Math 24: Winter 2021
Lecture 14

Dana P. Williams

Dartmouth College

Monday, February 8, 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?

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.

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

We stated this theorem on Friday and did the following examples to show off. We considered

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

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 we discovered B is not invertible since our computation showed that $\operatorname{rank}(B) = 2 < 3$.

Proof

Before giving the proof of the theorem, we need a little observation that will prove useful for the proof and down the road!

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 $(A \mid B) = [v_1 \ \cdots \ v_n \ w_1 \ \cdots \ w_p]$. But we also know that matrix multiplication works "column by column": $C[v_1 \ \cdots \ v_n \ w_1 \ \cdots \ w_p] = [Cv_1 \ \cdots \ Cv_n \ Cw_1 \ \cdots \ Cw_p]$. 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 $E_1,\dots,E_m$ 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 $BA = CA = I_n$. The latter implies that A is invertible with $A^{-1} = B$. 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 $(A \mid I_n)$ via elementary row operations corresponding to the elementary matrices $E_1, \dots, E_m$, then we get an augmented matrix

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

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

Break Time

The previous theorem justifies our calculations and assertions from the end of the previous lecture.

Now let's take a break see if there are some questions before we move on §3.3 and the careful study of systems we promised in §1.4.

Systems

Recall that a system of m linear equations in n unknowns $x_1, \dots, x_n$ over a field $\mathbf{F}$ can be written as follows:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where the a_{ij} and b_k are scalars in $\mathbf{F}$. The $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

is called the coefficient matrix of the system.

Matrix Version

Then if we let $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$, we can write our system in matrix form:

$$Ax = b.$$

where $x \in \mathbf{F}^n$ is viewed as a $n \times 1$-matrix and $b \in \mathbf{F}^m$ is viewed as a $m \times 1$-matrix. A solution to our system is just a vector $s = (s_1, \dots, s_n)$ so that $x = s$ satisfies each equation simultaneously. The set K of all solutions is called the solution set of the system. If the solution set is nonempty, then the system is called consistent. If the solution set is empty, then the system is called inconsistent.

Example

Example

The system

$$\begin{aligned} x_1 + x_2 &= 3 \\ x_2 - 3x_2 &= -1 \\ 2x_1 - x_2 &= 3 \end{aligned}$$

has matrix form

$$\begin{pmatrix} 1 & 1 \\ 1 & -3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix}.$$

I leave it to you to check that this system it consistent and has the unique solution $(2, 1)$. What happens if we change $b = (3, -1, 3)$ to $(3, -1, 2)$? (Document Camera)

Homogeneous Equations

Definition

A system $Ax = b$ of m linear equations in n unknowns is called homogeneous if $b = 0$. (Here, $0 = 0_{\mathbf{F}^m}$.) Otherwise the systems is called nonhomogeneous.

Remark (The Trivial Solution)

Every homogeneous system $Ax = 0$ has at least one solution, 0, called the trivial solution. (Note that the trivial solution is $0_{\mathbf{F}^n}$! Then $A0_{\mathbf{F}^n} = 0_{\mathbf{F}^m}$.)

The Kernel

Theorem

Let $Ax = 0$ be a homogeneous system of m equations and n unknowns over a field $\mathbf{F}$. Let K be the set of all solutions to $Ax = 0$. Then $\mathrm{K} = \mathrm{N}(L_A)$. Hence K is a subspace of $\mathbf{F}^n$ with $\dim(\mathrm{K}) = n - \mathrm{rank}(A)$.

Proof.

We clearly have $\mathrm{K} = \{\, x \in \mathbf{F}^n : Ax = 0 \,\} = \mathrm{N}(L_A)$. Then we already know $\mathrm{N}(L_A)$ is a subspace. The rest follows from the Dimension Theorem. $\square$

Corollary (Fewer Equations Than Unknowns)

Suppose that $m < n$ and that A is a $m \times n$-matrix. Then the homogeneous system $Ax = 0$ has a nontrivial solution.

