Lecture 11

Definition. Let $f : \boldsymbol{R}^n \to \boldsymbol{R}^k$ be a function. We say that f is a linear function (or linear mapping) if for $\alpha, \beta \in \boldsymbol{R}$, $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{R}^n$, we have that

$$f(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) = \alpha f(\boldsymbol{x}_1) + \beta f(\boldsymbol{x}_2).$$

Example. Let $f : \boldsymbol{R}^2 \to \boldsymbol{R}^2$ be the function

$$f(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 2x + 3y \\ -x + 4y \end{bmatrix}.$$

It is clear that f is a function. We claim that f is a linear function. To see this, let $\begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} z \\ w \end{bmatrix} \in \boldsymbol{R}^2, \alpha, \beta$ real numbers. Then:

$$\begin{aligned}
f(\alpha \begin{bmatrix} x \\ y \end{bmatrix} + \beta \begin{bmatrix} z \\ w \end{bmatrix}) &= f(\begin{bmatrix} \alpha x + \beta z \\ \alpha y + \beta w \end{bmatrix} \\
&= \begin{bmatrix} 2(\alpha x + \beta z) + 3(\alpha y + \beta w) \\ -(\alpha x + \beta z) + 4(\alpha y + \beta w) \end{bmatrix} \\
&= \begin{bmatrix} 2\alpha x + 3\alpha y + 2(\beta z + \beta w) \\ -(\alpha x + 4\alpha y) - \beta z + 4\beta w \end{bmatrix} \\
&= \alpha \begin{bmatrix} 2x + 3y \\ -x - 4y) \end{bmatrix} + \beta \begin{bmatrix} 2z + 3w \\ -z + 4w \end{bmatrix} \\
&= \alpha f(\begin{bmatrix} x \\ y \end{bmatrix}) + \beta f(\begin{bmatrix} z \\ w \end{bmatrix}).
\end{aligned}$$

Thus f is linear.

A useful observation regarding this simple example is the following: Let $A = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}$. Then

$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x+3y \\ -x+4y \end{bmatrix}$$
$$= A\begin{bmatrix} x \\ y \end{bmatrix}.$$

Observation: Let A bean $n \times k$ matrix. Then A "induces" a linear function $f_A, f_A : \boldsymbol{R}^k \to \boldsymbol{R}^n$ by the rule:

$$\text{for } \boldsymbol{x} \in R^k, f_A(\boldsymbol{x}) = A\boldsymbol{x}.$$

The fact that it is a linear function is a consequence of the properties of matrix multiplication.

Of course the last observation brings up the following question: if $f : \boldsymbol{R}^n \to \boldsymbol{R}^k$ is a linear function, can we find a matrix A such that $f(\boldsymbol{x}) = A\boldsymbol{x}$ for all $\boldsymbol{x} \in \boldsymbol{R}^n$?

We will answer this question. In our discussion the fact that any vector in $\boldsymbol{R}^n, \boldsymbol{R}^k$ can be written as a linear combination of the basis consisting of the standard unit vectors in each space will play a crucial role.

From now on the basis of $\boldsymbol{R}^n$ which is the set of standard unit vectors will be called the standard basis.

As a first observation we note that since $\boldsymbol{x}$ is an $n \times 1$ matrix, A must have n columns and since $f(\boldsymbol{x}) \in \boldsymbol{R}^k$, it follows that A must have k rows. Thus A must be a $k \times n$ matrix.

Next, we observe that if $\boldsymbol{x} \in \boldsymbol{R}^n, \boldsymbol{x} = \begin{matrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{matrix}$ then:

$$\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{e}_j.$$

By the linearity of f we get:

$$f(\boldsymbol{x}) = \sum_{j=1}^{n} x_j f(\boldsymbol{e}_j).$$

We set $A = [f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), \ldots, f(\boldsymbol{e}_n)]$, the matrix written by column vectors.

For each $j \in \{1, 2, \ldots, n\}$ we let

$$f(\boldsymbol{e}_j) = \begin{matrix} a_{1,j} \\ a_{2,j} \\ . \\ . \\ a_{n,j} \end{matrix} .$$

This is the j^{th} column vector written as a linear combination of the standard basis of $\boldsymbol{R}^k$.

The matrix A we have obtained has the desired property. It is called the matrix of f with respect to the standard bases.

