Linear maps

We define the notion of linear maps between vector spaces. We show that, after choosing bases of its domain and target, a linear map is uniquely represented by a matrix. We continue to consider right vector spaces over a field or skewfield.

Definition 1. Let F be a field and let V and W be two F-vector spaces. A map $f\colon V \to W$ is linear if the following hold:

(L1) For all $x, y \in V$, $f(x+y) = f(x) + f(y)$.
(L2) For every $x \in V$ and $a \in F$, $f(x \cdot a) = f(x) \cdot a$.

We remark that being linear is a property of the map $f\colon V \to W$. We also note that a linear map always satisfies $f(0) = 0$. Indeed, using property (L1), we have

$$f(0) + f(0) = f(0+0) = f(0),$$

and subtracting $f(0)$ on both sides, we conclude that $f(0) = 0$ as claimed. A map $g\colon V \to W$ is said to be affine if the map $f\colon V \to W$ given by $f(x) = g(x) - g(0)$ is linear. Examples of linear maps are the identity map $\mathrm{id}_V\colon V \to V$ and the zero map $0\colon V \to W$. The constant map $b\colon V \to W$ with value $b \in W$ is affine; it is linear if and only if $b = 0$.

Proposition 2. *Let F be a field, let U, V, and W be three F-vector spaces. If two maps $f\colon V \to W$ and $g\colon U \to V$ are linear, then their composition $f \circ g\colon U \to W$ again is linear.*

Proof. We verify that $f \circ g$ satisfies (L1)–(L2), using that the maps f and g do so. First, for all $x, y \in U$,

$$\begin{aligned}(f \circ g)(x+y) &= f(g(x+y)) = f(g(x) + g(y)) = f(g(x)) + f(g(y)) \\ &= (f \circ g)(x) + (f \circ g)(y),\end{aligned}$$

which shows that $f \circ g$ has property (L1). Similarly, for all $x \in U$ and $a \in F$,

$$(f \circ g)(xa) = f(g(xa)) = f(g(x)a) = f(g(x))a = (f \circ g)(x)a,$$

which shows that $f \circ g$ also has property (L2). □

Proposition 3. *Let F be a field and let $f\colon V \to W$ be a linear map between two F-vector spaces. The following* (1)–(2) *are equivalent.*

(1) *The map $f\colon V \to W$ is a bijection.*
(2) *There exists a linear map $g\colon W \to V$ such that $f \circ g = \mathrm{id}_W$ and $g \circ f = \mathrm{id}_V$.*

Proof. We first suppose that (2) holds. Since the identity map is a bijection, the equality $f \circ g = \mathrm{id}_W$ shows that f is surjective, and the equality $g \circ f = \mathrm{id}_V$ shows that f is injective. This shows that (1) holds. Conversely, if (1) holds, then we may define a map $g\colon W \to V$ by declaring that $g(y) = x$ if and only if $y = f(x)$. It follows immediately from the definition that $f \circ g = \mathrm{id}_W$ and $g \circ f = \mathrm{id}_V$, but we must show that $g\colon W \to V$ is linear. Since $f\colon V \to W$ is linear, we have

$$\begin{aligned}y + z &= f(g(y)) + f(g(z)) = f(g(y) + g(z)) \\ ya &= f(g(y))a = f(g(y)a)\end{aligned}$$

which shows that $g(y+z) = g(y) + g(z)$ and $g(ya) = g(y)a$, respectively. This shows that $g\colon W \to V$ is linear, so (2) holds. □

Definition 4. Let F be a field, let V and W be two F-vector spaces of finite dimension n and m, respectively, and let $f\colon V \to W$ be a linear map. The matrix of $f\colon V \to W$ with respect to bases $(v_1, \dots, v_n)$ of V and $(w_1, \dots, w_m)$ of W is the unique $m \times n$ matrix $A = (a_{ij}) \in M_{m,n}(F)$ such that

$$f(v_j) = w_1 a_{1j} + w_2 a_{2j} + \dots + w_m a_{mj}$$

for all $1 \leqslant j \leqslant n$.

We stress that the matrix of $f\colon V \to W$ with respect to the bases $(v_1, \dots, v_n)$ of V and $(w_1, \dots, w_m)$ of W depends not only on the map $f\colon V \to W$ but also on the chosen bases of V and W.