Proof.

We have $m \geq \operatorname{rank}(A) = \operatorname{rank}(L_A)$. Hence $\dim(\mathrm{K}) = \dim(\mathrm{N}(L_A)) = n - \operatorname{rank}(L_A) \geq n - m > 0$ by the Dimension Theorem. Therefore $K \neq \{0\}$ and there is a nontrivial solution. $\square$

Examples

Example

Consider the system

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

over $\mathbf{R}$. Then $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \end{pmatrix}$ is our coefficient matrix. Clearly $\operatorname{rank}(A) = 2$. Hence $\dim(\mathsf{K}) = 1$. Since $x = (1, 1, -1)$ is a solution, we know that $\mathsf{K} = \{\, t(1, 1, -1) : t \in \mathbf{R} \,\}$.

Return

Example

Example

Consider the one equation system $x_1 - x_2 + x_3 - x_4 = 0$. Here $A = (1 \ \ -1 \quad 1 \ \ -1)$ has rank 1 and $\dim(\mathsf{K}) = 3$. It is not hard to see that $\beta = \{\, (1,0,0,1), (-1,0,1,0), (1,1,0,0) \,\}$ is a set of linearly independent solutions. Thus β is a basis for K and

$$\begin{aligned}\mathsf{K} &= \operatorname{Span}(\beta) \\ &= \{\, t_1(1,0,0,1) + t_2(-1,0,1,0) + t_3(1,1,0,0) : t_1, t_2, t_3 \in \mathbf{R} \,\} \\ &= \{\, (t_1 - t_2 + t_3, t_3, t_2, t_1) : t_1, t_2, t_3 \in \mathbf{R} \,\}.\end{aligned}$$

Break Time

Time for a break and questions.

Nonhomogeneous Equations

Definition

If $Ax = b$ is a nonhomogeneous system of m equations in n unknown, then $Ax = 0$ is the corresponding homogeneous system.

Theorem

Suppose that $Ax = b$ is a consistent nonhomogeneous system with solution set K. Let K_H be the solution set to the corresponding homogeneous system $Ax = 0$. Then for any $s \in K$,

$$\mathsf{K} = \{s\} + \mathsf{K}_H = \{\, s + h : h \in \mathsf{K}_H \,\}.$$

Proof.

Fix $s \in \mathsf{K}$ as above. If $w \in K$, then $A(w - s) = Aw - As = b - b = 0$. Thus $w - s \in \mathsf{K}_H$ and $w = s + (w - s) \in \{s\} + \mathsf{K}_H$. Therefore $\mathsf{K} \subset \{s\} + \mathsf{K}_H$.

Proof

Proof Continued.

On the other hand, if $h \in K_H$, then $A(s+h) = As + Ah = b + 0 = b$ and $s + h \in \mathrm{K}$. This shows $\{s\} + \mathrm{K}_H \subset \mathrm{K}$. Therefore $\mathrm{K} = \{s\} + \mathrm{K}_H$ as claimed. □

Remark

The upshot here is that if we know the solutions K_H to the homogeneous system $Ax = 0$, then to completely solve the nonhomogeneous system $Ax = b$, we just need to find a "particular solution", say s_0, to $Ax = b$. Then the solution set K to the nonhomogeneous system is $\{s_0\} + \mathrm{K}_H$.

Example

Example

Consider the nonhomogeneous system

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 6 \\ 2x_1 - x_2 + x_3 &= 2. \end{aligned}$$

over $\mathbf{R}$. Then $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \end{pmatrix}$ is our coefficient matrix as in ▸ our example. Clearly $s = (1,1,1)$ is a solution. Thus the solution set is

$$\begin{aligned} \{s\} + \mathsf{K}_H &= \{\, (1,1,1) + t(1,1,-1) : t \in \mathbf{R} \,\} \\ &= \{\, (1+t, 1+t, 1-t) : t \in \mathbf{R} \,\}. \end{aligned}$$

Invertible Coefficient Matrix

Theorem

Let $Ax = b$ be a system of n equations in n unknowns. If A is invertible, then the system has a unique solution—namely, $A^{-1}b$. Conversely, if the system $Ax = b$ has a unique solution, then A is invertible.