Definition. Let $f : \boldsymbol{R}^n \to \boldsymbol{R}^k$ be a linear function. We define the kernel or null space,

of f ,$Ker(f)$and the range of $f, R(f)$ as

$$Ker(f) = \{\boldsymbol{x} \in \boldsymbol{R}^n : f(\boldsymbol{x}) = 0_k\}$$

$$R(f) = \{\boldsymbol{y} \in \boldsymbol{R}^k : \text{ there exists } \boldsymbol{x} \in \boldsymbol{R}^n \text{ such that } f(\boldsymbol{x}) = \boldsymbol{y}\}.$$

Observation. $Ker(f)$ is a subspace of $\boldsymbol{R}^n$ and $R(f)$ is a subspace of $\boldsymbol{R}^k$.

Observation. If $f : \boldsymbol{R}^n \to \boldsymbol{R}^k$ is a linear function and if A is its matrix with respect to the standard bases, then $Ker(f)$ is the solution set of the homogeneous linear system given by

$$A\boldsymbol{x} = 0_k.$$

Lemma. Let $f : \boldsymbol{R}^n \to \boldsymbol{R}^k$ be a linear function. The function f is one to one if and only if $Ker(f) = \{0_n\}$.

Proof.

If f is one to one, then, since $f(0_n) = 0_k$, there cannot be any other element $\boldsymbol{x} \in \boldsymbol{R}^n$ such that $f(\boldsymbol{x}) = 0_k$.

It follows that $Ker(f) = \{0_n\}$.

Now assume that $Ker(f) = \{0_n\}$. If $\boldsymbol{x},\boldsymbol{y} \in \boldsymbol{R}^n$ and $f(\boldsymbol{x}) = f(\boldsymbol{y})$, then

$$f(\boldsymbol{x}) - f(\boldsymbol{y}) = 0_n$$

$$f(\boldsymbol{x} - \boldsymbol{y}) = 0_n$$

$$\boldsymbol{x} - \boldsymbol{y} \in Ker(f)$$

$$\boldsymbol{x} - \boldsymbol{y} = 0_n$$

$$\boldsymbol{x} = \boldsymbol{y}.$$

Thus f is one to one.

Lemma. Let $f : \boldsymbol{R}^n \to \boldsymbol{R}^k$ be a linear function and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ a basis for $\boldsymbol{R}^n$. If f is one to one, then $\{f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)\}$ is a basis for the range of f.

Proof.

We already know that this set spans the range of f. To complete the proof we only need to show that it is a linearly independent set.

Assume that:

$$\sum_{i=1}^{n} \alpha_i f(\boldsymbol{v}_i) = 0_k.$$

Then, by the linearity of f:

$$f\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i\right) = 0_k.$$

Since f is one to one, we must have that:

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i = 0_n.$$

Since $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis, it is linearly independent and so we must have that $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.

Thus $\{f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)\}$ is a linearly independent set.

Definition. Let $f : \boldsymbol{R}^n \to \boldsymbol{R}^k$ be a linear function. We define the nullity of f, $nullity(f)$, as he dimension of $Ker(f)$, and we define the rank of f, $rank(f)$, as the dimension of the range of f.

In the proof of the following result we will make use of the exercises in the homework assignment that will be published today.

Theorem. Let $f : \boldsymbol{R}^n \to \boldsymbol{R}^k$ be a linear function. Then

$$nullity(f) + rank(f) = n.$$

Proof.

Since the kernel of f is a subspace of $\boldsymbol{R}^n$, either $Ker(f) = \{0_n\} (\dim(Ker(f)) = 0)$ or $\dim(Ker(f)) \geq 1$.

If $nullity(f) = 0$ then f is one to one and $\{f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), f(\boldsymbol{e}_3), \ldots, f(\boldsymbol{e}_n)\}$ is linearly independent, so $rank(f) = n$ and the result holds.

Now assume that $nullity(f) > 0$, say $nullity(f) = p, 1 \leq p \leq n$. If $nullity(f) = n, f(\boldsymbol{x}) = 0_k$ for all $\boldsymbol{x} \in \boldsymbol{R}^n$. The result is true in this particular case.

The last case is: $1 \leq p < n$. Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ be a basis for $Ker(f)$. Then we know that there exist $\boldsymbol{v}_{p+1}, \boldsymbol{v}_{p+2}, \ldots, \boldsymbol{v}_n$ in $\boldsymbol{R}^n$ such that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p, \boldsymbol{v}_{p+1}, \ldots, \boldsymbol{v}_n\}$ is a basis for $\boldsymbol{R}^n$.

