Special kinds of maps

November 13, 2019

1 Special maps from V to V

The subject of linear algebra is (first, or possibly second) vector spaces and (second, or possibly first) interesting functions ("maps") taking one vector space to another. I want to collect here in one place a number of the ideas we've discussed about maps, to see the formal similiarities among them. All of these ideas make sense for maps between two *different* vector spaces, and all of them are interesting also for *infinite-dimensional* vector spaces; but to simplify, I'll first look at them for a single finite-dimensional vector space. So for this section, I'll always assume

$$V \text{ is a finite-dimensional vector space over any field } F. \tag{1.1}$$

Definition 1.2. A function $T\colon V \to V$ (that is, something that takes one vector v and hands you another vector $T(v)$) is called *linear* if T respects the vector space structure:

$$T(v+w) = T(v) + T(w), \qquad T(av) = aT(v) \qquad (v, w \in V,\ a \in F).$$

I'll write $n = \dim V$, and sometimes choose

$$(e_1, \ldots, e_n) = \text{basis of } V \quad (e_i \in V). \tag{1.3a}$$

Once we've chosen a basis, elements of V can be written uniquely as

$$v = x_1e_1 + x_2e_2 + \cdots + x_ne_n \qquad (x_i \in F). \tag{1.3b}$$

In this way V is identified with $n \times 1$ column vectors

$$v \longleftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{1.3c}$$

. Again in the presence of a basis, giving a linear map T is the same thing as giving n vectors in V

$$t^1 = T(e_1) \quad t^2 = T(e_2) \quad \cdots \quad t^n = T(e_n) \tag{1.3d}$$

or equivalently n column vectors

$$\begin{pmatrix} t_1^1 \\ t_2^1 \\ \vdots \\ t_n^1 \end{pmatrix} \quad \begin{pmatrix} t_1^2 \\ t_2^2 \\ \vdots \\ t_n^2 \end{pmatrix} \cdots \begin{pmatrix} t_1^n \\ t_2^n \\ \vdots \\ t_n^n \end{pmatrix} \tag{1.3e}$$

or equivalently the $n \times n$ matrix

$$\begin{pmatrix} t_1^1 & t_1^2 & \cdots & t_1^n \\ t_2^1 & t_2^2 & \cdots & t_2^n \\ \vdots & \vdots & \ddots & \vdots \\ t_n^1 & t_n^2 & \cdots & t_n^n \end{pmatrix}. \tag{1.3f}$$

What these definitions say is that *applying the linear map T to a vector v means taking a certain linear combination of the columns of T*:

$$T \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = T(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n) = x_1 t^1 + x_2 t^2 + \cdots + x_n t^n. \tag{1.3g}$$

The point of these notes is to talk about various special kinds of linear maps, and what kinds of matrices they correspond to.

Definition 1.4. The linear map $T \in \mathcal{L}(V)$ is called *injective* or *one-to-one* if its null space is zero:

$$\text{Null}(T) = 0.$$

Because of (1.3g), it is equivalent to require *the columns of the matrix of T are linearly independent.*

We could sharpen this equivalence a bit. How *much* T fails to be injective is measured by the *size* of the null space. If $(w_1, w_2, \cdots, w_m)$ is any list of vectors in a vector space W over F, we could define a subspace

$$\begin{aligned} &D(w_1, \ldots, w_m) \subset F^m, \\ &D(w_1, \ldots, w_m) = \{(x_1, \ldots, x_m) \in F^m \mid x_1 w_1 + \cdots + x_m w_m = 0\} \end{aligned} \tag{1.5}$$

The D stands for "dependence:" $(w_1, \ldots, w_m)$ is linearly independent if and only if $D(w_1, \ldots, w_m) = 0$. The larger D is, the more dependent is the list of vectors. What's clear from the definitions is

$$\mathrm{Null}(T) = D(\text{columns of } T). \tag{1.6}$$

That is, *T fails to be injective exactly as much as its columns fail to be linearly independent.*

Surjectivity is "dual."

Definition 1.7. The linear map $T \in \mathcal{L}(V)$ is called *surjective* or *onto* if its range is all of V:

$$\mathrm{Range}(T) = V.$$

Because of (1.3g), it is equivalent to require *the columns of the matrix of T span V.*

Again we can make this more precise:

$$\mathrm{Range}(T) = \mathrm{span}(\text{columns of } T). \tag{1.8}$$

That is, *T fails to be surjective exactly as much as its columns fail to span V.*

A list of n vectors in the n-dimensional space V is linearly independent if and only if it spans V; so Definitions 1.4 and 1.7 are equivalent. That is, a linear map on an n-dimensional vector space is injective if and only if it is surjective. When we work with $\mathcal{L}(V, W)$, injectivity and surjectivity will become two different properties.

