Queen Mary
University of London

MTH6140 Linear Algebra II

Notes 3 **28th October 2010**

---

3 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.

3.1 Definition and basic properties

Definition 3.1 Let V and W be vector spaces over a field K. A function T from V to W is a *linear map* if it preserves addition and scalar multiplication, that is, if

- $T(v+v') = T(v)+T(v')$ for all $v, v' \in V$;
- $T(cv) = cT(v)$ for all $v \in V$ and $c \in K$.

Remark We can combine the two conditions into one as follows:

$$T(cv+c'v') = cT(v)+c'T(v').$$

Definition 3.2 Let $T : V \to W$ be a linear map. The *image* of T is the set

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

and the *kernel* of T is

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

Proposition 3.1 Let $T : V \to W$ be a linear map. Then the image of T is a subspace of W and the kernel is a subspace of V.

Proof We have to show that both $\mathrm{Im}(T)$ and $\mathrm{Ker}(T)$ are closed under addition and scalar multiplication. For the image, if $w, w' \in \mathrm{Im}(T)$ then we can find $v, v' \in V$ such that $w = T(v)$ and $w' = T(v')$. Then

$$w + w' = T(v) + T(v') = T(v + v') \in \mathrm{Im}(T),$$

and if $c \in K$ then

$$cw = cT(v) = T(cv),$$

so $\mathrm{Im}(T)$ is indeed a subspace of W.

For the kernel, if $v, v' \in \mathrm{Ker}(T)$ then $T(v) = T(v') = 0$. Therefore

$$T(v + v') = T(v) + T(v') = 0 + 0 = 0,$$

so $v + v' \in \mathrm{Ker}(T)$. Also, if $c \in K$ then

$$T(cv) = cT(v) = c0 = 0,$$

so $cv \in \mathrm{Ker}(T)$.

Definition 3.3 We define the *rank* of T to be $\mathrm{rank}(T) = \dim(\mathrm{Im}(T))$ and the *nullity* of T to be $\mathrm{nul}(T) = \dim(\mathrm{Ker}(T))$. (Recall that in Chapter 2 we defined the rank of a matrix. This will turn out to be closely related to the rank of a linear map!)

Theorem 3.2 (Rank–Nullity Theorem) Let $T : V \to W$ be a linear map. Then $\mathrm{rank}(T) + \mathrm{nul}(T) = \dim(V)$.

Proof Choose a basis $u_1, u_2, \ldots, u_q$ for $\mathrm{Ker}(T)$, where $q = \dim(\mathrm{Ker}(T)) = \mathrm{nul}(T)$. The vectors $u_1, \ldots, u_q$ are linearly independent vectors of V, so by Corollary 1.3(b) 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 $T(v_1), \ldots, T(v_s)$ form a basis for $\mathrm{Im}(T)$. We have to show that they are linearly independent and spanning.

Linearly independent: Suppose that $c_1T(v_1) + \cdots + c_sT(v_s) = 0$. Then $T(c_1v_1 + \cdots + c_sv_s) = 0$, so that $c_1v_1 + \cdots + c_sv_s \in \mathrm{Ker}(T)$. But then this vector can be expressed in terms of the basis for $\mathrm{Ker}(T)$:

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

whence

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

But the us and vs form a basis for V, so they are linearly independent. So this equation implies that all the as and cs are zero. The fact that $c_1 = \cdots = c_s = 0$ shows that the vectors $T(v_1), \ldots, T(v_s)$ are linearly independent.

Spanning: Take any vector in $\mathrm{Im}(T)$, say w. Then $w = T(v)$ for some $v \in V$. Write v in terms of the basis for V:

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

for some $a_1, \ldots, a_q, c_1, \ldots, c_s$. Applying T, we get

$$\begin{aligned} w &= T(v) \\ &= a_1T(u_1) + \cdots + a_qT(u_q) + c_1T(v_1) + \cdots + c_sT(v_s) \\ &= c_1w_1 + \cdots + c_sw_s, \end{aligned}$$

since $T(u_i) = 0$ (as $u_i \in \mathrm{Ker}(T)$) and $T(v_i) = w_i$. So the vectors $w_1, \ldots, w_s$ span $\mathrm{Im}(T)$.

