Chapter 4

Linear maps between vector spaces

We return to the setting of vector spaces in order to define linear maps between them. We will see that these maps can be represented by matrices, decide when two matrices represent the same linear map, and give another proof of the canonical form for equivalence.

4.1 Definition and basic properties

Definition 4.1. Let V and W be vector spaces over a field $\mathbb{K}$. A function α from V to W is a *linear map* if it preserves addition and scalar multiplication, that is, if

- $\alpha(v_1 + v_2) = \alpha(v_1) + \alpha(v_2)$ for all $v_1, v_2 \in V$;
- $\alpha(cv) = c\alpha(v)$ for all $v \in V$ and $c \in \mathbb{K}$.

Remark 4.2. (a) We can combine the two conditions into one as follows:

$$\alpha(c_1v_1 + c_2v_2) = c_1\alpha(v_1) + c_2\alpha(v_2).$$

(b) In other literature the term "linear transformation" is often used instead of "linear map".

Definition 4.3. Let $\alpha : V \to W$ be a linear map. The *image* of α is the set

$$\mathrm{Im}(\alpha) = \{w \in W : w = \alpha(v) \text{ for some } v \in V\},$$

and the *kernel* of α is

$$\mathrm{Ker}(\alpha) = \{v \in V : \alpha(v) = \mathbf{0}\}.$$

Proposition 4.4. *Let $\alpha : V \to W$ be a linear map. Then the image of α is a subspace of W and the kernel is a subspace of V.*

Proof. We have to show that each subset is closed under addition and scalar multiplication, and is non-empty. Non-emptiness is immediate: the zero vector $\mathbf{0}$ is in both $\mathrm{Im}(\alpha)$ and $\mathrm{Ker}(\alpha)$ since $\alpha(\mathbf{0}) = \mathbf{0}$.

Suppose that w_1, w_2 are vectors in the image of α. By definition of $\mathrm{Im}(\alpha)$, there exist $v_1, v_2 \in V$ such that $w_1 = \alpha(v_1)$ and $w_2 = \alpha(v_2)$. Then

$$w_1 + w_2 = \alpha(v_1) + \alpha(v_2) = \alpha(v_1 + v_2),$$

by linearity of α. It follows that $w_1 + w_2 \in \mathrm{Im}(\alpha)$. Now suppose $w \in \mathrm{Im}(\alpha)$ and $c \in \mathbb{K}$. By definition of $\mathrm{Im}(\alpha)$, there exists $v \in V$ such that $w = \alpha(v)$. Then

$$cw = c\alpha(v) = \alpha(cv),$$

demonstrating that $cw \in \mathrm{Im}(\alpha)$.

Next suppose v_1, v_2 are vectors in the kernel of α. By definition of $\mathrm{Ker}(\alpha)$, we know that $\alpha(v_1) = \alpha(v_2) = \mathbf{0}$. Thus

$$\alpha(v_1 + v_2) = \alpha(v_1) + \alpha(v_2) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

from which it follows that $v_1 + v_2 \in \mathrm{Ker}(\alpha)$. Finally, suppose $v \in \mathrm{Ker}(\alpha)$ and $c \in \mathbb{K}$. Then $\alpha(v) = \mathbf{0}$ and

$$\alpha(cv) = c\alpha(v) = c\,\mathbf{0} = \mathbf{0},$$

demonstrating that $cv \in \mathrm{Ker}(\alpha)$. □

Definition 4.5. We define the *rank* of α to be $\varrho(\alpha) = \dim(\mathrm{Im}(\alpha))$ and the *nullity* of α to be $\nu(\alpha) = \dim(\mathrm{Ker}(\alpha))$. (We use the Greek letters 'rho' and 'nu' here to avoid confusing the rank of a linear map with the rank of a matrix, though they will turn out to be closely related!)

Theorem 4.6 (Rank–Nullity Theorem)**.** *Let $\alpha : V \to W$ be a linear map. Then $\varrho(\alpha) + \nu(\alpha) = \dim(V)$.*

Proof. Choose a basis $u_1, u_2, \ldots, u_q$ for $\mathrm{Ker}(\alpha)$, where $q = \dim(\mathrm{Ker}(\alpha)) = \nu(\alpha)$. The vectors $u_1, \ldots, u_q$ are linearly independent vectors of V, so we can add further vectors to get a basis for V, say $u_1, \ldots, u_q, v_1, \ldots, v_s$, where $q + s = \dim(V)$.

We claim that the vectors $\alpha(v_1), \ldots, \alpha(v_s)$ form a basis for $\mathrm{Im}(\alpha)$. We have to show that they are linearly independent and spanning.