Next, suppose that

$$W \subset V, \qquad \dim W = p, \quad q = n - p \tag{1.9a}$$

is a p-dimensional subspace of V. We can choose a basis of W

$$(f_1, \ldots, f_p) = \text{basis of } W \quad (f_i \in W) \tag{1.9b}$$

and extend it to a basis of V.

$$(f_1, \ldots, f_p, g_1, \ldots g_q) = \text{basis of } V. \tag{1.9c}$$

Recall that then

$$(g_1 + W, \ldots g_q + W) = \text{basis of } V/W \tag{1.9d}$$

Definition 1.10. In the setting (1.9), we say that $T \in \mathcal{L}(V)$ *preserves* W (or that W *is an invariant subspace for* T) if

$$Tw \in W, \qquad \text{all } w \in W.$$

If the basis $(f_1, \ldots, f_p, g_1, \ldots, g_q)$ is chosen as in (1.9c), then it is equivalent to require that

$$T(f_1), \ldots, T(f_p) \textit{ all belong to } W.$$

In terms of the matrix of T in the basis $(f_1, \ldots, f_p, g_1, \ldots, g_q)$, this is

$$\textit{the first } p \textit{ columns belong to } F^p \subset F^n;$$

that is, that *the last* q *entries of each of the first* p *columns are all zero.*

The conditions in the definition say that the matrix of T is *block upper triangular*:

$$T = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}, \tag{1.11}$$

with A a $p \times p$ matrix, B a $p \times q$ matrix, 0 the $q \times p$ zero matrix, and D a $q \times q$ matrix. Furthermore

$$A = \text{matrix of } T \text{ restricted to } W \text{ in the basis } (f_1, \ldots, f_p), \tag{1.12}$$

$$D = \text{matrix of } T \text{ on } V/W \text{ in basis } (g_1 + W, \ldots, g_q + W). \tag{1.13}$$

Finally we discuss isometries. For this suppose F is $\mathbb{R}$ or $\mathbb{C}$, and that V is an n-dimensional inner product space over F. Recall that we can fix an *orthonormal* basis

$$(e_1, \ldots, e_n) = \text{orthonormal basis of } V \quad (e_i \in V), \tag{1.14}$$

meaning that

$$\langle e_i, e_j \rangle = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j). \end{cases}$$

Definition 1.15. Recall that $T \in \mathcal{L}(V)$ is called an *isometry* if

$$\langle Tv, Tw \rangle = \langle v, w \rangle \qquad (\text{all } v, w \in V).$$

Assuming that $(e_1, \ldots, e_n)$ is an orthonormal basis of V, it is equivalent to require that *the columns of* T *are also an orthonormal basis of* V. Another equivalent statement is that $T^*T = I$*; that is, that* $T^* = T^{-1}$.

It's a fact from calculus that any vector in $\mathbb{R}^2$ of length 1 is of the form

$$\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} \tag{1.16}$$

for some real number θ that is determined up to a multiple of 2π. It's an elementary geometric fact (still in $\mathbb{R}^2$) that $\begin{pmatrix}-b\\ a\end{pmatrix}$ is perpendicular to $\begin{pmatrix}a\\ b\end{pmatrix}$; and that if a and b are not both zero, then

$$\begin{pmatrix}a\\ b\end{pmatrix}^{\perp} = \left\{ r\begin{pmatrix}-b\\ a\end{pmatrix} \mid r\in\mathbb{R}\right\}.$$

It follows easily that every orthonormal basis of $\mathbb{R}^2$ is of the form

$$\left(\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}, \begin{pmatrix}\mp\sin\theta\\ \pm\cos\theta\end{pmatrix}\right), \tag{1.17}$$

with θ a real number determined up to a multiple of 2π and a single choice of sign. (That is, the two signs $\mp$ and $\pm$ must be opposite.) According to Definition 1.15, this means that every isometry of $\mathbb{R}^2$ is of the form

$$\begin{pmatrix}\cos\theta & \mp\sin\theta\\ \sin\theta & \pm\cos\theta\end{pmatrix}. \tag{1.18}$$

In the case of $\begin{pmatrix}-\\ +\end{pmatrix}$ (signs in the second column) this matrix represents rotation of $\mathbb{R}^2$ counterclockwise by an angle of θ. In the case of $\begin{pmatrix}+\\ -\end{pmatrix}$, the matrix is a reflection fixing the line through $\begin{pmatrix}\cos\theta/2\\ \sin\theta/2\end{pmatrix}$, and acting by -1 on the perpendicular line through $\begin{pmatrix}-\sin\theta/2\\ \cos\theta/2\end{pmatrix}$.

