganske elegant ved hjælp af determinanter.

Sætning 3.4.6 (Cramers regel) Lad A være en invertibel $n \times n$ -matrix med indgange i et legeme $\mathbb{F}$, lad $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{F}^n$ være søjlevektorerne i A og lad $\mathbf{b} \in \mathbb{F}^n$ være en vilkårlig søjlevektor. Da er den entydige løsning $\mathbf{x} = A^{-1}\mathbf{b}$ til ligningen $A\mathbf{x} = \mathbf{b}$ givet ved

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{med} \quad x_i = \frac{\det(\mathbf{a}_1 \dots \mathbf{a}_{i-1} \mathbf{b} \mathbf{a}_{i+1} \dots \mathbf{a}_n)}{\det(\mathbf{a}_1 \dots \mathbf{a}_{i-1} \mathbf{a}_i \mathbf{a}_{i+1} \dots \mathbf{a}_n)}.$$

Bevis Formlen for $A^{-1} = C = (c_{ij})$ fra sætning 3.4.6 viser, at

$$x_i = (c_{i1} \ c_{i2} \ \dots \ c_{in}) \mathbf{b} = \sum_{j=1}^n c_{ij} b_j = \sum_{j=1}^n (-1)^{i+j} \det(A_{ji}) \det(A)^{-1} b_j,$$

3 Determinant

og vi skal derfor vise, at

$$\sum_{j=1}^m (-1)^{i+j} b_j \det(A_{ji}) = \det(\mathbf{a}_1 \dots \mathbf{a}_{i-1} \mathbf{b} \mathbf{a}_{i+1} \dots \mathbf{a}_n).$$

Summen på venstre side er præcis udviklingen langs i 'te søjle af determinanten på højre side, så den ønskede identitet følger fra sætning 3.2.8. $\square$

Eksempel 3.4.7 Vi betragter det reelle ligningssystem $A\mathbf{x} = \mathbf{b}$, hvor

$$A = \begin{pmatrix} 5 & 1 \\ 3 & 4 \end{pmatrix} \quad \text{og} \quad \mathbf{b} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Da $\det(A) = 17$ er invertibel, så er A ifølge sætning 3.4.1 også invertibel, og vi kan derfor anvende sætning 3.4.6 til at finde den entydige løsning $\mathbf{x}$. Sætningen siger, at

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{med} \quad x_1 = \frac{\det \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}}{\det \begin{pmatrix} 5 & 1 \\ 3 & 4 \end{pmatrix}} = \frac{5}{17} \quad \text{og} \quad x_2 = \frac{\det \begin{pmatrix} 5 & 2 \\ 3 & 3 \end{pmatrix}}{\det \begin{pmatrix} 5 & 1 \\ 3 & 4 \end{pmatrix}} = \frac{9}{17}.$$

3.5 Polynomier

Vi har allerede bemærket, at for at definere determinanten er det nødvendigt, at den kommutative lov (P3) gælder for multiplikation af skalarer, men vi har ikke brugt, at enhver skalar $a \in \mathbb{F}$ med $a \neq 0$ er invertibel. Vi kan derfor mere generelt tillade skalarer i en kommutativ ring, som vi nu definerer. Alle resultater i dette kapitel gælder stadig, hvis legemet $\mathbb{F}$ erstattes af den kommutativ ring $\mathbb{Z}$ af hele tal eller $\mathbb{F}[t]$ af polynomier med koefficienter i $\mathbb{F}$. Vi definerer sidstnævnte, som vi anvender i kapitel 5, nedenfor.

Definition 3.5.1 En *kommutativ ring* er en triple $(R, +, \cdot)$ af en mængde R og to afbildninger $+: R \times R \rightarrow R$ og $\cdot: R \times R \rightarrow R$, sådan at:

(A1) For alle $a, b, c \in R$ er $(a + b) + c = a + (b + c)$.

(A2) Der findes et element $0 \in R$, sådan at $a + 0 = a = 0 + a$ for alle $a \in R$.

(A3) For alle $a \in R$, findes $b \in R$, sådan at $a + b = 0 = b + a$.

(A4) For alle $a, b \in R$ er $a + b = b + a$.

(P1) For alle $a, b, c \in R$ er $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(P2) Der findes et element $1 \in R$, sådan at $a \cdot 1 = a = 1 \cdot a$ for alle $a \in R$.

(P4) For alle $a, b \in R$ er $a \cdot b = b \cdot a$.

(D1) For alle $a, b, c \in R$ er $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$

(D2) For alle $a, b, c \in R$ er $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$.

Vi misbruger ofte notation og skiver blot R for den kommutative ring $(R, +, \cdot)$. Vi siger, at $a \in R$ er *invertibel*, hvis der findes $b \in R$, sådan at $a \cdot b = 1 = b \cdot a$. Et legeme er således en kommutativ ring, hvor $0 \neq 1$, og hvor ethvert $a \in R$ med $a \neq 0$ er invertibelt.

Eksempel 3.5.2 (1) Mængden $\mathbb{Z}$ af hele tal sammen med afbildningerne $+: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ og $\cdot: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ givet ved sædvanlig sum og produkt af hele tal udgør en kommutativ ring. Denne kommutative ring er ikke et legeme, da $\pm 1 \in \mathbb{Z}$ er de eneste invertible elementer. For eksempel findes der ikke et $b \in \mathbb{Z}$, sådan at $2 \cdot b = 1 = b \cdot 2$.

(2) Mængden $C^0(\mathbb{R})$ af kontinuerte funktioner $f: \mathbb{R} \rightarrow \mathbb{R}$ sammen med afbildningerne $+: C^0(\mathbb{R}) \times C^0(\mathbb{R}) \rightarrow C^0(\mathbb{R})$ og $\cdot: C^0(\mathbb{R}) \times C^0(\mathbb{R}) \rightarrow C^0(\mathbb{R})$ defineret ved

$$(f + g)(x) = f(x) + g(x) \quad \text{og} \quad (f \cdot g)(x) = f(x) \cdot g(x)$$

udgør en kommutativ ring. Elementerne “0” og “1” i denne kommutative ring er de konstante funktioner med værdi 0 og 1, mens de invertible elementer er de kontinuerte funktioner $f: \mathbb{R} \rightarrow \mathbb{R}$, sådan at $f(x) \neq 0$ for *alle* $x \in \mathbb{R}$. For eksempel er $\text{id}: \mathbb{R} \rightarrow \mathbb{R}$ ikke invertibel, da $\text{id}(0) = 0$.

Vi definerer et *polynomium* med koefficienter i en kommutativ ring R til at være en følge $(a_0, a_1, a_2, \dots)$ af elementer i en kommutativ ring R , sådan at $a_i \neq 0$ for højst endeligt mange i . Vi vil dog altid indføre en variabel “ t ” og skrive

$$\sum_{i \geq 0} a_i t^i$$

i stedet for $(a_0, a_1, a_2, \dots)$, og vi siger da, at a_i er koefficienten til t^i . Vi definerer nu henholdsvis *summen* og *produktet* af to polynomier ved

$$\begin{aligned} \left(\sum_{i \geq 0} a_i t^i \right) + \left(\sum_{i \geq 0} b_i t^i \right) &= \sum_{i \geq 0} (a_i + b_i) t^i \\ \left(\sum_{i \geq 0} a_i t^i \right) \cdot \left(\sum_{j \geq 0} b_j t^j \right) &= \sum_{k \geq 0} \left(\sum_{i+j=k} a_i \cdot b_j \right) t^k. \end{aligned}$$

3 Determinant

Vi bemærker, at sum og produkt af polynomier er defineret ved, at vi lader som om, den variable “ t ” var et element i R .

Eksempel 3.5.3 Normalt undlader vi at skrive de koefficienter a_i , der er lig med 0. Så en typisk sum og et typisk produkt af to polynomier er for eksempel

$$\begin{aligned}(2t^3 - t^2 + 3t + 1) + (t^2 + 2t + 1) &= 2t^3 + 5t + 2, \\ (2t^3 - t^2 + 3t + 1) \cdot (t^2 + 2t + 1) &= 2t^5 + 3t^4 + 3t^3 + 5t^2 + 5t + 1.\end{aligned}$$

Specielt skriver vi 0 for det polynomium, hvis koefficienter alle er 0, og vi skriver 1 for det polynomium, hvis koefficient til t^i er 1 for $i = 0$ og 0 for $i > 0$.

Sætning 3.5.4 Hvis R er en kommutativ ring, da udgør mængden $R[t]$ af polynomier med koefficienter i R sammen med sum og produkt af polynomier en kommutativ ring.

Bevis Vi viser (P1) og (D1); (D2) vises ligesom (D1), mens de øvrige egenskaber følger umiddelbart fra de tilsvarende egenskaber for R . Ifølge definitionen af produktet af polynomier samt henholdsvis (D2) og (D1) for skalarer gælder det, at

$$\begin{aligned}\left(\sum_{i \geq 0} a_i t^i \cdot \sum_{j \geq 0} b_j t^j\right) \cdot \sum_{k \geq 0} c_k t^k &= \sum_{l \geq 0} \left(\sum_{i+j+k=l} (a_i b_j) c_k\right) t^l, \\ \sum_{i \geq 0} a_i t^i \cdot \left(\sum_{j \geq 0} b_j t^j \cdot \sum_{k \geq 0} c_k t^k\right) &= \sum_{l \geq 0} \left(\sum_{i+j+k=l} a_i (b_j c_k)\right) t^l,\end{aligned}$$

og dermed følger (P1) for polynomier fra (P1) for skalarer.

$$\begin{aligned}\sum_{i \geq 0} a_i t^i \cdot \left(\sum_{j \geq 0} b_j t^j + \sum_{j \geq 0} c_j t^j\right) &= \sum_{k \geq 0} \left(\sum_{i+j=k} a_i (b_j + c_j)\right) t^k, \\ \left(\sum_{i \geq 0} a_i t^i \cdot \sum_{j \geq 0} b_j t^j\right) + \left(\sum_{i \geq 0} a_i t^i \cdot \sum_{j \geq 0} c_j t^j\right) &= \sum_{k \geq 0} \left(\sum_{i+j=k} (a_i b_j + a_i c_j)\right) t^k,\end{aligned}$$

således at (D1) for polynomier følger fra (D1) for skalarer. □

Hvis d er et helt tal, så siger vi, at et polynomium

$$p = p(t) = \sum_{i \geq 0} a_i t^i \in R[t]$$

har grad højst d og skriver $\deg(p) \leq d$, hvis $a_i = 0$ for alle $i > d$. Vi skriver også

$$R[t]_{\leq n} = \{p \in R[t] \mid \deg(p) \leq d\} \subset R[t]$$

for delmængden af polynomier af grad højst $d \geq 0$. Specielt er nulpolynomiet det eneste polynomium, der har grad højst d for alle d , mens de konstante polynomier præcis er de polynomier, der har grad højst 0.

Lemma 3.5.5 *Lad $p, q \in R[t]$ være polynomier med koefficienter i en kommutativ ring R . Da gælder følgende udsagn:*

(1) *Hvis $\deg(p) \leq d$ og $\deg(q) \leq e$, da er $\deg(p + q) \leq \max\{d, e\}$.*

(2) *Hvis $\deg(p) \leq d$ og $\deg(q) \leq e$, da er $\deg(p \cdot q) \leq d + e$.*

Bevis Både (1) og (2) følger direkte fra definitionerne. $\square$

Hvis $p(t) = a_d t^d + \cdots + a_1 t^1 + a_0 t^0$ med $a_d \neq 0$, da siger vi, at $p(t)$ har *grad* d , og vi skriver da $\deg(p(t)) = d$. Med andre ord er graden $p(t)$ det mindste hele tal d , sådan at $\deg(p(t)) \leq d$. Vi bemærker, at graden af nulpolynomiet ikke er defineret.

Eksempel 3.5.6 Polynomierne $p(t) = t^2 + 2t - 1$ og $q(t) = -t^2 + 2$ har begge grad 2. Deres produkt $p(t) \cdot q(t) = -t^4 - 2t^3 + 3t^2 + 4t - 2$ har grad 4, mens deres sum $p(t) + q(t) = 2t + 1$ har grad 1.

Generelt er udsagnene i lemma 3.5.5 det bedste, man kan sige om graden af sum og produkt af polynomier. For generelt kan R indeholde nuldivisorer, hvilke er elementer $a, b \in R$, sådan at $a \neq 0$ og $b \neq 0$, men $a \cdot b = 0$. Men hvis R ikke indeholder nuldivisorer, da gælder det mere præcise udsagn, at $\deg(p \cdot q) = \deg(p) + \deg(q)$ for alle $p, q \in R[t]$, der begge er forskellige fra nulpolynomiet.

Vi betragter nu determinantafbildningen

$$\det: M_n(R[t]) \rightarrow R[t]$$

for $n \times n$ -matricer med indgange i $R[t]$. Denne findes ifølge sætning 3.2.1, der som nævnt gælder for enhver kommutativ ring.

Sætning 3.5.7 *Lad $R[t]$ være den kommutative ring af polynomier med koefficienter i en kommutativ ring R , og lad $P = (p_{ij}) \in M_n(R[t])$ være en matrix med indgange i $R[t]$. Hvis $\deg(p_{ij}) \leq d$ for alle $1 \leq i, j \leq n$, da er*

$$\deg(\det(P)) \leq dn.$$

Bevis Beviset er ved induktion på $n \geq 0$, og tilfældet $n = 0$ er trivielt. (Det er tilfældet $n = 1$ naturligvis også.) Så vi antager, at sætningen gælder for $n = r - 1$ og viser, at den gælder for $n = r$. Ved Laplace-udvikling langs første række får vi, at

$$\det(P) = \sum_{j=1}^r (-1)^{1+j} p_{1j} \det(P_{1j}).$$

3 Determinant

Her er $\deg(p_{1j}) \leq d$ per antagelse, og $\deg(\det(P_{1j})) \leq d(r-1)$ per induktion, og derfor viser lemma 3.5.5, at $\deg(\det(P)) \leq d + d(r-1) = dr$ som ønsket. $\square$

Eksempel 3.5.8 Ved hjælp af sætning 3.1.1 udregner vi, at

$$\det \begin{pmatrix} 5-t & 2 \\ 1 & 3-t \end{pmatrix} = (5-t)(3-t) - 2 = t^2 - 8t + 13.$$

Vi ser, at indgangene i matricen er polynomier af grad højst 1, mens determinanten er et polynomium af grad højst $1 \cdot 2 = 2$.

Vi forklarer nu, hvordan den variable “ t ” kan tilordnes en værdi, dvs. hvordan man kan “evaluere” et polynomium på et element. En afbildning

$$f: R \rightarrow S$$

mellem to kommutative ringe siges at være en *ringhomomorfi*, hvis den opfylder:

(R1) For alle $a, b \in R$ er $f(a+b) = f(a) + f(b)$.

(R2) For alle $a, b \in R$ er $f(a \cdot b) = f(a) \cdot f(b)$.

(R3) $f(1_R) = 1_S$.

En ringhomomorfi $f: R \rightarrow S$ og et element $s \in S$ giver anledning til en ringhomomorfi

$$\text{ev}_{f,s}: R[t] \rightarrow S,$$

der er defineret ved formelen

$$\text{ev}_{f,s}(\sum_{i \geq 0} a_i t^i) = \sum_{i \geq 0} f(a_i) s^i.$$

Bemærk, at sum og produkt af polynomier netop er defineret, sådan at denne afbildning er en ringhomomorfi. Hvis $f: R \rightarrow S$ er underforstået, så skriver vi også $p(s)$ i stedet for $\text{ev}_{f,s}(p)$ og siger, at $p(s)$ fremkommer fra $p(t)$ ved at *substituere* $s \in S$ for t .

Eksempel 3.5.9 (1) Lad $f: R \rightarrow S$ være identitetsafbildningen $\text{id}: \mathbb{R} \rightarrow \mathbb{R}$ af de reelle tal. Givet $p(t) \in \mathbb{R}[t]$, kan vi da substituere $a \in \mathbb{R}$ for t , hvilket giver $p(a) = \text{ev}_a(p) \in \mathbb{R}$. Hvis for eksempel $p(t) = 3t^2 + t - 4$, og $a = 5$, så er $p(5) = \text{ev}_5(p) = 3 \cdot 5^2 + 5 - 4 = 76$.

(2) Lad $f: R \rightarrow S$ være den kanoniske inklusion $f: \mathbb{R} \rightarrow \mathbb{C}$ af de reelle tal i de komplekse tal. Givet $p(t) \in \mathbb{R}[t]$, kan vi da substituere $a \in \mathbb{C}$ for t , hvilket giver $p(a) = \text{ev}_a(p) \in \mathbb{C}$. Hvis $p(t) = 3t^2 + t - 4$, og $a = 1 + i \in \mathbb{C}$, så er $p(1+i) = 3(1+i)^2 + (1+i) - 4 = -3 + 7i$.

(3) Lad $R = \mathbb{R}$ og lad $S = C^0(\mathbb{R})$ være den kommutative ring af kontinuerte funktioner fra eksempel 3.5.2. Vi lader $f: \mathbb{R} \rightarrow C^0(\mathbb{R})$ være afbildningen, der til $c \in \mathbb{R}$ tilordner den konstante funktion $f(c): \mathbb{R} \rightarrow \mathbb{R}$ med værdi c , og bemærker, at identitetsafbildningen $\text{id}: \mathbb{R} \rightarrow \mathbb{R}$ er et element i $S = C^0(\mathbb{R})$. Vi skriver “ x ” for dette element, hvilket giver

$$\text{ev}_{f,x}: \mathbb{R}[t] \rightarrow C^0(\mathbb{R}),$$

der er en ringhomomorfi og afbilder $p(t)$ til den kontinuerte funktion $p(x)$, der fås ved at substituere $x \in S$ for t . Hvis $p(t) = 3t^2 + t - 4 \in \mathbb{R}[t]$, så er $p(x) \in C^0(\mathbb{R})$ den kontinuerte funktion givet ved $p(x) = 3x^2 + x - 4$.

Vi siger, at $s \in S$ er en *rod* i polynomiet $p(t) \in R[t]$, hvis $p(s) = 0$. For eksempel har polynomiet $p(t) = t^2 + 1 \in \mathbb{R}[t]$ ingen rødder i $\mathbb{R}$, mens $\pm i$ er rødder i $\mathbb{C}$. Den følgende vigtige sætning kaldes for algebraens fundamentalsætning. Der gik mere end 250 år, fra Cardano først indførte komplekse tal i 1545, indtil Gauss i 1799 og Argand i 1806 beviste denne sætning. Undervejs var der mange forsøg på både at bevise og modbevise sætningen.

Sætning 3.5.10 (Gauss-Argand) *Ethvert ikke-konstant polynomium med koefficienter i $\mathbb{C}$ har mindst en rod i $\mathbb{C}$.*

Bevis Vi henviser til en af de følgende lærebøger for et bevis.

- (1) David S. Dummit og Richard M. Foote. Abstract algebra. Third Edition. John Wiley & Sons, Inc., Hoboken, NJ, 2004.
- (2) Serge Lang. Algebra. Revised third edition. Graduate Texts in Mathematics, 211. Springer-Verlag, New York, 2002.

De to bøger præsenterer det samme bevis. I den først bog, er beviset givet i Theorem 35 på side 615–617; og i den anden bog, er beviset givet i Example 5 på side 272–273. $\square$

Man viser da ved polynomiumsdivision, at ethvert polynomium $p(t) \in \mathbb{C}[t]$ af grad præcis $d \geq 0$ har præcis d rødder i $\mathbb{C}$, talt med multiplicitet.

Sætning 3.5.11 *Lad $p(t) = a_d t^d + a_{d-1} t^{d-1} + \cdots + a_1 t + a_0 \in \mathbb{C}[t]$ være et polynomium med koefficienter i $\mathbb{C}$, sådan at $d \geq 0$ og $a_d \neq 0$. Da findes $z_1, \dots, z_d \in \mathbb{C}$, sådan at*

$$p(t) = a_d(t - z_1) \cdots (t - z_d).$$

3 Determinant

Bevis Påstanden følger fra sætning 3.5.10 ved brug af den euklidiske algoritme. Denne er bevist i Theorem 1.1 i Lang's bog (2) på side 173–174. $\square$

Eksempel 3.5.12 Vi anvender formlen for rødderne i et andengradspolynomium til at faktorisere polynomiet $p(t) = t^2 - 4t + 13$. Diskriminanten er

$$\Delta = (-4)^2 - 4 \cdot 1 \cdot 13 = (-4) \cdot (-4 + 13) = -4 \cdot 9 = (6i)^2,$$

så de to rødder er $t = \frac{1}{2}(-(-4) \pm 6i) = 2 \pm 3i$. Den euklidiske algoritme viser da, at

$$p(t) = (t - (2 + 3i))(t - (2 - 3i)).$$

Vi bemærker, at de to rødder er konjugerede, fordi $p(t)$ har reelle koefficienter.

Eksempel 3.5.13 Vi illustrerer Euklids algoritme. Polynomiet

$$p(t) = t^3 - 6t^2 + 5t + 12$$

har $a = -1$ som rod, og den euklidiske algoritme viser da, at der findes et entydigt bestemt polynomium $q(t)$, sådan at $p(t) = (t + 1)q(t)$. Vi gennemfører divisionen

$$\begin{array}{r} t^2 - 7t + 12 \\ t+1 \overline{) t^3 - 6t^2 + 5t + 12} \\ \underline{t^3 + t^2} \\ -7t^2 \\ \underline{-7t^2 - 7t} \\ 12t \\ \underline{12t + 12} \\ 0 \end{array}$$

og ser, at $t + 1$ går op i $p(t)$ med kvotient $q(t) = t^2 - 7t + 12$. Vi kan yderligere faktorisere $q(t)$ ved hjælp af formlen for rødderne i et andengradspolynomium og får derved, at

$$p(t) = (t + 1)(t - 3)(t - 4).$$

Vi bemærker, at beviset for algebraens fundamentalsætning ikke er konstruktivt, så sætningen siger altså blot, at der findes en rod, men den siger ikke noget om, hvordan man bærer sig ad med at finde en rod. For et generelt polynomium $p(t) \in \mathbb{C}[t]$, kan man kun undtagelsesvis finde $z \in \mathbb{C}$, der opfylder $p(z) = 0$ eksakt. For ligegyldigt hvor mange

cifre af $p(z)$, man udregner, kan man generelt ikke vide, at der ikke vil dukke ikke-nul cifre op i $p(z)$, hvis man regner længere. Man skal derfor nøje vurdere den fejl, man begår ved at lade som om ligheden $p(z) = 0$ gælder eksakt.

3.6 Opgaver

3.1 Udregn determinanten af følgende matricer:

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & -2 & 3 \\ 4 & 3 & 1 \\ 2 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 1 \\ 5 & \pi & 5 \\ 2 & 1/2 & 2 \end{pmatrix}.$$

3.2 Bestem determinanterne af følgende matricer:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad C = \begin{pmatrix} 1+i & 2 \\ 1 & 1-i \end{pmatrix}.$$

3.3 Bestem determinanterne af følgende matricer:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 0 & 4 \\ 4 & 0 & 0 & 8 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 2 & 4 & 8 \end{pmatrix}.$$

3.4 Afgør om følgende udsagn er sande eller falske.

a)

$$\det \begin{pmatrix} 8 & 8 & 8 \\ 2 & 3 & 2 \\ 5 & 6 & 7 \end{pmatrix} = 8 \cdot \det \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 5 & 6 & 7 \end{pmatrix}$$

b)

$$\det \begin{pmatrix} 0 & 0 & c \\ 0 & b & e \\ a & d & f \end{pmatrix} = -abc.$$

c)

$$\det \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix} = abcd.$$

3 Determinant

d)

$$\det \begin{pmatrix} 0 & 0 & 0 & a \\ 0 & 0 & b & 0 \\ 0 & c & 0 & 0 \\ d & 0 & 0 & 0 \end{pmatrix} = -abcd.$$

3.5 Udregn determinanten af den følgende matrix:

$$A = \begin{pmatrix} a & 0 & 0 & 0 & 1 \\ 0 & a & 0 & 1 & 0 \\ 0 & 0 & b & 0 & 0 \\ 0 & 1 & 0 & a & 0 \\ 1 & 0 & 0 & 0 & a \end{pmatrix}$$

3.6 Udregn determinanter af de følgende matricer:

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

3.7 (★) Udregn determinanten af matricen

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Beskriv afbildningen $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, der er defineret ved $f(\mathbf{x}) = A\mathbf{x}$, geometrisk.

3.8 Bestem determinanten af matricerne

$$A = \begin{pmatrix} 2+i & 4 \\ 3i & 1-i \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 2 & 3-4i & -2+i \\ 0 & 1+i & 0 \\ 3i & 3+4i & 1-i \end{pmatrix}.$$

3.9 (★) For alle $n \geq 0$ betragter vi den lineære isomorfi $f_n: \mathbb{R}^n \rightarrow \mathbb{R}^n$ defineret ved

$$f_n \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_n \\ \vdots \\ x_2 \\ x_1 \end{pmatrix}.$$

a) Angiv for alle $n \geq 0$ matricen A_n , der repræsenterer $f_n: \mathbb{R}^n \rightarrow \mathbb{R}^n$ med hensyn til standardbasen for både domænet og codomænet.

- b) Beregn determinanten af A_n for alle $0 \leq n \leq 9$, gerne ved brug af Maple, og opstil på baggrund heraf en hypotese for værdien af $\det(A_n)$ for alle $n \geq 0$.
- c) Bevis den hypotese, der blev opstillet i (b).
- d) En lineær isomorfi $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ siges at være *orienteringsbevarende*, hvis determinanten af matricen A , der repræsenterer afbildningen med hensyn til standardbasen for både domænet og codomænet, er positiv. For hvilke $n \geq 0$ er afbildningen $f_n: \mathbb{R}^n \rightarrow \mathbb{R}^n$ orienteringsbevarende?

3.10 De beregninger, der skal udføres for at udregne determinanten af en matrix, kan opdeles i tre typer, nemlig multiplikation; addition; og fortegnsskift. For at forstå effektiviteten af beregning ved Laplace-udvikling, vil vi udelukkende se på antallet af multiplikationer.

- a) Redegør for, at der ved beregning af determinanten af en generel $n \times n$ -matrix ved Laplace-udvikling (sætning 3.2.8) skal udføres $n!$ multiplikationer.
- b) Udfør følgende kommandoer i Maple:

```
>with(LinearAlgebra):
>detTime:= [seq([n,time(Determinant(RandomMatrix(n,n)))] ,n=1..100)]:
>plot(detTime);
```

og forklar (eventuelt ved at slå kommandoerne op med ?time osv.) hvilket eksperiment, der herved er udført.

- c) Redegør for, at (a) og (b) viser, at Maple ikke benytter Laplace-udvikling til beregning af determinanter.

3.11 Besvar følgende spørgsmål:

- a) Hvad kan man sige om determinanten af en matrix, hvori to rækker er ens?
- b) Hvordan ændres værdien af determinanten, hvis man skifter fortegn på alle indgange i en række?
- c) Hvordan ændres værdien af determinanten, hvis man skifter fortegn på alle indgange i matricen?
- d) Er det rigtigt, at hvis indgange i en matrix er forskellige fra nul, da er også determinanten forskellig fra nul?

3.12 Vi lader t være et reelt tal og betragter matricen

$$A = \begin{pmatrix} 1 & t & 0 \\ t & 1 & t \\ 0 & t & 1 \end{pmatrix}.$$

- a) Beregn $\det(A)$.
- b) Gør rede for, at A er invertibel for $t \neq \pm\sqrt{1/2}$.

3 Determinant

c) Anvend for $t \neq \pm\sqrt{1/2}$ Cramers regel til at løse ligningssystemet $A\mathbf{x} = \mathbf{b}$, hvor

$$\mathbf{b} = \begin{pmatrix} t \\ 0 \\ t \end{pmatrix}.$$

3.13 Vi betragter matricerne

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 2 & 4 & 5 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 4 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 \\ 6 & 7 & 2 & 0 \\ 8 & 9 & 0 & 1 \end{pmatrix}.$$

a) Udregn $\det(A)$ og $\det(B)$.

b) Udregn $\det(AB)$.

c) Udregn $\det(AB^{-1})$.

3.14 For ethvert reelt tal c betragter vi matricen:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & c+1 & 1 \\ 1 & 2 & c+1 \end{pmatrix}.$$

Afgør for hvilke værdier af c , matricen A er invertibel, og bestem for disse værdier af c den inverse matrix A^{-1} .

3.15 Omform ved hjælp af rækkeoperationer matricen

$$A = \begin{pmatrix} 1 & 0 & -a & 0 \\ 0 & 1 & 0 & 2 \\ -1 & 0 & 1 & 0 \\ 0 & 1+a & 0 & 1 \end{pmatrix}$$

til en øvre triangulær matrix og afgør herved, for hvilke værdier af a den givne matrix er invertibel.

3.16 Afgør ved hjælp af determinanten, om de følgende ligningssystemer har en entydig løsning:

$$\begin{array}{lcl} x_1 + 2x_2 - x_3 & = & 5 \\ \text{a)} \quad 3x_1 + 2x_3 & = & -4 \\ x_1 + 3x_2 + x_3 & = & 19 \end{array}$$

$$\begin{array}{lcl} x_1 + 2x_2 - x_3 & = & 51 \\ \text{b)} \quad 3x_1 + x_2 + 2x_3 & = & 7 \\ -2x_1 - 4x_2 + 2x_3 & = & 25 \end{array}$$

- 3.17 Vis, at hvis alle indgange i en kvadratisk matrix A er hele tal, og hvis $\det(A) = \pm 1$, så er alle indgange i A^{-1} ligeledes hele tal. [Vink: brug sætning 3.4.1.]
- 3.18 (★) Lad A være en $n \times n$ -matrix og antag, at ligningssystemet $A\mathbf{x} = \mathbf{0}$ har $\mathbf{0}$ som eneste løsning. Vis, at ligningssystemet $A^k\mathbf{x} = \mathbf{0}$ ligeledes har $\mathbf{0}$ som eneste løsning.
- 3.19 Den følgende opgave viser, at en ringhomomorfi mellem to legemer altid er injektiv. Vi definerer *kernen* af en ringhomomorfi $\varphi: R \rightarrow S$ til at være delmængden

$$\ker(\varphi) = \{r \in R \mid \varphi(r) = 0\}.$$

- a) Vis, at $\ker(\varphi) = \{0\}$, hvis og kun hvis φ er injektiv.
- b) Vis, at kernen $\ker(\varphi)$ for en ringhomomorfi opfylder, at hvis $u \in \ker(\varphi)$ og $r \in R$, da er også $u \cdot r \in \ker(\varphi)$.
- c) Vis, at en ringhomomorfi $\varphi: \mathbb{F} \rightarrow \mathbb{F}'$ mellem to legemer altid er injektiv.
[Vink: Antag, at $\ker(\varphi) \neq \{0\}$ og anvend (b) til at vise, at $1 \in \ker(\varphi)$.]
- 3.20 Find to 2×2 -matricer A og B så at $\det(A+B) \neq \det(A) + \det(B)$.
- 3.21 (★) Lad $\mathbb{F}$ være et legeme, og lad $V = \mathbb{F}^n$. Lad $k \geq 0$, og lad V^k være mængden af k -tupler af vektorer i V . Givet k lineære afbildninger

$$f_1, \dots, f_k: V \rightarrow \mathbb{F},$$

betrakter vi afbildningen $\omega: V^k \rightarrow \mathbb{F}$ defineret ved

$$\omega(\mathbf{v}_1, \dots, \mathbf{v}_k) = \det \begin{pmatrix} f_1(\mathbf{v}_1) & f_1(\mathbf{v}_2) & \dots & f_1(\mathbf{v}_k) \\ f_2(\mathbf{v}_1) & f_2(\mathbf{v}_2) & \dots & f_2(\mathbf{v}_k) \\ \vdots & \vdots & \ddots & \vdots \\ f_k(\mathbf{v}_1) & f_k(\mathbf{v}_2) & \dots & f_k(\mathbf{v}_k) \end{pmatrix}.$$

Afbildningen er et eksempel på en alternerende form, og opgaven går ud på at eftertervisse de følgende egenskaber.

- (a) Vis, at afbildningen $\omega: V^k \rightarrow \mathbb{F}$ er antisymmetrisk i den forstand, at det for alle $1 \leq i < j \leq k$ gælder, at

$$\omega(\dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots) = -\omega(\dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots).$$

- (b) Vis, at hvis der findes $1 \leq i < j \leq k$, sådan at $f_i = f_j$, da er $\omega: V^k \rightarrow \mathbb{F}$ lig med nulafbildningen.
- (c) Vis, at hvis $\mathbf{v}_1 = \mathbf{v}'_1 + \mathbf{v}''_1$, så er

$$\omega(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \omega(\mathbf{v}'_1, \mathbf{v}_2, \dots, \mathbf{v}_k) + \omega(\mathbf{v}''_1, \mathbf{v}_2, \dots, \mathbf{v}_k).$$

4 Vektorrum

Lad A være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$. I kapitel 1 har vi beskrevet en algoritme, der producerer en parametrisering af løsningsmængden

$$N_A = \{\mathbf{x} \in \mathbb{F}^n \mid A\mathbf{x} = \mathbf{0}\} \subset \mathbb{F}^n$$

til det homogene lineære ligningssystem $A\mathbf{x} = \mathbf{0}$; se sætning 1.2.17. Denne algoritme afhænger dog af nogle temmeligt arbitrære valg såsom definitionen af en matrix på reduceret echelonform. I modsætning til det afhænger løsningsmængden N_A selv ikke af nogle valg, og vi ønsker derfor at udvikle vores begrebsverden, sådan at vi direkte kan udtrykke, hvad denne løsningsmængde “er” for en størrelse. Hertil indfører vi begrebet *vektorrum*, som vi allerede har stiftet bekendtskab med i afsnit 2.2.

4.1 Vektorrum

Vi lader $\mathbb{F}$ være et legeme, og som før tænker vi på $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. Hele dette kapitel virker dog ligeså godt for skævlegemer, så $\mathbb{F} = \mathbb{H}$ er også tilladt. Et (abstrakt) vektorrum består af en mængde, hvis elementer vi kalder vektorer, udstyret med to operationer “+” og “ $\cdot$ ”, som vi kalder vektorsum og skalarmultiplikation.

Definition 4.1.1 Lad $\mathbb{F}$ være et legeme. Et *højre $\mathbb{F}$ -vektorrum* er en triple $(V, +, \cdot)$, der består af en mængde V og to afbildninger $+: V \times V \rightarrow V$ og $\cdot: V \times \mathbb{F} \rightarrow V$, sådan at:

- (A1) For alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ er $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.
- (A2) Der findes et element $\mathbf{0} \in V$, sådan at $\mathbf{x} + \mathbf{0} = \mathbf{x} = \mathbf{0} + \mathbf{x}$ for alle $\mathbf{x} \in V$.
- (A3) For alle $\mathbf{x} \in V$, findes $\mathbf{y} \in V$, sådan at $\mathbf{x} + \mathbf{y} = \mathbf{0} = \mathbf{y} + \mathbf{x}$.
- (A4) For alle $\mathbf{x}, \mathbf{y} \in V$ er $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
- (V1) For alle $\mathbf{x} \in V$ og $a, b \in \mathbb{F}$ er $(\mathbf{x} \cdot a) \cdot b = \mathbf{x} \cdot (ab)$.
- (V2) For alle $\mathbf{x}, \mathbf{y} \in V$ og $a \in \mathbb{F}$ er $(\mathbf{x} + \mathbf{y}) \cdot a = (\mathbf{x} \cdot a) + (\mathbf{y} \cdot a)$.
- (V3) For alle $\mathbf{x} \in V$ og $a, b \in \mathbb{F}$ er $\mathbf{x} \cdot (a + b) = (\mathbf{x} \cdot a) + (\mathbf{x} \cdot b)$.
- (V4) For alle $\mathbf{x} \in V$ er $\mathbf{x} \cdot 1 = \mathbf{x}$.

4 Vektorrum

Man definerer *venstre* $\mathbb{F}$ -vektorrum tilsvarende; den eneste forskel er, at vektorerne kan multipliceres med skalarer fra venstre i stedet for fra højre.

Bemærkning 4.1.2 Lad $\mathbb{F}$ være et legeme og lad $(V, +, \cdot)$ være et $\mathbb{F}$ -vektorrum.

(1) Vektoren $\mathbf{0} \in V$ i (A2) er entydigt bestemt. For hvis $\mathbf{0}$ og $\mathbf{0}'$ begge opfylder (A2), da viser (A2) for henholdsvis $\mathbf{0}$ og $\mathbf{0}'$, at $\mathbf{0}' = \mathbf{0} + \mathbf{0}' = \mathbf{0}$. Vi kalder $\mathbf{0} \in V$ for *nulvektoren*.

(2) Givet $\mathbf{x} \in V$, da er vektoren $\mathbf{y} \in V$ i (A3) entydigt bestemt. For hvis $\mathbf{y}$ og $\mathbf{y}'$ begge opfylder (A3), så viser (A3) for henholdsvis $\mathbf{y}$ og $\mathbf{y}'$ samt (A1) og (A2), at

$$\mathbf{y} = \mathbf{y} + \mathbf{0} = \mathbf{y} + (\mathbf{x} + \mathbf{y}') = (\mathbf{y} + \mathbf{x}) + \mathbf{y}' = \mathbf{0} + \mathbf{y}' = \mathbf{y}'.$$

Vi skriver $-\mathbf{x}$ for vektoren $\mathbf{y}$ og kalder den for den *modsatte vektor* af $\mathbf{x}$.

(3) Vi definerer *differencen* af to vektorer $\mathbf{x}$ og $\mathbf{y}$ til at være vektoren $\mathbf{x} - \mathbf{y} = \mathbf{x} + (-\mathbf{y})$.

(4) Man viser som i sætning 0.2.3, at $\mathbf{x} \cdot 0 = \mathbf{0}$ og $\mathbf{x} \cdot (-1) = -\mathbf{x}$.

Der findes rigtig mange vektorrum inden for matematik. Vi giver nogle eksempler.

Eksempel 4.1.3 (1) Mængden $M_{m,1}(\mathbb{F})$ af $m \times 1$ -matricer med indgange i $\mathbb{F}$ sammen med afbildningerne $+: M_{m,1}(\mathbb{F}) \times M_{m,1}(\mathbb{F}) \rightarrow M_{m,1}(\mathbb{F})$ og $\cdot: M_{m,1}(\mathbb{F}) \times M_{1,1}(\mathbb{F}) \rightarrow M_{m,1}(\mathbb{F})$, der er givet ved matrixsum og matrixprodukt, udgør et højre $\mathbb{F}$ -vektorrum. Her har vi identificeret $M_{1,1}(\mathbb{F})$ med $\mathbb{F}$. Vi skriver $(\mathbb{F}^m, +, \cdot)$ for dette højre $\mathbb{F}$ -vektorrum og kalder det for $\mathbb{F}$ -vektorrummet af søjlevektorer af dimension m . Vi har allerede betragtet dette vektorrum i afsnit 2.2.

(2) Mængden $M_{1,n}(\mathbb{F})$ af $1 \times n$ -matricer med indgange i $\mathbb{F}$ sammen med afbildningerne $+: M_{1,n}(\mathbb{F}) \times M_{1,n}(\mathbb{F}) \rightarrow M_{1,n}(\mathbb{F})$ og $\cdot: M_{1,1}(\mathbb{F}) \times M_{1,n}(\mathbb{F}) \rightarrow M_{1,n}(\mathbb{F})$ givet ved matrixsum og matrixprodukt udgør et venstre $\mathbb{F}$ -vektorrum, hvor vi identificerer $M_{1,1}(\mathbb{F})$ med $\mathbb{F}$. Vi indfører ikke anden notation for dette vektorrum.

(3) Mængden $M_{m,n}(\mathbb{F})$ af $m \times n$ -matricer med indgange i $\mathbb{F}$ sammen med afbildningerne $+: M_{m,n}(\mathbb{F}) \times M_{m,n}(\mathbb{F}) \rightarrow M_{m,n}(\mathbb{F})$ og $\cdot: M_{m,n}(\mathbb{F}) \times \mathbb{F} \rightarrow M_{m,n}(\mathbb{F})$, hvor den første er matrixsum, og hvor den anden afbilder (A, a) med $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$ og $a \in \mathbb{F}$ til

$$A \cdot a = (a_{ij} \cdot a) \in M_{m,n}(\mathbb{F}),$$

er et højre vektorrum. Vi bemærker, at skalarmultiplikationen i dette vektorrum er ikke givet ved matrixmultiplikation.

(4) Hvis triplen $(\mathbb{F}, +, \cdot)$ er et legeme, da er den samme triple både et højre og et venstre vektorrum. For aksiomerne for et legeme i Definition 0.2.1 medfører aksiomerne for et vektorrum i Definition 4.1.1.

(5) Vi har i (4) givet mængden $\mathbb{C}$ af komplekse tal en struktur af højre $\mathbb{C}$ -vektorrum. Vi kan også give denne en struktur af højre $\mathbb{R}$ -vektorrum med vektorsum $+: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ og skalarmultiplikation $\cdot: \mathbb{C} \times \mathbb{R} \rightarrow \mathbb{C}$ defineret ved

$$\begin{aligned}(x_1 + ix_2) + (y_1 + iy_2) &= (x_1 + y_1) + i(x_2 + y_2), \\ (x_1 + ix_2) \cdot a &= x_1 a + ix_2 a.\end{aligned}$$

(6) Mængden $\mathbb{C}$ af komplekse tal kan ligeledes gives en struktur af højre $\mathbb{Q}$ -vektorrum med vektorrum $+: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ og skalarmultiplikation $\cdot: \mathbb{C} \times \mathbb{Q} \rightarrow \mathbb{C}$ defineret ved

$$\begin{aligned}(x_1 + ix_2) + (y_1 + iy_2) &= (x_1 + y_1) + i(x_2 + y_2), \\ (x_1 + ix_2) \cdot a &= x_1 a + ix_2 a.\end{aligned}$$

(7) Mængden $V = \{|\uparrow\rangle z_1 + |\downarrow\rangle z_2 \mid z_1, z_2 \in \mathbb{C}\}$ af formelle komplekse linearkombinationer af symbolerne $|\uparrow\rangle$ og $|\downarrow\rangle$, som læses “spin op” og “spin ned”, sammen med afbildningerne $+: V \times V \rightarrow V$ og $\cdot: V \times \mathbb{C} \rightarrow V$ defineret ved

$$\begin{aligned}(|\uparrow\rangle z_1 + |\downarrow\rangle z_2) + (|\uparrow\rangle w_1 + |\downarrow\rangle w_2) &= |\uparrow\rangle (z_1 + w_1) + |\downarrow\rangle (z_2 + w_2) \\ (|\uparrow\rangle z_1 + |\downarrow\rangle z_2) \cdot w &= |\uparrow\rangle z_1 w + |\downarrow\rangle z_2 w\end{aligned}$$

er et højre $\mathbb{C}$ -vektorrum. Vektorerne i dette vektorrum repræsenterer spintilstandene af en enkelt partikel.

(8) Mængden $\mathbb{F}[t]_{\leq d}$ af polynomier med koefficienter i $\mathbb{F}$ og af grad højst d sammen med afbildningerne $+: \mathbb{F}[t]_{\leq d} \times \mathbb{F}[t]_{\leq d} \rightarrow \mathbb{F}[t]_{\leq d}$ og $\cdot: \mathbb{F}[t]_{\leq d} \times \mathbb{F}[t]_{\leq 0} \rightarrow \mathbb{F}[t]_{\leq d}$ defineret til at være henholdsvis sum og produkt af polynomier udgør et højre $\mathbb{F}$ -vektorrum, idet vi identificerer $\mathbb{F}[t]_{\leq 0}$ med $\mathbb{F}$.

(9) Mængden $C^0(\mathbb{R})$ af kontinuerte funktioner $f: \mathbb{R} \rightarrow \mathbb{R}$ sammen med afbildningerne $+: C^0(\mathbb{R}) \times C^0(\mathbb{R}) \rightarrow C^0(\mathbb{R})$ og $\cdot: C^0(\mathbb{R}) \times \mathbb{R} \rightarrow C^0(\mathbb{R})$ defineret ved

$$(f + g)(x) = f(x) + g(x) \quad \text{og} \quad (f \cdot a)(x) = f(x) \cdot a$$

udgør et højre $\mathbb{R}$ -vektorrum.

Vi vil i denne bog altid betragte højre $\mathbb{F}$ -vektorrum, og vi vil derfor forkorte og sige $\mathbb{F}$ -vektorrum i stedet for højre $\mathbb{F}$ -vektorrum. Hvis $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, så siger vi også, at et $\mathbb{F}$ -vektorrum er henholdsvis et reelt vektorrum og et komplekst vektorrum. Hvis $\mathbb{F}$ er underforstået, så vil vi sommetider sige vektorrum i stedet for $\mathbb{F}$ -vektorrum. Desuden vil vi, som det er sædvane, misbruge notation og skrive V for vektorrummet $(V, +, \cdot)$.

Definition 4.1.4 Lad $\mathbb{F}$ være et legeme, og lad $(V, +, \cdot)$ være et $\mathbb{F}$ -vektorrum. En delmængde $U \subset V$ siges at være *stabil med hensyn til vektorrumssstrukturen på V* , hvis følgende er opfyldt:

- (1) Nulvektoren $\mathbf{0} \in V$ tilhører $U \subset V$.
- (2) Hvis $\mathbf{x}, \mathbf{y} \in U \subset V$, da er også $\mathbf{x} + \mathbf{y} \in U \subset V$.
- (3) Hvis $\mathbf{x} \in U \subset V$ og $a \in \mathbb{F}$, da er også $\mathbf{x} \cdot a \in U \subset V$.

4 Vektorrum

Vi bemærker, at (1) medfører, at $U \subset V$ er ikke den tomme delmængde. Omvendt, hvis $U \subset V$ ikke er den tomme delmængde og opfylder (3), da opfylder $U \subset V$ også (1). For vi kan da vælge $\mathbf{u} \in U$, og (3) viser dermed, at $\mathbf{0} = \mathbf{u} \cdot 0 \in U$.

Eksempel 4.1.5 (1) Vi påstår, at delmængden

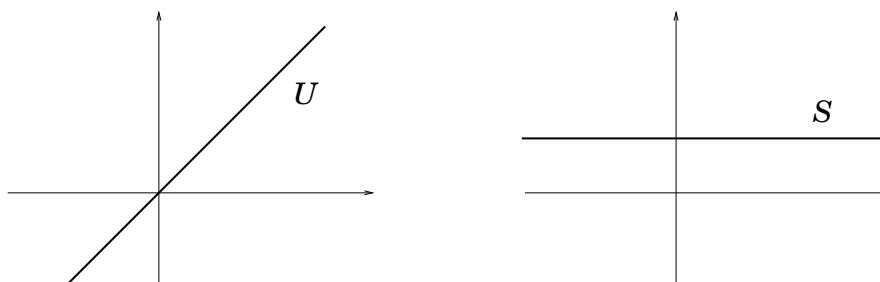
$$U = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 = x_2 \right\} \subset V = \mathbb{R}^2$$

er stabil med hensyn til vektorrumsstrukturen på $\mathbb{R}^2$. For $\mathbf{0} \in U$, da $0 = 0$, så (1) gælder; hvis $\mathbf{x}, \mathbf{y} \in U$, så er også $\mathbf{x} + \mathbf{y} \in U$, fordi " $x_1 = x_2$ " og " $y_1 = y_2$ " medfører " $x_1 + y_1 = x_2 + y_2$ ", så (2) gælder; og hvis $\mathbf{x} \in U$ og $a \in \mathbb{R}$, da er $\mathbf{x} \cdot a \in U$, idet " $x_1 = x_2$ " medfører " $x_1 a = x_2 a$ ", så (3) gælder også.

(2) Vi påstår modsat, at delmængden

$$S = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_2 = 1 \right\} \subset V = \mathbb{R}^2$$

ikke er stabil med hensyn til vektorrumsstrukturen på $\mathbb{R}^2$. Vi påstår faktisk, at *ingen* af betingelserne (1)–(3) i definition 4.1.4 er opfyldt af $S \subset V$. For $\mathbf{0} \notin S$, da $0 \neq 1$; hvis $\mathbf{x}, \mathbf{y} \in S$, da er $\mathbf{x} + \mathbf{y} \notin S$, da $x_2 + y_2 = 1 + 1 \neq 1$; og hvis endelig $\mathbf{x} \in S$ og $1 \neq a \in \mathbb{R}$, da er $\mathbf{x} \cdot a \notin S$, idet $x_2 \cdot a = 1 \cdot a \neq 1$.



Figur 4.1: $U \subset \mathbb{R}^2$ er stabil under vektorrumsstrukturen på $\mathbb{R}^2$, mens $S \subset \mathbb{R}^2$ ikke er det.

Lad $U \subset V$ være en delmængde af et $\mathbb{F}$ -vektorrum $(V, +, \cdot)$, der er stabil med hensyn til vektorrumsstrukturen på V . Da giver vektorsummen og skalarmultiplikationen på V anledning til afbildningerne $+' : U \times U \rightarrow U$ og $\cdot' : U \times \mathbb{F} \rightarrow U$ givet ved

$$\mathbf{x} +' \mathbf{y} = \mathbf{x} + \mathbf{y} \quad \text{og} \quad \mathbf{x} \cdot' a = \mathbf{x} \cdot a,$$

hvor $\mathbf{x}, \mathbf{y} \in U$ og $a \in \mathbb{F}$. For (2) sikrer netop, at $\mathbf{x} + \mathbf{y} \in U$, mens (3) sikrer, at $\mathbf{x} \cdot a \in U$. Vi siger, at afbildningerne $+' : U \times U \rightarrow U$ og $\cdot' : U \times \mathbb{F} \rightarrow U$ fremkommer ved restriktion af

afbildningerne $+: V \times V \rightarrow V$ og $\cdot: V \times \mathbb{F} \rightarrow V$. Vi vil ofte misbruge notation og ligeledes skrive $+: U \times U \rightarrow U$ og $\cdot: U \times \mathbb{F} \rightarrow U$ for de restringerede afbildninger.

Sætning 4.1.6 *Lad $\mathbb{F}$ være et legeme, lad $(V, +, \cdot)$ være et $\mathbb{F}$ -vektorrum, lad $U \subset V$ være en delmængde, der er stabil med hensyn til vektorrumssstrukturen på V , og lad $+' : U \times U \rightarrow U$ og $\cdot' : U \times \mathbb{F} \rightarrow U$ være afbildningerne, der fremkommer ved restriktion af afbildningerne $+: V \times V \rightarrow V$ og $\cdot: V \times \mathbb{F} \rightarrow V$. Da er også $(U, +', \cdot')$ et $\mathbb{F}$ -vektorrum.*

Bevis Vi ønsker at vise, at de restringerede afbildninger $+'$ og $\cdot'$ opfylder (A1)–(A4) og (V1)–(V4) i definition 4.1.4. Da $+$ og $\cdot$ opfylder (A1)–(A4) og (V1)–(V4), følger det umiddelbart, at $+'$ og $\cdot'$ opfylder (A1), (A4) og (V1)–(V4), og vi viser nu, at (A2) og (A3) er opfyldt. Så lad $\mathbf{0} \in V$ være nulvektoren. Ifølge (1) er $\mathbf{0} \in U$, og for alle $\mathbf{x} \in U$ gælder det, at

$$\mathbf{x} +' \mathbf{0} = \mathbf{x} + \mathbf{0} = \mathbf{x} = \mathbf{0} + \mathbf{x} = \mathbf{0} +' \mathbf{x},$$

idet (A2) gælder i V . Dette viser, at også $(U, +', \cdot)$ opfylder (A2). Hvis endelig $\mathbf{x} \in U$, så viser (3), at også $\mathbf{x} \cdot (-1) \in U$, og udregningen

$$\mathbf{x} +' (\mathbf{x} \cdot (-1)) = \mathbf{x} + (\mathbf{x} \cdot (-1)) = \mathbf{x} + (-\mathbf{x}) = \mathbf{0} = (-\mathbf{x}) + \mathbf{x} = (\mathbf{x} \cdot (-1)) + \mathbf{x} = (\mathbf{x} \cdot (-1)) +' \mathbf{x}$$

viser derfor, at $(U, +', \cdot')$ opfylder (A3). $\square$

Definition 4.1.7 *Lad $\mathbb{F}$ være et legeme. Et $\mathbb{F}$ -vektorrum $(U, +', \cdot')$ siges at være et *underrum* af et $\mathbb{F}$ -vektorrum $(V, +, \cdot)$, hvis U er en delmængde af V , der er stabil med hensyn til vektorrumssstrukturen på V , og hvis $+' : U \times U \rightarrow U$ og $\cdot' : U \times \mathbb{F} \rightarrow U$ netop er afbildningerne, der fremkommer ved restriktion af $+: V \times V \rightarrow V$ og $\cdot: V \times \mathbb{F} \rightarrow V$.*

Vi vil herefter misbruge notation og blot skrive, at $U \subset V$ er et underrum i stedet for at skrive, at $U \subset V$ er stabil under vektorrumssstrukturen på V , og at $(U, +', \cdot')$ er et underrum af $(V, +, \cdot)$. Vi siger således, at i eksempel 4.1.5 er $U \subset \mathbb{R}^2$ et underrum, mens $S \subset \mathbb{R}^2$ ikke er det.

Lemma 4.1.8 *Lad A være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$. Delmængden*

$$N_A = \{\mathbf{x} \in \mathbb{F}^n \mid A\mathbf{x} = \mathbf{0}\} \subset \mathbb{F}^n,$$

der består af løsninger til det homogene ligningssystem “ $A\mathbf{x} = \mathbf{0}$ ”, er et underrum af $\mathbb{F}^n$. Tilsvarende er delmængden

$$R_A = \{A\mathbf{x} \in \mathbb{F}^m \mid \mathbf{x} \in \mathbb{F}^n\} \subset \mathbb{F}^m,$$

der består af de vektorer $\mathbf{b} \in \mathbb{F}^m$, for hvilke ligningen “ $A\mathbf{x} = \mathbf{b}$ ” har mindst én løsning, et underrum af $\mathbb{F}^m$.

4 Vektorrum

Bevis Vi viser, at $N_A \subset \mathbb{F}^n$ opfylder (1)–(3) i definition 4.1.4. Da $A\mathbf{0} = \mathbf{0}$, er (1) opfyldt; hvis $A\mathbf{x} = \mathbf{0}$ og $A\mathbf{y} = \mathbf{0}$, da er $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$, så (2) er opfyldt; og hvis $A\mathbf{x} = \mathbf{0}$, da er også $A(\mathbf{x}a) = (A\mathbf{x})a = \mathbf{0}a = \mathbf{0}$, så (3) er opfyldt. Dermed er $N_A \subset \mathbb{F}^n$ stabil med hensyn til vektorrumstrukturen på $\mathbb{F}^n$ og definerer derfor ifølge sætning 4.1.6 et underrum af $\mathbb{F}^n$.

