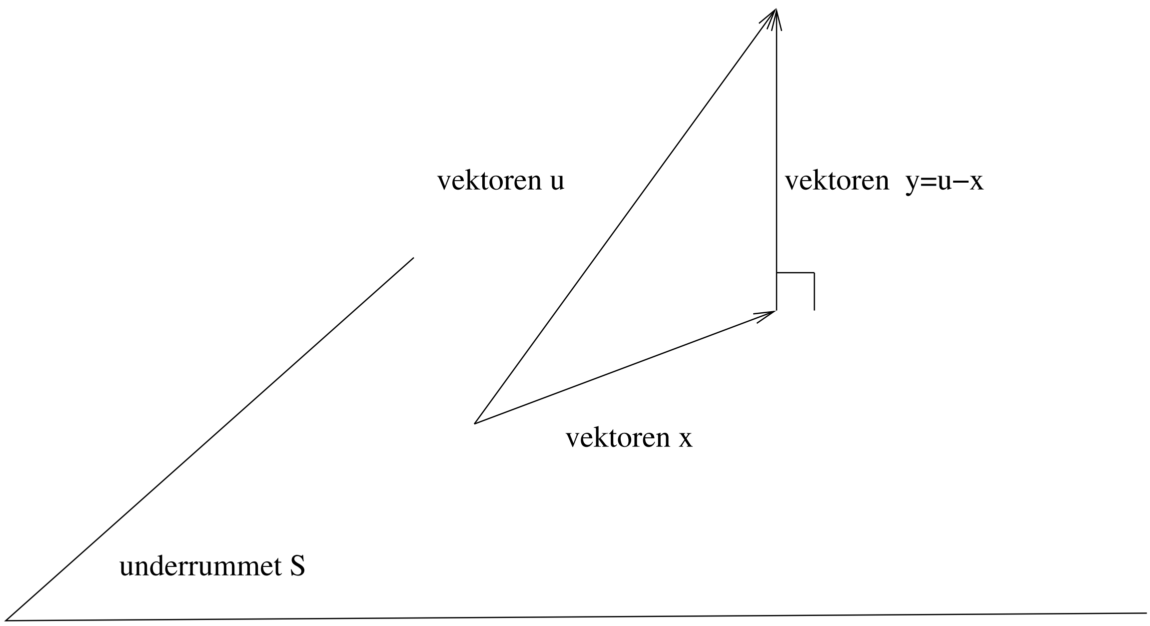
Ortogonalitet, projektion

Flemming Topsøe

Mat 1GA-notat, 3.12.03



1 Her er nogle fiduser!

Ganske kort kan du i denne appetitvækker få nogle fiduser, der ofte letter arbejdet med at bestemme projektionen af en vektor på et underrum og lignende sager. I de følgende afsnit kan du se en mere udførlig redegørelse.

Lad S være et underrum af et vektorrum V med indre produkt og lad $\mathbf{u} \in V$. Med $\mathrm{proj}_S\,\mathbf{u}$ betegner vi som i Messer projektionen af $\mathbf{u}$ på S. Her er en kendsgerning, du kan have glæde af:

Vektoren $\mathbf{x} = \operatorname{proj}_S \mathbf{u}$ *er entydigt bestemt som den vektor, der opfylder følgende to krav:*
1. $\mathbf{x} \in S$
2. $\langle \mathbf{x}, \mathbf{t} \rangle = \langle \mathbf{u}, \mathbf{t} \rangle$ *for alle* $\mathbf{t} \in S$.

I ord viser betingelse 2, at den søgte vektor, $\mathbf{x}$, har samme indre produkt med enhver vektor i S som $\mathbf{u}$.

Er $S = \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_n)$, gælder følgende nyttige variant:

Vektoren $\mathbf{x} = \operatorname{proj}_S \mathbf{u}$ *er entydigt bestemt som den vektor, der opfylder følgende to krav:*
1. $\mathbf{x} \in S$
2'. $\langle \mathbf{x}, \mathbf{u}_i \rangle = \langle \mathbf{u}, \mathbf{u}_i \rangle$ *for* $i = 1, \cdots, n$.

Dernæst en fidus vedrørende Gram-Schmidt ortonormalisering. Fidusen består i at vente med normeringen. Det er det sidste, man gør. Så slipper man typisk for at slæbe rundt på irriterende kvadratrødder eller fjollede brøker.

Lad $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ *være givne og sæt* $S_k = \operatorname{span}(\mathbf{u}_1, \cdots, \mathbf{u}_k)$ *for* $k = 1, \cdots, n$. *Bestem et nyt sæt* $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ *af vektorer ved*
$\mathbf{v}_1 = \mathbf{u}_1$, $\mathbf{v}_k = \mathbf{u}_k - \operatorname{proj}_{S_{k-1}} \mathbf{u}_k$ *for* $k = 2, \cdots, n$.
Så er $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ *et ortogonalsæt. Sættes*
$\mathbf{e}_k = \frac{1}{\|\mathbf{v}_k\|} \mathbf{v}_k$ *for* $k = 1, \cdots, n$,
er $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ *et ortonormalsæt.*