Thus, $\mathrm{rank}(T) = \dim(\mathrm{Im}(T)) = s$. Since $\mathrm{nul}(T) = q$ and $q + s = \dim(V)$, the theorem is proved.

3.2 Representation by matrices

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

Let $T : V \to W$ be a linear map, where $\dim(V) = m$ and $\dim(W) = n$. As we saw in Chapter 1, we can take V and W in their coordinate representation: $V = K^m$ and $W = K^n$ (the elements of these vector spaces being represented as column vectors). Let $v_1, \ldots, v_m$ be the standard basis for V (so that v_i is the vector with ith coordinate 1 and all other coordinates zero), and $w_1, \ldots, w_n$ the standard basis for W. Then for $i = 1, \ldots, m$, the vector $T(v_i)$ belongs to W, so we can write it as a linear combination of $w_1, \ldots, w_n$.

Definition 3.4 The matrix representing the linear map $T : V \to W$ relative to the bases $B = (v_1, \ldots, v_m)$ for V and $C = (w_1, \ldots, w_n)$ for W is defined to be the $n \times m$ matrix A whose (i, j) entry is A_{ij}, where

$$T(v_i) = \sum_{j=1}^{n} A_{ji}w_j$$

for $i = 1, \ldots, n$.

In practice this means the following. Take $T(v_i)$ and write it as a column vector

$$\begin{bmatrix} A_{1i} \\ A_{2i} \\ \vdots \\ A_{ni} \end{bmatrix}.$$

This vector is the ith column of the matrix representing T. So, for example, if $m = 3$, $n = 2$, and

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

then the vectors $T(v_i)$ as column vectors are

$$T(v_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad T(v_2) = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad T(v_3) = \begin{bmatrix} 3 \\ -1 \end{bmatrix},$$

and so the matrix representing T is

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

The most important thing about this representation is that the action of T is now easily described:

Proposition 3.3 Let $T : V \to W$ be a linear map. Choose bases for V and W and let A be the matrix representing T. Then, if we represent vectors of V and W as column vectors relative to these bases, we have

$$T(v) = Av.$$

Proof Let $v_1, \ldots, v_m$ be the basis for V, and $w_1, \ldots, w_n$ for W. Take $v = \sum_{i=1}^{m} c_i v_i \in V$, so that in coordinates

$$v = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}.$$

Then

$$T(v) = T\left(\sum_{i=1}^{m} c_i v_i\right) = \sum_{i=1}^{m} c_i T(v_i) = \sum_{i=1}^{m} c_i \sum_{j=1}^{n} A_{ji} w_j = \sum_{j=1}^{n} \sum_{i=1}^{m} c_i A_{ji} w_j,$$

so the jth coordinate of $T(v)$ is $\sum_{i=1}^{m} A_{ji} c_i$, which is precisely the jth coordinate in the matrix product Av.

In our example, if $v = 2v_1 + 3v_2 + 4v_3 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$, then

$$T(v) = Av = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 20 \\ 13 \end{bmatrix}.$$

One of the important things about this representation is that addition and composition of linear maps correspond to addition and multiplication of the matrices representing them.

Definition 3.5 Let S and T be linear maps from V to W. Define their sum $S+T$ by the rule

$$(S+T)(v) = S(v)+T(v)$$

for all $v \in V$. It is easy to check that $S+T$ is a linear map.

Proposition 3.4 If S and T are linear maps represented by matrices A and B respectively, then $S+T$ is represented by the matrix $A+B$.

The proof of this is not difficult: just use the definitions.

Definition 3.6 Let U,V,W be vector spaces over K, and let $S : U \to V$ and $T : V \to W$ be linear maps. The composition $T \circ S$ (also sometimes denoted by TS) is the function $U \to W$ defined by the rule

$$(T \circ S)(u) = T(S(u))$$

for all $u \in U$. Again it is easily checked that $T \circ S$ is a linear map. Note that the order is important: we take a vector $u \in U$, apply S to it to get a vector in V, and then apply T to get a vector in W. So $T \circ S$ means "apply S, then T".

Proposition 3.5 If $S : U \to V$ and $T : V \to W$ are linear maps represented by matrices A and B respectively, then $T \circ S$ is represented by the matrix BA.

Again the proof is tedious but not difficult.

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 3.5 hold. We could 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.

