Math 24: Winter 2021
Lecture 13

Dana P. Williams

Dartmouth College

Friday, February 5, 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. But first, are there any questions from last time?

Review

Theorem

Let $A \in M_{m\times n}(\mathbf{F})$ and let P and Q be invertible $m \times m$- and $n \times n$-matrices, respectively. Then

$$\operatorname{rank}(A) = \operatorname{rank}(PA) = \operatorname{rank}(AQ) = \operatorname{rank}(PAQ).$$

Corollary

Elementary row and column operations preserve the rank of a matrix.

Rank

Example

Consider $A = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 1 & 2 & 1 & 3 \\ 1 & 1 & 0 & 1 \end{pmatrix}$. It is a worthwhile exercise to see that we can use elementary row and column operations to transform A to the matrix $D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ and that D has the "block" form $D = \begin{pmatrix} I_2 & O_1 \\ O_2 & O_3 \end{pmatrix}$ where the O_k are zero matrices of the appropriate form. Of course, it is now abundantly clear that $\operatorname{rank}(A) = 2$.

Rank Theorem

Theorem (Rank Theorem)

Let A be a $m \times n$-matrix of rank r.

1. *We have* $r = \operatorname{rank}(A) \leq \min\{\, m, n \,\}$.
2. *We can transform A by finitely many elementary row and column operations into a $m \times n$-matrix*
$$D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix} \tag{‡}$$
where I_r is the $r \times r$ identity matrix and O_1, O_2, and O_3 are zero matrices of the appropriate shape.
3. *In fact there is a $m \times m$ invertible matrix B and a $n \times n$ invertible matrix C such that $BAC = D$ where D is as in (‡)*

Proof

Proof.

We will start by proving item (2).

If $A = O$, the zero matrix, then $r = 0$ and we can just let $D = A$.

So we assume $A \neq 0$ and proceed by induction on m the number of rows of A.

If $m = 1$, then since A is non-zero, we can use an elementary column operation of type 2 to transform A to transform A to a matrix with a nonzero entry in the $(1, 1)$ position. Then we can transform A to the matrix $D = [1\ 0\ \cdots\ 0]$ using a type 2 operation followed by multiple type 3 operations. This established the result when $m = 1$.

Now we assume that the result is true for matrices with $m \geq 1$ rows and consider a matrix A with $m + 1$ rows.

Proof

Proof Continued.

Since $A = (A_{ij}) \neq O$, some $A_{ij} \neq 0$. Using type 1 operations we can transform A to a matrix of the form

$$A' = \begin{pmatrix} * & R \\ C & B' \end{pmatrix}$$

where $*$ is a nonzero scalar, B' is a $m \times (n-1)$ matrix, R is a $1 \times (n-1)$ matrix, and C is a $m \times 1$ matrix. Then we an multiply the first row by a nonzero scalar so that the $(1,1)$-entry is a 1.

Proof

Proof Continued.

Then we can use type 3 operations to transform this matrix to one of the form

$$A'' = \begin{pmatrix} 1 & O \\ O & B'' \end{pmatrix}$$

and B'' is a $m \times (n-1)$ matrix. It is an exercise to check that $\operatorname{rank}(B'') = \operatorname{rank}(A'') - 1$. But then $\operatorname{rank}(A'') - 1 = \operatorname{rank}(A') - 1 = \operatorname{rank}(A) - 1 = r - 1$.

Now by induction, we can transform B'' via elementary row and column operations into

$$D' = \begin{pmatrix} I_{r-1} & O_4 \\ O_5 & O_6 \end{pmatrix}.$$

Then we can perform the corresponding operations to A'' to get

$$D = \begin{pmatrix} 1 & O & O \\ O & I_{r-1} & O_4 \\ O & O_5 & O_6 \end{pmatrix} = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}.$$

This is exactly the matrix desired. This proves item (2).

Proof

Proof Continued.

(1) Item (1) follows since $\operatorname{rank}(A) = \operatorname{rank}(D)$ and we clearly have $r \leq \min\{ m, n \}$ in D.

(3) Since D was obtained from A via a finite sequence of elementary row operations and elementary column operations, there are elementary matrices E_k and D_j such that

$$D = E_p E_{p-1} \cdots E_1 A D_1 D_2 \cdots D_q.$$

But each E_k and each D_j is invertible. Hence so are $B = E_p \cdots E_1$ and $C = D_1 \cdots D_q$ □

Break Time

Ok, that was a big result. We get to the pay off soon, but first a break and time for questions.

