Math 131: Linear Algebra

Zhixian (Susan) Zhu

3.C. Linear maps as matrices

- Representing a linear map by a matrix
- Sum and scalar multiplication
- Compositions of linear maps

9.1 $\mathtt{Mat}_{n\times m}(\mathbb{F})$

Recall that $\mathtt{Mat}_{n\times m}(\mathbb{F})$ is the vector space of all matrices of size $n \times m$ over the field $\mathbb{F}$. The addition is defined to be the matrix addition (which is adding the corresponding entries) and the scalar product is the scalar product for matrices (which is multiply the number to each entry of the matrix).

Let E_{ij} be the matrix that has 1 in the (i, j)-th entry and is zero elsewhere. Then

$$\{E_{ij}\}_{\substack{1\le i\le n\\ 1\le j\le m}}$$

is a basis of $\mathtt{Mat}_{n\times m}(\mathbb{F})$ which is called the *standard basis*.

Example 1 *The standard basis of $\mathtt{Mat}_{2\times 2}(\mathbb{F})$ is*

$$\left\{E_{11} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, E_{12} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, E_{21} = \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}, E_{22} = \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right\}.$$

Any 2×2 matrix $\begin{bmatrix}a & b\\ c & d\end{bmatrix}$ can be expressed uniquely as $aE_{11} + bE_{12} + cE_{21} + dE_{22}$. Therefore we can write $\begin{bmatrix}a & b\\ c & d\end{bmatrix}$ as a row vector with respect to the standard basis: (a, b, c, d).

In MATH 31, we know how to define sum, scalar multiplication and the product of matrices.

9.2 Matrix representations

Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be a basis of a finite dimensional vector space V and $v \in V$. We can associate a column matrix to v, denoted by $v_{\mathcal{E}}$. We write v as the unique combination of $\{e_1, \ldots, e_n\}$, i.e.,

$$v = c_1e_1 + c_2v_2 \ldots + c_ne_n.$$

Then

$$[v]_{\mathcal{E}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}_{\mathcal{E}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

The c_i is called the i–th coordinate of v with respect to the basis $\mathcal{E}$. Later we will drop $\mathcal{E}$ once it is clear.

Lemma 1 *Let $v = c_1e_1 + c_2v_2 \ldots + c_ne_n$, $u = d_1e_1 + d_2v_2 \ldots + d_ne_n$, and $\lambda \in \mathbb{F}$. then*

$$[v+u]_{\mathcal{E}} = [v]_{\mathcal{E}} + [u]_{\mathcal{E}}.$$

and

$$[\lambda v]_{\mathcal{E}} = \lambda [v]_{\mathcal{E}}.$$

Sketch of the proof:

$$[v+u]_{\mathcal{E}} = \begin{bmatrix} c_1 + d_1 \\ c_2 + d_2 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = [v]_{\mathcal{E}} + [u]_{\mathcal{E}}.$$

$$[\lambda v]_{\mathcal{E}} = \begin{bmatrix} \lambda c_1 \\ \lambda c_2 \\ \vdots \\ \lambda c_n \end{bmatrix} = \lambda [v]_{\mathcal{E}}.$$

Definition 1 Let $L : V \to W$ be a linear map. Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be a basis of V, and $\mathcal{F} = \{f_1, \ldots, f_m\}$ be a basis of W. The matrix representation of L under this two bases $\mathcal{E}$ of V and $\mathcal{F}$ of W is given in the following steps:

1. Apply L to the basis vector e_1: $L(e_1)$.
2. Since L is a map from V to W, $L(e_1)$ should be a vector in W.
3. Since $L(e_1) \in W$, and $\mathcal{F} = \{f_1, \ldots, f_m\}$ be a basis of W, we can write $L(e_1)$ as a linear combination of $\mathcal{F}$.

$$L(e_1) = a_{11}f_1 + a_{21}f_2 + \ldots + a_{m1}f_m = \sum_{i=1}^{m} a_{i1}f_i.$$

$$[L(e_1)]_{\mathcal{F}} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}_{\mathcal{F}}.$$

4. Repeat the process to other basis vectors of V. We have

$$[L(e_2)_{\mathcal{F}}] = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}_{\mathcal{F}}, \dots, [L(e_n)]_{\mathcal{F}} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}_{\mathcal{F}}.$$

5. Put all these column vectors together, i.e., the k–th column is the coordinates of $L(e_i)$ with respect to basis $\mathcal{F}$. we have:

$$[L]_{\mathcal{F}\leftarrow\mathcal{E}} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

It is clear that the matrix depends on the choices of bases $\mathcal{E}$ and $\mathcal{F}$. If there is no confusion, we can simply denote it by $[L]$. In the book, they use the notation $\mathcal{M}(L)$.