Tilsvarende viser vi, at $R_A \subset \mathbb{F}^m$ opfylder (1)–(3) i definition 4.1.4. Da $A\mathbf{0} = \mathbf{0}$ er (1) opfyldt; hvis $A\mathbf{x} = \mathbf{b}$ og $A\mathbf{y} = \mathbf{c}$, da er $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{b} + \mathbf{c}$, så (2) er opfyldt; og hvis $A\mathbf{x} = \mathbf{b}$, så er $A(\mathbf{x}a) = (A\mathbf{x})a = \mathbf{b}a$, så også (3) er opfyldt. $\square$

Definition 4.1.9 Lad A være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$. Da kaldes underrummene $N_A \subset \mathbb{F}^n$ og $R_A \subset \mathbb{F}^m$ for *nulrummet* og *søjlerummet* af matricen A .

Vi bemærker, at mængden af løsninger til den inhomogene ligning “ $A\mathbf{x} = \mathbf{b}$ ” ikke er et underrum af vektorrummet $\mathbb{F}^n$, hvis $\mathbf{b} \neq \mathbf{0}$. For da er $A\mathbf{0} \neq \mathbf{b}$, så (1) i definition 4.1.4 ikke opfyldt. Vi skal dog se i korollar 4.6.6, at løsningsmængden til $A\mathbf{x} = \mathbf{b}$ altid er et *affint underrum*.

Eksempel 4.1.10 Underrummet $U \subset \mathbb{R}^2$ fra eksempel 4.1.5 er netop nulrummet

$$N_A = \{\mathbf{x} \in \mathbb{R}^2 \mid A\mathbf{x} = \mathbf{0}\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \right\} \subset \mathbb{R}^2$$

for $A = \begin{pmatrix} 1 & -1 \end{pmatrix} \in M_{1,2}(\mathbb{R})$. Og givet et vilkårligt $b \in \mathbb{R}$, da opfylder

$$\mathbf{x} = \begin{pmatrix} b \\ 0 \end{pmatrix} \in \mathbb{R}^2,$$

at $A\mathbf{x} = b$, hvilket viser, at søjlerummet $R_A \subset \mathbb{R}$ er hele $\mathbb{R}$.

Eksempel 4.1.11 Per definition består søjlerummet af matricen

$$A = \begin{pmatrix} 1 & -i \\ -1 & i \end{pmatrix} \in M_{2,2}(\mathbb{C})$$

af de vektorer $\mathbf{b} \in \mathbb{C}^2$, for hvilke der findes $\mathbf{z} \in \mathbb{C}^2$, sådan at $A\mathbf{z} = \mathbf{b}$. Da

$$A\mathbf{z} = \begin{pmatrix} 1 & -i \\ -1 & i \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_1 - iz_2 \\ -z_1 + iz_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} (z_1 - iz_2) \in \mathbb{C}^2,$$

konkluderer vi, at

$$R_A = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot z \mid z \in \mathbb{C} \right\} \subset \mathbb{C}^2.$$

4.2 Lineære afbildninger

Vi har allerede defineret og studeret lineære afbildninger $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ i sektion 2.3, og vi udvider nu dette begreb til afbildninger mellem generelle vektorrum.

Definition 4.2.1 Lad $\mathbb{F}$ være et legeme og lad V og W være $\mathbb{F}$ -vektorrum. En afbildning $f: V \rightarrow W$ er *lineær*, hvis den opfylder følgende betingelser:

(L1) For alle $\mathbf{x}, \mathbf{y} \in V$ gælder det, at $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$.

(L2) For alle $\mathbf{x} \in V$ og $a \in \mathbb{F}$ gælder det, at $f(\mathbf{x} \cdot a) = f(\mathbf{x}) \cdot a$.

Vi understreger, at en lineær afbildning $f: V \rightarrow W$ mellem to vektorrum altid afbilder nulvektoren i V til nulvektoren i W . Det følger nemlig fra (L1), at

$$f(\mathbf{0}) + f(\mathbf{0}) = f(\mathbf{0} + \mathbf{0}) = f(\mathbf{0}),$$

hvorfra udsagnet følger ved at trække $f(\mathbf{0})$ fra på begge sider af lighedstegnet. Vi vil også sige, at en afbildning $g: V \rightarrow W$ er *affin*¹, hvis afbildningen $f: V \rightarrow W$ defineret ved $f(\mathbf{x}) = g(\mathbf{x}) - g(\mathbf{0})$ er lineær. For eksempel er identitetsafbildningen $\text{id}: V \rightarrow V$ og nulafbildningen $\mathbf{0}: V \rightarrow W$ begge lineære. Den konstante afbildning $\mathbf{b}: V \rightarrow W$ med værdi $\mathbf{b} \in W$ er altid affin, og den er lineær hvis og kun hvis $\mathbf{b} = \mathbf{0}$.

Eksempel 4.2.2 (1) Lad $\mathbb{F}$ være et legeme og lad $\mathbb{F}[t]$ være $\mathbb{F}$ -vektorrummet af polynomier med koefficienter i $\mathbb{F}$. Da er afbildningen $D: \mathbb{F}[t] \rightarrow \mathbb{F}[t]$, der til et polynomium

$$p(t) = \sum_{n \geq 0} a_n t^n$$

tilordner det formelt afledte polynomium

$$p'(t) = \sum_{n \geq 1} n a_n t^{n-1},$$

en lineær afbildning.

(2) Hvis V er et $\mathbb{F}$ -vektorrum, og hvis $U \subset V$ er et underrum, da er den kanoniske inklusion $i: U \rightarrow V$, som er defineret ved $i(\mathbf{u}) = \mathbf{u}$, en lineær afbildning. Lad også W være et $\mathbb{F}$ -vektorrum. Da gælder det endvidere, at en afbildning $f: W \rightarrow U$ er lineær, hvis og kun hvis den sammensatte afbildning $i \circ f: W \rightarrow V$ er lineær.

(3) Lad $C^0(\mathbb{R})$ være $\mathbb{R}$ -vektorrummet af kontinuerte afbildninger $f: \mathbb{R} \rightarrow \mathbb{R}$ fra eksempel 4.1.3 (9). Da er afbildningen $i: \mathbb{R} \rightarrow C^0(\mathbb{R})$, der er defineret ved $i(a)(x) = ax$, lineær.

¹I fysiklitteraturen anvendes "lineær" ofte i betydningen "affin".

4 Vektorrum

Vi viser nu, at lineære afbildninger $h: \mathbb{F}^n \rightarrow V$ fra $\mathbb{F}$ -vektorrummet $\mathbb{F}^n$ til et vilkårligt $\mathbb{F}$ -vektorrum V er i 1-1 korrespondance med n -tupler $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ af vektorer i V .

Lemma 4.2.3 *Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$ -vektorrum, og lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ være en n -tuple af vektorer i V . Da findes der præcis én lineær afbildning $h: \mathbb{F}^n \rightarrow V$, sådan at $h(\mathbf{e}_i) = \mathbf{v}_i$ for alle $1 \leq i \leq n$, nemlig afbildningen givet ved*

$$h(\mathbf{x}) = \mathbf{v}_1 x_1 + \mathbf{v}_2 x_2 + \dots + \mathbf{v}_n x_n.$$

Bevis Hvis $h: \mathbb{F}^n \rightarrow V$ er lineær og opfylder, at $h(\mathbf{e}_i) = \mathbf{v}_i$ for alle $1 \leq i \leq n$, da er

$$h(\mathbf{x}) = h(\mathbf{e}_1 x_1 + \dots + \mathbf{e}_n x_n) = h(\mathbf{e}_1) x_1 + \dots + h(\mathbf{e}_n) x_n = \mathbf{v}_1 x_1 + \dots + \mathbf{v}_n x_n,$$

hvilket viser entydighedsudsagnet. Omvendt opfylder afbildningen $h: \mathbb{F}^n \rightarrow V$ defineret ved denne formel, at $h(\mathbf{e}_i) = \mathbf{v}_i$ for alle $1 \leq i \leq n$, og vi viser nu, at den er lineær. For at vise, at $h: \mathbb{F}^n \rightarrow V$ opfylder (L1), lader vi $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ og udregner, at

$$\begin{aligned} h(\mathbf{x} + \mathbf{y}) &= \mathbf{v}_1(x_1 + y_1) + \dots + \mathbf{v}_n(x_n + y_n) \\ &= (\mathbf{v}_1 x_1 + \dots + \mathbf{v}_n x_n) + (\mathbf{v}_1 y_1 + \dots + \mathbf{v}_n y_n) = h(\mathbf{x}) + h(\mathbf{y}), \end{aligned}$$

som ønsket. Her følger den anden lighed ved gentagen anvendelse af (V3). For at vise, at $h: \mathbb{F}^n \rightarrow V$ opfylder (L2), lader vi $\mathbf{x} \in \mathbb{F}^n$ og $a \in \mathbb{F}$ og udregner, at

$$h(\mathbf{x}a) = \mathbf{v}_1(x_1 a) + \dots + \mathbf{v}_n(x_n a) = (\mathbf{v}_1 x_1 + \dots + \mathbf{v}_n x_n) a = h(\mathbf{x}) a$$

som ønsket. Her følger den anden lighed ved gentagen anvendelse af (V1) og (V2). $\square$

Eksempel 4.2.4 Lad $V = C^0([0, 2\pi])$ være vektorrummet af kontinuerte afbildninger $f: [0, 2\pi] \rightarrow \mathbb{R}$ fra eksempel 4.1.3 (9). Da er afbildningen $h: \mathbb{R}^2 \rightarrow C^0([0, 2\pi])$ givet ved

$$h \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \cos(x) a_1 + \sin(x) a_2$$

den entydigt bestemte lineære afbildning, sådan at $h(\mathbf{e}_1) = \cos(x)$ og $h(\mathbf{e}_2) = \sin(x)$.

Lemma 4.2.5 *Lad $\mathbb{F}$ være et legeme, og lad U , V og W være $\mathbb{F}$ -vektorrum. Givet to lineære afbildninger $f: V \rightarrow W$ og $g: U \rightarrow V$, da er $f \circ g: U \rightarrow W$ igen lineær.*

Bevis Vi viser, at $f \circ g: U \rightarrow W$ opfylder (L1) og (L2). For alle $\mathbf{x}, \mathbf{y} \in U$, gælder det, at

$$(f \circ g)(\mathbf{x} + \mathbf{y}) = f(g(\mathbf{x} + \mathbf{y})) = f(g(\mathbf{x}) + g(\mathbf{y})) = f(g(\mathbf{x})) + f(g(\mathbf{y})) = (f \circ g)(\mathbf{x}) + (f \circ g)(\mathbf{y}),$$

hvilket viser, at $f \circ g: U \rightarrow W$ opfylder (L1). Ligeledes gælder det for alle $\mathbf{x} \in U$ og $a \in \mathbb{F}$, at

$$(f \circ g)(\mathbf{x} \cdot a) = f(g(\mathbf{x} \cdot a)) = f(g(\mathbf{x}) \cdot a) = f(g(\mathbf{x})) \cdot a,$$

hvilket viser, at $f \circ g: U \rightarrow W$ opfylder (L2). $\square$

Mens lighed giver god mening for tal, så er det ikke noget godt begreb for vektorrum. Vi kan tænke på to vektorrum V og W som værende to sprog og ønsker da i stedet at spørge, om der findes en måde at oversætte fra et sprog til et andet. Endvidere er det vigtigt at kende den præcise oversættelse og ikke blot vide, om en oversættelse findes eller ej. Dette formaliseres i det generelle begreb isomorfi, som vi nu indfører.

Definition 4.2.6 Lad $\mathbb{F}$ være et legeme. En lineær afbildning $f: V \rightarrow W$ mellem to $\mathbb{F}$ -vektorrum er en *isomorfi*, hvis der findes en lineær afbildning $g: W \rightarrow V$, sådan at $f \circ g = \text{id}_W$ og $g \circ f = \text{id}_V$.

Vi viser, at en lineær afbildning er en isomorfi, hvis og kun hvis den er bijektiv.

Sætning 4.2.7 Lad $\mathbb{F}$ være et legeme og lad $f: V \rightarrow W$ være en lineær afbildning mellem to $\mathbb{F}$ -vektorrum. Da er følgende udsagn ækvivalente:

- (1) Afbildningen $f: V \rightarrow W$ er en bijektion.
- (2) Afbildningen $f: V \rightarrow W$ er en isomorfi.

Bevis Vi antager først (1) og viser (2). Da f er bijektiv, findes der ifølge lemma 0.1.3 en afbildning $g: W \rightarrow V$, som opfylder, at $f \circ g = \text{id}_W$ og $g \circ f = \text{id}_V$. Vi skal vise, at $g: W \rightarrow V$ er lineær. Da $f: V \rightarrow W$ opfylder (L1), gælder det for alle $\mathbf{y}, \mathbf{z} \in W$, at

$$\mathbf{y} + \mathbf{z} = f(g(\mathbf{y})) + f(g(\mathbf{z})) = f(g(\mathbf{y}) + g(\mathbf{z})),$$

hvilket netop viser, at $g(\mathbf{y} + \mathbf{z}) = g(\mathbf{y}) + g(\mathbf{z})$. Og da $f: V \rightarrow W$ også opfylder (L2), gælder det endvidere for alle $\mathbf{y} \in W$ og $a \in \mathbb{F}$, at

$$\mathbf{y} \cdot a = f(g(\mathbf{y})) \cdot a = f(g(\mathbf{y}) \cdot a),$$

hvilket viser, $g(\mathbf{y} \cdot a) = gf(g(\mathbf{y}) \cdot a) = g(\mathbf{y}) \cdot a$. Altså er g lineær, hvilket viser (2).

Vi antager dernæst (2) og viser (1). Da f er en isomorfi, findes der per definition en lineær afbildning $g: W \rightarrow V$, sådan at $f \circ g = \text{id}_W$ og $g \circ f = \text{id}_V$. Specielt følger det fra lemma 0.1.3, at $f: V \rightarrow W$ er en bijektion. $\square$

4 Vektorrum

Eksempel 4.2.8 (1) Lad os betragte $M_{2,3}(\mathbb{F})$ som et $\mathbb{F}$ -vektorrum med vektorsum og skalarmultiplikation defineret som i eksempel 4.1.3 (3). Da er afbildningen

$$h: \mathbb{F}^6 \rightarrow M_{2,3}(\mathbb{F})$$

defineret ved

$$h(\mathbf{x}) = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \end{pmatrix}$$

en lineær bijektion. Ifølge sætning 4.2.7 er den dermed en isomorfi.

(2) Lad $V = \{ |\uparrow\rangle z_1 + |\downarrow\rangle z_2 \mid z_1, z_2 \in \mathbb{C} \}$ være det komplekse vektorrum af formelle linearkombinationer af symbolerne $|\uparrow\rangle$ og $|\downarrow\rangle$ fra eksempel 4.1.3 (7). Da er afbildningen

$$h: \mathbb{C}^2 \rightarrow V$$

defineret ved $h(\mathbf{z}) = |\uparrow\rangle z_1 + |\downarrow\rangle z_2$ en lineær bijektion. Ifølge sætning 4.2.7 er den dermed en isomorfi.

Vi viser nu, at det for lineære afbildninger er særligt let at afgøre, hvorvidt de er injektive eller ej.

Lemma 4.2.9 Hvis $\mathbb{F}$ er et legeme, og hvis $f: V \rightarrow W$ er en lineær afbildning mellem to $\mathbb{F}$ -vektorrum, da er de følgende udsagn ækvivalente:

(1) Afbildningen $f: V \rightarrow W$ er injektiv.

(2) Hvis $\mathbf{v} \in V$ og $f(\mathbf{v}) = \mathbf{0}$, så er $\mathbf{v} = \mathbf{0}$.

Bevis Vi minder om, at $f: V \rightarrow W$ er injektiv, hvis der for alle $\mathbf{w} \in W$, højst findes ét $\mathbf{v} \in V$, sådan at $f(\mathbf{v}) = \mathbf{w}$. Så (1) medfører per definition (2). Vi antager derfor, at (2) gælder, og viser (1). Vi skal vise, at hvis $\mathbf{v}_1, \mathbf{v}_2 \in V$ og $f(\mathbf{v}_1) = f(\mathbf{v}_2)$, da er $\mathbf{v}_1 = \mathbf{v}_2$. Vi lader derfor $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$ og udregner, at

$$\begin{aligned} f(\mathbf{v}) &= f(\mathbf{v}_1 - \mathbf{v}_2) = f(\mathbf{v}_1 + \mathbf{v}_2 \cdot (-1)) \stackrel{(L1)}{=} f(\mathbf{v}_1) + f(\mathbf{v}_2 \cdot (-1)) \\ &\stackrel{(L2)}{=} f(\mathbf{v}_1) + f(\mathbf{v}_2) \cdot (-1) = f(\mathbf{v}_1) - f(\mathbf{v}_2) = \mathbf{0}. \end{aligned}$$

Ifølge vores antagelse, at $f: V \rightarrow W$ opfylder (2), medfører dette, at $\mathbf{v} = \mathbf{0}$. Men da er $\mathbf{v}_1 = \mathbf{v}_2$, hvilket viser, at $f: V \rightarrow W$ opfylder (1) som ønsket. $\square$

Eksempel 4.2.10 Vi anvender lemma 4.2.9 til at vise, at den lineære afbildning

$$h: \mathbb{R}^2 \rightarrow C^0([0, 2\pi])$$

fra eksempel 4.2.4, som er defineret ved $h(\mathbf{a}) = \cos(x)a_1 + \sin(x)a_2$, er injektiv. Vi skal vise, at hvis $\mathbf{a} \in \mathbb{R}^2$ og $h(\mathbf{a}) = \mathbf{0}$, da gælder det nødvendigvis, at $\mathbf{a} = \mathbf{0}$. Vi husker på, at nulvektoren $\mathbf{0} \in C^0([0, 2\pi])$ er nulafbildningen $\mathbf{0}: [0, 2\pi] \rightarrow \mathbb{R}$, som er givet ved $\mathbf{0}(x) = 0$ for alle $x \in [0, 2\pi]$. Vi skal derfor vise, at hvis

$$\cos(x)a_1 + \sin(x)a_2 = 0$$

for alle $x \in [0, 2\pi]$, så er både $a_1 = 0$ og $a_2 = 0$. Vi kan tænke på denne ligning som et lineært ligningssystem i de to variable a_1 og a_2 og med uendeligt mange ligninger, der er indicerede ved $x \in [0, 2\pi]$. For $x = 0$ har vi ligningen

$$a_1 = \cos(0)a_1 + \sin(0)a_2 = 0,$$

som viser, at $a_1 = 0$. Og for $x = \frac{\pi}{2}$ har vi ligningen

$$a_2 = \cos(\frac{\pi}{2})a_1 + \sin(\frac{\pi}{2})a_2 = 0,$$

som viser, at $a_2 = 0$. Dette viser, at $h: \mathbb{R}^2 \rightarrow C^0([0, 2\pi])$ er injektiv.

4.3 Basis for et vektorrum

Vi indfører nu tre helt centrale egenskaber. Disse begreber knytter sig ikke til enkelte vektorer i et vektorrum, men derimod til *familier* af vektorer, og de kræver derfor nogen tilvænning. De tre centrale egenskaber af en familie af vektorer i et vektorrum V er, at den kan være *lineært uafhængig*, at den kan *frembringe* V , eller at den kan være en *basis for* V . For lethedens skyld vil vi for det meste kun betragte *endelige* familier; vi betragter uendelige familier sidst i afsnittet.

Lad $(x_i)_{i \in I}$ være en familie af elementer i en mængde X indiceret ved en mængde I . Hvis I er endelig, så siger vi, at familien $(x_i)_{i \in I}$ er endelig, og hvis $I = \{1, 2, \dots, n\}$, så skriver vi også $(x_1, x_2, \dots, x_n)$ i stedet for $(x_i)_{i \in I}$ og siger, at denne familie er en n -tuple af elementer i X . Standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ fra definition 2.2.4 er således en n -tuple af vektorer i $\mathbb{F}^n$, og for enhver mængde X er den tomme familie $()$ den entydigt bestemte 0-tuple af elementer i X . Vi understreger, at en familie af elementer ikke er det samme som en mængde af elementer. For eksempel er 1-tuplen (1) og 2-tuplen $(1, 1)$ forskellige, mens mængderne $\{1\}$ og $\{1, 1\}$ er ens. Vi betegner altid familier med runde parenteser, mens vi betegner mængder med Tuborgparenteser.

4 Vektorrum

Definition 4.3.1 Lad $\mathbb{F}$ være et legeme, og lad V være et $\mathbb{F}$ -vektorrum. En vektor $\mathbf{v} \in V$ siges at være en *linearkombination* af en endelig familie $(\mathbf{v}_i)_{i \in I}$ af vektorer i V , hvis der findes en familie af skalarer $(a_i)_{i \in I}$ med samme index mængde, sådan at

$$\mathbf{v} = \sum_{i \in I} \mathbf{v}_i a_i.$$

Vi bemærker, at hvis $\mathbf{v}$ er en linearkombination af $(\mathbf{v}_i)_{i \in I}$, da vil der generelt være flere familier af skalarer $(a_i)_{i \in I}$, sådan at $\mathbf{v} = \sum_{i \in I} \mathbf{v}_i a_i$.

Eksempel 4.3.2 (1) I ethvert vektorrum V er nulvektoren $\mathbf{0}$ en linearkombination af den tomme familie $(\)$. For den tomme sum i V er per definition lig med $\mathbf{0}$.

(2) Vi viste i lemma 2.2.6, at enhver vektor $\mathbf{x} \in \mathbb{F}^n$ på entydig vis kan skrives som en linearkombination af standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$, nemlig,

$$\mathbf{x} = \mathbf{e}_1 x_1 + \dots + \mathbf{e}_n x_n.$$

(3) Enhver vektor $\mathbf{x} \in \mathbb{R}^2$ er en linearkombination af familien

$$(\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}),$$

idet det for familien af skalarer $(a_1, a_2) = (x_1 - x_2, -x_1 + 2x_2)$ gælder, at

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} (x_1 - x_2) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} (-x_1 + 2x_2) = \mathbf{v}_1 a_1 + \mathbf{v}_2 a_2.$$

Endvidere er familien af skalarer (a_1, a_2) med denne egenskab entydigt bestemt af $\mathbf{x}$.

(4) Enhver vektor $\mathbf{x} \in \mathbb{R}^2$ er en linearkombination af familien

$$(\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}),$$

idet $\mathbf{x}$, som vi netop har set, kan skrives som en linearkombination af familien $(\mathbf{v}_1, \mathbf{v}_2)$. Imidlertid er familien af skalarer (a_1, a_2, a_3) , sådan at $\mathbf{x} = \mathbf{v}_1 a_1 + \mathbf{v}_2 a_2 + \mathbf{v}_3 a_3$, ikke entydigt bestemt af $\mathbf{x}$. For eksempel er både

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot 1 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot (-1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot 0 \quad \text{og} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot 0 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot 1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot (-1).$$

Faktisk er der uendeligt mange måder at skrive $\mathbf{x} \in \mathbb{R}^2$ som en linearkombination af familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.

(5) Lad $A \in M_{m,n}(\mathbb{F})$, og lad $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ være den familie af vektorer i $\mathbb{F}^m$, der består af søjlerne i A . En vektor $\mathbf{b} \in \mathbb{F}^m$ er da en linearkombination

$$\mathbf{b} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n$$

af denne familie, hvis og kun hvis ligningssystemet $A\mathbf{x} = \mathbf{b}$ har mindst én løsning $\mathbf{x} \in \mathbb{F}^n$. I givet fald er familien af skalarer $(x_1, \dots, x_n)$ entydigt bestemt af $\mathbf{b}$, hvis og kun hvis ligningssystemet $A\mathbf{x} = \mathbf{b}$ har netop én løsning $\mathbf{x} \in \mathbb{F}^n$.

(6) Lad $\mathbb{F}[t]_{\leq d}$ være vektorrummet af polynomier med koefficienter i $\mathbb{F}$ af grad højst d fra eksempel 4.1.3 (8). Enhver vektor $p(t)$ i dette vektorrum er en linearkombination af familien af vektorer $(t^0, t^1, \dots, t^d)$. For hvis $p(t) = \sum_{0 \leq n \leq d} a_n t^n$, da er

$$p(t) = \sum_{0 \leq n \leq d} t^n \cdot a_n.$$

Vi bemærker, at familien af skalarer $(a_0, a_1, \dots, a_d)$ er entydigt bestemt af $p(t)$.

(4) Lad $C^0([0, 2\pi], \mathbb{C})$ være $\mathbb{C}$ -vektorrummet af kontinuerte funktioner $f: [0, 2\pi] \rightarrow \mathbb{C}$. Vektoren $\exp(ix)$ i dette vektorrum er en linearkombination

$$\exp(ix) = \cos(x) \cdot 1 + \sin(x) \cdot i$$

af familien af vektorer $(\cos(x), \sin(x))$. Familien af skalarer $(1, i)$ er igen entydig.

Delmængden af et vektorrum V , der består af alle linearkombinationer af en endelig familie af vektorer $(\mathbf{v}_i)_{i \in I}$ kaldes for *spannet* af $(\mathbf{v}_i)_{i \in I}$ og betegnes $\text{span}((\mathbf{v}_i)_{i \in I})$. Vi viser nu, at denne delmængde udgør et underrum $U \subset V$, og vi kalder dette for *underrummet frembragt af familien $(\mathbf{v}_i)_{i \in I}$* .

Lemma 4.3.3 *Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$ -vektorrum, og lad $(\mathbf{v}_i)_{i \in I}$ være en endelig familie af vektorer i V . Da udgør delmængden*

$$U = \text{span}((\mathbf{v}_i)_{i \in I}) \subset V,$$

der består af alle linearkombinationer af familien $(\mathbf{v}_i)_{i \in I}$, et underrum af V .

Bevis Vi skal vise, at delmængden $U \subset V$ opfylder (1)–(3) i definition 4.1.4. Vi kan udtrykke nulvektoren som $\mathbf{0} = \sum_{i \in I} \mathbf{v}_i \cdot 0$. Derfor er $\mathbf{0} \in U$, så (1) er opfyldt. Og hvis både $\mathbf{x} = \sum_{i \in I} \mathbf{v}_i a_i$ og $\mathbf{y} = \sum_{i \in I} \mathbf{v}_i b_i$ er i U , da er $\mathbf{x} + \mathbf{y} = \sum_{i \in I} \mathbf{v}_i a_i + \sum_{i \in I} \mathbf{v}_i b_i = \sum_{i \in I} \mathbf{v}_i (a_i + b_i)$ også i U , så (2) er også opfyldt. Hvis endelig $\mathbf{x} = \sum_{i \in I} \mathbf{v}_i a_i$ er i U og $a \in \mathbb{F}$, så er også $\mathbf{x}a = (\sum_{i \in I} \mathbf{v}_i a_i)a = \sum_{i \in I} \mathbf{v}_i (a_i a)$ i U , hvilket viser, at (3) er opfyldt. $\square$

4 Vektorrum

Definition 4.3.4 Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$ -vektorrum og lad $(\mathbf{v}_i)_{i \in I}$ være en endelig familie af vektorer i V .

- (1) Familien af vektorer $(\mathbf{v}_i)_{i \in I}$ siges at være *lineært uafhængig*, hvis den eneste familie af skalarer $(a_i)_{i \in I}$, for hvilken det gælder, at

$$\sum_{i \in I} \mathbf{v}_i a_i = \mathbf{0},$$

er nulfamilien $(a_i)_{i \in I}$, hvor $a_i = 0$ for alle $i \in I$. Ellers siges familien $(\mathbf{v}_i)_{i \in I}$ at være *lineært afhængig*.

- (2) Familien af vektorer $(\mathbf{v}_i)_{i \in I}$ siges at *frembringe* eller *udspænde* V , hvis enhver vektor $\mathbf{v} \in V$ er en linearkombination

$$\mathbf{v} = \sum_{i \in I} \mathbf{v}_i a_i$$

af familien $(\mathbf{v}_i)_{i \in I}$.

- (3) Familien af vektorer $(\mathbf{v}_i)_{i \in I}$ siges at være en *basis* for V , hvis den både er lineært uafhængig og frembringer V .

Det er ofte lettest at bruge definitionen ovenfor til at afgøre om en given familie af vektorer i et vektorrum er en basis eller ej. Det følgende lemma giver en alternativ karakterisering af baser, og antyder samtidig, hvorfor dette begreb er så nyttigt.

Lemma 4.3.5 Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$ -vektorrum, og lad $(\mathbf{v}_i)_{i \in I}$ være en endelig familie af vektorer i V . Da er følgende ækvivalent:

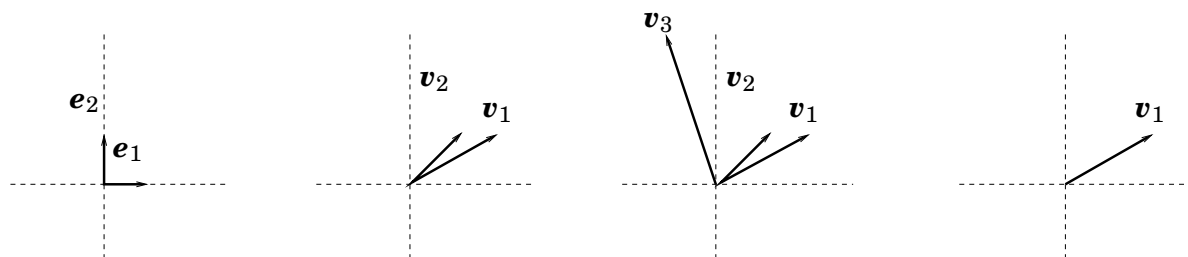
- (i) Familien $(\mathbf{v}_i)_{i \in I}$ er en basis for V .
- (ii) Enhver vektor $\mathbf{v} \in V$ kan skrives som en linearkombination af familien $(\mathbf{v}_i)_{i \in I}$, og denne opskrivning er entydig.

Bevis Vi antager først (i) og viser (ii). Da $(\mathbf{v}_i)_{i \in I}$ frembringer V , kan enhver vektor $\mathbf{v} \in V$ per definition skrives som en linearkombination $\mathbf{v} = \sum_{i \in I} \mathbf{v}_i a_i$ af familien $(\mathbf{v}_i)_{i \in I}$. Vi skal derfor vise, at hvis også $\mathbf{v} = \sum_{i \in I} \mathbf{v}_i b_i$, da er $a_i = b_i$ for alle $i \in I$. Men

$$\sum_{i \in I} \mathbf{v}_i (a_i - b_i) = \sum_{i \in I} \mathbf{v}_i a_i - \sum_{i \in I} \mathbf{v}_i b_i = \mathbf{v} - \mathbf{v} = \mathbf{0},$$

og da $(\mathbf{v}_i)_{i \in I}$ er lineært uafhængig, konkluderer vi, at $a_i - b_i = 0$ for alle $i \in I$. Dette viser som ønsket, at (ii) gælder.

Vi antager dernæst (ii) og viser (i). Da enhver vektor $\mathbf{v} \in V$ kan skrives som en linearkombination af familien $(\mathbf{v}_i)_{i \in I}$, så frembringer denne familie V . Endvidere kan nulvektoren $\mathbf{0} \in V$ altid skrives som linearkombinationen $\mathbf{0} = \sum_{i \in I} \mathbf{v}_i \cdot 0$, og per antagelse er denne opskrivning entydig. Derfor er $(\mathbf{v}_i)_{i \in I}$ også lineært uafhængig, hvilket viser, at (i) gælder. $\square$



Figur 4.2: Familierne $(\mathbf{e}_1, \mathbf{e}_2)$ og $(\mathbf{v}_1, \mathbf{v}_2)$ er baser for $\mathbb{R}^2$, mens familierne $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ og $(\mathbf{v}_1)$ ikke er der; se eksempel 4.3.6.

Eksempel 4.3.6 (1) Vi betragter de følgende tre vektorer i $\mathbb{R}^2$:

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Familien $(\mathbf{v}_1)$ er lineært uafhængig, da $\mathbf{v}_1 a_1 = \mathbf{0}$ medfører, at $a_1 = 0$. Den frembringer imidlertid ikke $\mathbb{R}^2$, da vektoren $\mathbf{v}_2$ for eksempel ikke er en linearkombination af $(\mathbf{v}_1)$. Familien $(\mathbf{v}_1, \mathbf{v}_2)$ er også lineært uafhængig, da $\mathbf{v}_1 a_1 + \mathbf{v}_2 a_2 = \mathbf{0}$, hvis og kun hvis

$$\begin{aligned} 2a_1 + a_2 &= 0 \\ a_1 + a_2 &= 0, \end{aligned}$$

og da dette ligningssystem har kun den ene løsning, hvor både $a_1 = 0$ og $a_2 = 0$. Vi så eksempel 4.3.2, at $(\mathbf{v}_1, \mathbf{v}_2)$ frembringer $\mathbb{R}^2$, så dermed er $(\mathbf{v}_1, \mathbf{v}_2)$ en basis for $\mathbb{R}^2$. Familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ frembringer også $\mathbb{R}^2$, fordi delfamilien $(\mathbf{v}_1, \mathbf{v}_2)$ gør det, men den er ikke lineært uafhængig. For ligningssystemet $\mathbf{v}_1 a_1 + \mathbf{v}_2 a_2 + \mathbf{v}_3 a_3 = \mathbf{0}$ har løsninger, hvor a_1, a_2 og a_3 ikke alle er nul, for eksempel $a_1 = 4, a_2 = -7$ og $a_3 = 1$.

(2) I ethvert vektorrum V er den tomme familie $()$ lineært uafhængig, da betingelsen i definition 4.3.4 (1) trivielt er opfyldt. Den tomme familie er dermed en basis for V , hvis og kun hvis $V = \{\mathbf{0}\}$. For den tomme sum er per definition lig med $\mathbf{0}$.

(3) Ifølge lemma 2.2.6 er standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ i definition 2.2.4 en basis for $\mathbb{F}^m$.

(4) En endelig familie af vektorer $(\mathbf{v}_i)_{i \in I}$, hvor der findes et $j \in I$, sådan at $\mathbf{v}_j = \mathbf{0}$, er lineært afhængig. For lad $(a_i)_{i \in I}$ være familien af skalarer, hvor $a_j = 1$ og $a_i = 0$ for $i \neq j$. Da er $\sum_{i \in I} \mathbf{v}_i a_i = \mathbf{0}$, selvom $(a_i)_{i \in I}$ ikke er nulfamilien.

4 Vektorrum

Ifølge lemma 4.1.8 er nulrummet $N_A \subset \mathbb{F}^n$ af en $m \times n$ -matrix A et underrum, og vi finder nu i et eksempel en basis herfor. Vi giver en generel algoritme i sætning 4.5.11.

Eksempel 4.3.7 (Basis for nulrum) Vi har tidligere i eksempel 1.2.16 betragtet den følgende matrix A og udregnet den entydigt bestemte matrix A' , der er på reduceret echelonform og rækkeækvivalent med A .

$$A = \begin{pmatrix} 0 & 0 & 2 & -1 & 8 \\ 1 & -2 & 3 & 2 & 1 \\ 3 & -6 & 10 & 6 & 5 \end{pmatrix} \quad A' = \begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -4 \end{pmatrix}$$

Vi aflæser heraf, at den fælles løsningsmængde til “ $A\mathbf{x} = \mathbf{0}$ ” og “ $A'\mathbf{x} = \mathbf{0}$ ” er givet ved

$$N_A = N_{A'} = \left\{ \begin{pmatrix} 2t_1 & -3t_2 \\ t_1 & -2t_2 \\ 0 & 4t_2 \\ 0 & t_2 \end{pmatrix} \mid t_1, t_2 \in \mathbb{F} \right\}.$$

Med andre ord har vi vist, at familien

$$\left(\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} -3 \\ 0 \\ -2 \\ 4 \\ 1 \end{pmatrix} \right)$$

udgør en basis for nulrummet $N_A = N_{A'} \subset \mathbb{F}^5$.

Lemma 4.1.8 viser tilsvarende, at søjlerummet $R_A \subset \mathbb{F}^m$ af en $m \times n$ -matrix A er et underrum. Vi viser nu, at familien af søjler i A frembringer dette underrum, hvilket forklarer navnet.

Eksempel 4.3.8 (Frembringere for søjlerummet) Hvis $A \in M_{m,n}(\mathbb{F})$ er en matrix, da frembringer familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$, der består af søjlerne i matricen A , søjlerummet $R_A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^n\} \subset \mathbb{F}^m$. For hvis $\mathbf{x} = \mathbf{e}_1x_1 + \dots + \mathbf{e}_nx_n$, da er

$$A\mathbf{x} = A(\mathbf{e}_1x_1 + \dots + \mathbf{e}_nx_n) = (A\mathbf{e}_1)x_1 + \dots + (A\mathbf{e}_n)x_n = \mathbf{a}_1x_1 + \dots + \mathbf{a}_nx_n.$$

Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ er generelt ikke lineært uafhængig og er derfor i almindelighed ikke en basis for $R_A \subset \mathbb{F}^m$. Vi angiver senere i sætning 4.5.8 en generel algoritme, der viser, hvordan denne familie udtyndes til en basis.

Lemma 4.3.9 *Lad $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ være den n -tuple af vektorer i $\mathbb{F}^m$, der består af søjlerne i en $m \times n$ -matrix $A \in M_{m,n}(\mathbb{F})$.*

- (1) *Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ er lineært uafhængig, hvis og kun $\mathbf{x} = \mathbf{0}$ er den eneste løsning til " $A\mathbf{x} = \mathbf{0}$." Dette gælder endvidere, hvis og kun hvis " $A\mathbf{x} = \mathbf{b}$ " højst har én løsning for alle $\mathbf{b} \in \mathbb{F}^m$.*
- (2) *Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ frembringer $\mathbb{F}^m$, hvis og kun " $A\mathbf{x} = \mathbf{b}$ " har mindst én løsning for alle $\mathbf{b} \in \mathbb{F}^m$.*
- (3) *Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ er en basis for $\mathbb{F}^m$, hvis og kun hvis " $A\mathbf{x} = \mathbf{b}$ " har præcis én løsning for alle $\mathbf{b} \in \mathbb{F}^m$. Dette gælder endvidere, hvis og kun hvis A er invertibel.*

Bevis En løsning $\mathbf{x} \in \mathbb{F}^n$ til ligningssystemet " $A\mathbf{x} = \mathbf{b}$ " udtrykker netop $\mathbf{b} \in \mathbb{F}^m$ som en linearkombination $A\mathbf{x} = \mathbf{a}_1x_1 + \dots + \mathbf{a}_nx_n$ af familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$. Den første påstand i (1) samt påstanden (2) følger derfor fra definition 4.3.4. Den sidste del af (1) fås igen ved at bemærke, at hvis både $A\mathbf{y} = \mathbf{b}$ og $A\mathbf{z} = \mathbf{b}$, så er $\mathbf{x} = \mathbf{y} - \mathbf{z}$ en løsning til " $A\mathbf{x} = \mathbf{0}$." Endelig udtrykker den første påstand i (3) netop ækvivalensen af de to udsagn i lemma 4.3.5, mens den sidste påstand følger fra sætning 2.4.9. $\square$

Eksempel 4.3.10 Vi udregnede i eksempel 3.2.9, at matricen

$$A = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} \in M_3(\mathbb{R})$$

har determinant $\det(A) = 132 \neq 0$. Da $\det(A)$ er invertibel i $\mathbb{R}$, er A derfor invertibel ifølge sætning 3.4.1. Derfor viser lemma 4.3.9 altså, at familien

$$\left(\mathbf{a}_1 = \begin{pmatrix} 5 \\ 0 \\ 2 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} -1 \\ 2 \\ 6 \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix} \right)$$

er en basis for $\mathbb{R}^3$.

Ifølge lemma 4.2.3 bestemmer en n -tuple $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ af vektorer i et $\mathbb{F}$ -vektorrum V en entydig lineær afbildning $h: \mathbb{F}^n \rightarrow V$, sådan at $h(\mathbf{e}_i) = \mathbf{v}_i$ for alle $1 \leq i \leq n$. Vi viser nu, at denne afbildning præcis er injektiv, surjektiv eller bijektiv eftersom n -tuplen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er lineært uafhængig, frembringer V eller er en basis for V .

Sætning 4.3.11 *Lad $\mathbb{F}$ være et legeme, og lad V være et $\mathbb{F}$ -vektorrum. Lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ være en n -tuple af vektorer i V , og lad $h: \mathbb{F}^n \rightarrow V$ være den entydigt bestemte lineære afbildning, sådan at $h(\mathbf{e}_i) = \mathbf{v}_i$ for alle $1 \leq i \leq n$. Da gælder følgende udsagn:*

- (1) *Familien $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er lineært uafhængig, hvis og kun hvis $h: \mathbb{F}^n \rightarrow V$ er injektiv.*
- (2) *Familien $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ frembringer V , hvis og kun hvis $h: \mathbb{F}^n \rightarrow V$ er surjektiv.*
- (3) *Familien $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er en basis for V , hvis og kun hvis $h: \mathbb{F}^n \rightarrow V$ er bijektiv.*

Bevis Per definition er familien $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ lineært uafhængig, hvis og kun hvis det for alle $\mathbf{x} \in \mathbb{F}^n$ gælder, at hvis $h(\mathbf{x}) = \mathbf{0}$, så er $\mathbf{x} = \mathbf{0}$. Og ifølge lemma 4.2.9 er dette tilfældet, hvis og kun hvis $h: \mathbb{F}^n \rightarrow V$ er injektiv. Dette viser (1).

Familien $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ frembringer per definition V , hvis og kun hvis enhver vektor $\mathbf{v} \in V$ er en linearkombination $\mathbf{v} = \mathbf{v}_1 x_1 + \dots + \mathbf{v}_n x_n$ af $(\mathbf{v}_1, \dots, \mathbf{v}_n)$. Dette betyder præcist, at $h: \mathbb{F}^n \rightarrow V$ er surjektiv, hvorfor (2) følger.

Endelig er $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ en basis, hvis og kun hvis den både er lineært uafhængig og frembringer V , og vi har netop vist, at dette er tilfældet, hvis og kun hvis afbildningen $h: \mathbb{F}^n \rightarrow V$ både er injektiv og surjektiv. Dette viser (3). $\square$

Hvis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er en basis for V , så viser sætning 4.3.11, at den entydige lineære afbildning $h: \mathbb{F}^n \rightarrow V$, sådan at $h(\mathbf{e}_i) = \mathbf{v}_i$ for alle $1 \leq i \leq n$, er en isomorfi. Vi skal i næste afsnit se, at dette tillader os at oversætte alle spørgsmål i lineær algebra, der angår V , til tilsvarende spørgsmål, der angår $\mathbb{F}^n$. Det er dog vigtigt at huske på, at denne oversættelse *afhænger af valget af basis* for V .

Korollar 4.3.12 *Lad A være en $m \times n$ -matrix af rang $0 \leq r \leq \min\{m, n\}$ med indgange i et legeme $\mathbb{F}$, og lad $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ være familien af søjlevektorer i A .*

- (1) *Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ er lineært uafhængig, hvis og kun hvis $r = n$.*
- (2) *Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ frembringer $\mathbb{F}^m$, hvis og kun hvis $r = m$.*
- (3) *Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ er en basis for $\mathbb{F}^m$, hvis og kun hvis $r = m = n$.*

Særligt er familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ en basis hvis og kun hvis A er invertibel.

Bevis Da den entydigt bestemte lineære afbildning $h: \mathbb{F}^n \rightarrow \mathbb{F}^m$, sådan at $h(\mathbf{e}_i) = \mathbf{a}_i$ for $1 \leq i \leq n$, er givet ved $h(\mathbf{x}) = \mathbf{a}_1 x_1 + \dots + \mathbf{a}_n x_n = A\mathbf{x}$, følger påstanden umiddelbart fra sætning 4.3.11 og sætning 2.5.15. $\square$

Eksempel 4.3.13 Vi betragter igen matricen

$$A = \begin{pmatrix} 0 & 0 & 2 & -1 & 8 \\ 1 & -2 & 3 & 2 & 1 \\ 3 & -6 & 10 & 6 & 5 \end{pmatrix}$$

fra eksempel 4.3.7. Den entydigt bestemte matrix A' , der er rækkeækvivalent til A og på reduceret echelonform, udregnes til at være

$$A' = \begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -4 \end{pmatrix}.$$

Matricen A' har 3 ledende indgange, og derfor er $\text{rank}(A) = 3$. Da A er en 3×5 -matrix, sådan at $m = 3$ og $n = 5$, og da $r = \text{rank}(A) = 3 = m$, så viser korollar 4.3.12, at søjlerne i A frembringer $\mathbb{R}^3$.

Korollar 4.3.14 *Lad A være en $n \times n$ -matrix med indgange i et legeme $\mathbb{F}$, og lad $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ være familien af søjlevektorer i A . Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ er en basis for $\mathbb{F}^n$, hvis og kun hvis $\det(A) \neq 0$.*

Bevis Dette er en direkte konsekvens af korollar 3.4.2 and korollar 4.3.12. $\square$

Vi skal nu vise, at et vektorrum, der er frembragt af en endelig familie af vektorer, har en basis. Vi viser et mere præcist resultat, som er mere anvendeligt. Generelt, hvis $(x_i)_{i \in I}$ er en familie af elementer i en mængde X , og hvis $J \subset I$ er en delmængde, så siger vi, at familien $(x_i)_{i \in J}$ er en *delfamilie* af familien $(x_i)_{i \in I}$. Særligt er alle delfamilier af en n -tuple $(x_1, \dots, x_n)$ af formen $(x_{i_1}, \dots, x_{i_p})$, hvor $0 \leq p \leq n$ og $1 \leq i_1 < \dots < i_p \leq n$. For eksempel er den tomme familie $()$ og familierne $(\mathbf{v}_1, \mathbf{v}_3)$ og $(\mathbf{v}_2)$ alle delfamilier af familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. Den følgende sætning er den lineære algebras hovedsætning.

Sætning 4.3.15 *Lad $\mathbb{F}$ være et legeme og lad V være et $\mathbb{F}$ -vektorrum. Hvis $(\mathbf{v}_i)_{i \in I}$ er en endelig familie af vektorer i V , som frembringer V , og hvis $(\mathbf{v}_i)_{i \in K}$ er en lineært uafhængig delfamilie deraf, da findes en mængde J med $K \subset J \subset I$, sådan at delfamilien $(\mathbf{v}_i)_{i \in J}$ er en basis for V .*

Bevis Vi ønsker at finde $K \subset J \subset I$, sådan at $(\mathbf{v}_i)_{i \in J}$ er en basis for V . Vi ved, at der findes $K \subset J \subset I$, sådan at $(\mathbf{v}_i)_{i \in J}$ er lineært uafhængig, for det er rigtigt for $J = K$.

4 Vektorrum

Vi vælger derfor et $K \subset J \subset I$, sådan at $(\mathbf{v}_i)_{i \in J}$ er lineært uafhængig, og sådan at J er maksimal i den forstand, at hvis $J \subset S \subset I$ og $(\mathbf{v}_i)_{i \in S}$ er lineært uafhængig, da er $S = J$. Vi vil nu vise, at familien $(\mathbf{v}_i)_{i \in J}$ også frembringer V .

Så vi antager modsætningsvis, at dette ikke er tilfældet og viser, at det strider mod maksimaliteten af J . Ifølge vores antagelse findes der en vektor $\mathbf{v} \in V$, som ikke er en linearkombination af $(\mathbf{v}_i)_{i \in J}$, og da familien $(\mathbf{v}_i)_{i \in I}$ frembringer V , kan vi yderligere antage, at $\mathbf{v} = \mathbf{v}_h$, hvor $h \in I$ og $h \notin J$. Vi lader $S = J \cup \{h\} \subset I$ og viser, at familien $(\mathbf{v}_i)_{i \in S}$ er lineært uafhængig. Så lad

$$\sum_{i \in S} \mathbf{v}_i a_i = \mathbf{0}$$

være en linearkombination, der er lig med nulvektoren. Hvis $a_h \neq 0$, da er

$$\mathbf{v}_h = \sum_{i \in J} \mathbf{v}_i \cdot (-a_i a_h^{-1})$$

en linearkombination af $(\mathbf{v}_i)_{i \in J}$. Men da vi netop har antaget, at $\mathbf{v}_h$ ikke er en linearkombination af $(\mathbf{v}_i)_{i \in J}$, konkluderer vi, at $a_h = 0$. Deraf følger, at

$$\sum_{i \in J} \mathbf{v}_i a_i = \sum_{i \in S} \mathbf{v}_i a_i = \mathbf{0},$$

og da $(\mathbf{v}_i)_{i \in J}$ er lineært uafhængig, følger det, at $a_i = 0$ for alle $i \in J$. Vi har altså hermed vist, at $a_i = 0$ for alle $i \in S$, hvilket viser, at $(\mathbf{v}_i)_{i \in S}$ er lineært uafhængig. Men $S \neq J$, så dette strider mod maksimaliteten af J , og vi konkluderer derfor, at vores antagelse, at familien $(\mathbf{v}_i)_{i \in J}$ ikke frembringer V er forkert. Så familien $(\mathbf{v}_i)_{i \in J}$ frembringer V og er samtidigt lineært uafhængig og derfor en basis for V . Dette viser sætningen. $\square$

Definition 4.3.16 Et vektorrum V er *endeligt frembragt*, hvis der findes en endelig familie $(\mathbf{v}_i)_{i \in I}$ af vektorer i V , der frembringer V .

Sætning 4.3.15 viser særligt, at et endeligt frembragt vektorrum V har en endelig basis. Lad nemlig $(\mathbf{v}_i)_{i \in I}$ være en endelig familie af vektorer i V , der frembringer V . Da den tomme delfamilie $(\)$ er lineært uafhængig, viser sætning 4.3.15, at der findes $\emptyset \subset J \subset I$, sådan at $(\mathbf{v}_i)_{i \in J}$ er en basis for V .

Eksempel 4.3.17 Vi betragter igen familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ af vektorer i $\mathbb{R}^2$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Som vi bemærkede i eksempel 4.3.6 (1), så frembringer denne familie $\mathbb{R}^2$, men den er ikke lineært uafhængig. De tre delfamilier $(\mathbf{v}_1, \mathbf{v}_2)$, $(\mathbf{v}_1, \mathbf{v}_3)$ og $(\mathbf{v}_2, \mathbf{v}_3)$ er alle lineært uafhængige og frembringer $\mathbb{R}^2$, og de udgør dermed alle baser for $\mathbb{R}^2$. De første to af disse indeholder den lineært uafhængige familie $(\mathbf{v}_1)$ som en delfamilie, mens dette ikke er tilfældet for den tredje.

4.3 Basis for et vektorrum

Sætning 4.3.15 siger specielt, at enhver endelig familie af vektorer, der frembringer et vektorrum, har en delfamilie, der er en basis for dette vektorrum. Sætningen siger dog *ikke*, at en sådan basis er entydig, og eksempel 4.3.17 viser da også, at dette ikke er tilfældet. Sætningen og dens bevis siger heller ikke noget om, *hvordan* man bærer sig ad med at finde en sådan basis. Dette vender vi tilbage til i afsnit 4.5.

Sætning 4.3.18 *Lad $\mathbb{F}$ være et legeme, og lad V være et $\mathbb{F}$ -vektorrum, der er frembragt af en familie $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ af m vektorer. Hvis $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ er en lineært uafhængig familie af n vektorer i V , så er $n \leq m$.*

Bevis Ifølge sætning 4.3.15 findes der en delfamilie $(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_k})$ af $(\mathbf{v}_1, \dots, \mathbf{v}_m)$, som er en basis for V . Ifølge sætning 4.3.11 er den lineære afbildning $h: \mathbb{F}^k \rightarrow V$ defineret ved

$$h(\mathbf{x}) = \mathbf{v}_{i_1}x_1 + \dots + \mathbf{v}_{i_k}x_k$$

derfor bijektiv. Samme sætning viser, at den lineære afbildning $g: \mathbb{F}^n \rightarrow V$ defineret ved

$$g(\mathbf{y}) = \mathbf{w}_1y_1 + \dots + \mathbf{w}_ny_n$$

er injektiv. Dermed er den sammensatte afbildning $h^{-1} \circ g: \mathbb{F}^n \rightarrow \mathbb{F}^k$ lineær og injektiv, og korollar 2.5.16 viser derfor, at $n \leq k$. Da også $k \leq m$, følger det, at $n \leq m$ som ønsket. $\square$

Eksempel 4.3.19 (1) I eksempel 4.3.6 (1) viste vi, at familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ af vektorer i $\mathbb{R}^2$ ikke er lineært uafhængig. Men dette følger også fra sætning 4.3.18, som viser, at en lineært uafhængig familie af vektorer $\mathbb{R}^2$ aldrig kan bestå af mere end 2 vektorer. For familien $(\mathbf{e}_1, \mathbf{e}_2)$ består af 2 vektorer og frembringer $\mathbb{R}^2$.
(2) Familien $(1, t, t^2, t^3, \dots)$ i vektorrummet $\mathbb{F}[t]$ af alle polynomier er lineært uafhængig. Fordi denne familie er uendelig, viser sætning 4.3.18, at vektorrummet $\mathbb{F}[t]$ *ikke* er endeligt frembragt.

Sætning 4.3.20 *Lad $\mathbb{F}$ være et legeme og V et endeligt frembragt $\mathbb{F}$ -vektorrum. Hvis både $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ er baser for V , så er $m = n$.*

Bevis Da familien $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ frembringer V , og da familien $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ er lineært uafhængig, viser sætning 4.3.18, at $m \geq n$, og uligheden $m \leq n$ vises tilsvarende. $\square$

Sætning 4.3.20 viser, at det følgende dimensionsbegreb er veldefineret.

4 Vektorrum

Definition 4.3.21 Lad V være et endeligt frembragt vektorrum over et legeme $\mathbb{F}$, og lad $(\mathbf{v}_1, \dots, \mathbf{v}_d)$ være en basis for V . Da kaldes det naturlige tal

$$\dim_{\mathbb{F}}(V) = d$$

for dimensionen af V .

Hvis $\mathbb{F}$ er underforstået, så skriver vi $\dim(V)$ i stedet for $\dim_{\mathbb{F}}(V)$.