More Words

Definition

Let A be a $m \times n$ matrix with columns $v_1, \dots, v_n \in \mathbf{F}^m$ and rows $r_1, \dots, r_m \in \mathbf{F}^n$. Then we call $\mathrm{Col}(A) = \mathrm{Span}(\{ v_1, \dots, v_n \})$ the Column space of A and $\mathrm{Row}(A) = (\{ r_1, \dots, r_m \})$ the Row space of A.

Remark

We saw last time that $\mathrm{Col}(A) = \mathrm{R}(L_A)$. Keep in mind that $\mathrm{Col}(A) \subset \mathbf{R}^m$ and $\mathrm{Row}(A) \subset \mathbf{R}^n$. Then $\dim(Col(A))$ is the maximal number of linearly independent columns in A and $\dim(\mathrm{Row}(A))$ is the maximal number of linearly independent rows in A.

The Payoff

Corollary

Let A be a $m \times n$-matrix.

1. $\text{rank}(A^t) = \text{rank}(A)$.
2. $\dim(\text{Row}(A)) = \text{rank}(A) = \dim(\text{Col}(A))$.

Remark

Keep in mind that while $\dim(\text{Row}(A)) = \dim(\text{Col}(A))$ they potentially live in different vector spaces: one in $\mathbf{F}^n$ and one in $\mathbf{F}^m$.

Proof

Proof.

Using the Rank Theorem, we have $D = BAC$ where $D = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ where $r = \text{rank}(A)$ and both B and C are invertible. It is clear that D^t is of the same form but with different sized zero blocks. Hence $\text{rank}(D^t) = r$. But $D^t = C^t A^t B^t$. Since C^t and B^t are also invertible, $\text{rank}(A^t) = \text{rank}(C^t D^t B^t) = \text{rank}(D^t) = r$. This proves item (1).

(2) Note that $\text{Row}(A) = \text{Col}(A^t)$ and $\text{rank}(A^t) = \dim(\text{Col}(A^t)) = \dim(\text{R}(L_{A^t}))$. Since $\text{rank}(A^t) = \text{rank}(A)$ by item (1), the assertion is proved. □

Remark

I kind of like this result. For example, if I have a 3×15-matrix then of its 15 column vectors, at most 3 are linearly independent!

Invertible Matrices Again

Corollary

A $n \times n$ matrix is invertible if and only if it is the product of elementary matrices.

Proof.

Clearly the product of elementary matrices is invertible because the product of invertible matrices is invertible.

Suppose that A is invertible. Then $\operatorname{rank}(A) = n$. Therefore it follows from the rank theorem that we can transform A into $D = I_n$ via a sequence of elementary row and column operations. Hence there are elementary matrices E_k and D_j such that

$$\underbrace{E_p \cdots E_1}_{B} \, A \, \underbrace{D_1 \cdots D_q}_{C} = I_n.$$

Therefore

$$A = B^{-1} I_n C^{-1} = B^{-1} C^{-1} = E_1^{-1} \cdots E_p^{-1} D_q^{-1} \cdots D_1^{-1},$$

and we know that the inverse of an elementary matrix is elementary. □

Ranks of Products

Theorem

Suppose that $T : V \to W$ and $S : W \to Z$ are linear transformations between finite-dimensional vector spaces. Suppose also that A and B are matrices so that AB is defined.

1. $\operatorname{rank}(ST) \leq \operatorname{rank}(S)$.
2. $\operatorname{rank}(ST) \leq \operatorname{rank}(T)$.
3. $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$.
4. $\operatorname{rank}(AB) \leq \operatorname{rank}(B)$.

Proof.

(1) We have

$$\mathrm{R}(ST) = ST(V) = S(T(V)) \subset S(W) = \mathrm{R}(S).$$

Hence $\operatorname{rank}(ST) \leq \operatorname{rank}(S)$.

(3) Here we an use item (1) to see that

$$\operatorname{rank}(AB) = \operatorname{rank}(L_{AB}) \leq \operatorname{rank}(L_A) = \operatorname{rank}(A).$$

This proves item (3).

Proof

Proof.

(4) We can apply item (3) to see that

$$\operatorname{rank}(AB) = \operatorname{rank}((AB)^t) = \operatorname{rank}(B^t A^t) \leq \operatorname{rank}(B^t) = \operatorname{rank}(B).$$

This proves item (4).

(2) Let α, β, and γ be bases for V, W, and Z, respectively. Let $A' = [S]_\beta^\gamma$ and $B = [T]_\alpha^\beta$. Then $A'B' = [ST]_\alpha^\gamma$. Now can use item (4) to see that

$$\operatorname{rank}(ST) = \operatorname{rank}([ST]_\alpha^\gamma) = \operatorname{rank}(A'B') \leq \operatorname{rank}(B') = \operatorname{rank}(T).$$

This completes the proof. □

Break Time

Time for a well-deserved break and some questions.