Example 2 *Compute a matrix representation for $L \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$ is the map*

$$L(q) = 5q'' + 3q',$$

with respect to the standard bases $\mathcal{E} = \{1, t, t^2, t^3\}$ and $\mathcal{F} = \{1, t, t^2\}$.

Solution:

1. *Apply L to the basis vector $e_1 = 1$,*

$$L(e_1) = 0.$$

2. *Since L is a map from V to W, $L(e_1)$ should be a vector in W. Since $L(e_1) \in W$, and $\mathcal{F} = \{1, t, t^2\}$ be a basis of W, we can write $L(e_1)$ as a linear combination of $\mathcal{F}$:*

$$[L(e_1)] = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{F}}.$$

3. *Repeat this process to other basis vectors of V.*

$$L(e_2) = L(t) = 3,\ L(e_3) = L(t^2) = 10 + 6t,\ L(e_4) = L(t^3) = 30t + 9t^2.$$

We have

$$[L(e_2)] = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{F}}, [L(e_3)] = \begin{bmatrix} 10 \\ 6 \\ 0 \end{bmatrix}_{\mathcal{F}}, [L(e_4)] = \begin{bmatrix} 0 \\ 30 \\ 9 \end{bmatrix}_{\mathcal{F}}.$$

4. *Put all these column vectors together, we have:*

$$[L]_{\mathcal{F}\leftarrow\mathcal{E}} = \begin{bmatrix} 0 & 3 & 10 & 0 \\ 0 & 0 & 6 & 30 \\ 0 & 0 & 0 & 9 \end{bmatrix}.$$

Theorem 1 *Given bases of $\mathcal{E}$ and $\mathcal{F}$ as above. Then we can associate a linear map from V to W to any matrix of size $m \times n$. For any matrix A of size $m \times n$, we can associate a linear map $L \in \mathcal{L}(V, W)$ such that*

$$[L]_{\mathcal{F}\leftarrow\mathcal{E}} = A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

Sketch of the proof: For any vector $v \in V$, we define $L(v)$ by its coordinates with respect to the $\mathcal{F}$ basis. That means

$$[L(v)]_{\mathcal{F}} = A[v]_{\mathcal{E}}.$$

We first interpret the notations.

1. *There are three notations here: $[L]_{\mathcal{F}\leftarrow\mathcal{E}}$, $[L(v)]_{\mathcal{F}}$ and $[v]_{\mathcal{E}}$. $[L(v)]_{\mathcal{F}}$ and $[v]_{\mathcal{E}}$ are the column vectors related to vectors.*

2. *Since we know that any linear map $L(v)$ is determined by the images of e_i under L. We know that if the desired L exists, then the column vector associated to $L(e_i)$ should be the i-th column of A.*

$$L(e_i) = \sum_{j=1}^{m} a_{ji} f_j.$$

   Using linearity, we can define $L(v)$ for any element $v \in V$. We write v as a linear combination of $\mathcal{E}$:

$$v = c_1 e_1 + c_2 e_2 + \dots + c_n e_n = \sum_{i=1}^{n} c_i e_i \text{ for some } c_1, c_2, \dots, c_n \in \mathbb{F}.$$

3. *Then we define*

$$\begin{aligned} L(v) &= \sum_{i=1}^{n} c_i L(e_i) \\ &= \sum_{i=1}^{n} c_i \sum_{j=1}^{m} a_{ji} f_j \\ &= \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ji} c_i \right) f_j \end{aligned}$$

   We can use matrix multiplication to memorize the above formula. Note that

$$[v]_{\mathcal{E}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}_{\mathcal{E}}.$$

We define

$$\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}_{\mathcal{F}} = A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}_{\mathcal{E}} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}_{\mathcal{E}}.$$