Eksempel 4.3.22 (1) For ethvert legeme $\mathbb{F}$ har vektorrummet $\mathbb{F}^m$ dimension

$$\dim_{\mathbb{F}}(\mathbb{F}^m) = m.$$

For standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ er en basis for $\mathbb{F}^m$.

(2) Underrummet $N_A \subset \mathbb{F}^5$, som vi betragtede i eksempel 4.3.7, har dimension

$$\dim_{\mathbb{F}}(N_A) = 2.$$

For familien $(\mathbf{v}_1, \mathbf{v}_2)$ fra eksempel 4.3.7 er en basis.

(3) Det reelle vektorrum $\mathbb{C}$, som vi betragtede i eksempel 4.1.3 (5), har dimension

$$\dim_{\mathbb{R}}(\mathbb{C}) = 2.$$

For $(1, i)$ er en basis, da ethvert $z \in \mathbb{C}$ kan skrives entydigt som $z = 1 \cdot a + i \cdot b$ med $a, b \in \mathbb{R}$.

(4) Lad $\mathbb{F}[t]_{\leq d}$ være $\mathbb{F}$ -vektorrummet af polynomier med koefficienter i $\mathbb{F}$ og af grad højst d , som vi betragtede i eksempel 4.1.3 (8). Det har dimension

$$\dim_{\mathbb{F}}(\mathbb{F}[t]_{\leq d}) = d + 1.$$

For familien $(t^0, t^1, \dots, t^d)$ er en basis, da ethvert polynomium $p(t)$ af grad højst d kan skrives entydigt som en linearkombination $p(t) = t^0 \cdot a_0 + t^1 \cdot a_1 + \dots + t^d \cdot a_d$, og indeks-mængden $\{0, 1, \dots, d\}$, som har $d + 1$ elementer.

Vi ønsker nu at vise, at hvis V er et endeligt frembragt vektorrum, og hvis $U \subset V$ er et underrum, så er $\dim(U) \leq \dim(V)$. Det er dog ikke trivielt at vise, at vektorrummet U er endeligt frembragt, så vi viser først, at dette faktisk er tilfældet.

Sætning 4.3.23 Lad $\mathbb{F}$ være et legeme, lad V være et endeligt frembragt $\mathbb{F}$ -vektorrum, og lad $U \subset V$ være et underrum. Da er U endeligt frembragt, og $\dim(U) \leq \dim(V)$.

Bevis Vi viser først, at U er endeligt frembragt. Så vi lader $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ være en familie af vektorer, der frembringer V , og antager, at U ikke er endeligt frembragt. Vi viser da, at der for alle $n \geq 0$ findes en lineært uafhængig familie $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ af vektorer i U , hvilket er i modstrid med sætning 4.3.18. Beviset for denne påstand er ved induktion på $n \geq 0$, og tilfældet $n = 0$ gælder, da den tomme familie er lineært uafhængig. Så vi antager, at påstanden er bevist for $n = p - 1$ og viser den for $n = p$. Per induktion findes der en lineært uafhængig familie $(\mathbf{u}_1, \dots, \mathbf{u}_{p-1})$ af $p - 1$ vektorer i U , og da U per antagelse ikke er endeligt frembragt, er underrummet $U' \subset U$ frembragt af denne familie ikke hele U . Vi kan derfor vælge en vektor $\mathbf{u}_p \in U \setminus U'$, og vi viser nu, at familien $(\mathbf{u}_1, \dots, \mathbf{u}_{p-1}, \mathbf{u}_p)$ af p vektorer i U er lineært uafhængig. Så vi lader

$$\mathbf{u}_1 a_1 + \dots + \mathbf{u}_{p-1} a_{p-1} + \mathbf{u}_p a_p = \mathbf{0}$$

være en linearkombination af $(\mathbf{u}_1, \dots, \mathbf{u}_p)$, som er lig med nulvektoren, og skal da vise, at $a_i = 0$ for alle $1 \leq i \leq p$. Hvis $a_p \neq 0$, da er²

$$\mathbf{u}_p = -(\mathbf{u}_1 a_1 + \dots + \mathbf{u}_{p-1} a_{p-1}) a_p^{-1},$$

hvilket strider mod, at $\mathbf{u}_p \notin U'$. Så $a_p = 0$, og derfor er

$$\mathbf{u}_1 a_1 + \dots + \mathbf{u}_{p-1} a_{p-1} = \mathbf{0},$$

hvoraf det følger, at $a_i = 0$ for $1 \leq i \leq p - 1$, idet $(\mathbf{u}_1, \dots, \mathbf{u}_{p-1})$ er lineært uafhængig. Dette viser induktionsskridtet og dermed påstanden, at der for alle $n \geq 0$ findes en lineært uafhængig familie $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ af vektorer i U . Men denne familie er da også en lineært uafhængig familie af vektorer i V , hvilket er i modstrid med sætning 4.3.18. Vi slutter derfor, at vores antagelse, at U ikke er endeligt frembragt var forkert, hvilket viser, at U er endeligt frembragt som ønsket. Vi anvender endelig sætning 4.3.15 til at vælge baser $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ for U og $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ for V . Familien $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ er da specielt en lineær uafhængig familie af vektorer i V , mens familien $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ af vektorer i V frembringer V . Derfor viser sætning 4.3.18, at

$$\dim(U) = n \leq m = \dim(V)$$

som ønsket. □

Sætning 4.3.24 Lad $\mathbb{F}$ være et legeme, lad V være et endeligt frembragt $\mathbb{F}$ -vektorrum, og lad $d = \dim_{\mathbb{F}}(V)$. Da gælder følgende udsagn:

- (1) Enhver lineært uafhængig familie $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ af vektorer i V er en delfamilie af en basis $(\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}, \dots, \mathbf{v}_d)$ for V .
- (2) Enhver familie $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ af vektorer i V , der frembringer V , har en delfamilie $(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_d})$, der er en basis for V .

² Bemærk, at vi anvender, at $\mathbb{F}$ er et legeme, sådan at det for alle elementer $a \in \mathbb{F}$ gælder, at hvis $a \neq 0$, da er a invertibel.

4 Vektorrum

Bevis For at bevise (1) lader vi $(\mathbf{v}_{m+1}, \dots, \mathbf{v}_p)$ være en endelig familie af vektorer i V , der frembringer V . Da frembringer familien $(\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}, \dots, \mathbf{v}_p)$ ligeledes V , og den indeholder den lineært uafhængige familie $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ som en delfamilie. Sætning 4.3.15 viser derfor, at der findes en delmængde

$$\{1, 2, \dots, m\} \subset J \subset \{1, 2, \dots, m, m+1, \dots, p\},$$

sådan at delfamilien $(\mathbf{v}_i)_{i \in J}$ er en basis for V , og ved om nødvendigt at reindicere, kan vi antage, at $J = \{1, 2, \dots, m, m+1, \dots, d\}$. Dette viser (1).

Endelig følger (2) umiddelbart fra sætning 4.3.15, som viser, at der findes

$$\emptyset \subset \{i_1, \dots, i_d\} \subset \{1, 2, \dots, n\},$$

sådan at delfamilien $(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_d})$ er en basis for V . □

Vi illustrerer de sidste to sætninger med følgende simple eksempler.

Eksempel 4.3.25 Vi betragter igen underrummet fra eksempel 4.1.5,

$$U = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 = x_2 \right\} \subset V = \mathbb{R}^2.$$

Da V er endeligt frembragt af dimension 2, så viser sætning 4.3.23, at U er endeligt frembragt af dimension højst 2, og det er da heller ikke svært at se, at familien

$$(\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix})$$

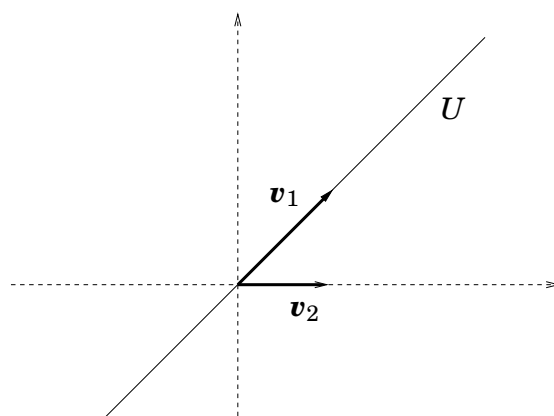
er en basis for U . Så U har dimension 1. Hvis vi betragter $(\mathbf{v}_1)$ som en familie af vektorer i V , da er denne familie derfor lineært uafhængig. Ifølge sætning 4.3.24 findes der derfor en basis for V , som indeholder $(\mathbf{v}_1)$ som en delfamilie. Standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for V har ikke denne egenskab, da $\mathbf{v}_1$ hverken er lig med $\mathbf{e}_1$ eller $\mathbf{e}_2$. Derimod er

$$(\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix})$$

en basis for V , som indeholder $(\mathbf{v}_1)$ som en delfamilie. For eksempel er

$$\det(\mathbf{v}_1 \ \mathbf{v}_2) = \det \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = -1 \neq 0,$$

og lemma 4.3.9 viser derfor, at $(\mathbf{v}_1, \mathbf{v}_2)$ er en basis for V . Se også figur 4.3.



Figur 4.3: Underrum $U \subset V = \mathbb{R}^2$ og basis for V som indholder en basis for U .

Bemærkning 4.3.26 Vi skitserer her, hvordan begreber og sætninger i dette afsnit udvides til vektorrum, der ikke antages at være endeligt frembragte.

Vi definerer en linearkombination af en familie af vektorer i et vektorrum som følger. Hvis $a = (a_i)_{i \in I}$ er en familie af skalarer, da definerer vi dens *support* til at være

$$\text{supp}(a) = \{i \in I \mid a_i \neq 0\} \subset I,$$

og vi siger, at $a = (a_i)_{i \in I}$ har *endelig support*, hvis $\text{supp}(a) \subset I$ er en endelig delmængde. Vi lader nu V være et vektorrum og lader $(v_i)_{i \in I}$ være en familie af vektorer i V . Hvis $(a_i)_{i \in I}$ er en familie af skalarer, der er indiceret ved den samme mængde I og som har endelig support, da definerer vi

$$\sum_{i \in I} v_i a_i = \sum_{i \in \text{supp}(a)} v_i a_i$$

og kalder denne sum for en *linearkombination* af familien $(v_i)_{i \in I}$. Vi bemærker, at mens højresiden er en endelig sum og derfor meningsfuld, er venstresiden blot et symbol, der ved denne ligning defineres til at være lig med højresiden. Med denne definition af en linearkombination kan definition 4.3.4 nu udvides til vilkårlige familier mutatis mutandis.

Den lineære algebras hovedsætning gælder generelt og siger, at hvis $(v_i)_{i \in I}$ er en familie af vektorer i V , der frembringer V , og hvis $(v_i)_{i \in K}$ er en lineært uafhængig delfamilie, da findes $K \subset J \subset I$, sådan at delfamilien $(v_i)_{i \in J}$ er en basis for V . Man kan dog ikke bevise denne sætning uden at anvende Zorns lemma.³ Et vektorrum V har altid en familie af vektorer, der frembringer V , nemlig identitetsfamilien $(v)_{v \in V}$, og den tomme familie er en lineært uafhængig delfamilie heraf. Dermed har V en basis, som kan være (vilkårligt) uendelig.

³ A. Blass. *Existence of bases implies the axiom of choice*. Axiomatic Set Theory (Boulder, CO, 1983), 31–33, Contemp. Math., 31, Amer. Math. Soc., Providence, RI, 1984.

4 Vektorrum

Man kan også vise, at to baser for det samme vektorrum nødvendigvis består af det samme “antal” vektorer. Mere præcist, hvis $(\mathbf{v}_i)_{i \in I}$ og $(\mathbf{w}_j)_{j \in J}$ er to baser for V , da har mængderne I og J samme kardinalitet⁴ α , og man definerer da dimensionen af V til at være $\dim_{\mathbb{F}}(V) = \alpha$. For eksempel har $\mathbb{F}[t]$ dimension $\aleph_0$, hvilket er kardinaliteten af de naturlige tal, mens vektorrummet, der består af alle følger $(a_i)_{i \in \mathbb{N}}$ af skalarer $a_i \in \mathbb{F}$, har dimension $2^{\aleph_0}$, hvilket er kardinaliteten af de reelle tal.

4.4 Matrixrepræsentation af lineære afbildninger

Som Descartes har lært os, er det lettest at lave udregninger ved at bruge koordinater. Vi indfører nu koordinater for endeligt frembragte vektorrum og matrixrepræsentation for lineære afbildninger mellem endeligt frembragte vektorrum. Det er dog vigtigt at huske, at begge begreber afhænger af valg af baser.

Vi viste i lemma 4.2.3, at en n -tuple $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ af vektorer i et vektorrum V entydigt bestemmer en lineær afbildning $h: \mathbb{F}^n \rightarrow V$, sådan at $h(\mathbf{e}_i) = \mathbf{v}_i$ for alle $1 \leq i \leq n$, og vi viste i sætning 4.3.11, at denne afbildning er en isomorfi, hvis og kun hvis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er en basis for V .

Definition 4.4.1 Lad V være et endeligt frembragt vektorrum over et legeme $\mathbb{F}$. Lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ være en basis for V , og lad $h: \mathbb{F}^n \rightarrow V$ være den lineære isomorfi givet ved

$$h(\mathbf{x}) = \mathbf{v}_1 x_1 + \mathbf{v}_2 x_2 + \dots + \mathbf{v}_n x_n.$$

Givet en vektor $\mathbf{v} \in V$, da kaldes den entydigt bestemte søjlevektor $\mathbf{x} \in \mathbb{F}^n$, sådan at $h(\mathbf{x}) = \mathbf{v}$ for koordinaterne af $\mathbf{v} \in V$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$.

Eksempel 4.4.2 (1) Vi betragter igen basen

$$\left(\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$$

for $V = \mathbb{R}^2$ fra eksempel 4.3.6 (1). Den entydigt bestemte lineære isomorfi $h: \mathbb{R}^2 \rightarrow V$, som opfylder, at $h(\mathbf{e}_1) = \mathbf{v}_1$ og $h(\mathbf{e}_2) = \mathbf{v}_2$, er da givet ved

$$h(\mathbf{x}) = \mathbf{v}_1 x_1 + \mathbf{v}_2 x_2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ x_1 + x_2 \end{pmatrix}.$$

Vi ønsker at bestemme koordinaterne $\mathbf{x} \in \mathbb{R}^2$ af vektoren

$$\mathbf{v} = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \in V$$

⁴ Kenneth Kunen. Set Theory. Studies in Logic (London), 34. College Publications, London, 2011. Kardinaliteten af en mængde er defineret i Definition I.10.11.

4.4 Matrixrepræsentation af lineære afbildninger

med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2)$. Per definition er $\mathbf{x} \in \mathbb{R}^2$ den entydigt bestemte søjlevektor, sådan at $h(\mathbf{x}) = \mathbf{v}$. Denne søjlevektor er altså løsningen til ligningssystemet

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix},$$

som er

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Vi har dermed udregnet, at $\mathbf{v} = \mathbf{v}_1 \cdot 2 + \mathbf{v}_2 \cdot 1$.

(2) Vi betragter vektorrummet $V = \mathbb{F}[t]_{\leq d}$ af polynomier med koefficienter i $\mathbb{F}$ og af grad højst d fra eksempel 4.1.3. Familien $(t^0, t^1, \dots, t^d)$ er en basis for V , og den entydige lineære isomorfi $h: \mathbb{F}^{d+1} \rightarrow V$, sådan at $h(\mathbf{e}_i) = t^{i-1}$ for alle $1 \leq i \leq d+1$, er givet ved

$$h(\mathbf{x}) = t^0 \cdot x_1 + t^1 \cdot x_2 + \dots + t^d \cdot x_{d+1} = x_1 + x_2 t + \dots + x_{d+1} t^d.$$

Vi aflæser heraf, at koordinaterne af vektoren $p(t) = a_0 + a_1 t + \dots + a_d t^d \in V$ med hensyn til basen $(t^0, t^1, \dots, t^d)$ for V er lig med søjlevektoren

$$\mathbf{x} = \begin{pmatrix} a_0 \\ \vdots \\ a_d \end{pmatrix} \in \mathbb{F}^{d+1}.$$

Vi bemærker det forvirrende indeksskift, som skyldes, at vi i dette afsnit kun tillader baser at være indiceret ved $I = \{1, 2, \dots, n\}$ og ikke ved for eksempel $I = \{0, 1, \dots, d\}$.

(3) Vi lader $V = \{|\uparrow\rangle z_1 + |\downarrow\rangle z_2 \mid z_1, z_2 \in \mathbb{C}\}$ være $\mathbb{C}$ -vektorrummet fra eksempel 4.1.3 (7) og betragter basen $(\mathbf{v}_1 = |\uparrow\rangle^{1/\sqrt{2}} + |\downarrow\rangle^{1/\sqrt{2}}, \mathbf{v}_2 = |\uparrow\rangle^{1/\sqrt{2}} - |\downarrow\rangle^{1/\sqrt{2}})$ herfor. Den entydige lineære afbildning $h: \mathbb{C}^2 \rightarrow V$, sådan at $h(\mathbf{e}_1) = \mathbf{v}_1$ og $h(\mathbf{e}_2) = \mathbf{v}_2$, er da givet ved

$$\begin{aligned} h(\mathbf{w}) &= \mathbf{v}_1 w_1 + \mathbf{v}_2 w_2 = (|\uparrow\rangle \frac{1}{\sqrt{2}} + |\downarrow\rangle \frac{1}{\sqrt{2}}) w_1 + (|\uparrow\rangle \frac{1}{\sqrt{2}} - |\downarrow\rangle \frac{1}{\sqrt{2}}) w_2 \\ &= |\uparrow\rangle (\frac{1}{\sqrt{2}} w_1 + \frac{1}{\sqrt{2}} w_2) + |\downarrow\rangle (\frac{1}{\sqrt{2}} w_1 - \frac{1}{\sqrt{2}} w_2). \end{aligned}$$

Koordinaterne $\mathbf{w} \in \mathbb{C}^2$ af vektoren $\mathbf{v} = |\uparrow\rangle^{1/\sqrt{2}} + |\downarrow\rangle^{i/\sqrt{2}}$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2)$ er dermed den entydige løsning til ligningssystemet

$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix},$$

som er

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}^{-1} \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1/2 + i/2 \\ 1/2 - i/2 \end{pmatrix}.$$

4 Vektorrum

Vi skal nu vise, at givet to endeligt frembragte vektorrum samt baser for disse, da kan enhver lineær afbildning mellem disse vektorrum på entydig vis repræsenteres ved en matrix. Det er vigtigt at huske denne sætning *helt præcist*, da den danner grundlaget for alle beregninger, der har med lineære afbildninger at gøre.

Sætning 4.4.3 *Lad $\mathbb{F}$ være et legeme, og lad $f: V \rightarrow W$ være en lineær afbildning fra et n -dimensionalt $\mathbb{F}$ -vektorrum til et m -dimensionalt $\mathbb{F}$ -vektorrum. Lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ være baser for henholdsvis V og W , og lad*

$$A = (a_{ij}) \in M_{m,n}(\mathbb{F})$$

være matricen, hvis j 'te søjle er koordinaterne for $f(\mathbf{v}_j) \in W$ med hensyn til $(\mathbf{w}_1, \dots, \mathbf{w}_m)$. Hvis $\mathbf{v} \in V$ og $\mathbf{w} = f(\mathbf{v}) \in W$, og hvis $\mathbf{x} \in \mathbb{F}^n$ og $\mathbf{y} \in \mathbb{F}^m$ er koordinaterne for $\mathbf{v} \in V$ med hensyn til $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og for $\mathbf{w} \in W$ med hensyn til $(\mathbf{w}_1, \dots, \mathbf{w}_m)$, da er

$$\mathbf{y} = A\mathbf{x}.$$

Matricen $A \in M_{m,n}(\mathbb{F})$ er endvidere entydigt bestemt ved denne egenskab.

Bevis Ifølge definition 4.4.1 kan vi skrive $\mathbf{v} \in V$ entydigt som

$$\mathbf{v} = \sum_{j=1}^n \mathbf{v}_j x_j,$$

hvor $\mathbf{x} \in \mathbb{F}^n$ er koordinaterne af $\mathbf{v}$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$. Vi anvender f på denne ligning og udnytter henholdsvis (L1) og (L2) til at omskrive højresiden:

$$f(\mathbf{v}) = f\left(\sum_{j=1}^n \mathbf{v}_j x_j\right) = \sum_{j=1}^n f(\mathbf{v}_j x_j) = \sum_{j=1}^n f(\mathbf{v}_j) x_j.$$

Ifølge definition 4.4.1 kan vi ligeledes skrive vektoren $f(\mathbf{v}_j) \in W$ entydigt som

$$f(\mathbf{v}_j) = \sum_{i=1}^m \mathbf{w}_i a_{ij},$$

hvor

$$\mathbf{a}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{F}^m$$

er koordinaterne af $f(\mathbf{v}_j)$ med hensyn til basen $(\mathbf{w}_1, \dots, \mathbf{w}_m)$. Vi substituerer disse sidste ligninger i udtrykket for $\mathbf{w} = f(\mathbf{v})$ overfor, hvilket giver

$$\mathbf{w} = \sum_{j=1}^n \left(\sum_{i=1}^m \mathbf{w}_i a_{ij} \right) x_j = \sum_{i=1}^m \mathbf{w}_i \left(\sum_{j=1}^n a_{ij} x_j \right).$$

4.4 Matrixrepræsentation af lineære afbildninger

Da koordinaterne $\mathbf{y} \in \mathbb{F}^m$ af vektoren $\mathbf{w}$ med hensyn til basen $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ er entydigt bestemt, aflæser vi af denne ligning, at

$$y_i = \sum_{j=1}^n a_{ij} x_j$$

for alle $1 \leq i \leq m$. Dermed har vi vist, at $\mathbf{y} = A\mathbf{x}$ som ønsket, hvor $A = (a_{ij})$. Endelig er matricen A entydigt bestemt, idet koordinaterne $\mathbf{a}_j \in \mathbb{F}^m$ for vektoren $f(\mathbf{v}_j) \in W$ med hensyn til basen $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ er entydigt bestemte. $\square$

Sætning 4.4.3 giver en 1-1-korrespondance mellem lineære afbildninger $f: V \rightarrow W$ og matricer $A \in M_{m,n}(\mathbb{F})$. Denne 1-1-korrespondance *afhænger af valget af baser* $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V og $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for W , og vi indfører derfor følgende terminologi.

Definition 4.4.4 Lad $\mathbb{F}$ være et legeme, og lad $f: V \rightarrow W$ være en lineær afbildning fra et n -dimensionalt $\mathbb{F}$ -vektorrum til et m -dimensionalt $\mathbb{F}$ -vektorrum. Lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ være baser for henholdsvis V og W . Da kaldes matricen

$$A = (a_{ij}) \in M_{m,n}(\mathbb{F}),$$

hvis j 'te søjle er koordinaterne for $f(\mathbf{v}_j) \in W$ med hensyn til basen $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for W , for matricen, der repræsenterer $f: V \rightarrow W$ med hensyn til baserne $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V og $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for W .

Bemærkning 4.4.5 Vi kan reformulere definitionen af den matrix $A \in M_{m,n}(\mathbb{F})$, der repræsenterer den lineære afbildning $f: V \rightarrow W$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V og basen $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for W som følger. Lad $h: \mathbb{F}^n \rightarrow V$ og $k: \mathbb{F}^m \rightarrow W$ være de entydigt bestemte lineære isomorfier, sådan at $h(\mathbf{e}_j) = \mathbf{v}_j$ for alle $1 \leq j \leq n$ og $k(\mathbf{e}_i) = \mathbf{w}_i$ for alle $1 \leq i \leq m$. Der findes netop én lineær afbildning $g: \mathbb{F}^n \rightarrow \mathbb{F}^m$, sådan at diagrammet

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \uparrow h & & \uparrow k \\ \mathbb{F}^n & \xrightarrow{g} & \mathbb{F}^m \end{array}$$

kommutterer, nemlig, $g = k^{-1} \circ f \circ h$, og vi påstår, at $A \in M_{m,n}(\mathbb{F})$ også repræsenterer afbildningen $g: \mathbb{F}^n \rightarrow \mathbb{F}^m$ med hensyn til standardbaserne. For da $f \circ h = k \circ g$, så gælder det for alle $1 \leq j \leq n$, at

$$g(\mathbf{e}_j) = \mathbf{e}_1 a_{1j} + \mathbf{e}_2 a_{2j} + \dots + \mathbf{e}_m a_{mj},$$

hvis og kun hvis

$$f(\mathbf{v}_j) = \mathbf{w}_1 a_{1j} + \mathbf{w}_2 a_{2j} + \dots + \mathbf{w}_m a_{mj},$$

hvilket viser påstanden.

Eksempel 4.4.6 (1) Lad $\mathbb{F}$ være et legeme, og betragt vektorerne $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{F}^2$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 7 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Familien $(\mathbf{v}_1, \mathbf{v}_2)$ er en basis for $\mathbb{F}^2$, og matricen $P \in M_2(\mathbb{F})$, der repræsenterer identitetsafbildningen $\text{id}: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for codomænet, er per definition givet ved

$$P = \begin{pmatrix} 7 & 3 \\ 2 & 1 \end{pmatrix}.$$

Søjlene i P er nemlig koordinaterne af vektorerne $\mathbf{v}_1 = \text{id}(\mathbf{v}_1)$ og $\mathbf{v}_2 = \text{id}(\mathbf{v}_2)$ med hensyn til standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for $\mathbb{F}^2$. Denne matrix er altså *ikke* identitetsmatricen, hvilket skyldes, at vi ikke anvender den samme basis for domænet og codomænet.

(2) Matricen $A \in M_n(\mathbb{F})$, der repræsenterer identitetsafbildning $\text{id}: V \rightarrow V$ med hensyn til den *samme* basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for både domænet og codomænet, er identitetsmatricen $A = I_n$. Per definition er den j 'te søjle i A nemlig koordinaterne af $\text{id}(\mathbf{v}_j) = \mathbf{v}_j$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$, og udregningen

$$\mathbf{v}_j = \sum_{i < j} \mathbf{v}_i \cdot 0 + \mathbf{v}_j \cdot 1 + \sum_{i > j} \mathbf{v}_i \cdot 0$$

viser, at koordinaterne af $\mathbf{v}_j$ med hensyn til denne basis er lig med $\mathbf{e}_j \in \mathbb{F}^n$.

(3) Vi betragter den lineære afbildning $D: \mathbb{F}[t]_{\leq 2} \rightarrow \mathbb{F}[t]_{\leq 1}$, der til et polynomium

$$p(t) = a_0 + a_1 t + a_2 t^2$$

tilordner det formelt afledte polynomium

$$D(p(t)) = p'(t) = a_1 + 2a_2 t.$$

Vi aflæser da fra udregningen

$$D(t^0) = 0 = t^0 \cdot 0 + t^1 \cdot 0$$

$$D(t^1) = 1 = t^0 \cdot 1 + t^1 \cdot 0$$

$$D(t^2) = 2t = t^0 \cdot 0 + t^1 \cdot 2,$$

at den matrix $A \in M_{2,3}(\mathbb{F})$, der repræsenterer D med hensyn til basen (t^0, t^1, t^2) for domænet og (t^0, t^1) for codomænet, er givet ved

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

4.4 Matrixrepræsentation af lineære afbildninger

Bemærkning 4.4.7 Vi vil symbolisere, at $f: V \rightarrow W$ er repræsenteret af $A \in M_{m,n}(\mathbb{F})$ med hensyn til baserne $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V og $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for W med figuren

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ (\mathbf{v}_1, \dots, \mathbf{v}_n) & \xrightarrow{A} & (\mathbf{w}_1, \dots, \mathbf{w}_m). \end{array}$$

Denne notation er ikke standard, men som vi skal se nedenfor, gør den det lettere at overskue, hvordan sammensætning af lineære afbildninger oversættes til matrixsprog. Vi understreger, at den prikkede pil ikke betegner en afbildning.

Eksempel 4.4.8 Med notationen fra bemærkning 4.4.7 udtrykker vi den situation, vi betragtede i eksempel 4.4.6, som følger.

$$\begin{array}{ccc} \mathbb{F}^2 & \xrightarrow{\text{id}} & \mathbb{F}^2 \\ (\mathbf{v}_1, \mathbf{v}_2) & \xrightarrow{P} & (\mathbf{e}_1, \mathbf{e}_2) \end{array} \qquad \begin{array}{ccc} \mathbb{F}[t]_{\leq 2} & \xrightarrow{D} & \mathbb{F}[t]_{\leq 1} \\ (t^0, t^1, t^2) & \xrightarrow{A} & (t^0, t^1) \end{array}$$

Sætning 4.4.9 Lad $\mathbb{F}$ være et legeme, lad U , V og W være $\mathbb{F}$ -vektorum af dimension henholdsvis p , n og m , og lad $(\mathbf{u}_1, \dots, \mathbf{u}_p)$, $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ være baser for henholdsvis U , V og W . Lad $f: V \rightarrow W$ og $g: U \rightarrow V$ være lineære afbildninger, og lad $A \in M_{m,n}(\mathbb{F})$, $B \in M_{n,p}(\mathbb{F})$ og $C \in M_{m,p}(\mathbb{F})$ være matricerne, der repræsenterer henholdsvis $f: V \rightarrow W$, $g: U \rightarrow V$ og $f \circ g: U \rightarrow W$ med hensyn til de givne baser. Da er

$$C = AB.$$

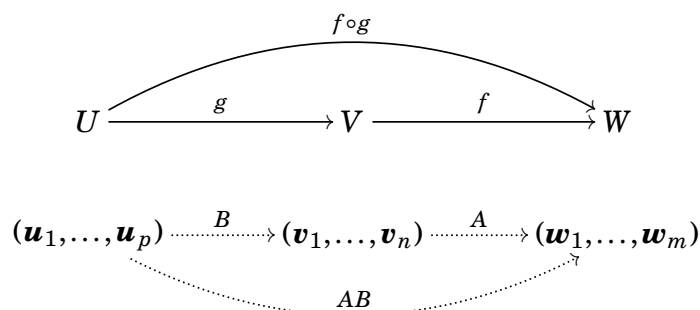
Bevis For $\mathbf{u} \in U$ lader vi $\mathbf{x} \in \mathbb{F}^p$, $\mathbf{y} \in \mathbb{F}^n$ og $\mathbf{z} \in \mathbb{F}^m$ være koordinaterne for henholdsvis $\mathbf{u} \in U$ med hensyn til basen $(\mathbf{u}_1, \dots, \mathbf{u}_p)$, $\mathbf{v} = g(\mathbf{u}) \in V$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$, og $\mathbf{w} = f(g(\mathbf{u})) \in W$ med hensyn til basen $(\mathbf{w}_1, \dots, \mathbf{w}_m)$. Ifølge sætning 4.4.3 er

$$\mathbf{z} = A\mathbf{y} = A(B\mathbf{x}) = (AB)\mathbf{x},$$

og da matricen C , der repræsenterer $f \circ g: U \rightarrow W$ med hensyn til baserne $(\mathbf{u}_1, \dots, \mathbf{u}_p)$ for U og $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for W , er entydigt bestemt, konkluderer vi, at $C = AB$ som ønsket. Vi bemærker, at vi i udregningen ovenfor har anvendt det ikke-trivielle faktum, at den associative lov gælder for matrixmultiplikation, hvilket vi viste i sætning 2.1.10. $\square$

4 Vektorrum

Bemærkning 4.4.10 Med notationen fra bemærkning 4.4.7 kan vi udtrykke udsagnet i sætning 4.4.9 ved følgende diagram.



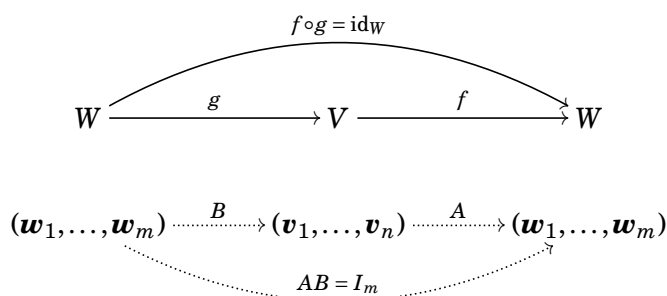
Vi bemærker, at den rækkefølge, hvori de to afbildninger f og g sammensættes, er den samme som den rækkefølge, hvori de to matricer A og B , der repræsenterer dem med hensyn til de valgte baser, multipliceres.

Korollar 4.4.11 Lad $\mathbb{F}$ være et legeme, og lad $f: V \rightarrow W$ være en lineær afbildning fra et n -dimensionalt $\mathbb{F}$ -vektorrum V til et m -dimensionalt $\mathbb{F}$ -vektorrum W . Hvis $A \in M_{m,n}(\mathbb{F})$ er matricen, der repræsenterer $f: V \rightarrow W$ med hensyn til baser $(v_1, \dots, v_n)$ for V og $(w_1, \dots, w_m)$ for W , da er de følgende udsagn ækvivalente:

- (1) Der findes en lineær afbildning $g: W \rightarrow V$, sådan at $f \circ g = \text{id}_W$ og $g \circ f = \text{id}_V$.
- (2) Der findes en matrix $B \in M_{n,m}(\mathbb{F})$, sådan at $AB = I_m$ og $BA = I_n$.

I givet fald repræsenterer matricen $B \in M_{n,m}(\mathbb{F})$ afbildningen $g: W \rightarrow V$ med hensyn til basen $(w_1, \dots, w_m)$ for W og $(v_1, \dots, v_n)$ for V .

Bevis Vi antager først (1), og for at vise (2) betragter vi følgende figur.



Lad $B \in M_{n,m}(\mathbb{F})$ være matricen, der repræsenterer $g: W \rightarrow V$ med hensyn til baserne $(w_1, \dots, w_m)$ for W og $(v_1, \dots, v_n)$ for V . Ifølge sætning 4.4.9 er $AB \in M_m(\mathbb{F})$ da matricen, der repræsenterer $f \circ g = \text{id}_W: W \rightarrow W$ med hensyn til den samme basis $(w_1, \dots, w_m)$ for både domænet og codomænet, og denne matrix er derfor lig med identitetsmatricen I_m . Vi ser tilsvarende, at $g \circ f = \text{id}_V$ medfører, at $BA = I_n$, så (2) gælder.

4.4 Matrixrepræsentation af lineære afbildninger

Omvendt, hvis (2) gælder, så lader vi $g: W \rightarrow V$ være den lineære afbildning, der er repræsenteret af $B \in M_{n,m}(\mathbb{F})$ med hensyn til baserne $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for W og $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V . Ifølge sætning 4.4.9 er $f \circ g: W \rightarrow W$ den lineære afbildning, der er repræsenteret af $AB = I_m$ med hensyn til den *samme* basis $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for domænet og codomænet, og denne lineære afbildning er derfor lig med identitetsafbildningen id_W . Beviset for at $BA = I_n$ medfører, at $g \circ f = \text{id}_V$, er tilsvarende, så (1) følger. $\square$

Hvis de ækvivalente udsagn (1) og (2) i korollar 4.4.11 gælder, da er $m = n$, idet en invertibel matrix altid er kvadratisk. Vi viser nu følgende mere præcise resultat.

Sætning 4.4.12 *Lad $\mathbb{F}$ være et legeme, og lad $f: V \rightarrow W$ være en lineær afbildning mellem $\mathbb{F}$ -vektorum. Da er f en isomorfi hvis og kun hvis det for enhver basis $(\mathbf{v}_i)_{i \in I}$ for V gælder, at familien $(f(\mathbf{v}_i))_{i \in I}$ er en basis for W . I givet fald har V og W samme dimension.*

Bevis Vi antager først, at f er en isomorfi. Vi lader $(\mathbf{v}_i)_{i \in I}$ være en basis for V og skal da vise, at $(f(\mathbf{v}_i))_{i \in I}$ er en basis for W . Så lad $g: W \rightarrow V$ være den inverse afbildning af $f: V \rightarrow W$. Givet en vektor $\mathbf{w} \in W$, da kan vi skrive $\mathbf{v} = g(\mathbf{w}) \in V$ som en linearkombination $\mathbf{v} = \sum_{i \in I} \mathbf{v}_i a_i$. Det følger derfor, at

$$\mathbf{w} = f(\mathbf{v}) = f\left(\sum_{i \in I} \mathbf{v}_i a_i\right) = \sum_{i \in I} f(\mathbf{v}_i) a_i,$$

hvilket viser, at familien $(f(\mathbf{v}_i))_{i \in I}$ frembringer W . Og hvis en linearkombination

$$\sum_{i \in I} f(\mathbf{v}_i) a_i$$

er lig med nulvektoren, da er også

$$\sum_{i \in I} \mathbf{v}_i a_i = \sum_{i \in I} g(f(\mathbf{v}_i)) a_i = g\left(\sum_{i \in I} f(\mathbf{v}_i) a_i\right)$$

lig med nulvektoren, og da $(\mathbf{v}_i)_{i \in I}$ er lineært uafhængig, er alle a_i derfor lig med nul. Dette viser, at familien $(f(\mathbf{v}_i))_{i \in I}$ også er lineært uafhængig og dermed en basis for W . Specielt har V og W derfor samme dimension. Omvendt, hvis $(\mathbf{v}_i)_{i \in I}$ er en basis for V og $(f(\mathbf{v}_i))_{i \in I}$ en basis for W , da giver formelen $g(\sum_{i \in I} f(\mathbf{v}_i) a_i) = \sum_{i \in I} \mathbf{v}_i a_i$ en veldefineret og lineær afbildning $g: W \rightarrow V$. Per definition er $f \circ g = \text{id}_W$ og $g \circ f = \text{id}_V$, hvilket som ønsket viser, at $f: V \rightarrow W$ er en isomorfi. $\square$

Eksempel 4.4.13 Vi har i eksempel 4.4.6 bestemt matricen $P \in M_2(\mathbb{F})$, der repræsenterer identitetsafbildningen $\text{id}: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2)$ for domænet og

4 Vektorrum

standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for codomænet. Eftersom identitetsafbildningen er sin egen inverse afbildning, viser korollar 4.4.11, at matricen, der repræsenterer $\text{id}: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ med hensyn til standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for domænet og basen $(\mathbf{v}_1, \mathbf{v}_2)$ for codomænet er

$$P^{-1} = \begin{pmatrix} 1 & -3 \\ -2 & 7 \end{pmatrix}.$$

Her har vi brugt eksempel 3.4.4 til at udregne den inverse matrix.

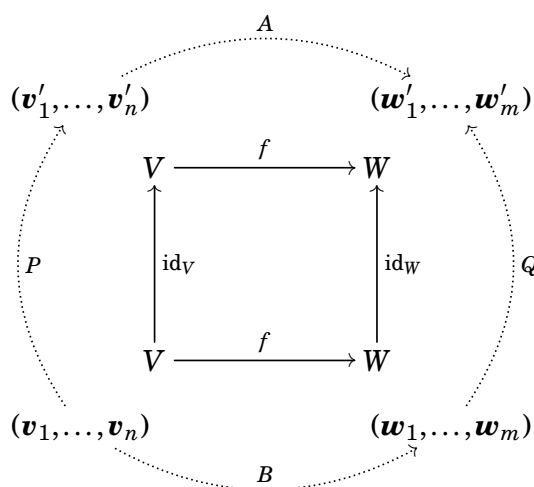
Vi skal nu se, hvordan matricen, der repræsenterer en lineær afbildning med hensyn til ét valg af baser, bestemmer matricen, der repræsenterer den samme afbildning med hensyn til et andet valg af baser. Vi opfordrer læseren til at huske det simple bevis for den følgende sætning i stedet for det komplicerede udsagn.

Sætning 4.4.14 *Lad $\mathbb{F}$ være et legeme, og lad $f: V \rightarrow W$ være en lineær afbildning fra et n -dimensionalt $\mathbb{F}$ -vektorrum V til et m -dimensionalt $\mathbb{F}$ -vektorrum W . Lad $A \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f: V \rightarrow W$ med hensyn til basen $(\mathbf{v}'_1, \dots, \mathbf{v}'_n)$ for V og $(\mathbf{w}'_1, \dots, \mathbf{w}'_m)$ for W , og lad $B \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f: V \rightarrow W$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V og $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for W . Da er*

$$B = Q^{-1}AP,$$

hvor $P \in M_n(\mathbb{F})$ repræsenterer $\text{id}_V: V \rightarrow V$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for domænet og basen $(\mathbf{v}'_1, \dots, \mathbf{v}'_n)$ for codomænet, og hvor $Q \in M_m(\mathbb{F})$ repræsenterer $\text{id}_W: W \rightarrow W$ med hensyn til basen $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ for domænet og basen $(\mathbf{w}'_1, \dots, \mathbf{w}'_m)$ for codomænet.

Bevis Den følgende figur illustrerer den information, vi har til rådighed.



4.4 Matrixrepræsentation af lineære afbildninger

Ifølge sætning 4.4.9 er matricen C , der repræsenterer $\text{id}_W \circ f = f \circ \text{id}_V: V \rightarrow W$ med hensyn til baserne $(v'_1, \dots, v'_n)$ for V og $(w_1, \dots, w_m)$ for W , lig med

$$C = QB = AP,$$

og da identitetsafbildningen $\text{id}_W: W \rightarrow W$ er en isomorfi, følger det fra korollar 4.4.11, at matricen Q er invertibel. Derfor er $B = Q^{-1}AP$ som ønsket. $\square$

Bemærkning 4.4.15 Vi aflæser udsagnet i sætning 4.4.14 fra figuren i beviset som følger. Vi ønsker at skrive den prikkede pil “B” som en sammensætning af de øvrige prikkede pile og bemærker, at dette kræver, at pilen “Q” vendes om. Pilen i den modsatte retning er da “ Q^{-1} ”, og den findes, fordi Q er invertibel.

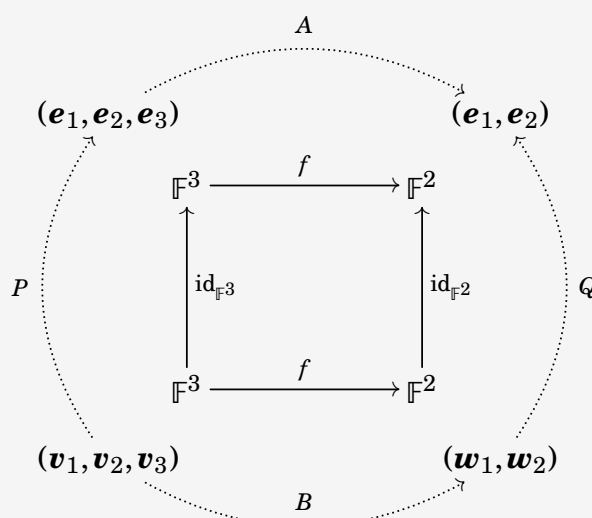
Eksempel 4.4.16 Vi betragter den lineære afbildning $f: \mathbb{F}^3 \rightarrow \mathbb{F}^2$, der er repræsenteret med hensyn til standardbaserne (e_1, e_2, e_3) for $\mathbb{F}^3$ og (e_1, e_2) for $\mathbb{F}^2$ af matricen

$$A = \begin{pmatrix} 2 & 4 & 1 \\ 1 & -1 & 0 \end{pmatrix} \in M_{2,3}(\mathbb{F}),$$

og vi ønsker at bestemme matricen $B \in M_{2,3}(\mathbb{F})$, der repræsenterer $f: \mathbb{F}^3 \rightarrow \mathbb{F}^2$ med hensyn til de nye baser

$$\left(v_1 = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right) \quad \text{og} \quad \left(w_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right)$$

for henholdsvis $\mathbb{F}^3$ og $\mathbb{F}^2$. Vi betragter derfor den følgende figur.



4 Vektorrum

Vi ser direkte fra definitionen af de nye basisvektorer, at

$$P = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{F}) \quad \text{og} \quad Q = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} \in M_2(\mathbb{F}),$$

og vi konkluderer derfor fra sætning 4.4.14, at

$$B = Q^{-1}AP = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -8 & -13 & -3 \\ 5 & 6 & 2 \end{pmatrix}.$$

Her har vi anvendt eksempel 3.4.4 til at udregne Q^{-1} .

Korollar 4.4.17 *Lad $\mathbb{F}$ være et legeme, og lad V være et n -dimensionalt $\mathbb{F}$ -vektorrum. Lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ være baser for V , og lad $P \in M_n(\mathbb{F})$ være matricen, der repræsenterer $\text{id}_V: V \rightarrow V$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for domænet og basen $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ for codomænet. Hvis $\mathbf{x} \in \mathbb{F}^n$ og $\mathbf{y} \in \mathbb{F}^n$ er koordinaterne af en vektor $\mathbf{v} \in V$ med hensyn til henholdsvis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_n)$, da er*

$$\mathbf{y} = P\mathbf{x}.$$

Bevis Påstanden følger umiddelbart af sætning 4.4.3 anvendt på $\text{id}_V: V \rightarrow V$. □

Eksempel 4.4.18 Vi betragter en vektor

$$\mathbf{y} = \mathbf{e}_1 y_1 + \mathbf{e}_2 y_2 + \mathbf{e}_3 y_3 \in \mathbb{F}^3$$

og ønsker at bestemme koordinaterne $\mathbf{x} \in \mathbb{F}^3$ af denne vektor med hensyn til den nye basis $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for $\mathbb{F}^3$ fra eksempel 4.4.16. Ifølge korollar 4.4.17 gælder det da, at

$$\mathbf{y} = P\mathbf{x},$$

hvor P er matricen, der repræsenterer $\text{id}: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for

4.4 Matrixrepræsentation af lineære afbildninger

domænet og basen (e_1, e_2, e_3) for codomænet. Vi bestemmer derfor P^{-1} ,

$$\begin{aligned} (P | I) &= \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ \color{red}{3} & 1 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} \\ \\ +(-1)R_1 \end{array} \\ &= \left(\begin{array}{ccc|ccc} \color{red}{2} & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 & 0 & 1 \end{array} \right) \begin{array}{l} \\ +(-2)R_3 \\ \end{array} \\ &= \left(\begin{array}{ccc|ccc} 0 & \color{red}{-2} & 1 & 3 & 0 & -2 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & \color{red}{1} & 0 & -1 & 0 & 1 \end{array} \right) \begin{array}{l} +2R_2 \\ \\ +(-1)R_2 \end{array} \\ &= \left(\begin{array}{ccc|ccc} \color{red}{0} & \color{red}{0} & \color{red}{1} & 3 & 2 & -2 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ \color{red}{1} & \color{red}{0} & \color{red}{0} & -1 & -1 & 1 \end{array} \right) \begin{array}{l} \\ R_1 \leftrightarrow R_3 \\ \end{array} \\ (I | P^{-1}) &= \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 3 & 2 & -2 \end{array} \right), \end{aligned}$$

og konkluderer, at

$$\mathbf{x} = P^{-1}\mathbf{y} = \begin{pmatrix} -1 & -1 & 0 \\ 0 & 1 & 0 \\ 3 & 2 & -2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -y_1 - y_2 \\ y_2 \\ 3y_1 + 2y_2 - 2y_3 \end{pmatrix}.$$

Bemærkning 4.4.19 For at kunne skifte koordinater og lave lignende udregninger er det kun nødvendigt at huske følgende:

- Definitionen af matricen, der repræsenterer en lineær afbildning med hensyn til givne baser for dens domæne og codomæne.
- Hvis to lineære afbildning sammensættes, så multipliceres de matricer, der repræsenterer dem, i samme rækkefølge.

Koordinatskift svarer til at anvende disse principper på identitetsafbildningen. Det er en god idé at tegne figurer, som vi har gjort ovenfor, for ikke at blive forvirret over, hvilken vej afbildningerne går, og hvilken rækkefølge, de sammensættes i.

Bemærkning 4.4.20 Lineære afbildninger mellem uendeligt dimensionale vektorrum forekommer naturligt i mange sammenhænge. For eksempel har mængden $C^0([a, b])$ af kontinuerte reelle funktioner $f: [a, b] \rightarrow \mathbb{R}$ en reel vektorrumsstruktur med vektorsum og skalarmultiplikation defineret ved $(f + g)(x) = f(x) + g(x)$ og $(f \cdot a)(x) = f(x) \cdot a$, og

4 Vektorrum

afbildningen $I: C^0([a, b]) \rightarrow \mathbb{R}$ defineret ved

$$I(f) = \int_a^b f(x) dx$$

er da en lineær afbildning. Dimensionen af $C^0([a, b])$ er imidlertid $2^{\aleph_0}$, så det er ikke muligt at studere dette vektorrum med de algebraiske metoder, vi har udviklet her. I stedet betragter vi $C^0([a, b])$ som et *topologisk* vektorrum, hvilket gør det muligt at tale om (visse) uendelige summer, og som sådan udgør den tællelige familie $(x^n)_{n \in \mathbb{N}}$ ifølge Weierstrass' approksimationssætning en slags basis.

4.5 Kerne og billede

Vi definerer kernen og billedet af en lineær afbildning og viser, hvordan disse begreber giver anledning til en bedre forståelse af lineære ligningssystemer.

Definition 4.5.1 Lad $\mathbb{F}$ være et legeme. Hvis $f: V \rightarrow W$ er en lineær afbildning mellem to $\mathbb{F}$ -vektorrum, da kaldes

$$\ker(f) = \{v \in V \mid f(v) = \mathbf{0}\} \subset V \quad \text{og} \quad \text{im}(f) = \{f(v) \in W \mid v \in V\} \subset W$$

henholdsvis *kernen* af $f: V \rightarrow W$ og *billedet* af $f: V \rightarrow W$.

Lemma 4.5.2 Lad $f: V \rightarrow W$ være en lineær afbildning mellem to $\mathbb{F}$ -vektorrum. Da er $\ker(f) \subset V$ underrum af V og $\text{im}(f) \subset W$ underrum af W .

Bevis Vi skal vise, at $\ker(f) \subset V$ og $\text{im}(f) \subset W$ opfylder (1)–(3) i definition 4.1.4 og begynder med $\ker(f) \subset V$. Da $f(\mathbf{0}) = \mathbf{0}$, er $\mathbf{0} \in \ker(f)$, så (1) gælder. Hvis $f(u) = \mathbf{0}$ og $f(v) = \mathbf{0}$, da er også $f(u + v) = f(u) + f(v) = \mathbf{0} + \mathbf{0} = \mathbf{0}$, så (2) gælder. Endelig, hvis $f(v) = \mathbf{0}$, da er $f(v \cdot a) = f(v) \cdot a = \mathbf{0} \cdot a = \mathbf{0}$, så også (3) gælder.

Vi betragter dernæst $\text{im}(f) \subset W$. Da $\mathbf{0} = f(\mathbf{0})$, er $\mathbf{0} \in \text{im}(f)$, så (1) gælder. Hvis $w = f(u)$ og $z = f(v)$, da er $w + z = f(u) + f(v) = f(u + v)$, så (2) gælder. Og hvis $w = f(v)$, da er $w \cdot a = f(v) \cdot a = f(v \cdot a)$, så også (3) gælder. Vi har nu bevist lemmaet. $\square$

Eksempel 4.5.3 (1) For nulafbildningen $\mathbf{0}: V \rightarrow W$ er $\ker(\mathbf{0}) = V$ og $\text{im}(\mathbf{0}) = \{\mathbf{0}\}$, og for identitetsafbildningen $\text{id}_V: V \rightarrow V$ er $\ker(\text{id}_V) = \{\mathbf{0}\}$ og $\text{im}(\text{id}_V) = V$.

(2) Hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er repræsenteret af matricen $A \in M_{m,n}(\mathbb{F})$ med hensyn til de respektive standardbaser, da er $\ker(f) = N_A$ lig med nulrummet for A , og $\text{im}(f) = R_A$ er lig med søjlerummet af A .

Vi ser af eksempel 4.5.3, at det for både nulafbildningen og identitetsafbildningen gælder, at summen af dimensionerne af kernen og billedet er lig med dimensionen af domænet. Vi skal nu vise, at denne identitet, der kaldes Grassmanns dimensionsformel, gælder helt generelt.

Sætning 4.5.4 *Lad $\mathbb{F}$ være et legeme. Hvis $f: V \rightarrow W$ er en lineær afbildning mellem endeligt dimensionale $\mathbb{F}$ -vektorrum, da er*

$$\dim(\ker(f)) + \dim(\operatorname{im}(f)) = \dim(V).$$

Bevis Vi vælger baser $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_r)$ for henholdsvis $\ker(f)$ og $\operatorname{im}(f)$, hvilket er muligt ifølge sætning 4.3.15, og vælger yderligere en familie $(\mathbf{v}_{p+1}, \dots, \mathbf{v}_{p+r})$ af vektorer i V , sådan at $f(\mathbf{v}_{p+i}) = \mathbf{w}_i$, for alle $1 \leq i \leq r$. Vi påstår, at $(\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{v}_{p+1}, \dots, \mathbf{v}_{p+r})$ er en basis for V , hvorfra sætningen umiddelbart følger. For at vise påstanden, viser vi først, at familien $(\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{v}_{p+1}, \dots, \mathbf{v}_{p+r})$ er lineært uafhængig. Så vi lader

$$\mathbf{v}_1 a_1 + \dots + \mathbf{v}_p a_p + \mathbf{v}_{p+1} a_{p+1} + \dots + \mathbf{v}_{p+r} a_{p+r} = \mathbf{0}$$

være en linearkombination, der er lig med $\mathbf{0}$, og skal vise at $a_i = 0$ for alle $1 \leq i \leq p+r$. Vi anvender den lineære afbildning $f: V \rightarrow W$ på begge sider af ligningen og får

$$\mathbf{w}_1 a_{p+1} + \dots + \mathbf{w}_r a_{p+r} = \mathbf{0},$$

idet $f(\mathbf{v}_i) = \mathbf{0}$ for $1 \leq i \leq p$, og $f(\mathbf{v}_{p+i}) = \mathbf{w}_i$ for $1 \leq i \leq r$. Da familien $(\mathbf{w}_1, \dots, \mathbf{w}_r)$ er lineært uafhængig, konkluderer vi, at $a_{p+i} = 0$ for $1 \leq i \leq r$. Den oprindelige ligning er altså

$$\mathbf{v}_1 a_1 + \dots + \mathbf{v}_p a_p = \mathbf{0},$$

og da også $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ er lineært uafhængig, følger det, at $a_i = 0$ for $1 \leq i \leq p$. Dette viser, at familien $(\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{v}_{p+1}, \dots, \mathbf{v}_{p+r})$ er lineært uafhængig som ønsket. Vi mangler at vise, at denne familie også frembringer V . Så lad $\mathbf{v} \in V$ være en vilkårlig vektor, og lad $\mathbf{w} = f(\mathbf{v}) \in W$. Da $(\mathbf{w}_1, \dots, \mathbf{w}_r)$ frembringer W , kan vi skrive $\mathbf{w} = \mathbf{w}_1 b_1 + \dots + \mathbf{w}_r b_r$ som en linearkombination heraf. Da er

$$f(\mathbf{v} - (\mathbf{v}_{p+1} b_1 + \dots + \mathbf{v}_{p+r} b_r)) = \mathbf{w} - (\mathbf{w}_1 b_1 + \dots + \mathbf{w}_r b_r) = \mathbf{0},$$

hvilket viser, at $\mathbf{v} - (\mathbf{v}_{p+1} b_1 + \dots + \mathbf{v}_{p+r} b_r) \in \ker(f)$, og vi kan derfor skrive denne vektor som en linearkombination

$$\mathbf{v} - (\mathbf{v}_{p+1} b_1 + \dots + \mathbf{v}_{p+r} b_r) = \mathbf{v}_1 a_1 + \dots + \mathbf{v}_p a_p$$

af familien $(\mathbf{v}_1, \dots, \mathbf{v}_p)$, der per antagelse frembringer $\ker(f)$. Men dermed er

$$\mathbf{v} = \mathbf{v}_1 a_1 + \dots + \mathbf{v}_p a_p + \mathbf{v}_{p+1} b_1 + \dots + \mathbf{v}_{p+r} b_r$$

en linearkombination af $(\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{v}_{p+1}, \dots, \mathbf{v}_{p+r})$ som ønsket. Dette viser påstanden og dermed sætningen. $\square$

4 Vektorrum

Bemærkning 4.5.5 Grassmanns dimensionsformel giver et kvantitativt udtryk for, at desto mindre billedet af en lineær afbildning $f: V \rightarrow W$ er, desto større må dens kerne være og vice versa. Vi bemærker også, at dimensionen af codomænet W ikke spiller nogen rolle i sætningen, udover at det nødvendigvis altid gælder, at $\dim(\text{im}(f)) \leq \dim(W)$.