Proof.

Suppose that A is invertible. Then $A(A^{-1}b) = b$ and $s = A^{-1}b$ is a solution to $Ax = b$. On the other hand, if $As = b$, then $s = A^{-1}(As) = A^{-1}b$. So $A^{-1}b$ is the unique solution to $Ax = b$.

Conversely, suppose that $Ax = b$ has the unique solution $s \in \mathbf{F}^n$. Then $\{s\} = \{s\} + \mathsf{K}_H$ where K_H is the set of solutions to $Ax = 0$. Then $\mathsf{K}_H = \{0\}$ and $\mathsf{N}(L_A) = \{0\}$. Then L_A is one-to-one and hence invertible. Since A is a square, this implies A is invertible

. $\square$

Example

Example

Consider the system

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 1 \\ x_1 + x_2 &= 2 \\ 2x_1 + x_2 + 2x_3 &= 3. \end{aligned}$$

Here $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 1 & 2 \end{pmatrix}$. We saw earlier that
$A^{-1} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}$. Thus the unique solution is
$$\begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -\frac{1}{2} \end{pmatrix}.$$

Consistency

Theorem

The system $Ax = b$ is consistent if and only if $\operatorname{rank}(A) = \operatorname{rank}(A \mid b)$.

Proof.

Let $A = [v_1 \ \cdots \ v_n]$. If $Ax = b$ is consistent, then $b \in \mathrm{R}(L_A) = \operatorname{Span}\bigl(\{\, v_1, \dots, v_n \,\}\bigr)$. Therefore

$$\operatorname{Span}\bigl(\{\, v_1, \dots, v_n \,\}\bigr) = \operatorname{Span}\bigl(\{\, v_1, \dots, v_n, b \,\}\bigr).$$

Then

$$\dim\bigl(\operatorname{Span}\bigl(\{\, v_1, \dots, v_n \,\}\bigr)\bigr) = \dim\bigl(\operatorname{Span}\bigl(\{\, v_1, \dots, v_n, b \,\}\bigr)\bigr)$$

and $\operatorname{rank}(A) = \operatorname{rank}(A \mid b)$.

Proof

Proof Continued.

Conversely, if $\operatorname{rank}(A) = \operatorname{rank}(A \mid b)$, then

$$\dim\bigl(\operatorname{Span}\bigl(\{\, v_1, \dots, v_n \,\}\bigr)\bigr) = \dim\bigl(\operatorname{Span}\bigl(\{\, v_1, \dots, v_n, b \,\}\bigr)\bigr).$$

Since $\bigl(\operatorname{Span}\bigl(\{\, v_1, \dots, v_n \,\}\bigr)$ is a subspace of $\operatorname{Span}\bigl(\{\, v_1, \dots, v_n, b \,\}\bigr)$, this implies

$$\operatorname{Span}\bigl(\{\, v_1, \dots, v_n \,\}\bigr) = \operatorname{Span}\bigl(\{\, v_1, \dots, v_n, b \,\}\bigr).$$

But then $b \in \operatorname{Span}\bigl(\{\, v_1, \dots, v_n \,\}\bigr) = \mathrm{R}(L_A)$ and $Ax = b$ is consistent. $\square$

Example

Example

Consider the system

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 - x_2 + x_3 &= 0 \\ x_1 \qquad + x_3 &= 0. \end{aligned}$$

We have $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ while $(A \mid b) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$.
We see immediately that $\operatorname{rank}(A) = 2$. (Consider $\operatorname{Col}(A)$.) With just a bit of thought, it is clear that $\operatorname{rank}((A \mid b)) = 3$. Hence the system is inconsistent.

Enough

1. That is enough for today.