Augmented Matrices

Definition

Let A be a $m \times n$-matrix and B a $m \times p$-matrix, then the augmented matrix $(A \mid B)$ is the $m \times (n + p)$-matrix $[A\ B]$ whose first n columns are the columns of A and the last p columns the columns of B.

Example

Let $A = \left(\begin{smallmatrix} a & b & c \\ d & e & f \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right)$. Then $(A \mid B) = \left(\begin{smallmatrix} a & b & c & x & y \\ d & e & f & z & w \end{smallmatrix}\right)$.
Sometimes we "keep the bar" to remind us what we added on:
$(A \mid B) = \left(\begin{smallmatrix} a & b & c \\ d & e & f \end{smallmatrix} \,\middle|\, \begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right)$

Finding Inverses

Theorem

Let A be a $n \times n$ matrix. Then A is invertible if and only if we can use elementary row operations to transform the augmented matrix $(A \mid I_n)$ to a matrix of the form $(I_n \mid B)$. In that case, $B = A^{-1}$. On the other hand, if we can use elementary row operations to transform $(A \mid I_n)$ into a matrix of the form $(C \mid D)$ with $\operatorname{rank}(C) < n$, then A is not invertible.

▸ Return1 ▸ Return3

Remark

Keep in mind that only elementary row operations are allowed here.

Examples First

Example

Before diving into the proof, let's consider the theorem's practical aspects with a few examples. Consider

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 1 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 3 & 0 & 4 \end{pmatrix}.$$

What can we say about the inverses of these matrices?

Solution

As we saw on the document camera, $A^{-1} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}.$

But B *is not invertible since* $\operatorname{rank}(B) = 2 < 3$.

Proof

Before giving the proof of the theorem, we need a little observation.

Lemma

If $(A \mid B)$ is an augmented $m \times (n + p)$-matrix and C is a $r \times m$-matrix, then $C(A \mid B) = (CA \mid CB)$.

Proof of the Lemma.

Suppose that the columns of A are $v_1, \dots, v_n$ and the columns of B are $w_1, \dots, w_p$. Then $(A \mid B) = [v_1 \ \cdots \ v_n \ w_1 \ \cdots \ w_p]$. Then we proved that
$C[[v_1 \ \cdots \ v_n \ w_1 \ \cdots \ w_p] = [Cv_1 \ \cdots \ Cv_n \ Cw_1 \ \cdots \ Cw_p]$. But the same observation implies this is equal to $(CA \mid CB)$. □

Proof

Proof of the Theorem.

Suppose that we can transform $(A \mid I_n)$ into $(I_n \mid B)$ via a finite sequence of elementary row operations. Then there are elementary matrices $E_1, \ldots, E_m$ such that

$$\underbrace{E_m E_{m-1} \cdots E_1}_{C}(A \mid I_n) = (I_n \mid B)$$

Then the Lemma implies $(CA \mid C) = (I_n \mid B)$. Therefore $B = C$ and $CA = I_n$. The latter implies that A is invertible with $A^{-1} = C = B$. This proves half of the first assertion and the second assertion Go.

Proof

Proof Continued.

Now suppose that A is invertible. Then we know that A is the product $D_1 D_2 \cdots D_m$ of elementary matrices D_k. But then

$$\begin{aligned} D_m^{-1} \cdots D_1^{-1}(A \mid I_n) &= D_m^{-1} \cdots D_1^{-1}(D_1 D_2 \cdots D_m \mid I_n) \\ &= D_m^{-1} \cdots D_2^{-1}(D_2 \cdots D_m | D_1^{-1}) \\ &\vdots \\ &(I_n \mid D_m^{-1} \cdots D_1^{-1}). \end{aligned}$$

This says precisely that we can transform $(A \mid I_n)$ into the form $(I_n \mid B)$ via elementary row operations. This proves the remaining half of the first assertion. Go

Proof

Proof Continued.

If we preform elementary row operations on $(A \mid B)$ via elementary row operations corresponding to the elementary matrices $E_1, \dots, E_m$, then we get an augmented matrix

$$(C \mid D) = (UA \mid UB)$$

where $U = E_n \cdots E_1$. Since U is invertible, $\text{rank}(A) = \text{rank}(UA) = \text{rank}(C)$. Hence if $\text{rank}(C) < n$, then $\text{rank}(A) < n$ and A is not invertible. $\square$

Enough

1. That is enough for today.