Example 5. The following example is extremely useful to remember. Let $(e_1, \dots, e_n)$ be the standard basis of F^n and let $(v_1, \dots, v_n)$ be another basis of F^n. Then the matrix of $\mathrm{id}_{F^n}\colon F^n \to F^n$ with respect to the basis $(v_1, \dots, v_n)$ of the domain F^n and the basis $(e_1, \dots, e_n)$ of the target F^n is $A = (a_{ij})$, where

$$v_j = e_1 a_{1j} + e_2 a_{2j} + \dots + e_n a_{nj}.$$

In other words, the jth column in the matrix A consists of the coordinates of the vector v_j with respect to the standard basis.

Proposition 6. *Let F be a field; let V and W be two F-vector spaces of finite dimension n and m, respectively; and let $f\colon V \to W$ be a linear map. Let*

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

be the matrix of $f\colon V \to W$ with respect to bases $(v_1, \dots, v_n)$ of V and $(w_1, \dots, w_m)$ of W; and let

$$x = v_1 x_1 + v_2 x_2 + \dots + v_n x_n$$
$$f(x) = w_1 y_1 + w_2 y_2 + \dots + w_m y_m$$

be the unique expressions of $x \in V$ and $f(x) \in W$ as linear combinations of the bases $(v_1, \dots, v_n)$ of V and $(w_1, \dots, w_m)$ of W. In this situation,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Proof. By using the linearity of $f\colon V \to W$ and the definition of the matrix A that represents f with respect to the bases $(v_1, \dots, v_n)$ and $(w_1, \dots, w_m)$, we find

$$f(x) = f(\sum_{j=1}^n v_j x_j) = \sum_{j=1}^n f(v_j) x_j = \sum_{j=1}^n (\sum_{i=1}^m w_i a_{ij}) x_j = \sum_{i=1}^m w_i (\sum_{j=1}^n a_{ij} x_j).$$

Therefore, by the uniqueness of the coordinates of the vector $f(x)$ with respect to the basis $(w_1, \dots, w_m)$ of W, we conclude that for all $1 \leqslant i \leqslant m$,

$$y_i = \sum_{j=1}^n a_{ij} x_j.$$

The proposition now follows from the definition of matrix multiplication. □

Addendum 7. *Let F be a field; let U, V and W be three F-vector spaces of finite dimension p, n, and m, respectively. Let $f\colon V \to W$ and $g\colon U \to V$ be a linear maps, let A be the matrix of $f\colon V \to W$ with respect to bases $(v_1, \dots, v_n)$ of V and $(w_1, \dots, w_m)$ of W, and let B the matrix of $g\colon U \to V$ with respect to a basis $(u_1, \dots, u_p)$ of U and the same basis $(v_1, \dots, v_n)$ of V. In this situation, the matrix C of the composite map $f \circ g\colon U \to W$ with respect to the bases $(u_1, \dots, u_p)$ of U and $(w_1, \dots, w_m)$ is the product matrix $C = AB$.*

Proof. We consider the unique expressions

$$\begin{aligned} x &= u_1x_1 + u_2x_2 + \cdots + u_px_p \\ g(x) &= v_1y_1 + v_2y_2 + \cdots + v_ny_n \\ f(g(x)) &= w_1z_1 + w_2z_2 + \cdots + w_mz_m \end{aligned}$$

of $x \in U$ as a linear combination of the basis $(u_1, \dots, u_p)$ of U; $g(x) \in V$ as a linear combination of the basis $(v_1, \dots, v_n)$ of V; and $f(g(x)) \in W$ as a linear combination of the basis $(w_1, \dots, w_m)$ of W. We write out

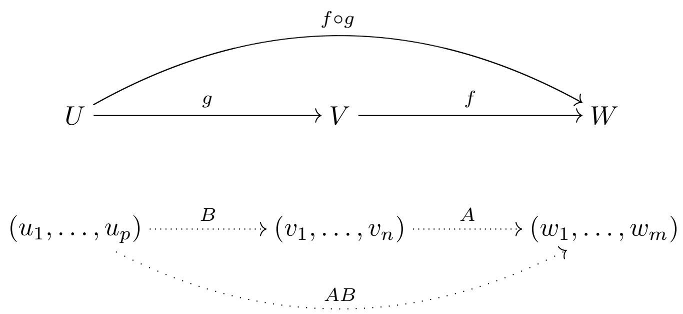
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots a_{np} & \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

and use Proposition 6 in the case of the maps f and g to see that

$$\begin{aligned} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} &= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}, \end{aligned}$$

from which the statement follows by applying Proposition 6 to the map $f \circ g$. We also use that, by the associativity of matrix multiplication, the meaning of the triple product at the right-hand side of the last equality is unambiguous. ☐

Remark 8. The following figure may help memorizing Addendum 7.



