L4Axb

February 7, 2015

1 $A\mathbf{x} = \mathbf{b}$

1.1 Multiplying a Matrix by a Vector

Last lecture we studied the vector equation:

$$x_1\mathbf{a_1} + x_2\mathbf{a_2} + ... + x_n\mathbf{a_n} = \mathbf{b}.$$

We also talked about how one can form a matrix from a set of vectors:

$$A = [\mathbf{a_1} \ \mathbf{a_2} \ ... \ \mathbf{a_n}].$$

In particular, if we have n vectors $\mathbf{a_1}, \mathbf{a_2}, ..., \mathbf{a_n} \in \mathbb{R}^m$, then A will be an $m \times n$ matrix:

We're going to express a vector equation in a new way. We are going to say that

$$x_1\mathbf{a_1} + x_2\mathbf{a_2} + ... + x_n\mathbf{a_n}$$

is the same as:

$$A\mathbf{x}$$

What we have done is to define **Matrix-vector multiplication**. That is, $A\mathbf{x}$ denotes multiplying a matrix and a vector, and we define it to be the same as a linear combination of the $\{\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n}\}$ with weights $x_1, x_2, \ldots, x_n$.

Here is the definition:

If A is an $m \times n$ matrix, with columns $\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n}$, and if $\mathbf{x} \in \mathbb{R}^n$, then the **product of A and $\mathbf{x}$,** denoted $A\mathbf{x}$, is **the linear combination of the columns of A using the corresponding entries in x as weights;** that is,

$$A\mathbf{x} = [\mathbf{a_1} \ \mathbf{a_2} \ \ldots \ \mathbf{a_n}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + \ldots + x_n\mathbf{a_n}.$$

Notice that if A is an $m \times n$ matrix, then $\mathbf{x}$ must be in $\mathbb{R}^n$ in order for this operation to be defined. That is, the number of **columns** of A must match the number of **rows** of $\mathbf{x}$.

If the number of columns of A does not match the number of rows of $\mathbf{x}$, then $A\mathbf{x}$ has no meaning.

1.2 Example.

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} =$$

$$4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -1 \\ 3 \end{bmatrix} =$$

$$\begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} =$$

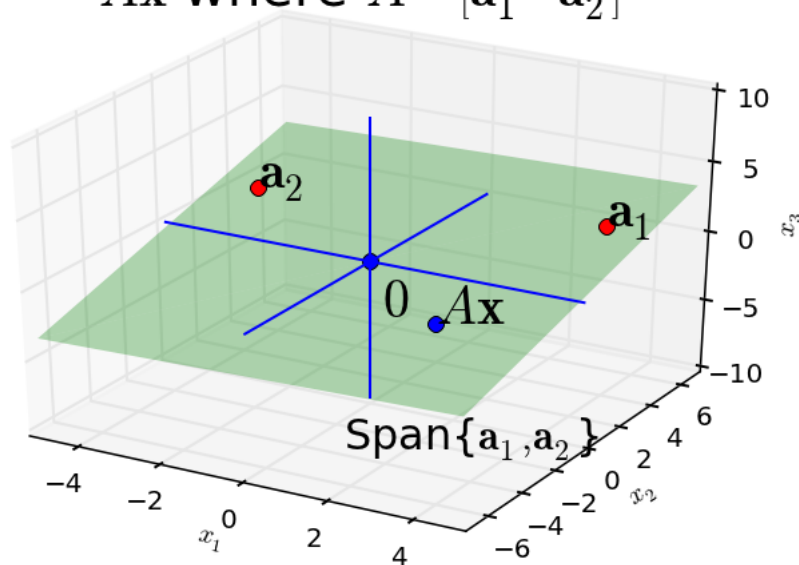$$\begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

1.3 Example.

For $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3} \in \mathbb{R}^m$, write the linear combination $3\mathbf{v_1} + 5\mathbf{v_2} + 7\mathbf{v_3}$ as a matrix times a vector.

