LINEAR ALGEBRA- LECTURE 17

1. Matrix of a linear transformation

Recall that a linear transformation[1] $T : V \longrightarrow W$ is a function that satisfies

$$T(u+v) = T(u) + T(v)$$
$$T(au) = aT(u)$$

for all $u, v \in V$ and all scalars a. We have already seen several examples of linear transformations. Here are two more.

Example 1.1. Let A be a $m \times n$ matrix with entries in a field F. This matrix gives rise to a function, which we again denote by A

$$A : F^n \longrightarrow F^m$$

as follows. We set

$$A(X) = AX = x_1A_1 + x_2A_2 + \cdots + x_nA_n \tag{1.1.1}$$

where

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in F^n$$

is a column vector, $A_i \in F^m$ denotes the i-th column of A and AX denotes the matrix product. In particular note that the product AX is a linear combination of the column vectors of A. This shows that the vector AX belongs to the column space of the matrix A.

Before discussing the next example, we recall the definition of a coordinate vector. Let V be a vector space over F and let $v \in V$. Let $B = (v_1, \ldots, v_n)$ be a basis of V. Then we may write v uniquely as

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n.$$

We recall that the column vector

$$v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in F^n$$

is called the coordinate vector of v. Note that v may be written as the product[2]

$$v = (v_1, v_2, \ldots, v_n) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = BX.$$

[1] Whenever we talk of a linear transformation $T : V \longrightarrow W$, it is implicit that V, W are vector spaces over the same field F.

[2] B is a hypervector.

Example 1.2. Let $B = (v_1, v_2, \ldots, v_n)$ be a basis of the vector space V over F. We then have a map, which we denote by B

$$B : F^n \longrightarrow V$$

by

$$B(X) = BX$$

where the product on the right is the matrix product. Note that B is linear and infact an isomorphism of vector spaces. The inverse linear transformation

$$B^{-1} : V \longrightarrow F^n$$

maps a vector $v \in V$ to its coordinate vector.

Associated to any linear transformation $T : V \longrightarrow W$ are two vector spaces, namely, the kernel $\ker(T)$ of T and the image $\mathrm{im}(T)$ of T. We note the following important fact.

Theorem 1.3. Let $T : V \longrightarrow W$ be a linear transformation with V finite dimensional. Then

$$\dim(V) = \dim(\ker(T)) + \dim(\mathrm{im}(T)). \tag{1.3.1}$$

Proof. Since V is finite dimensional so is $\ker(T)$. Fix a basis $(u_1, \ldots, u_k)$ of $\ker(T)$. Extend this to a basis

$$B = (u_1, \ldots, u_k, v_1, \ldots, v_{n-k})$$

of V so that V is n-dimensional. We claim that the set

$$B' = (T(v_1), \ldots, T(v_{n-k}))$$

is a basis of $\mathrm{im}(T)$.

Let $w \in \mathrm{im}(T)$. Let $v \in V$ be such that $T(v) = w$. Since B is a basis of V we can write

$$v = a_1u_1 + \cdots + a_ku_k + b_1v_1 + \cdots + b_{n-k}v_{n-k}$$

and hence

$$w = T(v) = b_1T(v_1) + \cdots + b_{n-k}(v_{n-k}).$$

Thus B' spans $\mathrm{im}(T)$. Next suppose that there exist scalars $b_1, \ldots, b_{n-k}$ such that

$$T(\sum_i b_iv_i) = b_1T(v_1) + \cdots + b_{n-k}T(v_{n-k}) = 0.$$

Thus $\sum_i b_iv_i \in \ker(T)$ and hence we may write

$$\sum_i b_iv_i = \sum_j a_ju_j$$

for some scalars $a_1, \ldots, a_k$. But as B is linearly independent we must have $a_i = 0$, $b_j = 0$ for all i, j. Thus B' linearly independent and hence a basis of $\mathrm{im}(T)$. This completes the proof of the theorem. □

This theorem has several interesting consequences which we now note. First a definition.