We will prove that $\{f(\boldsymbol{v}_{p+1}), \ldots, f(\boldsymbol{v}_n)\}$ is a basis for the range of f. The vectors

certainly span the range, so all we need to prove is that the set is linearly independent.

We have:

$$\sum_{i=p+1}^{n} \alpha_i f(\boldsymbol{v}_i) = 0_k$$

$$f(\sum_{i=p+1}^{n} \alpha_i \boldsymbol{v}_i) = 0_k.$$

Thus $\sum_{i=p+1}^{n} \alpha_i \boldsymbol{v}_i \in Ker(f)$.

Since $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ is a basis for $Ker(f)$, we must be able to write

$$\sum_{i=p+1}^{n} \alpha_i \boldsymbol{v}_i = \sum_{i=1}^{p} \alpha_1 \boldsymbol{v}_i.$$

Now we have that

$$\sum_{i=1}^{p} \alpha_1 \boldsymbol{v}_i - \sum_{i=p+1}^{n} \alpha_i \boldsymbol{v}_i = 0_n.$$

The linear independence of the basis implies that $\alpha_i = 0, i = 1\ldots,p,p+1,\ldots,n$ and thus $\alpha_i = 0, i = p+1,\ldots,n$ and the result follows.

Definition. If A is an $n \times k$ matrix, we consider it as a linear mapping from $\boldsymbol{R}^k$ into $\boldsymbol{R}^n$ and then the nullity of A and the rank of A are those if the linear transformation $f_A : \boldsymbol{R}^k \to \boldsymbol{R}^n$ it induces.

Observation. If A is an $n \times k$ matrix, $A = [a_{i,j}]_{\substack{i=1\ldots n\\ j=1\ldots k}}$ then, writing A as an array of column vectors we have that

$$A = [A\boldsymbol{e}_1, A\boldsymbol{e}_2, \ldots, A\boldsymbol{e}_k].$$

Since $f_A(\boldsymbol{x}) = A\boldsymbol{x}$, we see that the rank of A is the dimension of the subspace of $\boldsymbol{R}^n$ spanned by the column vectors. We can also consider A as the column matrix of its row vectors,

$$A = \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ . \\ . \\ \boldsymbol{r}_n \end{bmatrix}.$$

The set of row vectors spans a subspace of $\boldsymbol{R}^k$. The dimension of this subspace of $\boldsymbol{R}^k$ also provides useful information about A. The fact that the dimension of the subspace spanned by the row vectors is equal to the rank of A is the content of the next result.

Theorem. Let $A[a_{i,j}]_{\substack{i=1\ldots n\\ j=1\ldots k}}$. Then $rank(A)$ is equal to the dimension of the subspace of $\boldsymbol{R}^k$ spanned by the row vectors of A.

Proof.

We know that $rank(A) + nullity(A) = k$.

We consider the homogeneous linear system of n equations in k unknowns that has A as the matrix of coefficients, that is:

$$\begin{aligned}
a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,k}x_k &= 0\\
a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,k}x_k &= 0\\
&.\\
&.\\
&.\\
a_{n,1}x_1 + a_{n,2}x_2 + \ldots + a_{n,k}x_k &= 0.
\end{aligned}$$

The subspace of the solutions of this system is the kernel of the linear function induced by A.

We know that we can get an equivalent linear system that is triangular. If the equivalent triangular system has p nonzero equations then we can write it as

$$\begin{array}{ccccccc}
b_{1,m_1}x_{m_1} + & b_{1,m_1+1}x_{m_1+1} & . & .+ & b_{1,k}x_k & = & 0\\
0 + & b_{2,m_2}x_{m_2} + & . & .+ & b_{2,p}x_p & = & 0\\
. & . & . & . & . & . & .\\
. & . & . & . & . & . & .\\
. & . & . & . & . & . & .\\
0 & 0 & b_{p,m_p}x_{m_p} + & .+ & b_{2.k}x_k & = & 0
\end{array}$$

From this we see two things: the dimension of the kernel of A is p and $x_{m_1}, x_{m_2},, \ldots, x_{m_p}$ can be written in terms of the remaining variables, which can be assigned any value. It follows that the dimension of the kernel is also equal to $k - p$. Thus:

$$\begin{aligned}
nullity(A) &= k - p\\
nullity(A) + rank(A) &= k\\
k - p + rank(A) &= k\\
rank(A) &= p.
\end{aligned}$$

Thus the result has been proved.