De vektorer, $\mathbf{v}_k$ og $\mathbf{e}_k$, der bestemmes her, er eksakt de samme som de vektorer, der optræder i Messer Theorem 4.17 og betegnelserne er de samme. I Messer står en præcis formel til at bestemme $\mathbf{v}$'erne. Normalt er det lettere i stedet at benytte en bestemmelse af de optrædende projektioner via betingelserne 1 og 2' ovenfor. Når, f.eks. vektoren $\mathbf{v}_3$ skal bestemmes, sætter man
$\operatorname{proj}_{S_2} \mathbf{u}_3 = \alpha \mathbf{u}_1 + \beta \mathbf{u}_2$
(herved sikres, at betingelse 1 er opfyldt) og koefficienterne α og β bestemmes så af ligningerne, der fås af 2', altså af
$\langle \alpha \mathbf{u}_1 + \beta \mathbf{u}_2, \mathbf{u}_1 \rangle = \langle \mathbf{u}_3, \mathbf{u}_1 \rangle$
$\langle \alpha \mathbf{u}_1 + \beta \mathbf{u}_2, \mathbf{u}_2 \rangle = \langle \mathbf{u}_3, \mathbf{u}_2 \rangle$.
Den søgte vektor er så
$\mathbf{v}_3 = \mathbf{u}_3 - \operatorname{proj}_{S_2} \mathbf{u}_3 = \mathbf{u}_3 - \big(\alpha \mathbf{u}_1 + \beta \mathbf{u}_2\big)$.

2 Hvorfor dette notat?

Fordi Messer's behandling af projektion og ortonormalisering ikke er optimal. På ét punkt skjuler han noget væsentligt, på et andet fører hans anvisninger ofte til unødigt komplicerede regninger.

Vil du bare have nogle fiduser, der kan lette løsning af opgaver omkring projektion og ortogonalisering – eller orto*norma*lisering – kan du nøjes med at kigge på Korollar 1 og sætningerne 6 og 7, der ofte er en hjælp i forhold til Messer's Theorem 4.17 (Gram-Schmidt ortonormalisering).

Vil du lære mere, kan du jo læse hele notatet, der stort set er uafhængig af Messer, Chapter 4 (men dog ikke erstatter alt herfra).

I de følgende afsnit er V et vektorrum med indre produkt og S et underrum.

3 Ortogonal projektion

Ved $S^\perp$ forstås mængden af vektorer, der er ortogonale på *enhver* vektor i S.

Sætning 1 *$S^\perp$ er et underrum.*

BEVIS: Følger af det indre produkts linearitetsegenskaber. Overvej ! $\square$

BEMÆRKNING, UDFORDRING: Var det vigtigt i sætningen at forudsætte, at S er et underrum? Nej, vel?

EKSEMPLER: Er $S = \{\mathbf{0}\}$, er $S^\perp = V$ og er $S = V$, er $S^\perp = \{\mathbf{0}\}$. $\square$

Øvelse 1. Vis, at $S \subseteq S^{\perp\perp}$ (her er $S^{\perp\perp}$ en kort betegnelse for $(S^\perp)^\perp$). *Bemærkning*: Eksemplerne ovenfor kunne tyde på at der gælder det stærkere resultat $S = S^{\perp\perp}$. Det er nu ikke altid rigtigt, men vigtige positive resultater herom findes, se bl.a. Øvelse 3. $\square$

Sætning 2 *Hvis $S = \mathrm{span}(\mathbf{u}_1, \ldots, \mathbf{u}_n)$, er $S^\perp$ mængden af vektorer, der er ortogonale på enhver af vektorerne $\mathbf{u}_1, \ldots, \mathbf{u}_n$.*

BEVIS: Følger af det indre produkts linearitetsegenskaber. Overvej! $\square$

Morale: Nok at checke ortogonalitet for vektorerne i et udspændende sæt.

Lad der nu, udover underrummet S, være givet en vektor $\mathbf{u} \in V$. Vi siger da, at vektoren $\mathbf{x}$ er den *ortogonale projektion af $\mathbf{u}$ på S* såfremt de associationer, figuren på side 1 giver os, er rigtige, altså såfremt dels $\mathbf{x} \in S$, dels $\mathbf{x}$ virkelig fås ved at "gå vinkelret ned fra $\mathbf{u}$ på S", altså såfremt vektoren $\mathbf{u} - \mathbf{x}$ er ortogonal på S.[1]

Hvordan kan vi tænke på vektoren $\mathbf{u} - \mathbf{x}$? Først et banalt kneb: Kald den noget![2] Lad os sige $\mathbf{y}$. Altså $\mathbf{y} = \mathbf{u} - \mathbf{x}$ og dermed

$$\mathbf{u} = \mathbf{x} + \mathbf{y}\,. \tag{1}$$

Hvis jeg påstår, at $\mathbf{u} = \mathbf{x}$, begår jeg en fejl. Jeg må lægge $\mathbf{y}$ til $\mathbf{x}$ for at få $\mathbf{u}$. Vi kalder $\mathbf{y}$ for *fejlvektoren.* Vektoren $\mathbf{x}$ er, kort udtrykt, *projektionen* og betegnes $\mathrm{proj}_S\,\mathbf{u}$. Når vi taler om projektionen er det altså underforstået, at det drejer sig om den ortogonale projektion.