Eksempel 4.5.6 (1) For nulafbildningen $\mathbf{0}: V \rightarrow W$ har kernen samme dimension som V , mens billedet har dimension 0. For identitetsafbildningen $\text{id}_V: V \rightarrow V$ har kernen dimension 0, mens billedet har samme dimension som V .

(2) For den lineære afbildning $f: \mathbb{F}^4 \rightarrow \mathbb{F}^3$ defineret ved

$$f(\mathbf{x}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{x}$$

er $\ker(f) = \{\mathbf{x} \in \mathbb{F}^4 \mid x_1 = 0 \text{ og } x_2 = 0\}$, mens $\text{im}(f) = \{\mathbf{y} \in \mathbb{F}^3 \mid y_3 = 0\}$. Så $\ker(f)$ har basis $(\mathbf{e}_3, \mathbf{e}_4)$ og dimension 2, mens $\text{im}(f)$ har basis $(\mathbf{e}_1, \mathbf{e}_2)$ og dimension 2. Dermed er

$$\dim(\ker(f)) + \dim(\text{im}(f)) = 2 + 2 = 4 = \dim(\mathbb{F}^4),$$

som sætning 4.5.4 viser.

Definition 4.5.7 Lad $\mathbb{F}$ være et legeme, og lad $f: V \rightarrow W$ være en lineær afbildning mellem endeligt frembragte $\mathbb{F}$ -vektorrums. Da kaldes dimensionen

$$\text{rank}(f) = \dim(\text{im}(f))$$

af billedet af $f: V \rightarrow W$ for *rangen* af $f: V \rightarrow W$.

Vi vil nu relatere rang af lineære afbildninger til rang af matricer fra definition 2.5.13. Hvis A er en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$, og hvis $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ er den lineære afbildning defineret ved $f(\mathbf{x}) = A\mathbf{x}$, da er

$$\ker(f) = N_A = \{\mathbf{x} \in \mathbb{F}^n \mid A\mathbf{x} = \mathbf{0}\} \subset \mathbb{F}^n$$

lig med nulrummet for A , mens

$$\text{im}(f) = R_A = \{A\mathbf{x} \in \mathbb{F}^m \mid \mathbf{x} \in \mathbb{F}^n\} \subset \mathbb{F}^m$$

er lig med søjlerummet for A . Vi har set, at Gauss-elimination giver en algoritme til at bestemme en basis for nulrummet, og vi skal nu se, at Gauss-elimination ligeledes giver en algoritme til at bestemme en basis for søjlerummet.

Sætning 4.5.8 *Lad A være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$ og lad A' være den entydigt bestemte matrix på reduceret echelonform, der er rækkeækvivalent med A . Hvis $1 \leq j_1 < \dots < j_r \leq n$ indicerer de søjler i A' , der indeholder en ledende indgang, da er familien $(\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_r})$ bestående af de tilsvarende søjler i A en basis for $R_A \subset \mathbb{F}^m$.*

Bevis Vi vælger en invertibel matrix $P \in M_m(\mathbb{F})$, sådan at $A' = PA$, hvilket er muligt ifølge sætning 2.5.11. Hvis $\mathbf{y} = A\mathbf{x}$, da er $P\mathbf{y} = PA\mathbf{x} = A'\mathbf{x}$, hvilket viser, at afbildningen $f: R_A \rightarrow R_{A'}$ defineret ved $f(\mathbf{y}) = P\mathbf{y}$ er veldefineret. Vi lader $Q \in M_m(\mathbb{F})$ være den inverse matrix til P og $g: R_{A'} \rightarrow R_A$ den lineære afbildning defineret ved $g(\mathbf{z}) = Q\mathbf{z}$. Vi bemærker, at f og g er hinandens inverse.

Vi minder om, at for en vilkårlig $m \times n$ -matrix C er dens j 'te søjle givet ved $\mathbf{c}_j = C\mathbf{e}_j$. Specielt gælder det for søjlerne $\mathbf{a}_j$ i A og $\mathbf{a}'_j$ i A' , at

$$f(\mathbf{a}_j) = P\mathbf{a}_j = PA\mathbf{e}_j = A'\mathbf{e}_j = \mathbf{a}'_j \quad \text{og} \quad g(\mathbf{a}'_j) = Q\mathbf{a}'_j = QA'\mathbf{e}_j = A\mathbf{e}_j = \mathbf{a}_j.$$

Vi minder ligeledes om, at vi fra eksempel 4.3.8 ved, at familien $(\mathbf{a}'_1, \dots, \mathbf{a}'_n)$ frembringer $R_{A'}$. Vi påstår, at familien $(\mathbf{e}_1, \dots, \mathbf{e}_r)$ er en basis for $R_{A'}$. Denne familie er en familie af vektorer i $R_{A'}$, idet det ifølge definitionen af en matrix på reduceret echelonform gælder, at $\mathbf{a}'_{j_s} = \mathbf{e}_s$ for alle $1 \leq s \leq r$. Familien er endvidere lineært uafhængig, da den er lineært uafhængig som en familie af vektorer i $\mathbb{F}^m$. Endelig frembringer den $R_{A'}$, idet det ifølge definitionen af en matrix på reduceret echelonform gælder, at $a'_{ij} = 0$ for alle $r < i \leq m$ og $1 \leq j \leq n$. Dette viser påstanden, og sætning 4.4.12 viser da, at familien

$$(g(\mathbf{e}_1), \dots, g(\mathbf{e}_r)) = (g(\mathbf{a}'_{j_1}), \dots, g(\mathbf{a}'_{j_r})) = (\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_r})$$

er en basis for R_A som ønsket. $\square$

Sætningen giver specielt en algoritme, der til en given n -tuple af vektorer i $\mathbb{F}^m$, der frembringer $\mathbb{F}^m$, giver en delfamilie, der er en basis for $\mathbb{F}^m$. Vi illustrerer nu dette.

Eksempel 4.5.9 Vi betragter 5-tuplen af vektorer i $\mathbb{F}^3$ givet ved

$$\left(\mathbf{a}_1 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} 1 \\ 9 \\ 5 \end{pmatrix}, \mathbf{a}_4 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}, \mathbf{a}_5 = \begin{pmatrix} 2 \\ 7 \\ 8 \end{pmatrix} \right),$$

og vi anvender da eksempel 4.3.8 og sætning 4.5.8 til at finde en delfamilie heraf, der er en basis for $\mathbb{F}^3$, hvis dette er muligt. Vi omdanner derfor matrixen

$$A = \begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 3 & 4 & 9 & 0 & 7 \\ 2 & 3 & 5 & 1 & 8 \end{pmatrix},$$

4 Vektorrum

hvis j 'te søjle er $\mathbf{a}_j$, til en matrix A' på reduceret echelonform, hvilket giver

$$A' = \begin{pmatrix} 1 & 0 & 7 & 0 & -39 \\ 0 & 1 & -3 & 0 & 31 \\ 0 & 0 & 0 & 1 & -7 \end{pmatrix}.$$

De ledende indgange i A' findes i søjlerne 1, 2 og 4, så ifølge sætning 4.5.8 er delfamilien $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4)$ derfor en basis $\mathbb{F}^3$. Vi bemærker, at grunden til at en sådan basis findes i dette tilfælde er, at $\text{rank}(A) = 3 = m$. Hvis dette ikke havde været tilfældet, så ville den oprindelige tuple ikke frembringe hele $\mathbb{F}^3$, men et underrum deraf, og vi ville da i stedet have fundet en basis for dette underrum.

Korollar 4.5.10 *Lad $\mathbb{F}$ være et legeme, lad A være en $m \times n$ -matrix med indgange i $\mathbb{F}$, og lad $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ være den lineære afbildning defineret ved $f(\mathbf{x}) = A\mathbf{x}$. Da er*

$$\text{rank}(f) = \text{rank}(A).$$

Bevis Sætning 4.5.8 viser, at dimensionen af søjlerummet R_A er lig med antallet r af ledende indgange i den entydigt bestemte matrix A' , der er på reduceret echelonform og rækkeækvivalent med A . Det gælder altså, at

$$\text{rank}(f) = \dim(\text{im}(f)) = \dim(R_A) = r = \text{rank}(A)$$

som ønsket. □

Sætning 4.5.11 *Lad A være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$ og lad A' være den entydigt bestemte matrix på reduceret echelonform, der er rækkeækvivalent med A . Lad $1 \leq j_1 < \dots < j_r \leq n$ og $1 \leq k_1 < \dots < k_p \leq n$ indicere de søjler i A' , der henholdsvis indeholder og ikke indeholder ledende indgange, og for $1 \leq i \leq p$ lad*

$$\mathbf{c}_i = \mathbf{e}_{k_i} - \sum_{s=1}^r \mathbf{e}_{j_s} a'_{s,k_i}.$$

Da udgør familien $(\mathbf{c}_1, \dots, \mathbf{c}_p)$ en basis for nulrummet $N_A = N_{A'} \subset \mathbb{F}^n$.

Bevis Vi viser først, at $(\mathbf{c}_1, \dots, \mathbf{c}_p)$ er en familie af vektorer i $N_{A'}$. Vi påstår nemlig, at

$$A' \mathbf{c}_i = A'(\mathbf{e}_{k_i} - \sum_{s=1}^r \mathbf{e}_{j_s} a'_{s,k_i}) = \mathbf{a}'_{k_i} - \sum_{s=1}^r \mathbf{a}'_{j_s} a'_{s,k_i} = \mathbf{a}'_{k_i} - \sum_{s=1}^r \mathbf{e}_s a'_{s,k_i} = \mathbf{0}.$$

Her følger den første identitet fra definitionen af $\mathbf{c}_i$, mens den anden skyldes identiteten $A'\mathbf{e}_{j_s} = \mathbf{a}'_{j_s}$. Den tredje og fjerde identitet følger af, at for $1 \leq s \leq r$ er den j_s 'te søjle i A' lig med $\mathbf{e}_s$, og at for alle $r < s \leq m$ er $a'_{s,k_i} = 0$. Dette viser, at $\mathbf{c}_i \in N_{A'}$ for alle $1 \leq i \leq p$.

Vi viser dernæst, at familien $(\mathbf{c}_1, \dots, \mathbf{c}_p)$ er lineært uafhængig. Så lad

$$\mathbf{c}_1 a_1 + \mathbf{c}_2 a_2 + \dots + \mathbf{c}_p a_p = \mathbf{0}$$

være en linearkombination, der er lig med nulvektoren. For $1 \leq i \leq p$ udtrykker den k_i 'te ligning i dette ligningssystem, at $a_i = 0$, da $\mathbf{e}_{k_i}$ kun optræder i vektoren $\mathbf{c}_{k_i}$. Dette viser som ønsket, at $(\mathbf{c}_1, \dots, \mathbf{c}_p)$ er lineært uafhængig.

Vi mangler at vise, at $(\mathbf{c}_1, \dots, \mathbf{c}_p)$ frembringer $N_{A'}$. Sætning 4.5.4 og korollar 4.5.10 viser, at $\dim(N_{A'}) = n - r = p$. Dermed er $(\mathbf{c}_1, \dots, \mathbf{c}_p)$ altså en lineært uafhængig p -tuple af vektorer i det p -dimensionale vektorrum $N_{A'}$, og sætning 4.3.24 viser derfor som ønsket, at denne p -tuple er en basis for $N_{A'}$. $\square$

Vi illustrerer endelig algoritmerne i sætning 4.5.11 og sætning 4.5.8, der producerer baser for henholdsvis kernen og billedet af en lineær afbildning $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$.

Eksempel 4.5.12 Vi betragter afbildningen $f: \mathbb{F}^4 \rightarrow \mathbb{F}^3$ givet ved $f(\mathbf{x}) = A\mathbf{x}$, hvor

$$A = \begin{pmatrix} 2 & 4 & 3 & 7 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix}$$

og ønsker at bestemme en basis for $\ker(f)$ og $\operatorname{im}(f)$. Vi omdanner derfor først A til en matrix A' på reduceret echelonform ved hjælp af rækkeoperationer.

$$\begin{aligned} A &= \begin{pmatrix} 2 & 4 & 3 & 7 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix} \begin{array}{l} +(-2)R_3 \\ \\ \end{array} \\ &\quad \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix} \begin{array}{l} +(-1)R_2 \\ \\ +(-1)R_2 \end{array} \\ &\quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{pmatrix} R_1 \leftrightarrow R_3 \\ A' &= \begin{pmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Her har vi igen markeret de indgange, vi ønsker at ændre med rødt, og vi har markeret de $r = 2$ ledende indgange i $A' = B$ med blå. Så $\operatorname{im}(f)$ har dimension $r = 2$, og

4 Vektorrum

Grassmanns dimensionsformel viser, at $\ker(f)$ har dimension $n - r = 4 - 2 = 2$. Da de ledende indgange i A' findes i søjle 1 og 3, viser sætning 4.5.8, at

$$\left(\mathbf{a}_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \right)$$

er en basis for $\operatorname{im}(f)$. Og ifølge sætning 4.5.11 er $(\mathbf{c}_1, \mathbf{c}_2)$ en basis for $\ker(f)$, hvor

$$\mathbf{c}_1 = \mathbf{e}_2 - (\mathbf{e}_{j_1} b_{12} + \mathbf{e}_{j_2} b_{22}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} 2 - \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} 0 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

og

$$\mathbf{c}_2 = \mathbf{e}_4 - (\mathbf{e}_{j_1} b_{14} + \mathbf{e}_{j_2} b_{24}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} 2 - \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} 1 = \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

4.6 Affine underrum og kvotientrum

Lad V være et vektorrum over et legeme $\mathbb{F}$, og lad $U \subset V$ være et underrum. I dette afsnit vil vi sige, at $U \subset V$ er et lineært underrum for at huske på, at det altid gælder, at $\mathbf{0} \in U$. Givet et vilkårligt $\mathbf{v} \in V$ er delmængden

$$\mathbf{v} + U = \{\mathbf{v} + \mathbf{u} \mid \mathbf{u} \in U\} \subset V$$

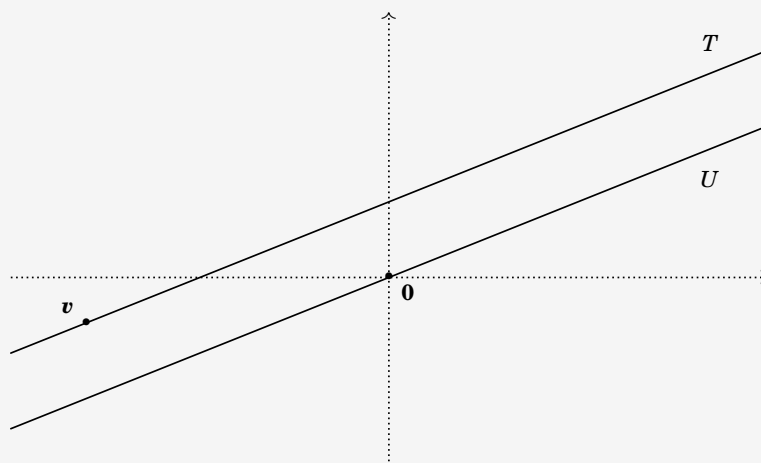
ikke et lineært underrum, medmindre $\mathbf{v} \in U$. Vi viser nedenfor, at løsningsmængden til et inhomogent ligningssystem " $A\mathbf{x} = \mathbf{b}$ " enten er tom eller på formen $\mathbf{v} + N_A \subset \mathbb{F}^n$, hvor $N_A \subset \mathbb{F}^n$ er det lineære underrum af løsninger til det homogene ligningssystem " $A\mathbf{x} = \mathbf{0}$ ". Vi indfører følgende begreb.

Definition 4.6.1 Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$ -vektorrum, og lad $U \subset V$ være et lineært underrum. En delmængde $T \subset V$ er et *affint underrum parallelt med* $U \subset V$, hvis der findes $\mathbf{v} \in T$, sådan at

$$T = \mathbf{v} + U = \{\mathbf{v} + \mathbf{u} \mid \mathbf{u} \in U\} \subset V.$$

Eksempel 4.6.2 Den følgende figur illustrerer det lineære underrum $U \subset \mathbb{R}^2$ og det affine underrum $T \subset \mathbb{R}^2$ parallelt med U , hvor

$$U = \left\{ \begin{pmatrix} 8 \\ 3 \end{pmatrix} t \mid t \in \mathbb{R} \right\} \quad \text{og} \quad T = \begin{pmatrix} -8 \\ -1 \end{pmatrix} + U.$$



Lemma 4.6.3 Lad V være et vektorrum, lad $U \subset V$ være et lineært underrum, og lad $T \subset V$ være et affint underrum parallelt med $U \subset V$. Da gælder det for alle $\mathbf{v} \in T$, at $T = \mathbf{v} + U$. Desuden er $T \subset V$ et lineært underrum, hvis og kun hvis $T = U$.

Bevis Hvis T er et affint underrum, da er $T = \mathbf{v} + U$ for en vektor $\mathbf{v} \in V$. Idet $\mathbf{0} \in U$, ved vi, at $\mathbf{v} \in T$. Hvis også $\mathbf{v}' \in T = \mathbf{v} + U$, så findes der $\mathbf{u} \in U$, sådan at $\mathbf{v}' = \mathbf{v} + \mathbf{u}$. Derfor er

$$\mathbf{v}' + U = (\mathbf{v} + \mathbf{u}) + U = \mathbf{v} + (\mathbf{u} + U) \subset \mathbf{v} + U.$$

Ved at ombytte $\mathbf{v}$ og $\mathbf{v}'$ i argumentet, ser vi ligeledes, at $\mathbf{v} + U \subset \mathbf{v}' + U$, hvilket viser den første påstand i lemmaet. For at vise den anden påstand antager vi, at $T \subset V$ er et lineært underrum parallelt med $U \subset V$. Da er $\mathbf{0} \in T$ per definition, så ifølge den første påstand er $T = \mathbf{0} + U = U$. $\square$

Eksempel 4.6.4 Det affine underrum T i eksempel 4.6.2 kan også skrives som

$$T = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + U,$$

idet

$$\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -8 \\ -1 \end{pmatrix} + \begin{pmatrix} 8 \\ 3 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 8 \\ 3 \end{pmatrix} \in U.$$

4 Vektorrum

Lemma 4.6.5 *Lad $\mathbb{F}$ være et legeme, lad $f: V \rightarrow W$ være en lineær afbildning mellem to $\mathbb{F}$ -vektorrums, og lad $U = \ker(f) \subset V$ være kernen af $f: V \rightarrow W$. Lad $\mathbf{w} \in W$, og lad*

$$T = f^{-1}(\mathbf{w}) = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{w}\}$$

være det inverse billede ved $f: V \rightarrow W$. Da er $T \subset V$ enten den tomme delmængde eller et affint underrum parallelt med $U \subset V$.

Bevis Hvis $T \subset V$ er tom, er der ikke noget at vise, så vi antager, at $T \subset V$ ikke er tom. Vi kan da vælge $\mathbf{v} \in T$ og skal vise, at $T = \mathbf{v} + U$. Hvis $\mathbf{u} \in U$, da er

$$f(\mathbf{v} + \mathbf{u}) = f(\mathbf{v}) + f(\mathbf{u}) = \mathbf{w} + \mathbf{0} = \mathbf{w},$$

hvilket viser, at $\mathbf{v} + U \subset T$. Omvendt gælder det for alle $\mathbf{v}' \in T$, at

$$f(\mathbf{v}' - \mathbf{v}) = f(\mathbf{v}') - f(\mathbf{v}) = \mathbf{w} - \mathbf{w} = \mathbf{0},$$

hvilket viser, at $\mathbf{u} = \mathbf{v}' - \mathbf{v} \in U$. Så $\mathbf{v}' = \mathbf{v} + \mathbf{u} \in \mathbf{v} + U$, og derfor er også $T \subset \mathbf{v} + U$. $\square$

Korollar 4.6.6 *Lad $A \in M_{m,n}(\mathbb{F})$ være en $m \times n$ -matrix med indgange i et legeme $\mathbb{F}$. Løsningsmængden til det lineære ligningssystem $A\mathbf{x} = \mathbf{b}$ er et affint underrum parallelt med løsningsmængden til det homogene ligningssystem $A\mathbf{x} = \mathbf{0}$.*

Bevis Løsningsmængden til det homogene ligningssystem $A\mathbf{x} = \mathbf{0}$ er netop nulrummet $N_A \subset \mathbb{F}^n$, som ifølge lemma 4.1.8 er et lineært underrum. Lad endvidere $f: \mathbb{F}^n \rightarrow \mathbb{F}^m$ være den lineære afbildning defineret ved $f(\mathbf{x}) = A\mathbf{x}$. Løsningsmængden til det lineære ligningssystem $A\mathbf{x} = \mathbf{b}$ er netop delmængden $f^{-1}(\mathbf{b}) \subset \mathbb{F}^n$, mens $N_A = \ker(f)$. Påstanden følger derfor fra lemma 4.6.5. $\square$

Eksempel 4.6.7 Vi betragter det lineære ligningssystem “ $A\mathbf{x} = \mathbf{b}$,” hvor

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad \text{og} \quad \mathbf{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix},$$

og for at løse det omdanner vi totalmatricen $(A \mid \mathbf{b})$ til matricen

$$(A' \mid \mathbf{b}') = \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 0 & 0 \end{array} \right),$$

der er på reduceret echelonform. Vi aflæser heraf, at den fælles løsningsmængde til ligningssystemerne “ $A\mathbf{x} = \mathbf{b}$ ” og “ $A'\mathbf{x} = \mathbf{b}'$ ” er lig med det affine underrum

$$T = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \end{pmatrix} t \mid t \in \mathbb{F} \right\} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} t \mid t \in \mathbb{F} \right\} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + N_A \subset \mathbb{F}^2$$

parallelt med løsningsmængden $N_A \subset \mathbb{F}^2$ til det homogene ligningssystem “ $A\mathbf{x} = \mathbf{0}$.”

4.6 Affine underrum og kvotientrum

Givet et lineært underrum $U \subset V$ af et vektorrum V , danner vi nu et vektorrum V/U , hvis vektorer er de affine underrum $T \subset V$, der er parallelle med $U \subset V$.

Definition 4.6.8 Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$ -vektorrum, og lad $U \subset V$ være et lineært underrum. Kvotientrummet af V med hensyn til $U \subset V$ er vektorrummet

$$(V/U, +, \cdot),$$

hvor V/U er mængden af affine underrum $T \subset V$, der er parallelle med $U \subset V$, og hvor vektorsum og skalarmultiplikation er givet ved henholdsvis

$$(\mathbf{v} + U) + (\mathbf{v}' + U) = (\mathbf{v} + \mathbf{v}') + U \quad \text{og} \quad (\mathbf{v} + U) \cdot a = (\mathbf{v} \cdot a) + U.$$

Et affint underrum $T \subset V$, der er parallelt med det lineære underrum $U \subset V$, udgør altså en enkelt vektor i kvotientvektorrummet V/U . Specielt er $U \subset V$ nulvektoren i dette vektorrum. Vi definerer den *kanoniske*⁵ *projektion* til at være afbildningen

$$p: V \rightarrow V/U$$

givet ved $p(\mathbf{v}) = \mathbf{v} + U$ og bemærker, at vektorsum og skalarmultiplikation på V/U netop er defineret, sådan at denne afbildning er lineær.

Sætning 4.6.9 Lad $\mathbb{F}$ være et legeme, lad V være et endeligt frembragt $\mathbb{F}$ -vektorrum, og lad $U \subset V$ være et lineært underrum. Da er

$$\dim(V/U) = \dim(V) - \dim(U).$$

Bevis Den kanoniske projektion $p: V \rightarrow V/U$ er surjektiv og $\ker(p) = U$. Udsagnet følger derfor af Grassmanns dimensionsformel (sætning 4.5.4). $\square$

Hvis $U \subset V$ er et lineært underrum af et vektorrum V , så kalder vi dimensionen af kvotientrummet V/U for *codimensionen* af det lineære underrum $U \subset V$. Da er $U = V$, hvis og kun hvis codimensionen af $U \subset V$ er lig med nul. Vi siger endvidere, at $U \subset V$ er en *hyperplan*, hvis codimensionen af $U \subset V$ er lig med 1. Vi bemærker endelig, at det er muligt for codimensionen af et lineært underrum $U \subset V$ at være endelig, selvom U og V begge er uendeligdimensionale vektorrum.

⁵Med "kanonisk" udtrykker vi blot, at vi foretrækker projektionsafbildningen $p(\mathbf{v}) = \mathbf{v} + U$ fremfor andre projektionsafbildninger, såsom $q(\mathbf{v}) = -\mathbf{v} + U$. Ordet har ikke noget matematisk indhold.

4 Vektorrum

Eksempel 4.6.10 Hvis $X \subset \mathbb{R}^n$ er en åben delmængde, da udgør mængden $C^\infty(X, \mathbb{R}^m)$ af de vilkårligt ofte kontinuert differentiable afbildninger $f: X \rightarrow \mathbb{R}^m$ med vektorsum og skalarmultiplikation defineret ved

$$(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}) \quad \text{og} \quad (f \cdot a)(\mathbf{x}) = f(\mathbf{x}) \cdot a$$

et reelt vektorrum. Hvis $m = n$, så siger vi, at $g \in C^\infty(X, \mathbb{R}^n)$ er et vektorfelt på X . Vi lader $n = 2$ og definerer *gradienten* og *rotationen* til at være de lineære afbildninger

$$C^\infty(X, \mathbb{R}) \xrightarrow{\text{grad}} C^\infty(X, \mathbb{R}^2) \xrightarrow{\text{rot}} C^\infty(X, \mathbb{R}),$$

der er givet ved

$$\text{grad}(f) = \begin{pmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \end{pmatrix} \quad \text{og} \quad \text{rot} \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \frac{\partial g_1}{\partial x_2} - \frac{\partial g_2}{\partial x_1}.$$

Et vektorfelt $g \in \text{im}(\text{grad})$ siges at være *konservativt*, mens et vektorfelt $g \in \text{ker}(\text{rot})$ siges at være *rotationsfrit*. Ethvert konservativt vektorfelt er rotationsfrit, idet

$$(\text{rot} \circ \text{grad})(f) = \frac{\partial^2 f}{\partial x_2 \partial x_1} - \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0,$$

men det omvendte er generelt ikke tilfældet. Så

$$\text{im}(\text{grad}) \subset \text{ker}(\text{rot})$$

er et underrum, men det er i almindelighed ikke hele rummet. Kvotientrummet

$$H^1(X) = \text{ker}(\text{rot}) / \text{im}(\text{grad})$$

kaldes for *de Rham-cohomologien* af X i grad 1. Skønt vektorrummene $\text{ker}(\text{rot})$ og $\text{im}(\text{grad})$ begge er uendeligdimensionale (medmindre X er den tomme mængde), så er kvotientrummet $H^1(X)$ typisk endeligdimensionalt. Dimensionen af dette vektorrum er lig med antallet af "huller" i X . For eksempel er

$$\dim(H^1(\mathbb{R}^2)) = 0,$$

hvilket er det (måske) velkendte resultat, at ethvert rotationsfrit vektorfelt defineret på hele $\mathbb{R}^2$ er konservativt, mens man kan vise, at

$$\dim(H^1(\mathbb{R}^2 \setminus \{\mathbf{0}\})) = 1.$$

Lad os vise, at vektorfeltet $g: \mathbb{R}^2 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}^2$ defineret ved

$$g(\mathbf{x}) = \begin{pmatrix} \frac{-x_2}{x_1^2 + x_2^2} \\ \frac{x_1}{x_1^2 + x_2^2} \end{pmatrix}$$

er rotationsfrit, men at det ikke er konservativt. Dette medfører, at elementet

$$g + \text{im}(\text{grad}) \in H^1(\mathbb{R}^2 \setminus \{\mathbf{0}\})$$

ikke er nul. Vi antager, at $g = \text{grad}(f)$ er konservativt, og viser, at denne antagelse fører til en modstrid. Antagelsen viser, at

$$\int_0^{2\pi} \frac{d}{dt} f \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} dt = f \begin{pmatrix} \cos(2\pi) \\ \sin(2\pi) \end{pmatrix} - f \begin{pmatrix} \cos(0) \\ \sin(0) \end{pmatrix} = f \begin{pmatrix} 1 \\ 0 \end{pmatrix} - f \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0,$$

mens en direkte udregning ved brug af kædereglen viser, at

$$\int_0^{2\pi} \frac{d}{dt} f \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} dt = \int_0^{2\pi} \left\langle g \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \right\rangle dt = \int_0^{2\pi} dt = 2\pi,$$

hvilket giver den ønskede modstrid. At $\text{rot}(g) = 0$ eftervises ved direkte udregning.

Vi betragter igen de lineære afbildninger

$$U \xrightarrow{i} V \xrightarrow{p} V/U$$

defineret ved $i(\mathbf{u}) = \mathbf{u}$ og $p(\mathbf{v}) = \mathbf{v} + U$, som vi kalder for den kanoniske inklusion og den kanoniske projektion. Den sammensatte afbildning $p \circ i$ er da lig med nulafbildningen, idet $U \in V/U$ er nulvektoren i kvotientrummet.

Sætning 4.6.11 *Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$ -vektorrum, og lad $U \subset V$ være et lineært underrum. Hvis $f: V \rightarrow W$ er en lineær afbildning, sådan at $f \circ i: U \rightarrow W$ er lig med nulafbildningen, da findes der præcis én lineær afbildning*

$$V/U \xrightarrow{\bar{f}} W,$$

sådan at $f = \bar{f} \circ p$.

Bevis Da vi ønsker, at $f = \bar{f} \circ p$, er vi tvunget til at definere $\bar{f}: V/U \rightarrow W$ ved

$$\bar{f}(\mathbf{v} + U) = f(\mathbf{v}),$$

og vi skal derfor vise, at denne definition er meningsfuld. Hvis $\mathbf{v} + U = \mathbf{v}' + U$, så findes der $\mathbf{u} \in U$, sådan at $\mathbf{v} = \mathbf{v}' + \mathbf{u}$, og derfor er

$$f(\mathbf{v}) = f(\mathbf{v}' + \mathbf{u}) = f(\mathbf{v}') + f(\mathbf{u}) = f(\mathbf{v}') + \mathbf{0} = f(\mathbf{v}'),$$

hvilket viser, at afbildningen $\bar{f}: V/U \rightarrow W$ er veldefineret. Endelig er denne afbildning lineær, fordi $f: V \rightarrow W$ er lineær. $\square$

4.7 Opgaver

4.1 Beskriv underrummet $U \subset \mathbb{R}^3$ frembragt af familien $(\mathbf{v}_1, \mathbf{v}_2)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

4.2 Undersøg for de følgende familier af vektorer i $\mathbb{R}^3$ om de er lineært uafhængige, om de frembringer $\mathbb{R}^3$, og om de udgør en basis for $\mathbb{R}^3$. Hvis familien er lineært afhængig, angiv da også en linearkombination af familien, som er lig med nulvektoren $\mathbf{0}$.

a) Familien $(\mathbf{u}_1, \mathbf{u}_2)$, hvor

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

b) Familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

c) Familien $(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$, hvor

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

d) Familien $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4)$, hvor

$$\mathbf{a}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \text{og} \quad \mathbf{a}_4 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

4.3 Undersøg i hvert af de følgende tilfælde, om den givne familie af vektorer i $\mathbb{R}^4$ er lineært uafhængig, om den frembringer $\mathbb{R}^4$, og om den er en basis for $\mathbb{R}^4$. Hvis familien er lineært afhængig, angiv da også en linearkombination af familien, som er lig med nulvektoren $\mathbf{0}$.

a) Familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 4 \\ 6 \\ 8 \\ 6 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 0 \\ 3 \\ 2 \\ 1 \end{pmatrix}.$$

b) Familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}.$$

c) Familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

d) Familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 4 \\ 2 \\ -1 \\ 3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 6 \\ 5 \\ -5 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 5 \end{pmatrix}.$$

4.4 Vis, at delmængden

$$\mathbb{R}_{\geq 0}^n = \{\mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0 \text{ for alle } 1 \leq i \leq n\} \subset \mathbb{R}^n$$

ikke er stabil med hensyn til vektorrumstrukturen på $\mathbb{R}^n$.

4.5 Lad $(\mathbb{C}^n, +, *)$ være triplen, hvor $+: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}^n$ er den sædvanlige vektorsum, og hvor $*: \mathbb{C}^n \times \mathbb{C} \rightarrow \mathbb{C}^n$ er defineret ved

$$\mathbf{x} * z = \mathbf{x} \cdot |z|.$$

Her er $|z| = \sqrt{\bar{z}z} \in \mathbb{R}$ absolutværdien af $z \in \mathbb{C}$. Vis, at $(\mathbb{C}^n, +, *)$ ikke er et $\mathbb{C}$ -vektorrum.

4.6 (★) Lad $(M_4(\mathbb{R}), +, *)$ være triplen, hvor $+: M_4(\mathbb{R}) \times M_4(\mathbb{R}) \rightarrow M_4(\mathbb{R})$ er matrixsum, og hvor $*: M_4(\mathbb{R}) \times \mathbb{H} \rightarrow M_4(\mathbb{R})$ er defineret ved

$$A * (a + bi + cj + dk) = A \cdot \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}.$$

Vis at $(M_4(\mathbb{R}), +, \cdot)$ er et (højre) $\mathbb{H}$ -vektorrum.

[Vink: Vi ved allerede, at matrixsum opfylder (A1)–(A4) i definition 4.1.1, så det er tilstrækkeligt at vise, at også (V1)–(V4) er opfyldt.]

4 Vektorrum

4.7 Lad $V = M_2(\mathbb{F})$ være mængden af 2×2 -matricer med indgange i et legeme $\mathbb{F}$, og betragt triplen $(V, +, *)$, hvor $+: V \times V \rightarrow V$ er matrixsum, mens $*: V \times \mathbb{F} \rightarrow V$ er defineret ved

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} * a = \begin{pmatrix} x_{11}a & x_{12}a \\ x_{21}a & x_{22}a \end{pmatrix}.$$

Vi bemærker, at “ $*$ ” ikke er matrixprodukt.

a. Eftersis, at $(V, +, *)$ er et $\mathbb{F}$ -vektorrum.

b. Givet vektorerne

$$\mathbf{x} = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{og} \quad \mathbf{z} = \begin{pmatrix} 4 & 5 \\ -4 & -3 \end{pmatrix}$$

i V , find alle skalarer $a, b \in \mathbb{F}$, sådan at

$$\mathbf{x} * a + \mathbf{y} * b = \mathbf{z}.$$

4.8 Afgør i hvert af de følgende tilfælde, om den angivne delmængde af $\mathbb{R}^4$ er stabil med hensyn til den sædvanlige vektorrumstruktur på $\mathbb{R}^4$ eller ej.

a) Delmængden af alle vektorer $\mathbf{x}$, hvis førstekoordinat x_1 er et helt tal.

b) Delmængden af alle vektorer $\mathbf{x}$, hvis førstekoordinat x_1 er lig med 0.

c) Delmængden bestående af alle vektorer $\mathbf{x}$, hvis førstekoordinat x_1 eller andenkoordinat x_2 eller begge er lig med 0.

d) Delmængden bestående af alle vektorer $\mathbf{x}$, hvis førstekoordinat x_1 og andenkoordinat x_2 tilfredsstiller ligningen $x_1 + 2x_2 = 0$.

e) Delmængden bestående af alle vektorer, hvis førstekoordinat x_1 og andenkoordinat x_2 tilfredsstiller ligningen $x_1 + 2x_2 = 1$.

4.9 Lad U være et vektorrum, og lad $V, W \subset U$ være to underrum.

a) Vis, at

$$V \cap W = \{\mathbf{u} \in U \mid \mathbf{u} \in V \text{ og } \mathbf{u} \in W\} \subset U$$

er et underrum. [Vink: Eftersis (1)–(3) i definition 4.1.4.]

b) Vis, at

$$V + W = \{\mathbf{v} + \mathbf{w} \in U \mid \mathbf{v} \in V \text{ og } \mathbf{w} \in W\} \subset U$$

er et underrum.

4.10 Undersøg, om familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix},$$

er en basis for $\mathbb{R}^4$.

4.11 Lad $U \subset \mathbb{R}^5$ være underrummet frembragt af $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ -5 \\ 0 \\ 5 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 4 \\ -5 \\ 1 \\ 3 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 2 \\ 5 \\ 2 \\ 2 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 2 \\ -2 \\ 10 \\ 1 \\ -1 \end{pmatrix}.$$

Angiv en delfamilie af $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$, der er en basis for U , og skriv de af vektorerne $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$, der ikke tilhører den angivne basis, som en linearkombination af denne.

4.12 Betragt de følgende vektorer i $\mathbb{R}^6$:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{w}_1 = \begin{pmatrix} 4 \\ -5 \\ -1 \\ -5 \\ -1 \\ 5 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} -3 \\ 2 \\ -1 \\ 2 \\ -1 \\ -2 \end{pmatrix}.$$

Undersøg, om de to familier $(\mathbf{v}_1, \mathbf{v}_2)$ og $(\mathbf{w}_1, \mathbf{w}_2)$ frembringer det samme underrum af $\mathbb{R}^6$ eller ej.

4.13 Lad $U \subset \mathbb{R}^3$ være underrummet frembragt af $(\mathbf{v}_1, \mathbf{v}_2)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 4 \\ -7 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ -5 \\ 4 \end{pmatrix}.$$

Bestem alle $a \in \mathbb{R}$, sådan at vektoren

$$\begin{pmatrix} 1 \\ a \\ 5 \end{pmatrix}$$

tilhører underrummet U .

4.14 Vis, at familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 1 \\ 5 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 4 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 2 \\ 9 \\ 0 \\ 7 \end{pmatrix},$$

4 Vektorrum

er en basis for $\mathbb{R}^4$, og skriv vektoren

$$\mathbf{v} = \begin{pmatrix} 2 \\ 4 \\ 1 \\ 3 \end{pmatrix}$$

som en linearkombination af denne basis.

4.15 Lad $U \subset \mathbb{R}^4$ være underrummet frembragt af $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5)$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_5 = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 1 \end{pmatrix}.$$

- a) Find en delfamilie af familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5)$, der indeholder $\mathbf{v}_1$ og $\mathbf{v}_5$ og er en basis for U .
- b) Find en delfamilie af familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5)$, der indeholder $\mathbf{v}_3$ og $\mathbf{v}_4$, og er en basis for U .

4.16 Denne opgave viser, at en lineær afbildning $f: V \rightarrow W$ mellem $\mathbb{F}$ -vektorrum kan defineres ved at angive dens værdi på vektorerne i en basis for V . Dette kan ses som en generalisering af lemma 4.2.3.

- a) Hvis $(\mathbf{v}_i)_{i \in I}$ er en basis for V , og hvis $(\mathbf{w}_i)_{i \in I}$ er en familie af vektorer i W , der er indiceret ved den samme indekssmængde I , da definerer formlen

$$f\left(\sum_{i \in I} \mathbf{v}_i a_i\right) = \sum_{i \in I} \mathbf{w}_i a_i$$

en veldefineret afbildning $f: V \rightarrow W$. Vis, at denne afbildning er lineær.

- b) Lad $g: V \rightarrow W$ være en lineær afbildning, og lad $(\mathbf{v}_i)_{i \in I}$ være en basis for V . Vis, at den lineære afbildning $f: V \rightarrow W$, der fremkommer som i a) fra familien $(g(\mathbf{v}_i))_{i \in I}$ af vektorer i W , er lig med $g: V \rightarrow W$.

4.17 (★) Lad $(\mathbb{C}, +, \cdot)$ være det *reelle* vektorrum, hvor $\mathbb{C}$ er mængden af komplekse tal, hvor $+: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ er sum af komplekse tal, og hvor $\cdot: \mathbb{C} \times \mathbb{R} \rightarrow \mathbb{C}$ er defineret ved

$$(x + iy) \cdot a = xa + iya.$$

Lad $(V, +, \cdot)$ være det reelle vektorrum af reelle 2×2 -matricer fra opgave 4.7.

- a) Vis, at afbildningen $f: \mathbb{C} \rightarrow V$ defineret ved

$$f(x + iy) = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

er lineær.

- b) Find en basis for det reelle vektorrum $\mathbb{C}$ og bestem dets dimension.
 c) Find ligeledes en basis for det reelle vektorrum V og bestem dets dimension.

4.18 (★) Lad $\mathbb{F}$ være et legeme. Vis, at vektorrummet af polynomier $\mathbb{F}[t]$ *ikke* er endeligt frembragt.

[Vink: Vis, at $(t^0, t^1, t^2, \dots)$ er en lineært uafhængig familie i $\mathbb{F}[t]$, hvor man skal bruge bemærkning 4.3.26.]

4.19 Lad V være et vektorrum af dimension n , og lad $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ være en lineært uafhængig familie af vektorer i V . Afgør om følgende udsagn er korrekte eller ej.

- a) Det gælder nødvendigvis, at $k \leq n$.
 b) Hvis $k = n$, da er $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ en basis for V .
 c) Hvis $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ er en basis for V , da er $k = n$.
 d) Familien $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ er en delfamilie af en basis

$$(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n)$$

for vektorrummet V .

- e) Familien $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ er en delfamilie af præcis én basis

$$(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n)$$

for vektorrummet V .

4.20 Vi betragter familien

$$\left(\mathbf{v}_1 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 7 \end{pmatrix} \right)$$

af vektorer i $\mathbb{F}^2$.

- a) Vis, at $(\mathbf{v}_1, \mathbf{v}_2)$ er en basis for $\mathbb{F}^2$.
 b) Angiv matricen P , der repræsenterer $\text{id}: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for codomænet.
 c) Bestem matricen, der repræsenterer $\text{id}: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ med hensyn til basen $(\mathbf{e}_1, \mathbf{e}_2)$ for domænet og basen $(\mathbf{v}_1, \mathbf{v}_2)$ for codomænet.
 d) Find koordinaterne $\mathbf{y} \in \mathbb{F}^2$ for vektoren

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{F}^2$$

med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2)$.

4 Vektorrum

4.21 Vi betragter familien

$$\left(\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right)$$

af vektorer i $\mathbb{F}^3$.

- Vis, at $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ er en basis for $\mathbb{F}^3$.
- Angiv matricen P , der repræsenterer $\text{id}: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for codomænet.
- Bestem matricen, der repræsenterer $\text{id}: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til baserne $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for domænet og $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for codomænet.
- Find koordinaterne $\mathbf{y} \in \mathbb{F}^3$ af vektoren

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{F}^3$$

med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.

4.22 Vi betragter den lineære afbildning $f: \mathbb{F}^3 \rightarrow \mathbb{F}^2$, der er repræsenteret af matricen

$$A = \begin{pmatrix} 2 & 7 & 3 \\ 1 & 14 & 6 \end{pmatrix}$$

med hensyn til de respektive standardbaser. Vi betragter også basen

$$\left(\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$$

for $\mathbb{F}^3$ og basen

$$\left(\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right)$$

for $\mathbb{F}^2$.

- Find matricen P , der repræsenterer identitetsafbildningen $\text{id}: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til basen $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for codomænet.
- Find ligeledes matricen Q , der repræsenterer $\text{id}: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for codomænet.
- Angiv sammenhængen mellem matricerne A , P og Q , og den matrix B , der repræsenterer $f: \mathbb{F}^3 \rightarrow \mathbb{F}^2$ med hensyn til basen $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ for domænet og basen $(\mathbf{v}_1, \mathbf{v}_2)$ for codomænet.

d) Bestem matricen B .

e) Tegn en figur som i eksempel 4.4.16.

4.23 Vi betragter den lineære endomorfi $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$, der er repræsenteret af matricen

$$A = \begin{pmatrix} -4 & -3 & 3 \\ -2 & 3 & 0 \\ -7 & -1 & 4 \end{pmatrix}$$

med hensyn til standardbasen for både domænet og codomænet. Vi betragter endvidere den nye basis

$$\left(\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \right)$$

for $\mathbb{F}^3$.

a) Angiv matricen P , der repræsenterer $\text{id}: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for codomænet.

b) Udtryk matricen B , der repræsenterer $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for både domænet og codomænet, ved hjælp af matricerne A og P .

c) Bestem matricen B .

d) Tegn en figur som i eksempel 4.4.16.

4.24 Lad $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ være familien af vektorer i $\mathbb{F}^3$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

a) Vis, at $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ er en basis for $\mathbb{F}^3$.

b) Angiv matricen P , der repræsenterer $\text{id}: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for codomænet.

c) Om en lineær endomorfi $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ vides det, at matricen B , der repræsenterer endomorfien med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for både domænet og codomænet, er givet ved

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}.$$

Udtryk matricen A , der repræsenterer $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for både domænet og codomænet ved hjælp af matricerne B og P .

d) Bestem matricen A .

4 Vektorrum

e) Tegn en figur som i eksempel 4.4.16.

4.25 Vi betragter den lineære endomorfi $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$, $f(\mathbf{x}) = A\mathbf{x}$, hvor

$$A = \begin{pmatrix} -1 & 5 & -5 \\ 1 & 3 & -1 \\ -4 & 4 & -2 \end{pmatrix},$$

samt basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for $\mathbb{F}^3$, hvor

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

- a) Angiv matricen P , der repræsenterer $\text{id}: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for codomænet.
- b) Udtryk matricen B , der repræsenterer $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for både domænet og codomænet, ved hjælp af matricerne A og P .
- c) Bestem matricen B .
- d) Afgør om $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ er isomorfi eller ej.
- e) Tegn en figur som i eksempel 4.4.16.

4.26 Lad $f: \mathbb{F}^4 \rightarrow \mathbb{F}^3$ være afbildningen givet ved $f(\mathbf{x}) = A\mathbf{x}$, hvor

$$A = \begin{pmatrix} 2 & 1 & 8 & -1 \\ 2 & 0 & 4 & -2 \\ 0 & 1 & 4 & 1 \end{pmatrix}.$$

- a) Bestem dimensionerne af $\ker(f)$ og $\text{im}(f)$.
- b) Angiv baser for $\ker(f)$ og $\text{im}(f)$.

4.27 Lad $f: \mathbb{F}^5 \rightarrow \mathbb{F}^3$ være afbildningen givet ved $f(\mathbf{x}) = A\mathbf{x}$, hvor

$$A = \begin{pmatrix} 2 & -4 & -1 & 5 & -4 \\ 1 & -2 & 0 & 2 & -1 \\ 1 & -2 & 1 & 1 & 1 \end{pmatrix}.$$

- a) Bestem dimensionerne af $\ker(f)$ og $\text{im}(f)$.
- b) Angiv baser for $\ker(f)$ og $\text{im}(f)$.

4.28 Vis, at hvis $A \in M_{m,n}(\mathbb{F})$ og $B \in M_{n,p}(\mathbb{F})$, da er

$$\text{rank}(AB) \leq \text{rank}(A).$$

[Vink: Oversæt til et spørgsmål om lineære afbildninger.]

4.29 Den direkte sum af to $\mathbb{F}$ -vektorrums $(U, +, \cdot)$ og $(V, +, \cdot)$ er $\mathbb{F}$ -vektorrums

$$(U, +, \cdot) \oplus (V, +, \cdot) = (U \times V, +, \cdot),$$

hvor $U \times V = \{(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \in U, \mathbf{v} \in V\}$ er mængden af ordnede par af en vektor i U og en vektor i V , og hvor $(\mathbf{u}, \mathbf{v}) + (\mathbf{u}', \mathbf{v}') = (\mathbf{u} + \mathbf{u}', \mathbf{v} + \mathbf{v}')$ og $(\mathbf{u}, \mathbf{v}) \cdot a = (\mathbf{u} \cdot a, \mathbf{v} \cdot a)$. Vi forkorter normalt og skriver $U \oplus V$ i stedet for $(U, +, \cdot) \oplus (V, +, \cdot)$.

- a) Vis, at hvis familierne $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ og $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er baser for henholdsvis U og V , da er familien

$$((\mathbf{u}_1, \mathbf{0}), \dots, (\mathbf{u}_m, \mathbf{0}), (\mathbf{0}, \mathbf{v}_1), \dots, (\mathbf{0}, \mathbf{v}_n))$$

en basis for $U \oplus V$. Bemærk, at da gælder det, at $\dim(U \oplus V) = \dim(U) + \dim(V)$.

Vi antager nu, at U og V begge er underrum af et $\mathbb{F}$ -vektorrums W .

- b) Vis, at afbildningen

$$f: U \oplus V \rightarrow W$$

defineret ved $f(\mathbf{u}, \mathbf{v}) = \mathbf{u} + \mathbf{v}$ er lineær.

- c) Vis, at afbildningen

$$g: U \cap V \rightarrow U \oplus V$$

defineret ved $g(\mathbf{w}) = (\mathbf{w}, -\mathbf{w})$ er lineær og injektiv.

- d) Vis, at $\ker(f) = \text{im}(g)$.

- e) Konkluder, at hvis U og V er endeligt frembragte, da er

$$\dim(U + V) + \dim(U \cap V) = \dim(U) + \dim(V),$$

hvor $U + V = \{\mathbf{u} + \mathbf{v} \in W \mid \mathbf{u} \in U, \mathbf{v} \in V\} \subset W$. [Vink: Bemærk, at $U + V = \text{im}(f)$.]

4.30 Vi betragter vektorrummet $\mathbb{F}[t]_{\leq 3}$ af polynomier

$$p(t) = a_0 t^0 + a_1 t^1 + a_2 t^2 + a_3 t^3$$

af grad højst 3 med koefficienter i $\mathbb{F}$ samt basen (t^0, t^1, t^2, t^3) for dette vektorrum; se eksempel 4.3.22.

- a) Angiv matricen A for den lineære afbildning

$$f: \mathbb{F}[t]_{\leq 3} \rightarrow \mathbb{F}[t]_{\leq 3}$$

givet ved $f(p(t)) = p(t+1)$ med hensyn til basen (t^0, t^1, t^2, t^3) for både domænet og codomænet.

4 Vektorrum

- b) Angiv matricen B for den lineære afbildning

$$g: \mathbb{F}[t]_{\leq 3} \rightarrow \mathbb{F}[t]_{\leq 3}$$

givet ved $g(p(t)) = p(t-1)$ med hensyn til basen (t^0, t^1, t^2, t^3) for både domænet og codomænet.

- c) Hvad er sammenhængen mellem matricerne A og B ?

4.31 Lad $C^0(\mathbb{R})$ være vektorrummet af kontinuerte funktioner $f: \mathbb{R} \rightarrow \mathbb{R}$ fra eksempel 4.1.3 (9), og lad $C^1(\mathbb{R})$ være underrummet af de funktioner $\varphi: \mathbb{R} \rightarrow \mathbb{R}$, der er differentiable, og hvis afledte funktion $\varphi': \mathbb{R} \rightarrow \mathbb{R}$ er kontinuert. Lad $a \in \mathbb{R}$ være et reelt tal.

- a) Vis, at afbildningen $D: C^1(\mathbb{R}) \rightarrow C^0(\mathbb{R})$ defineret ved

$$D(\varphi)(x) = \varphi'(x)$$

er lineær.

- b) Vis, at afbildningen $I_a: C^0(\mathbb{R}) \rightarrow C^1(\mathbb{R})$ defineret ved

$$I_a(f)(x) = \int_a^x f(t) dt$$

er lineær.

- c) Bestem de sammensatte afbildninger

$$D \circ I_a: C^0(\mathbb{R}) \rightarrow C^0(\mathbb{R}) \quad \text{og} \quad I_a \circ D: C^1(\mathbb{R}) \rightarrow C^1(\mathbb{R}).$$

- d) (★) Konkluder heraf, at $C^1(\mathbb{R})$ ikke er endeligt frembragt.

[Vink: Anvend sætning 4.3.18.]

4.32 Lad $\mathbb{R}[t]_{\leq d}$ være vektorrummet af polynomier af grad højst d med reelle koefficienter fra eksempel 4.1.3 (8). Givet $d+1$ forskellige reelle tal $x_0, x_1, \dots, x_d$, definerer vi afbildningen $f: \mathbb{R}[t]_{\leq d} \rightarrow \mathbb{R}^{d+1}$ ved

$$f(p(t)) = \begin{pmatrix} p(x_0) \\ p(x_1) \\ \vdots \\ p(x_d) \end{pmatrix}.$$

- a) Vis, at afbildningen $f: \mathbb{R}[t]_{\leq d} \rightarrow \mathbb{R}^{d+1}$ er lineær.
 b) Vis, at afbildningen $f: \mathbb{R}[t]_{\leq d} \rightarrow \mathbb{R}^{d+1}$ er injektiv. [Vink: Hvor mange rødder kan et polynomium af grad $\leq d$ have?]
 c) Konkluder, at afbildningen $f: \mathbb{R}[t]_{\leq d} \rightarrow \mathbb{R}^{d+1}$ er en isomorfi.

d) (★) Er det samme tilfældet, hvis $\mathbb{R}$ erstattes af et vilkårlig legeme $\mathbb{F}$?

