Math 24: Winter 2021
Lecture 9

Dana P. Williams

Dartmouth College

Wednesday, January 27, 2021

Let's Get Started

1. We should be recording.
2. Remember, it is more comfortable for me if you turn on your video so that I feel like I am talking to real people.
3. Our preliminary exam will be available from Thursday at 11am (after office hours) and must be submitted by Saturday at 10pm EST. You will have 150 minutes to work on the exam with an extra 30 minutes for scanning and submitting via `gradescope`. The exam will cover through and including §2.2 in the text (which I will finish Today).
4. But first, are there any questions from last time?

Review

Definition

An ordered basis for a finite dimensional vector space V of dimension n is a basis $\{ v_1, v_2, \dots, v_n \}$ considered as an (ordered) n-tuple.

Notation

Let $\beta = \{ v_1, \dots, v_n \}$ be an ordered basis for a vector space V. Then for each $x \in V$, there is a unique vector $(a_1, \dots, a_n) \in \mathbf{F}^n$—that is, an ordered n-tuple—$(a_1, \dots, a_n) \in \mathbf{F}^n$ such that

$$x = \sum_{k=1}^{n} a_k v_k.$$

We call $(a_1, \dots, a_n)$ the coordinate vector of x relative to β. We use the notation $[x]_\beta = (a_1, \dots, a_n)$ or $[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$. You proved on homework that the map $x \mapsto [x]_\beta$ is a one-to-one and onto linear transformation from V to $\mathbf{F}^n$ (where $n = \dim(V)$).

The Matrix of a Linear Transformation

Definition

Suppose that V and W are finite-dimensional vector spaces with ordered bases $\beta = \{ v_1, \dots, v_n \}$ and $\gamma = \{ w_1, \dots, w_m \}$, respectively. (Note that $\dim(V) = n$ and $\dim(W) = m$.) Then the matrix of T with respect to β and γ is the $m \times n$-matrix $[T]_\beta^\gamma$ whose j^{th}-column is the coordinate vector $[T(v_j)]_\gamma$. When $V = W$ and $\beta = \gamma$, then we usually write $[T]_\beta$ in place of $[T]_\beta^\beta$.

Remark

This is easier to make sense of if we agree—as do the authors of our text—to think $[T(v_j)]$ as a column vector. Then

$$[T]_\beta^\gamma = \big[[T(v_1)]_\gamma \; [T(v_2)]_\gamma \; \cdots \; [T(v_n)]_\gamma \big].$$

Old Friends

Definition

We let the symbol δ_{ij} be the Kronecker delta where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \text{ and} \\ 0 & \text{if } i \neq j. \end{cases}$$

Then the $n \times n$ identity matrix is the matrix $I_n \in M_{n\times n}(\mathbf{F})$ with $(I_n)_{ij} = \delta_{ij}$.

Example

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The Identity Transformation

Remark

It is a simple matter of untangling definitions to see that if $I_V : V \to V$ is the identity transformation and β is any (finite) ordered basis for V, then $[I_V]_\beta = I_n$ (where $n = \dim(V)$). If $T_0 : V \to W$ is the zero transformation between finite-dimensional vector spaces, then $[T_0]_\beta^\gamma = O$ for any ordered bases β and γ of V and W, respectively. Here of course, O is the $\dim(W) \times \dim(V)$ zero matrix.

Linear Transformations

Proposition

Suppose V and W are vector spaces over $\mathbf{F}$. We let $\mathcal{L}(V, W)$ be the set of all linear transformations $T : V \to W$. Suppose that $T, S \in \mathcal{L}(V, W)$ and $a \in \mathbf{F}$. If we define $(T + S)(x) = T(x) + S(x)$ and $(aT)(x) = aT(x)$, then $T + S \in \mathcal{L}(V, W)$ and $aT \in \mathcal{L}(V, W)$. With respect to these operations, $\mathcal{L}(V, W)$ is a vector space over $\mathbf{F}$ with zero vector given by the zero transformation and additive inverse given by $(-T)(x) = -T(x)$. We usually write $\mathcal{L}(V)$ in place of $\mathcal{L}(V, V)$.