We note that the composition of the maps f and g and the product of the matrices A and B that represent them with respect to the indicated bases are formed in the same order. (This is true, because we use right vector spaces; if we were using left vectors spaces, the order would be reversed.)

Corollary 9. *Let F be a field and let $f\colon V \to W$ be a linear map between two F-vector spaces V and W. Let $A \in M_{m,n}(F)$ be matrix of $f\colon V \to W$ with respect to a basis $(v_1, \dots, v_n)$ of V and a basis $(w_1, \dots, w_m)$ of W. In this situation, the following* (1)–(2) *are equivalent.*

(1) *There exists a linear map $g\colon W \to V$ such that $f \circ g = \mathrm{id}_W$ and $g \circ f = \mathrm{id}_V$.*
(2) *There exists a matrix $B \in M_{n,m}(F)$ such that $AB = E_m$ and $BA = E_n$.*

Proof. We first suppose that (1) holds and define $B \in M_{n,m}(F)$ to be the matrix of $g\colon W \to V$ with respect to the bases $(w_1, \dots, w_m)$ of W and $(v_1, \dots, v_n)$ of V. By Addendum 7, $AB \in M_{m,m}(F)$ is the matrix of $\mathrm{id}_W\colon W \to W$ with respect to the *same* basis $(w_1, \dots, w_m)$ of the domain and target, and this matrix, by definition, is the identity matrix E_m. Similarly, Addendum 7 shows that $BA \in M_{n,n}(F)$ is the matrix of $\mathrm{id}_V\colon V \to V$ with respect to the *same* basis $(v_1, \dots, v_n)$ of the domain and target, and, by definition, this matrix is the identity matrix E_n. So (2) holds. Conversely, if (2) holds, then we define $g\colon W \to V$ to be the map that to $y = w_1y_1 + \dots + w_my_m \in W$ assigns $x = v_1x_1 + \dots + v_nx_n \in V$ with

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

where $B = (b_{ji}) \in M_{n,m}(F)$. It is linear, so (1) holds by Addendum 7. □

Remark 10. If the equivalent conditions (1)–(2) in Corollary 9 hold, then $m = n$. Indeed, if $(v_1, \dots, v_n)$ if a basis of V, then $(f(v_1), \dots, f(v_n))$ is a basis of W.

We next apply Addendum 7 to prove the following change-of-basis result, and encourage the reader to memorize the easy proof instead of the more complicated statement.

Corollary 11. *Let F be a field and let $f\colon V \to W$ be a linear map between two F-vector spaces V and W. Let A be the matrix of $f\colon V \to W$ with respect to bases $(v_1, \dots, v_n)$ of V and $(w_1, \dots, w_m)$ of W, and let B be the matrix of $f\colon V \to W$ with respect to bases $(v_1', \dots, v_n')$ of V and $(w_1', \dots, w_m')$ of W. Let P be the matrix of $\mathrm{id}_V\colon V$ with respect to the bases $(v_1', \dots, v_n')$ of the domain and $(v_1, \dots, v_n)$ of the target, and let Q be the matrix of $\mathrm{id}_W\colon W \to W$ with respect to the bases $(w_1', \dots, w_m')$ of the domain and $(w_1, \dots, w_m)$ of the target. In this situation,*