4.33 (★) Betragt vektorrummet V af alle C^∞ -afbildninger fra $\mathbb{R}$ til $\mathbb{R}$ med vektorsum og skalarmultiplikation defineret ved

$$(f + g)(x) = f(x) + g(x) \quad \text{og} \quad (f \cdot a)(x) = f(x) \cdot a,$$

hvor $f, g \in V$ og $a \in \mathbb{R}$.

a) Givet $a_1, \dots, a_n \in V$, vis at delmængden af V , der består af alle løsninger $f \in V$ til den homogene differentialligning

$$a_n(x) \frac{d^n}{dx^n} f(x) + \dots + a_1(x) \frac{d}{dx} f(x) = 0$$

er et underrum af V .

b) Vis, at det samme gælder for delmængden af V , der består af alle løsninger til ligningen

$$a_n(x)L_n(f) + \dots + a_1(x)L_1(f) = 0,$$

hvor $a_1, \dots, a_n \in V$ og $L_1, \dots, L_n: V \rightarrow V$ er lineære afbildninger.

c) Vis, at underrummet i (b) kan beskrives som kernen for en lineær afbildning $L: V \rightarrow V$.

d) Forklar, at (a) er et specialtilfælde af (b).

4.34 (★) Fibonacci-følgen⁶ $(f_1, f_2, f_3, \dots)$ er den følge af naturlige tal, der er defineret rekursivt ved at sætte $f_1 = f_2 = 1$ og kræve, at det for alle $n \geq 1$ gælder, at

$$f_{n+2} = f_{n+1} + f_n.$$

De første få led i følgen er således

$$f_1 = 1, f_2 = 1, f_3 = 2, f_4 = 3, f_5 = 5, f_6 = 8, f_7 = 13, f_8 = 21, \dots$$

Vi betragter denne følge som en følge af reelle tal og vil nu anvende lineær algebra til at udlede følgende lukkede formel for det n 'te led i Fibonacci-følgen:

$$f_n = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n.$$

Hertil definerer vi V til at være mængden af alle de følger af reelle tal

$$\mathbf{x} = (x_1, x_2, x_3, \dots),$$

⁶ Følgen optræder i Fibonacci's bog *Liber Abaci* fra 1202, der som den første indførte det hindu-arabiske talsystem i vesten. Formålet med bogen var at demonstrere fordelene ved dette talsystem fremfor det romerske talsystem.

4 Vektorrum

for hvilke $x_{n+2} = x_{n+1} + x_n$ for alle $n \geq 1$. Så Fibonacci-følgen $\mathbf{f} = (f_1, f_2, f_3, \dots)$ er et element i V . Vi definerer $+: V \times V \rightarrow V$ og $\cdot: V \times \mathbb{R} \rightarrow V$ ved

$$\begin{aligned}\mathbf{x} + \mathbf{y} &= (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots), \\ \mathbf{x} \cdot a &= (x_1 a, x_2 a, x_3 a, \dots).\end{aligned}$$

- a) Vis, at $(V, +, \cdot)$ er et reelt vektorrum.
- b) Lad $\mathbf{x} = (x_1, x_2, x_3, \dots)$ og $\mathbf{y} = (y_1, y_2, y_3, \dots)$ tilhøre V . Vis, at $\mathbf{x} = \mathbf{y}$, hvis og kun hvis $x_1 = y_1$ og $x_2 = y_2$.
- c) Vi betragter de to elementer $\mathbf{u}, \mathbf{v} \in V$ defineret ved

$$\begin{aligned}\mathbf{u} &= (1, 0, 1, 1, 2, 3, 5, \dots), \\ \mathbf{v} &= (0, 1, 1, 2, 3, 5, 8, \dots).\end{aligned}$$

Vis, at familien $(\mathbf{u}, \mathbf{v})$ er en basis for V og konkluder, at V har dimension 2.

- d) For alle $t \in \mathbb{R}$ betragter vi følgen

$$\mathbf{x}(t) = (t, t^2, t^3, \dots, t^n, \dots)$$

af reelle tal. Vis, at der foruden $t = 0$ netop findes to reelle tal $t = u$ og $t = v$, sådan at $\mathbf{x}(t) \in V$, og bestem disse to tal.

- e) Vis, at familien $(\mathbf{x}(u), \mathbf{x}(v))$ er en basis for V .
- f) Find endelig de entydigt bestemte reelle tal a og b , sådan at

$$\mathbf{f} = \mathbf{x}(u) \cdot a + \mathbf{x}(v) \cdot b,$$

og konkluder herfra, at den lukkede formel for Fibonacci-tallene gælder.

- 4.35 Lad $C^0(\mathbb{R})$ være $\mathbb{R}$ -vektorrummet af kontinuerte funktioner $f: \mathbb{R} \rightarrow \mathbb{R}$. For alle $N \geq 0$, definerer vi

$$\text{Trig}_N(\mathbb{R}) \subset C^0(\mathbb{R})$$

til at være underrummet, der er frembragt af familien

$$(1, \cos(x), \sin(x), \cos(2x), \sin(2x), \dots, \cos(Nx), \sin(Nx)).$$

Her betegner “1” den konstante funktion med værdi 1, mens “ $\cos(nx)$ ” og “ $\sin(nx)$ ” betegner de trigonometriske funktioner $x \mapsto \cos(nx)$ og $x \mapsto \sin(nx)$. Med andre ord består $\text{Trig}_N(\mathbb{R}) \subset C^0(\mathbb{R})$ af alle kontinuerte funktioner $f: \mathbb{R} \rightarrow \mathbb{R}$, der kan skrives på formen

$$f(x) = a_0 + \sum_{n=1}^N (\cos(nx) \cdot a_n + \sin(nx) \cdot b_n)$$

med $a_i, b_i \in \mathbb{R}$.

- a. Vis, at familien $(1, \cos(x), \sin(x))$ er lineært uafhængig, og bestem dimensionen af $\text{Trig}_1(\mathbb{R})$.

[Vink: Man kan tænke på $a_0 + \cos(x)a_1 + \sin(x)b_1 = 0$ som et lineært liningssystem i variable a_0 , a_1 og b_1 med uendeligt mange ligninger, der er parametriseret ved $x \in \mathbb{R}$. Det, der skal vises, er, at dette liningssystem kun har den ene løsning $a_0 = a_1 = b_1 = 0$.]

I det næste spørgsmål kan det anvendes uden bevis, at familien

$$(1, \cos(x), \sin(x), \cos(2x), \sin(2x))$$

er lineært uafhængig og dermed en basis for $\text{Trig}_2(\mathbb{R})$.

- b. Vis, at de kontinuerte funktioner $\cos^2(x)$ og $\sin^2(x)$ tilhører $\text{Trig}_2(\mathbb{R})$ og bestem deres koordinater med hensyn til basen ovenfor.

Trigonometriske polynomier har et væld af anvendelser, for i digital repræsentation af lyd. Hvis et signal repræsenteres af et trigonometrisk polynomium, da angiver koefficienterne a_n og b_n en alternativ og komprimeret repræsentation af signalet.

5 Egenværdier og egenrum

Vi skal nu betragte *lineære endomorfier*, som er lineære afbildninger $f: V \rightarrow V$, der har *samme* domæne og codomæne. For disse introducerer vi begreberne *egenværdi* og *egenrum*, som i sidste ende fører til den såkaldte Iwasawa-dekomposition af lineære endomorfier. Den simpleste del er de lineære endomorfier, der er *diagonaliserbare* i den forstand, at der findes en basis, hvori de er repræsenteret af en diagonalmatrix. De nye begreber egenværdi og egenrum er også vigtige begreber i sig selv. Således er observable i kvantemekanik repræsenteret af lineære endomorfier, mens de mulige værdier, som disse observable kan antage ved måling, er disse lineære endomorfiers egenværdier.

Vi antager igen i dette kapitel, at den kommutative lov gælder for multiplikation af skalarer.

5.1 Egenværdier og egenrum for kvadratiske matricer

Vi definerer først begreberne egenværdi og egenrum for kvadratiske matricer. Givet to kvadratiske matricer $A = (a_{ij}) \in M_n(\mathbb{F})$ og $B = (b_{ij}) \in M_n(\mathbb{F})$ med indgange i et legeme $\mathbb{F}$, definerer vi $A + Bt \in M_n(\mathbb{F}[t])$ til at være den kvadratiske matrix med indgange i polynomiumsringen $\mathbb{F}[t]$, der er givet ved

$$A + Bt = (a_{ij} + b_{ij}t) \in M_n(\mathbb{F}[t]).$$

Vi kan da betragte dens determinant

$$\det(A + Bt) \in \mathbb{F}[t],$$

som ifølge sætning 3.5.7 er et polynomium af grad højst n . Vi skal her kun betragte tilfældet, hvor $B = -I$ er den modsatte matrix af identitetsmatricen.

Definition 5.1.1 Hvis $\mathbb{F}$ er et legeme og $A \in M_n(\mathbb{F})$, da kaldes polynomiet

$$\chi_A(t) = \det(A - It) = \det \begin{pmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{pmatrix}$$

for det *karakteristiske polynomium* af A .

5 Egenverdier og egenrum

For eksempel er det karakteristiske polynomium af en 2×2 -matrix givet ved

$$\begin{aligned}\chi_A(t) &= \det \begin{pmatrix} a_{11} - t & a_{12} \\ a_{21} & a_{22} - t \end{pmatrix} = (a_{11} - t)(a_{22} - t) - a_{12}a_{21} \\ &= t^2 - (a_{11} + a_{22})t + (a_{11}a_{22} - a_{12}a_{21}),\end{aligned}$$

hvilket vi kan skrive mere kompakt som $\chi_A(t) = t^2 - \operatorname{tr}(A)t + \det(A)$, hvor $\operatorname{tr}(A)$ er sporet af A , som er defineret til at være summen af diagonalindgangene.

Eksempel 5.1.2 (1) For den komplekse matrix

$$A = \begin{pmatrix} 3 - i & 1 - 2i \\ 1 + 2i & 3 + i \end{pmatrix}$$

giver ovenstående formel for det karakteristiske polynomium, at

$$\chi_A(t) = t^2 - \operatorname{tr}(A)t + \det(A) = t^2 - 6t + 5 = (1 - t)(5 - t).$$

(2) Vi udregner det karakteristiske polynomium af den reelle matrix

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \\ 0 & 2 & 2 \end{pmatrix}$$

ved Laplace-udvikling af determinanten efter første søjle.

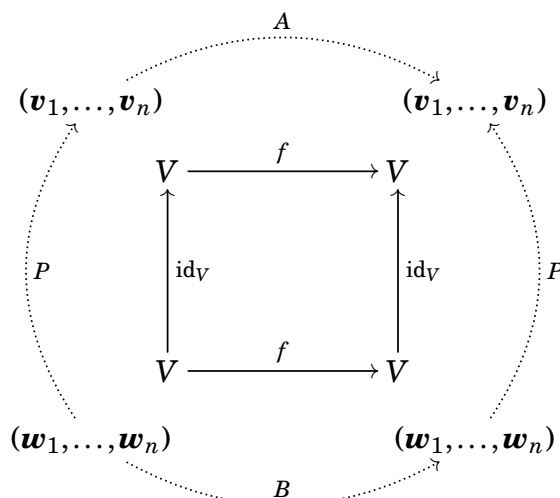
$$\begin{aligned}\chi_A(t) &= \det(A - It) = \det \begin{pmatrix} 1 - t & 2 & 0 \\ 0 & 3 - t & 1 \\ 0 & 2 & 2 - t \end{pmatrix} = (1 - t) \det \begin{pmatrix} 3 - t & 1 \\ 2 & 2 - t \end{pmatrix} \\ &= (1 - t)((3 - t)(2 - t) - 2) = (1 - t)(4 - 5t + t^2) = (1 - t)^2(4 - t)\end{aligned}$$

Bemærkning 5.1.3 Lad $f: V \rightarrow V$ være en lineær endomorfi af et endeligt frembragt $\mathbb{F}$ -vektorum, og lad $A \in M_n(\mathbb{F})$ være matricen, der repræsenterer $f: V \rightarrow V$ med hensyn til den *samme* basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for både domænet og codomænet. Vi skal nu vise, at det karakteristiske polynomium $\chi_A(t)$ kun afhænger af $f: V \rightarrow V$. Hvis også $B \in M_n(\mathbb{F})$ repræsenterer $f: V \rightarrow V$ med hensyn til den *samme* basis $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ for både domænet og codomænet, da gælder det ifølge sætning 4.4.14, at

$$B = P^{-1}AP,$$

hvor $P \in M_n(\mathbb{F})$ er matricen, der repræsenterer identitetsafbildningen $\operatorname{id}_V: V \rightarrow V$ med hensyn til baserne $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ for domænet og $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for codomænet. Den følgende figur illustrerer koordinatskiftet mellem de to baser.

5.1 Egenverdier og egenrum for kvadratiske matricer



Vi viser nu, at de karakteristiske polynomier $\chi_A(t)$ og $\chi_B(t)$ er ens, hvorfor det fælles polynomium kun afhænger af den lineære endomorfi $f: V \rightarrow V$.

Sætning 5.1.4 *Lad $A \in M_n(\mathbb{F})$ være en kvadratisk matrix med indgange i et legeme $\mathbb{F}$. For enhver invertibel matrix $P \in M_n(\mathbb{F})$ gælder det, at*

$$\chi_{P^{-1}AP}(t) = \chi_A(t).$$

Bevis Vi påstår, at det for alle $Q \in M_n(\mathbb{F}[t])$ gælder, at

$$Q(It) = (It)Q.$$

Denne påstand viser, at

$$P^{-1}(A - It)P = P^{-1}AP - P^{-1}(It)P = P^{-1}AP - P^{-1}P(It) = P^{-1}AP - It,$$

og det følger derfor fra sætning 3.2.17, at

$$\begin{aligned} \chi_A(t) &= \det(A - It) = \det(P)^{-1} \det(A - It) \det(P) = \det(P^{-1}(A - It)P) \\ &= \det(P^{-1}AP - It) = \chi_{P^{-1}AP}(t) \end{aligned}$$

som ønsket. Vi mangler at vise påstanden. Den (i, k) 'te indgang i $Q(It)$ lig med

$$\sum_{j=1}^n q_{ij} \cdot \delta_{jk} t = q_{ik} \cdot t,$$

mens den (i, k) 'te indgang i $(It) \cdot Q$ er lig med

$$\sum_{j=1}^n \delta_{ij} t \cdot q_{jk} = t \cdot q_{ik}.$$

Men $q_{ik} \cdot t = t \cdot q_{ik}$, da den kommutative lov gælder for multiplikation af polynomier. Vi har nu vist påstanden og dermed sætningen. $\square$

5 Egenverdier og egenrum

Lad $A \in M_n(\mathbb{F})$ være en kvadratisk matrix. Ved at substituere skalaren $\lambda \in \mathbb{F}$ for t i det karakteristiske polynomium $\chi_A(t) \in \mathbb{F}[t]$ få vi skalaren

$$\chi_A(\lambda) = \det(A - I\lambda) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} \in \mathbb{F}.$$

Vi skriver derfor også $I\lambda$ for diagonalmatricen $\text{diag}(\lambda, \dots, \lambda)$.

Lemma 5.1.5 Hvis $\mathbb{F}$ er et legeme, og hvis $A \in M_n(\mathbb{F})$ er en kvadratisk matrix, da er

$$N_{A-I\lambda} = \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \mathbf{v}\lambda\} \subset \mathbb{F}^n.$$

Bevis Per definition er $\mathbf{v} \in N_{A-I\lambda}$, hvis og kun hvis $(A - I\lambda)\mathbf{v} = \mathbf{0}$, og dette gælder, hvis og kun hvis $A\mathbf{v} = (I\lambda)\mathbf{v}$. Vi bemærker nu, at det for alle $\mathbf{v} \in \mathbb{F}^n$ gælder, at

$$(I\lambda)\mathbf{v} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix} = \begin{pmatrix} v_1 \lambda \\ \vdots \\ v_n \lambda \end{pmatrix} = \mathbf{v}\lambda,$$

idet vi har antaget, at den kommutative lov gælder for multiplikation af skalarer. $\square$

Vi udtrykker lemma 5.1.5 ved at sige, at på underrummet $N_{A-I\lambda} \subset \mathbb{F}^n$ er den venstre multiplikation med A givet ved højre *skalering* med vægt λ .

Eksempel 5.1.6 Vi udregner nulrummet N_{A-I7} for matricen

$$A = \begin{pmatrix} 5 & 8 \\ 1 & 3 \end{pmatrix}$$

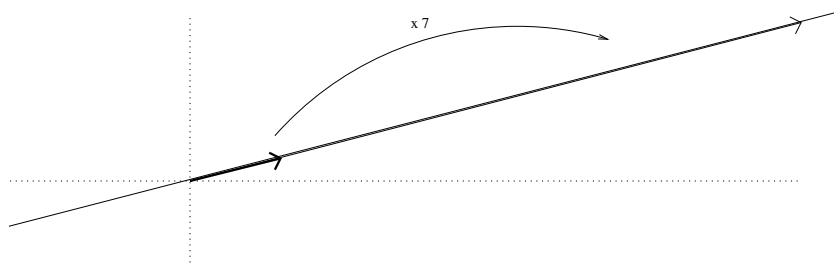
ved at anvende rækkeoperationer til at omdanne matricen $B = A - I7$ til en matrix B' på reduceret echelonform.

$$\begin{aligned} B &= \begin{pmatrix} -2 & 8 \\ 1 & -4 \end{pmatrix} \quad (-\tfrac{1}{2}) \cdot R_1 \\ &\quad \begin{pmatrix} 1 & -4 \\ 1 & -4 \end{pmatrix} \quad +(-1)R_1 \\ B' &= \begin{pmatrix} 1 & -4 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Vi aflæser, at nulrummet $N_{A-I7} \subset \mathbb{F}^2$ er det en-dimensionale underrum med basis

$$(\mathbf{v}_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}).$$

5.1 Eigenverdier og egenrum for kvadratiske matricer



Figur 5.1: På nulrummet N_{A-I7} fra eksempel 5.1.6 er venstre multiplikation med A lig med højre skalering med 7.

Definition 5.1.7 Lad A være en $n \times n$ -matrix med indgange i et legeme $\mathbb{F}$.

- (1) En skalar $\lambda \in \mathbb{F}$ er en *eigenverdi* for A , hvis $N_{A-I\lambda} \neq \{\mathbf{0}\}$.
- (2) Hvis $\lambda \in \mathbb{F}$ er en eigenverdi for A , så kaldes nulrummet $N_{A-I\lambda}$ for *egenrummet* hørende til eigenverdien $\lambda \in \mathbb{F}$.
- (3) Hvis $\lambda \in \mathbb{F}$ er en eigenverdi for A , så kaldes enhver vektor $\mathbf{v} \in N_{A-I\lambda}$, der ikke er nulvektoren, for en *eigenvektor* for A hørende til eigenverdien $\lambda \in \mathbb{F}$.

Ifølge lemma 5.1.5 er en vektor $\mathbf{v}$ en eigenvektor for A hørende til eigenverdien λ , hvis og kun hvis den ikke er nulvektoren og opfylder $A\mathbf{v} = \mathbf{v}\lambda$. Vi bemærker, at mens egenrummet $N_{A-I\lambda}$ hørende til eigenverdi $\lambda \in \mathbb{F}$ er entydigt bestemt, så er der mange eigenvektorer $\mathbf{v}$ hørende til eigenverdi λ . Hvis nemlig $\mathbf{v}$ er en eigenvektor hørende til eigenverdi λ , og $a \in \mathbb{F}$ er forskellig fra nul, da er $a\mathbf{v}$ også en eigenvektor hørende til λ .

Sætning 5.1.8 Eigenverdierne af en kvadratisk matrix $A \in M_n(\mathbb{F})$ med indgange i et legeme $\mathbb{F}$ er præcis rødderne i $\mathbb{F}$ af det karakteristiske polynomium $\chi_A(t) \in \mathbb{F}[t]$.

Bevis Per definition er $\lambda \in \mathbb{F}$ en eigenverdi for A , hvis den homogene ligning

$$(A - I\lambda)\mathbf{x} = \mathbf{0}$$

har en ikke-triviell løsning $\mathbf{x} = \mathbf{v} \neq \mathbf{0}$, og da matricen $A - I\lambda$ er kvadratisk, er dette tilfældet, hvis og kun hvis matricen $A - I\lambda$ ikke er invertibel. Ifølge sætning 3.4.1 er matricen $A - I\lambda$ invertibel, hvis og kun hvis dens determinant $\chi_A(\lambda) = \det(A - I\lambda)$ er invertibel, og da $\mathbb{F}$ er et legeme, er $\chi_A(\lambda)$ invertibel, hvis og kun hvis $\chi_A(\lambda) \neq 0$. Vi har hermed vist, at $\lambda \in \mathbb{F}$ er en eigenverdi for A , hvis og kun hvis $\chi_A(\lambda) = 0$ som ønsket. $\square$

5 Egenverdier og egenrum

Definition 5.1.9 Lad $\lambda \in \mathbb{F}$ være en egenverdi for en kvadratisk matrix $A \in M_n(\mathbb{F})$. Den *geometriske multiplicitet* af λ er dimensionen af egenrummet $N_{A-I\lambda}$ hørende til λ , mens den *algebraiske multiplicitet* af λ er det største hele tal $k \geq 1$, for hvilket der findes $q(t) \in \mathbb{F}[t]$, sådan at

$$\chi_A(t) = (\lambda - t)^k \cdot q(t).$$

Vi beviser senere i sætning 5.3.7, at den algebraiske multiplicitet af en egenverdi altid er større end eller lig med den geometriske multiplicitet.

Eksempel 5.1.10 (1) Identitetsmatricen $I \in M_n(\mathbb{F})$ har karakteristisk polynomium

$$\chi_I(t) = \det(I - It) = (1 - t)^n.$$

Sætning 5.1.8 viser derfor, at $\lambda = 1$ er eneste egenverdi. Den algebraiske multiplicitet er lig med n , og da $N_{I-I1} = N_O = \mathbb{F}^n$, er også den geometriske multiplicitet lig med n . Bemærk nemlig, at $I\mathbf{x} = \mathbf{x} = \mathbf{x} \cdot 1$ for alle $\mathbf{x} \in \mathbb{F}^n$.

(2) Nulmatricen $O \in M_n(\mathbb{F})$ har karakteristisk polynomium

$$\chi_O(t) = \det(O - It) = \det(-It) = (-t)^n.$$

Så $\lambda = 0$ er eneste egenverdi med algebraisk multiplicitet n , og da $N_{O-I0} = N_O = \mathbb{F}^n$, er den geometriske multiplicitet også lig med n . For $O\mathbf{x} = \mathbf{0} = \mathbf{x} \cdot 0$ for alle $\mathbf{x} \in \mathbb{F}^n$.

(3) Det karakteristiske polynomium af matricen

$$A = \begin{pmatrix} 5 & 8 \\ 1 & 3 \end{pmatrix}$$

fra eksempel 5.1.6 er

$$\chi_A(t) = \det \begin{pmatrix} 5-t & 8 \\ 1 & 3-t \end{pmatrix} = t^2 - 8t + 7 = (t-1)(t-7).$$

Så A har egenverdier $\lambda_1 = 1$ og $\lambda_2 = 7$, der begge har algebraisk multiplicitet 1. Vi udregnede egenrummet hørende til egenverdien $\lambda_2 = 7$ i eksempel 5.1.6. Da det er en-dimensionalt, har $\lambda_2 = 1$ geometrisk multiplicitet 1. Vi udregner tilsvarende, at

$$N_{A-I\lambda_1} = N_{A-I} = \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \cdot t \mid t \in \mathbb{F} \right\} \subset \mathbb{F}^2,$$

hvilket viser, at også $\lambda_1 = 1$ har geometrisk multiplicitet 1.

(4) Det karakteristiske polynomium af matricen

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{F})$$

5.1 Egenverdier og egenrum for kvadratiske matricer

udregnes til at være

$$\chi_A(t) = \det \begin{pmatrix} -t & -1 \\ 1 & -t \end{pmatrix} = t^2 + 1.$$

Hvis $\mathbb{F} = \mathbb{R}$, så har matricen A altså ingen egenverdier, men hvis $\mathbb{F} = \mathbb{C}$, da er

$$\chi_A(t) = (t - i)(t + i),$$

så A har egenverdierne $\lambda_1 = +i$ og $\lambda_2 = -i$, begge med algebraisk multiplicitet 1. For at bestemme egenrummet hørende til $\lambda_1 = +i$ omdanner vi matricen $B = A - Ii$ ved hjælp af rækkeoperationer til en matrix B' på reduceret echelonform.

$$\begin{aligned} B &= \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \quad i \cdot R_1 \\ &\quad \begin{pmatrix} 1 & -i \\ 1 & -i \end{pmatrix} \quad + (-1)R_1 \\ B' &= \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Egenrummet for $A \in M_2(\mathbb{C})$ hørende til $\lambda_1 = +i$ er altså givet ved

$$N_{A-Ii} = \left\{ \begin{pmatrix} i \\ 1 \end{pmatrix} \cdot t \mid t \in \mathbb{C} \right\} \subset \mathbb{C}^2.$$

For $\lambda_2 = -i$, omdanner vi tilsvarende $C = A - I(-i)$ ved hjælp af rækkeoperationer til en matrix C' på reduceret echelonform.

$$\begin{aligned} C &= \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \quad (-i) \cdot R_1 \\ &\quad \begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix} \quad + (-1)R_1 \\ C' &= \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Vi aflæser egenrummet for $A \in M_2(\mathbb{C})$ hørende til egenverdien $\lambda_2 = -i$ til at være

$$N_{A+Ii} = \left\{ \begin{pmatrix} -i \\ 1 \end{pmatrix} \cdot t \mid t \in \mathbb{C} \right\} \subset \mathbb{C}^2.$$

Egenverdierne $\lambda_1 = +i$ og $\lambda_2 = -i$ har dermed også begge geometrisk multiplicitet 1.

(5) Det karakteristiske polynomium af matricen

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(\mathbb{F})$$

5 Egenverdier og egenrum

er $\chi_A(t) = (1-t)^2$, hvilket viser, at A har $\lambda = 1$ som eneste egenverdi med algebraisk multiplicitet 2. Det tilhørende egenrum

$$N_{A-I1} = \left\{ \mathbf{x} \in \mathbb{F}^2 \mid \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mathbf{x} = \mathbf{0} \right\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot t \mid t \in \mathbb{F} \right\} \subset \mathbb{F}^2$$

har imidlertid dimension 1, så egenverdien $\lambda = 1$ har altså geometrisk multiplicitet 1, hvilket er mindre end den algebraiske multiplicitet 2.

Vi viser nu, at en familie af egenvektorer hørende til parvis forskellige egenverdier nødvendigvis er lineært uafhængig.

Sætning 5.1.11 *Lad $\mathbb{F}$ være et legeme, og lad A være en $n \times n$ -matrix med indgange i $\mathbb{F}$. Lad $(\lambda_1, \dots, \lambda_p)$ være en familie af parvis forskellige egenverdier for A , og lad $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ være en familie af vektorer i $\mathbb{F}^n$, sådan at $\mathbf{v}_i$ er en egenvektor hørende til λ_i for alle $1 \leq i \leq p$. Da er familien $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ lineært uafhængig. Specielt er $p \leq n$.*

Bevis Vi beviser udsagnet ved induktion på $p \geq 0$. Hvis $p = 0$, så gælder udsagnet, da den tomme familie altid er lineært uafhængig. Vi antager derfor, at udsagnet allerede er bevist for $p = r - 1$ og beviser, at det gælder for $p = r$. Så vi lader

$$\mathbf{v}_1 c_1 + \dots + \mathbf{v}_r c_r = \mathbf{0}$$

være en linearkombination af $(\mathbf{v}_1, \dots, \mathbf{v}_r)$, som er lig med nulvektoren, og skal vise, at $c_i = 0$ for alle $1 \leq i \leq r$. Hertil bemærker vi, at

$$(A - I\lambda_i)(\mathbf{v}_j) = \mathbf{v}_j \cdot (\lambda_j - \lambda_i),$$

idet $A\mathbf{v}_j = \mathbf{v}_j \lambda_j$ og $(I\lambda_i)\mathbf{v}_j = \mathbf{v}_j \lambda_i$. Dette viser, at

$$(A - I\lambda_1) \dots (A - I\lambda_{r-1})(\mathbf{v}_1 c_1 + \dots + \mathbf{v}_r c_r) = \mathbf{v}_r c_r (\lambda_r - \lambda_1) \dots (\lambda_r - \lambda_{r-1}).$$

Da venstresiden per antagelse er lig med $\mathbf{0}$, konkluderer vi, at $c_r = 0$. Dermed er

$$\mathbf{v}_1 c_1 + \dots + \mathbf{v}_{r-1} c_{r-1} = \mathbf{0},$$

og den induktive hypotese viser derfor, at $c_1 = \dots = c_{r-1} = 0$. Altså er $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ lineært uafhængig, hvilket viser induktionsskridtet. Endelig følger den sidste påstand, at det ifølge sætning 4.3.18 nødvendigvis gælder, at $p \leq n$. $\square$

Korollar 5.1.12 Lad $\mathbb{F}$ være et legeme, og lad $A \in M_n(\mathbb{F})$ være en kvadratisk matrix med indgange i $\mathbb{F}$. Hvis $(\lambda_1, \dots, \lambda_n)$ er en familie af n parvis forskellige egenverdier for A , da er enhver familie $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ af tilhørende egenvektorer en basis for $\mathbb{F}^n$.

Bevis Familien af egenvektorer $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er lineært uafhængig ifølge sætning 5.1.11, og da den består af $n = \dim_{\mathbb{F}}(\mathbb{F}^n)$ vektorer, viser sætning 4.3.24, at den er en basis. $\square$

Eksempel 5.1.13 Vi så i eksempel 5.1.10, at matricen

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{C})$$

har egenverdier $\lambda_1 = +i$ og $\lambda_2 = -i$, og at vektorerne

$$\mathbf{v}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

er egenvektorer for henholdsvis λ_1 og λ_2 . Dermed er $(\mathbf{v}_1, \mathbf{v}_2)$ en familie af egenvektorer hørende til $n = 2$ forskellige egenverdier for A . Ifølge korollar 5.1.12 er $(\mathbf{v}_1, \mathbf{v}_2)$ dermed en basis for $\mathbb{C}^2$. (Dette følger også af, at $\det(\mathbf{v}_1 \ \mathbf{v}_2) = 2i \neq 0$.)

5.2 Diagonaliserbare matricer

I dette afsnit studerer vi de diagonaliserbare matricer, som vi vil definere.

Definition 5.2.1 En kvadratisk matrix $A \in M_n(\mathbb{F})$ kaldes *diagonaliserbar*, hvis der findes en invertibel matrix $P \in M_n(\mathbb{F})$, sådan at $P^{-1}AP$ er en diagonalmatrix D .

Lad $A \in M_n(\mathbb{F})$ være en diagonaliserbar matrix, og lad $P \in M_n(\mathbb{F})$ være en invertibel matrix, sådan at $D = P^{-1}AP$ er en diagonalmatrix. Da matricen P er invertibel, udgør den familie $(\mathbf{v}_1, \dots, \mathbf{v}_n)$, der består af søjler i P , en basis for $\mathbb{F}^n$. Den lineære endomorfi $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ defineret ved $f(\mathbf{x}) = A\mathbf{x}$ er da repræsenteret af matricen A med hensyn til standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ for både domænet og codomænet. Bemærkning 5.1.3 viser, at $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ er repræsenteret af $D = P^{-1}AP$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for både domænet og codomænet. At A er diagonaliserbar udtrykker altså, at der findes en basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for $\mathbb{F}^n$, sådan at matricen, der repræsenterer $f: \mathbb{F}^n \rightarrow \mathbb{F}^n$ med hensyn til denne basis for både domænet og codomænet, er en diagonalmatrix.

5 Egenverdier og egenrum

Sætning 5.2.2 Lad $\mathbb{F}$ være et legeme, lad $A, P \in M_n(\mathbb{F})$ være kvadratiske matricer, og lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ være familien af søjlevektorer i P . Da er følgende udsagn ækvivalente:

- (1) Matricen P er invertibel og $P^{-1}AP$ er en diagonalmatrix.
- (2) Familien $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er en basis for $\mathbb{F}^n$ og består af egenvektorer for A .

Hvis dette er tilfældet og hvis $P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n)$, så gælder det endvidere, at $\mathbf{v}_i$ er en egenvektor for A hørende til egenværdien λ_i for alle $1 \leq i \leq n$.

Bevis Ifølge lemma 4.3.9 er $P \in M_n(\mathbb{F})$ invertibel, hvis og kun hvis familien

$$(\mathbf{v}_1, \dots, \mathbf{v}_n) = (P\mathbf{e}_1, \dots, P\mathbf{e}_n)$$

af søjlevektorer i P er en basis for $\mathbb{F}^n$. Endvidere er $P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n)$, hvis og kun hvis $P^{-1}AP\mathbf{e}_i = \mathbf{e}_i\lambda_i$ for alle $1 \leq i \leq n$, hvilket gælder, hvis og kun hvis $AP\mathbf{e}_i = P\mathbf{e}_i\lambda_i$ for alle $1 \leq i \leq n$. Idet $P\mathbf{e}_i = \mathbf{v}_i$, gælder dette, hvis og kun hvis $A\mathbf{v}_i = \mathbf{v}_i\lambda_i$ for alle $1 \leq i \leq n$, hvilket netop betyder, at basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ består af egenvektorer for A . $\square$

Korollar 5.2.3 En kvadratisk matrix $A \in M_n(\mathbb{F})$ er diagonaliserbar, hvis og kun hvis der findes en basis for $\mathbb{F}^n$ bestående af egenvektorer for A .

Bevis Dette følger umiddelbart fra definition 5.2.1 og sætning 5.2.2. $\square$

Eksempel 5.2.4 Vi undersøger om matricen

$$A = \begin{pmatrix} -4 & -6 \\ 1 & 1 \end{pmatrix}$$

er diagonaliserbar. Dette er ifølge sætning 5.2.2 tilfældet, hvis og kun hvis der findes en basis $(\mathbf{v}_1, \mathbf{v}_2)$ for $\mathbb{F}^2$, der består af egenvektorer for A . Vi udregner

$$\chi_A(t) = t^2 - \text{tr}(A)t + \det(A) = t^2 + 3t + 2 = (-2 - t)(-1 - t).$$

Altså har A to forskellige egenverdier $\lambda_1 = -2$ og $\lambda_2 = -1$ og er derfor diagonaliserbar ifølge korollar 5.1.12. Vi bestemmer egenrummet hørende til $\lambda_1 = -2$ og omdanner derfor matricen $B = A - I(-2)$ til en matrix B' på reduceret echelonform.

$$\begin{aligned} B &= \begin{pmatrix} -2 & -6 \\ 1 & 3 \end{pmatrix} \quad (-\tfrac{1}{2}) \cdot R_1 \\ &\quad \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix} \quad + (-1)R_1 \\ B' &= \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Heraf ser vi, at

$$N_{A-I(-2)} = \left\{ \begin{pmatrix} -3 \\ 1 \end{pmatrix} \cdot t \mid t \in \mathbb{F} \right\},$$

og tilsvarende ses, at

$$N_{A-I(-1)} = \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \cdot t \mid t \in \mathbb{F} \right\}.$$

Sætning 5.2.2 viser derfor, at

$$P = (\mathbf{v}_1 \quad \mathbf{v}_2) = \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}$$

er invertibel, og at

$$P^{-1}AP = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix} = \text{diag}(-2, -1),$$

hvilket let eftervises.

Eksempel 5.2.5 Vi undersøger om matricen

$$A = \begin{pmatrix} -1 & 0 & -2 \\ 3 & 2 & 2 \\ 1 & -1 & 3 \end{pmatrix}$$

er diagonaliserbar. Vi begynder med at udregne det karakteristiske polynomium.

$$\begin{aligned} \chi_A(t) &= \det \begin{pmatrix} -1-t & 0 & -2 \\ 3 & 2-t & 2 \\ 1 & -1 & 3-t \end{pmatrix} \stackrel{(D6')}{=} \det \begin{pmatrix} 2-t & 2-t & 0 \\ 3 & 2-t & 2 \\ 1 & -1 & 3-t \end{pmatrix} \\ &\stackrel{(D2')}{=} (2-t) \det \begin{pmatrix} 1 & 1 & 0 \\ 3 & 2-t & 2 \\ 1 & -1 & 3-t \end{pmatrix} \stackrel{(D6')}{=} (2-t) \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1-t & 2 \\ 0 & -2 & 3-t \end{pmatrix} \\ &\stackrel{(D6)}{=} (2-t) \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1-t & 2 \\ 0 & -2 & 3-t \end{pmatrix} \stackrel{(D6)}{=} (2-t) \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-t & 2 \\ 0 & 1-t & 3-t \end{pmatrix} \\ &\stackrel{(D2)}{=} (2-t)(1-t) \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 3-t \end{pmatrix} \stackrel{(D6)}{=} (2-t)(1-t) \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1-t \end{pmatrix} \\ &= (2-t)(1-t)^2. \end{aligned}$$

Vi ser herfra, at $\lambda_1 = 1$ og $\lambda_2 = 2$ er de eneste egenverdier af A , og at deres algebraiske multipliciteter er henholdsvis 2 og 1.

5 Egenverdier og egenrum

For at udregne egenrummet N_{A-I} hørende til $\lambda_1 = 1$ anvender vi rækkeoperationer til at omdanne $B = A - I$ til en matrix B' på reduceret echelonform.

$$\begin{aligned}
 B &= \begin{pmatrix} -2 & 0 & -2 \\ 3 & 1 & 2 \\ 1 & -1 & 2 \end{pmatrix} \begin{array}{l} +2 \cdot R_3 \\ +(-3) \cdot R_3 \end{array} \\
 &\begin{pmatrix} 0 & -2 & 2 \\ 0 & 4 & -4 \\ 1 & -1 & 2 \end{pmatrix} R_1 \leftrightarrow R_3 \\
 &\begin{pmatrix} 1 & -1 & 2 \\ 0 & 4 & -4 \\ 0 & -2 & 2 \end{pmatrix} \frac{1}{4} \cdot R_2 \\
 &\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -1 \\ 0 & -2 & 2 \end{pmatrix} \begin{array}{l} +R_2 \\ +3R_2 \end{array} \\
 B' &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

Vi ser heraf, at egenrummet hørende til $\lambda_1 = 1$ er givet ved

$$N_{A-I} = \left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} t \mid t \in \mathbb{F} \right\} \subset \mathbb{F}^3,$$

og vi ser tilsvarende, at egenrummet N_{A-I_2} hørende til $\lambda_2 = 2$ er givet ved

$$N_{A-I_2} = \left\{ \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} t \mid t \in \mathbb{F} \right\} \subset \mathbb{F}^3.$$

Den geometriske multiplicitet af både $\lambda = 1$ og $\lambda = 2$ er altså lig med 1. De to egenrum udspænder et underrum af $\mathbb{F}^3$ af dimension 2, og derfor har $\mathbb{F}^3$ ikke nogen basis, der består af egenvektorer for A . Så A er ikke diagonaliserbar.

Bemærkning 5.2.6 Ved udregning af det karakteristiske polynomium er der altid risiko for at lave fejl, hvilket vil gøre alle følgende udregninger forkerte. En sådan fejl er dog let at opdage. For hvis λ ikke er en egenverdi for A , så er $B = A - I\lambda$ en invertibel matrix, og Gauss-elimination vil da bestemme den tilhørende rækkeækvivalente matrix på reduceret echelonform til at være $B' = I$. Så hvis Gauss-elimination af $B = A - I\lambda$ giver $B' = I$, så ved vi altså, at λ ikke er en egenverdi.

5.2 Diagonaliserbare matricer

Vi viser nu, at hvis $A, B \in M_n(\mathbb{F})$ er *samtidigt diagonaliserbare* i den forstand, at der findes en basis for $\mathbb{F}^n$, der består af egenvektorer for dem begge, da kommuterer A og B nødvendigvis med hinanden.

Sætning 5.2.7 *Lad A og B være to $n \times n$ -matricer med indgange i et legeme $\mathbb{F}$. Hvis der findes en basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for $\mathbb{F}^n$ bestående af vektorer, der både er egenvektorer for A og B , så gælder det nødvendigvis, at $AB = BA$.*

Bevis Ifølge sætning 5.2.2 er matricen $P = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n)$ invertibel, og matricerne

$$P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n)$$

og

$$P^{-1}BP = \text{diag}(\mu_1, \dots, \mu_n)$$

er diagonalmatricer. Vi konkluderer derfor, at

$$AB = P \text{diag}(\lambda_1, \dots, \lambda_n) P^{-1} P \text{diag}(\mu_1, \dots, \mu_n) P^{-1} = P \text{diag}(\lambda_1 \mu_1, \dots, \lambda_n \mu_n) P^{-1}$$

og

$$BA = P \text{diag}(\mu_1, \dots, \mu_n) P^{-1} P \text{diag}(\lambda_1, \dots, \lambda_n) P^{-1} = P \text{diag}(\mu_1 \lambda_1, \dots, \mu_n \lambda_n) P^{-1}.$$

Heraf følger som ønsket, at $AB = BA$, idet vi har antaget, at den kommutative lov gælder for multiplikation af skalarer. $\square$

Bemærkning 5.2.8 Til trods for, at sætning 5.2.7 er ganske let at vise, så har den enorme konsekvenser, idet den er grundlaget for Heisenbergs ubestemthedsrelationer. I kvantemekanik svarer observable og tilstande til henholdsvis matricer og vektorer, og en observabel A har en veldefineret værdi λ i tilstanden $\mathbf{v}$, hvis og kun hvis $\mathbf{v}$ er en egenvektor for A med egenværdi λ . Sætning 5.2.7 siger derfor, at hvis $AB \neq BA$, så er værdierne af de observable A og B ikke samtidigt veldefinerede.

Hvis matricen A er diagonaliserbar, og hvis $P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n)$, da følger det fra sætning 5.1.4, at det karakteristiske polynomium faktoriserer som et produkt

$$\begin{aligned} \chi_A(t) &= \chi_{P^{-1}AP}(t) = \det(\text{diag}(\lambda_1 - t, \dots, \lambda_n - t)) \\ &= (\lambda_1 - t) \cdots (\lambda_n - t) \end{aligned}$$

af førstegradspolynomier. Det omvendte udsagn gælder imidlertid ikke, medmindre egenværdierne $\lambda_1, \dots, \lambda_n$ er parvis forskellige. Vi har i stedet følgende resultat.

5 Egenverdier og egenrum

Sætning 5.2.9 Lad $A \in M_n(\mathbb{F})$ være en kvadratisk matrix med indgange i et legeme $\mathbb{F}$. Da er følgende udsagn ækvivalente:

- (1) Det karakteristiske polynomium er et produkt af førstegradspolynomier,

$$\chi_A(t) = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t).$$

- (2) Der findes en invertibel matrix $P \in M_n(\mathbb{F})$, sådan at matricen $P^{-1}AP$ er en øvre triangulær matrix.

Bevis Vi antager først (1) og beviser (2). Beviset er ved induktion på $n \geq 0$. Tilfældet $n = 0$ gælder, da den tomme matrix både er invertibel og øvre triangulær. Så vi antager, at (1) medfører (2) for $n = r - 1$ og viser, at (1) medfører (2) for $n = r$. Vi vælger en egenvektor $\mathbf{v}_1$ hørende til λ_1 og supplerer ved hjælp af sætning 4.3.24 til en basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for $\mathbb{F}^n$. Matricen $Q = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n)$ er da invertibel (lemma 4.3.9), og

$$B = Q^{-1}AQ = \left(\begin{array}{c|ccc} \lambda_1 & b_{12} & \dots & b_{1p} \\ \hline 0 & & & \\ \vdots & & B_{11} & \\ 0 & & & \end{array} \right),$$

fordi $Q^{-1}AQ\mathbf{e}_1 = Q^{-1}A\mathbf{v}_1 = Q^{-1}\mathbf{v}_1\lambda_1 = \mathbf{e}_1\lambda_1$. Ifølge sætning 5.1.4 er $\chi_B(t) = \chi_A(t)$, og ved Laplace-udvikling af determinanten efter første søjle får vi endvidere, at

$$\chi_B(t) = \det(B - I_r t) = (\lambda_1 - t) \det(B_{11} - I_{r-1} t) = (\lambda_1 - t) \chi_{B_{11}}(t).$$

Per entydighed af polynomiumsdivision¹ følger derfor, at

$$\chi_{B_{11}}(t) = (\lambda_2 - t) \cdot \dots \cdot (\lambda_r - t).$$

Den induktive antagelse, at (1) medfører (2) for $n = r - 1$, viser nu, at der findes en invertibel matrix $R_{11} \in M_{r-1}(\mathbb{F})$, sådan at matricen $C_{11} = R_{11}^{-1}B_{11}R_{11} \in M_{r-1}(\mathbb{F})$ er øvre triangulær. Vi lader da

$$R = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & R_{11} & \\ 0 & & & \end{array} \right) \in M_r(\mathbb{F})$$

¹ Dette følger fra Euklids algoritme. Et bevis kan findes på side 173–174 i Serge Lang, Algebra. Revised third edition. Graduate Texts in Mathematics, 211. Springer-Verlag, New York, 2002.

og $P = QR \in M_r(\mathbb{F})$. Da er matricen

$$C = P^{-1}AP = R^{-1}Q^{-1}AQR = R^{-1}BR = \left(\begin{array}{c|ccc} \lambda_1 & c_{12} & \cdots & c_{1r} \\ \hline 0 & & & \\ \vdots & & C_{11} & \\ 0 & & & \end{array} \right) \in M_r(\mathbb{F})$$

øvre triangulær, hvilket viser induktionsskridtet og derfor (2). Endelig følger den omvendte implikation, at (2) medfører (1), fra sætning 5.1.4 og sætning 3.3.3. $\square$

Vi bemærker, at algebraens fundamentalsætning (sætning 3.5.10) medfører, at den første betingelse (1) i sætning 5.2.9 altid er opfyldt for $\mathbb{F} = \mathbb{C}$. Dette gælder dog ikke nødvendigvis for $\mathbb{F} = \mathbb{R}$. Forskellen er afspejlet i det følgende resultat.

Sætning 5.2.10 (1) Hvis $A \in M_n(\mathbb{R})$ er en symmetrisk reel kvadratisk matrix, så findes der en invertibel reel matrix $P \in M_n(\mathbb{R})$, sådan at $P^{-1}AP \in M_n(\mathbb{R})$ er øvre triangulær.
 (2) Hvis $A \in M_n(\mathbb{C})$ er en vilkårlig kompleks kvadratisk matrix, så findes der en invertibel kompleks matrix $P \in M_n(\mathbb{C})$, sådan at $P^{-1}AP \in M_n(\mathbb{C})$ er øvre triangulær.

Bevis Vi begynder med at vise (2). Ifølge korollar 3.5.11 er ethvert polynomium med komplekse koefficienter et produkt af førstegradspolynomier. Dette gælder specielt for det karakteristiske polynomium $\chi_A(t)$, og udsagnet (2) følger derfor fra sætning 5.2.9.

Vi viser dernæst (1). Vi lader $B \in M_n(\mathbb{C})$ være matricen $A \in M_n(\mathbb{R})$ betragtet som en kompleks matrix. Det følger som før fra korollar 3.5.11, at

$$\chi_B(t) = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t) \in \mathbb{C}[t],$$

hvor $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ er egenverdierne af B . Vi påstår, at disse egenverdier alle er reelle tal. Da den reelle matrix A er symmetrisk, er den komplekse matrix B hermitisk. Lad nu $\lambda \in \mathbb{C}$ være en egenverdi for B , og lad $\mathbf{0} \neq \mathbf{z} \in \mathbb{C}^n$ være en dertil hørende egenvektor. Da er $B\mathbf{z} = \mathbf{z}\lambda$, og derfor er $\mathbf{z}^*B^* = \lambda^*\mathbf{z}^*$, hvilket medfører, at

$$\lambda^*\mathbf{z}^*\mathbf{z} = \mathbf{z}^*B^*\mathbf{z} = \mathbf{z}^*B\mathbf{z} = \mathbf{z}^*\mathbf{z}\lambda,$$

idet $B^* = B$. Men $\mathbf{z} \neq \mathbf{0}$, så $\mathbf{z}^*\mathbf{z} \neq 0$, og vi konkluderer derfor, at $\lambda^* = \lambda$, hvilket viser påstanden. Polynomiumsidentiteten ovenfor viser derfor, at

$$\chi_A(t) = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t) \in \mathbb{R}[t]$$

med $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, og derfor følger (1) ved at anvende sætning 5.2.9 som før. $\square$

5 Egenverdier og egenrum

Eksempel 5.2.11 Vi undersøger, om den komplekse matrix

$$A = \begin{pmatrix} 1 & 0 & -i \\ i & 1 & -2+2i \\ 0 & 0 & 1 \end{pmatrix}$$

er diagonaliserbar. Vi udregner det karakteristiske polynomium ved Laplace-udvikling langs sidste række.

$$\chi_A(t) = \det \begin{pmatrix} 1-t & 0 & -i \\ i & 1-t & -2+2i \\ 0 & 0 & 1-t \end{pmatrix} = (1-t) \det \begin{pmatrix} 1-t & 0 \\ i & 1-t \end{pmatrix} = (1-t)^3$$

Så $\lambda = 1$ er den eneste egenverdi, og vi anvender Gauss-elimination på $B = A - I$ for at bestemme det tilhørende egenrum N_{A-I} .

$$\begin{aligned} B &= \begin{pmatrix} 0 & 0 & -i \\ i & 0 & -2+2i \\ 0 & 0 & 0 \end{pmatrix} & R_1 \leftrightarrow R_2 \\ &\begin{pmatrix} i & 0 & -2+2i \\ 0 & 0 & -i \\ 0 & 0 & 0 \end{pmatrix} & \begin{array}{l} R_1 \cdot (-i) \\ R_2 \cdot i \end{array} \\ &\begin{pmatrix} 1 & 0 & 2+2i \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} & +R_2 \cdot (-\frac{2-2i}{8}) \\ B' &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Heraf aflæser vi, at N_{A-I} er det en-dimensionale underrum

$$N_{A-I} = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} t \mid t \in \mathbb{C} \right\} \subset \mathbb{C}^3.$$

Der findes altså ikke en basis af egenvektorer for A , og derfor er A ifølge korollar 5.2.3 ikke diagonaliserbar. Sætning 5.2.9 viser imidlertid, at der findes en invertibel matrix P , sådan at $P^{-1}AP$ er øvre triangulær. For vi har med

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{at} \quad P^{-1}AP = \begin{pmatrix} 1 & i & -2+2i \\ 0 & 1 & -i \\ 0 & 0 & 1 \end{pmatrix}.$$

Bemærkning 5.2.12 Det er let at udregne potenser af diagonaliserbare matricer. Hvis nemlig $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, da er $D^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$, og hvis $D = P^{-1}AP$, da er

$$D^k = (P^{-1}AP)^k = P^{-1}APP^{-1}AP \dots P^{-1}AP = P^{-1}A^kP,$$

idet vi fjerner de $k - 1$ forekomster af $PP^{-1} = I$. Heraf følger, at

$$A^k = PD^kP^{-1},$$

så op til et koordinatskift er det ligeså let at udregne potenser af en diagonaliserbar matrix som af en diagonalmatrix. Vi bemærker også, at A^k igen er diagonaliserbar.

Eksempel 5.2.13 Vi anvender bemærkning 5.2.12 på den diagonaliserbare matrix A i eksempel 5.2.4 og får, at

$$\begin{aligned} A^k &= P \text{diag}((-2)^k, (-1)^k) P^{-1} = \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (-2)^k & 0 \\ 0 & (-1)^k \end{pmatrix} \begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} -3(-2)^k & -2(-1)^k \\ (-2)^k & (-1)^k \end{pmatrix} \begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3(-2)^k - 2(-1)^k & 6(-2)^k - 6(-1)^k \\ -(-2)^k + (-1)^k & -2(-2)^k + 3(-1)^k \end{pmatrix}. \end{aligned}$$

Her har vi anvendt eksempel 3.4.4 til at udregne P^{-1} .

Diagonalisering kan bruges til at forstå lineære processer, der gentager sig. Dette vil vi nu illustrere med følgende simple eksempel.

Eksempel 5.2.14 I et land L bor der x personer på landet og y personer i byerne. Man har observeret, at hvert år flytter 15% af befolkningen fra landet til byerne, mens 10% af befolkningen i byerne flytter på landet. Vi ønsker at bestemme, hvor mange personer der vil være på landet og i byerne efter 20 år, hvis denne udvikling forsætter.

Vi observerer først, at i denne model er ændringen i befolkningssammensætning fra år til år givet ved en lineær afbildning: efter et år er der $x - 15\%x + 10\%y$ personer på landet og $y - 10\%y + 15\%x$ personer i byerne. Så de nye antal personer efter et år er

$$\begin{pmatrix} 85/100 x + 10/100 y \\ 15/100 x + 90/100 y \end{pmatrix} = \begin{pmatrix} 17/20 & 1/10 \\ 3/20 & 9/10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Efter 20 år vil befolkningssammensætningen i vores model derfor være

$$A^{20} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{hvor} \quad A = \begin{pmatrix} 17/20 & 1/10 \\ 3/20 & 9/10 \end{pmatrix}.$$

Vi udregner nu, at

$$\chi_A(t) = t^2 - \frac{7}{4}t + \frac{3}{4} = (1-t)\left(\frac{3}{4} - t\right),$$

5 Egenverdier og egenrum

hvilket viser, at A har egenverdier $\lambda_1 = 1$ og $\lambda_2 = 3/4$. Korollar 5.1.12 og korollar 5.2.3 viser da, at A er diagonaliserbar. Vi udregner endvidere, at

$$N_{A-I} = \left\{ \begin{pmatrix} 2 \\ 3 \end{pmatrix} t \mid t \in \mathbb{R} \right\} \subset \mathbb{R}^2 \quad \text{og} \quad N_{A-I\frac{3}{4}} = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} t \mid t \in \mathbb{R} \right\} \subset \mathbb{R}^2.$$

Det følger, at

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3/4 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}^{-1},$$

og derfor viser bemærkning 5.2.12, at

$$\begin{aligned} A^k &= \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3/4 \end{pmatrix}^k \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1^k & 0 \\ 0 & (3/4)^k \end{pmatrix} \begin{pmatrix} 1/5 & 1/5 \\ -3/5 & 2/5 \end{pmatrix} \\ &= \begin{pmatrix} 2/5 + 3/5(3/4)^k & 2/5 - 2/5(3/4)^k \\ 3/5 - 3/5(3/4)^k & 3/5 + 2/5(3/4)^k \end{pmatrix}. \end{aligned}$$

For $k = 20$, kan vi helt ignorere de led, der indeholder $(3/4)^k$, idet $(3/4)^{20} \sim 0,003$, hvilket er ubetydeligt sammenlignet med $2/5 = 0,4$ og $3/5 = 0,6$. Denne simple model viser altså, at efter 20 år er befolkningssammensætningen i landet L stabiliseret med omkring $40\% x + 40\% y$ på landet og $60\% x + 60\% y$ i byerne.

