LECTURE 11

Linear Transformations and Matrices

1. Matrix Multiplication

We have put this off long enough; let us now define matrix multiplication.

Let $Mat_{n,m}(\mathbb{F})$ be the set of $n \times m$ matrices with entries in $\mathbb{F}$. As usual, we will write a typical element of $Mat_{n,m}(\mathbb{F})$ as an array with n rows and m columns

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

Definition 11.1. *Let $\mathbf{A} \in Mat_{n,m}(\mathbb{F})$ and let $\mathbf{B} \in Mat_{m,\ell}(\mathbb{F})$. The matrix product $\mathbf{AB}$ is defined as the matrix in $Mat_{n,\ell}(\mathbb{F})$ with entries*

$$(\mathbf{AB})_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} \qquad , \qquad 1 \leq i \leq n \quad , \quad 1 \leq j \leq \ell \quad .$$

Remark 11.2. Another formulation of the rule for matrix multiplication is this: the $(ij)^{th}$ element of the matrix product $\mathbf{AB}$ is the dot product for the i^{th} row vector of $\mathbf{A}$ with the j^{th} column vector of $\mathbf{B}$.

Recall that the set $Mat_{n,m}(\mathbb{F})$ is naturally a vector space over $\mathbb{F}$ with scalar multiplication and vector addition defined by

$$\begin{aligned} (\lambda \cdot \mathbf{A})_{ij} &= \lambda a_{ij} \qquad , \qquad 1 \leq i \leq n \quad , \quad 1 \leq j \leq m \\ (\mathbf{A} + \mathbf{B})_{ij} &= a_{ij} + b_{ij} \qquad , \qquad 1 \leq i \leq n \quad , \quad 1 \leq j \leq m \end{aligned}$$

Theorem 11.3. *Let $\mathbf{A}, \mathbf{C} \in Mat_{n,m}(\mathbb{F})$, $\mathbf{B}, \mathbf{D} \in Mat_{m,\ell}(\mathbb{F})$ and let $\alpha, \beta \in \mathbb{F}$.*

- $\mathbf{A}(\alpha \mathbf{B} + \beta \mathbf{D}) = \alpha(\mathbf{AB}) + \beta(\mathbf{AD})$
- $(\alpha \mathbf{A} + \beta \mathbf{C})\mathbf{B} = \alpha(\mathbf{AB}) + \beta(\mathbf{CB})$

Corollary 11.4. *Let $\mathbf{A}$ be an $n \times m$ matrix. Then for each $\ell = 1, 2, 3, \ldots$, $\mathbf{A}$ defines a family of linear transformations $L_\ell : Mat_{m,\ell}(\mathbb{F}) \to Mat_{n,\ell}(\mathbb{F})$ via left multiplication*

$$Mat_{m,\ell}(\mathbb{F}) \ni \mathbf{B} \quad \longmapsto \quad \mathbf{AB} \in Mat_{n,\ell}(\mathbb{F})$$

and for each $k = 1, 2, 3, \ldots$, $\mathbf{A}$ defines a family of linear transformations $R_k : Mat_{k,n}(\mathbb{F}) \to Mat_{k,m}(\mathbb{F})$ via right multiplication

$$Mat_{k,n}(\mathbb{F}) \ni \mathbf{B} \quad \longmapsto \quad \mathbf{BA} \in Mat_{k.m}(\mathbb{F}) \quad .$$

Remark 11.5. Because we can think of a vectors in $\mathbb{F}^m$ either as $m \times 1$ matrices (column vectors) or as $1 \times m$ matrices (row vectors), we actually have two possibilities for implementing matrix multiplication on $\mathbb{F}^m$. This ambiguity is usually resolved by *adopting the convention* that elements of $\mathbb{F}^m$ correspond to $m \times 1$ matrices.

