18.700 - LINEAR ALGEBRA, DAY 8
MATRICES

SAM SCHIAVONE

CONTENTS

I. PRE-CLASS PLANNING

I.1. Goals for lesson.

(1) Students will learn the definition of the range of a linear map.
(2) Students will learn the Rank-Nullity Theorem.

I.2. Methods of assessment.

(1) Student responses to questions posed during lecture
(2) Student responses to worksheet

I.3. Materials to bring.
(1) Laptop + adapter (2) Worksheets

(0:00)

II. Lesson Plan

Announcements: • Final exam: Monday, December 16, 1:30 - 4:30pm, 6-120

II.1. **Last time.**

- Gave the definition of dimension.
- Proved some basic properties about dimension.
- Defined linear maps.
- Show that a linear map is uniquely determined by its action on a basis.
- Defined the null space of a linear map.
- T one-to-one $\iff \ker(T) = \{0\}$.

II.2. **The Rank-Nullity Theorem.**

Definition 1. Let X and Y be sets and $f : X \to Y$ be a function. The *range* or *image* of f is

$$\operatorname{range}(f) = \operatorname{img}(f) = f(X) := \{f(x) : x \in X\} .$$

[Draw picture of blobs, showing that $\operatorname{img}(f)$ need not fill up all of Y.]

Definition 2. If $\operatorname{img}(f) = Y$, then f is *onto* or *surjective*.

Remark 3. Warning: You must specify the codomain for the notion of surjectivity to make sense! E.g., $f(x) = x^2$ as a function $\mathbb{R} \to \mathbb{R}$ or $\mathbb{R} \to [0, \infty)$.

Definition 4. Let $T : V \to W$ be linear. The dimension of $\operatorname{range}(T)$ is called the *rank* of T.

Lemma 5. *If $T : V \to W$ is linear, then $\operatorname{img}(T)$ is a subspace of W.*

Proof. Exercise. Apply subspace criterion. □

II.2.1. *Rank-nullity theorem.* The sizes of the kernel and the image are inversely correlated. E.g., the zero map $0 : V \to W$ has large null space [ask students]—all of V—and small range—just $\{0\}$. On the other hand, the identity map $I : V \to V$ has small kernel [ask students]—just $\{0\}$—and large image—all of V. This relationship is captured precisely in the following result.

Theorem 6 (Rank-Nullity Theorem)**.** *Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then $\operatorname{img}(T)$ is also finite-dimensional and*

$$\dim(V) = \dim(\ker(T)) + \dim(\operatorname{img}(T)) .$$

In words, the dimension of the domain of T is equal to the sum of the nullity and rank of T.

Proof. Let $u_1, \ldots, u_m$ be a basis of $\ker(T)$. By the Extension Theorem, we can extend this to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V. Thus $\dim(\ker(T)) = m$ and $\dim(V) = m + n$, so it suffices to show that $\dim(\operatorname{img}(T)) = n$.

We claim that $T(v_1), \ldots, T(v_n)$ is a basis for $\operatorname{img}(T)$. [Ask students why we don't include any u_i.] Given $v \in V$, then

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n$$

for some scalars $a_i, b_j \in \mathbb{F}$. Applying T, we have

$$T(v) = T\left(\sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} b_j v_j\right) = \sum_i a_i \cancelto{0}{T(u_i)} + \sum_j b_j T(v_j)$$

since the u_i all map to 0 since they are in $\ker(T)$. Thus $T(v_1), \ldots, T(v_n)$ spans $\operatorname{img}(T)$, hence it is finite-dimensional.

It remains to show they are linearly independent. Suppose there exists $c_1, \ldots, c_n \in \mathbb{F}$ such that

$$0 = c_1 T(v_1) + \cdots + c_n T(v_n) = T(c_1 v_1 + \cdots + c_n v_n) .$$

Then $\sum_{k=1}^{n} c_k v_k \in \ker(T)$, so there exist $d_1, \ldots, d_m \in \mathbb{F}$ such that

$$c_1 v_1 + \cdots + c_n v_n = d_1 u_1 + \cdots d_m u_m .$$

Then

$$0 = d_1 u_1 + \cdots d_m u_m - c_1 v_1 - \cdots - c_n v_n .$$

Since $u_1, \ldots, u_m, v_1, \ldots, v_n$ is a basis of V, hence linearly independent, then $0 = c_1 = \cdots, c_n = d_1 = \cdots = d_m$. Thus $T(v_1), \ldots, T(v_n)$ is linearly independent, hence is a basis of $\operatorname{img}(T)$. □