Sætning 3 *Projektionen er entydigt bestemt.*

BEVIS: Påstanden er, at der til den givne vektor $\mathbf{u}$ og det givne underrum S højst kan findes een vektor $\mathbf{x}$, så $\mathbf{x}$ er projektionen af $\mathbf{u}$ på S. Antag, både at $\mathbf{x} = \mathrm{proj}_S\,\mathbf{u}$ og at $\mathbf{x}' = \mathrm{proj}_S\,\mathbf{u}$. Lad de tilhørende fejlvektorer være $\mathbf{y}$ og $\mathbf{y}'$. Så er

$$\mathbf{u} = \mathbf{x} + \mathbf{y} \text{ og } \mathbf{u} = \mathbf{x}' + \mathbf{y}'\,. \tag{2}$$

Specielt må differensvektoren $\mathbf{x} - \mathbf{x}'$ være den samme som differensvektoren $\mathbf{y}' - \mathbf{y}$. Kald den $\mathbf{z}$. Af $\mathbf{z} = \mathbf{x} - \mathbf{x}'$ følger det, at $\mathbf{z} \in S$. Af $\mathbf{z} = \mathbf{y}' - \mathbf{y}$ følger det, at $\mathbf{z} \in S^{\perp}$ (Sætning 1!). Så må $\langle \mathbf{z}, \mathbf{z} \rangle = \mathbf{0}$. Overvej! Så er $\mathbf{z} = \mathbf{0}$ og det følger, som ønsket, at $\mathbf{x}' = \mathbf{x}$ (og at $\mathbf{y}' = \mathbf{y}$). $\square$

Fint! projektionen er altså entydig. Men findes den? Nej, faktisk eksisterer projektionen ikke altid. Men er S endelig dimensional, er det i orden. Det viser vi i Sætning 5. Her følger først en nem sætning, der tillader os at bestemme projektionen i mange tilfælde.

Sætning 4 *Lad* $\mathbf{u}$ *og* S *være givne og lad* $\mathbf{x}$ *være en eller anden vektor. Der gælder da, at* $\mathbf{x} = \mathrm{proj}_S\,\mathbf{u}$*, hvis og kun hvis* $\mathbf{x}$ *opfylder følgende to betingelser:*

1. $\mathbf{x} \in S$*,*

2. Vektoren $\mathbf{x}$ *har samme indre produkt med enhver vektor i* S *som* $\mathbf{u}$*.*

[1]Bemærk, at denne definition er forskellig fra den i Messer. Faktisk definerer Messer kun projektion i en meget speciel situation (M, Definition 4.18). Dybt utilfredsstillende! – bl.a. da I senere (Mat 2MA, vist nok) skal arbejde med begrebet i den generalitet, der her er givet. Og så er det da synd at snyde sig for den tiltalende geometrisk inspirerede definition der straks viser, hvad meningen er.

[2]Opgavebesvarelserne fra klasserne viser, at I – Mat 1GA-studerende – udnytter dette kneb alt for sjældent. Kald ting noget og alt bliver klarere! Ikke dybt, men effektivt.

BEVIS: Ifølge definitionen er betingelsen for at $\mathbf{x}$ er projektionen af $\mathbf{u}$ på S, dels, at $\mathbf{x} \in S$, dels at $\mathbf{u} - \mathbf{x} \in S^{\perp}$. Sidste betingelse er, at det for alle $\mathbf{t} \in S$ gælder, at $\langle \mathbf{u} - \mathbf{x}, \mathbf{t} \rangle = 0$ ($\mathbf{t}$ er en "testvektor"). På grund af det indre produkts linearitetsegenskaber kan denne betingelse også skrives på formen

$$\langle \mathbf{x}, \mathbf{t} \rangle = \langle \mathbf{u}, \mathbf{t} \rangle \ \text{ for alle } \mathbf{t} \in S\,. \tag{3}$$

Resultatet følger. □

Ofte er følgende variant mere nyttig end sætningen:

Korollar 1 *Hvis vektorerne* $\mathbf{u}_1, \ldots, \mathbf{u}_n$ *udspænder* S*, er* $\mathbf{x} = \mathrm{proj}_S\, \mathbf{u}$*, hvis og kun hvis* $\mathbf{x} \in S$ *og*

$$\langle \mathbf{x}, \mathbf{u}_i \rangle = \langle \mathbf{u}, \mathbf{u}_i \rangle \ \textit{ for } i = 1, \cdots, n. \tag{4}$$

BEVIS: "kun hvis": Klart (og ret unyttigt!). "hvis": Er betingelserne opfyldt, er $\langle \mathbf{u} - \mathbf{x}, \mathbf{u}_i \rangle = 0$ for $i = 1, \cdots, n$ og dermed vil $\mathbf{u} - \mathbf{x} \in S^{\perp}$ ifølge Sætning 2. Derfor! □

Det er altså tilstrækkeligheden (hvis-delen), der er det vigtigste. For at checke, at en vektor $\mathbf{x}$ er projektionen af $\mathbf{u}$ på S er det altså nok at checke, at $\mathbf{x} \in S$ og at $\mathbf{x}$ har samme indre produkt med enhver vektor i et udspændende sæt som $\mathbf{u}$. Særdeles nyttigt!