Definition 1.4. Given a linear transformation $T : V \longrightarrow W$, the dimension of the image of T is by definition the rank of the map T, that is,

$$\mathrm{rank}(T) = \dim(\mathrm{im}(T)).$$

Thus for a linear transformation $T : V \longrightarrow W$, with V finite dimensional, the equality in (1.3.1) may be written as

$$\dim(V) = \dim(\ker(T)) + \mathrm{rank}(T). \tag{1.4.1}$$

Example 1.5. Let $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be a linear transformation. Thus $\mathrm{rank}(T) \leq 2$. This forces

$$\dim(\ker(T)) \geq 1.$$

In particular $\ker(T) \neq \{0\}$. Thus the linear map T can never be $1-1$. More generally if $T : V \longrightarrow W$ is a linear map with V finite dimensional and $\dim(V) > \dim(W)$, T can never be injective. Notice that if $\dim(V) \geq \dim(W)$ then one can always define a surjective linear map

$$S : V \longrightarrow W.$$

This verification is left as an exercise.

Example 1.6. Dually, let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ be a linear map. Then as

$$2 = \dim(\ker(T)) + \mathrm{rank}(T)$$

we have that $\mathrm{rank}(T) \leq 2$. Hence T can never be onto. More generally, if $T : V \longrightarrow W$ is a linear map with V finite dimensional and $\dim(V) < \dim(W)$, then T cannot be surjective. Notice that if $\dim(V) \leq \dim(W)$, then one can always define an injective linear map

$$S : V \longrightarrow W.$$

This verification is left as an exercise.

Example 1.7. Let $T : V \longrightarrow W$ be a linear transformation where both are finite dimensional and $\dim(V) = \dim(W) = n$. Assume that T is surjective so that $\mathrm{rank}(T) = n$. This forces $\dim(\ker(T)) = 0$ and hence T is injective. Since one can also argue back we see that if T is injective, then T is surjective. Thus if $T : V \longrightarrow W$ is a linear transformation between two vector spaces of the same finite dimension, then T is injective if and only if T is surjective. This is not true when the spaces are not finite dimensional.

We note the following facts.

Lemma 1.8. Let A me a $m \times n$ matrix with entries in a field F. Let $A : F^n \longrightarrow F^m$ be the linear transformation defined by $A(X) = AX$.

(1) Given $B \in F^m$, there exists $X \in F^n$ with $A(X) = B$ if and only if the system of equations $AX = B$ has a solution.
(2) $A : F^n \longrightarrow F^m$ is onto if and only if the system of equations $AX = B$ has a solution for each $B \in F^m$.
(3) $A : F^n \longrightarrow F^m$ is injective if and only if the system of equations $AX = 0$ has a unique solution.

Proof. Is left as an exercise. □

Example 1.9. Let A be a $m \times n$ matrix with entries in a field F. Let, as usual, A also denote the associated linear map

$$A : F^n \longrightarrow F^m, \quad A(X) = AX.$$

We look at several cases.

(1) Suppose that $m < n$. Then we know (by Example 1.5) that the linear map

$$A : F^n \longrightarrow F^m$$

cannot be injective. Hence there exists a nonzero vector $X \in F^n$ with

$$AX = A(X) = 0.$$

Thus the homogeneous system of equations $AX = 0$ has a nonzero solution. Something that we had proved earlier using different methods. We now know more. Let W be the solution space of the homogeneous system $AX = 0$ so that $W = \ker(A)$. Then

$$\dim(W) = n - \operatorname{rank}(A).$$

(2) Assume that $n > m$. In this case we know (Example 1.6) that the linear transformation

$$A : F^n \longrightarrow F^m$$

is not surjective. Thus there exists $B \in F^m$ which is not in the image of A. This is the same as saying (by the above lemma) that there exists $B \in F^m$ such that the system of equations

$$AX = B$$

has no solutions.