Corollary 7. *Suppose V and W are finite-dimensional vector spaces with $\dim(V) > \dim(W)$. Then no linear map $T : V \to W$ is injective.*

Proof. By the Rank-Nullity Theorem, then

$$\dim(\ker(T)) = \dim(V) - \dim(\operatorname{img}(T)) .$$

Since $\operatorname{img}(T) \subseteq W$, then $\dim(\operatorname{img}(T)) \leq \dim(W)$, so $-\dim(W) \leq -\dim(\operatorname{img}(T))$. Then

$$\dim(\ker(T)) = \dim(V) - \dim(\operatorname{img}(T)) \geq \dim(V) - \dim(W) > 0 .$$

Thus $\ker(T) \neq \{0\}$, so T is not injective. □

Corollary 8. *Suppose V and W are finite-dimensional vector spaces with $\dim(V) < \dim(W)$. Then no linear map $T : V \to W$ is surjective.*

Proof. Exercise. Similar to the above. □

II.2.2. *Linear maps and systems of equations.* Fix $m, n \in \mathbb{Z}_{\geq 0}$ and suppose $A_{ij} \in \mathbb{F}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Consider the homogeneous system of equations

$$\sum_{k=1}^{n} A_{1k} x_k = 0$$

$$\vdots$$

$$\sum_{k=1}^{n} A_{mk} x_k = 0 .$$

We can interpret this system using linear maps by defining

$$T : \mathbb{F}^n \to \mathbb{F}^m$$

$$(x_1, \ldots, x_n) \mapsto \left(\sum_{k=1}^{n} A_{1k} x_k, \ldots, \sum_{k=1}^{n} A_{mk} x_k \right) .$$

One can show that T is linear. Thus the above linear system is equivalent to the equation $T(x) = 0$. Therefore computing the solutions to the system is the same as computing [ask students] $\ker(T)$.

Let $b_1, \dots, b_m \in \mathbb{F}$. Similarly, we can reinterpret searching for solutions to the linear system

$$\sum_{k=1}^{n} A_{1k} x_k = b_1$$
$$\vdots$$
$$\sum_{k=1}^{n} A_{mk} x_k = b_m$$

as asking whether $b := (b_1, \dots, b_m)$ is in $\operatorname{img}(T)$ or not.

II.3. **Worksheet.**

II.4. **The matrix of a linear map.**

Definition 9. Let $\mathcal{B} := (u_1, \dots, u_n)$ be a basis for the finite-dimensional vector space V. Given $x \in V$, let $a_1, \dots, a_n \in \mathbb{F}$ be the unique scalars such that

$$x = \sum_{i=1}^{n} a_i u_i \,.$$

The *coordinate vector* or *matrix of x with respect to $\mathcal{B}$*, denoted by $[x]_{\mathcal{B}}$, is

$$[x]_{\mathcal{B}} := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{F}^n \,.$$

[Ask students: what is $[u_1]_{\mathcal{B}}$?]

Example 10 (Polynomial example with respect to $1, x, x^2$.)**.**

Similarly, we can encode linear maps as matrices.

Definition 11. Let $m, n \in \mathbb{Z}_{\geq 0}$. An *$m$-by-$n$ matrix A* is a rectangular array of elements of $\mathbb{F}$ with m rows and n columns:

$$\begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} .$$

The entry in row i, column j is called the (i, j)-entry and is denoted A_{ij}.

Suppose V and W are finite-dimensional vector spaces with bases $\mathcal{B} := (v_1, \dots, v_n)$ and $\mathcal{C} := (w_1, \dots, w_m)$, respectively. Let $T : V \to W$ be linear. We have seen that the values $T(v_1), \dots, T(v_n)$ uniquely determine T.

