LECTURE 11

Matrices and Linear Transformations

1. Mappings between Sets

Let **A** be an $m \times n$ matrix. The goal of this lecture is to develop a geometric interpretation for homogeneous linear systems of the form $\mathbf{Ax} = \mathbf{b}$.

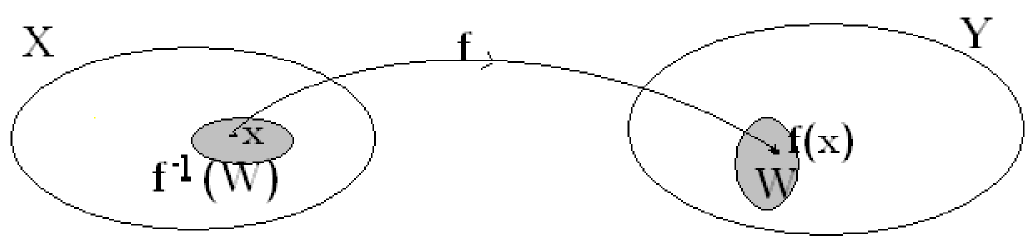
First let me recall some basic notions about maps between two sets. Let X and Y be sets. A **function** $f : X \to Y$ is a rule that associates with each element $x \in X$ an element $f(y) \in Y$. The set X is called the **domain** of the function f and the set Y is called the **codomain** of f. The set

$$\{y \in Y \mid y = f(x) \text{ for some } x \in X\}$$

is called the **image** of the function f, and if W is a subset of Y, then the set

$$f^{-1}(W) = \{x \in X \mid f(x) \in W\}$$

is called the **inverse image of** W under f.



2. Linear Transformations

We shall now restrict our attention to the following kinds of maps.

DEFINITION 11.1. *A function* $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ *is called a* ***linear transformation*** *if it satisfies*

1. $\mathbf{T}(\mathbf{u} + \mathbf{v}) = \mathbf{T}(\mathbf{u}) + \mathbf{T}(\mathbf{v})$ *(i.e. the function* $\mathbf{T}$ *preserves vector addition)*
2. $\mathbf{T}(r\mathbf{v}) = r\mathbf{T}(\mathbf{v})$ *(i.e., the function* $\mathbf{T}$ *preserves scalar multiplication)*

for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *and all scalars* $r \in \mathbb{R}$.

It is easy to see that if a mapping preserves both vector addition and scalar addition, then it will also preserve a combination of such operations; that is to say, it will preserve genearl linear combinations

$$\mathbf{T}(r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \cdots + r_k\mathbf{v}_k) = r_1\mathbf{T}(\mathbf{v}_1) + r_2\mathbf{T}(\mathbf{v}_2) + \cdots + r_k\mathbf{T}(\mathbf{v}_k)$$

EXAMPLE 11.2. Show that the transformation $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3 : (s,t) \to (t, s, 1 + t + s)$ is not a linear transformation.

- Let $\mathbf{v} = (s,t)$ Then

$$\begin{aligned} \mathbf{T}(\mathbf{v}) &= \mathbf{T}(s,t) = (t,s,1+t+s) \\ \mathbf{T}(r\mathbf{v}) &= \mathbf{T}(rs,rt) = (rt,rs,1+rs+rt) \\ &\neq r(t,s,1+t+s) = r\mathbf{T}(\mathbf{v}) \end{aligned}$$

and so $\mathbf{T}$ does not preserve scalar multiplication: hence it is not a linear transformation.

3. Linear Transformations and Matrices

Note that in the preceding example, despite the fact that the coordinates of the image points are linear functions in the s and t, the mapping $\mathbf{T}$ is not a linear transformation. What then constitutes a linear mapping?

Lemma 11.3. *Let $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping and let $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ be a basis for $\mathbb{R}^n$. Then every vector in the image of $\mathbf{T}$ can be written as a linear combination of the vectors $\mathbf{T}(\mathbf{b}_1), \mathbf{T}(\mathbf{b}_2), \dots, \mathbf{T}(\mathbf{b}_n)$.*

Proof. Since B is a basis for $\mathbb{R}^n$, any vector $\mathbf{v} \in \mathbb{R}^n$ can be expressed as

$$\mathbf{v} = r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \cdots + r_n\mathbf{b}_n$$

And so the image of a vector $\mathbf{v}$ by $\mathbf{T}$ will be expressible as

$$\begin{aligned} \mathbf{T}(\mathbf{v}) &= \mathbf{T}(r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \cdots + r_n\mathbf{b}_n) \\ &= r_1\mathbf{T}(\mathbf{b}_1) + r_2\mathbf{T}(\mathbf{b}_2) + \cdots + r_k\mathbf{T}(\mathbf{b}_k) \quad \text{(since } \mathbf{T} \text{ is a linear transformation)} \end{aligned}$$