Lad os først se et eksempel på anvendelse af Sætning 4, dernæst et eksempel, hvor korollaret kommer i sving.

Eksempel 1. Opgaven 1GA, S03, nr.5 ser på underrummet $S \subseteq \mathbb{R}^4$, der består af vektorer (x_1, x_2, x_3, x_4), så $x_2 = 0$ og $x_3 = x_4$. Vi skal bestemme projektionen $\mathbf{w}$ af $\mathbf{v} = (3, 0, 2, 1)$ på S.

Har man først gættet svaret, er det let at checke. Og et kvalificeret gæt kan man f.eks. finde ved at bemærke, at begge vektorer $\mathbf{u}_1 = (1, 0, 0, 0)$ og $\mathbf{u}_2 = (0, 0, 1, 1)$ ligger i S. Så prøver vi med $\mathbf{w} = \alpha\mathbf{u}_1 + \beta\mathbf{u}_2$. Ligningerne $\langle \mathbf{w}, \mathbf{u}_1 \rangle = \langle \mathbf{v}, \mathbf{u}_1 \rangle$ og $\langle \mathbf{w}, \mathbf{u}_2 \rangle = \langle \mathbf{v}, \mathbf{u}_2 \rangle$, som vi jo må forlange opfyldt i henhold til Sætning 4, giver $\alpha = 3$ og $2\beta = 3$. Derfor er $\mathbf{w} = (3, 0, \frac{3}{2}, \frac{3}{2})$ et rigtig godt gæt. At det virker, er let at checke: Den fundne vektor $\mathbf{w}$ tilhører klart S og for en vilkårlig vektor $\mathbf{x} = (x_1, x_2, x_3, x_4) \in S$ får vi

$$\langle \mathbf{w}, \mathbf{x} \rangle = 3x_1 + \frac{3}{2}(x_3 + x_4) = 3x_1 + 3x_3\,, \tag{5}$$

der er det samme som $\langle \mathbf{v}, \mathbf{x} \rangle$. Sætning 4 viser så umiddelbart, at $\mathbf{w}$ er den ønskede projektion.

Er man fræk, kan løsningen på opgaven skrives på to linier:
Da $\mathbf{w} = (3, 0, \frac{3}{2}, \frac{3}{2}) \in S$ og da $\langle \mathbf{w}, \mathbf{x} \rangle = 3x_1 + \frac{3}{2}(x_3 + x_4) = 3 + 3x_3 = \langle \mathbf{v}, \mathbf{x} \rangle$ for enhver vektor $\mathbf{x} = (x_1, x_2, x_3, x_4) \in S$, er $\mathbf{w}$ den søgte projektion.
- selvfølgelig, en så fræk (men perfekt!) løsning kræver, at man først har grisset lidt på sit kladdepapir og fundet frem til det rigtige gæt. $\square$

Eksempel 2. Lad $V = \mathbb{C}([-1, 1])$ udstyret med det indre produkt

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx \tag{6}$$

og lad S være underrummet af alle polynomier af grad højst 3.[3] Lad f være givet ved $f(x) = x^4\,; x \in [-1, 1]$ og lad os bestemme $\mathrm{proj}_S\, f$.

Først bemærkes, at S er underrummet udspændt af de fire første elementarpolynomier p_0, p_1, p_2 og p_3 (så faktisk er $f = p_4$).[4] Projektionen, vi søger, kalder vi g.

Vi ser på betingelserne i Sætning 4. Første krav sikrer vi os ved at søge efter en linearkombination af p_0, p_1, p_2 og p_3, lad os sige:

$$g = ap_0 + bp_1 + cp_2 + dp_3\,. \tag{7}$$

Anden betingelse bliver nu $\langle g, p_0 \rangle = \langle f, p_0 \rangle$, $\langle g, p_1 \rangle = \langle f, p_1 \rangle$, $\langle g, p_2 \rangle = \langle f, p_2 \rangle$ og $\langle g, p_3 \rangle = \langle f, p_3 \rangle$ eller:

$$a\frac{2}{1} + c\frac{2}{3} = \frac{2}{5} \tag{8}$$

$$b\frac{2}{3} + d\frac{2}{5} = 0 \tag{9}$$

$$a\frac{2}{3} + c\frac{2}{5} = \frac{2}{7} \tag{10}$$

$$b\frac{2}{5} + d\frac{2}{7} = 0\,. \tag{11}$$

[3]en spidsfindighed: Idet polynomier normalt er defineret på hele $\mathbb{R}$, bør vi her pinligt korrekt tale om restriktionen af disse polynomier til det aktuelle definitionsinterval, $[-1, 1]$. Det er jo ikke til at holde ud, og man tillader sig da også at undlade at være *så* omhyggelige med notationen i den slags tilfælde.

[4]*Elementarpolynomierne* er polynomierne $p_k\,; k \geq 0$ givet ved $p_k(x) = x^k$.

Af (8) og (10) ses, at $b = d = 0$ og af (7) og (9) fås $a = -\frac{3}{35}\,, c = \frac{6}{7}$. Vi konkluderer, at projektionen findes i dette tilfælde (ikke overraskende!) og at den er givet ved

$$g(x) = -\frac{3}{35} + \frac{6}{7}x^2 \ , \ x \in [-1, 1]\,. \tag{12}$$

□

Øvelse 2 Find projektionen i situationen fra eksemplet, hvis f i stedet er givet ved $f(x) = (x-1)^4\,; x \in [-1, 1]$. Gør dette under udnyttelse af resultatet i eksemplet og uden at regne alt for meget.