Definition 12. With notation as above, for each $j = 1, \dots, n$ there are unique scalars $A_{1j}, \dots, A_{mj} \in \mathbb{F}$ such that

$$T(v_j) = A_{1j}w_1 + \cdots + A_{mj}w_m .$$

The matrix whose i, j entry is A_{ij} is the *matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$* and is denoted ${}_{\mathcal{C}}[T]_{\mathcal{B}}$ or $\mathcal{M}(T; \mathcal{B}, \mathcal{C})$. When the bases are clear from context, we simply write $[T]$ or $\mathcal{M}(T)$.

Note that the j^{th} column of $[T]$ is the coordinate vector of $T(v_j)$:

$$ {}_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{pmatrix} | & & | \\ [T(v_1)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & & | \end{pmatrix}$$

- For $V = \mathbb{F}^n$, let $\mathcal{E}_n$ be the basis

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} .$$

  We call this the *standard basis* for $\mathbb{F}^n$ and use this basis unless otherwise specified.
- For $V = \mathcal{P}_m(\mathbb{F})$, we similarly use the standard monomial basis $1, x, x^2, \dots, x^m$ unless otherwise specified.

Example 13. Let $T : \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_1(\mathbb{R})$ be the differentiation map $T(f) = f'$. Then

$$\begin{aligned} T(1) &= 0 = 0 \cdot 1 + 0 \cdot x \\ T(x) &= 1 = 1 \cdot 1 + 0 \cdot x \\ T(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x \end{aligned}$$

so

$$[T] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} .$$

We define addition and scalar multiplication of matrices to agree with the corresponding operations for linear maps.

Definition 14.
- Given matrices A and B of the same size, we define their sum entrywise, i.e., $A + B$ is the matrix whose i, j entry is $A_{ij} + B_{ij}$.
- Given a matrix A and $\lambda \in \mathbb{F}$, we define λA to be the matrix whose i, j entry is λA_{ij}.

Lemma 15. *Suppose $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$. Then*
- $[S + T] = [S] + [T]$, *and*
- $[\lambda T] = \lambda[T]$.

Proof. Exercise. ∎

Denote the set of all $m \times n$ matrices by $\mathrm{M_{m\times n}}(\mathbb{F})$ or $\mathrm{Mat}_{m\times n}$ or $\mathbb{F}^{m,n}$.

Lemma 16. *With the operations of addition and scalar multiplication above,* $\mathrm{M_{m\times n}}(\mathbb{F})$ *is a vector space of dimension mn.*

Proof. Exercise. □

II.4.1. *Matrix multiplication.* Suppose that U, V, W are finite-dimensional vector spaces with bases

$$\begin{aligned} \mathcal{B} &:= (u_1, \dots, u_p) \\ \mathcal{C} &:= (v_1, \dots, v_n) \\ \mathcal{D} &:= (w_1, \dots, w_m) . \end{aligned}$$

Suppose $T : U \to V$ and $S : V \to W$ are linear maps. We previously saw that the composition $ST : U \to W$ is linear.We define matrix multiplication in such a way that

$$[ST] = [S][T] .$$

Let $A := [S]$ and $B := [T]$. Then for each $j = 1, \dots, p$ we have

$$(ST)(u_j) = S\left(\sum_{k=1}^{n} B_{kj} v_k\right) = \sum_{k=1}^{n} B_{kj} S(v_k) = \sum_{k=1}^{n} B_{kj} \sum_{i=1}^{m} A_{ik} w_k = \sum_{i=1}^{m}\sum_{k=1}^{n} \left(A_{ik}B_{kj}\right) w_k .$$

Thus $[ST]$ is the $m \times p$ matrix whose i, j entry is $\sum_{k=1}^{n} \left(A_{ik}B_{kj}\right)$.

Definition 17. Given an $m \times n$ matrix A and a $n \times p$ matrix B, their product AB is defined to be the $m \times p$ matrix whose i, j entry is $\sum_{k=1}^{n} \left(A_{ik}B_{kj}\right)$.

So we multiply the entries of row j of A by those of column k of B, then add these together.

Example 18.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 6 & 5 & 0 & 0 \\ 2 & 1 & 0 & -1 \end{pmatrix} = \cdots$$

[Ask students about the other order BA.]

Proposition 19. *If* $T \in \mathcal{L}(U, V)$ *and* $S \in \mathcal{L}(V, W)$*, then* $[ST] = [S][T]$.

Proof. This is true by the definition of matrix multiplication and the earlier calculation done as motivation. □

Let A be an $m \times n$ matrix.