Equivalently,for every j,

$$b_j = \sum_{i=1}^{n} a_{ji} c_i.$$

Then in practice, we can use the following formula to calculate $T(v)$.

$$\begin{aligned} L(v) &= \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ji} c_i \right) f_j \\ &= \sum_{j=1}^{m} b_j f_j. \end{aligned}$$

4. *Now we should check that L defined as above is linear. It follows from the matrix multiplications.*

$$[L(v+u)]_{\mathcal{F}} = A[v+u]_{\mathcal{E}} = A[v]_{\mathcal{E}} + A[u]_{\mathcal{E}}.$$

This implies that $L(u+v) = L(u) + L(v)$.

Similarly, let $\lambda \in \mathbb{F}$. Then

$$[L(\lambda v)]_{\mathcal{F}} = A[\lambda v]_{\mathcal{E}} = A\lambda[v]_{\mathcal{E}} = \lambda A[v]_{\mathcal{E}} = \lambda[L(v)]_{\mathcal{F}}.$$

This implies that $L(\lambda v) = \lambda L(v)$.

9.3 Sum, scalar multiplication, and composition of linear maps in terms of matrices

Let $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$. Then the associated matrices has the following relations.

Proposition 1 *Let $\mathcal{E} = \{e_1, \dots, e_n\}$ be a basis of V, and $\mathcal{F} = \{f_1, \dots, f_m\}$ be a basis of W. Then*

$$[S+T]_{\mathcal{F}\leftarrow\mathcal{E}} = [S]_{\mathcal{F}\leftarrow\mathcal{E}} + [T]_{\mathcal{F}\leftarrow\mathcal{E}},$$

and

$$[\lambda S]_{\mathcal{F}\leftarrow\mathcal{E}} = \lambda[S]_{\mathcal{F}\leftarrow\mathcal{E}}.$$

We leave the proof to the reader as a homework problem.

We have seen that there is a one-to-one correspondence between the linear maps from V to W to the matrix space of size $m \times n$, after we choose a basis of V and W. The above argument shows that this correspondence also preserves the addition and scalar multiplication. In section §3.D., we show that $\mathcal{L}(V, W)$ and $\mathtt{Mat}_{m\times n}(\mathbb{F})$ are isomorphic as vector spaces.

Proposition 2 *Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be a basis of U, $\mathcal{F} = \{f_1, \ldots, f_m\}$ be a basis of V and $\mathcal{G} = \{g_1, \ldots, g_l\}$ be a basis of W. Let $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$. Then $ST \in \mathcal{L}(U, W)$ has the matrix*

$$[ST]_{\mathcal{G}\leftarrow\mathcal{E}} = [S]_{\mathcal{G}\leftarrow\mathcal{F}}[T]_{\mathcal{F}\leftarrow\mathcal{E}}.$$

Sketch of the proof: Clearly these two matrices has the same size, $l \times n$. We just need to compare the i–th column for each $1 \leq i \leq n$. Let

$$[S]_{\mathcal{G}\leftarrow\mathcal{F}} = B = \begin{bmatrix} b_{11} & \ldots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{l1} & \ldots & b_{lm} \end{bmatrix},$$

$$[T]_{\mathcal{F}\leftarrow\mathcal{E}} = A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix},$$

and

$$[ST]_{\mathcal{G}\leftarrow\mathcal{E}} = C = \begin{bmatrix} c_{11} & \ldots & c_{ln} \\ \vdots & \ddots & \vdots \\ c_{m1} & \ldots & c_{ln} \end{bmatrix},$$

Our assertion follows from $C = BA$. Note that

$$T(e_i) = \sum_{j=1}^{m} a_{ji} f_j,$$

$$S(f_j) = \sum_{k=1}^{l} b_{kj} g_k,$$

$$ST(e_i) = \sum_{k=1}^{l} c_{ki} g_k.$$

Hence

$$\begin{aligned} \sum_{k=1}^{l} c_{ki} g_k = (ST)(e_i) = S(T(e_i)) &= S(\sum_{j=1}^{m} a_{ji} f_j) \\ &= \sum_{j=1}^{m} a_{ji} S(f_j) \\ &= \sum_{j=1}^{m} a_{ji} (\sum_{k=1}^{l} b_{kj} g_k) \\ &= \sum_{k=1}^{l} (\sum_{j=1}^{m} b_{kj} a_{ji}) g_k \end{aligned}$$

Hence

$$c_{ki} = \sum_{j=1}^{m} b_{kj} a_{ji}.$$

This implies that $C = BA$.