2 Special maps from V to W

In this section I'll see how to extend the ideas from Section 1 to maps between two different vector spaces. So for this section, I'll always assume

$$V \text{ and } W \text{ are finite-dimensional vector spaces over any field } F. \tag{2.1}$$

Definition 2.2. A function $T\colon V \to W$ (that is, something that takes a vector $v \in V$ and hands you a vector $T(v) \in W$) is called *linear* if T respects the vector space structure:

$$T(v+v') = T(v) + T(v'), \qquad T(av) = aT(v) \qquad (v, v' \in V,\ a \in F).$$

I'll write $n = \dim V$, $m = \dim W$ and sometimes choose

$$\begin{aligned}(e_1, \ldots, e_n) &= \text{basis of } V \quad (e_j \in V),\\ (f_1, \ldots, f_m) &= \text{basis of } W \quad (f_i \in W),\end{aligned} \tag{2.3a}$$

Once we've chosen bases, elements of V and W can be written uniquely as

$$\begin{aligned}v &= x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \qquad (x_j \in F),\\ w &= y_1 f_1 + y_2 f_2 + \cdots + y_m f_m \qquad (y_i \in F),\end{aligned} \tag{2.3b}$$

In this way V is identified with $n \times 1$ column vectors, and W with $m \times 1$ column vectors:

$$v \longleftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad w \longleftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \tag{2.3c}$$

. Again in the presence of these bases, giving a linear map T is the same thing as giving n vectors in W

$$t^1 = T(e_1) \quad t^2 = T(e_2) \ \cdots \ t^n = T(e_n) \tag{2.3d}$$

or equivalently n column vectors

$$\begin{pmatrix} t^1_1 \\ t^1_2 \\ \vdots \\ t^1_m \end{pmatrix} \quad \begin{pmatrix} t^2_1 \\ t^2_2 \\ \vdots \\ t^2_m \end{pmatrix} \cdots \begin{pmatrix} t^n_1 \\ t^n_2 \\ \vdots \\ t^n_m \end{pmatrix} \tag{2.3e}$$

or equivalently the $m \times n$ matrix

$$\begin{pmatrix} t^1_1 & t^2_1 & \cdots & t^n_1 \\ t^1_2 & t^2_2 & \cdots & t^n_2 \\ \vdots & \vdots & \ddots & \vdots \\ t^1_m & t^2_m & \cdots & t^n_m \end{pmatrix}. \tag{2.3f}$$

What these definitions say is that *applying the linear map T to a vector v means taking a certain linear combination of the columns of T*:

$$T\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = T(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1t^1 + x_2t^2 + \cdots + x_nt^n. \quad (2.3g)$$

We now begin to extend the definitions from Section 1.

Definition 2.4. The linear map $T \in \mathcal{L}(V, W)$ is called *injective* or *one-to-one* if its null space is zero:

$$\mathrm{Null}(T) = 0.$$

Because of (2.3g), it is equivalent to require *the columns of the matrix of T are linearly independent*. It is also equivalent to require that *T has a left inverse $S \in \mathcal{L}(W, V)$:*

$$ST = I_V.$$

These conditions can be satisfied only if $n = \dim V \leq \dim W = m$.

We could sharpen this equivalence a bit. How *much* T fails to be injective is measured by the *size* of the null space. If $(w_1, w_2, \cdots, w_m)$ is any list of vectors in a vector space W over F, we could define a subspace

$$\begin{aligned} &D(w_1, \ldots, w_m) \subset F^m, \\ &D(w_1, \ldots, w_m) = \{(x_1, \ldots, x_m) \in F^m \mid x_1w_1 + \cdots + x_mw_m = 0\} \end{aligned} \quad (2.5)$$

The D stands for "dependence:" $(w_1, \ldots, w_m)$ is linearly independent if and only if $D(w_1, \ldots, w_m) = 0$. The larger D is, the more dependent is the list of vectors. What's clear from the definitions is

$$\mathrm{Null}(T) = D(\text{columns of } T). \quad (2.6)$$

That is, *T fails to be injective exactly as much as its columns fail to be linearly independent.*

Surjectivity is "dual."

Definition 2.7. The linear map $T \in \mathcal{L}(V, W)$ is called *surjective* or *onto* if its range is all of W:

$$\mathrm{Range}(T) = W.$$

Because of (2.3g), it is equivalent to require *the columns of the matrix of T span W*. It is also equivalent to require that T *has a right inverse* $S \in \mathcal{L}(W,V)$*:*

$$TS = I_W.$$

These conditions can be satisfied only if $m = \dim W \leq \dim V = n$.

Again we can make this more precise:

$$\text{Range}(T) = \text{span}(\text{columns of } T). \tag{2.8}$$

That is, T *fails to be surjective exactly as much as its columns fail to span* W.