2. Linear Transformations, Bases, and Matrix Multiplication

Let V and W be finitely generated vector spaces and let $T : V \to W$ be a linear transformation from V to W. Since V and W are finitely generated vector spaces they both have bases. Let $B = \{v_1, \ldots, v_m\}$ be a basis for V and let $B' = \{w_1, \ldots, w_n\}$ be a basis for W. From the data T, B, and B' we can form a matrix in $Mat_{\mathbb{F}}(n, m)$; for each basis vector $v_i \in V$ gets mapped by T to a vector $T(v_i) \in W$, which in turn must be expressible in terms of the basis B' of W. Say

$$T(v_i) = \lambda_1^{(i)} w_1 + \lambda_2^{(i)} w_2 + \cdots + \lambda_n^{(i)} w_n$$

Then a more general vector $v = a_1 v_1 + \cdots + a_m v_m$ gets mapped to

$$\begin{aligned}
T(v) &= T(a_1 v_1 + \cdots + a_m v_m) \\
&= a_1 T(v_1) + \cdots + a_m T(v_m) \qquad \text{because } T \text{ is a linear transformation} \\
&= a_1 \lambda_1^{(1)} w_1 + \cdots + a_1 \lambda_n^{(1)} w_n \\
&\quad + a_2 \lambda_1^{(2)} w_1 + \cdots + a_2 \lambda_n^{(2)} w_n \\
&\quad + \cdots \\
&\quad + a_m \lambda_1^{(m)} w_1 + \cdots + a_m \lambda_n^{(m)} w_n \\
&= \left(\lambda_1^{(1)} a_1 + \lambda_1^{(2)} a_2 + \cdots + \lambda_1^{(m)} a_m\right) w_1 \\
&\quad + \left(\lambda_2^{(1)} a_1 + \lambda_2^{(2)} a_2 + \cdots + \lambda_2^{(m)} a_m\right) w_2 \\
&\quad + \cdots \\
&\quad + \left(\lambda_n^{(1)} a_1 + \lambda_n^{(2)} a_2 + \cdots + \lambda_n^{(m)} a_m\right) w_n
\end{aligned}$$

Put another way, a vector $v \in V$ with coordinate vector $[a_1, \ldots, a_m] \in \mathbb{F}^m$ with respect to the basis $B \subset V$ gets mapped to $T(v)$ which has coordinate vector

$$\left[\lambda_1^{(1)} a_1 + \lambda_1^{(2)} a_2 + \cdots + \lambda_1^{(m)} a_m, \lambda_1^{(1)} a_2 + \lambda_2^{(2)} a_2 + \cdots + \lambda_2^{(m)} a_m, \ldots, \lambda_n^{(1)} a_1 + \lambda_n^{(2)} a_2 + \cdots + \lambda_n^{(m)} a_m\right]$$

with respect to the basis $B' = \{w_1, \ldots, w_n\}$ of W. Let us now write the coordinate vector $[a_1, \ldots, a_m]$ as a $m \times 1$ matrix, and coordinate vector of $T(v)$ as an $n \times 1$ matrix. Then we have

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \xrightarrow{\quad T \quad} \begin{bmatrix} \lambda_1^{(1)} a_1 + \lambda_1^{(2)} a_2 + \cdots + \lambda_1^{(m)} a_m \\ \lambda_1^{(1)} a_2 + \lambda_2^{(2)} a_2 + \cdots + \lambda_2^{(m)} a_m \\ \vdots \\ \lambda_n^{(1)} a_1 + \lambda_n^{(2)} a_2 + \cdots + \lambda_n^{(m)} a_m \end{bmatrix} = \begin{bmatrix} \lambda_1^{(1)} & \lambda_1^{(2)} & \cdots & \lambda_1^{(m)} \\ \lambda_2^{(1)} & \lambda_2^{(2)} & \cdots & \lambda_2^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n^{(1)} & \lambda_n^{(2)} & \cdots & \lambda_n^{(m)} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \tag{1}$$

where on the right we have simply recognized that the coordinate vector for $T(v)$ can also be obtained via a matrix multiplication.

Let me summarize this result. Let $B = \{v_1, \ldots, v_m\}$ be a basis for V, let $B' = \{w_1, \ldots, w_n\}$ be a basis for W, and let $T : V \to W$ be a linear transformation. Form an $n \times m$ matrix $\mathbf{T}_{B,B'}$ by utilizing the coordinate vectors for $T(v_i)$ with respect to B' as columns. Then a vector $v \in V$ with coordinate vector $\mathbf{v}_B \in \mathbb{F}^m$ with respect to B gets mapped by T to the vector in W with coordinate vector $\mathbf{T}_{B,B'} \mathbf{v}_B$.