3.3 Change of basis

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

Remember the notion of *transition matrix* from Chapter 1. If $B = (v_1, \ldots, v_m)$ and $B' = (v'_1, \ldots, v'_m)$ are two bases for a vector space V, the transition matrix $P_{B,B'}$ is the matrix whose jth column is the coordinate representation of v'_j in the basis B. Then (see Proposition 1.5) we have

$$[v]_B = P[v]_{B'},$$

where $[v]_B$ is the coordinate representation of an arbitrary vector in the basis B, and similarly for B'. Recall that the inverse of $P_{B,B'}$ is $P_{B',B}$ (see Corollary 1.6 (b)). Let p_{ij} denote the (i, j) entry of $P = P_{B,B'}$.

Now let $C = (w_1, \ldots, w_n)$ and $C' = (w'_1, \ldots, w'_n)$ be two different bases for the same vector space W, with transition matrix $Q_{C,C'}$ and inverse $Q_{C',C}$. Let $Q = Q_{C,C'}$ and let $R = Q_{C',C}$ be its inverse, with (i, j) entry denoted by r_{ij}.

Let T be a linear map from V to W. Then T is represented by a matrix A using the bases B and C, and by a matrix A' using the bases B' and C'. What is the relation between A and A'?

We just do it and see. To get A', we have to represent the vectors $T(v'_i)$ in the basis C'. We have

$$v'_j = \sum_{i=1}^{m} p_{ij} v_i,$$

so

$$\begin{aligned}
T(v'_j) &= \sum_{i=1}^{m} p_{ij} T(v_i) \\
&= \sum_{i=1}^{m} \sum_{k=1}^{m} p_{ij} A_{ki} w_k \\
&= \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{l=1}^{n} p_{ij} A_{ki} r_{lk} w'_l.
\end{aligned}$$

This means, on turning things around, that

$$(A')_{lj} = \sum_{k=1}^{n} \sum_{i=1}^{m} r_{lk} A_{ki} p_{ij},$$

so, according to the rules of matrix multiplication,

$$A' = RAP = Q^{-1}AP.$$

So we have proved the following result:

Proposition 3.6 Let $T : V \to W$ be a linear map represented by matrix A relative to the bases B for V and C for W, and by the matrix A' relative to the bases B' for V and C' for W. If $P = P_{B,B'}$ and $Q = P_{C,C'}$ are the transition matrices from the unprimed to the primed bases, then

$$A' = Q^{-1}AP.$$

This is rather technical; you need it for explicit calculations, but for theoretical purposes the importance is the following corollary. Recall (see Definition 2.5) 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 3.7 Two matrices represent the same linear map with respect to different bases if and only if they are equivalent.

This holds because transition matrices are always invertible, and any invertible matrix can be regarded as a transition matrix.

3.4 Canonical form revisited

Now we can give a simpler proof of Theorem 2.3 about canonical form for equivalence. First, we make the following observation.

Theorem 3.8 Let $T : V \to W$ be any linear map. Then there are bases for V and W such that the matrix representing T is, in block form,

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

where $r = \mathrm{rank}(T)$ is the rank of T.

Proof As in the proof of Theorem 3.2, choose a basis $u_1, \ldots, u_s$ for $\mathrm{Ker}(T)$, and extend to a basis $u_1, \ldots, u_s, v_1, \ldots, v_r$ for V. Then $T(v_1), \ldots, T(v_r)$ is a basis for $\mathrm{Im}(T)$, and so can be extended to a basis $T(v_1), \ldots, T(v_r), x_1, \ldots, x_t$ for W. Now we will use the bases

$$\begin{aligned} v_1, \ldots, v_r, v_{r+1} = u_1, \ldots, v_{r+s} = u_s \quad &\text{for} \quad V, \\ w_1 = T(v_1), \ldots, w_r = T(v_r), w_{r+1} = x_1, \ldots, w_{r+s} = x_s \quad &\text{for} \quad W. \end{aligned}$$

We have

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

so the matrix of T 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. Combining Theorem 3.8 with Proposition 3.7, we see:

Theorem 3.9 A matrix of rank r is equivalent to the matrix

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

We also see, by the way, that the rank of a linear map (that is, the dimension of its image) is equal to the rank of any matrix which represents it. So all our definitions of rank agree!

The conclusion is that

> two matrices are equivalent if and only if they have the same rank.

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 $\min\{m,n\}+1$.