(3) Finally, let $n = m$. Then there are two subcases. First assume that the matrix A is invertible. Then the system of equations $AX = B$ has a unique solution for each $B \in F^n$. In particular,

$$A(X) = 0$$

if and only if $X = 0$. Thus $A : F^n \longrightarrow F^n$ is an isomorphism. Thus A is an isomorphism if and only if

$$n = \operatorname{rank}(A) = \dim(\text{column space of } A).$$

Next assume that the matrix A is not invertible. Thus the homogeneous system $AX = 0$ has a non zero solution. In particular,

$$\ker(A) \neq 0$$

so that $\dim(\ker(A)) > 0$. This implies that $A : F^n \longrightarrow F^n$ is not onto. Hence there exists $B \in F^n$ such that

$$AX = A(X) = B$$

has no solution. If $AX = B$ has a solution then clearly it has more than one solution. For if $Y \in \ker(A)$, then

$$A(X + Y) = A(X + Y) = A(X) + A(Y) = 0 + B = B.$$

The above example makes it clear that solutions to a system of equations $AX = B$ can be well understood in terms of the nature of the linear transformation

$$A : F^n \longrightarrow F^m.$$

Let F be a field and let

$$T : F^n \longrightarrow F^m$$

be a linear transformation. Let $B = (e_1, \ldots, e_n)$ and $B' = (e'_1 \ldots, e'_m)$ denote the standard bases of F^n and F^m respectively. We may now write

$$T(e_j) = a_{1j}e'_1 + \cdots + a_{mj}e'_m$$

for $1 \leq j \leq n$. Now let A_j be the column vector

$$A_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

and let A be the $m \times n$ matrix whose j-th column is the column vector A_j. By Example 1.1, the matrix A induces a linear map

$$A : F^n \longrightarrow F^m, \quad A(X) = AX.$$

It is clear that, as maps, the linear transformations T and A are identical, that is,

$$T(X) = A(X) = AX$$

for all $X \in F^n$. Thus in particular

$$\text{rank}(T) = \text{rank}(A) = \dim(\text{column space of } A). \tag{1.9.1}$$

The last equality follows from the discussion in Example 1.1. The matrix A is called the matrix of the linear transformation T relative to the bases B and B'.

Here are two examples.

Example 1.10. Let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the linear map defined by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ y \end{pmatrix}$$

Let $B = (e_1, e_2) = B'$ e the bases of the domain and target of T respectively. Then

$$T(e_1) = T(1,0)^t = (1,0)^t = 1 \cdot e_1 + 0 \cdot e_2$$
$$T(e_2) = T(0,1)^t = (1,1)^t = 1 \cdot e_1 + 1 \cdot e_2$$

Thus the matrix of T is

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

relative to the above bases. Now let $w_1 = e_2$ and $w_2 = e_1$. Let us now determine the matrix of T relative to the bases

$$B = (w_1, w_2) = B'$$

of the domain and the target of T. We note

$$T(w_1) = T(0,1)^t = (1,1)^t = 1 \cdot w_1 + 1 \cdot w_2$$
$$T(w_2) = T(1,0)^t = (1,0)^t = 0 \cdot w_1 + 1 \cdot w_2$$

and thus the matrix of T is

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

relative to the second set of bases. Thus this example shows the importance of keeping tranc of the order of the basis vectors.

Example 1.11. Let $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be the linear transformation defined by

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y \\ x + z \end{pmatrix}$$

Let us determine the matrix of T relative to the standard ordered bases of $\mathbb{R}^3$ and $\mathbb{R}^2$. We note

$$T(e_1) = T(1,0,0)^t = (1,1)^t = e_1 + e_2$$
$$T(e_2) = T(0,1,0)^t = (1,0)^t = e_1 + 0 \cdot e_2$$
$$T(e_3) = T(0,0,1)^t = (0,1)^t = 0 \cdot e_1 + e_2.$$

Thus the matrix of T is

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

relative to the standard bases.

Here are some problems.

Exercise 1.12. Prove the claims made in Examples 1.51.6,1.7.

Exercise 1.13. Prove Lemma 1.8

Exercise 1.14. Let A be a $m \times n$ matrix and B a $n \times m$ matrix. If $BA = I$ prove that $m \geq n$.

Exercise 1.15. Let A be a $\ell \times m$ matrix and B a $n \times p$ matrix. Show that the map

$$M_{m \times n}(\mathbb{R}) \longrightarrow M_{\ell \times p}(\mathbb{R}), \quad X \mapsto AXB$$

is a linear transformation.

Exercise 1.16. Let A be a $m \times n$ matrix of reals. Show that the space of solution to the homogeneous system $AX = 0$ has dimension at least $n - m$.

Exercise 1.17. Find all linear transformations $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ that carries the subspace $x = y$ onto the subspace $y = 3x$.