Theorem 11.4. *Let $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, let $\{\mathbf{e}_i \mid i = 1, \dots, n\}$ be the standard basis for $\mathbb{R}^n$:*

$$(\mathbf{e}_i)_j = \begin{cases} 1 & , \quad j = i \\ 0 & , \quad j \neq i \end{cases}$$

and let $\mathbf{A}$ be the $m \times n$ matrix whose i^{th} column coincides with $\mathbf{T}(\mathbf{e}_i) \in \mathbb{R}^m$. Then

$$\mathbf{T}(\mathbf{x}) = \mathbf{A}\mathbf{x}$$

In other words, every linear transformation $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ is equivalent to the matrix mulitiplication of the vectors $\mathbf{x} \in \mathbb{R}^n$ by an $m \times n$ matrix $\mathbf{A}$. The converse of this fact is also true, if $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ is the mapping defined by

$$\mathbf{x} \in \mathbb{R}^n \to \mathbf{A}\mathbf{x} \in \mathbb{R}^m$$

then $\mathbf{T}$ is a linear transformation.

Example 11.5. Find the matrix corresponding to the linear transformation $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$ given by $\mathbf{T}(x_1, x_2) = (x_1 - x_2, x_1 + x_2, x_1)$.

- We have

$$\begin{aligned} \mathbf{T}(\mathbf{e}_1) &= \mathbf{T}(1,0) = (1-0, 1+0, 1) = (1,1,1) \\ \mathbf{T}(\mathbf{e}_2) &= \mathbf{T}(0,1) = (0-1, 0+1, 0) = (-1,1,0) \end{aligned}$$

Hence

$$\mathbf{A} = [\mathbf{T}(\mathbf{e}_1), \mathbf{T}(\mathbf{e}_2)] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

We confirm

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \\ x_1 \end{bmatrix} = \mathbf{T}\mathbf{x}$$

DEFINITION 11.6. *The **kernel** of a linear mapping $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ is the set of all $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Tx} = \mathbf{0} \in \mathbb{R}^m$. The **range** of $\mathbf{T}$ is the set of all $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y} = \mathbf{T}(\mathbf{x})$ for some $\mathbf{x} \in \mathbb{R}^n$.*

Now let $\mathbf{A}$ be the $m \times n$ matrix corresponding to a linear transformation $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$. Then

$$\begin{aligned} \ker(\mathbf{T}) &= \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{T}(\mathbf{x}) = \mathbf{0}\} \\ &= \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{0}\} = \text{ Null space of } \mathbf{A} \end{aligned}$$

$$\begin{aligned} \mathit{range}(\mathbf{T}) &= \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = \mathbf{T}(\mathbf{x}) \quad , \quad \text{for some } \mathbf{x} \in \mathbb{R}^n\} \\ &= \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = \mathbf{Ax} \quad , \quad \text{for some } \mathbf{x} \in \mathbb{R}^n\} = \text{ column space of } \mathbf{A} \end{aligned}$$

Note also that the dimension n of the domain $\mathbb{R}^n$ of $\mathbf{T}$ is same as the number of column in the corresponding matrix $\mathbf{A}$. Now from Theorem 10.6 of Lecture 10 we know

$$(\text{number of columns of } \mathbf{A}) = (\text{dimension of null space of } \mathbf{A}) + (\text{dimension of column space of } \mathbf{A})$$

In terms of notions of linear transformations this translates to

$$(\text{dimension of domain of } \mathbf{T}) = (\text{dimension of kernel of } \mathbf{T}) + (\text{dimension of range of } \mathbf{T})$$

EXAMPLE 11.7. Consider the linear transformation $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^2$ given by $\mathbf{T}(x_1, x_2, x_3) = (x_1 + x_3, x_1 + x_2 + 2x_3, -x_1 +$ Find a basis for the kernel of $\mathbf{T}$ and a basis for the range of $\mathbf{T}$.