Linearly independent: Suppose that $c_1\alpha(v_1) + \cdots + c_s\alpha(v_s) = \mathbf{0}$. We need to show that $c_1 = \cdots = c_s = 0$. Applying the linear map α we have

$$\alpha(c_1v_1 + \cdots + c_sv_s) = c_1\alpha(v_1) + \cdots + c_s\alpha(v_s) = \mathbf{0},$$

so that $c_1v_1 + \cdots + c_sv_s \in \mathrm{Ker}(\alpha)$. The vector $c_1v_1 + \cdots + c_sv$ can be expressed in terms of the basis for $\mathrm{Ker}(\alpha)$:

$$c_1v_1 + \cdots + c_sv_s = a_1u_1 + \cdots + a_qu_q,$$

whence

$$-a_1u_1 - \cdots - a_qu_q + c_1v_1 + \cdots + c_sv_s = \mathbf{0}.$$

But the list $(u_1, \ldots, u_q, v_1, \ldots, v_s)$ is a basis for V, and hence is linearly independent. It follows that $c_1 = \cdots = c_s = 0$ (and incidentally $a_1 = \cdots = a_q = 0$), as required.

Spanning: Take any vector in $\mathrm{Im}(\alpha)$, say w. We need to show that $w \in \langle \alpha(v_1), \dots, \alpha(v_s) \rangle$. Since $w \in \mathrm{Im}(\alpha)$ we know that $w = \alpha(v)$ for some $v \in V$. Write v in terms of the basis for V:

$$v = a_1 u_1 + \cdots + a_q u_q + c_1 v_1 + \cdots + c_s v_s$$

for some $a_1, \dots, a_q, c_1, \dots, c_s$. Applying α, we get

$$\begin{aligned} w &= \alpha(v) \\ &= a_1 \alpha(u_1) + \cdots + a_q \alpha(u_q) + c_1 \alpha(v_1) + \cdots + c_s \alpha(v_s) \\ &= c_1 \alpha(v_1) + \cdots + c_s \alpha(v_s), \end{aligned}$$

where we used the fact that $u_i \in \mathrm{Ker}(\alpha)$ and hence $\alpha(u_i) = \mathbf{0}$. So the vectors $\alpha(v_1), \dots, \alpha(v_s)$ span $\mathrm{Im}(\alpha)$.

Thus, $\varrho(\alpha) = \dim(\mathrm{Im}(\alpha)) = s$. Since $\nu(\alpha) = q$ and $q + s = \dim(V)$, the theorem is proved. □

4.2 Representation by matrices

We come now to the second role of matrices in linear algebra: **they represent linear maps between vector spaces**.

Let $\alpha : V \to W$ be a linear map, where $\dim(V) = n$ and $\dim(W) = m$. Let $v_1, \dots, v_n$ be a basis for V and $w_1, \dots, w_m$ a basis for W. Then for $j = 1, \dots, n$, the vector $\alpha(v_j)$ belongs to W, so we can write it as a linear combination of $w_1, \dots, w_m$.

Definition 4.7. The matrix representing the linear map $\alpha : V \to W$ relative to the bases $(v_1, \dots, v_n)$ for V and $(w_1, \dots, w_m)$ for W is the $m \times n$ matrix whose (i, j) entry is a_{ij}, where

$$\alpha(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

for $j = 1, \dots, n$. (The indices on the right hand side are reversed from what you might expect by analogy with matrix multiplication, but it will all turn out right in the end!)

In practice this means the following. Take $\alpha(v_j)$ and write it as a as a linear combination $\alpha(v_j) = a_{1j} w_1 + \cdots + a_{mj} w_m$ of basis vectors of W. Then the column vector $\begin{bmatrix} a_{1j} & a_{2j} & \cdots & a_{mj} \end{bmatrix}^\top$ is the jth column of the matrix representing α. So, for example, if $n = 3$, $m = 2$, and

$$\alpha(v_1) = w_1 + w_2, \quad \alpha(v_2) = 2w_1 + 5w_2, \quad \alpha(v_3) = 3w_1 - w_2,$$

then the matrix representing α is

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & -1 \end{bmatrix}.$$

Now the most important thing about this representation is that the action of α is now easily described:

Proposition 4.8. *Let $\alpha : V \to W$ be a linear map. Choose bases $\mathcal{B}$ for V and $\mathcal{B}'$ for W and let A be the matrix representing α with respect to these bases. Then*

$$[\alpha(v)]_{\mathcal{B}'} = A[v]_{\mathcal{B}}.$$

Proof. Let $\mathcal{B} = (v_1, \ldots, v_n)$ be the basis for V, and $\mathcal{B}' = (w_1, \ldots, w_m)$ the basis for W. Suppose $v = \sum_{j=1}^{n} c_j v_j \in V$, so that in coordinates

$$[v]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Then

$$\alpha(v) = \sum_{j=1}^{n} c_j \alpha(v_j) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} w_i \sum_{j=1}^{n} a_{ij} c_j,$$

so the ith coordinate of $[\alpha(v)]_{\mathcal{B}'}$ is $\sum_{j=1}^{n} a_{ij} c_j$, which is precisely the ith coordinate in the matrix product $A[v]_{\mathcal{B}}$. □

