Solutions of Linear Systems
Reading: Lay 1.5

September 6, 2013

The amount of extra theory in this section is small. What theory we need to introduce can mostly be explained by examples. So I am going to run the lecture on the following pattern: I do one example, then you try a similar one. We will do this for three different examples. Along the way, I will explain what theory we need.

The first section is material that was technically covered last class (from Lay 1.4).

1 From last time

Recall that last time, we discussed matrix equations. The big question was whether a set of vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ spanned $\mathbb{R}^m$:

- Does $\text{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} = \mathbb{R}^m$?

An equivalent way to state this problem was in terms of the matrix A whose first column is $\mathbf{a}_1$, second column is $\mathbf{a}_2$, etc:

$$A = \begin{bmatrix} \mathbf{a}_1 \, \mathbf{a}_2 \, \ldots \mathbf{a}_n \end{bmatrix}.$$

With this definition of A, we can restate the above "big question" in this form:

- Does the equation $A\mathbf{x} = \mathbf{b}$ have a solution for every $\mathbf{b} \in \mathbb{R}^m$?

Example 1.1. *(I will work this example in class.)* Does the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ with vectors

$$\mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

span $\mathbb{R}^3$?

Write the matrix whose columns are $\mathbf{u}, \mathbf{v}, \mathbf{w}$:

$$\begin{bmatrix} 3 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

How can we tell from this matrix whether the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ spans $\mathbb{R}^3$? The answer: they span $\mathbb{R}^3$ if and only if the matrix above has a pivot position in each row. So we start computing the RREF:

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 2/3 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

You can actually see at this point that there is a pivot in each row (ask yourself why), so we will stop. So the answer is that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ spans $\mathbb{R}^3$.

Example 1.2. *(I will have you work this during class.)* Let

$$\mathbf{t} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}.$$

Does $\{\mathbf{t}, \mathbf{u}, \mathbf{v}, \mathbf{w}\}$ span $\mathbb{R}^3$? We write the matrix whose columns are these vectors and start row reduction:

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

At this point, it is already clear there will be no pivot in the bottom row! So the answer is no, the vectors do not span.

2 Back to one vector

We just talked about spanning. Now to the "easy" problem of whether $A\mathbf{x} = \mathbf{b}$ has a solution for one particular $\mathbf{b}$. In fact, let's make it easier on ourselves and take $\mathbf{b} = \mathbf{0}$.

- Given a matrix A does $A\mathbf{x} = \mathbf{0}$ have a solution? Or more than one solution?

It is very easy to see that no matter what A is, there is always at least one $\mathbf{x}$ that solves this equation. Specifically, the zero vector is always a solution: $A\mathbf{0} = \mathbf{0}$ for any matrix A. We call this the **trivial solution**.

So our question should actually be

- Given a matrix A does $A\mathbf{x} = \mathbf{0}$ have a solution which is not the trivial solution?

We call solutions which are not $\mathbf{0}$ "**nontrivial solutions**". We know how to answer this question already by using linear systems. We illustrate via examples.

Example 2.1. Given

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & 1 & 1 \end{bmatrix},$$

does the equation $A\mathbf{x} = 0$ have a nontrivial solution?

We note that if we label the entries of $\mathbf{x}$ by the names x_1, x_2, x_3, this is equivalent to asking whether the linear system

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

has a solution other than the trivial solution $x_1 = x_2 = x_3 = 0$. So we write the augmented matrix and row reduce:

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 1/2 & 3 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 1/2 & 3 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

This linear system has no free variables, so there is only one solution: $x_1 = x_2 = x_3 = 0$. So there is no nontrivial solution to our matrix equation.

Example 2.2. Given

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix},$$

does the equation $A\mathbf{x} = 0$ have a nontrivial solution? We again write the augmented matrix and compute the RREF:

$$\begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 3 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 3 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

From here it is easy to see that the variable x_3 will be free. If you work out the rest of the RREF calculation, you get

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

For any choice of x_3, the choices $x_1 = -x_3$, $x_2 = -x_3$ give $A\mathbf{x} = \mathbf{0}$. So there is a nontrivial solution.