Hvis det påstås, at $f = g$, begås en fejl. Hvor stor er den målt med normen af fejlvektoren? Sammenlign med den tilsvarende fejl svarende til f fra eksemplet.

4 Eksistens af projektion

Først en hjælpesætning:

Lemma 1 *Antag, at* $\mathbf{u}_1, \ldots, \mathbf{u}_n$ *udspænder* S *og lad* $\mathbf{u} \in V$. *Der gælder da, at projektionen eksisterer, hvis og kun hvis ligningssystemet*

$$\begin{array}{l} r_1 \langle \mathbf{u}_1, \mathbf{u}_1 \rangle + \cdots + r_n \langle \mathbf{u}_n, \mathbf{u}_1 \rangle = \langle \mathbf{u}, \mathbf{u}_1 \rangle \ , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ r_1 \langle \mathbf{u}_1, \mathbf{u}_n \rangle + \cdots + r_n \langle \mathbf{u}_n, \mathbf{u}_n \rangle = \langle \mathbf{u}, \mathbf{u}_n \rangle \ , \end{array} \tag{13}$$

har en løsning. I bekræftende fald er

$$\mathrm{proj}_S\, \mathbf{u} = r_1 \mathbf{u}_1 + \cdots + r_n \mathbf{u}_n\,. \tag{14}$$

Bevis: Udnyt, at ligningerne (13) udtrykker, at $\langle \mathbf{v}, \mathbf{u}_k \rangle = \langle \mathbf{u}, \mathbf{u}_k \rangle$ for $k = 1, \cdots, n$ med $\mathbf{v}$ defineret ved linearkombinationen i (14). Resultatet følger så af Korollar 1. □

Advarsel: Projektionen er jo entydig iflg. Sætning 3, men der kan sagtens være flere løsninger til (13). De vil bare allesammen definere samme vektor via linearkombinationen på højre-siden i (14).

Sætning 5 *Er* S *endelig-dimensional, eksisterer projektionen* $\mathrm{proj}_S\, \mathbf{u}$ *for enhver vektor* $\mathbf{u} \in V$.

Bevis: Antag $dim(S) = n$ og vælg en endelig basis $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ for S. Se på ligningssystemet (13). Vi påstår, at rækkevektorerne i den tilhørende koefficientmatrix er lineært uafhængige. Med henblik på et indirekte bevis herfor antager vi, at de er afhængige. Ifølge Messer, Theorem 3.5 er en af rækkerne så en linearkombination af de øvrige. Vi kan gerne antage, at det drejer sig om den sidste række. Så findes skalarer $c_1, \cdots, c_{n-1}$ således, at

$$\begin{aligned}
&c_1\big(\langle \mathbf{u}_1, \mathbf{u}_1\rangle, \cdots, \langle \mathbf{u}_n, \mathbf{u}_1\rangle\big) \\
&+c_2\big(\langle \mathbf{u}_1, \mathbf{u}_2\rangle, \cdots, \langle \mathbf{u}_n, \mathbf{u}_2\rangle\big) \\
&+\cdots \\
&+c_{n-1}\big(\langle \mathbf{u}_1, \mathbf{u}_{n-1}\rangle, \cdots, \langle \mathbf{u}_n, \mathbf{u}_{n-1}\rangle\big) \\
&=c_1\big(\langle \mathbf{u}_1, \mathbf{u}_n\rangle, \cdots, \langle \mathbf{u}_n, \mathbf{u}_n\rangle\big),
\end{aligned}$$

eller

$$\langle \mathbf{u}_k, \mathbf{v}\rangle = \langle \mathbf{u}_k, \mathbf{u}_n\rangle \text{ for } k = 1, \cdots, n, \tag{15}$$

hvor vi har sat

$$\mathbf{v} = c_1\mathbf{u}_1 + \cdots + c_{n-1}\mathbf{u}_{n-1}. \tag{16}$$

Overvej! Af (15) følger, at $\mathbf{u}_k \perp (\mathbf{u}_n - \mathbf{v})$ for $k = 1, \cdots, n$. Så viser Sætning 2, at $\mathbf{u}_n - \mathbf{v} \in S^\perp$. Da $\mathbf{u}_n - \mathbf{v} \in S$, sluttes heraf, at $\langle \mathbf{u}_n - \mathbf{v}, \mathbf{u}_n - \mathbf{v}\rangle = 0$. Så er $\mathbf{u}_n - \mathbf{v} = \mathbf{0}$ eller $\mathbf{u}_n = \mathbf{v}$. Dette er en modstrid, da $\mathbf{u}_n$ vides ikke at være en linearkombination af $\mathbf{u}_1, \cdots, \mathbf{u}_{n-1}$.