In our example, if $v = 2v_1 + 3v_2 + 4v_3$, so that the coordinate representation of v relative to the basis (v_1, v_2, v_3) is $\begin{bmatrix} 2 & 3 & 4 \end{bmatrix}^\top$, then

$$[\alpha(v)]_{\mathcal{B}'} = A[v]_{\mathcal{B}} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 20 \\ 13 \end{bmatrix}.$$

The column vector on the right gives the coordinate representation of $\alpha(v)$ relative to the basis (w_1, w_2), that is, $\alpha(v) = 20w_1 + 13w_2$.

Addition and multiplication of linear maps correspond to addition and multiplication of the matrices representing them.

Definition 4.9. Let α and β be linear maps from V to W. Define their sum $\alpha + \beta$ by the rule

$$(\alpha + \beta)(v) = \alpha(v) + \beta(v)$$

for all $v \in V$. It is routine to check that $\alpha + \beta$ is a linear map.

Proposition 4.10. *If α and β are linear maps represented relative to some basis by matrices A and B, respectively, then $\alpha + \beta$ is represented by the matrix $A + B$, relative to the same basis.*

The proof of this is not too difficult: just apply the definitions as in the Proof of Proposition 4.12 below.

Definition 4.11. Let U, V and W be vector spaces over $\mathbb{K}$, and let $\alpha : U \to V$ and $\beta : V \to W$ be linear maps. The product $\beta\alpha$ is the function $U \to W$ defined by the rule

$$(\beta\alpha)(u) = \beta(\alpha(u))$$

for all $u \in U$. Again it is routine to check that $\beta\alpha$ is a linear map. Note that the order is important: we take a vector $u \in U$, apply α to it to get a vector in V, and then apply β to get a vector in W. So $\beta\alpha$ means "apply α, then β".

Proposition 4.12. *If $\alpha : U \to V$ and $\beta : V \to W$ are linear maps represented by matrices A and B respectively, then $\beta\alpha$ is represented by the matrix BA.*

Proof. Suppose linear maps α and β are represented by matrices A and B relative to bases $\mathcal{B}$ of U, $\mathcal{B}'$ of V, and $\mathcal{B}''$ of W. Then

$$[(\beta\alpha)u]_{\mathcal{B}''} = [\beta(\alpha(u))]_{\mathcal{B}''} = B\,[\alpha(u)]_{\mathcal{B}'} = B(A[u]_{\mathcal{B}}) = (BA)[u]_{\mathcal{B}},$$

where we have used, in turn, the definition of product of maps, Proposition 4.8 (twice) and associatively of matrix multiplication. □

Remark Let $l = \dim(U)$, $m = \dim(V)$ and $n = \dim(W)$, then A is $m \times l$, and B is $n \times m$; so the product BA is defined, and is $n \times l$, which is the right size for a matrix representing a map from an l-dimensional to an n-dimensional space.

The significance of all this is that the strange rule for multiplying matrices is chosen so as to make Proposition 4.12 hold. The definition of multiplication of linear maps is the natural one (composition), and we could then say: what definition of matrix multiplication should we choose to make the Proposition valid? We would find that the usual definition was forced upon us.

4.3 Change of basis

The matrix representing a linear map depends on the choice of bases we used to represent it. We briefly discuss what happens if we change the basis.

Recall the notion of *transition matrix* from Chapter 1. If $\mathcal{B} = (v_1, \dots, v_n)$ and $\mathcal{B}' = (v'_1, \dots, v'_n)$ are two bases for a vector space V of dimension n, then the transition matrix $P_{\mathcal{B},\mathcal{B}'}$ is the matrix whose jth column is the coordinate representation of v'_j relative to the basis $\mathcal{B}$. We saw that

$$[v]_{\mathcal{B}} = P_{\mathcal{B},\mathcal{B}'}[v]_{\mathcal{B}'},$$

where $[v]_{\mathcal{B}}$ is the coordinate representation of an arbitrary vector v relative to the basis $\mathcal{B}$, and similarly for $\mathcal{B}'$. The transition matrix $P_{\mathcal{B}',\mathcal{B}}$ that transforms $[v]_{\mathcal{B}}$ back to $[v]_{\mathcal{B}'}$ is just the inverse of the matrix $P_{\mathcal{B},\mathcal{B}'}$.