What does the solution we just got mean? It means that the solution set of $A\mathbf{x} = \mathbf{0}$ is the collection of all vectors of the form

$$\mathbf{x} = \begin{bmatrix} -x_3 \\ -x_3 \\ x_3 \end{bmatrix}$$

or equivalently, all vectors of the form

$$\mathbf{x} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

for real x_3. This means that the solution set is equal to

$$\operatorname{span} \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

This is a general fact:

Theorem 2.3. *Let A be a matrix. The solution set of $A\mathbf{x} = \mathbf{0}$ always has the form*

$$\operatorname{span} \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$

for some vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Let's show one more quick example, which is already in RREF for us.

Example 2.4. What is the solution set of $A\mathbf{x} = \mathbf{0}$, where

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}?$$

There are two free variables and so for any choice of x_2 and x_3, taking

$$x_1 = -x_2 - 2x_3$$

gives a solution. This means that the solution set is all vectors of the form

$$\begin{aligned} \begin{bmatrix} -x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} -x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} -2x_3 \\ 0 \\ x_3 \end{bmatrix} \\ &= x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}. \end{aligned} \tag{1}$$

This is just another way of writing

$$\operatorname{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\},$$

so the theorem above makes sense.

3 Parametric form

Any vector equation of the form

$$\mathbf{v} = c\mathbf{u} + d\mathbf{w}$$

where c and d are allowed to vary over real numbers is called a **parametric vector equation** (in principle, we could have more vectors than just two). So any solution like (1) is called a solution in **parametric vector form**. This is just a fancy name.

4 Nonhomogeneous systems

We've dealt with equations of the form

$$A\mathbf{x} = \mathbf{b}$$

for $\mathbf{b} = \mathbf{0}$, called a homogeneous system. We're now going to talk about the case where $\mathbf{b} \neq \mathbf{0}$, which we call a **nonhomogenous system**.

First, a question: say that $\mathbf{y}$ solves the homogeneous problem $A\mathbf{y} = \mathbf{0}$ and $\mathbf{x}$ solves the nonhomegeneous problem $A\mathbf{x} = \mathbf{b}$. Then what does $A(\mathbf{x} + \mathbf{y})$ equal?

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

On the other hand, say $\mathbf{z}$, like $\mathbf{x}$, satisfies $A\mathbf{z} = \mathbf{b}$ Then $A(\mathbf{x} - \mathbf{z}) = A\mathbf{x} - A\mathbf{z} = 0$. So $\mathbf{x} - \mathbf{z}$ solves the homogeneous problem. We summarize:

- If $\mathbf{x}$ solves the non homogeneous problem $A\mathbf{x} = \mathbf{b}$ and $\mathbf{y}$ solves the homogeneous problem, then $A(\mathbf{x} + c\mathbf{y}) = \mathbf{b}$ for every scalar c.

- If $\mathbf{x}$ and $\mathbf{z}$ both solve the non homogeneous problem (for the same vector $\mathbf{b}$), then
$$A(\mathbf{x} - \mathbf{z}) = \mathbf{0}.$$

By the above, the following fact is true: given the non homogeneous problem $A\mathbf{x} = \mathbf{b}$, we can write the general solution (i.e., describe the *entire* solution set) in the form

$$\{\mathbf{z} + c_1\mathbf{y}_1 + \ldots + c_2\mathbf{y}_k\}, \tag{2}$$

Where $\mathbf{z}$ is one particular solution ($A\mathbf{z} = \mathbf{b}$) and where each $\mathbf{y}_i$ satisfies the homogeneous problem. This is the reason why we introduced the "parametric form" notation above.

In class, if there is time, I will stop here and have you try Ex. 3 from Lay 1.5. If you are reading these notes, stop here and try to do Ex 3 from Lay without reading his solution. Note that you get a general solution of the form (2) as I claimed above.

5 Summary

The procedure here describes what we have learned in this section about writing the solution set of a consistent system $A\mathbf{x} = \mathbf{b}$:

1. Write the augmented matrix and find the RREF;
2. Solve for the basic variables (variables coming from the pivot columns) in terms of the free variables;
3. Write the general solution as a linear combination of vectors with the free variables as parameters.