5.3 Egenverdier og egenrum for lineære endomorfier

Vi overfører nu de begreber, vi har indført i de sidste to afsnit, til lineære endomorfier af endeligt frembragte abstrakte vektorrum, og vi begynder med det karakteristiske polynomium.

Definition 5.3.1 Lad $\mathbb{F}$ være et legeme, og lad V være et $\mathbb{F}$ -vektorrum V af endelig dimension n . Det *karakteristiske polynomium* af en lineær endomorfi $f: V \rightarrow V$ er da det karakteristiske polynomium

$$\chi_f(t) = \chi_A(t) = \det(A - It) \in \mathbb{F}[t]$$

af den kvadratiske matrix $A \in M_n(\mathbb{F})$, der repræsenterer $f: V \rightarrow V$ med hensyn til en fælles basis $(v_1, \dots, v_n)$ for domænet og codomænet.

Vi viser nu, at denne definition er meningsfuld i den forstand, at $\chi_f(t)$ ikke afhænger

5.3 Egenverdier og egenrum for lineære endomorfier

af valget af basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V . Så lad også $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ være en basis for V , og lad $B \in M_n(\mathbb{F})$ være matricen, der repræsenterer $f: V \rightarrow V$ med hensyn til denne basis for både domænet og codomænet. Som vi forklarede i bemærkning 5.1.3 gælder det da, at

$$B = P^{-1}AP,$$

hvor $P \in M_n(\mathbb{F})$ er matricen, der repræsenterer identitetsafbildningen $\text{id}_V: V \rightarrow V$ med hensyn til basen $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ for domænet og basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for codomænet. Derfor viser sætning 5.1.4, at $\chi_A(t) = \chi_B(t)$, og dermed er $\chi_f(t)$ uafhængig af valget af basis. Det er dog afgørende, at vi anvender den *samme* basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for både domænet og codomænet af $f: V \rightarrow V$.

Lad nu V være et vilkårligt $\mathbb{F}$ -vektorrum. Hvis $\lambda \in \mathbb{F}$, da definerer formlen

$$(\text{id}_V \lambda)(\mathbf{v}) = \mathbf{v}\lambda$$

en lineær endomorfi $\text{id}_V \lambda: V \rightarrow V$. For hvis $\mathbf{v}, \mathbf{w} \in V$ og $\mu \in \mathbb{F}$, da er

$$\begin{aligned} (\text{id}_V \lambda)(\mathbf{v} + \mathbf{w}) &= (\mathbf{v} + \mathbf{w})\lambda = \mathbf{v}\lambda + \mathbf{w}\lambda = (\text{id}_V \lambda)(\mathbf{v}) + (\text{id}_V \lambda)(\mathbf{w}) \\ (\text{id}_V \lambda)(\mathbf{v}\mu) &= (\mathbf{v}\mu)\lambda = \mathbf{v}(\mu\lambda) = \mathbf{v}(\lambda\mu) = (\mathbf{v}\lambda)\mu = (\text{id}_V \lambda)(\mathbf{v})\mu, \end{aligned}$$

hvor vi har brugt den kommutative lov for multiplikation af skalarer. Særligt gælder det for enhver lineær endomorfi $f: V \rightarrow V$ og $\lambda \in \mathbb{F}$, at

$$V_\lambda = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{v}\lambda\} = \ker(f - \text{id}_V \lambda) \subset V$$

er et underrum af vektorrummet V .

Definition 5.3.2 Lad V være et vektorrum over et legeme $\mathbb{F}$, lad $f: V \rightarrow V$ være en lineær endomorfi, og lad $V_\lambda = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{v}\lambda\}$.

- (1) En skalar $\lambda \in \mathbb{F}$ er en *egenverdi* for $f: V \rightarrow V$, hvis $V_\lambda \neq \{\mathbf{0}\}$.
- (2) Hvis $\lambda \in \mathbb{F}$ er en egenverdi for $f: V \rightarrow V$, så kaldes underrummet $V_\lambda \subset V$ for *egenrummet* for $f: V \rightarrow V$ hørende til egenværdien $\lambda \in \mathbb{F}$.
- (3) Hvis $\lambda \in \mathbb{F}$ er en egenverdi for $f: V \rightarrow V$, så kaldes enhver vektor $\mathbf{v} \in V_\lambda$, der ikke er nulvektoren, for en *egenvektor* for $f: V \rightarrow V$ hørende til egenværdien $\lambda \in \mathbb{F}$.

En vektor $\mathbf{v} \in V$, der ikke er nulvektoren, er altså en egenvektor for $f: V \rightarrow V$ hørende til egenværdien λ , hvis og kun hvis $f(\mathbf{v}) = \mathbf{v}\lambda$.

5 Egenverdier og egenrum

Eksempel 5.3.3 (1) Nulafbildningen $\mathbf{0}: V \rightarrow V$ har $\lambda = 0$ som eneste egenverdi, og

$$V_0 = \ker(\mathbf{0} - \text{id}_V \cdot 0) = \ker(\mathbf{0}) = V.$$

(2) Identitetsafbildningen $\text{id}_V: V \rightarrow V$ har $\lambda = 1$ som eneste egenverdi, og

$$V_1 = \ker(\text{id}_V - \text{id}_V \cdot 1) = \ker(\mathbf{0}) = V.$$

(3) Mængden $V = C^\infty(\mathbb{R})$ af funktioner $f: \mathbb{R} \rightarrow \mathbb{R}$, der er vilkårligt ofte differentiable, har en reel vektorrumstruktur, hvor vektorsum og skalarmultiplikation er givet ved henholdsvis $(f + g)(x) = f(x) + g(x)$ og $(f \cdot \lambda)(x) = f(x) \cdot \lambda$, og afbildningen $D: V \rightarrow V$, der til $f \in V$ tilordner den afledte $f' \in V$, er en lineær endomorfi. Ethvert $\lambda \in \mathbb{R}$ er en egenverdi for $D: V \rightarrow V$, og det tilhørende egenrum er det en-dimensionale underrum

$$V_\lambda = \{e^{\lambda x} \cdot c \mid c \in \mathbb{R}\} \subset V,$$

da funktionerne $f(x) = e^{\lambda x} \cdot c$ udgør alle løsninger til differentiaalligningen $Df = f\lambda$. Så endomorfien $D: V \rightarrow V$ har specielt uendeligt mange egenverdier. Sætning 5.1.11 viser, at dette kun er muligt, fordi vektorrummet V er uendeligdimensionalt.

Vi viser i sætning 5.3.6, at for et endeligt frembragt $\mathbb{F}$ -vektorrum V , er egenverdierne af en lineær endomorfi $f: V \rightarrow V$ netop rødderne i $\mathbb{F}$ af dets karakteristiske polynomium $\chi_f(t)$. Vi begynder med følgende resultat.

Lemma 5.3.4 *Lad $\mathbb{F}$ være et legeme, lad V være et endeligt frembragt $\mathbb{F}$ -vektorrum, og lad $g: V \rightarrow V$ være en lineær endomorfi. Lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ være en basis for V , lad $h: \mathbb{F}^n \rightarrow V$ være den entydigt bestemte lineære isomorfi, sådan at $h(\mathbf{e}_i) = \mathbf{v}_i$ for $1 \leq i \leq n$, og lad $B \in M_n(\mathbb{F})$ være matricen, der repræsenterer $g: V \rightarrow V$ med hensyn til denne basis for både domænet og codomænet. Da gælder det, at $h(N_B) = \ker(g)$.*

Bevis Vi husker fra bemærkning 4.4.5, at den lineære endomorfi $k: \mathbb{F}^n \rightarrow \mathbb{F}^n$, der er givet ved $k(\mathbf{x}) = B\mathbf{x}$, gør det følgende diagram kommutativt.

$$\begin{array}{ccc} V & \xrightarrow{g} & V \\ \uparrow h & & \uparrow h \\ \mathbb{F}^n & \xrightarrow{k} & \mathbb{F}^n \end{array}$$

Vi påstår, at dette medfører, at $h(N_B) = \ker(g)$. For hvis $\mathbf{x} \in N_B = \ker(k)$, så er

$$g(h(\mathbf{x})) = h(k(\mathbf{x})) = h(B\mathbf{x}) = h(\mathbf{0}) = \mathbf{0},$$

5.3 Egenverdier og egenrum for lineære endomorfier

så $h(\mathbf{x}) \in \ker(f)$. Omvendt, hvis $\mathbf{v} \in \ker(g)$, da findes der $\mathbf{x} \in \mathbb{F}^n$, sådan at $\mathbf{x} = h(\mathbf{x})$, idet $h: \mathbb{F}^n \rightarrow V$ er surjektiv. Og da $h: \mathbb{F}^n \rightarrow V$ også er injektiv, så viser udregningen

$$h(h(\mathbf{x})) = g(h(\mathbf{x})) = g(\mathbf{v}) = \mathbf{0},$$

at $\mathbf{x} \in \ker(h) = N_B$. Dermed er $\mathbf{v} = h(\mathbf{x}) \in h(N_B)$, hvilket viser påstanden. $\square$

Korollar 5.3.5 *Lad $\mathbb{F}$ være et legeme, lad V være et endeligt frembragt $\mathbb{F}$ -vektorrum, og lad $f: V \rightarrow V$ være en lineær endomorfi. Lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ være en basis for V , lad $h: \mathbb{F}^n \rightarrow V$ være den entydigt bestemte lineære isomorfi, sådan at $h(\mathbf{e}_i) = \mathbf{v}_i$ for $1 \leq i \leq n$, og lad $A \in M_n(\mathbb{F})$ være matricen, der repræsenterer $f: V \rightarrow V$ med hensyn til denne basis for både domænet og codomænet. Da gælder det for alle $\lambda \in \mathbb{F}$, at*

$$h(N_{A-I\lambda}) = V_\lambda.$$

Særligt er λ en egenverdi for $f: V \rightarrow V$, hvis og kun hvis λ er en egenverdi for A .

Bevis Identiteten $h(N_{A-I\lambda}) = V_\lambda$ følger fra lemma 5.3.4 med $g = f - \text{id}_V \lambda: V \rightarrow V$. For matricen B , der repræsenterer $g: V \rightarrow V$ med hensyn til den fælles basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for domænet og codomænet, er $B = A - I\lambda$, så $N_B = N_{A-I\lambda}$, og $\ker(g) = \ker(f - \text{id}_V \lambda) = V_\lambda$.

Endelig er λ per definition en egenverdi for $f: V \rightarrow V$, hvis og kun hvis $V_\lambda \neq \{\mathbf{0}\}$, og da $h: \mathbb{F}^n \rightarrow V$ er en isomorfi, så viser identiteten ovenfor, at dette er tilfældet, hvis og kun hvis $N_{A-I\lambda} \neq \{\mathbf{0}\}$, hvilket er definitionen på, at λ er en egenverdi for A . $\square$

Sætning 5.3.6 *Lad $\mathbb{F}$ være et legeme, og lad V være et endeligt frembragt $\mathbb{F}$ -vektorrum. Da er egenverdierne af en lineær endomorfi $f: V \rightarrow V$ præcis rødderne i $\mathbb{F}$ af det karakteristiske polynomium $\chi_f(t) \in \mathbb{F}[t]$.*

Bevis Lad $A \in M_n(\mathbb{F})$ være matricen, der repræsenterer $f: V \rightarrow V$ med hensyn til en fælles basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for domænet og codomænet. Ifølge korollar 5.3.5 er $\lambda \in \mathbb{F}$ en egenverdi for $f: V \rightarrow V$, hvis og kun hvis $\lambda \in \mathbb{F}$ er en egenverdi for A , og ifølge sætning 5.1.8 er dette tilfældet, hvis og kun hvis $\lambda \in \mathbb{F}$ er en rod i $\chi_A(t) = \chi_f(t)$. $\square$

Hvis $\lambda \in \mathbb{F}$ er en egenverdi for den lineære endomorfi $f: V \rightarrow V$, så definerer vi igen den *geometriske multiplicitet* til at være dimensionen af egenrummet V_λ hørende til λ ; og vi definerer den *algebraiske multiplicitet* til at være det maksimale hele tal $k \geq 1$, for hvilket der findes $q(t) \in \mathbb{F}[t]$, sådan at $\chi_f(t) = (\lambda - t)^k \cdot q(t) \in \mathbb{F}[t]$.

Sætning 5.3.7 *Lad $\mathbb{F}$ være et legeme og V et endeligt frembragt $\mathbb{F}$ -vektorrum. Givet en egenverdi $\lambda \in \mathbb{F}$ for en lineær endomorfi $f: V \rightarrow V$, da er den algebraiske multiplicitet af λ større end eller lig med den geometriske multiplicitet af λ .*

Bevis Vi vælger først en basis $(v_1, \dots, v_n)$ for egenrummet V_λ og supplerer dernæst denne til en basis $(v_1, \dots, v_n, w_1, \dots, w_p)$ for V . Da vektorerne $v_1, \dots, v_n$ er egenvektorer for $f: V \rightarrow V$ hørende til λ , er matricen $A \in M_{n+p}(\mathbb{F})$, der repræsenterer $f: V \rightarrow V$ med hensyn til basen $(v_1, \dots, v_n, w_1, \dots, w_p)$ for både domænet og codomænet, på formen

$$A = \left(\begin{array}{c|c} I_n \lambda & B \\ \hline O_{p,n} & C \end{array} \right),$$

hvor $I_n \in M_{n,n}(\mathbb{F})$ er identitetsmatricen og $O_{p,n} \in M_{p,n}(\mathbb{F})$ er nulmatricen, og hvor de to matricer $B \in M_{n,p}(\mathbb{F})$ og $C \in M_{p,p}(\mathbb{F})$ kan være vilkårlige. Ved at foretage Laplace-udvikling af determinanten efter første søjle n gange får vi da, at

$$\chi_f(t) = \det(A - I_{n+p}t) = \left(\begin{array}{c|c} I_n \lambda - I_n t & B \\ \hline O_{p,n} & C - I_p t \end{array} \right) = (\lambda - t)^n \det(C - I_p t).$$

Dette viser, at den algebraiske multiplicitet af λ er større end eller lig med den geometriske multiplicitet n som ønsket. $\square$

Bemærkning 5.3.8 Lad os definere, at en lineær endomorfi $f: V \rightarrow V$ af et endeligt frembragt $\mathbb{F}$ -vektorrum er *diagonaliserbar*, hvis der findes en basis for V , der består af egenvektorer for $f: V \rightarrow V$. Bemærkning 5.1.3 viser, at dette er tilfældet, hvis og kun hvis matricen A , der repræsenterer $f: V \rightarrow V$ med hensyn til en vilkårlig, men fælles, basis for domænet og codomænet, er diagonaliserbar. Vi har i sætning 5.1.11 vist, at egenvektorer hørende til forskellige egenverdier er lineært uafhængige. Derfor er $f: V \rightarrow V$ diagonaliserbar, hvis og kun hvis det gælder, at

$$\sum_{\lambda} \dim(V_{\lambda}) = \dim(V),$$

hvor summen til venstre løber over de endeligt mange egenverdier af $f: V \rightarrow V$. Vi har netop vist i sætning 5.3.7, at det gælder følgende ulighed:

$$\sum_{\lambda} \dim V_{\lambda} = \sum_{\lambda} \text{geometrisk multiplicitet}(\lambda) \leq \sum_{\lambda} \text{algebraisk multiplicitet}(\lambda).$$

Og da et polynomium af grad højst d med koefficienter i et legeme $\mathbb{F}$ højst kan have d rødder i $\mathbb{F}$, får vi endvidere følgende ulighed:

$$\sum_{\lambda} \text{algebraisk multiplicitet}(\lambda) \leq \deg(\chi_f(t)) = \dim(V).$$

Vi konkluderer altså, at $f: V \rightarrow V$ er diagonaliserbar, hvis og kun hvis de følgende betingelser begge er opfyldt:

(1) Den følgende identitet gælder:

$$\sum_{\lambda} \text{algebraisk multiplicitet}(\lambda) = \dim(V).$$

(2) For enhver egen værdi λ af $f: V \rightarrow V$ er

$$\text{geometrisk multiplicitet}(\lambda) = \text{algebraisk multiplicitet}(\lambda).$$

Endelig viser algebraens fundamentalsætning (sætning 3.5.10), at betingelsen (1) altid er opfyldt, hvis $\mathbb{F} = \mathbb{C}$.

5.4 Opgaver

5.1 Vi betragter polynomiet

$$p(t) = \det \begin{pmatrix} 3-t & -5 & 4 \\ 1 & -t & 1 \\ 1 & 1 & -t \end{pmatrix}.$$

a) Bestem $p(t)$.

b) Udtryk $p(t)$ som et produkt af førstegradspolynomier.

5.2 For hver af de nedenstående reelle matricer $A \in M_2(\mathbb{R})$, definerer vi $B \in M_2(\mathbb{C})$ til at være matricen A betragtet som en kompleks matrix. Find i hvert af de tre tilfælde eventuelle egen værdier og tilhørende egenrum for henholdsvis A og B . Undersøg endvidere, om matricen A (resp. B) er diagonaliserbar, og find i givet fald en basis for $\mathbb{R}^2$ (resp. $\mathbb{C}^2$), der består af egenvektorer for A (resp. B) samt en invertibel matrix $P \in M_2(\mathbb{R})$ (resp. $Q \in M_2(\mathbb{C})$), sådan at $P^{-1}AP$ (resp. $Q^{-1}BQ$) er en diagonalmatrix.

a)

$$A = \begin{pmatrix} 2 & 4 \\ 5 & 1 \end{pmatrix}.$$

b)

$$A = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}.$$

c)

$$A = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}.$$

5.3 Betragt matricen

$$A = \begin{pmatrix} 3 & 0 & 1 \\ -1 & 4 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

5 Egenverdier og egenrum

- a) Find egenverdierne for matricen A .
- b) Find baser for de tilhørende egenrum.
- c) Find en invertibel reel matrix P , sådan at $P^{-1}AP$ er en diagonalmatrix.

5.4 Betragt matricen

$$A = \begin{pmatrix} 1 & 1 \\ -4 & 5 \end{pmatrix}$$

- a) Find egenverdierne for matricen A .
- b) Find baser for de tilhørende egenrum.
- c) Redegør for, at der *ikke* findes en invertibel reel matrix P , sådan at $P^{-1}AP$ er en diagonalmatrix.

5.5 Vi betragter den lineære endomorfi $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$, der er repræsenteret af matricen

$$A = \begin{pmatrix} 1 & 0 & 0 & -3 \\ 2 & 3 & 0 & 3 \\ -2 & -1 & 2 & -3 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

med hensyn til standardbasen for både domænet og codomænet.

- a) Bestem egenverdier og egenvektorer for A .
- b) Angiv en basis $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ for $\mathbb{R}^4$, sådan at matricen D , der repræsenterer $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ med hensyn til denne basis for både domænet og codomænet, er en diagonalmatrix.
- c) Angiv en invertibel matrix P , sådan at $D = P^{-1}AP$.

5.6 Betragt matricen

$$A = \begin{pmatrix} 2 & 1+2i \\ 0 & i \end{pmatrix}.$$

- a) Bestem egenverdier og egenvektorer for A .
- b) Angiv en invertibel matrix P , sådan at $P^{-1}AP$ er en diagonalmatrix.

5.7 Betragt matricen

$$\begin{pmatrix} 3+3i & 0 \\ -1+2i & 6-3i \end{pmatrix}.$$

- a) Bestem egenverdier og egenvektorer for A .
- b) Angiv en invertibel matrix P , sådan at $P^{-1}AP$ er en diagonalmatrix.

5.8 Bevis at egenverdierne for en (øvre eller nedre) triangulær matrix A netop er diagonalindgangene.

5.9 Lad $\mathbb{F}$ være et legeme.

(a) Vis, at hvis $\lambda \in \mathbb{F}$, da er

$$B \cdot I\lambda = I\lambda \cdot B$$

for alle $B \in M_n(\mathbb{F})$. [Vink: Udregn begge matrixprodukter.]

(b) Lad $A \in M_n(\mathbb{F})$ være en diagonaliserbar matrix med en enkelt egen værdi λ . Bevis, at da er $A = I\lambda$.

5.10 Vi har i opgave 2.9 defineret sporet $\text{tr}(A)$ af en kvadratisk matrix A som summen af diagonalindgangene og vist, at $\text{tr}(AB) = \text{tr}(BA)$.

a) Lad $f: V \rightarrow V$ være en lineær endomorfi, og lad A og B være matrixerne, der repræsenterer $f: V \rightarrow V$ med hensyn til baser $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ for V . Vis, at $\text{tr}(A) = \text{tr}(B)$.

b) Lad $A \in M_n(\mathbb{C})$, og lad $\lambda_1, \dots, \lambda_k$ være egen værdierne af A . Vis, at hvis A er diagonaliserbar, da er

$$\text{tr}(A) = \sum_{i=1}^k d_i \lambda_i$$

hvor d_i er den algebraiske multiplicitet af λ_i .

5.11 Vis, at det for enhver kompleks matrix $A \in M_n(\mathbb{C})$ gælder, at

$$\det(A) = \prod_{i=1}^k \lambda_i^{d_i}$$

hvor $\lambda_1, \dots, \lambda_k$ er egen værdierne af A , og hvor d_i er den algebraiske multiplicitet af egen værdien λ_i .

[Vink: Anvend sætning 5.2.10.]

5.12 I vektorrummet $\mathbb{R}^3$ er der givet den lineært uafhængige familie af vektorer

$$\left(\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right).$$

En lineær endomorfi $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ er fastlagt ved

$$f(\mathbf{v}_1) = \mathbf{v}_2, \quad f(\mathbf{v}_2) = \mathbf{v}_1 \quad \text{og} \quad f(\mathbf{v}_3) = \mathbf{0}.$$

a) Opskriv den matrix, der repræsenterer $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ i basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.

b) Gør rede for, at $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ er diagonaliserbar.

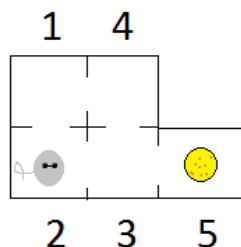
5 Egenverdier og egenrum

5.13 Vi betragter de lineære endomorfier $f, g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, der er repræsenteret af

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

med hensyn til standardbasen for både domænet og codomænet.

- a) Find en basis $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for $\mathbb{R}^3$, som består af vektorer, der er egenvektorer for både A og B .
- b) Find en invertibel matrix P , sådan at $P^{-1}AP$ og $P^{-1}BP$ begge er diagonalmatricer.
- 5.14 (★) En blind mus befinder sig i et rum på tegningen herunder og leder efter en ost. Den går hvert minut til et nyt rum i et tilfældigt mønster, indtil den finder osten. Musen bevæger sig med lige stor sandsynlighed ind i hvert af de tilstødende lokaler. (F.eks. er sandsynligheden for at musen går fra rum 2 til 1 og fra rum 2 til 3, begge lig med $1/2$.) Og sådan fortsætter musen med at vandre indtil den finder osten, hvorefter den forbliver i rum 5 og tilbringer resten af sine dage med at spise den store ost. Til at beskrive musens bevægelse kan vi opskrive en overgangsmatrix P , hvis elementer P_{ij} er sandsynlighederne for at musen går **til** rum i **fra** rum j . (Så for eksempel er $P_{12} = \frac{1}{2} = P_{32}$, mens $P_{22} = P_{42} = P_{52} = 0$.)



- a) Opskriv matricen P ud fra musens mulige bevægelser. Gør herunder rede for, hvorfor $P_{55} = 1$, mens $P_{i5} = 0$ for $i = 1, 2, 3, 4$.
- b) Vi kan repræsentere startrummet for musen ved en standard enhedsvektor $\mathbf{e}_j$, $j = 1, \dots, 5$, da musen starter i netop ét af rummene 1–5. Hvis musen befinder sig i rum j , vil sandsynligheden for at træffe musen i næste minut i de forskellige rum altså være givet ved vektoren $P\mathbf{e}_j$, som er den j 'te søjle i matricen P . Her er den i 'te indgang i vektoren $P\mathbf{e}_j$ lig med sandsynligheden for, at musen er i det i 'te rum efter et minut. Tjek, at matricen P fundet i a) har den egenskab.

Mere generelt vil sandsynligheden for at træffe musen i de forskellige rum efter k minutter, hvis den starter i det j 'te rum, være givet ved vektoren $P^k \mathbf{e}_j$. For hver $j = 1, 2, 3, 4$, bestem (gerne ved brug af Maple) sandsynligheden for, at musen finder osten på højst 10 minutter, givet at den starter i det j 'te rum. Hvilken af de 4 sandsynligheder er størst, og hvilken er mindst? Svarer det til dine forventninger?

- c) Musen starter i rum 1. Efter at have vandret i noget tid og stadig ikke have fundet osten, kommer musen til at tænke, at den måske aldrig vil finde osten. Vis, at sandsynligheden for at musen finder osten, dvs. befinder sig i rum 5 efter k minutter, går mod 1 når k går mod uendelig.

[Hint: Diagonalisering].

(Til denne del af opgaven må der gerne bruges Maple til at finde egenverdier, egenvektorer og inverse matricer.)

- 5.15 (★) Vi definerer en reel matrix $A \in M_n(\mathbb{R})$ til at være *skævsymmetrisk*, hvis den opfylder, at $A^t = -A$. Vis, at det for en sådan matrix gælder, at

$$\chi_A(t) = (-1)^n \chi_A(-t).$$

- 5.16 Betragt matricen

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

- Udregn A^2, A^3 .
- Udregn $\det(A)$.
- Udregn egenverdierne og egenrummerne for matricen A . Er A diagonaliserbar?

Lad $\mathbb{F}$ være et legeme. En matrix $N \in M_n(\mathbb{F})$ kaldes *nilpotent*, hvis der findes et naturligt tal k sådan at, hvis man ganger matricen N med sig selv k gange, så får man nulmatricen, dvs. hvis

$$N^k = O_{n,n}.$$

- Vis, at hvis N er en nilpotent matrix, så er $\det(N) = 0$.
- Vis, at hvis N er en diagonaliserbar nilpotent matrix i $M_n(\mathbb{F})$, så er matricen $N = O_{n,n}$.
- Hvad er det karakteristiske polynomium for N ?

6 Vektorrum med indre produkt

Dette kapitel omhandler geometriske strukturer på reelle og komplekse vektorrum. Disse strukturer er ikke lineære, men derimod kvadratiske, af natur, og de afhænger af et valg af en yderligere struktur, der kaldes for et *indre produkt*. Et indre produkt giver således anledning til en *norm* eller længdemål af vektorer, samt til begrebet, at to vektorer er *ortogonale* eller vinkelrette. I det reelle tilfælde er det også muligt at definere *vinklen* mellem to vektorer. Vi viser den meget anvendelige Cauchy-Schwarz-ulighed, hvorfra de fleste egenskaber for indre produkt og norm følger. Vi definerer og studerer også *lineære isometrier*, som er de lineære afbildninger mellem vektorrum med indre produkt, der bevarer denne ekstra geometriske struktur.

Vi siger, at en basis for et vektorrum med indre produkt er *ortonormal* med hensyn til det indre produkt, hvis vektorerne i basen er parvis ortogonale og har norm 1. En stor fordel ved en ortonormal basis er, at koordinaterne af en vektor med hensyn til denne basis er givet ved indre produkt med basisvektorerne. Vi angiver endvidere en algoritme, der kaldes for *Gram-Schmidt-ortogonalisering*, der erstatter en vilkårlig basis med en basis, der er ortonormal med hensyn til det indre produkt.

Endelig viser vi de såkaldte *spektralsætninger* for henholdsvis *selvadjungerede* lineære endomorfier af reelle vektorrum med indre produkt og *normale* lineære endomorfier af komplekse vektorrum. I begge tilfælde udtaler sætningen, at der findes en basis af egenvektorer for den lineære endomorfi, der endvidere er ortonormal med hensyn til det indre produkt. Vi anvender til sidst spektralsætningerne til at bevise Sylvesters klassifikation af hermitiske former på reelle og komplekse vektorrum.

I dette kapitel antager vi, at $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, men i de tre første afsnit gælder alle udsagn og definitioner også for $\mathbb{F} = \mathbb{H}$ med kvaternionisk konjugation.

6.1 Indre produkt

Vi betragter vektorrum over $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og for $a \in \mathbb{F}$ definerer vi

$$a^* = \begin{cases} a, & \text{hvis } \mathbb{F} = \mathbb{R}, \\ \bar{a}, & \text{hvis } \mathbb{F} = \mathbb{C}. \end{cases}$$

For reelle tal giver det mening at spørge, om et tal er positivt eller negativt, hvorimod dette spørgsmål ikke giver mening for komplekse tal generelt. Vi bemærker imidlertid, at hvis $a \in \mathbb{F}$ opfylder, at $a = a^*$, så er $a \in \mathbb{R}$, uanset om $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. I dette tilfælde er udsagnet " $a > 0$ " derfor meningsfuldt, hvilket vi anvender i følgende definition.

6 Vektorrum med indre produkt

Definition 6.1.1 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. Et *indre produkt* på et $\mathbb{F}$ -vektorrum V er en afbildning $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$, der opfylder følgende: For alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ og $a \in \mathbb{F}$ er

$$(H1) \quad \langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle.$$

$$(H2) \quad \langle \mathbf{u}, \mathbf{v}a \rangle = \langle \mathbf{u}, \mathbf{v} \rangle a.$$

$$(H3) \quad \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle.$$

$$(H4) \quad \langle \mathbf{u}a, \mathbf{v} \rangle = a^* \langle \mathbf{u}, \mathbf{v} \rangle.$$

$$(H5) \quad \langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle^*.$$

$$(P) \quad \text{For alle } \mathbf{0} \neq \mathbf{u} \in V \text{ er } \langle \mathbf{u}, \mathbf{u} \rangle > 0.$$

Et $\mathbb{F}$ -vektorrum med indre produkt er et par $(V, \langle -, - \rangle)$ bestående af et $\mathbb{F}$ -vektorrum V og et indre produkt $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$.

Vi minder om, at vi i afsnit 2.6 definerede en hermitisk form på V til at være en afbildning $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$, der opfylder (H1)–(H5), idet vi dog kun betragtede $V = \mathbb{F}^n$. Mens hermitiske former er defineret generelt for et skævlegeme $\mathbb{F}$ med skævinvolution $(-)^*$, så giver *positivitetsbetingelsen* (P) i definitionen af et indre produkt kun mening, hvis $\mathbb{F}$ er $\mathbb{R}$, $\mathbb{C}$ eller $\mathbb{H}$ og $(-)^*$ de respektive konjugeringsafbildninger.

Eksempel 6.1.2 (1) Standard-indreproduktet $\langle -, - \rangle: \mathbb{F}^n \times \mathbb{F}^n \rightarrow \mathbb{F}$ defineret ved

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y} = x_1^* y_1 + \cdots + x_n^* y_n$$

er et indre produkt på $\mathbb{F}^n$. Vi efterviser, at positivitetsbetingelsen (P) er opfyldt. Hvis $\mathbb{F} = \mathbb{R}$, da er $\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + \cdots + x_n^2 \geq 0$ med lighed, hvis og kun hvis $\mathbf{x} = \mathbf{0}$. Og hvis $\mathbb{F} = \mathbb{C}$, da er $\langle \mathbf{x}, \mathbf{x} \rangle = |x_1|^2 + \cdots + |x_n|^2 \geq 0$, og der gælder igen lighed, hvis og kun hvis $\mathbf{x} = \mathbf{0}$.

(2) Vi lader $A = A^* \in M_n(\mathbb{F})$ være en hermitisk matrix og husker fra sætning 2.6.10, at afbildningen $\langle -, - \rangle: \mathbb{F}^n \times \mathbb{F}^n \rightarrow \mathbb{F}$ defineret ved

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* A \mathbf{y}$$

er en hermitisk form, hvilket vil sige, at den opfylder (H1)–(H5). Hvis denne afbildning også opfylder (P), så siger vi, at den hermitiske matrix $A \in M_n(\mathbb{F})$ er *positiv definit*. En diagonalmatrix $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ er således hermitisk, hvis og kun hvis $\lambda_1, \dots, \lambda_n$ alle er reelle tal, og den er hermitisk og positiv definit, hvis og kun hvis $\lambda_1, \dots, \lambda_n$ alle er positive reelle tal. For eksempel er $A = \text{diag}(-1, 1, \dots, 1)$, der svarer til Minkowskis hermitiske form, ikke positiv definit, så Minkowski-formen er ikke et indre produkt.

(3) Lad $V = C^0([a, b])$ være det uendeligdimensionale reelle vektorrum af kontinuerte funktioner $f: [a, b] \rightarrow \mathbb{R}$ fra bemærkning 4.4.20. Da definerer formen

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

et indre produkt $\langle -, - \rangle: V \times V \rightarrow \mathbb{R}$.

Definition 6.1.3 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad V være et $\mathbb{F}$ -vektorrum. Givet et indre produkt $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$, da kaldes afbildningen $\|-\|: V \rightarrow \mathbb{R}$ defineret ved

$$\|v\| = \sqrt{\langle v, v \rangle}$$

for *normen* hørende til $\langle -, - \rangle$.

Vi kalder også $\|v\|$ for *længden* af v med hensyn til $\langle -, - \rangle$. Vi bemærker, at normen af v er veldefineret, fordi $\langle v, v \rangle$ er et ikke-negativt reelt tal, og at normen af v selv er et ikke-negativt reelt tal. Det gælder endvidere, at

$$\|va\| = \|v\||a|$$

for alle $v \in V$ og $a \in \mathbb{F}$. Her er $|a| = \sqrt{a^*a} \in \mathbb{R}$ absolutværdien af $a \in \mathbb{F}$.

Vi viser nu, at absolutværdien af det indre produkt af to vektorer altid er mindre eller lig med produktet af deres norm. Denne ulighed, der kendes som *Cauchy-Schwarz' ulighed* og skyldes Cauchy, Schwarz og Bunyakovsky, gælder i ethvert vektorrum med indre produkt. Det er en enormt anvendelig ulighed, hvilket vi illustrerer med beviset for trekantsuligheden nedenfor.

Sætning 6.1.4 (Cauchy-Schwarz) Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt. For alle $v, w \in V$ er

$$|\langle v, w \rangle| \leq \|v\|\|w\|.$$

Bevis Hvis $v = 0$, er påstanden triviell, så vi antager, at $v \neq 0$. Beviset i dette tilfælde beror på følgende snedige trick: Vi betragter vektoren

$$z = \text{pr}_v(w) = v \cdot \frac{\langle v, w \rangle}{\langle v, v \rangle},$$

6 Vektorrum med indre produkt

som vi kalder for projektionen af $\mathbf{w}$ på $\mathbf{v}$. (Se også definition 6.2.7.) For udregningen

$$\begin{aligned}\langle \mathbf{z}, \mathbf{w} - \mathbf{z} \rangle &= \left\langle \mathbf{v} \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}, \mathbf{w} - \mathbf{v} \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \right\rangle \\ &= \left(\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \right)^* \cdot \langle \mathbf{v}, \mathbf{w} \rangle - \left(\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \right)^* \cdot \langle \mathbf{v}, \mathbf{v} \rangle \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = 0\end{aligned}$$

viser, at $\mathbf{z}$ og $\mathbf{w} - \mathbf{z}$ er ortogonale; se definition 6.2.1 nedenfor. Vi skriver nu $\mathbf{w}$ som $\mathbf{z} + (\mathbf{w} - \mathbf{z})$ og udregner

$$\langle \mathbf{w}, \mathbf{w} \rangle = \langle \mathbf{z} + (\mathbf{w} - \mathbf{z}), \mathbf{z} + (\mathbf{w} - \mathbf{z}) \rangle = \langle \mathbf{z}, \mathbf{z} \rangle + 0 + 0 + \langle \mathbf{w} - \mathbf{z}, \mathbf{w} - \mathbf{z} \rangle \geq \langle \mathbf{z}, \mathbf{z} \rangle,$$

hvor uligheden til højre gælder fordi $\langle \mathbf{w} - \mathbf{z}, \mathbf{w} - \mathbf{z} \rangle \geq 0$. Vi bemærker, at uligheden giver mening, fordi begge sider af uligheden er reelle tal. Vi udregner endelig, at

$$\langle \mathbf{z}, \mathbf{z} \rangle = \left\langle \mathbf{v} \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}, \mathbf{v} \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \right\rangle = \left(\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \right)^* \cdot \langle \mathbf{v}, \mathbf{v} \rangle \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle^*}{\langle \mathbf{v}, \mathbf{v} \rangle} \cdot \langle \mathbf{v}, \mathbf{w} \rangle = \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\langle \mathbf{v}, \mathbf{v} \rangle},$$

idet $\langle \mathbf{v}, \mathbf{v} \rangle^* = \langle \mathbf{v}, \mathbf{v} \rangle$. Så uligheden ovenfor er derfor ækvivalent med uligheden

$$\langle \mathbf{v}, \mathbf{v} \rangle \cdot \langle \mathbf{w}, \mathbf{w} \rangle \geq |\langle \mathbf{v}, \mathbf{w} \rangle|^2,$$

fordi $\langle \mathbf{v}, \mathbf{v} \rangle > 0$. Sætningen følger nu ved at uddrage kvadratroden på begge sider. $\square$

Lad V være et $\mathbb{F}$ -vektorrum med indre produkt $\langle -, - \rangle$ og tilhørende norm $\| - \|$. Givet to vektorer $\mathbf{v}$ og $\mathbf{w}$ i V , kan vi danne trekanten med hjørner $\mathbf{0}$, $\mathbf{v}$ og $\mathbf{v} + \mathbf{w}$, og vi kan betragte $\|\mathbf{v}\|$, $\|\mathbf{w}\|$ og $\|\mathbf{v} + \mathbf{w}\|$ som længderne af de tre sider i denne trekant. Trekantsuligheden, som vi nu viser, siger altså, at længden af én side i en trekant altid er mindre end eller lig med summen af længderne af de to resterende sider.

Sætning 6.1.5 (Trekantsuligheden) *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt. Da gælder det for alle $\mathbf{v}, \mathbf{w} \in V$, at*

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$

Bevis Vi bemærker først, at det for alle $a = x + iy \in \mathbb{C}$ gælder, at

$$a + a^* = (x + iy) + (x - iy) = 2x \leq 2\sqrt{x^2 + y^2} = 2|a|,$$

og at samme ulighed gælder trivielt også for alle $a \in \mathbb{R}$. Vi udregner nu

$$\begin{aligned}\|\mathbf{v} + \mathbf{w}\|^2 &= \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\ &= \|\mathbf{v}\|^2 + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle^* + \|\mathbf{w}\|^2 \leq \|\mathbf{v}\|^2 + 2|\langle \mathbf{v}, \mathbf{w} \rangle| + \|\mathbf{w}\|^2 \\ &\leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2 = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2,\end{aligned}$$

hvor den første af de to uligheder følger af den ulighed, vi viste i begyndelsen af beviset, og den anden er Cauchy-Schwarz' ulighed. Den ønskede ulighed følger ved at uddrage kvadratroden. $\square$

For *reelle* vektorrum med indre produkt er Cauchy-Schwarz' ulighed ækvivalent med udsagnet, at det for alle vektorer $\mathbf{v}$ og $\mathbf{w}$ i vektorrummet gælder, at

$$-\|\mathbf{v}\|\|\mathbf{w}\| \leq \langle \mathbf{v}, \mathbf{w} \rangle \leq \|\mathbf{v}\|\|\mathbf{w}\|.$$

Hvis $\mathbf{v} \neq \mathbf{0}$ og $\mathbf{w} \neq \mathbf{0}$, så kan vi ækvivalent skrive denne ulighed som

$$-1 \leq \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\|\|\mathbf{w}\|} \leq 1,$$

fordi $\|\mathbf{v}\| > 0$ og $\|\mathbf{w}\| > 0$. Vi definerer nu *vinklen* mellem $\mathbf{v}$ og $\mathbf{w}$ til at være det entydigt bestemte reelle tal $0 \leq \theta \leq \pi$, sådan at

$$\cos \theta = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\|\|\mathbf{w}\|}.$$

Se eksempel 6.2.9 for relation til trigonometri og en definition af $\sin \theta$.

Eksempel 6.1.6 (1) Vi betragter $\mathbb{R}^3$ med standard-indreproduktet og udregner vinklen mellem de to vektorer

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Per definition er denne det entydigt bestemte reelle tal $0 \leq \theta \leq \pi$, sådan at

$$\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}} = \frac{2}{\sqrt{2}\sqrt{3}} = \frac{\sqrt{6}}{2},$$

hvilket approksimativt er $\theta \doteq 35,26^\circ$.

(2) Vi lader $(C^0([0, 1]), \langle -, - \rangle)$ være det reelle vektorrum med indre produkt fra eksempel 6.1.2 (3), og lader $f, g: [0, 1] \rightarrow \mathbb{R}$ være de to vektorer heri defineret ved henholdsvis $f(x) = x$ og $g(x) = x^2$. For at bestemme vinklen θ mellem f og g udregner vi:

$$\begin{aligned} \langle f, f \rangle &= \int_0^1 f(x)f(x)dx = \int_0^1 x^2 dx = \frac{1}{3} \\ \langle f, g \rangle &= \int_0^1 f(x)g(x)dx = \int_0^1 x^3 dx = \frac{1}{4} \\ \langle g, g \rangle &= \int_0^1 g(x)g(x)dx = \int_0^1 x^4 dx = \frac{1}{5} \end{aligned}$$

Det gælder altså, at

$$\cos \theta = \frac{\langle f, g \rangle}{\sqrt{\langle f, f \rangle} \cdot \sqrt{\langle g, g \rangle}} = \frac{\sqrt{3} \cdot \sqrt{5}}{4} = \frac{\sqrt{15}}{4}$$

og dermed, at $\theta \doteq 14,48^\circ$.

6.2 Ortogonalitet

Mens vinkelbegrebet kun giver mening for reelle vektorrum med indre produkt, så giver begrebet ortogonalitet mening for vilkårlige $\mathbb{F}$ -vektorrum med indre produkt.

Definition 6.2.1 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt.

- (1) En vektor $\mathbf{u} \in V$ er en *enhedsvektor* med hensyn til $\langle -, - \rangle$, hvis $\|\mathbf{u}\| = 1$.
- (2) To vektorer $\mathbf{v}, \mathbf{w} \in V$ er *ortogonale* med hensyn til $\langle -, - \rangle$, hvis $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.
- (3) En familie $(\mathbf{v}_i)_{i \in I}$ af vektorer i V er *ortogonal* med hensyn til $\langle -, - \rangle$, hvis den består af vektorer, der er parvis ortogonale med hensyn til $\langle -, - \rangle$.
- (4) En familie $(\mathbf{u}_i)_{i \in I}$ af vektorer i V er *ortonormal* med hensyn til $\langle -, - \rangle$, hvis den både er ortogonal med hensyn til $\langle -, - \rangle$ og består af enhedsvektorer med hensyn til $\langle -, - \rangle$.

Vi bemærker, at $\|\mathbf{u}\| = 1$, hvis og kun hvis $\langle \mathbf{u}, \mathbf{u} \rangle = 1$.

Eksempel 6.2.2 (1) Standardbasen $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ for $\mathbb{F}^m$ er ortonormal med hensyn til standard-indreproduktet, fordi $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$.

(2) Den følgende familie $(\mathbf{v}_1, \mathbf{v}_2)$ af vektorer i $\mathbb{F}^2$ er ortogonal, men ikke ortonormal, med hensyn til standard-indreproduktet.

$$(\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix})$$

For selvom $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$, så er $\langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 2 \neq 1$ og $\langle \mathbf{v}_2, \mathbf{v}_2 \rangle = 2 \neq 1$.

(3) Ved en udregning af integraler ses, at i vektorrummet $C^0([0, 2\pi])$ af kontinuerte funktioner $f: [0, 2\pi] \rightarrow \mathbb{R}$ er familien

$$\left(\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos(x), \frac{1}{\sqrt{\pi}} \sin(x), \frac{1}{\sqrt{\pi}} \cos(2x), \frac{1}{\sqrt{\pi}} \sin(2x), \frac{1}{\sqrt{\pi}} \cos(3x), \frac{1}{\sqrt{\pi}} \sin(3x), \dots \right)$$

ortonormal med hensyn til det indre produkt, som vi definerede i eksempel 6.1.2 (3). Underrummet af $C^0([0, 2\pi])$, der er frembragt af denne familie af funktioner, kaldes for vektorrummet af *trigonometriske polynomielle funktioner*. Disse anvendes blandt andet til at modelere signaler.



Figur 6.1: Den trigonometriske polynomielle funktion $\sin(x) + \frac{1}{3}\sin(3x) + \frac{1}{5}\sin(5x)$.

Lemma 6.2.3 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt. Enhver ortonormal familie af vektorer i V er lineært uafhængig.*

Bevis Lad $(v_i)_{i \in I}$ være en ortonormal familie af vektorer i V , og lad

$$\sum_{i \in I} v_i a_i = \mathbf{0}$$

være en linearkombination af $(v_i)_{i \in I}$, der er lig med nulvektoren. (Hvis familien ikke er endelig, så anvendes definitionen givet i bemærkning 4.3.26.) Vi skal da vise, at $a_i = 0$ for alle $i \in I$. Ligningen ovenfor viser, at

$$\langle v_i, \sum_{j \in I} v_j a_j \rangle = \sum_{j \in I} \langle v_i, v_j \rangle a_j = \sum_{j \in I} \delta_{ij} a_j = a_i$$

er lig med $\langle v_i, \mathbf{0} \rangle = 0$ for alle $i \in I$ som ønsket. $\square$

For at bestemme koordinaterne $\mathbf{x} \in \mathbb{F}^n$ af en vektor $\mathbf{w}$ i et $\mathbb{F}$ -vektorrum V med hensyn til en basis $(v_1, \dots, v_n)$, er det generelt nødvendigt at bestemme den entydige løsning til ligningen $v_1 x_1 + \dots + v_n x_n = \mathbf{w}$ blandt vektorer i V . Vi skal nu se, at i modsætning hertil, så kan koordinaterne $\mathbf{a} \in \mathbb{F}^n$ af $\mathbf{w} \in V$ med hensyn til en basis $(u_1, \dots, u_n)$, der er ortonormal med hensyn til et indre produkt $\langle -, - \rangle$ på V , umiddelbart angives ved hjælp af det indre produkt.

Sætning 6.2.4 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt, og lad $(u_i)_{i \in I}$ være en ortonormal basis for V . Da er*

$$\mathbf{w} = \sum_{i \in I} u_i \cdot \langle u_i, \mathbf{w} \rangle$$

for alle $\mathbf{w} \in V$.

6 Vektorrum med indre produkt

Bevis Da enhver vektor $\mathbf{w} \in V$ kan skrives entydigt som en linearkombination

$$\mathbf{w} = \sum_{j \in I} \mathbf{u}_j \cdot a_j$$

af basen $(\mathbf{u}_j)_{j \in I}$, følger sætningen af udregningen

$$\langle \mathbf{u}_i, \mathbf{w} \rangle = \langle \mathbf{u}_i, \sum_{j \in I} \mathbf{u}_j \cdot a_j \rangle = \sum_{j \in I} \langle \mathbf{u}_i, \mathbf{u}_j \cdot a_j \rangle = \sum_{j \in I} \langle \mathbf{u}_i, \mathbf{u}_j \rangle \cdot a_j = a_i.$$

Her følger den anden og tredje identitet af henholdsvis (H1) og (H2), mens den sidste identitet følger af $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$. $\square$

Eksempel 6.2.5 Ved at normere vektorerne i basen

$$(\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix})$$

fra eksempel 6.2.2, får vi den nye basis

$$(\mathbf{u}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}),$$

der er ortonormal med hensyn til standard-indreproduktet. Sætning 6.2.4 viser da for eksempel, at koordinaterne $\mathbf{a} \in \mathbb{F}^2$ af vektoren

$$\mathbf{w} = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \in \mathbb{F}^2$$

med hensyn til basen $(\mathbf{u}_1, \mathbf{u}_2)$ er givet ved

$$\mathbf{a} = \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{w} \rangle \\ \langle \mathbf{u}_2, \mathbf{w} \rangle \end{pmatrix} = \begin{pmatrix} 8/\sqrt{2} \\ 2/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \cdot \sqrt{2}.$$

Vi viser nu omvendt, at hvis koordinaterne af to vektorer med hensyn til en basis, der er ortonormal med hensyn til et indre produkt, er kendte, da er det indre produkt af de to vektorer lig med standard-indreproduktet af deres koordinatvektorer.

Sætning 6.2.6 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et endeligt frembragt $\mathbb{F}$ -vektorrum med indre produkt, og lad $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ være en basis for V , der er ortonormal med hensyn til $\langle -, - \rangle$. Lad $\mathbf{v}, \mathbf{w} \in V$, og lad $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ være koordinaterne for henholdsvis $\mathbf{v}$ og $\mathbf{w}$ med hensyn til basen $(\mathbf{u}_1, \dots, \mathbf{u}_n)$. Da er

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{x}^* \mathbf{y} = x_1^* y_1 + \dots + x_n^* y_n.$$

Bevis Den ønskede formel følger fra udregningen

$$\begin{aligned}\langle \mathbf{v}, \mathbf{w} \rangle &= \left\langle \sum_{i \in I} \mathbf{u}_i \cdot x_i, \sum_{j \in J} \mathbf{u}_j \cdot y_j \right\rangle = \sum_{i \in I} \langle \mathbf{u}_i \cdot x_i, \sum_{j \in J} \mathbf{u}_j \cdot y_j \rangle = \sum_{i \in I} \sum_{j \in I} \langle \mathbf{u}_i \cdot x_i, \mathbf{u}_j \cdot y_j \rangle \\ &= \sum_{i \in I} \sum_{j \in I} x_i^* \cdot \langle \mathbf{u}_i, \mathbf{u}_j \cdot y_j \rangle = \sum_{i \in I} \sum_{j \in I} x_i^* \cdot \langle \mathbf{u}_i, \mathbf{u}_j \rangle \cdot y_j = \sum_{i \in I} x_i^* y_i,\end{aligned}$$

hvor anden, tredje, fjerde og femte identitet fås fra egenskaberne ved et indre produkt, mens den sidste identitet fås fra $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$. $\square$

Definition 6.2.7 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorum med indre produkt. Givet vektorer $\mathbf{v}, \mathbf{w} \in V$, kaldes vektoren

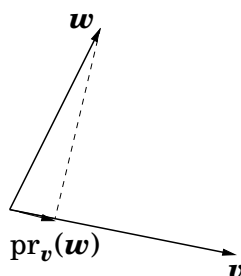
$$\text{pr}_{\mathbf{v}}(\mathbf{w}) = \mathbf{v} \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$$

for den *ortogonale projektion* af $\mathbf{w}$ på $\mathbf{v}$ med hensyn til $\langle -, - \rangle$.

Vi bemærker, at $\mathbf{w} - \text{pr}_{\mathbf{v}}(\mathbf{w})$ er ortogonal på $\mathbf{v}$, idet

$$\langle \mathbf{v}, \mathbf{w} - \text{pr}_{\mathbf{v}}(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{w} - \mathbf{v} \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{v} \rangle \cdot \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = 0,$$

samt at $\mathbf{w} = \text{pr}_{\mathbf{v}}(\mathbf{w}) + (\mathbf{w} - \text{pr}_{\mathbf{v}}(\mathbf{w}))$, hvilket er trivielt, men ofte ganske brugbart.



Figur 6.2: Ortogonal projektion $\text{pr}_{\mathbf{v}}(\mathbf{w})$ af vektoren $\mathbf{w}$ på vektoren $\mathbf{v}$

Eksempel 6.2.8 Figur 6.2 illustrerer det følgende eksempel på ortogonal projektion i $\mathbb{R}^2$ med hensyn til standard-indreproduktet:

$$\mathbf{v} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \quad \text{og} \quad \text{pr}_{\mathbf{v}}(\mathbf{w}) = \begin{pmatrix} 5 \\ -1 \end{pmatrix} \cdot \frac{10 - 4}{25 + 1} = \begin{pmatrix} 5 \\ -1 \end{pmatrix} \cdot \frac{3}{13}.$$

6 Vektorrum med indre produkt

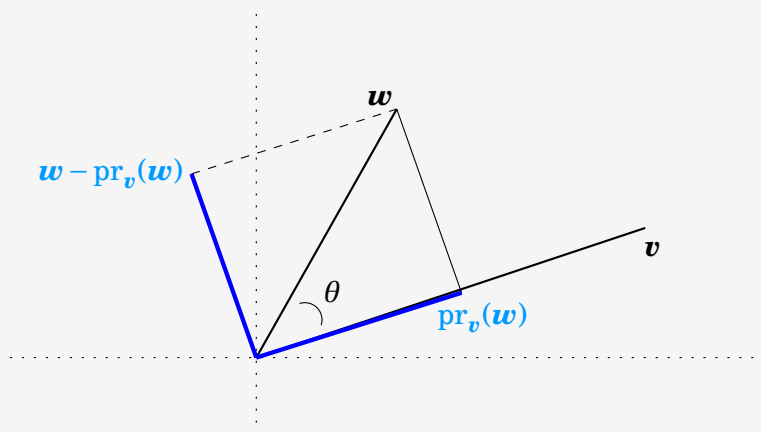
Eksempel 6.2.9 Givet to vektorer $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, der begge er forskellige fra nulvektoren, definerede vi i afsnit 6.1 vinklen θ mellem $\mathbf{v}$ og $\mathbf{w}$ til at være det entydigt bestemte reelle tal $0 \leq \theta \leq \pi$, for hvilket det gælder, at

$$\cos \theta = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \|\mathbf{w}\|}.$$

Vi bemærker nu, at $\cos \theta$ også kan skrives som

$$\cos(\theta) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \|\mathbf{w}\|} = \frac{\|\mathbf{v}\| \langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle \|\mathbf{w}\|} = \frac{\|\text{pr}_{\mathbf{v}}(\mathbf{w})\|}{\|\mathbf{w}\|}.$$

Hvis vektorerne $\mathbf{v}$ og $\mathbf{w}$ ikke er parallelle, da udspænder de en plan i $\mathbb{R}^n$, og figuren nedenfor illustrerer, at $\cos \theta$ som forventet er lig med længden af den hosliggende katete $\text{pr}_{\mathbf{v}} \mathbf{w}$ divideret med længden af hypotenusen $\mathbf{w}$ i den retvinklede trekant, som dannes af vektorerne $\text{pr}_{\mathbf{v}}(\mathbf{w})$, $\mathbf{w} - \text{pr}_{\mathbf{v}}(\mathbf{w})$ og $\mathbf{w}$.



