Math 24: Winter 2021
Lecture 15

Dana P. Williams

Dartmouth College

Wednesday, February 10, 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

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 $s \in K$,*

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

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.

Theorem

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

Solving Systems of Equations

1. Until now we have treated actually solving a system of linear equations in a *ad hoc* basis letting you more of less figure it out on you own.
2. Section 3.4 is concerned with systematically solving systems of equations.
3. The goal is not just to solve the system, but do it in a way that allows us to describe the solution set with a minimal number of linearly independent vectors.
4. I only want to spend a lecture on this, so we may cut a corner or two by sketching some proofs and even just stating a result or two.

Equivalent Systems

Definitions

Two systems of linear equations in n unknowns are said to be equivalent if they have the same solution set.

Theorem

Let $Ax = b$ be a system of m equations in n unknowns. (This means A is a $m \times n$-matrix!) If C is an invertible $m \times m$ matrix, then $Ax = b$ is equivalent to $(CA)x = Cb$.

Proof.

Let K be the solution set for $Ax = b$ and K' the solution set for $(CA)x = Cb$. If $s \in \mathsf{K}$, then $As = b$. But then $(CA)s = C(As) = Cb$, so $s \in \mathsf{K}'$. On the other hand, if $s' \in \mathsf{K}'$, then $(CA)s' = Cb$. But then $As' = C^{-1}(CA)s' = C^{-1}(Cb) = b$ and $As' = b$. That is $s' \in \mathsf{K}$. Therefore $\mathsf{K} = \mathsf{K}'$ and the systems are equivalent. □

Elementary Row Operations

Corollary

Let $Ax = b$ be a system of m equations in n unknowns. If $(A' \mid b')$ is obtained from $(A \mid b)$ by a sequence of elementary row operations, then $A'x = b'$ is equivalent to $Ax = b$.

Proof.

If $(A' \mid b')$ is obtained from $(A \mid b)$ by a sequence of elementary row operations, then there are $m \times m$-elementary matrices E_k such that $(A' \mid b') = E_p \cdots E_1(A \mid b)$. Let $C = E_p \cdots E_1$. Then C is invertible and $(A' \mid b') = (CA \mid Cb)$. Thus $A'x = b'$ is the system $(CA)x = Cb$ and $A'x = b'$ is equivalent to $Ax = b$. □

An Example

Consider the system

$$\begin{aligned} x_1 + 2x_2 + x_3 - x_4 + 3x_5 &= 2 \\ x_1 + x_2 + x_3 \qquad - 3x_5 &= 3 \\ 3x_1 + 2x_2 + 3x_3 - 2x_4 \qquad &= 1. \end{aligned}$$

Then form the augmented matrix $(A \mid b)$ for $Ax = b$.

$$\begin{pmatrix} 1 & 2 & 1 & -1 & 3 & 2 \\ 1 & 1 & 1 & 0 & -3 & 3 \\ 3 & 2 & 3 & -2 & 0 & 1 \end{pmatrix}$$

Now we mimic what we were told to do in §1.4! We want a leading coefficient of 1 in each row with all zeros below it.

Step One

Then we get

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

Now we use the leading 1's to zero out the rest of their column to get $(A' \mid b')$:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & -4 & 1 \\ 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & -5 & 3 \end{pmatrix}.$$

Thus our new equivalent system is

$$\begin{aligned} x_1 + x_3 - 4x_5 &= 1 \\ x_2 + x_5 &= 2 \\ x_4 - 5x_5 &= 3. \end{aligned}$$

I suggest re-writing this as

$$\begin{aligned} x_1 &= 1 - x_3 + 4x_5 \\ x_2 &= 2 - x_5 \\ x_4 &= 3 + 5x_5. \end{aligned}$$

Now it is clear we are free to pick x_3 and x_5 as we please. I like to say that x_3 and x_5 are "free variables". Hence if $x_3 = t$ and $x_5 = s$, then we have

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 - t + 4s \\ 2 - s \\ t \\ 3 + 5s \\ s \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \\ 0 \end{pmatrix}}_{s_0} + t \underbrace{\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{u_1} + s \underbrace{\begin{pmatrix} 4 \\ -1 \\ 0 \\ 5 \\ 1 \end{pmatrix}}_{u_2}.$$

Therefore $\mathsf{K} = \{ s_0 \} + \mathsf{K}_H$ where $\mathsf{K}_H = \operatorname{Span}\big(\{ u_1, u_2 \}\big)$ and K_H is the solution set for the associated homogeneous equation $Ax = 0$.