$$3\mathbf{v_1} + 5\mathbf{v_2} + 7\mathbf{v_3} = [\mathbf{v_1} \ \mathbf{v_2} \ \mathbf{v_3}] \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} = A\mathbf{x}.$$

Notice that if $A = [\mathbf{a_1} \ \mathbf{a_2} \ \ldots \ \mathbf{a_n}]$ then $A\mathbf{x}$ is some point that lies in $\text{Span}\{\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n}\}$.

$$A\mathbf{x} \text{ where } A = [\mathbf{a}_1 \quad \mathbf{a}_2]$$

1.4 Question Time! Q4.1

1.5 The Matrix Equation

Now let's write an equation. Let's start with a linear system:

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

We saw last lecture that this linear system is equivalent to a vector equation, namely

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

Now, today we have learned that this vector equation is also equivalent to this **matrix equation**:

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}.$$

which has the form $A\mathbf{x} = \mathbf{b}$.

Notice how simple it is to go from the linear system to the matrix equation: we put the coefficients of the linear system into A, and the right hand side values into $\mathbf{b}$, and we can then write the matrix equation $A\mathbf{x} = \mathbf{b}$ immediately.

OK, let's write this out formally.

Theorem. If A is an $m \times n$ matrix, with columns $\mathbf{a_1}, \mathbf{a_2}, \dots, \mathbf{a_n}$, and if $\mathbf{b}$ is in $\mathbb{R}^m$, the matrix equation

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

has the same solution set as the vector equation

$$x_1\mathbf{a_1} + x_2\mathbf{a_2} + \dots + x_n\mathbf{a_n} = \mathbf{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

3

$$[\mathbf{a_1}\ \mathbf{a_2}\ ...\ \mathbf{a_n}\ \mathbf{b}].$$

This is an absolutely key result, because it gives us power to solve problems from three different viewpoints. If we are looking for a solution set $x_1, x_2, \ldots, x_n$, there are three different questions we can ask:

1. What is the solution set of the linear system $[\mathbf{a_1}\ \mathbf{a_2}\ ...\ \mathbf{a_n}\ \mathbf{b}]$?
2. What is the solution set of the vector equation $x_1\mathbf{a_1} + x_2\mathbf{a_2} + ... + x_n\mathbf{a_n} = \mathbf{b}$?
3. What is the solution set of the matrix equation $A\mathbf{x} = \mathbf{b}$?

We now know that all these questions have the same answer. And all of them can be solved by row reducing the augmented matrix $[\mathbf{a_1}\ \mathbf{a_2}\ ...\ \mathbf{a_n}\ \mathbf{b}]$.

In particular, one of the most common questions you will encounter is, given a matrix A and a vector $\mathbf{b}$, is there an $\mathbf{x}$ that makes $A\mathbf{x} = \mathbf{b}$ true?

You will encounter this form of question quite often in fields like computer graphics, data mining, algorithms, or quantum computing.

1.6 Question Time! Q4.2

1.7 Existence of Solutions

We can see that $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of A – that is, if $\mathbf{b}$ lies in $\mathrm{Span}\{\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n}\}$. We can adopt the term "consistent" here too – i.e., is $A\mathbf{x} = \mathbf{b}$ consistent?

Now we will ask a more general question, which is specifically about the matrix A. That is:

Does $A\mathbf{x} = \mathbf{b}$ have a solution for *any* $\mathbf{b}$?

Example.

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

We know how to test for consistency: form the augmented matrix and row reduce it.

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 7 & 5 & b_3 + 3b_1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{bmatrix}.$$

The third entry in the column 4, simplified, is $b_1 - \frac{1}{2}b_2 + b_3$. This expression can be nonzero.

For example, take $b_1 = 1, b_2 = 1, b_3 = 1$; then the lower right entry is 1.5. This gives a row of the form $0 = k$ which indicates an inconsistent system.