Endvidere er $\sin(\theta)$, op til et fortegn, lig med længden af den modstående katete divideret med længden af hypotenusen i denne trekant,

$$\sin(\theta) = \pm \frac{\|\mathbf{w} - \text{pr}_{\mathbf{v}}(\mathbf{w})\|}{\|\mathbf{w}\|},$$

mens bestemmelsen af, om fortegnet er $+1$ eller -1 , kræver en *orientering* af planen udspændt af $\mathbf{v}$ og $\mathbf{w}$. Ved hjælp af trekantsuligheden konkluderer vi, at

$$0 \leq \frac{\|\mathbf{w} - \text{pr}_{\mathbf{v}}(\mathbf{w})\|}{\|\mathbf{w}\|} \leq 1,$$

og vi bemærker endvidere, at

$$\left(\frac{\|\text{pr}_{\mathbf{v}}(\mathbf{w})\|}{\|\mathbf{w}\|} \right)^2 + \left(\frac{\|\mathbf{w} - \text{pr}_{\mathbf{v}}(\mathbf{w})\|}{\|\mathbf{w}\|} \right)^2 = 1.$$

Vi skal nu vise, at ethvert endeligt frembragt vektorrum med indre produkt har en ortonormal basis. Mere præcist angiver vi en algoritme, der til en given basis tilordner en ny basis, der er ortonormal med hensyn til det indre produkt. Algoritmen kaldes for Gram-Schmidt-ortogonalisering.

Sætning 6.2.10 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt, og lad $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ være en lineært uafhængig familie af vektorer i V . Lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ være familien af vektorer i V , hvori $\mathbf{v}_j$ er defineret rekursivt ved*

$$\mathbf{v}_j = \mathbf{w}_j - \sum_{1 \leq k < j} \text{pr}_{\mathbf{v}_k}(\mathbf{w}_j)$$

for alle $1 \leq j \leq n$; og lad endvidere $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ være familien af vektorer i V , hvor

$$\mathbf{u}_j = \mathbf{v}_j \cdot \|\mathbf{v}_j\|^{-1}$$

for alle $1 \leq j \leq n$. Da gælder følgende udsagn:

- (1) *Familierne $(\mathbf{u}_1, \dots, \mathbf{u}_n)$, $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ frembringer alle det samme underrum af V .*
- (2) *Familien $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er ortogonal med hensyn til $\langle -, - \rangle$.*
- (3) *Familien $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ er ortonormal med hensyn til $\langle -, - \rangle$.*

Hvis $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ er en basis for V , da er $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ og $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ dermed baser for V , som er henholdsvis ortogonale og ortonormale med hensyn til $\langle -, - \rangle$.

Vi bemærker specielt, at $\mathbf{v}_1 = \mathbf{w}_1$ og $\mathbf{u}_1 = \mathbf{w}_1 \cdot \|\mathbf{w}_1\|^{-1}$.

Bevis Vi beviser påstanden ved induktion på $n \geq 0$. Tilfældet $n = 0$ er trivielt, så vi antager, at påstanden er vist for $n = r - 1$, og viser den for $n = r$. Ifølge den induktive antagelse frembringer $(r - 1)$ -tuplerne $(\mathbf{v}_1, \dots, \mathbf{v}_{r-1})$ og $(\mathbf{w}_1, \dots, \mathbf{w}_{r-1})$ samme underrum af V , og derfor frembringer r -tuplerne $(\mathbf{v}_1, \dots, \mathbf{v}_{r-1}, \mathbf{w}_r)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_{r-1}, \mathbf{w}_r)$ også samme underrum W af V . Fra definitionen af $\mathbf{v}_r$ ved vi endvidere, at

$$\mathbf{w}_r = \mathbf{v}_r + \sum_{1 \leq k < r} \mathbf{v}_k \cdot \frac{\langle \mathbf{v}_k, \mathbf{w}_r \rangle}{\langle \mathbf{v}_k, \mathbf{v}_k \rangle},$$

hvilket viser, at r -tuplerne $(\mathbf{v}_1, \dots, \mathbf{v}_{r-1}, \mathbf{v}_r)$ og $(\mathbf{v}_1, \dots, \mathbf{v}_{r-1}, \mathbf{w}_r)$ begge frembringer W , og vi har dermed vist, at $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_r)$ begge frembringer W . Da r -tuplen $(\mathbf{w}_1, \dots, \mathbf{w}_r)$ er lineært uafhængig, udgør denne en basis for W , og sætning 4.3.24 viser da, at r -tuplen $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ ligeledes er en basis for W . Vektorerne $\mathbf{v}_1, \dots, \mathbf{v}_r$ er derfor

6 Vektorrum med indre produkt

alle forskellige fra $\mathbf{0}$, hvilket viser, at skalarerne $\|\mathbf{v}_1\|, \dots, \|\mathbf{v}_r\|$ alle er forskellige fra 0, således at vektorerne $\mathbf{u}_1, \dots, \mathbf{u}_r$ er veldefinerede. Endvidere viser

$$\mathbf{v}_1 \cdot a_1 + \dots + \mathbf{v}_r \cdot a_r = \mathbf{u}_1 \cdot (\|\mathbf{v}_1\| \cdot a_1) + \dots + \mathbf{u}_r \cdot (\|\mathbf{v}_r\| \cdot a_r),$$

at også $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ frembringer W . Dette viser induktionsskridtet for (1).

Ifølge den induktive antagelse er $(\mathbf{v}_1, \dots, \mathbf{v}_{r-1})$ endvidere ortogonal med hensyn til $\langle -, - \rangle$. For at vise, at også $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ er ortogonal med hensyn til $\langle -, - \rangle$, skal vi derfor blot vise, at $\langle \mathbf{v}_i, \mathbf{v}_r \rangle = 0$ for alle $1 \leq i \leq r-1$. Så vi lader $1 \leq i \leq r-1$ være fastholdt og udregner, at

$$\begin{aligned} \langle \mathbf{v}_i, \mathbf{v}_r \rangle &= \left\langle \mathbf{v}_i, \mathbf{w}_r - \sum_{1 \leq k < r} \mathbf{v}_k \cdot \frac{\langle \mathbf{v}_k, \mathbf{w}_r \rangle}{\langle \mathbf{v}_k, \mathbf{v}_k \rangle} \right\rangle \\ &\stackrel{(H2)}{=} \langle \mathbf{v}_i, \mathbf{w}_r \rangle - \sum_{1 \leq k < r} \langle \mathbf{v}_i, \mathbf{v}_k \rangle \cdot \frac{\langle \mathbf{v}_k, \mathbf{w}_r \rangle}{\langle \mathbf{v}_k, \mathbf{v}_k \rangle} \\ &= \langle \mathbf{v}_i, \mathbf{w}_r \rangle - \langle \mathbf{v}_i, \mathbf{v}_i \rangle \cdot \frac{\langle \mathbf{v}_i, \mathbf{w}_r \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} = 0, \end{aligned}$$

hvor den sidste lighed følger af, at $\langle \mathbf{v}_i, \mathbf{v}_k \rangle = 0$ for alle $1 \leq k < r$ med $i \neq k$. Dermed er $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ ortogonal med hensyn til $\langle -, - \rangle$, hvilket viser induktionsskridtet for (2).

Endelig viser udregningen

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \langle \mathbf{v}_i \cdot \|\mathbf{v}_i\|^{-1}, \mathbf{v}_j \cdot \|\mathbf{v}_j\|^{-1} \rangle = \|\mathbf{v}_i\|^{-1} \cdot \langle \mathbf{v}_i, \mathbf{v}_j \rangle \cdot \|\mathbf{v}_j\|^{-1} = \begin{cases} 1 & \text{hvis } i = j, \\ 0 & \text{hvis } i \neq j, \end{cases}$$

at $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ er ortonormal med hensyn til $\langle -, - \rangle$. Dette viser induktionsskridtet for (3) og dermed sætningen. Den sidste bemærkning følger af (1)–(3) og sætning 4.3.24. $\square$

Korollar 6.2.11 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. Ethvert endeligt frembragt $\mathbb{F}$ -vektorrum med indre produkt $(V, \langle -, - \rangle)$ har en basis, der er ortonormal med hensyn til $\langle -, - \rangle$.*

Bevis Ifølge sætning 4.3.15 findes der en endelig basis for V . Vi vælger derfor en sådan basis $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ og anvender Gram-Schmidt-ortogonalisering til at omdanne denne til en basis $(\mathbf{u}_1, \dots, \mathbf{u}_n)$, der er ortonormal med hensyn til $\langle -, - \rangle$. $\square$

Eksempel 6.2.12 Vi anvender Gram-Schmidt-ortogonalisering til at omdanne basen

$$\left(\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \mathbf{w}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$$

for $\mathbb{F}^3$ til en basis $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, der er ortogonal med hensyn til standard-indreproduktet $\langle -, - \rangle$, og til en basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, der er ortonormal med hensyn til $\langle -, - \rangle$.

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ \mathbf{v}_2 &= \mathbf{w}_2 - \mathbf{v}_1 \cdot \frac{\langle \mathbf{v}_1, \mathbf{w}_2 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1+0+2}{1+1+1} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \\ \mathbf{v}_3 &= \mathbf{w}_3 - \mathbf{v}_1 \cdot \frac{\langle \mathbf{v}_1, \mathbf{w}_3 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} - \mathbf{v}_2 \cdot \frac{\langle \mathbf{v}_2, \mathbf{w}_3 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \\ &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{0+1+1}{1+1+1} - \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cdot \frac{0-1+1}{0+1+1} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{3}.\end{aligned}$$

Vi har hermed fundet den nye basis

$$\left(\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{3} \right),$$

der er ortogonal med hensyn til $\langle -, - \rangle$. Endelig normerer vi vektorerne i denne basis og får derved basen

$$\left(\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{3}}, \mathbf{u}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}, \mathbf{u}_3 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{6}} \right),$$

som er ortonormal med hensyn til $\langle -, - \rangle$. Her kan vi i stedet for at normere $\mathbf{v}_3$ ligeså godt normere $\mathbf{v}'_3 = \mathbf{v}_3 \cdot 3$, sådan at vi slipper for den ubehagelige faktor $1/3$.

Definition 6.2.13 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt. Hvis $U \subset V$ er et underrum, så kaldes

$$U^\perp = \{ \mathbf{v} \in V \mid \langle \mathbf{u}, \mathbf{v} \rangle = 0 \text{ for alle } \mathbf{u} \in U \} \subset V$$

for det *ortogonale komplement* af $U \subset V$ med hensyn til $\langle -, - \rangle$.

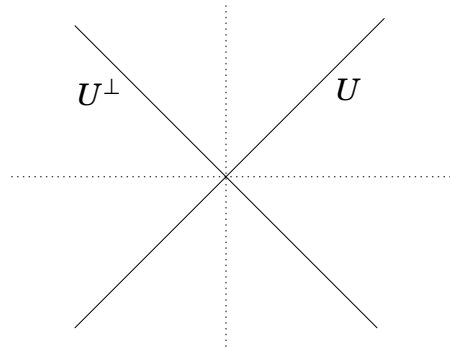
6 Vektorrum med indre produkt

Lemma 6.2.14 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt. Hvis $U \subset V$ er et underrum, da er dets ortogonale komplement $U^\perp \subset V$ med hensyn til $\langle -, - \rangle$ ligeledes et underrum. Desuden er $U \cap U^\perp = \{\mathbf{0}\}$.*

Bevis Vi efterviser, at $U^\perp \subset V$ opfylder definition 4.1.4 (1)–(3). Vi har $\mathbf{0} \in U^\perp$, idet $\langle \mathbf{u}, \mathbf{0} \rangle = 0$ for alle $\mathbf{u} \in U$, så (1) er opfyldt. Hvis $\mathbf{v}_1, \mathbf{v}_2 \in U^\perp$, så gælder det for alle $\mathbf{u} \in U$, at $\langle \mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2 \rangle = \langle \mathbf{u}, \mathbf{v}_1 \rangle + \langle \mathbf{u}, \mathbf{v}_2 \rangle = 0 + 0 = 0$, så (2) er opfyldt. Hvis $\mathbf{v} \in U^\perp$ og $a \in \mathbb{F}$, så gælder det for alle $\mathbf{u} \in U$, at $\langle \mathbf{u}, \mathbf{v} \cdot a \rangle = \langle \mathbf{u}, \mathbf{v} \rangle \cdot a = 0 \cdot a = 0$, så også (3) er opfyldt. Dermed er $U^\perp \subset V$ et underrum. Endelig følger $U \cap U^\perp = \{\mathbf{0}\}$ af positivitetsbetingelsen (P). $\square$

Eksempel 6.2.15 For underrummet $U = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 = x_2\} \subset \mathbb{R}^2$ fra eksempel 4.1.5 er dets ortogonale komplement med hensyn til standard-indreproduktet givet ved

$$U^\perp = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 + x_2 = 0 \right\} \subset \mathbb{R}^2.$$



Figur 6.3: Underrummene $U, U^\perp \subset \mathbb{R}^2$ fra eksempel 6.2.15.

Sætning 6.2.16 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt, lad $U \subset V$ være et endeligt frembragt underrum, og lad $U^\perp \subset V$ være dets ortogonale komplement. Enhver vektor $\mathbf{w} \in V$ kan da skrives entydigt som*

$$\mathbf{w} = \text{pr}_U(\mathbf{w}) + (\mathbf{w} - \text{pr}_U(\mathbf{w}))$$

med $\text{pr}_U(\mathbf{w}) \in U$ og $\mathbf{w} - \text{pr}_U(\mathbf{w}) \in U^\perp$. Hvis endvidere $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ er en basis for U , der er ortogonal med hensyn til $\langle -, - \rangle$, så er

$$\text{pr}_U(\mathbf{w}) = \mathbf{v}_1 \cdot \frac{\langle \mathbf{v}_1, \mathbf{w} \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} + \dots + \mathbf{v}_n \cdot \frac{\langle \mathbf{v}_n, \mathbf{w} \rangle}{\langle \mathbf{v}_n, \mathbf{v}_n \rangle}.$$

Bevis Vi bemærker først, at hvis $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ er en vilkårlig basis for U , og hvis $\mathbf{v} \in V$ opfylder, at $\langle \mathbf{u}_i, \mathbf{v} \rangle = 0$ for alle $1 \leq i \leq n$, så er $\mathbf{v} \in U^\perp$. For vi kan skrive en vilkårlig vektor $\mathbf{u} \in U$ som en linearkombination $\mathbf{u} = \mathbf{u}_1 a_1 + \dots + \mathbf{u}_n a_n$, og

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}_1 a_1 + \dots + \mathbf{u}_n a_n, \mathbf{v} \rangle = a_1^* \langle \mathbf{u}_1, \mathbf{v} \rangle + \dots + a_n^* \langle \mathbf{u}_n, \mathbf{v} \rangle = 0.$$

Vi lader nu $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ være en basis for U , der er ortonormal med hensyn til $\langle -, - \rangle$. En sådan basis findes ifølge korollar 6.2.11. Hvis $\mathbf{w} = \mathbf{u} + \mathbf{v}$ med $\mathbf{u} \in U$ og $\mathbf{v} \in U^\perp$, da er

$$\langle \mathbf{u}_i, \mathbf{w} \rangle = \langle \mathbf{u}_i, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}_i, \mathbf{u} \rangle + \langle \mathbf{u}_i, \mathbf{v} \rangle = \langle \mathbf{u}_i, \mathbf{u} \rangle.$$

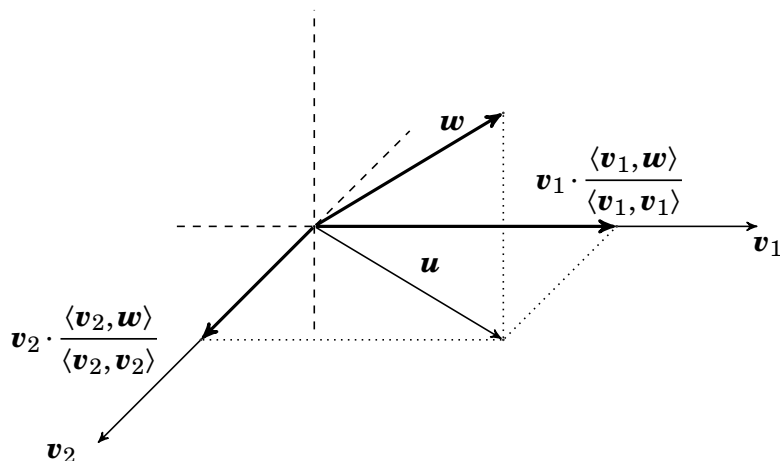
Derfor er vektoren $\mathbf{u} \in U$ nødvendigvis givet ved

$$\mathbf{u} = \mathbf{u}_1 \cdot \langle \mathbf{u}_1, \mathbf{w} \rangle + \dots + \mathbf{u}_n \cdot \langle \mathbf{u}_n, \mathbf{w} \rangle.$$

Vi definerer nu $\mathbf{u} \in U$ til at være denne vektor og definerer $\mathbf{v} = \mathbf{w} - \mathbf{u}$. Vi mangler at vise, at $\mathbf{v} \in U^\perp$. Da $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$, så gælder det for alle $1 \leq i \leq n$, at

$$\langle \mathbf{u}_i, \mathbf{v} \rangle = \langle \mathbf{u}_i, \mathbf{w} - (\mathbf{u}_1 \cdot \langle \mathbf{u}_1, \mathbf{w} \rangle + \dots + \mathbf{u}_n \cdot \langle \mathbf{u}_n, \mathbf{w} \rangle) \rangle = \langle \mathbf{u}_i, \mathbf{w} \rangle - \langle \mathbf{u}_i, \mathbf{w} \rangle = 0.$$

Altså er $\mathbf{v} \in U^\perp$ som ønsket. □



Figur 6.4: Den ortogonale projektion $\mathbf{u} = \text{pr}_U(\mathbf{w})$ af $\mathbf{w}$ på underrummet $U \subset \mathbb{R}^3$ med basen $(\mathbf{v}_1, \mathbf{v}_2)$, der er ortogonal med hensyn til standard-indreproduktet.

Definition 6.2.17 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$ -vektorrum med indre produkt, og lad $U \subset V$ være et endeligt frembragt underrum. Afbildningen

$$\text{pr}_U: V \rightarrow V$$

kaldes for den *ortogonale projektion* på $U \subset V$ med hensyn til $\langle -, - \rangle$.

6 Vektorrum med indre produkt

Vi bemærker, at formelen for $\text{pr}_U(\mathbf{w})$ i sætning 6.2.16 samt egenskaberne (H1)–(H2) af et indre produkt viser, at den ortogonale projektion $\text{pr}_U: V \rightarrow V$ er en lineær afbildning.

Eksempel 6.2.18 Vi ønsker at bestemme den ortogonale projektion $\text{pr}_U: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ på underrummet $U \subset \mathbb{R}^3$ frembragt af familien

$$\left(\mathbf{w}_1 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right)$$

med hensyn til standard-indreproduktet på $\mathbb{R}^3$, og vi anvender derfor Gram-Schmidt algoritmen til at omdanne basen $(\mathbf{w}_1, \mathbf{w}_2)$ for U til en ny basis $(\mathbf{v}_1, \mathbf{v}_2)$, der er ortogonal med hensyn til standard-indreproduktet.

$$\mathbf{v}_1 = \mathbf{w}_1$$

$$\mathbf{v}_2 = \mathbf{w}_2 - \mathbf{v}_1 \cdot \frac{\langle \mathbf{v}_1, \mathbf{w}_2 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cdot \frac{0 - 2 + 0}{0 + 1 + 1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Dermed er den ortogonale projektion $\text{pr}_U: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ givet ved

$$\text{pr}_U(\mathbf{x}) = \mathbf{v}_1 \cdot \frac{\langle \mathbf{v}_1, \mathbf{x} \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} + \mathbf{v}_2 \cdot \frac{\langle \mathbf{v}_2, \mathbf{x} \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cdot \frac{0 - x_2 + x_3}{0 + 1 + 1} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{x_1 + x_2 + x_3}{1 + 1 + 1},$$

hvoraf vi aflæser, at $\text{pr}_U(\mathbf{x}) = A\mathbf{x}$, hvor

$$A = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 5/6 & -1/6 \\ 1/3 & -1/6 & 5/6 \end{pmatrix}.$$

Med andre ord er $A \in M_3(\mathbb{R})$ den matrix, der repræsenterer $\text{pr}_U: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ med hensyn til standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for både domænet og codomænet.

6.3 Lineære isometrier

Vi vil nu betragte de lineære afbildninger mellem vektorrum med indre produkt, der bevarer det indre produkt. Disse afbildninger bevarer da de geometriske strukturer, som det indre produkt repræsenterer, såsom længder og vinkler. Vi viser specielt, at en lineær isomorfi $f: V \rightarrow V$ af et endeligt frembragt vektorrum bevarer et indre produkt $\langle -, - \rangle$ på V , hvis og kun hvis den matrix A , der repræsenterer $f: V \rightarrow V$ med hensyn til

en fælles ortonormal basis for domæne og codomæne, opfylder, at $A^* = A^{-1}$. Her er A^* den adjungerede matrix defineret i afsnit 2.6.

Definition 6.3.1 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være $\mathbb{F}$ -vektorrum med indre produkt. En lineær afbildning $f: W \rightarrow V$, der opfylder

$$\langle f(\mathbf{w}), f(\mathbf{w}') \rangle_V = \langle \mathbf{w}, \mathbf{w}' \rangle_W$$

for alle $\mathbf{w}, \mathbf{w}' \in W$, kaldes for en *lineær isometri* med hensyn til $\langle -, - \rangle_W$ og $\langle -, - \rangle_V$.

Eksempel 6.3.2 (1) Vi betragter igen rotationen $r_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gennem θ radianer imod urets retning fra eksempel 2.3.13, der er givet ved $r_\theta(\mathbf{x}) = B_\theta \mathbf{x}$ med

$$B_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Udregningen

$$\begin{aligned} \langle r_\theta(\mathbf{x}), r_\theta(\mathbf{y}) \rangle &= \left\langle \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \mathbf{x}, \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \mathbf{y} \right\rangle \\ &= \cos(\theta)x_1 - \sin(\theta)x_2)(\cos(\theta)y_1 - \sin(\theta)y_2) \\ &\quad + (\sin(\theta)x_1 + \cos(\theta)x_2)(\sin(\theta)y_1 + \cos(\theta)y_2) \\ &= (\cos(\theta)^2 + \sin(\theta)^2)x_1y_1 + (\cos(\theta)^2 + \sin(\theta)^2)x_2y_2 \\ &= x_1y_1 + x_2y_2 = \langle \mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

viser da, at $r_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ er en lineær isometri med hensyn til standard-indreproduktet.

(2) Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $f: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ være den vægtede skalering

$$f(\mathbf{x}) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{x}.$$

Da viser udregningen

$$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle = \left\langle \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{x}, \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{y} \right\rangle = |\lambda_1|^2 x_1 y_1 + |\lambda_2|^2 x_2 y_2,$$

at $f: \mathbb{F}^2 \rightarrow \mathbb{F}^2$ er en lineær isometri med hensyn til standard-indreproduktet på $\mathbb{F}^2$, hvis og kun hvis $|\lambda_1| = 1$ og $|\lambda_2| = 1$.

Navnet “isometri” er noget forvirrende, idet det kunne antyde, at en lineær isometri

6 Vektorrum med indre produkt

altid er en isomorfi, hvilket *ikke* er tilfældet. Derimod har vi følgende resultat.

Lemma 6.3.3 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være to $\mathbb{F}$ -vektorrum med indre produkt.*

- (1) *En lineær isometri $f: W \rightarrow V$ er nødvendigvis injektiv.*
- (2) *Hvis en lineær isometri $f: W \rightarrow V$ er en isomorfi med invers afbildning $g: V \rightarrow W$, da er den inverse afbildning også en lineær isometri.*

Bevis Vi viser først (1). Så vi lader $w \in W$ og antager, at $f(w) = 0$. Da er

$$\langle w, w \rangle_W = \langle f(w), f(w) \rangle_V = \langle 0, 0 \rangle_V = 0,$$

og positivitetsegenskaben (P) af $\langle -, - \rangle_W$ medfører derfor, at $w = 0$. Da afbildningen $f: W \rightarrow V$ er lineær, konkluderer vi fra lemma 4.2.9, at den er injektiv.

Vi viser dernæst (2). Den inverse afbildning $g: V \rightarrow W$ er lineær ifølge sætning 4.2.7. Så vi lader $v, v' \in V$ og udregner, at

$$\langle g(v), g(v') \rangle_W = \langle f(g(v)), f(g(v')) \rangle_V = \langle v, v' \rangle_V,$$

hvor den første gælder, da $f: W \rightarrow V$ er en lineær isometri, mens den anden gælder, da afbildningerne $f: W \rightarrow V$ og $g: V \rightarrow W$ er hinandens inverse. $\square$

En lineær isometri som er også en isomorfi kaldes en *isometrisk isomorfi*.

Eksempel 6.3.4 Lad $u \in \mathbb{F}^n$ være en enhedsvektor med hensyn til standard-indreproduktet. Da er afbildningen $\sigma_u: \mathbb{F} \rightarrow \mathbb{F}^n$ defineret ved $\sigma_u(a) = ua$ en lineær isometri, men den er kun en isomorfi, hvis $n = 1$.

Vi indfører nu begrebet *adjunktion* for lineære afbildninger mellem vektorrum med indre produkt. Vi anvender senere dette begreb til bedre at forstå lineære isometrier.

Definition 6.3.5 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være $\mathbb{F}$ -vektorrum med indre produkt. To lineære afbildninger $f: W \rightarrow V$ og $g: V \rightarrow W$ siges at være *adjungerede* med hensyn til $\langle -, - \rangle_V$ og $\langle -, - \rangle_W$, hvis

$$\langle v, f(w) \rangle_V = \langle g(v), w \rangle_W$$

for alle $v \in V$ og $w \in W$.

Vi skal vise, at enhver lineær afbildning mellem to endeligt frembragte vektorrum med indre produkt har en entydigt bestemt adjungeret afbildning. Vi begynder med at vise entydighedsudsagnet, som gælder mere generelt.

Lemma 6.3.6 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være to $\mathbb{F}$ -vektorrum med indre produkt. Lad $f: W \rightarrow V$ være en lineær afbildning, og lad $g, h: V \rightarrow W$ være lineære afbildninger, der begge er adjungerede til $f: W \rightarrow V$ med hensyn til $\langle -, - \rangle_V$ og $\langle -, - \rangle_W$. Da er $g = h$.*

Bevis Vi antager, at både $g: V \rightarrow W$ og $h: V \rightarrow W$ er adjungerede til $f: W \rightarrow V$ med hensyn til $\langle -, - \rangle_V$ og $\langle -, - \rangle_W$. For alle $v \in V$ og $w \in W$ gælder det da, at

$$\langle g(w) - h(w), v \rangle_W = \langle g(w), v \rangle_W - \langle h(w), v \rangle_W = \langle w, f(v) \rangle_V - \langle w, f(v) \rangle_V = 0.$$

Dette gælder specielt for $v = g(w) - h(w)$, og positivitetsegenskaben (P) ved et indre produkt medfører derfor, at $g(w) - h(w) = \mathbf{0}$ som ønsket. $\square$

Sætning 6.3.7 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være to endeligt frembragte $\mathbb{F}$ -vektorrum med indre produkt, og lad $f: W \rightarrow V$ være en lineær afbildning. Lad $(v_1, \dots, v_m)$ og $(w_1, \dots, w_n)$ være baser for V og W , der er ortonormale med hensyn til henholdsvis $\langle -, - \rangle_V$ og $\langle -, - \rangle_W$.*

- (1) *Der findes en entydigt bestemt lineær afbildning $g: V \rightarrow W$, sådan at $f: W \rightarrow V$ og $g: V \rightarrow W$ er adjungerede med hensyn til $\langle -, - \rangle_V$ og $\langle -, - \rangle_W$.*
- (2) *Hvis matricen $A \in M_{m,n}(\mathbb{F})$ repræsenterer afbildningen $f: W \rightarrow V$ med hensyn til de givne ortonormale baser, så repræsenterer den adjungerede matrix $A^* \in M_{n,m}(\mathbb{F})$ den adjungerede afbildning $g: V \rightarrow W$ med hensyn til disse baser.*

Bevis Vi påstår, at hvis to lineære afbildninger $f: W \rightarrow V$ og $g: V \rightarrow W$ er repræsenteret ved adjungerede matricer $A \in M_{m,n}(\mathbb{F})$ og $A^* \in M_{n,m}(\mathbb{F})$ med hensyn til ortonormale baser $(v_1, \dots, v_m)$ for V og $(w_1, \dots, w_n)$ for W , så er de adjungerede lineære afbildninger. Dette viser både (1) og (2), da entydighedsudsagnet i (1) allerede er vist i lemma 6.3.6. Så lad $v \in V$ og $w \in W$, og lad $x \in \mathbb{F}^m$ og $y \in \mathbb{F}^n$ være koordinaterne for v og w med hensyn til de givne baser. Vi har da som ønsket, at

$$\langle v, f(w) \rangle_V = x^*(A y) = (A^* x)^* y = \langle g(v), w \rangle_W,$$

hvor de to identiteter til højre og venstre følger af sætning 6.2.6, mens den midterste identitet følger af sætning 2.6.7. $\square$

Eksempel 6.3.8 Vi lader $W = \mathbb{R}^2$ og $V = \mathbb{R}^3$ med standard-indreprodukterne og betragter den lineære afbildning $f: W \rightarrow V$ defineret ved $f(\mathbf{x}) = A\mathbf{x}$, hvor

$$A = \begin{pmatrix} 2 & 3 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Matricen A repræsenterer altså $f: W \rightarrow V$ med hensyn til standardbaserne $(\mathbf{e}_1, \mathbf{e}_2)$ og $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, og da disse er ortogonale med hensyn til standard-indreprodukterne, så repræsenterer den adjungerede matrix

$$A^* = A^t = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & -1 \end{pmatrix}$$

ifølge sætning 6.3.7 den adjungerede afbildning $g: V \rightarrow W$. Udregningen

$$\begin{aligned} \langle \mathbf{x}, f(\mathbf{y}) \rangle &= x_1(2y_1 + 3y_2) + x_2y_2 + x_3(y_1 - y_2) \\ &= (2x_1 + x_3)y_1 + (3x_1 + x_2 - x_3)y_2 = \langle g(\mathbf{x}), \mathbf{y} \rangle \end{aligned}$$

illustrerer dette forhold.

Udsagnet, at en lineær afbildning er en isometri, kan udtrykkes ved hjælp af dens adjungerede lineære afbildning på følgende vis.

Lemma 6.3.9 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være to $\mathbb{F}$ -vektorrum med indre produkt. Hvis to lineære afbildninger $f: W \rightarrow V$ og $g: V \rightarrow W$ er adjungerede med hensyn til $\langle -, - \rangle_V$ og $\langle -, - \rangle_W$, da er $f: W \rightarrow V$ en lineær isometri med hensyn til $\langle -, - \rangle_W$ og $\langle -, - \rangle_V$, hvis og kun hvis $g \circ f = \text{id}_W: W \rightarrow W$.*

Bevis Da $f: W \rightarrow V$ og $g: V \rightarrow W$ er adjungerede, så er

$$\langle f(\mathbf{w}), f(\mathbf{w}') \rangle_V = \langle (g \circ f)(\mathbf{w}), \mathbf{w}' \rangle_W$$

for alle $\mathbf{w}, \mathbf{w}' \in W$. Derfor er $f: W \rightarrow V$ en lineær isometri, hvis og kun hvis

$$\langle (g \circ f)(\mathbf{w}), \mathbf{w}' \rangle_W = \langle \mathbf{w}, \mathbf{w}' \rangle_W$$

for alle $\mathbf{w}, \mathbf{w}' \in W$. Hvis dette gælder, så får vi for $\mathbf{w}' = (g \circ f)(\mathbf{w}) - \mathbf{w}$, at

$$\langle \mathbf{w}', \mathbf{w}' \rangle_W = \langle (g \circ f)(\mathbf{w}) - \mathbf{w}, \mathbf{w}' \rangle_W = \langle (g \circ f)(\mathbf{w}), \mathbf{w}' \rangle_W - \langle \mathbf{w}, \mathbf{w}' \rangle_W = 0,$$

hvorfra vi konkluderer, at $\mathbf{w}' = \mathbf{0}$ og dermed, at $g \circ f = \text{id}_W$. Hvis omvendt $g \circ f = \text{id}_W$, så er betingelsen ovenfor trivielt opfyldt. $\square$

Sætning 6.3.10 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være to endeligt frembragte $\mathbb{F}$ -vektorrums med indre produkt. Lad $f: W \rightarrow V$ være en lineær afbildning, og lad $A \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f: W \rightarrow V$ med hensyn til ortonormale baser $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ og $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ for henholdsvis V og W . De følgende udsagn (1)–(3) er ækvivalente.*

- (1) *Afbildningen $f: W \rightarrow V$ er en lineær isometri.*
- (2) *Det gælder, at $A^*A = I_n$.*
- (3) *Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ af vektorer i $\mathbb{F}^m$, der består af søjlerne i A , er ortonormal med hensyn til standard indre-produktet.*

Tilsvarende er de følgende udsagn (4)–(6) ækvivalente.

- (4) *Afbildningen $f: W \rightarrow V$ er en isometrisk isomorfi.*
- (5) *Matricen A er invertibel, og $A^{-1} = A^*$.*
- (6) *Familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ af vektorer i $\mathbb{F}^m$, der består af søjlerne i A , udgør en basis, der er ortonormal med hensyn til standard indre-produktet.*

Bevis Ifølge sætning 6.3.7 er den adjungerede afbildning $g: V \rightarrow W$ repræsenteret af den adjungerede matrix $A^* \in M_{n,m}(\mathbb{F})$ med hensyn til de givne baser. Så $g \circ f = \text{id}_W$, hvis og kun hvis $A^*A = I_n$, og ækvivalensen af (1) og (2) følger derfor af lemma 6.3.9. Vi bemærker endvidere, at den (i, j) 'te indgang i A^*A er lig med $\mathbf{a}_i^* \mathbf{a}_j$, som netop er standard indre-produktet $\langle \mathbf{a}_i, \mathbf{a}_j \rangle$ af vektorerne $\mathbf{a}_i$ og $\mathbf{a}_j$ i $\mathbb{F}^m$. Derfor er $A^*A = I_n$, hvis og kun hvis $\langle \mathbf{a}_i, \mathbf{a}_j \rangle = \delta_{ij}$ for alle $1 \leq i, j \leq n$, hvilket viser, at (2) og (3) er ækvivalente.

Vi viser nu, at (4) og (5) er ækvivalente. Hvis (4) er opfyldt, da er $f: W \rightarrow V$ både en lineær isometri, hvilket ifølge ækvivalensen af (1) og (2) medfører, at $A^*A = I_n$, og en isomorfi, hvilket medfører, at A er invertibel. Endvidere er

$$A^* = A^*(AA^{-1}) = (A^*A)A^{-1} = A^{-1},$$

så (5) er opfyldt. Omvendt, hvis (5) er opfyldt, da medfører ækvivalensen af (1) og (2), at $f: W \rightarrow V$ er en lineær isometri, og da A er invertibel, er $f: W \rightarrow V$ endvidere en isomorfi, så (4) er opfyldt. Endelig viser korollar 4.3.12, at A er invertibel, hvis og kun hvis familien $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ af vektorer i $\mathbb{F}^m$, der består af søjlerne i A , er en basis, og ækvivalensen af (5) og (6) følger derfor fra ækvivalensen af (2) og (3). $\square$

6 Vektorrum med indre produkt

Eksempel 6.3.11 (1) Afbildningen $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ fra eksempel 6.3.8 er ikke en isometri med hensyn til standard-indreprodukterne. For udregningen

$$A^*A = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 5 & 11 \end{pmatrix}$$

viser, at A^*A er forskellig fra I_2 .

(2) Vi har allerede vist, at rotationsafbildningen $r_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ fra eksempel 6.3.2 er en isometrisk isomorfi med hensyn til standard-indreprodukterne, men udregningen

$$B_\theta^* = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} = B_\theta^{-1}$$

bekræfter, at dette er tilfældet.

(3) Familien $(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1)$ af vektorer i $\mathbb{F}^3$ udgør en basis, der er ortonormal med hensyn til standard indre-produktet. Sætning 6.3.10 viser derfor, at matricen

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

er invertibel med invers

$$A^{-1} = A^* = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

samt at afbildningen $f: \mathbb{F}^3 \rightarrow \mathbb{F}^3$ givet ved $f(\mathbf{x}) = A\mathbf{x}$ er en isometrisk isomorfi. Denne afbildning er en rotation omkring linjen gennem $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$.

Vi vil i resten af dette afsnit udelukkende betragte lineære endomorfier, hvilket er lineære afbildninger, hvis domæne og codomæne er identitiske.

Definition 6.3.12 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$ og lad $(V, \langle -, - \rangle)$ være et endeligt frembragt $\mathbb{F}$ -vektorrum med indre produkt. Lad $f: V \rightarrow V$ være en lineær endomorfi, og lad $g: V \rightarrow V$ være den adjungerede lineære endomorfi med hensyn til $\langle -, - \rangle$.

(1) Endomorfien $f: V \rightarrow V$ er *selvadjungeret* med hensyn til $\langle -, - \rangle$, hvis $f = g$.

(2) Endomorfien $f: V \rightarrow V$ er *normal* med hensyn til $\langle -, - \rangle$, hvis $f \circ g = g \circ f$.

Vi benytter de resultater, vi har vist ovenfor, til at vise følgende resultat for lineære endomorfier.

Sætning 6.3.13 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et endeligt frembragt $\mathbb{F}$ -vektorrum med indre produkt, og lad $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ være en basis for V , der er ortonormal med hensyn til $\langle -, - \rangle$. Lad $f: V \rightarrow V$ være en lineær endomorfi, og lad $A \in M_n(\mathbb{F})$ være matricen, der repræsenterer $f: V \rightarrow V$ med hensyn til den givne ortonormale basis.

(1) Endomorfien $f: V \rightarrow V$ er en isometri med hensyn til $\langle -, - \rangle$, hvis og kun hvis

$$A^* A = I.$$

Dette er endvidere tilfældet, hvis og kun hvis $A^* = A^{-1}$.

(2) Endomorfien $f: V \rightarrow V$ er selvadjungeret med hensyn til $\langle -, - \rangle$, hvis og kun hvis

$$A^* = A.$$

(3) Endomorfien $f: V \rightarrow V$ er normal med hensyn til $\langle -, - \rangle$, hvis og kun hvis

$$A^* A = A A^*.$$

Bevis Den første del af (1) følger fra ækvivalensen af (1) og (2) i sætning 6.3.10. Da A er kvadratisk, så medfører identiteten $A^* A = I$ ifølge sætning 2.5.22, at A er invertibel med invers $A^{-1} = A^*$. Påstandene (2) og (3) følger umiddelbart fra sætning 6.3.7 og definition 6.3.12. $\square$

De klassiske betegnelser for de kvadratiske matricer $A \in M_n(\mathbb{F})$, der opfylder enten $A^* = A^{-1}$ eller $A^* = A$, afhænger desværre af $\mathbb{F}$. Det er der ikke noget at gøre ved, så vi anfører disse klassiske betegnelser i den følgende definition, hvor vi også genkalder nogle definitioner fra afsnit 2.6.

Definition 6.3.14 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$.

- (1) En kvadratisk matrix $Q \in M_m(\mathbb{R})$, der opfylder $Q^* = Q^{-1}$, kaldes *ortogonal*; og en kvadratisk matrix $U \in M_m(\mathbb{C})$, der opfylder $U^* = U^{-1}$, kaldes *unitær*.
- (2) En kvadratisk matrix $A \in M_m(\mathbb{R})$, der opfylder $A^* = A$, kaldes *symmetrisk*; og en kvadratisk matrix $A \in M_m(\mathbb{C})$, der opfylder $A^* = A$, kaldes *hermitisk*.
- (3) En kvadratisk matrix $A \in M_m(\mathbb{F})$, der opfylder $A^* A = A A^*$, kaldes *normal*.

Vi bemærker, at den kvadratiske matrix P er invertibel med invers $P^{-1} = P^*$, hvis og kun hvis familien $(\mathbf{p}_1, \dots, \mathbf{p}_m)$ af vektorer i $\mathbb{F}^m$, der består af søjlerne i P , er ortonormal med hensyn til standard-indreproduktet; se beviset for sætning 6.3.10.

Eksempel 6.3.15 Enhver symmetrisk reel matrix er normal, men det omvendte er ikke tilfældet. For eksempel er matricen

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

normal, idet

$$A^* A = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} = A A^*,$$

men A er ikke symmetrisk. Vi bemærker også, at hvis B er en kvadratisk reel eller kompleks matrix, da opfylder matricen $A = B^* B$ altid, at $A^* = A$. I det reelle tilfælde er A således symmetrisk, mens den i det komplekse tilfælde er hermitisk.

Sætning 6.3.16 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$.

- (1) Hvis $A, B \in M_m(\mathbb{F})$ opfylder $A^* = A^{-1}$ og $B^* = B^{-1}$, da er også $(AB)^* = (AB)^{-1}$.
- (2) Hvis $A \in M_m(\mathbb{F})$ opfylder $A^* = A^{-1}$, da er også $(A^{-1})^* = (A^{-1})^{-1}$.

Bevis De to påstande følger af de to udregninger

$$\begin{aligned} (AB)^* &= B^* A^* = B^{-1} A^{-1} = (AB)^{-1} \\ A^{-1} (A^{-1})^* &= A^* (A^{-1})^* = (A^{-1} A)^* = I^* = I, \end{aligned}$$

hvor vi anvender sætning 2.6.7. □

Bemærkning 6.3.17 En kvadratisk matrix $S \in M_m(\mathbb{H})$, der opfylder $S^* = S^{-1}$ kaldes for en *unitær symplektisk* matrix, og sætning 6.3.16 gælder ligeledes i dette tilfælde. Man anvender klassisk betegnelserne

$$O(m) \subset GL_m(\mathbb{R}), \quad U(m) \subset GL_m(\mathbb{C}) \quad \text{og} \quad \text{Sp}(m) \subset GL_m(\mathbb{H})$$

for delmængden af de invertible $m \times m$ -matricer, der består af de $m \times m$ -matricer A , der opfylder $A^* = A^{-1}$. Sætning 6.3.16 viser, at $(O(m), \cdot)$, $(U(m), \cdot)$ og $(\text{Sp}(m), \cdot)$ er grupper, og de kaldes for den *ortogonale gruppe*, den *unitære gruppe* og den *symplektiske gruppe*. I modsætning hertil behøver en kvadratisk matrix A , der opfylder $A^* = A$, ikke at være invertibel; og hvis både $A^* = A$ og $B^* = B$, så er $(AB)^* = B^* A^* = BA$, hvilket generelt heller ikke er lig med AB . Så de selvadjungerede matricer udgør ikke en gruppe.

6.4 Spektralsætningen

I dette afsnit viser vi den såkaldte spektralsætning for lineære endomorfier af endeligt frembragte reelle og komplekse vektorrum med indre produkt. Sætningen siger, at der findes en ortonormal basis for vektorrummet, der består af egenvektorer for den lineære endomorfi, hvis og kun hvis denne er selvadjungeret i det reelle tilfælde og normal i det komplekse tilfælde. I matrixsprog siger sætningen, at der for en reel symmetrisk matrix A findes en ortogonal matrix Q , sådan at $Q^{-1}AQ$ er en reel diagonalmatrix, mens der for en normal kompleks matrix A findes en unitær matrix U , sådan at $U^{-1}AU$ er en kompleks diagonalmatrix. Beviset anvender algebraens fundamentalsætning til at faktorisere det karakteristiske polynomium hørende til den lineære endomorfi i et produkt af førstegradspolynomier. Denne del af beviset er derfor ikke konstruktiv, men givet denne faktorisering er resten af beviset algoritmisk, hvilket vi illustrerer med passende eksempler i slutningen af afsnittet.

Sætning 6.4.1 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. En matrix $A \in M_n(\mathbb{F})$ er både triangulær og normal, hvis og kun hvis den er en diagonalmatrix.*

Bevis Hvis $A = \text{diag}(a_1, \dots, a_n)$ er en diagonal matrix, så er A specielt en triangulær matrix, og udregningen

$$\begin{aligned} A^* A &= \text{diag}(a_1, \dots, a_n)^* \text{diag}(a_1, \dots, a_n) = \text{diag}(a_1^* a_1, \dots, a_n^* a_n), \\ A A^* &= \text{diag}(a_1, \dots, a_n) \text{diag}(a_1, \dots, a_n)^* = \text{diag}(a_1 a_1^*, \dots, a_n a_n^*) \end{aligned}$$

viser, at A er normal, idet den kommutative lov for multiplikation gælder i $\mathbb{F}$. Så vi antager omvendt, at $A = (a_{ij}) \in M_n(\mathbb{F})$ er en normal øvre triangulær matrix og viser, at A er en diagonal matrix; beviset i tilfældet, hvor A er en normal nedre triangulær matrix er tilsvarende. Da A er normal, gælder det for alle $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, at

$$\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle A^* A\mathbf{x}, \mathbf{y} \rangle = \langle A A^* \mathbf{x}, \mathbf{y} \rangle = \langle (A^*)^* A^* \mathbf{x}, \mathbf{y} \rangle = \langle A^* \mathbf{x}, A^* \mathbf{y} \rangle,$$

og ved at sætte $\mathbf{x} = \mathbf{y} = \mathbf{e}_k$, konkluderer vi derfor, at

$$|a_{1k}|^2 + \dots + |a_{nk}|^2 = \langle A\mathbf{e}_k, A\mathbf{e}_k \rangle = \langle A^* \mathbf{e}_k, A^* \mathbf{e}_k \rangle = |a_{k1}|^2 + \dots + |a_{kn}|^2$$

for alle $1 \leq k \leq n$. Da A endvidere er øvre triangulær, er $a_{ij} = 0$ for $i > j$, og det følger derfor af den identitet, vi netop har fundet, at

$$|a_{1k}|^2 + \dots + |a_{kk}|^2 = |a_{kk}|^2 + \dots + |a_{kn}|^2$$

for alle $1 \leq k \leq n$. Vi viser ved induktion på $1 \leq i \leq n$, at $a_{ij} = 0$ for $i < j$, og begynder med tilfældet $i = 1$. Den ovenstående identitet for $k = 1$ viser, at

$$|a_{11}|^2 = |a_{11}|^2 + |a_{12}|^2 + \dots + |a_{1n}|^2,$$

6 Vektorrum med indre produkt

hvorfor $a_{12} = \dots = a_{1n} = 0$ som ønsket. Så vi antager, at påstanden er vist for $i = r - 1$, og viser den for $i = r$. Det gælder da, at

$$|a_{rr}|^2 = |a_{1r}|^2 + \dots + |a_{rr}|^2 = |a_{rr}|^2 + \dots + |a_{rn}|^2,$$

idet den første identitet følger fra den induktive hypotese, mens den anden identitet er den identitet, vi viste ovenfor, for $k = r$. Derfor er $a_{r,r+1} = \dots = a_{rn} = 0$, hvilket viser induktionsskridtet og dermed påstanden. Det gælder altså, at $a_{ij} = 0$ for $i \neq j$, hvilket som ønsket viser, at A er en diagonalmatrix. $\square$

Den følgende sætning kaldes for spektralsætningen for symmetriske matricer.

Sætning 6.4.2 For en kvadratisk matrix $A \in M_n(\mathbb{R})$ er følgende udsagn ækvivalente:

- (1) Matricen A er symmetrisk.
- (2) Der findes en ortogonal matrix $Q \in M_n(\mathbb{R})$, sådan at $Q^{-1}AQ$ er en diagonalmatrix.

Bevis Vi antager først (1) og viser (2). Ifølge sætning 5.2.10 findes der en invertibel matrix $P \in M_n(\mathbb{R})$, sådan at matricen $P^{-1}AP \in M_n(\mathbb{R})$ er øvre triangulær. Endvidere producerer Gram-Schmidt-algoritmen en øvre triangulær matrix $C \in M_n(\mathbb{R})$, sådan at matricen $Q = PC \in M_n(\mathbb{R})$ er ortogonal. For hvis $(\mathbf{p}_1, \dots, \mathbf{p}_n)$ er basen, der består af søjlerne i P , så giver Gram-Schmidt algoritmen en basis $(\mathbf{u}_1, \dots, \mathbf{u}_n)$, der er ortonormal med hensyn til standard-indreproduktet på $\mathbb{R}^n$, hvor

$$\begin{aligned}\mathbf{u}_1 &= \mathbf{p}_1 c_{11} \\ \mathbf{u}_2 &= \mathbf{p}_1 c_{12} + \mathbf{p}_2 c_{22} \\ &\vdots \\ \mathbf{u}_n &= \mathbf{p}_1 c_{1n} + \mathbf{p}_2 c_{2n} + \dots + \mathbf{p}_n c_{nn},\end{aligned}$$

og Q er da den ortogonale matrix med søjler $\mathbf{u}_1, \dots, \mathbf{u}_n$, mens C er den øvre triangulære matrix med indgange c_{ij} . Vi påstår, at $Q^{-1}AQ$ er både normal og øvre triangulær. For

$$\begin{aligned}(Q^{-1}AQ)^* Q^{-1}AQ &= (Q^*AQ)^* Q^*AQ = Q^*A^*QQ^*AQ = Q^*A^*AQ \\ Q^{-1}AQ(Q^{-1}AQ)^* &= Q^*AQ(Q^*AQ)^* = Q^*AQQ^*A^*Q = Q^*AA^*Q,\end{aligned}$$

hvor vi bruger at $Q^* = Q^{-1}$. Da A er symmetrisk, er den normal, så vi har nu vist, at $Q^{-1}AQ$ er normal. Endvidere er

$$Q^{-1}AQ = (PC)^{-1}APC = C^{-1}P^{-1}APC,$$

hvor $P^{-1}AP$ og C er øvre triangulære matricer. Men produktet af to øvre triangulære matricer er igen en øvre triangulær matrix, og ved at bruge sætning 3.4.1 kan vi vise, at

også den inverse matrix af en invertibel øvre triangulær matrix igen er øvre triangulær. Dette viser vores påstand, og sætning 6.4.1 viser, at denne medfører (2).

Vi antager dernæst (2) og viser (1). Så lad $Q \in M_n(\mathbb{R})$ være en ortogonal matrix, sådan at $Q^{-1}AQ = D$ er en diagonal matrix. Da D er symmetrisk, så viser udregningen

$$A^* = (QDQ^{-1})^* = (QDQ^*)^* = QD^*Q^* = QDQ^{-1} = A,$$

at også A er symmetrisk. Dette viser (1) og dermed sætningen. $\square$

Den næste sætning kaldes for spektralsætningen for normale komplekse matricer.

Sætning 6.4.3 *For en kompleks kvadratisk matrix $A \in M_n(\mathbb{C})$ er følgende ækvivalent:*

(1) *Matricen A er normal.*

(2) *Der findes en unitær matrix $U \in M_n(\mathbb{C})$, sådan at $U^{-1}AU$ er en diagonal matrix.*

Bevis Beviset for, at (1) medfører (2) er helt som beviset for den tilsvarende implikation i sætning 6.4.2. Den eneste forskel er, at “ $\mathbb{R}$ ” og “ortogonal” erstattes med henholdsvis “ $\mathbb{C}$ ” og “unitær”. Vi antager derfor (2) og viser (1). Så lad $U \in M_n(\mathbb{C})$ være en unitær matrix, sådan at $U^{-1}AU = D$ er en diagonal matrix. Da er $A = UDU^{-1} = UDU^*$, og da diagonal matrixen D er normal, så viser udregningen

$$\begin{aligned} A^*A &= (UDU^*)^*UDU^* = UD^*U^*UDU^* = UD^*DU^*, \\ AA^* &= UDU^*(UDU^*)^* = UDU^*UD^*U^* = UDD^*U^* \end{aligned}$$

at også A er normal. Dette viser (1) og dermed sætningen. $\square$

Korollar 6.4.4 *En normal kompleks matrix $A \in M_n(\mathbb{C})$ er hermitisk, hvis og kun hvis dens egenverdier er reelle.*

Bevis Da A er normal, så findes der ifølge sætning 6.4.3 en unitær matrix $U \in M_n(\mathbb{C})$, sådan at $U^*AU = \text{diag}(\lambda_1, \dots, \lambda_n) \in M_n(\mathbb{C})$ er en diagonal matrix D . Hvis egenverdierne $\lambda_1, \dots, \lambda_n$ alle er reelle, da er $D^* = D$ hermitisk, og udregningen

$$A^* = (UDU^*)^* = UD^*U^* = UDU^* = A$$

viser da, at A er hermitisk. Omvendt, hvis $A = A^*$ er hermitisk, så viser udregningen

$$D^* = (U^*AU)^* = U^*A^*U = U^*AU = D,$$

at D er hermitisk, og dermed er egenverdierne $\lambda_1, \dots, \lambda_n$ alle reelle. $\square$

6 Vektorrum med indre produkt

Vi betragter konkrete eksempler nedenfor, men oversætter først sætning 6.4.2 og sætning 6.4.3 til udsagn om lineære endomorfier af henholdsvis reelle og komplekse vektorrum med indre produkt.