Vi må forkaste vores antagelse og kan dermed konkludere, som ønsket, at rækkerne i (13) er lineært uafhængige. Så har (13) en løsning. Overvej! [5] Nu følger resultatet af lemmaet. □

Øvelse 3. Bevis, at hvis S er endelig-dimensional, så gælder formlen $S^{\perp\perp} = S$. *Vejledning.* Brug et indirekte bevis for den ikke-trivielle del. □

[5] Messer kunne godt have været så venlig, efter definitionen i Chapter 3 af lineær uafhængighed, at udnytte dette begreb sammen med den indsigt, der er vundet i Chapter 2 vedrørende lineære ligningssystemer, til at påpege, at såfremt rækkevektorerne i koefficientmatricen til et lineært ligningssystem er lineært uafhængige, så har ligningssystemet, uanset højresiden, en løsning.

5 Konstruktion af ortogonal- og ortonormalsystemer

Udgangspunktet er et *helt almindeligt system* $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ af vektorer i V, hvilket blot vil sige, at der ikke er knyttet nogen særlige betingelser til vektorerne. Finere er det, at være et OGS, et *ortogonalsystem*, endnu finere at være et ONS, et *ortonormalsystem* og aller finest er det da at være en ONB, en *ortonormalbasis* (se definitioner i Messer).

Men nu til det helt almindelige system $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$. Vi vil interessere os for de *tilhørende underrum* $S_1, \cdots, S_n$ defineret ved

$$S_k = \operatorname{span}(\mathbf{u}_1, \cdots, \mathbf{u}_k) \text{ for } k = 1, \cdots, n\,. \tag{17}$$

Det viser sig bekvemt også at definere underrummet S_0 som det trivielle underrum $S_0 = \{\mathbf{0}\}$.

Vi er interesserede i at knytte et passende OGS til $\mathbf{u}_1, \ldots, \mathbf{u}_n$ og finder det naturligt at kigge på vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$ defineret ved

$$\mathbf{v}_k = \mathbf{u}_k - \operatorname{proj}_{S_{k-1}} \mathbf{u}_k \text{ for } k = 1, \cdots, n\,. \tag{18}$$

For $k = 1$ skal formlen tolkes $\mathbf{v}_1 = \mathbf{u}_1$ (svarende til $S_0 = \{\mathbf{0}\}$).

Sætning 6 (Konstruktion af OGS) *Med vektorerne $\mathbf{v}_k$ givet ved (18), er $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ et* OGS. *Endvidere er*

$$\operatorname{span}(\mathbf{v}_1, \cdots, \mathbf{v}_k) = S_k \textit{ for } k = 1, \cdots, n. \tag{19}$$

Bevis: Lad $1 \leq l < k \leq n$. Så vil $\mathbf{v}_l \in S_l \subseteqq S_{k-1}$ og da $\mathbf{v}_k \in S_{k-1}^{\perp}$ vil $\mathbf{v}_k \perp \mathbf{v}_l$. Dette viser, at $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ er et OGS.

Det er klart, at (19) gælder for $k = 1$. Bemærk nu, at gyldigheden af (19) for én k-værdi medfører gyldigheden for den næste (thi af (19) og af en antagelse om at $\operatorname{span}(\mathbf{v}_1, \cdots, \mathbf{v}_{k-1}) = S_{k-1}$ følger, at $\mathbf{v}_k \in S_k$ og at $\mathbf{u}_k \in \operatorname{span}(\mathbf{v}_1, \cdots, \mathbf{v}_k)$). Så må (19) gælde for alle $k = 1, \cdots, n$. Overvej![6]

Nu er det fristende simpelthen at normere $\mathbf{v}$'erne og på den måde få et ONS. Men kan vi det? Kunne vi risikere, at en eller flere af $\mathbf{v}$'erne er nulvektoren? Svaret er enkelt:

---

[6]den meget opmærksomme læser vil bemærke, at dette ræsonnement minder om et induktionsbevis. Det er det også! Blot induktion over den endelige mængde $\{1, \cdots, n\}$, ikke over $\mathbb{N}$. Vi bygger på at $\{1, \cdots, n\}$ er endelig og dermed har enhver ikke-tom delmængde et første element.

Sætning 7 (OGS → ONS → ONB) *Vektorerne* $\mathbf{u}_1, \ldots, \mathbf{u}_n$ *er lineært uafhængige, hvis og kun hvis ingen af vektorerne* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *givet ved (18) er nulvektoren.*

Hvis $\mathbf{u}$*'erne er lineært uafhængige og vi definerer* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *ved*

$$\mathbf{e}_k = \frac{1}{\|\mathbf{v}_k\|}\mathbf{v}_k \textit{ for } k = 1, \cdots, n\,,$$

så er $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ *et* ONS. *Dette* ONS *opfylder betingelserne*

$$span(\mathbf{e}_1, \cdots, \mathbf{e}_k) = S_k \textit{ for } k = 1, \cdots, n\,,$$

specielt er $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ *en* ONB *for* S_n.

Bevis: Antag først, at $\mathbf{u}_1, \ldots, \mathbf{u}_n$ er lineært uafhængige. Så er $\dim(S_n) = n$. Af (19) ses, at $\mathbf{v}_1, \ldots, \mathbf{v}_n$ udspænder S_n. Ifølge Messer, Theorem 3.15 er $\mathbf{v}_1, \ldots, \mathbf{v}_n$ så lineært uafhængige, specielt er ingen af disse vektorer nulvektoren.