So the answer is "no, $A\mathbf{x} = \mathbf{b}$ is **not** consistent for all possible b_1, b_2, b_3."

Here is the echelon form of A (note we are talking only about A here, not the augmented matrix):

$$\begin{bmatrix} 1 & 3 & 4 \\ 0 & 14 & 10 \\ 0 & 0 & 0 \end{bmatrix}.$$

The equation $A\mathbf{x} = \mathbf{b}$ fails to be consistent for all $\mathbf{b}$ because the echelon form of A has a row of zeros.

If A had a pivot in all three rows, we would not care about the calculations in the augmented column because no matter what is in the augmented column, an echelon form of the augmented matrix could not have a row indicating $0 = k$.

Question: OK, so $A\mathbf{x} = \mathbf{b}$ is not always consistent. When **is** $A\mathbf{x} = \mathbf{b}$ consistent?

Answer: When $b_1 - \frac{1}{2}b_2 + b_3 = 0$. This makes the last row of the augmented matrix all zeros. This is a single equation in three unknowns, so it defines a plane through the origin in $\mathbb{R}^3$.
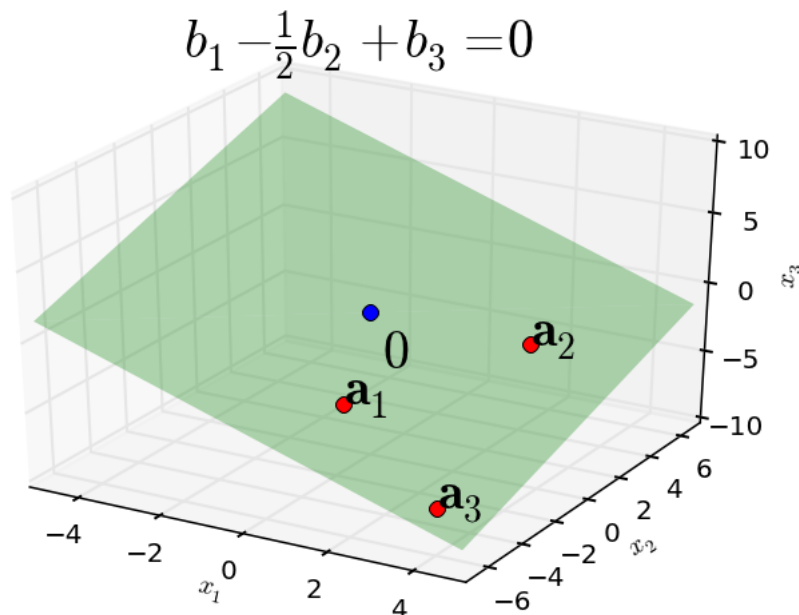
$$b_1 - \tfrac{1}{2}b_2 + b_3 = 0$$

The set of points that makes $b_1 - \frac{1}{2}b_2 + b_3 = 0$ is the set of $\mathbf{b}$ values that makes $A\mathbf{x} = \mathbf{b}$ consistent. But if $A\mathbf{x} = \mathbf{b}$ then $\mathbf{b}$ lies within $\mathrm{Span}\{\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}\}$.

So we have determined that $\mathrm{Span}\{\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}\}$ is the same point set as $b_1 - \frac{1}{2}b_2 + b_3 = 0$.

That means that $\mathrm{Span}\{\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}\}$ is a *plane* in this case. In the figure we have plotted $\mathbf{a_1}, \mathbf{a_2}$, and $\mathbf{a_3}$ and you can see that all three points, as well as the origin, lie in the same plane.

Now, in the most general case, four points will *not* line in a plane. In other words, the span of three vectors in $\mathbb{R}^3$ can be the whole space $\mathbb{R}^3$.

1.8 Question Time! Q4.3

Now we will tie together all three views of our question.