Sketch of the Proof.

Seeing that $T + S$ and aT are linear if both T and S are is routine. For example, $(T + S)(ax + y) = T(ax + y) + S(ax + y) = aT(x) + T(y) + aS(x) + S(y) = a(T + S)(x) + (T + S)(y)$. It remains to check axioms VS1–VS8, but I think that that is best done in private. (The proof is practically identical to that for $\mathscr{F}(X, \mathbf{F})$. For example, $0_{\mathcal{L}(V,W)}$ is the zero map T_0.) □

Matrices Rule

Theorem

Suppose that V and W are finite-dimensional vector spaces over $\mathbf{F}$ with ordered bases $\beta = \{ v_1, \dots, v_n \}$ and $\gamma = \{ w_1, \dots, w_m \}$, respectively. Then the map $T \mapsto [T]_\beta^\gamma$ is a one-to-one and onto linear transformation from $\mathcal{L}(V, W)$ to $M_{m\times n}(\mathbf{F})$.

Remark

Here, "linearity" amounts to showing that

$$[T + S]_\beta^\gamma = [T]_\beta^\gamma + [S]_\beta^\gamma \quad \text{and} \quad [aT]_\beta^\gamma = a[T]_\beta^\gamma.$$

Proof

Proof.

Note that if $T(v_j) = \sum_{i=1}^m a_{ij} w_i$ then $\big([T]_\beta^\gamma\big)_{ij} = a_{ij}$. Thus if $S(b_j) = \sum_{i=1}^m b_{ij} w_i$, then $(T+S)(v_j) = \sum_{i=1}^m (a_{ij} + b_{ij}) w_i$. But this just says

$$\big([T+S]_\beta^\gamma\big)_{ij} = \big([T]_\beta^\gamma\big)_{ij} + \big([S]_\beta^\gamma\big)_{ij}.$$

Similarly,

$$\big([aT]_\beta^\gamma\big)_{ij} = a\big([T]_\beta^\gamma\big)_{ij}.$$

This proves $T \mapsto [T]_\beta^\gamma$ is linear.
If $[T]_\beta^\gamma = O$, then clearly $T = T_0$ the zero transformation. Thus our map is one-to-one.
If $A = (A_{ij}) \in M_{m\times n}(\mathbf{F})$, we can let $T : V \to W$ be the unique linear transformation such that $T(v_j) = \sum_{i=1}^m A_{ij} w_i$. Then you can check that $[T]_\beta^\gamma = A$. So the map is onto as well.

Break Time

Time for a break and some questions.

Composition

Definition

If $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, Z)$, then their composition $S \circ T$ is the function $S \circ T(x) = S\big(T(x)\big)$. In Math 24, we will normally write ST in place of $S \circ T$.

Proposition

If $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, Z)$, then $ST \in \mathcal{L}(V, Z)$

Proof.

$$ST(ax + y) = S\big(T(ax + y)\big) = S\big(aT(x) + T(y)\big) = aS(T(x)) + S(T(y)) = aST(x) + ST(y).$$

□

Remark

This gives a nice "multiplication" operation in $\mathcal{L}(V) := \mathcal{L}(V, V)$.

Low Hanging Fruit

Theorem

Let V, W, and Z be vector spaces over $\mathbf{F}$. Suppose $T_1, T_2 \in \mathcal{L}(V, W)$ and $S_1, S_2 \in \mathcal{L}(W, Z)$. Then

1. $(S_1 + S_2)T_1 = S_1T_1 + S_2T_1$,
2. $S_1(T_1 + T_2) = S_1T_1 + S_1T_2$,
3. $I_W T_1 = T_1 = T_1 I_V$,
4. *for all* $a \in \mathbf{F}$, $a(S_1T_1) = (aS_1)T_1 = S_1(aT_1)$, *and*
5. *if* $U_1 \in \mathcal{L}(Z, Z')$, *we have* $U_1(S_1T_1) = (U_1S_1)T_1$.