Echelon and Reduced Echelon Form

Definition

A $m \times n$ matrix is said to be in row echelon form if

1. Any row containing a nonzero entry lies above any row of all zeros.
2. The first nonzero entry in any row—called a leading entry—is a 1 and all the entries below it in its column are zero.
3. The leading entry in any row occurs the right of any leading entry in any row above it.

If in addition each leading entry is the only nonzero entry in its column, then we say the matrix is in reduced row echelon form.

Break Time

Time for a break and questions.

Gaussian Elimination

Remark (Gaussian Elimination)

It is a theorem that every matrix can be put into reduced row echelon form using elementary row operations just as we did in our example. There are more details in the text, but we basically proceed as follows.

1. Forward Pass: We move a nonzero entry to the top of the first nonzero column, scale that entry to a 1 and zero out the entries below. Then we continue row by row until the matrix is in row echelon form.
2. Backward Pass: Moving left from the right-most leading entry we zero out the remaining entries in the column above each leading entry. The matrix is now in reduced row echelon form.
3. This process is called Gaussian Elimination.

Remark

It turns out that the reduced row echelon form of a matrix is unique. However, a matrix can have many echelon forms.

The General Solution

1. Starting with a system $Ax = b$, we form the augmented matrix $(A \mid b)$ and transform it to reduced row echelon form $(A' \mid b')$ via Gaussian elimination.
2. Since $(A' \mid b')$ is in reduced row echelon form, the non-zero rows are clearly linearly independent and the $\operatorname{rank}(A' \mid b')$ is the number r of nonzero rows.
3. Furthermore $\operatorname{rank}(A') = \operatorname{rank}(A' \mid b')$ if and only if there is no leading entry in the last column in which case $\operatorname{rank}(A) = \operatorname{rank}(A') = r$.
4. When solving the new system $A'x = b'$, we can ignore the zero rows. Furthermore, the variables corresponding to the r leading entries can be expressed in terms of the remaining $n - r$ variables (free variables) and we get a general solution of the form

$$s = s_0 + t_1 u_1 + \cdots + t_{n-r} u_{n-r}$$

where $t_1, \ldots, t_{n-r} \in \mathbf{R}$.

Summary

Theorem

Suppose that $Ax = b$ is a system of r nonzero equations such that $(A \mid b)$ is in reduced row echelon form. Suppose the system is consistent so that $\operatorname{rank}(A \mid b) = \operatorname{rank}(A)$. *Then*

1. $\operatorname{rank}(A) = r$, *and*
2. *If $s = s_0 + t_1 u_1 + \cdots + t_{n-r} u_{n-r}$ is the general solution as above, then $\{ u_1, \dots, u_{n-r} \}$ is a basis for the solution set K_H for the corresponding homogeneous system $Ax = 0$.*

Proof.

The previous discussion shows that $\operatorname{rank}(A) = r$ in this case. Let K be the set of solutions to $Ax = b$ and K_H the solutions to the homogeneous system $Ax = 0$. Clearly, $s_0 \in \mathsf{K}$. But then $\mathsf{K}_H = \{-s_0\} + \mathsf{K} = \operatorname{Span}\big(\{ u_1, \dots, u_{n-r} \}\big)$. Since $\operatorname{rank}(A) = r$, $\dim(K_H) = n - r$. Since $\{ u_1, \dots, u_{n-r} \}$ is a spanning set with $n - r$ elements, they must be linearly independent and therefore a basis. □

Break Time

Time for a break and some questions.

The Glories of Reduced Row Echelon Form

Theorem

Let $A = [a_1 \ \cdots \ a_n]$ be a $m \times n$ nonzero matrix with columns a_k and let $B = [b_1 \ \cdots \ b_n]$ be the the reduced row echelon form of A with columns b_k. Let $r \geq 1$ be the number of nonzero rows in B and $\{ e_1, \dots, e_m \}$ the standard ordered basis for $\mathbf{F}^m$.