Antag dernæst, at ingen af $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er nulvektoren. Da $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ ifølge Sætning 6 er et OGS, kan vi anvende Messer, Theorem 4.14 og slutte, at $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængige. Analogt med ræsonnementet ovenfor ses så, at $\mathbf{u}_1, \ldots, \mathbf{u}_n$ er lineært uafhængige.

Sidste del af sætningen er oplagt. □

6 Perspektivering, didaktiske kommentarer

Hvor står den vigtige Sætning 5 om projektionens eksistens hos Messer? I Theorem 4.17 om Gram-Schmidt ortonormalisering! Det står der altså, Messer laver ingen fejl. Men det er skjult godt. Messer har helt sikkert bevidst valgt at fremstille sagerne på sin måde. Det giver nu nogle ejendommeligheder.

Skal man anvende sætningen hos Messer, skal man strengt taget først kontrollere, at de givne vektorer er lineært uafhængige (det vil man opdage af sig selv, se Sætning 7). Dernæst er det dybt sært, at man strengt taget kun kan tale om projektion, hvis der er givet en ortonormalbasis i det underrum, man projicerer ned på (undtagelse: Messer behandler særskilt projektion på 1-dimensionalt underrum). Og som et tredje kritikpunkt fremføres, at det normalt giver en del lettelser, rent regnemæssigt, hvis man først ortogonaliserer (Sætning 6) og så til sidst normerer (Sætning 7). Herved undgås typisk at man hele vejen slæber rundt på en masse kvadratrødder eller sære brøker.

Så der er god grund til at “bessermache” på Messer. Men, vil læseren måske sige, er prisen ikke for høj, blev vi ikke nødt til at skrive en hel del ekstra for at få de gode resultater frem? Hmmm. Til dels er det en smagssag. Desuden kan man sige, at man *kunne* have gjort som Messer og blot efter den famøse Theorem 4.17 have suppleret med de “rigtige” bemærkninger. Jeg har valgt noget, der tager lidt mere plads. Det er især ræsonnementet i beviset for Sætning 5. På en vis måde er det overflødigt (det kan erstattes af Messer’s ræsonnement), men fordelen er, at det sætter fingeren på noget væsentligt i stedet for at skjule det og desuden muliggør det en flydende fremstilling byggende på “det rigtige” projektionsbegreb (ikke noget med at man partout skal have givet en ONB før man kan røre sig af stedet).

Det må vist være nok omkring mine pædagogisk-didaktiske overvejelser.

Lidt mere om perspektiverne.

Projektionen eksisterer altså altid når vi projicerer på endeligt dimensionale underrum. Det er uhyre vigtigt, at det går godt i en række andre tilfælde, altså for projektion på uendeligt dimensionale underrum. Dette kan kun klares ved indførelse af nye betragtningsmåder. Det væsentlige er, at man kombinerer de algebraiske, mere regnemæssige aspekter med grænseovervejelser, noget, der hører analysen til. Med det indre produkt har vi tilført en geometrisk dimension. Men faktisk kan man også via det indre produkt tilføre grænseovervejelser. Ideen er egentlig uhyre enkel: Vi siger, at en følge af vektorer $(\mathbf{u}_n)_{n\geq 1}$ *konvergerer* mod vektoren $\mathbf{u}$, såfremt $\|\mathbf{u}-\mathbf{u}_n\| \to 0$ for $n \to \infty$. Sådan! Sætningen om projektionens eksistens kan man så vise gælder, blot S er et *lukket* underrum, dvs. et underrum, man “ikke kan komme ud af” ved grænseovergang. Beviset bliver dog helt, helt anderledes (parallellogramreglen spiller ind!).

Sagen er den, at Sætning 5 er elementær i den forstand, at vi “kunne pege” på det ønskede objekt, den ønskede projektion, simpelthen ved at skrive en formel ned for projektionen. I det generelle tilfælde er dette ikke muligt, og man må betjene sig af ikke-konstruktive metoder. Disse ikke-konstruktive metoder er ejendommeligt nok helt analoge til de problemer og løsninger, der skal til, når man skal gå fra at arbejde i $\mathbb{Q}$ til at arbejde i $\mathbb{R}$. Mere om det på 2.det (eller 3.die?) studieår. Filosofisk set fascinerende, men jeg må slippe emnet her.

Derimod kunne jeg vende tilbage til det konstruktive. Ovenfor påstod jeg, at vi gav en færdig formel for projektionen på et endelig-dimensionalt underrum. Måske ikke helt, men følges skridtene frem til Sætning 7, får vi en færdig algoritme. Det er det, Messer fokuserer på i sin Theorem 4.17. Faktisk *kan* man bygge en færdig formel op, der på en måde er endnu bedre end Messers algoritme. Hertil er determinanterne bekvemme. Lad mig i form af en ikke-så-nem opgave vise, hvordan det kan gøres.

Opgave. Lad $\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \cdots$ være en følge af vektorer i et vektorrum med indre produkt. Definer *Gramdeterminanten* $\Gamma_n = \Gamma(\mathbf{u}_0, \cdots, \mathbf{u}_n)$ som determinanten af $(n+1 \times n+1)$-matricen

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle & \cdots & \langle \mathbf{u}_0, \mathbf{u}_n \rangle \\ & \cdots\cdots\cdots\cdots\cdots\cdots & \\ \langle \mathbf{u}_n, \mathbf{u}_0 \rangle & \cdots & \langle \mathbf{u}_n, \mathbf{u}_n \rangle \end{bmatrix} \tag{20}$$

for $n \geq 0$ og sæt $\Gamma_{-1} = 1$.