Proof.

I'll leave this as an exercise. □

Matrix Multiplication

Definition

Let $A \in M_{m\times n}(\mathbf{F})$ and $B \in M_{n\times p}(\mathbf{F})$, then AB is the matrix in $M_{m\times p}(\mathbf{F})$ whose $(i,j)^{\text{th}}$-entry is

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}.$$

Remark

Keep in mind the idea that $(m \times n)(n \times p) = (m \times p)$. Let's do some examples in the document camera.

Why Such a Crazy Definition?

Theorem

Suppose that V, W, and Z are finite-dimensional vector spaces with ordered bases β, γ, and α, respectively. Suppose that $T : V \to W$ and $S : W \to Z$ are linear. Then

$$[ST]_\beta^\alpha = [S]_\gamma^\alpha [T]_\beta^\gamma. \qquad (\dagger)$$

Proof.

Let $\beta = \{ v_1, \ldots, v_n \}$, $\gamma = \{ w_1, \ldots, w_m \}$, and $\alpha = \{ z_1, \ldots, w_p \}$. Then $[T]_\beta^\gamma$ is a $m \times n$-matrix and $[S]_\gamma^\alpha$ is a $p \times m$-matrix. Hence the right-hand side of $(\dagger)$ is defined and equals a $p \times n$-matrix which at the very least is the same shape as $[ST]_\beta^\alpha$. To keep the notation under control, let $[T]_\beta^\gamma = A = (A_{ij})$ and $[S]_\gamma^\alpha = B = (B_{ij})$.

Proof

Proof Continued.

We have $T(v_j) = \sum_{k=1}^{m} A_{kj} w_k$ and $S(w_k) = \sum_{i=1}^{p} B_{ik} z_i$. But then

$$\begin{aligned} ST(v_j) = S\Big(\sum_{k=1}^{m} A_{kj} w_k\Big) &= \sum_{k=1}^{m} A_{kj} S(w_k) \\ &= \sum_{k=1}^{m} \sum_{i=1}^{p} A_{kj} B_{ik} z_i \\ &= \sum_{i=1}^{p} \Big(\sum_{k=1}^{m} B_{ik} A_{kj}\Big) z_i \\ &= \sum_{i=1}^{p} (BA)_{ij} z_i \end{aligned}$$

Therefore $\big([ST]_{\beta}^{\alpha}\big)_{ij} = (BA)_{ij} = \big([S]_{\gamma}^{\alpha}[T]_{\beta}^{\alpha}\big)_{ij}$ as we wanted to show. □

Algebra of Matrix Multiplication

Theorem

The following always hold for matrices of the appropriate shapes.

1. $A(B + C) = AB + AC$ *and* $D(E + F) = DE + DF$.
2. $A(BC) = (AB)C$.
3. $aAB = (aA)B = A(aB)$ *for all* $a \in \mathbf{F}$.
4. $I_m A = A = A I_n$.
5. *If* AB *is defined, then so is* $B^t A^t$ *and* $(AB)^t = B^t A^t$.

Proof.

Except for item 2, these are routine exercises (sketched in the text). We will give a proof of item 2 later. Note for example, if A is a $m \times n$-matrix, then in item 1, we must have B and C $n \times p$-matrices for some $p \geq 1$. Similarly, in item 4, it is implicit that A is a $m \times n$-matrix. □

Hold on Tight . . . It Gets Bumpy from Here

Remark

For matrix multiplication, it is no longer true that $AB = BA$. In general, if one side is defined, the other may not be! If A is a 3×2-matrix and B is a 2×3-matrix, then AB is 3×3 while BA is 2×2. Worse, if $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, then $AB = O$ while $BA = A \neq O$. So even for 2×2-matrices, we usually have $AB \neq BA$. Also notice that $B(A - I_2) = O$, but neither B nor $A - I_2$ is the zero matrix. All this means that matrix algebra is more "interesting" that working with fields or vector spaces. For example, if A and B are square matrices of the same shape, then

$$(A + B)^2 := (A + B)(A + B) = A^2 + AB + BA + B^2$$

and this will usually not equal $A^2 + 2AB + B^2$.

