Professor Jennifer Balakrishnan, *jbala@bu.edu*

What is on today

1 The matrix equation $A\mathbf{x} = \mathbf{b}$

Lay–Lay–McDonald §1.4 pp. 37 – 43

The matrix equation $A\mathbf{x} = \mathbf{b}$ leads us to consider variations on the existence question we looked at earlier. We start with the following useful fact:

The equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of A.

Earlier we looked at the following existence question: "Is $\mathbf{b}$ in Span $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$?" An equivalent reformulation is "Is $A\mathbf{x} = \mathbf{b}$ consistent?" A more difficult existence problem is to determine whether the equation $A\mathbf{x} = \mathbf{b}$ is consistent for *all* possible $\mathbf{b}$, as below:

Example 1. *Let* $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$ *and* $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. *Is the equation* $A\mathbf{x} = \mathbf{b}$ *consistent for all possible* b_1, b_2, b_3*?*

Example 2. *The reduced matrix in the above example gives a description of all* $\mathbf{b}$ *for which the equation* $A\mathbf{x} = \mathbf{b}$ *is consistent. What is the condition, and what does this tell us?*

This leads us to the following theorem:

Theorem 3. *Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true or they are all false:*

1. *For each $\mathbf{b} \in \mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.*
2. *Each $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns of A.*
3. *The columns of A span $\mathbb{R}^m$.*
4. *A has a pivot position in every row.*

> Note that Theorem 3 is about a *coefficient* matrix, not an *augmented* matrix. If an augmented matrix $[A \quad \mathbf{b}]$ has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ may or may not be consistent.

Next we practice computing $A\mathbf{x}$ for various matrices A and vectors $\mathbf{x}$.

Example 4. *Compute the following:*

1. $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix}$
2. $\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix}$
3. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

We summarize the procedure we carried out above in the following rule:

> **Row-vector rule for computing $A\mathbf{x}$**
> If the product $A\mathbf{x}$ is defined, then the ith entry in $A\mathbf{x}$ is the sum of the products of corresponding entries from row i of A and from the vector $\mathbf{x}$.

We will use the following theorem throughout the course:

Theorem 5. *If A is an $m \times n$ matrix, $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$ and c is a scalar, then*

1. *$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;*
2. *$A(c\mathbf{u}) = c(A\mathbf{u})$.*

Proof. For simplicity, we prove the case of $n = 3$ (the general case follows very similarly, but just involves more notation). Let $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$ and let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$. For $i = 1, 2, 3$, let u_i, v_i be the ith entries in $\mathbf{u}, \mathbf{v}$, respectively. To prove the first statement, compute $A(\mathbf{u}+\mathbf{v})$ as a linear combination of the columns of A using the entries in $\mathbf{u}+\mathbf{v}$ as weights:

$$\begin{aligned} A(\mathbf{u}+\mathbf{v}) &= [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3] \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix} \\ &= (u_1 + v_1)\mathbf{a}_1 + (u_2 + v_2)\mathbf{a}_2 + (u_3 + v_3)\mathbf{a}_3 \\ &= (u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) + (v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3) \\ &= A\mathbf{u} + A\mathbf{v}. \end{aligned}$$

To prove the second statement, compute $A(c\mathbf{u})$ as a linear combination of the columns of A using the entries in $c\mathbf{u}$ as weights:

$$\begin{aligned} A(c\mathbf{u}) &= [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3] \begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} \\ &= (cu_1)\mathbf{a}_1 + (cu_2)\mathbf{a}_2 + (cu_3)\mathbf{a}_3 \\ &= c(u_1\mathbf{a}_1) + c(u_2\mathbf{a}_2) + c(u_3\mathbf{a}_3) \\ &= c(u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) \\ &= c(A\mathbf{u}). \end{aligned}$$

□

2 Solution sets of linear systems

Lay–Lay–McDonald §1.5 pp. 43 – 49

Solution sets of linear systems are important in the study of linear algebra, and we'll see them in various contexts throughout this course. Today we use vector notation to give geometric descriptions of these solution sets.