Proposition 4.13. *Let $\alpha : V \to W$ be a linear map represented by matrix A relative to the bases $\mathcal{B}$ for V and $\mathcal{C}$ for W, and by the matrix A' relative to the bases $\mathcal{B}'$ for V and $\mathcal{C}'$ for W. If $P = P_{\mathcal{C}',\mathcal{C}}$ and $Q = P_{\mathcal{B},\mathcal{B}'}$ are transition matrices relating the unprimed to the primed bases, then*

$$A' = PAQ.$$

Proof. At a high level the claim seems reasonable. Suppose we apply the matrix A' to a coordinate representation of some vector relative to the primed basis for V. Multiplication by Q will transform from the primed to unprimed basis, multiplication by A will apply the linear transformation relative to the unprimed bases, and finally P will transform back to the primed basis.

We just need to write that scheme down in symbols, which is not too difficult:

$$(PAQ)[v]_{\mathcal{B}'} = PA(Q\,[v]_{\mathcal{B}'}) = P(A\,[v]_{\mathcal{B}}) = P[\alpha(v)]_{\mathcal{C}} = [\alpha(v)]_{\mathcal{C}'}.$$

So $A' = PAQ$ is the representation of the linear map α relative to the primed bases. □

In practical terms, the above result is needed for explicit calculations. For theoretical purposes its importance lies the following corollary. Recall that two matrices A and B are equivalent if B is obtained from A by multiplying on the left and right by invertible matrices.

Proposition 4.14. *Two matrices represent the same linear map with respect to different bases if and only if they are equivalent.*

Proof. We saw in Proposition 4.13 that two matrices A and A' representing the same linear map are equivalent. This is the "only if" direction.

Nothing in the rest of the course rests on the "if " direction, but we include the proof for completeness. Suppose A and B are equivalent $n \times n$ matrices. Then there exist invertible matrices P and Q with $A' = PAQ$. The idea is that we can view the invertible matrix Q as a transition matrix $P_{\mathcal{B},\mathcal{B}'}$. Since P is invertible it has rank n and hence its columns are linearly independent. The columns can thus be interpreted as the basis vectors $\mathcal{B}'$ written in terms of the basis vectors $\mathcal{B}$. This is exactly the transition matrix from $\mathcal{B}'$ to $\mathcal{B}$. Similarly, P can be interpreted as a transition matrix $P_{\mathcal{C}',\mathcal{C}}$. Thus if A represents a linear map α with respect to the bases $\mathcal{B}$ and $\mathcal{C}$ then A' represents α with respect to $\mathcal{B}'$ and $\mathcal{C}'$. □

4.4 Canonical form revisited

We return Theorem 2.3 about canonical forms for equivalence with a view to showing that rank of a linear map and rank of a matrix are essentially the same thing.

Theorem 4.15. *Let $\alpha : V \to W$ be a linear map of rank $r = \varrho(\alpha)$. Then there are bases for V and W such that the matrix representing α is, in block form,*

$$\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}.$$

Proof. As in the proof of Theorem 4.6, choose a basis $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$ for V such that $v_{r+1}, \ldots, v_n$ is a basis for $\mathrm{Ker}(\alpha)$. (We can do this by choosing the basis $v_{r+1}, \ldots, v_n$ of $\mathrm{Ker}(\alpha)$ first, and then extending it to a basis for the whole space V.)

As we saw earlier, $w_1 = \alpha(v_1), \ldots, w_r = \alpha(v_r)$ is a basis for $\mathrm{Im}(\alpha)$, and can be extended to a basis $w_1, \ldots, w_r, w_{r+1}, \ldots, w_m$ of W. We have

$$\alpha(v_i) = \begin{cases} w_i, & \text{if } 1 \le i \le r; \\ \mathbf{0}, & \text{otherwise,} \end{cases}$$

so the matrix of α relative to these bases is

$$\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$$

as claimed. □

We recognise the matrix in the theorem as the canonical form for equivalence. It is now not difficult to see that rank of a linear map and rank of a matrix are consistent.

Corollary 4.16. *Suppose $\alpha : V \to W$ is a linear map of rank r. For any choice of bases $\mathcal{B}$ for V and $\mathcal{B}'$ for W, the rank of the matrix representing α relative to $\mathcal{B}$ and $\mathcal{B}'$ is also r.*

Proof. We know from Theorem 4.15 that there is some choice of bases for which the matrix A representing α takes the canonical form. In this case the rank of the linear map α and the matrix A certainly agree. Any other matrix A' representing α will be equivalent to A by Proposition 4.13. Equivalent matrices have the same rank, so the rank of A' is also r. □

So how many equivalence classes of $m \times n$ matrices are there, for given m and n? The rank of such a matrix can take any value from 0 up to the minimum of m and n; so the number of equivalence classes is $\min\{m, n\} + 1$. Thus the number of distinct linear maps from a vector space of dimension n to one of dimension m is also $\min\{m, n\} + 1$.