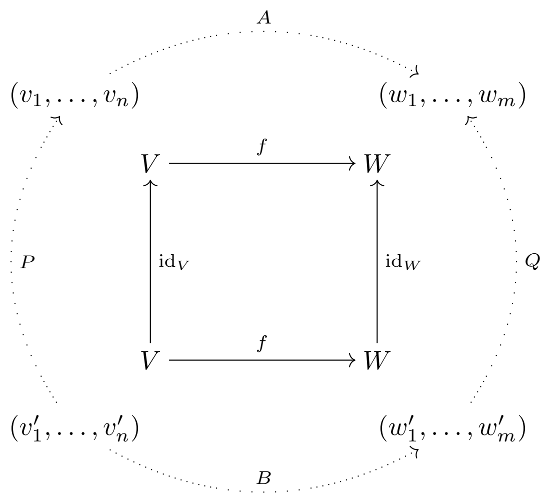
$$B = Q^{-1}AP.$$

Proof. Let C be the matrix of $f\colon V \to W$ with respect to the bases $(v_1', \dots, v_n')$ of V and $(w_1, \dots, w_m)$ of W. Now, on the one hand, since $f = f \circ \mathrm{id}_V\colon V \to W$, we conclude from Addendum 7 that $C = AP$. And on the other hand, since we also have $f = \mathrm{id}_W \circ f\colon V \to W$, we find that $C = QB$. Therefore,

$$QB = AP.$$

The statement follows, since Q is invertible by Corollary 9. □

Remark 12. The following figure may help memorizing the proof of Corollary 11.



We note that, to write the dotted arrow "B" as the composition of the remaining dotted arrows, the arrow "Q" must be reversed, whence $B = Q^{-1}AP$.

Example 13. To illustrate the material above, we will evaluate the matrix of the linear map $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2 & 4 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

with respect to the bases

$$\left(v'_1 = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}, v'_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, v'_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) \quad \text{and} \quad \left(w'_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, w'_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}\right)$$

of $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively. Now, the matrix A of $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ with respect to the standard bases (e_1, e_2, e_3) of $\mathbb{R}^3$ and (e_1, e_2) of $\mathbb{R}^2$ is

$$A = \begin{bmatrix} 2 & 4 & 1 \\ 1 & -1 & 0 \end{bmatrix};$$

the matrix P of $\mathrm{id}_{\mathbb{R}^3}\colon \mathbb{R}^3 \to \mathbb{R}^3$ with respect to the bases (v'_1, v'_2, v'_3) of the domain and (e_1, e_2, e_3) of the target is

$$P = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 1 \end{bmatrix};$$

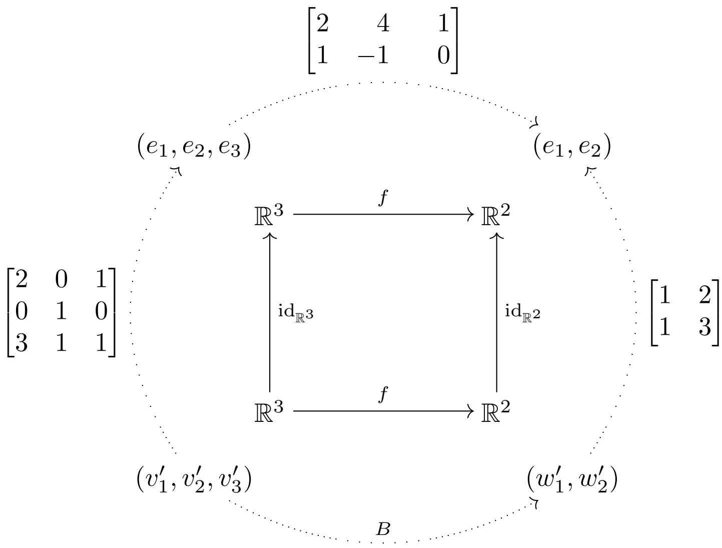
and the matrix Q of $\mathrm{id}_{\mathbb{R}^2}\colon \mathbb{R}^2 \to \mathbb{R}^2$ with respect to the bases (w'_1, w'_2) of the domain and (e_1, e_2) of the target is

$$Q = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}.$$

Therefore, the matrix B of $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ with respect to the bases (v_1', v_2', v_3') of $\mathbb{R}^3$ and (w_1', w_2') of $\mathbb{R}^2$ is

$$B = Q^{-1}AP = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 17 & 7 \\ -5 & -6 & -2 \end{bmatrix}.$$

The following figure illustrates the situation.



To make this kind of calculation and to always make it right, we need only remember two things, namely, (a) the definition of the matrix representing a linear map with respect to given bases of its domain and target, and (b) that the composition of linear maps corresponds to the product of the matrices that represent them.