Fun Facts

Remark

If A is a $m \times n$-matrix and $u \in \mathbf{F}^n$, then we can—and the text always does—view u as a $n \times 1$-matrix—aka a "column vector"—so that Au is a $m \times 1$-matrix whose i^{th}-entry is $\sum_{j=1}^n A_{ij} u_j$. (You should take the time to verify all this!!) In particular, if $(e_1, \dots, e_n)$ is the standard ordered basis for $\mathbf{F}^n$, then the j^{th} column of A is Ae_j. Thus we could write $A = \begin{bmatrix} Ae_1 & Ae_2 & \cdots & Ae_n \end{bmatrix}$.

Proposition

Suppose that $A = [u_1 \ \cdots \ u_n]$ *is a* $m \times n$ *matrix with columns* $u_1, \dots, u_n$. *If* B *is a* $p \times n$*-matrix, then* BA *is the* $p \times m$ *matrix with columns equal to* Bu_j. *That is, in my notation,*

$$BA = [Bu_1 \ Bu_2 \ \cdots \ Bu_n].$$

Proof

Proof.

The proof amounts to untangling notation. The j^{th}-column of BA is

$$\begin{pmatrix} (BA)_{1j} \\ \vdots \\ (BA)_{pj} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} B_{1k}A_{kj} \\ \vdots \\ \sum_{k=1}^{n} B_{pk}A_{kj} \end{pmatrix} = B \begin{pmatrix} A_{1j} \\ \vdots \\ A_{nj} \end{pmatrix} = Bu_j.$$

This completes the proof. □

Another Big Pay-Off

Theorem

Suppose that V and W are finite-dimensional vector spaces over $\mathbf{F}$ *and that $T : V \to W$ is a linear transformation. Let β and γ be ordered bases for V and W, respectively. Then for all $v \in V$, we have*

$$[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta$$

Remark

You would be wise to note carefully the placement of the superscripts and subscripts.

Proof

Proof.

Let $U : \mathbf{F} \to V$ be the linear map given by $U(a) = av$. Let $\alpha = \{1\}$ be the standard ordered basis for $\mathbf{F}$. Then $TU : \mathbf{F} \to W$ is linear and

$$\begin{aligned}[T(v)]_\gamma &= [TU(1)]_\gamma = [TU]_\alpha^\gamma \\ &= [T]_\beta^\gamma [U]_\alpha^\beta = [T]_\beta^\gamma [U(1)]_\beta \\ &= [T]_\beta^\gamma [v]_\beta.\end{aligned}$$

□

Example

Example

Let $T : P_2(\mathbf{R}) \to P_2(\mathbf{R})$ be the linear map $T(p) = p + p'$ from Monday. Let $\sigma = \{ 1, x, x^2 \}$ be the standard ordered basis for $P_n(\mathbf{R})$. Then

$$[T]_\sigma = \big[[T(1)]_\sigma \; [T(x)]_\sigma \; [T(x^2)]_\sigma \big] = \big[[1]_\sigma \; [1+x]_\sigma \; [2x + x^2]_\sigma \big]$$
$$= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then if $p(x) = 1 + 3x + 2x^2$,

$$[T(p)]_\sigma = [T]_\sigma [p]_\sigma = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 2 \end{pmatrix}.$$

Hence $T(p) = 4 + 3x + 2x^2$.

Enough

1. That is enough for today.