Definer en følge af vektorer $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \cdots$ ved

$$\mathbf{v}_n = \det \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle & \cdots & \langle \mathbf{u}_0, \mathbf{u}_{n-1} \rangle & \mathbf{u}_0 \\ & \cdots\cdots\cdots\cdots\cdots\cdots & & \\ \langle \mathbf{u}_n, \mathbf{u}_0 \rangle & \cdots & \langle \mathbf{u}_n, \mathbf{u}_{n-1} \rangle & \mathbf{u}_n \end{bmatrix} \tag{21}$$

(med $\mathbf{v}_0 = \mathbf{u}_0$).

Indse først, at denne følge af vektorer er veldefineret ved at se på opløsning efter sidste søjle. Dette viser, at $\mathbf{v}_n \in \operatorname{span}(\mathbf{u}_0, \cdots, \mathbf{u}_n)$. De koefficienter, man finder ved fremstillingen af $\mathbf{v}_n$ som en linearkombination af $\mathbf{u}_0, \cdots, \mathbf{u}_n$ benævnes α_{nj}, altså

$$\mathbf{v}_n = \sum_{j=0}^{n} \alpha_{nj} \mathbf{u}_j \,. \tag{22}$$

Vis nu følgende:

(i). $\alpha_{nn} = \Gamma_{n-1}\,,$

(ii). Hvis $\mathbf{u}_0, \cdots, \mathbf{u}_n$ er lineært afhængige, er $\Gamma_n = 0\,.$

Vejledning. Vis, at rækkevektorerne i matricen hørende til Γ_n er lineært afhængige.

(iii). Hvis $\mathbf{u}_0, \cdots, \mathbf{u}_n$ er lineært uafhængige og $\Gamma_{n-1} \neq 0$, så er $\mathbf{v}_n \neq \mathbf{0}\,,$

(iv). $\langle \mathbf{v}_n, \mathbf{u}_i \rangle = 0$ for $i = 0, \cdots, n-1$ og $\|\mathbf{v}_n\|^2 = \Gamma_{n-1}\Gamma_n\,,$

(v). Er $\mathbf{u}_0, \cdots, \mathbf{u}_n$ lineært afhængige, er $\Gamma_n = 0$ og
er $\mathbf{u}_0, \cdots, \mathbf{u}_n$ lineært uafhængige, er $\Gamma_n > 0$.

Vejledning. Induktion!

(vi). Antag dernæst, at $\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \cdots$ er lineært uafhængige (dvs. alle endelige delsæt er lineært uafhængige). Det er åbenbart det samme som at sige, at $\mathbf{0}$ ikke dukker op blandt $\mathbf{v}_i$'erne og også det samme som at alle Gramdeterminanterne er positive. Overvej! Definer $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \cdots$ ved

$$\mathbf{e}_n = \frac{1}{\Gamma_{n-1}\Gamma_n}\mathbf{v}_n \tag{23}$$

og bevis, at disse vektorer udgør et ONS, dvs. $\mathbf{e}_n \perp \mathbf{e}_m$ for $n \neq m$ og $\|\mathbf{e}_n\| = 1$ for alle $n \geq 0$.

(vii). Udover antagelser og brug af betegnelser som under (vi), er der nu givet en vektor $\mathbf{u}$. Definer vektorer $\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2, \cdots$ ved

$$\mathbf{y}_n = \frac{1}{\Gamma_n} \det \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle & \cdots & \langle \mathbf{u}_0, \mathbf{u}_n \rangle & \mathbf{u}_0 \\ & \cdots\cdots\cdots\cdots\cdots\cdots & & \\ \langle \mathbf{u}_n, \mathbf{u}_0 \rangle & \cdots & \langle \mathbf{u}_n, \mathbf{u}_n \rangle & \mathbf{u}_n \\ \langle \mathbf{u}, \mathbf{u}_0 \rangle & \cdots & \langle \mathbf{u}, \mathbf{u}_n \rangle & \mathbf{u} \end{bmatrix}. \tag{24}$$

Bevis, at $\mathbf{y}$'erne er fejlvektorerne svarende til problemet med at bestemme $\mathbf{u}$'ernes projektioner på underrummene S_n. Med andre ord,

$$\mathrm{proj}_{S_n}(\mathbf{u}) = \mathbf{u} - \mathbf{y}_n \text{ for } n \geq 1\,. \tag{25}$$

□

Vi kan også notere os, at uligheden $\Gamma_n \geq 0$ kan opfattes som en generalisation af Cauchy Schwartz' ulighed. Ikke sandt?

Resultaterne omkring projektion og ortogonalisering (ortonormalisering) kan udnyttes i en række sammenhænge i analyse, herunder numerisk analyse, og f.eks. i forbindelse med prediktion i sandsynlighedsteoretiske modeller. Vigtigst, må man nok sige, er anvendelserne i Fourieranalysen, anvendelser som Messer da også kommer noget ind på. Mere følger på 2.det studieår.