- For $i = 1, \dots, m$, let $A_{i,\cdot}$ denote row i of A, which is a $1 \times n$ matrix.
- For $j = 1, \dots, n$, let $A_{\cdot,j}$ denote column j of A, which is an $m \times 1$ matrix.

The next few results give different interpretations of matrix multiplication. Let A be an $m \times n$ matrix and B be an $n \times p$ matrix.

Lemma 20.

$$(AB)_{ij} = A_{i,\cdot}B_{\cdot,j}$$

for all $i = 1, \dots, m$ and all $j = 1, \dots, p$. [Draw picture of row and column.]

Proof. True by formula defining matrix multiplication. □

Lemma 21.

$$(AB)_{\cdot,j} = A(B_{\cdot,j})$$

for all $j = 1, \dots, p$.

Proof. Exercise. Both are $m \times 1$ matrices. Check that their i^{th} entries are equal using the formula. [Draw picture applying A to each of the columns of B.] □

Lemma 22. *Suppose A is $m \times n$ and $x = \begin{pmatrix} x_1 \\ \ddots\, x_n \end{pmatrix}$ is $n \times 1$. Then*

$$Ax = x_1 A_{\cdot 1} + \cdots + x_n A_{\cdot n}\,.$$

I.e., Ax is the linear combination of the columns of A with coefficients given by the entries of x.

Proof. Exercise. □

Lemma 23. *(a) For $j = 1, \dots, p$, $(AB)_{\cdot,j}$ (column j) is a linear combination of the columns of A with coefficients from $B_{\cdot,j}$ (column j).*

(b) For $i = 1, \dots, m$, $(AB)_{i,\cdot}$ (row i) is a linear combination of the rows of B with coefficients from $A_{i,\cdot}$ (row i).

Proof. Exercise. [Draw picture of second part.] □

Definition 24.

- The *column space* of A, denoted $\mathrm{Col}(A)$, is the span of the columns of A. The *column rank* is the dimension of $\mathrm{Col}(A)$.
- The *row space* of A, denoted $\mathrm{Row}(A)$, is the span of the rows of A. The *row rank* is the dimension of $\mathrm{Row}(A)$.

We'll see that these two quantities are actually equal!

Definition 25. The *transpose* of a matrix A, denoted A^t, is obtained from A by interchanging rows and columns. I.e.,

$$(A^t)_{ij} = A_{ji}\,.$$

Lemma 26 (Column-row factorization)**.** *Suppose A is $m \times n$ and has column rank $c \in \mathbb{Z}_{\geq 1}$. Then there exist an $m \times c$ matrix C and a $c \times n$ matrix R such that $A = CR$.*

Proof. The columns $A_{\cdot,1}, \dots, A_{\cdot,n}$, each an $m \times 1$ matrix, span $\mathrm{Col}(A)$. By a previous result, this list can be reduced to a basis $v_1, \dots, v_c$ of $\mathrm{Col}(A)$, which by definition must have length c. Use these as the columns of a $m \times c$ matrix C.

For $k = 1, \dots, n$, column k of A is a linear combination of the columns of C (since these are a basis), so there exist scalars $R_{1k}, \dots, R_{ck} \in \mathbb{F}$ such that

$$A_k = R_{1k} v_1 + \cdots + R_{ck} v_c\,.$$

Use the coefficients $R_{1k}, \dots, R_{ck}$ as the entries of the k^{th} column of a $c \times n$ matrix R. Then $A = CR$. □

Theorem 27 (Column rank = row rank)**.** *Suppose* $A \in \mathrm{M}_{\mathrm{m}\times\mathrm{n}}(\mathbb{F})$*. Then the column rank and row rank of* A *are equal.*

Proof. Let c be the column rank of A. Let $A = CR$ be the column-row factorization of A given by the previous lemma, where C is $m \times c$ and R is $c \times n$. Since every row of A can be written as a linear combination of the rows of R, and R has c rows, then the row rank of A is $\leq c$, which is the column rank of A.

We obtain the reverse inequality by applying the same argument to A^t, which yields

$$\text{column rank of } A = \text{row rank of } A^t \leq \text{column rank of } A^t = \text{row rank of } A.$$

□

Definition 28. The *rank* of a matrix is its column rank (= its row rank).