Next, suppose that

$$\begin{aligned} V_1 &\subset V, \qquad \dim V_1 = n_1, \quad n_2 = n - n_1 \\ W_1 &\subset W, \qquad \dim W_1 = m_1, \quad m_2 = m - m_1 \end{aligned} \tag{2.9a}$$

are subspaces of V and W. We can choose a basis of V_1

$$(e_1, \ldots, e_{n_1}) = \text{basis of } V_1 \quad (e_j \in V_1) \tag{2.9b}$$

and extend it to a basis of V

$$(e_1, \ldots, e_{n_1}, g_1, \ldots g_{n_2}) = \text{basis of } V. \tag{2.9c}$$

and we can choose a basis

$$(f_1, \ldots, f_{m_1}) = \text{basis of } V_1 \quad (f_i \in V_1) \tag{2.9d}$$

and extend it to a basis of W

$$(f_1, \ldots, f_{m_1}, h_1, \ldots h_{m_2}) = \text{basis of } W. \tag{2.9e}$$

Recall that then

$$(g_1 + V_1, \ldots g_{n_2} + V_1) = \text{basis of } V/V_1, \tag{2.9f}$$

and

$$(h_1 + W_1, \ldots h_{m_2} + W_1) = \text{basis of } W/W_1. \tag{2.9g}$$

Definition 2.10. In the setting (2.9), we say that $T \in \mathcal{L}(V,W)$ *carries* V_1 *to* W_1 if

$$Tv_1 \in W_1, \qquad \text{all } v_1 \in V_1.$$

If the bases

$$(e_1, \dots, e_{n_1}, g_1, \dots, g_{n_2}) \quad \text{and} \quad (f_1, \dots, f_{m_1}, h_1, \dots h_{m_2})$$

are chosen as in (2.9c) and (2.9e), then it is equivalent to require that

$$T(e_1), \dots, T(e_p) \textit{ all belong to } W_1.$$

In terms of the matrix of T in these bases, this is

$$\textit{the first } n_1 \textit{ columns belong to } F^{m_1} \subset F^m;$$

that is, that *the last* m_2 *entries of each of the first* n_1 *columns are all zero.*

The conditions in the definition say that the matrix of T is *block upper triangular*:

$$T = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}, \tag{2.11}$$

with A an $m_1 \times n_1$ matrix, B an $m_1 \times n_2$ matrix, 0 the $m_2 \times n_1$ zero matrix, and D an $m_2 \times n_2$ matrix. Furthermore

$$\begin{gathered} A = \text{matrix of } T|_{V_1} \in \mathcal{L}(V_1, W_1) \text{ in bases} \\ (e_1, \dots, e_{m_1}), \quad (f_1, \dots, f_{m_1}), \end{gathered} \tag{2.12}$$

$$\begin{gathered} D = \text{matrix of } T|_{V/V_1} \in \mathcal{L}(V/V_1, W/W_1) \text{ in bases} \\ (g_1 + V_1, \dots, g_{n_2} + V_1), \quad (h_1 + W_1, \dots, h_{m_2} + W_1). \end{gathered} \tag{2.13}$$

Finally we discuss isometries. For this suppose F is $\mathbb{R}$ or $\mathbb{C}$, and that V and W are inner product spaces over F, still of dimensions n and m. Recall that we can fix *orthonormal* bases

$$\begin{aligned} (e_1, \dots, e_n) &= \text{orthonormal basis of } V \quad (e_j \in V), \\ (f_1, \dots, f_m) &= \text{orthonormal basis of } W \quad (f_i \in W), \end{aligned} \tag{2.14}$$

meaning that

$$\langle e_j, e_{j'} \rangle = \begin{cases} 1 & (j = j') \\ 0 & (j \neq j') \end{cases}$$

and

$$\langle f_i, f_{i'} \rangle = \begin{cases} 1 & (i = i') \\ 0 & (i \neq i') \end{cases}$$

Definition 2.15. A linear map of inner product spaces $T \in \mathcal{L}(V,W)$ is called an *isometry* if

$$\langle Tv, Tv' \rangle = \langle v, v' \rangle \qquad (\text{all } v, v' \in V).$$

Assuming that $(e_1, \ldots, e_n)$ is an orthonormal basis of V, it is equivalent to require that *the columns of T are also an orthonormal set in W*. Another equivalent statement is that $T^*T = I_V$; that is, that T^* *is a left inverse of* T.

Because an orthonormal set is necessarily linearly independent, an isometry is automatically injective; the left inverse that must exist may be taken to be T^*. In particular, *isometries can exist in $\mathcal{L}(V,W)$ only if $n = \dim V \leq \dim W = m$.*

Any isometry from $\mathbb{R}^1$ to $\mathbb{R}^2$ is a 2×1 matrix

$$T = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}.$$