1. $\operatorname{rank}(A) = r$.
2. *For $1 \leq j \leq r$, there is a column b_{j_i} of B such that $b_{j_i} = e_i$.*
3. *Then $\{ a_{j_1}, \dots, a_{j_r} \}$ is a basis for $\operatorname{Col}(A) = \operatorname{R}(L_A)$.*
4. *The nonzero rows of B form a basis for $\operatorname{Row}(A)$ in $\mathbf{F}^n$.*

First An Example

Example

Before diving into the proof, let's try to see what the theorem says and how to use it. Let $A = [a_1\ a_2\ a_3\ a_4\ a_5] = \begin{pmatrix} 2 & 4 & 6 & 2 & 4 \\ 1 & 2 & 3 & 1 & 1 \\ 2 & 4 & 8 & 0 & 0 \\ 3 & 6 & 7 & 5 & 9 \end{pmatrix}$. You can check that the reduced row echelon form of A is

$B = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Now the theorem implies that $\operatorname{rank}(A) = 3$ and that $\{ a_1, a_3, a_5 \}$ is a basis for $\operatorname{Col}(A) = \mathrm{R}(L_A)$ in $\mathbf{R}^4$. Note that the first, third, and fifth columns of B are not necessarily in $\operatorname{Col}(A)$ let alone a basis. In this case, $(1, 0, 0, 0)$ is not in $\mathrm{R}(L_A)$! (You should check this.) Also we see that $\{ (1, 2, 0, 4, 0), (0, 0, 1, -1, 0), (0, 0, 0, 0, 1) \}$ is a basis for $\operatorname{Row}(A)$ in $\mathbf{R}^5$.

Now the Proof

Proof.

Let $A = [a_1 \ \cdots \ a_n]$ be a $m \times n$ matrix with columns a_k and let $B = [b_1 \ \cdots \ b_n]$ be the the reduced row echelon form of A with columns b_k. Because B is in reduced row echelon form, its nonzero rows must be linearly independent. Since $\operatorname{rank}(A) = \operatorname{rank}(B)$, $r = \operatorname{rank}(A)$ is the number of nonzero rows in B. This proves item (1).

Since $r \geq 1$, the the standard basis vectors $e_1, \ldots, e_r$ in $\mathbf{F}^m$ are among the columns of B. Let $b_{j_i} = e_i$. This proves item (2).

Since $\operatorname{rank}(A) = r = \dim\bigl(\operatorname{Col}(A)\bigr)$, to prove item (3), it suffices to see that $\{ a_{j_1}, \ldots, a_{j_r} \}$ is linearly independent. So we suppose

$$d_1 a_{j_1} + \cdots + d_r a_{j_r} = 0.$$

We need to see that $d_k = 0$ for all k.

Proof

Proof Continued.

Since we obtained B from A via a sequence or elementary row operations, there is an invertible matrix M such that $MA = B$. Since we know that matrix multiplication works "column by column", we have $Ma_{j_i} = e_i$. Therefore

$$\begin{aligned} 0 = M0 &= M(d_1 a_{j_1} + \cdots + d_r a_{j_r}) \\ &= d_1 M a_{j_1} + \cdots d_r M a_{j_r} = d_1 e_1 + \cdots + d_r e_r. \end{aligned}$$

Since $\{ e_1, \ldots, e_r \}$ are linearly independent in $\mathbf{F}^m$, we must have all the $d_k = 0$ as required.

(4) If we obtain A' from A via an elementary row operation, then it is not so hard to check that $\text{Row}(A) = \text{Row}(A')$. This very easy for a type 1 or type 2 operation and just a bit more work for a type 3. I leave it to you to check this. Since B is obtained from A via a finite sequence of elementary row operations, we have $\text{Row}(A) = \text{Row}(B)$. But the nonzero rows of B clearly span $\text{Row}(B)$ and are linearly independent. This proves item (4). □

Example

Example

Clearly $\beta = \{ (1,-1,1,-1), (2,1,-1,1) \}$ is linearly independent in $\mathbf{R}^4$. In theory, we should be able to extend β to a basis of $\mathbf{R}^4$. How?

Solution

Let $A = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$. *Clearly,* $\operatorname{rank}(A) = 4$. *If we put* A *into reduced row echelon form, then we get*
$U = \begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 & 0 & -\frac{2}{3} \\ 0 & 1 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$. *Therefore the first, second, fourth, and fifth columns of* A *for a basis for* $\mathbf{R}^4$. *Thus one such basis is*
$\gamma = \{ (1,-1,1,-1), (2,1,-1,1), (0,1,0,0), (0,0,1,0) \}$.

Enough

1. That is enough for today.