The upshot of this is: *once you coordinatize your vector spaces, a linear transformation is always implementable by matrix multiplication.*

3. The Kernel and Image of a Linear Transformation

Given a linear transformation $T : V \to W$, we have two associated subspaces of, respectively, the domain V and the codomain W:

$$\begin{array}{rcl} \ker(T) & = & \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}_W\} \\ range(T) & = & \{\mathbf{w} \in W \mid \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\} \end{array}$$

How does one compute these subspaces? Well, as is typical of abstract vector spaces, there little we can calculate without first positing some bases. So let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ be a basis for V and let $B' = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ be a basis for W. As we say in the preceding section, a linear transformation $T : V \to W$ is equivalent to defining a certain matrix multiplication that sends coordinate vectors for V to coordinate vectors for W. Schematically, the way this works is

$$\begin{array}{ccccccc} & & v \in V & \xrightarrow{T} & T(v) \in W & & \\ & \nearrow\swarrow & & & & \searrow\nwarrow & \\ \mathbf{v}_B \in \mathbb{F}^m & & & \xrightarrow{\mathbf{T}_{BB'}} & & & \mathbf{T}_{BB'}\mathbf{v}_B \in \mathbb{F}^n \end{array}$$

where $\mathbf{T}_{BB'}$ is the $n \times m$ matrix formed by using the coordinate vectors (w.r.t. B') of each $T(v_i)$ as columns.

Because of the tight one-to-one correspondence between vectors and their coordinate vectors, we should be able to identify $\ker(T)$ with a certain subspace of $\mathbb{F}^m$ and $range(T)$ with a certain subspace of $\mathbb{F}^n$. Indeed,

$$T(v) = 0 \quad \Longleftrightarrow \quad \mathbf{T}_{BB'}\mathbf{v}_B = \mathbf{0}_{\mathbb{F}^n}$$

which tells us that

$$\ker(T) \quad \Longleftrightarrow \quad NullSp(\mathbf{T}_{BB'}) \equiv \text{solution set of } \mathbf{T}_{BB'}\mathbf{x} = \mathbf{0} \tag{2}$$

Also, if we denote the components of $\mathbf{v}_B$ as $[a_1, \ldots, a_m]$ and the columns of $\mathbf{T}_{BB'}$ as $\mathbf{c}_1, \ldots, \mathbf{c}_m$ (still with $\mathbf{c}_i$ = coordinate vector for $T(v_i)$ with respect to B'), then $\mathbf{v}_B = [a_1]$

$$\mathbf{T}_{BB'}\mathbf{v}_B = a_1\mathbf{c}_1 + a_2\mathbf{c}_2 + \cdots + a_m\mathbf{c}_m$$

Thus, as we let $\mathbf{v}_B$ vary over $\mathbb{F}^m$, the vectors $\mathbf{T}_{BB'}\mathbf{v}_B$ run over the span of the columns of $\mathbf{T}_{BB'}$; and so

$$range(T) \quad \Longleftrightarrow \quad ColSp(\mathbf{T}_{BB'}) \tag{3}$$

Let's restate this result as a theorem:

THEOREM 11.6. *Suppose V and W are finitely generated vector spaces and $T : V \to W$ is a linear transformation. Let $B = \{b_1, \ldots, b_m\}$ be a basis for V and let $B' = \{b'_1, \ldots, b'_n\}$ be a basis for W. Let $i_B : V \to \mathbb{F}^m$ and $i_{B'} = W \to \mathbb{F}^n$ be the associated coordinatization isomorpohisms. Finally, let $\mathbf{A}_{T,B,B'}$ be associated $n \times m$ matrix of whose columns are given by*

$$\mathbf{c}_i = i_{B'}(T(b_i)) \in \mathbb{F}^n \quad , \quad i = 1, \ldots, m$$

Then

(i) *The linear transformation $i_{B'} \circ T \circ i_B^{-1} : \mathbb{F}^m \to \mathbb{F}^n$ coincides with matrix multiplication by $\mathbf{A}_{T,B,B'}$.*
(ii) $\ker(T) = i_B^{-1}(NullSp(\mathbf{A}_{T,B,B'}))$*, where $NullSp(\mathbf{A}_{T,B,B'}) \equiv$ the solution set of $\mathbf{A}_{T,B,B'}\mathbf{x} = \mathbf{0}$ in $\mathbb{F}^m$.*
(iii) $\mathrm{Im}(T) = i_B^{-1}(ColSp(\mathbf{A}_{T,B,B'})) = span(T(b_1,), \ldots, T(b_m))$.