Theorem: 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 $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
2. Each $\mathbf{b}$ in $\mathbb{R}^n$ 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.

This theorem is very powerful and general. Notice that we started by thinking about a matrix equation, but we ended by making statements about A alone. This is the first time we are studying the property of a matrix itself; we will study many properties of matrices on their own as we progress.

1.9 Inner Product

There is a simpler way to compute the matrix-vector product. It uses a concept called the **inner product** or **dot product.**

The inner product is defined for two sequences of numbers, and returns a single number. For example, the expression

$$[3\ 5\ 1]\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$$

is the same as

$$(3 \cdot 2) + (5 \cdot -1) + (1 \cdot 4)$$

which is the sum of the products of the corresponding entries, and yields the scalar value 5. This is the inner product of the two sequences.

The general definition of the inner product is:

$$[x_1\ x_2\ \ldots\ x_n]\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i$$

Let's see how this can be used to think about matrix-vector multiplication.

Consider the matrix-vector multiplication

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= x_1\begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 \\ -1x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -1x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix}$$

Now, what is the first entry in the result? It is the *inner product* of the first row of A and $\mathbf{x}$. That is,

$$[2\ 3\ 4]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [2x_1 + 3x_2 + 4x_3]$$

This is a fast and simple way to compute the result of a matrix-vector product. Each element of the result is the *inner product* of the corresponding row of A and the vector $\mathbf{x}$.

$$\begin{bmatrix} 2 & 3 & 4 \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 & +3x_2 & +4x_3 \\ & & \end{bmatrix}$$

$$\begin{bmatrix} -1 & 5 & -3 \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_1 & +5x_2 & -3x_3 \\ & & \end{bmatrix}$$

1.10   Example.

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \cdot 4 & +2 \cdot 3 & -1 \cdot 7 \\ 0 \cdot 4 & -5 \cdot 3 & +3 \cdot 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

1.11 An (I)nteresting matrix

Consider this matrix equation:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} 1 \cdot r & +0 \cdot s & +0 \cdot t \\ 0 \cdot r & +1 \cdot s & +0 \cdot t \\ 0 \cdot r & +0 \cdot s & +1 \cdot t \end{bmatrix} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}.$$

This sort of matrix – with 1s on the diagonal and 0s everywhere else – is called an **identity matrix** and is denoted by I.

You can see that for any $\mathbf{x}$, $I\mathbf{x} = \mathbf{x}$.

1.12 Algebraic Properties of the Matrix-Vector Product

Theorem. 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: Let us take $n = 3$, $A = [\mathbf{a_1}\ \mathbf{a_2}\ \mathbf{a_3}]$, and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$.
Then to prove (1.):

$$A(\mathbf{u} + \mathbf{v}) = [\mathbf{a_1}\ \mathbf{a_2}\ \mathbf{a_3}] \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix}$$

$$= (u_1 + v_1)\mathbf{a_1} + (u_2 + v_2)\mathbf{a_2} + (u_3 + v_3)\mathbf{a_3}$$

$$= (u_1\mathbf{a_1} + u_2\mathbf{a_2} + u_3\mathbf{a_3}) + (v_1\mathbf{a_1} + v_2\mathbf{a_2} + v_3\mathbf{a_3})$$

$$= A\mathbf{u} + A\mathbf{v}.$$

And to prove (2.):

$$A(c\mathbf{u}) = [\mathbf{a_1}\ \mathbf{a_2}\ \mathbf{a_3}] \begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix}$$

$$= (cu_1)\mathbf{a_1} + (cu_2)\mathbf{a_2} + (cu_3)\mathbf{a_3}$$

$$= c(u_1\mathbf{a_1}) + c(u_2\mathbf{a_2}) + c(u_3\mathbf{a_3})$$

$$= c(u_1\mathbf{a_1} + u_2\mathbf{a_2} + u_3\mathbf{a_3})$$

$$= c(A\mathbf{u}).$$