Sætning 6.4.5 *Lad $(V, \langle -, - \rangle)$ være et endeligt frembragt reelt vektorrum med indre produkt. For en lineær endomorfi $f: V \rightarrow V$ er følgende ækvivalent:*

- (1) *Endomorfien $f: V \rightarrow V$ er selvadjungeret.*
- (2) *Der findes en basis for V , der både er ortonormal med hensyn til $\langle -, - \rangle$ og består af egenvektorer for $f: V \rightarrow V$.*

Bevis Vi antager først (1) og viser (2). Vi vælger derfor en basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V , som er ortonormal med hensyn til $\langle -, - \rangle$, hvilket er muligt ifølge korollar 6.2.11. Da den lineære endomorfi $f: V \rightarrow V$ er selvadjungeret med hensyn til $\langle -, - \rangle$, viser sætning 6.3.13, at matricen $A \in M_n(\mathbb{R})$, der repræsenterer $f: V \rightarrow V$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for både domænet og codomænet, er symmetrisk. Derfor findes der ifølge sætning 6.4.2 en ortogonal matrix $Q = (q_{ij}) \in M_n(\mathbb{R})$, sådan at

$$Q^{-1}AQ = \text{diag}(a_1, \dots, a_n)$$

er en diagonalmatrix. Da $Q \in M_n(\mathbb{R})$ er ortogonal, er familien $(\mathbf{u}_1, \dots, \mathbf{u}_n)$, hvor

$$\mathbf{u}_j = \mathbf{v}_1 q_{1j} + \mathbf{v}_2 q_{2j} + \dots + \mathbf{v}_n q_{nj},$$

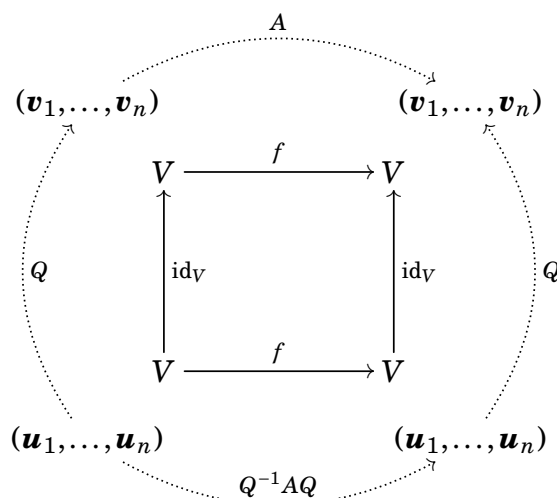
en basis for V , der er ortonormal med hensyn til $\langle -, - \rangle$. For

$$\begin{aligned} \langle \mathbf{u}_i, \mathbf{u}_j \rangle &= \langle \mathbf{v}_1 q_{1i} + \mathbf{v}_2 q_{2i} + \dots + \mathbf{v}_n q_{ni}, \mathbf{v}_1 q_{1j} + \mathbf{v}_2 q_{2j} + \dots + \mathbf{v}_n q_{nj} \rangle \\ &= q_{1i} q_{1j} + q_{2i} q_{2j} + \dots + q_{ni} q_{nj} = \mathbf{q}_i^* \mathbf{q}_j = \delta_{ij}. \end{aligned}$$

Desuden repræsenterer matricen Q identitetsafbildningen $\text{id}_V: V \rightarrow V$ med hensyn til basen $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ for domænet og basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for codomænet. Som illustreret af figuren på næste side, så viser sætning 4.4.14, at matricen, der repræsenterer $f: V \rightarrow V$ med hensyn til basen $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ for både domænet og codomænet, er diagonalmatricen

$$Q^{-1}AQ = \text{diag}(\lambda_1, \dots, \lambda_n) \in M_n(\mathbb{R}).$$

Så $\mathbf{u}_j$ er en egenvektor for $f: V \rightarrow V$ hørende til egenværdien λ_j , hvilket viser (2).



Vi antager omvendt (2) og viser (1). Så vi lader $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ være en basis for V , der er ortonormal med hensyn til $\langle -, - \rangle$ og består af egenvektorer for $f: V \rightarrow V$. Lad $\lambda_j \in \mathbb{R}$ være egenværdien, sådan at $f(\mathbf{u}_j) = \mathbf{u}_j \lambda_j$. Da er $D = \text{diag}(\lambda_1, \dots, \lambda_n) \in M_n(\mathbb{R})$ matricen, der repræsenterer $f: V \rightarrow V$ med hensyn til den valgte basis for både domænet og codomænet. Da D er symmetrisk, og da $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ er ortonormal med hensyn til $\langle -, - \rangle$, så viser sætning 6.3.7, at $f: V \rightarrow V$ er selvadjungeret med hensyn til $\langle -, - \rangle$. Dette viser (1) og dermed sætningen. $\square$

Sætning 6.4.6 *Lad $(V, \langle -, - \rangle)$ være et endeligt frembragt komplekst vektorrum med indre produkt. For en lineær endomorfi $f: V \rightarrow V$ er følgende ækvivalent:*

- (1) *Endomorfin $f: V \rightarrow V$ er normal.*
- (2) *Der findes en basis for V , der både er ortonormal med hensyn til $\langle -, - \rangle$ og består af egenvektorer for $f: V \rightarrow V$.*

Bevis Beviset ligner helt beviset for sætning 6.4.5, idet vi anvender sætning 6.4.3 i stedet for sætning 6.4.2 og erstatter “ $\mathbb{R}$ ” med “ $\mathbb{C}$ ”, “symmetrisk” og “selvadjungeret” med “normal”, og “ortogonal” med “unitær”. $\square$

Vi bemærker, at spektralsætningerne og deres beviser kun siger, at en ortonormal basis af egenvektorer findes, men de siger ikke, hvordan man bærer sig ad med at finde en sådan basis. Problemet er at finde rødder i det karakteristiske polynomium. Hvis det er gjort, så findes der en algoritme, som vi illustrerer ved nogle eksempler på de næste sider, der angiver den ønskede basis. Undervejs kan man spare sig en del arbejde ved at bemærke følgende.

6 Vektorrum med indre produkt

Korollar 6.4.7 Lad $A \in M_n(\mathbb{F})$ være enten en symmetrisk reel matrix eller en normal kompleks matrix. Egenvektorer $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, der hører til forskellige egenverdier $\lambda, \mu \in \mathbb{F}$, er nødvendigvis ortogonale med hensyn til standard-indreproduktet på $\mathbb{F}^n$.

Bevis Vi viser det komplekse tilfælde; det reelle tilfælde er helt tilsvarende. Vi lader $\lambda, \mu \in \mathbb{F}$ være to forskellige egenverdier for A og angiver baser for egenrummene $N_{A-I\lambda}$ og $N_{A-I\mu}$ som følger. Ifølge sætning 6.4.3 findes der en unitær matrix $U \in M_n(\mathbb{C})$, sådan at $U^*AU = \text{diag}(\lambda_1, \dots, \lambda_n)$ er en diagonal matrix. Så hvis $I = \{i \in \{1, \dots, n\} \mid \lambda_i = \lambda\}$ og $J = \{j \in \{1, \dots, n\} \mid \lambda_j = \mu\}$, da er $(U\mathbf{e}_i \mid i \in I)$ og $(U\mathbf{e}_j \mid j \in J)$ baser for henholdsvis $N_{A-I\lambda}$ og $N_{A-I\mu}$. Hvis nu $\mathbf{x} = \sum_{i \in I} U\mathbf{e}_i a_i \in N_{A-I\lambda}$ og $\mathbf{y} = \sum_{j \in J} U\mathbf{e}_j b_j \in N_{A-I\mu}$, så er

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i \in I, j \in J} a_i^* \langle U\mathbf{e}_i, U\mathbf{e}_j \rangle b_j = \sum_{i \in I, j \in J} a_i^* \langle \mathbf{e}_i, \mathbf{e}_j \rangle b_j = 0$$

som ønsket, idet $I \cap J = \emptyset$. Se også sætning 6.3.13. $\square$

Eksempel 6.4.8 Vi ønsker at finde en basis for $\mathbb{R}^3$, der er ortonormal med hensyn til standard-indreproduktet, og som endvidere består af egenvektorer for den lineære endomorfi $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ givet ved $f(\mathbf{x}) = A\mathbf{x}$, hvor $A \in M_3(\mathbb{R})$ er den symmetriske matrix

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -2 & 2 \\ -1 & 2 & 1 \end{pmatrix}.$$

Vi ved fra sætning 6.4.2, at dette er muligt. For at bestemme egenverdierne af A udregner vi først det karakteristiske polynomium.

$$\begin{aligned} \chi_A(t) &= \det(A - It) = \det \begin{pmatrix} 1-t & 2 & -1 \\ 2 & -2-t & 2 \\ -1 & 2 & 1-t \end{pmatrix} \\ &\stackrel{(D6')}{=} \det \begin{pmatrix} 0 & 2(2-t) & -t(2-t) \\ 0 & 2-t & 2(2-t) \\ -1 & 2 & 1-t \end{pmatrix} \stackrel{(D2')}{=} (2-t)^2 \det \begin{pmatrix} 0 & 2 & -t \\ 0 & 1 & 2 \\ -1 & 2 & 1-t \end{pmatrix} \\ &\stackrel{(D6')}{=} (2-t)^2 \det \begin{pmatrix} 0 & 0 & -4-t \\ 0 & 1 & 2 \\ -1 & 2 & 1-t \end{pmatrix} \stackrel{(D2')}{=} (2-t)^2(-4-t) \det \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & 2 & 1-t \end{pmatrix} \\ &\stackrel{(D5')}{=} -(2-t)^2(-4-t) \det \begin{pmatrix} -1 & 2 & 1-t \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = (2-t)^2(-4-t) \end{aligned}$$

Egenverdierne for A er altså $\lambda_1 = 2$ og $\lambda_2 = -4$. Her har vi benyttet sætning 3.3.3 til at udregne determinanten af den triangulære matrix.

Vi anvender dernæst Gauss-elimination til at finde baser for egenrummene hørende til $\lambda_1 = 2$ og $\lambda_2 = -4$. Så vi omdanner matricen $B = A - I \cdot 2$ til en rækkeækvivalent matrix B' på reduceret echelonform:

$$\begin{aligned}
 B &= \begin{pmatrix} -1 & 2 & -1 \\ \textcolor{red}{2} & -4 & 2 \\ \textcolor{red}{-1} & 2 & -1 \end{pmatrix} \begin{array}{l} \\ +2 \times R_1 \\ +(-1) \times R_1 \end{array} \\
 &\quad \begin{pmatrix} \textcolor{red}{-1} & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{array}{l} (-1) \times R_1 \\ \\ \end{array} \\
 B' &= \begin{pmatrix} \textcolor{blue}{1} & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

Egenrummet $N_{A-I \cdot 2}$ hørende til $\lambda_1 = 2$, som er løsningsmængden til ligningssystemet $B\mathbf{x} = \mathbf{0}$, kan derfor aflæses fra den rækkeækvivalente matrix B' til at have basis

$$(\mathbf{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}).$$

Tilsvarende omdanner vi matricen $C = A + I \cdot 4$ til en rækkeækvivalent matrix C' på reduceret echelonform:

$$\begin{aligned}
 C &= \begin{pmatrix} \textcolor{red}{5} & 2 & -1 \\ \textcolor{red}{2} & 2 & 2 \\ -1 & 2 & 5 \end{pmatrix} \begin{array}{l} +5 \times R_3 \\ +2 \times R_3 \\ \end{array} \\
 &\quad \begin{pmatrix} 0 & \textcolor{red}{12} & 24 \\ 0 & \textcolor{red}{6} & 12 \\ \textcolor{red}{-1} & 2 & 5 \end{pmatrix} \begin{array}{l} (1/12) \times R_1 \\ (1/6) \times R_2 \\ (-1) \times R_3 \end{array} \\
 &\quad \begin{pmatrix} 0 & \textcolor{red}{1} & 2 \\ 0 & 1 & 2 \\ 1 & \textcolor{red}{-2} & -5 \end{pmatrix} \begin{array}{l} +(-1) \times R_2 \\ \\ +2 \times R_2 \end{array} \\
 &\quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & -1 \end{pmatrix} R_1 \leftrightarrow R_3 \\
 C' &= \begin{pmatrix} \textcolor{blue}{1} & 0 & -1 \\ 0 & \textcolor{blue}{1} & 2 \\ 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

6 Vektorrum med indre produkt

Egenrummet $N_{A+I \cdot 4}$ hørende til $\lambda_2 = -4$, som er løsningsmængden til ligningssystemet $C\mathbf{x} = \mathbf{0}$, kan derfor aflæses fra C' til at have basis

$$(\mathbf{w}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix})$$

idet x_3 er eneste frie variabel.

Vi har nu fundet en basis $(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$ for $\mathbb{R}^3$, bestående af egenvektorer for A . Denne basis er dog ikke ortonormal med hensyn til standard-indreproduktet, og vi anvender derfor Gram-Schmidt-algoritmen til at omdanne den til en ortonormal basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$. Vi ved dog allerede fra korollar 6.4.7, at $\mathbf{w}_3$ er ortogonal til både $\mathbf{w}_1$ og $\mathbf{w}_2$, idet de hører til forskellige egenverdier for den symmetriske matrix A . Vi finder først den ortogonale basis $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, hvor

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \\ \mathbf{v}_2 &= \mathbf{w}_2 - \mathbf{v}_1 \cdot \frac{\langle \mathbf{v}_1, \mathbf{w}_2 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \cdot \frac{-2+0+0}{4+1+0} = \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix} \cdot \frac{1}{5} \\ \mathbf{v}_3 &= \mathbf{w}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}. \end{aligned}$$

Endelig får vi ved normalisering af vektorerne i denne basis den ortonormale basis

$$(\mathbf{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -1/\sqrt{30} \\ 2/\sqrt{30} \\ 5/\sqrt{30} \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}),$$

der består af egenvektorer for A . Den ortogonale matrix $Q \in M_3(\mathbb{R})$, der repræsenterer identitetsafbildningen $\text{id}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ med hensyn til den nye basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for codomænet, er da givet ved

$$Q = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{30} & 1/\sqrt{6} \\ 1/\sqrt{5} & 2/\sqrt{30} & -2/\sqrt{6} \\ 0 & 5/\sqrt{30} & 1/\sqrt{6} \end{pmatrix} \in M_3(\mathbb{R}),$$

og ifølge sætning 4.4.14 gælder det endvidere, at

$$Q^{-1}AQ = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix} \in M_3(\mathbb{R}),$$

idet $\mathbf{u}_1$, $\mathbf{u}_2$ og $\mathbf{u}_3$ er egenvektorer for $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ hørende til egenverdierne 2, 2 og -4.

Eksempel 6.4.9 Vi ønsker at finde en basis for $\mathbb{C}^2$, der er ortonormal med hensyn til standard-indreproduktet og består af egenvektorer for den lineære endomorfi $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ givet ved $f(\mathbf{x}) = A\mathbf{x}$, hvor $A \in M_2(\mathbb{C})$ er den normale matrix

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

fra eksempel 6.3.15. Det karakteristiske polynomium er givet ved

$$\chi_A(t) = \det(A - It) = \det \begin{pmatrix} 2-t & -1 \\ 1 & 2-t \end{pmatrix} = (2-t)^2 + 1 = t^2 - 4t + 5,$$

og det har diskriminant $\Delta = (-4)^2 - 4 \cdot 1 \cdot 5 = -4$ og rødder

$$\lambda_1 = \frac{-(-4) + \sqrt{-4}}{2 \cdot 1} = 2 + i \quad \text{og} \quad \lambda_2 = \frac{-(-4) - \sqrt{-4}}{2 \cdot 1} = 2 - i.$$

Vi anvender nu Gauss-elimination til at bestemme baser for de to egenrum, der begge nødvendigvis er en-dimensionale. Vi omdanner først $B = A - I\lambda_1$ til en rækkeækvivalent matrix B' på reduceret echelonform.

$$\begin{aligned} B &= \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} + i \times R_2 \\ &\quad \begin{pmatrix} 0 & 0 \\ 1 & -i \end{pmatrix} R_1 \leftrightarrow R_2 \\ B' &= \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Herved finder vi basen

$$(\mathbf{v}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix})$$

for egenrummet $N_{A-I\lambda_1}$ hørende til $\lambda_1 = 2 + i$. Vi omdanner ligeledes $C = A - I\lambda_2$ til en matrix C' på reduceret echelonform.

$$\begin{aligned} C &= \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} + (-i) \times R_2 \\ &\quad \begin{pmatrix} 0 & 0 \\ 1 & i \end{pmatrix} R_1 \leftrightarrow R_2 \\ C' &= \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Herved finder vi basen

$$(\mathbf{v}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix})$$

6 Vektorrum med indre produkt

for egenrummet $N_{A-I\lambda_2}$ hørende til $\lambda_2 = 2 - i$. Basen $(\mathbf{v}_1, \mathbf{v}_2)$ for $\mathbb{C}^2$ vides allerede at være ortogonal, hvilket også følger af korollar 6.4.7, så vi behøver ikke at anvende Gram-Schmidt. Endelig får vi ved normalisering direkte den ortonormale basis

$$(\mathbf{u}_1 = \begin{pmatrix} -i/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} i/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}),$$

for $\mathbb{C}^2$, der består af egenvektorer for $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$. Den unitære matrix $U \in M_2(\mathbb{C})$, der repræsenterer identitetsafbildningen $\text{id}: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ med hensyn til den nye basis $(\mathbf{u}_1, \mathbf{u}_2)$ for domænet og standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for codomænet, er da givet ved

$$U = \begin{pmatrix} i/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \in M_2(\mathbb{C}),$$

og sætning 4.4.14 viser derfor, at

$$U^{-1}AU = \begin{pmatrix} 2+i & 0 \\ 0 & 2-i \end{pmatrix} \in M_2(\mathbb{C}),$$

idet $\mathbf{u}_1$ og $\mathbf{u}_2$ er egenvektorer for $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ med egenverdier henholdsvis $2+i$ og $2-i$.

6.5 Klassifikation af hermitiske former

Vi lader fortsat $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$ og lader $(-)^*: \mathbb{F} \rightarrow \mathbb{F}$ være skævinvolutionen givet ved

$$a^* = \begin{cases} a & \text{hvis } \mathbb{F} = \mathbb{R}, \\ \bar{a} & \text{hvis } \mathbb{F} = \mathbb{C}. \end{cases}$$

Vi betragter hermitiske former $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$ på $\mathbb{F}$ -vektorrum V og bemærker, at valget af skævinvolution medfører, at $\langle \mathbf{v}, \mathbf{v} \rangle$ er et reelt tal for alle $\mathbf{v} \in V$. Dermed er den følgende definition meningsfuld.

Definition 6.5.1 Lad $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$ være en hermitisk form på et $\mathbb{F}$ -vektorrum V .

- (1) Et underrum $W \subset V$ er *positiv definit med hensyn til* $\langle -, - \rangle$, hvis $\langle \mathbf{w}, \mathbf{w} \rangle > 0$ for alle $\mathbf{0} \neq \mathbf{w} \in W$.
- (2) Et underrum $W \subset V$ er *negativ definit med hensyn til* $\langle -, - \rangle$, hvis $\langle \mathbf{w}, \mathbf{w} \rangle < 0$ for alle $\mathbf{0} \neq \mathbf{w} \in W$.

Eksempel 6.5.2 (1) En hermitisk form $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$ er et indre produkt, hvis og kun hvis ethvert underrum $W \subset V$ er positiv definit med hensyn til $\langle -, - \rangle$, hvis og kun hvis V er positiv definit med hensyn til $\langle -, - \rangle$.

(2) I tilfældet $V = \mathbb{R}^n$ med Minkowskis hermitiske form

$$\langle \mathbf{x}, \mathbf{y} \rangle = -x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

siger vi også, at et underrum $W \subset V$, der er positiv definit med hensyn til $\langle -, - \rangle$, er “spacelike”, og at et underrum $W \subset V$, der er negativ definit med hensyn til $\langle -, - \rangle$, er “timelike”. Det en-dimensionale underrum frembragt af $(\mathbf{e}_1a_1 + \cdots + \mathbf{e}_na_n)$ er således spacelike, hvis $a_1^2 < a_2^2 + \cdots + a_n^2$, mens det er timelike, hvis $a_1^2 > a_2^2 + \cdots + a_n^2$.

Lad n være et naturligt tal, og lad p, q og r være naturlige tal, sådan at $p + q + r = n$. Vi definerer da $D(p, q, r) \in M_n(\mathbb{F})$ til at være diagonalmatricen

$$D(p, q, r) = \text{diag}(-1, \dots, -1, 0, \dots, 0, 1, \dots, 1),$$

hvor de første p diagonalindgange er lig med -1 ; de næste q diagonalindgange er lig med 0 ; og de sidste r diagonalindgange er lig med 1 . Vi skal nu vise følgende sætning, der kaldes Sylvesters inertisætning.

Sætning 6.5.3 (Sylvester) *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$ være en hermitisk form på et n -dimensionalt $\mathbb{F}$ -vektorrum.*

(1) *Der findes en basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ for V og naturlige tal p, q og r , sådan at det for alle $\mathbf{v}, \mathbf{w} \in V$ med koordinater $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ gælder, at*

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{x}^* D(p, q, r) \mathbf{y}.$$

(2) *De naturlige tal p, q og r afhænger ikke af valget af basis og er karakteriseret som følger: p er den maksimale dimension af et underrum $W \subset V$, der er negativ definit med hensyn til $\langle -, - \rangle$; r er den maksimale dimension af et underrum $W \subset V$, der er positiv definit med hensyn til $\langle -, - \rangle$; og $q = n - p - r$.*

Bevis Vi viser først (1) for $\mathbb{F} = \mathbb{R}$. Vi vælger en vilkårlig basis $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ for V og lader $h: \mathbb{R}^n \rightarrow V$ være den entydigt bestemte isomorfi, sådan at $h(\mathbf{e}_i) = \mathbf{w}_i$ for alle $1 \leq i \leq n$. Da $h: \mathbb{R}^n \rightarrow V$ er lineær, så er $\langle h(-), h(-) \rangle$ en hermitisk form på $\mathbb{R}^n$, så matricen

$$A = (\langle h(\mathbf{e}_i), h(\mathbf{e}_j) \rangle) = (\langle \mathbf{w}_i, \mathbf{w}_j \rangle) \in M_n(\mathbb{R})$$

6 Vektorrum med indre produkt

er symmetrisk ifølge sætning 2.6.10. Hvis $\mathbf{v}, \mathbf{w} \in V$ har koordinater $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ med hensyn til den valgte basis, da gælder det endvidere, at

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle h(\mathbf{x}), h(\mathbf{y}) \rangle = \mathbf{x}^* A \mathbf{y}.$$

Sætning 6.4.2 viser nu, at der findes en ortogonal matrix $Q \in M_n(\mathbb{R})$, sådan at

$$Q^* A Q = \text{diag}(\lambda_1, \dots, \lambda_n) \in M_n(\mathbb{R})$$

er en diagonalmatrix. Vi påstår, at vi kan vælge $Q \in M_n(\mathbb{R})$, sådan at

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n.$$

For ellers kan vi vælge en permutation af n bogstaver $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$, sådan at $\lambda_{\sigma(1)} \leq \lambda_{\sigma(2)} \leq \dots \leq \lambda_{\sigma(n)}$. Permutationsmatricen $P(\sigma) \in M_n(\mathbb{F})$ er da ortogonal, idet dens søjler er standardenhedsvektorer, og

$$P(\sigma)^* Q^* A Q P(\sigma) = \text{diag}(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)}) \in M_n(\mathbb{R}).$$

Derfor kan og vil vi antage, at $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, idet vi ellers erstatter Q med $Q P(\sigma)$.

Vi lader nu p, q og r være de naturlige tal, sådan at $\lambda_i < 0$ for $1 \leq i \leq p$; $\lambda_i = 0$ for $p+1 \leq i \leq p+q$; og $\lambda_i > 0$ for $p+q+1 \leq i \leq n$. Vi lader dernæst $R = \text{diag}(\rho_1, \dots, \rho_n) \in M_n(\mathbb{R})$ være diagonalmatricen med diagonalindgange

$$\rho_i = \begin{cases} |\lambda_i|^{-\frac{1}{2}} & \text{hvis } \lambda_i \neq 0, \\ 1 & \text{hvis } \lambda_i = 0. \end{cases}$$

Matricen R er symmetrisk, og den er valgt, sådan at

$$P^* A P = R^* Q^* A Q R = D(p, q, r),$$

hvor vi skriver $P = (p_{ij}) = Q R \in M_n(\mathbb{R})$. Endelig definerer vi

$$\mathbf{v}_j = \mathbf{w}_1 p_{1j} + \mathbf{w}_2 p_{2j} + \dots + \mathbf{w}_n p_{nj}$$

for alle $1 \leq j \leq n$. Da er $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ en basis for V , og hvis $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ er koordinaterne for $\mathbf{v}, \mathbf{w} \in V$ med hensyn til basen $(\mathbf{v}_1, \dots, \mathbf{v}_n)$, så er $P\mathbf{x}, P\mathbf{y} \in \mathbb{F}^n$ koordinaterne for $\mathbf{v} \in V$ med hensyn til basen $(\mathbf{w}_1, \dots, \mathbf{w}_n)$. Derfor er

$$\langle \mathbf{v}, \mathbf{w} \rangle = (P\mathbf{x})^* A (P\mathbf{y}) = \mathbf{x}^* P^* A P \mathbf{y} = \mathbf{x}^* D(p, q, r) \mathbf{y}$$

som ønsket. Dette viser (1) for $\mathbb{F} = \mathbb{R}$; beviset for $\mathbb{F} = \mathbb{C}$ er helt analogt, idet man blot erstatter “symmetrisk” med “hermitisk”, “ortogonal” med “unitær” og sætning 6.4.2 med sætning 6.4.3.

Vi mangler at vise (2). Vi betragter igen $\mathbb{F} = \mathbb{R}$; tilfældet $\mathbb{F} = \mathbb{C}$ er helt analogt. Det er ikke klart, at de naturlige tal p, q og r i beviset for (1) er uafhængige af vores valg af

den ortogonal matrix $Q \in M_n(\mathbb{R})$. Vi træffer nu et valg af en sådan matrix $Q \in M_n(\mathbb{R})$ og finder derefter en basis $(v_1, \dots, v_n)$ for V som i beviset for (1). Vi lader $V_{<0} \subset V$ være underrummet frembragt af $(v_1, \dots, v_p)$ og lader $V_{\geq 0} \subset V$ være underrummet frembragt af $(v_{p+1}, \dots, v_n)$. Underrummet $V_{<0} \subset V$ har dimension p og er negativ definit med hensyn til $\langle -, - \rangle$. Vi ønsker at vise, at hvis $W \subset V$ er et underrum, der er negativ definit med hensyn til $\langle -, - \rangle$, så er dimensionen af W højst p . Men $W \cap V_{\geq 0} = \{0\}$, og derfor er

$$\dim(W) + \dim(V_{\geq 0}) \leq \dim(V),$$

hvilket viser, at $\dim(W) \leq p$ som ønsket. Hermed har vi vist, at p er den maksimale dimension af et underrum $W \subset V$, der er negativ definit med hensyn til $\langle -, - \rangle$. Beviset for, at r er den maksimale dimension af et underrum $W \subset V$, der er positiv definit med hensyn til $\langle -, - \rangle$ er helt tilsvarende. Dette afslutter beviset. $\square$

Bemærkning 6.5.4 Lad $(V, \langle -, - \rangle)$ være som i sætning 6.5.3. Vi understreger, at et maksimalt underrum $W \subset V$, der er enten negativ definit eller positiv definit med hensyn til den hermiteske form $\langle -, - \rangle$, *ikke* er entydigt. Det er kun de respektive dimensioner p og r af sådanne underrum, der er entydigt bestemte.

Eksempel 6.5.5 Vi betragter igen Minkowskiformen på $\mathbb{R}^n$ fra eksempel 6.5.2, hvor vi antager at $n \geq 2$. Med notation som i sætning 6.5.3, er denne givet ved

$$\langle x, y \rangle = x^* D(p, q, r) y,$$

hvor $p = 1$, $q = 0$ og $r = n - 1$. Sætning 6.5.3 viser derfor, at et timelike underrum højst kan være en-dimensionalt. Vi har allerede set i eksempel 6.5.2, at det en-dimensionale underrum frembragt af $(e_1 a_1 + \dots + e_n a_n)$ er timelike, hvis $a_1^2 > a_2^2 + \dots + a_n^2$. Så der er altså uendeligt mange forskellige maksimale timelike underrum. Vi bemærker også, at selvom $q = 0$, så findes der en-dimensionale underrum $U \subset \mathbb{R}^n$, sådan at det for alle $x \in U$ gælder, at $\langle x, x \rangle = 0$. Delmængden $C = \{x \in \mathbb{R}^n \mid \langle x, x \rangle = 0\} \subset \mathbb{R}^n$ er en tosidet kegle med keglepunkt 0 , som kaldes “lyskeglen” og som udgør grænsen mellem de “spacelike” og “timelike” dele af $\mathbb{R}^n$.

6.6 Opgaver

6.1 Vi lader $V = \mathbb{C}^2$ og definerer afbildningen $\langle -, - \rangle': V \times V \rightarrow \mathbb{C}$ ved

$$\langle x, y \rangle' = 3\bar{x}_1 y_1 + 4\bar{x}_2 y_2.$$

a) Vis, at $\langle -, - \rangle': V \times V \rightarrow \mathbb{C}$ er et indre produkt.

6 Vektorrum med indre produkt

b) Find normen af vektorerne

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} -4 \\ 3 \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} \sqrt{3} \\ 2i \end{pmatrix}$$

med hensyn til $\langle -, - \rangle'$.

c) Afgør, hvorvidt $\mathbf{x}$ og $\mathbf{y}$ er ortogonale med hensyn til $\langle -, - \rangle'$.

d) Afgør, hvorvidt $\mathbf{x}$ og $\mathbf{y}$ er ortogonale med hensyn til standard-indreproduktet $\langle -, - \rangle: V \times V \rightarrow \mathbb{C}$.

6.2 Vi betragter $\mathbb{R}$ -vektorrummet $V = C^0([0, 1])$ af kontinuerte funktioner $f: [0, 1] \rightarrow \mathbb{R}$ med det indre produkt $\langle -, - \rangle: V \times V \rightarrow \mathbb{R}$ defineret ved

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx.$$

a) Eftersis, at $\langle -, - \rangle$ er et indre produkt i V .

[Det følgende resultat kan anvendes uden bevis: Given en kontinuert afbildning $h: [0, 1] \rightarrow \mathbb{R}$, sådan at $h(x) \geq 0$ for alle $x \in [0, 1]$, da medfører $\int_0^1 h(x)dx = 0$, at $h(x) = 0$ for alle $x \in [0, 1]$.]

b) Vis, at de kontinuerte funktioner $f, g: [0, 1] \rightarrow \mathbb{R}$, der er defineret ved $f(x) = 1$ og $g(x) = \sqrt{3}(2x - 1)$, er ortogonale enhedsvektorer med hensyn til $\langle -, - \rangle$.

c) Bestem normen af den kontinuerte funktion $h: [0, 1] \rightarrow \mathbb{R}$, der er defineret ved $h(x) = x^n$, hvor $n \geq 0$ er et helt tal, med hensyn til $\langle -, - \rangle$.

6.3 Vis, at det for reelle tal $x_1, \dots, x_n$ gælder, at

$$(x_1 + \dots + x_n)^2 \leq n(x_1^2 + \dots + x_n^2).$$

(Vink: Anvend Cauchy-Schwarz' ulighed).

6.4 Find vinklen mellem vektorerne $\mathbf{x}, \mathbf{y} \in \mathbb{R}^4$ givet ved

$$\mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 1 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

med hensyn til standard-indreproduktet på $\mathbb{R}^4$.

6.5 Lad $U \subset \mathbb{R}^4$ være underrummet frembragt af den lineært uafhængige familie $(\mathbf{w}_1, \mathbf{w}_2)$ af vektorer i $\mathbb{R}^4$, hvor

$$\mathbf{w}_1 = \begin{pmatrix} 5 \\ 3 \\ 1 \\ 1 \end{pmatrix} \quad \text{og} \quad \mathbf{w}_2 = \begin{pmatrix} -1 \\ -3 \\ 3 \\ -1 \end{pmatrix}.$$

- a) Omdan basen $(\mathbf{w}_1, \mathbf{w}_2)$ for U til en basis $(\mathbf{v}_1, \mathbf{v}_2)$, der er ortogonal med hensyn til standard-indreproduktet på $\mathbb{R}^4$.
- b) Omdan endvidere basen $(\mathbf{v}_1, \mathbf{v}_2)$ for U til en basis $(\mathbf{u}_1, \mathbf{u}_2)$, der er ortonormal med hensyn til standard-indreproduktet på $\mathbb{R}^4$.

6.6 Vi betragter de følgende vektorer i $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

- a) Vis, at $(\mathbf{v}_1, \mathbf{v}_2)$ er en ortogonal familie med hensyn til standard-indreproduktet på $\mathbb{R}^3$ og konkluder, at den er lineært uafhængig.
- b) Find en vektor $\mathbf{w}_3$, sådan at $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_3)$ er en basis for $\mathbb{R}^3$.
- c) Anvend Gram-Schmidt-ortogonalisering til at omdanne basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_3)$ til en basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, der er ortonormal med hensyn til standard-indreproduktet.
- d) Vis, at matricen $P = (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3)$, er ortogonal.
- e) Bestem koordinaterne af standardenhedsvektorerne $\mathbf{e}_1$, $\mathbf{e}_2$ og $\mathbf{e}_3$ med hensyn til basen $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$.

6.7 Vi betragter de følgende vektorer i $\mathbb{R}^4$:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

- a) Eftersis, at familien $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ er ortogonal med hensyn til standard-indreproduktet $\langle -, - \rangle$ på $\mathbb{R}^4$ og konkluder, at den udgør en basis for $\mathbb{R}^4$.
- b) Vis, at afbildningen $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ defineret ved

$$f(\mathbf{x}) = \mathbf{v}_2 \langle \mathbf{v}_1, \mathbf{x} \rangle + \mathbf{v}_3 \langle \mathbf{v}_2, \mathbf{x} \rangle + \mathbf{v}_4 \langle \mathbf{v}_3, \mathbf{x} \rangle$$

er lineær.

- c) Bestem matricen A , der repræsenterer $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ med hensyn til standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ for både domænet og codomænet, samt matricen B , der repræsenterer $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ med hensyn til basen $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ for både domænet og codomænet.
- d) Beregn matricerne A^2 , A^3 , A^4 samt B^2 , B^3 , B^4 , gerne ved brug af Maple.
- e) Bestem dimensionen af billedet af $f^{\circ n}: \mathbb{R}^4 \rightarrow \mathbb{R}^4$, hvor

$$f^{\circ 1} = f, \quad f^{\circ 2} = f \circ f, \quad f^{\circ 3} = f \circ f \circ f, \quad \dots,$$

for alle $n \geq 1$.

6 Vektorrum med indre produkt

6.8 Lad V være et vektorrum af dimension n , og lad $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ være en lineært uafhængig familie af vektorer i V . Angiv med “ja” eller “nej” om følgende udsagn er korrekte eller ej.

- Hvis $f: V \rightarrow V$ er en lineær endomorfi af et endeligt frembragt komplekst vektorrum, så findes der en egenvektor for $f: V \rightarrow V$.
- Hvis $f: V \rightarrow V$ er en lineær endomorfi af et endeligt frembragt komplekst vektorrum, så har V en basis, der består af egenvektorer for $f: V \rightarrow V$.
- Ethvert endeligt frembragt reelt vektorrum med indre produkt $(V, \langle -, - \rangle)$ har en basis, der er ortogonal med hensyn til $\langle -, - \rangle$.
- Ethvert endeligt frembragt reelt vektorrum med indre produkt $(V, \langle -, - \rangle)$ har en basis, der er ortonormal med hensyn til $\langle -, - \rangle$.
- Lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være to reelle vektorrum med indre produkt, og lad $f: V \rightarrow W$ være en lineær afbildning. Hvis vektorerne $\mathbf{v}, \mathbf{v}' \in V$ er ortogonale med hensyn til $\langle -, - \rangle_V$, så er vektorerne $f(\mathbf{v}), f(\mathbf{v}') \in W$ ortogonale med hensyn til $\langle -, - \rangle_W$.

6.9 Lad $Q \in M_3(\mathbb{R})$ være matricen

$$Q = \begin{pmatrix} 2/3 & 2/3 & -1/3 \\ -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \end{pmatrix} \in M_3(\mathbb{R}).$$

- Eftervis, at Q er ortogonal.
- Find Q^{-1} .

6.10 Eftervis, at de tre Pauli-matricer

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

både er hermitiske og unitære komplekse matricer. Vis dernæst, at hver af de tre Pauli-matricer har egenverdierne $+1$ og -1 .

6.11 (★) Betragt triplen $(M_2(\mathbb{C}), +, \cdot)$, hvor $+: M_2(\mathbb{C}) \times M_2(\mathbb{C}) \rightarrow M_2(\mathbb{C})$ er matrixsum, og hvor $\cdot: M_2(\mathbb{C}) \times \mathbb{R} \rightarrow M_2(\mathbb{C})$ er givet ved

$$\begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} \cdot a = \begin{pmatrix} z_1 a & z_2 a \\ z_3 a & z_4 a \end{pmatrix}.$$

Det kan anvendes uden bevis, at $(M_2(\mathbb{C}), +, \cdot)$ er et $\mathbb{R}$ -vektorrum.

- Vis, at mængden V af hermitiske matricer udgør et underrum i $M_2(\mathbb{C})$
- Vis, at familien $(I, \sigma_1, \sigma_2, \sigma_3)$, hvor I er identitetsmatricen, og hvor σ_1, σ_2 og σ_3 er Pauli-matricerne, udgør en basis for V .

6.12 Lad $Q \in M_n(\mathbb{R})$ være en ortogonal matrix, og lad $U \in M_n(\mathbb{C})$ være en unitær matrix.

a) Vis, at $|\det(Q)| = 1$.

b) Vis, at $|\det(U)| = 1$.

(Vink: Husk, at for $z \in \mathbb{C}$ er $|z| = \sqrt{z^* z}$.)

6.13 Lad $Q \in M_n(\mathbb{R})$ være en ortogonal matrix.

a) Vis, at $Q^t(Q - I) = -(Q - I)^t$.

b) Vis, at hvis n er lige og $\det(Q) = -1$, så er $\lambda = 1$ en egenværdi for matricen Q .

c) Vis, at hvis n er ulige og $\det(Q) = 1$, så er $\lambda = 1$ en egenværdi for matricen Q .

(Vink: Hvis $A \in M_n(\mathbb{F})$, så er $\det(-A) = (-1)^n \det(A)$.)

6.14 Find et eksempel på to symmetriske matricer $A, B \in M_2(\mathbb{R})$, sådan at deres matrixprodukt $AB \in M_2(\mathbb{R})$ *ikke* er en symmetrisk matrix.

6.15 Lad $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ være den lineære endomorfi, der er repræsenteret med hensyn til standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for både domænet og codomænet af den symmetriske matrix

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{pmatrix} \in M_3(\mathbb{R}).$$

a) Udregn det karakteristiske polynomium $\chi_f(t)$ og bestem samtlige egenværdier af $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

b) Bestem en basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ for $\mathbb{R}^3$, der er ortonormal med hensyn til standardindreproduktet på $\mathbb{R}^3$ og består af egenvektorer for $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

c) Angiv en ortogonal matrix Q , sådan at

$$D = Q^{-1}AQ$$

er en diagonalmatrix.

6.16 Lad $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ være den lineære endomorfi, der er repræsenteret med hensyn til standardbasen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for både domænet og codomænet af den symmetriske matrix

$$A = \begin{pmatrix} 3 & -1 & -2 \\ -1 & 3 & 2 \\ -2 & 2 & 6 \end{pmatrix} \in M_3(\mathbb{R}).$$

a) Udregn det karakteristiske polynomium $\chi_f(t)$ og bestem samtlige egenværdier af $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

b) Bestem en basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ for $\mathbb{R}^3$, der er ortonormal med hensyn til standardindreproduktet på $\mathbb{R}^3$ og består af egenvektorer for $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

6 Vektorrum med indre produkt

- c) Angiv en ortogonal matrix Q , sådan at

$$D = Q^{-1}AQ$$

er en diagonalmatrix.

- 6.17 Lad $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ være den lineære endomorfi, der er repræsenteret med hensyn til standardbasen $(\mathbf{e}_1, \mathbf{e}_2)$ for både domænet og codomænet af den normale matrix

$$A = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \in M_2(\mathbb{C}).$$

- a) Udregn det karakteristiske polynomium $\chi_f(t)$ og bestem samtlige egenverdier af $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$.
b) Bestem en basis $(\mathbf{u}_1, \mathbf{u}_2)$ for $\mathbb{C}^2$, der er ortonormal med hensyn til standard-indreproduktet på $\mathbb{C}^2$ og består af egenvektorer for $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$.
c) Angiv en unitær matrix U , sådan at

$$D = U^{-1}AU$$

er en diagonalmatrix.

- 6.18 Lad $A \in M_n(\mathbb{R})$ være en symmetrisk matrix. Vis, at hvis A^2 er lig med nulmatricen, så er A nødvendigvis lig med nulmatricen.

(Vink: Anvend spektralsætningen.)

- 6.19 Lad $A \in M_n(\mathbb{C})$ være en normal kompleks matrix. Vis, at der findes en matrix $B \in M_n(\mathbb{C})$, sådan at $A = B^2$.

(Vink: Bemærk først, at der for alle $z \in \mathbb{C}$ findes $w \in \mathbb{C}$, sådan at $z = w^2$, og anvend derefter spektralsætningen.)

- 6.20 Vis, at hvis $A \in M_n(\mathbb{R})$ er en invertibel symmetrisk reel matrix, så er den inverse matrix $A^{-1} \in M_n(\mathbb{R})$ også symmetrisk.

(Vink: Brug spektralsætningen.)

- 6.21 (★) Lad $A, B \in M_n(\mathbb{R})$ være symmetriske matricer, der opfylder, at

$$AB = BA.$$

- a) Lad $\lambda \in \mathbb{R}$ være en egenverdi for A , og lad $(\mathbf{u}_1, \dots, \mathbf{u}_d)$ være en basis for egenrummet $N_{A-I\lambda}$, der er ortonormal med hensyn til standard-indreproduktet på $\mathbb{R}^n$. (Spektralsætningen viser, at disse findes.) Anvend antagelsen, at A og B kommuterer, til at vise, at det for alle $1 \leq i \leq d$ gælder, at

$$B\mathbf{u}_i \in N_{A-I\lambda}.$$

b) Vis, at matricen

$$C = (\langle B\mathbf{u}_i, \mathbf{u}_j \rangle) \in M_d(\mathbb{R})$$

er symmetrisk.

c) Lad $\mathbf{x} \in \mathbb{R}^d$ være en egenvektor for den symmetriske matrix C . (En sådan findes ifølge spektralsætningen.) Vis, at vektoren

$$\mathbf{v} = \mathbf{u}_1 x_1 + \dots + \mathbf{u}_d x_d$$

er en fælles egenvektor for A og B .

Man kan nu gentage beviset for sætning 6.4.2 og vise, at der findes en basis for $\mathbb{R}^n$, der er ortonormal med hensyn til standard-indreproduktet og består af egenvektorer for *både* A og B . Bemærk også, at antagelsen $AB = BA$ ifølge sætning 5.2.7 er nødvendig.

6.22 (★) Vi lader $C^0(\mathbb{R})$ være det reelle vektorrum bestående af kontinuerte funktioner $f: \mathbb{R} \rightarrow \mathbb{R}$, og vi definerer

$$\text{Trig}_N(\mathbb{R}) \subset C^0(\mathbb{R})$$

til at være underrummet frembragt af familien

$$(1, \cos(x), \cos(2x), \cos(3x), \dots, \cos(Nx)).$$

Vi kalder en vektor i $\text{Trig}_N(\mathbb{R})$ for et *signal* for at understrege anvendelsen til repræsentation af lyd. Det indre produkt

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)dx$$

på $C^0(\mathbb{R})$ definerer specielt et indre produkt på underrummet $\text{Trig}_N(\mathbb{R})$. Vi betragter i det følgende signalerne

$$\mathbf{v}_0 = \frac{1}{\sqrt{2}}, \quad \mathbf{v}_1 = \cos(x), \quad \mathbf{v}_2 = \cos(2x), \quad \mathbf{v}_3 = \cos(3x)$$

i $\text{Trig}_3(\mathbb{R})$.

a. Vis, eventuelt ved brug af Maple, at for $0 \leq i, j \leq 3$ er

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} \pi & \text{hvis } i = j, \\ 0 & \text{hvis } i \neq j, \end{cases}$$

og konkluder, at familien $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ er en ortogonal basis for underrummet $\text{Trig}_3(\mathbb{R})$ med hensyn til det indre produkt $\langle -, - \rangle$.

b. Angiv en basis $(\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ for $\text{Trig}_3(\mathbb{R})$, der er ortonormal med hensyn til det indre produkt $\langle -, - \rangle$.

6 Vektorrum med indre produkt

- c. Vis, at hvis $f, g \in \text{Trig}_3(\mathbb{R})$ har koordinater henholdsvis $\mathbf{x}, \mathbf{y} \in \mathbb{R}^4$ med hensyn til basen $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, så er

$$\langle f, g \rangle = \pi(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4).$$

- d. Betragt den lineære endomorfi $\Delta: \text{Trig}_3(\mathbb{R}) \rightarrow \text{Trig}_3(\mathbb{R})$ defineret ved

$$\Delta(f) = \frac{d^2f}{dx^2}.$$

Find den matrix $A \in M_4(\mathbb{R})$, der repræsenterer denne lineære endomorfi med hensyn til basen $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for både domænet og codomænet.

Den lineære endomorfi $\Delta: \text{Trig}_3(\mathbb{R}) \rightarrow \text{Trig}_3(\mathbb{R})$ kaldes for *Laplace operatoren*. Den er et eksempel på en differentialoperator.

Indeks

- ækvivalensrelation, 71
- absolutværdi, 229
- adjungeret lineær afbildning, 244, 245
- adjungeret matrix, 81, 245
 - determinant af, 109
 - egenskaber af, 83
- afbildning, 1
 - bijektiv, 1
 - injektiv, 1
 - inverse, 3
 - sammensatte, 2
 - surjektiv, 1
- affin afbildning, 53, 141
- affint underrum, 178
- algebraens fundamentalsætning, 127
- algebraisk multiplicitet, 204
- algebraisk multiplicitet, 219
- areal
 - af parallelogram, 98
- basis, 148
 - eksistens af, 153
- basisskift, 168
- bijektiv, 1, 75, 79
- billede, 172
 - basis for, 175
- Cauchy-Schwarz' ulighed, 229
- codimension
 - af underrum, 181
- codomæne, 1
- Cramers regel, 121
- delfamilie, 153
- determinant, 101, 111
 - af 2×2 -matrix, 99
 - af $n \times n$ -matrix, 101
 - af adjungeret matrix, 109
 - af invers matrix, 113, 120
 - af produkt, 112
 - af triangulær matrix, 116
 - rækkeudvikling af, 104, 107
 - søjleudvikling af, 107
- de Rham-cohomologi, 182
- diagonaliserbar
 - lineær endomorfi, 220
 - matrix, 207, 252, 253
 - samtidigt, 211
- diagonalmatrix, 65, 251
- diagram
 - kommutativt, 2
- dimension
 - af vektorrum, 156
- direkte sum, 193
- division, 5
- domæne, 1
- echelonform, 20
 - reduceret, 19
- egenrum
 - for kvadratisk matrix, 203
 - for lineær endomorfi, 217
- egenværdi
 - for kvadratisk matrix, 203
 - for lineær endomorfi, 217
- egenvektor

Indeks

- for kvadratisk matrix, 203
- for lineær endomorfi, 217
- endeligt frembragt, 154
- endomorf
 lineær, 199
- enhedsvektor, 232
- et-element, 5
- euklidisk transformation, 86
- Euklids algoritme, 128
- familie
 - af elementer, 39, 145
 - af skalarer, 159
 - frembringende, 148
 - lineært afhængig, 148
 - lineært uafhængig, 148
 - ortogonal, 232
 - ortonormal, 232
- fri variabel, 25, 28
- Gauss-elimination, 28, 72
- geometrisk multiplicitet, 204, 219
- grad
 - af determinant, 125
 - af polynomium, 124, 125
- gradient, 182
- Gram-Schmidt-ortogonalisering, 237
- Grassmanns dimensionsformel, 173
- gruppe, 67
 - ortogonale, 250
 - specielle lineære, 121
 - symplektiske, 250
 - unitære, 250
- gruppe, generelle lineære, 67
- Heisenbergs ubestemthedsrelationer, 211
- hermitisk form, 83
 - hørende til hermitisk matrix, 85
 - ikke-singulær, 85
 - klassifikation af, 261
- hermitisk matrix, 81, 249, 253
- hovedsætning
 - lineær algebras, 153
- hyperplan, 181
- identitetsafbildning, 2
- identitetsafbildningen, 58
- identitetsmatrix, 40
- indgang, 39
 - ledende, 20
- indre produkt, 228
 - standard, på $\mathbb{F}^n$, 228
- inhomogent, 12
- injektiv, 1, 75, 79, 144
- inklusion
 - kanonisk, 183
- invers
 - i gruppe, 67
- invers afbildning, 3
- invers matrix, 64
 - af 2×2 -matrix, 121
- invertibel, 5
 - skalar, 5
- isometri
 - lineær, 86
- isometrisk isomorfi, 86, 247
- isomorfi, 143
- kanonisk, 181
- kanonisk inklusion, 141, 183
- kanonisk projektion, 181, 183
- karakteristisk polynomium, 199, 216
- kardinalitet, 160
- kerne, 172
 - basis for, 176
- koefficientmatrix, 13
- kommutativ ring, 6, 122
 - af polynomier, 124
- kommutativt diagram, 2
- komplekse tal, 4
- koordinater
 - med hensyn til basis, 160
 - med hensyn til ortonormal basis, 233
- koordinatskift, 170, 171
- Kroneckers delta, 40, 46
- kvaternioner, 6

- kvotientrum, 181
 - universel egenskab, 183
- løsning til ligningssystem, 12
- løsningsmængde, 12
- Laplace-udvikling
 - af determinant, 107
- ledende indgang, 20, 28
- ledende variabel, 26, 28
- legeme, 3
 - skæv-, 6
- ligningssystem
 - homogent, 12
 - koefficientmatrix for, 13
 - rækkeoperationer på, 14
 - totalmatrix for, 13
- lineær afbildning, 52, 141
- lineær endomorfi, 199
- lineær isometri, 86, 243
- lineær isomorfi, 143
- lineær uafhængighed, 148
- lineært afhængig, 148
- lineært ligningssystem, 12
- linearkombination, 146, 159
 - af vektorer i $\mathbb{F}^m$, 50
- Lorenz-transformation, 86
- lyskegle, 263
- matrix, 39
 - adjungeret, 81
 - der repræsenterer en lineær afbildning, 57
 - diagonaliserbar, 207, 252
 - dimensioner af, 40
 - invers, 64
 - invertibel, 64
 - kvadratisk, 40
 - normal, 251, 253, 259
 - ortogonaliserbar, 253
 - produkt, 41
 - rækkeoperationer på, 15
 - sum, 41
 - symmetrisk, 82, 252, 256
- matrixrepræsentation
 - af hermitisk form, 85
 - af lineær afbildning, 162, 163
 - og basisskift, 168
 - og invers, 166
 - og sammensætning, 165
- Minkowskiformen, 80, 85, 261, 263
- misbrug af notation, 3, 48, 137, 139
- modsat element, 5
- multiplicitet
 - algebraisk, af egen værdi, 204, 219
 - geometrisk, af egen værdi, 204, 219
- multiplikativ invers, 5
- negativ definit, 260
- norm, 229
- normal
 - lineær endomorfi, 248, 255
 - matrix, 249, 251, 253, 259
- nul-element, 4
- nulmatrix, 40
- nulrum
 - af en matrix, 140
 - basis for, 150
- nulvektoren, 136
- operationsmatrix, 68
 - invers, 71
- ortogonal
 - familie, 232
 - matrix, 249
- ortogonal projektion, 235, 241
- ortogonalitet, 232
- ortogonalt komplement
 - med hensyn til indre produkt, 239
- ortonormal
 - familie, 232
- permutation
 - af n bogstaver, 114
- polynomium, 123
 - af grad d , 125
 - af grad højst d , 124, 125

Indeks

- karakteristisk, 199
- rod i, 127
- polynomiumsdivision, 128
- positiv definit, 228, 260
- positivitet
 - af indre produkt, 228
- projektion
 - kanonisk, 183
- rækkeækvivalens, 72
- rækkeækvivalensklasse, 74
- rækkeoperation
 - invers, 16
- rækkeoperationer, 14, 15
- rækkeudvikling
 - af determinant, 104, 107
- rang
 - af lineær afbildning, 174
 - af matrix, 74
 - af matrix på reduceret echelonform, 20
- rationale tal, 4
- reelle tal, 4
- restriktion af afbildning, 139
- ringhomomorfi, 126
- rotation, 182
- række, 13, 40
- søjlerum
 - af en matrix, 140
- søjleudvikling
 - af determinant, 107
- sammensatte afbildning, 2
- selvadjungeret
 - lineær endomorfi, 248, 254
- skævinvolution, 80
- skævlegeme, 6
- skalar, 1, 3
- skalarmultiplikation, 48, 135
- skalering, 202
- span
 - af familie af vektorer, 147
- sparse matrix, 118
- spektralsætningen
 - for normale afbildninger, 255
 - for normale matricer, 253
 - for selvadjungerede afbildninger, 254
 - for symmetriske matricer, 252
- spor, 89, 200
- stabil delmængde
 - af vektorrum, 137
- standard-indreproduktet på $\mathbb{F}^n$, 228
- standardbasen, 50
- standardenhedsvektor i $\mathbb{F}^n$, 50
- subtraktion, 5
- support, 159
- surjektiv, 1, 75, 79
- Sylvesters inertisætning, 261
- symmetrisk matrix, 82, 249, 252, 256
- søjle, 13, 40
- søjleoperationer, 79
- tomme familie, 145
 - linearkombination af, 146
- tomme sum, 50, 146
- topologisk vektorrum, 172
- totalmatrix, 13
- transponeret matrix, 82
- transposition, 114
- trekantsuligheden, 230
- triangulær matrix, 116, 251
 - øvre, 116
 - determinant af, 116
 - nedre, 116
- trigonometrisk polynomiell funktion, 232
- tuple, 49, 145
- undermatrix, 103
- underrum
 - affint, 178
 - frembragt af familie af vektorer, 147
 - lineært, 178
- unitær matrix, 249
- unitær symplektisk matrix, 250
- vektor

- rækkevektor, 47
- søjlevektor, 47
- vektorfelt, 182
 - konservativt, 182
 - rotationsfrit, 182
- vektorrum, 135
 - basis for, 148
 - endeligt frembragt, 154
 - højre, 135
 - med indre produkt, 228
 - topologisk, 172
 - underrum af, 139
 - venstre, 136
- vektorsum, 135
- vinkel
 - mellem vektorer i reelt vektorrum med indre produkt, 231
 - mellem vektorer i reelt vektorrum med indre produkt, 236
- volumen
 - af parallelepipedum, 101
- Weierstrass' approksimationssætning, 172
- Zorns lemma, 159