In addition, we also have

COROLLARY 11.7. *If $T : V \to W$ is a linear transformation between finitely generated vector spaces, then*

$$\dim(V) = \ker(T) + \dim \mathrm{Im}(T)$$

Proof. From the theorem,

$$\ker(T) = i_B^{-1}\left(NullSp\left(\mathbf{A}_{T,B,B'}\right)\right)$$

Since i_B is an isomorphism, $\ker(T)$ and the solution set of $\mathbf{A}_{T,B'B'}\mathbf{x} = \mathbf{0}$ must have the same dimension. On the other, by the theorem we also have that $\mathrm{Im}(T)$ is isomorphic to $ColSp\left(\mathbf{A}_{T,B,B'}\right)$ and so $\mathrm{Im}(T)$ must have the same dimension as the column space of $\mathbf{A}_{T,B,B'}$. In other words, $\dim(\mathrm{Im}(T)) = rank\left(\mathbf{A}_{T,B,B'}\right)$. But now from Corollary 8.4 of Lecture 8 we know

$$\dim\left(\text{solution set of } \mathbf{A}_{T,B,B'}\mathbf{x} = \mathbf{0}\right) = \#\text{ columns of } \mathbf{A}_{T,B,B} - rank\left(\mathbf{A}_{T,B,B'}\right) \tag{4}$$

Using

$$\begin{aligned}
\#\text{ columns of } \mathbf{A}_{T,B,B'} &= \#\text{ basis vectors in } B = \dim V \\
rank\left(\mathbf{A}_{T,B,B'}\right) &= \dim\left(ColSp\left(\mathbf{A}_{T,B,B'}\right)\right) = \dim \mathrm{Im}(T) \\
\dim\left(\text{solution set of } \mathbf{A}_{T,B,B'}\mathbf{x} = \mathbf{0}\right) &= \dim(\ker(T))
\end{aligned}$$

We get from (4)

$$\dim(\ker(T)) = \dim V - \dim(\mathrm{Im}(T))$$

and the statement of the corollary follows.

Example 11.8. Find the kernel and image of the following linear transformation acting on polynomials of degree ≤ 3

$$T(p) = x\frac{dp}{dx} - 2p$$

Using the standard basis $B = \left\{x^3, x^2, x, 1\right\}$ for the vector space of polynomial of degree 3 we see

$$\begin{aligned}
x^3 &\xrightarrow{T} \left(x\frac{d}{dx} - 2\right)x^3 &= 3x^3 - 2x^3 = x_3 &\rightarrow [1,0,0,0] \\
x^2 &\xrightarrow{T} \left(x\frac{d}{dx} - 2\right)x^2 &= 2x^2 - 2x^2 = 0 &\rightarrow [0,0,0,0] \\
x &\xrightarrow{T} \left(x\frac{d}{dx} - 2\right)x &= x - 2x = -x &\rightarrow [0,0,-1,0] \\
1 &\xrightarrow{T} \left(x\frac{d}{dx} - 2\right)x &= 0 - 2 = -2 &\rightarrow [0,0,0,-2]
\end{aligned}$$

and so

$$\mathbf{T}_{BB} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

From this is is fairly obvious that

$$\begin{aligned}
NullSp\left(\mathbf{T}_{BB}\right) &= span\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} &\longleftrightarrow\quad span\left(x^2\right) \\
ColSp\left(\mathbf{T}_{BB}\right) &= span\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right) &\longleftrightarrow\quad span\left(x^3, x, 1\right)
\end{aligned}$$

and so

$$\begin{aligned}
\ker(T) &= span\left(x^2\right) \\
range(T) &= span\left(x^3, x, 1\right)
\end{aligned}$$