This proposition explains why we use column vectors for matrix representations. The order of composition agrees with the order of matrix product.

9.4 Null space and range space related with matrices

We can also read the null space and range space from the associated matrix.

Definition 2 Let $A \in \mathtt{Mat}_{n\times m}(\mathbb{F})$. The *column space* of A is the subspace of $\mathbb{F}^n$ spanned by the column vectors of A and is denoted by $\mathtt{Col}(A)$.

The *null space* of A is the subspace of $\mathbb{F}^m$ which is the solution space of the equation $Ax = 0$ and is denoted by $\mathtt{null}(A)$.

Example 3 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} \in \mathtt{Mat}_{2\times 3}(\mathbb{R})$, which corresponding to a linear map $L : \mathbb{R}^3 \to \mathbb{R}^2$. Then $\mathtt{Col}(A)$ is the subspace of $\mathbb{R}^2$ spanned by $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \end{bmatrix} \right\}$. Since this set is linearly dependent. Then

$$\mathtt{Col}(A) = \mathtt{Span}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \left\{ \begin{bmatrix} a \\ a \end{bmatrix} : a \in \mathbb{R} \right\}.$$

To find $\mathtt{null}(A)$, we need to solve the equation $Ax = 0$, where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The solution is $x_1 = -2x_2 - 3x_3$. Then we can write

$$\begin{cases} x_1 = -2x_2 - 3x_3 \\ x_2 = x_2 \\ x_3 = x_3. \end{cases}$$

or

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

So $\mathtt{null}(A) = \mathtt{Span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Remark 1 Let A be a matrix of size $m \times n$ and L be a linear map from $\mathbb{F}^n$ to $\mathbb{F}^m$ defined by the matrix A. Then

$$\mathtt{range}(L) = \mathtt{Col}(A), \text{ and } \mathtt{null}(L) = \mathtt{null}(A).$$

In particular, $\texttt{rank}(A) = \dim \texttt{Col}\, A = \dim \texttt{range}(L)$ and

$$\texttt{nullity(A)} = \dim \texttt{null(A)} = \dim \texttt{null(L)}.$$

Example 4 *Let $V = \mathbb{R}^3$. Let $L : V \to V$ be an operator on V. Suppose L is defined by the matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Then the null space is $\texttt{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$, and the range is $\texttt{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$. Then we have $\dim \texttt{null}(L) = 1$, $\dim \texttt{range}(L) = 2$, $\dim V = 3$, and*

$$\dim_{\mathbb{F}} V = 3 = 1 + 2 = \dim_{\mathbb{F}} \texttt{null}(L) + \dim_{\mathbb{F}} \texttt{range}(L).$$

Remark 2 Note that the formula of Fundamental Theorem of Linear Map is a direct generalization of the rank formula discussed in Math 031: Let A be a matrix of size $m \times n$, then

$$n = \texttt{nullity(A)} + \texttt{rank(A)}.$$

Remark 3 In worksheet 8, problem 3, we will see that for any $L : V \to W$ if we choose bases $\mathcal{E}$ and $\mathcal{F}$ cleverly, then the associated matrix is in a simple form:

$$[L]_{\mathcal{F} \leftarrow \mathcal{E}} = \begin{bmatrix} I_{d \times d} & 0_{d \times (n-d)} \\ 0_{(m-d) \times d} & 0_{(m-d) \times (n-d)} \end{bmatrix},$$

where $d = \dim \texttt{range}(L)$ and $I_{d \times d}$ is the identity matrix of size d by d, $0_{d \times (n-d)}$ is the zero matrix of size d by $n - d$, etc.

When L is a linear operator on a vector space V, we usually only use one basis of V, say $\mathcal{E}$. Then we consider the associated matrix $[L]_{\mathcal{E} \leftarrow \mathcal{E}}$. In general, this matrix can't be as simply as the previous one. One of the important topics in Math 132 is to study the Jordan form associated to such a linear operator, see section §8.D.