- Let's first find the matrix representation of $\mathbf{T}$. We have

$$\begin{aligned} \mathbf{T}(\mathbf{e}_1) &= \mathbf{T}(1,0,0) = (1+0, 0+0+2(0), -1+1, 2(0)+2(0)) = (1,1,-1,0) \\ \mathbf{T}(\mathbf{e}_2) &= \mathbf{T}(0,1,0) = (0+0, 0+1+2(0), -0+1, 2(1)+2(0)) = (0,1,1,2) \\ \mathbf{T}(\mathbf{e}_3) &= \mathbf{T}(0,0,1) = (0+1, 0+0+2(1), -0+0, 2(0)+2(1)) = (1,2,0,2) \end{aligned}$$

and so the linear transformation $\mathbf{T}$ corresponds to the 4×3 matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ -1 & 1 & 0 \\ 0 & 2 & 2 \end{bmatrix}$$

As we pointed out above the kenel of $\mathbf{T}$ is the same as the null space of $\mathbf{A}$ and the range of $\mathbf{T}$ is the same thing as the column space of $\mathbf{A}$. To find the null space and column space of a matrix we first row reduce $\mathbf{A}$ to reduced row-echelon form

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ -1 & 1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{A}'$$

From Lecture 10 (Lemma 10.3) we know that the column space of the matrix $\mathbf{A}$ is spanned by the columns in $\mathbf{A}$ to which correspond columns in the row-echelon form $\mathbf{A}'$ that contain pivots. Thus,

$$\text{range of } \mathbf{T} = \text{column space of } \mathbf{A} = \text{ span}\left(\begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}\right)$$

The kernel of $\mathbf{T}$ can be identified with the null space of $\mathbf{A}$, which is equal to the null space of $\mathbf{A}'$: i.e, the solution set

$$\left.\begin{array}{r} x_1 + x_3 = 0 \\ x_2 + x_3 = 0 \\ 0 = 0 \\ 0 = 0 \end{array}\right\} \quad \Rightarrow \quad \left\{\begin{array}{l} x_1 = -x_3 \\ x_2 = -x_3 \end{array}\right.$$

so

$$\mathbf{x} = \begin{bmatrix} -x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

So

$$\ker(\mathbf{T}) = span\left(\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right)$$

4. Composition of Linear Transformations

Suppose we have two linear transformations

$$\begin{aligned} \mathbf{T}_1 &: \mathbb{R}^n \to \mathbb{R}^m \\ \mathbf{T}_2 &: \mathbb{R}^m \to \mathbb{R}^p \end{aligned}$$

Because every element in the range of $\mathbf{T}_1$ can be regarded as an element in the domain of $\mathbf{T}_2$ the composed mapping

$$\mathbf{T}_1 \circ \mathbf{T}_2 : \mathbb{R}^n \to \mathbb{R}^p \quad ; \quad \mathbf{x} \in \mathbb{R}^n \quad \mapsto \quad \mathbf{T}_2(\mathbf{T}_1(\mathbf{x})) \in \mathbb{R}^p$$

is well defined, and, in fact, is another linear transformation. Indeed, if we switch back to our matrix language, where the transformations $\mathbf{T}_1 : \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{T}_2 : \mathbb{R}^m \to \mathbb{R}^p$ are implemented by, respectively, an $m \times n$ matrix $\mathbf{A}_1$ and an $p \times m$ matrix $\mathbf{A}_2$, then to the composed transformation $\mathbf{T}_1 \circ \mathbf{T}_2 : \mathbb{R}^n \to \mathbb{R}^p$ we have the following matrix:

$$\mathbf{A}_{12} = \mathbf{A}_2 \mathbf{A}_1$$

Note that this matrix multiplication is also well-defined since the number m of columns of $\mathbf{A}_2$ is the same as the number m of rows of $\mathbf{A}_1$.

Example 11.8. Consider the linear transformation corresponding to a rotation in the xy plane by an angle θ

$$\begin{aligned} x &\to x' = x\cos(\theta) + y\sin(\theta) \\ y &\to y' = -x\sin(\theta) + y\cos(\theta) \end{aligned}$$

To this linear transformation corresponds the following 2×2 matrix:

$$\mathbf{A} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

If we apply this transformation twice, the effect should be that of a two rotations by the angle θ. We thus should have

$$\mathbf{AA} = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ -\sin(2\theta) & \cos(2\theta) \end{bmatrix} \tag{11.1}$$

Calculating the matrix multiplication on the left hand side:

$$\mathbf{AA} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} \cos^2(\theta) - \sin^2(\theta) & 2\cos(\theta)\sin(\theta)) \\ -2\cos(\theta)\sin(\theta) & \cos^2(\theta) - \sin^2(\theta) \end{bmatrix} \tag{11.2}$$

Comparing (11.1) with (11.2) we see we must have

$$\begin{aligned} \cos(2\theta) &= \cos^2(\theta) - \sin^2(\theta) \\ \sin(2\theta) &= \cos(\theta)\sin(\theta) \end{aligned}$$

We have thus, by a simple matrix calculation, rederived the double angle trig identities one learns in high school.