A system of linear equations is said to be *homogeneous* if it can be written in the form

$$A\mathbf{x} = \mathbf{0}$$

where A is an $m \times n$ matrix and $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$. Such a system *always* has at least one solution, namely $\mathbf{x} = \mathbf{0}$, the zero vector in $\mathbb{R}^n$. This solution is called the *trivial solution*. For a given equation $A\mathbf{x} = \mathbf{0}$, the important question is whether there exists a *nontrivial solution* to the system: that is, a nonzero vector $\mathbf{x}$ that satisfies the system.

Example 6. *Determine if the following homogeneous system has a nontrivial solution:*

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 0 \\ -3x_1 - 2x_2 + 4x_3 &= 0 \\ 6x_1 + x_2 - 8x_3 &= 0. \end{aligned}$$

Then describe the solution set.

What we just saw in the previous example is a consequence of the Existence and Uniqueness Theorem we saw in §1.2. That is, the Existence and Uniqueness Theorem leads us to the following fact:

The homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if and only if the equation has at least one free variable.

Example 7. *Describe all solutions of* $x_1 - 3x_2 - 2x_3 = 0$.

Examples 6 and 7 illustrate that the solution set of a homogenous equation $A\mathbf{x} = \mathbf{0}$ can always be expressed explicitly as Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ for suitable vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$. If the only solution is the zero vector, then the solution set is Span $\{\mathbf{0}\}$. If the equation $A\mathbf{x} = \mathbf{0}$ has only one free variable, the solution set is a line through the origin. A plane through the origin provides a good mental image for the solution set of $A\mathbf{x} = \mathbf{0}$ when there are two or more three variables.

The equation in Example 7 is an implicit description of the plane. The solution to the "system" (i.e., solving the equation) amounts to giving an explicit description as the set spanned by two vectors $\mathbf{u}$ and $\mathbf{v}$. The resulting solution is called a parametric vector equation of the plane. Sometimes such an equation is written as $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ (where $s, t \in \mathbb{R}$) to emphasize that the parameters vary over all real numbers. Whenever a solution is described explicitly with vectors as in Examples 6 and 7, we say that the solution is in *parametric vector form*.

When a nonhomogeneous linear system has many solutions, the general solution can be written in parametric vector form as one vector plus an arbitrary linear combination of vectors that satisfy the corresponding homogeneous system.

Example 8. *Describe all solutions of $A\mathbf{x} = \mathbf{b}$ where* $A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}$.

To describe the solution set of $A\mathbf{x} = \mathbf{b}$ geometrically, we can think of vector addition as a translation. Given $\mathbf{v}$ and $\mathbf{p}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$, the effect of adding $\mathbf{p}$ to $\mathbf{v}$ is to move $\mathbf{v}$ in a direction parallel to the line through $\mathbf{p}$ and $\mathbf{0}$. We say that $\mathbf{v}$ is translated by $\mathbf{p}$ to $\mathbf{v} + \mathbf{p}$. If each point on a line L in $\mathbb{R}^2$ or $\mathbb{R}^3$ is translated by a vector $\mathbf{p}$, the result is a line parallel to L.

Suppose L is the line through $\mathbf{0}$ and $\mathbf{v}$ described the equation $\mathbf{x} = t\mathbf{v}$ for $t \in \mathbb{R}$. Adding $\mathbf{p}$ to each point on L produces the translated line $\mathbf{x} = \mathbf{p} + t\mathbf{v}$. We call this the equation of the line through $\mathbf{p}$ parallel to $\mathbf{v}$. Thus the solution set of $A\mathbf{x} = \mathbf{b}$ is a line through $\mathbf{p}$ parallel to the solution set of $A\mathbf{x} = \mathbf{0}$. Below we state the more general result:

Theorem 9. *Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some given $\mathbf{b}$, and let $\mathbf{p}$ be a solution. Then the solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, where $\mathbf{v}_h$ is any solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.*

Put another way, what this theorem says is that if $A\mathbf{x} = \mathbf{b}$ has a solution, then the solution set is obtained by translating the solution set of $A\mathbf{x} = \mathbf{0}$, using any particular solution $\mathbf{p}$ of $A\mathbf{x} = \mathbf{b}$ for the translation.