Matrix-Vector Multiplication

Ryan C. Daileda

Trinity University

Linear Algebra

Recall

The linear system with coefficient matrix $A = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n)$ and "constant vector" $\mathbf{b}$ is represented by the augmented matrix

$$(A \quad \mathbf{b}),$$

and is equivalent to the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}.$$

We can put the list of variables $x_1, x_2, \ldots, x_n$ into a vector as well:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

The Matrix-Vector Product

We now seek to introduce notation that will express the entire linear system (coefficients, variables and constants) using only matrices and vectors.

With this in mind, we define the *matrix-vector product* to be

$$A\mathbf{x} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n.$$

That is, the product of a matrix A (on the left) with a vector (on the right) is the linear combination of the columns of A obtained by using the entries of $\mathbf{x}$ as weights.

Remarks

- We call $A\mathbf{x}$ a product and use multiplicative notation for reasons that will become clear shortly.
- We can only multiply an $m \times n$ matrix by a vector in $\mathbb{R}^n$. That is, in $A\mathbf{x}$ the matrix must have as many columns as the vector has entries.
- If we multiply an $m \times n$ matrix by a vector in $\mathbb{R}^n$, the result is a vector in $\mathbb{R}^m$.
- The linear system with augmented matrix $(A \quad \mathbf{b})$ can now be compactly represented as

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

Example

Let's multiply the matrix and vector

$$A = \begin{pmatrix} 1 & 4 & 0 \\ 4 & 2 & 0 \\ 1 & 5 & -2 \\ -1 & 4 & -4 \end{pmatrix} \quad \text{and} \quad \mathbf{v} = \begin{pmatrix} -3 \\ 3 \\ 4 \end{pmatrix}.$$

According to the definition we have

$$A\mathbf{v} = -3\begin{pmatrix} 1 \\ 4 \\ 1 \\ -1 \end{pmatrix} + 3\begin{pmatrix} 4 \\ 2 \\ 5 \\ 4 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 0 \\ -2 \\ -4 \end{pmatrix} = \begin{pmatrix} -3 \\ -12 \\ -3 \\ 3 \end{pmatrix} + \begin{pmatrix} 12 \\ 6 \\ 15 \\ 12 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -8 \\ -16 \end{pmatrix}$$

$$\begin{pmatrix} -3 \\ -12 \\ -3 \\ 3 \end{pmatrix} + \begin{pmatrix} 12 \\ 6 \\ 15 \\ 12 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -8 \\ -16 \end{pmatrix} = \begin{pmatrix} 9 \\ -6 \\ 4 \\ -1 \end{pmatrix}.$$

Notice that we first scaled the columns of A by the entries of $\mathbf{v}$ and then added the resulting vectors.

When working by hand, it is usually more efficient to work row by row.

That is, we scale the entries in the first row of A and add them, then scale the entries in the second row of A and add them, etc.

Thus, we would more quickly write

$$A\mathbf{v} = \begin{pmatrix} 1 & 4 & 0 \\ 4 & 2 & 0 \\ 1 & 5 & -2 \\ -1 & 4 & -4 \end{pmatrix} \begin{pmatrix} -3 \\ 3 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} -3 \cdot 1 + 3 \cdot 4 + 4 \cdot 0 \\ -3 \cdot 4 + 3 \cdot 2 + 4 \cdot 0 \\ -3 \cdot 1 + 3 \cdot 5 + 4 \cdot (-2) \\ (-3)(-1) + 3 \cdot 4 + 4(-4) \end{pmatrix} = \begin{pmatrix} 9 \\ -6 \\ 4 \\ -1 \end{pmatrix}.$$

The textbook calls this the *row-vector rule* for computing the matrix-vector product.

The Column Space

Given an $m \times n$ matrix $A = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n)$, we define its *column space* to be

$$\operatorname{Col} A = \operatorname{Span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\} \subseteq \mathbb{R}^m.$$

This gives us another way to talk about the existence of solutions to linear systems: the equation $A\mathbf{x} = \mathbf{b}$ has a solution iff $\mathbf{b} \in \operatorname{Col} A$.

Question. Given an $m \times n$ matrix A, when is it possible to solve $A\mathbf{x} = \mathbf{b}$ for *every* $\mathbf{b} \in \mathbb{R}^m$?

Equivalently, when do we have

$$\operatorname{Col} A = \mathbb{R}^m?$$

Notice that if we compute the RREF of the augmented matrix $(A \quad \mathbf{b})$, the first n columns (all but the last) will give the RREF of A itself.

To avoid the possibility of having a pivot in the last column of $(A \quad \mathbf{b})$ we *must* have a pivot in every row of A.

Therefore:

Theorem 1

Let A be an $m \times n$ matrix. The following are equivalent (TFAE):

1. *For every $\mathbf{b} \in \mathbb{R}^m$ the equation $A\mathbf{x} = \mathbf{b}$ has a solution.*
2. *Every vector in $\mathbb{R}^m$ is a linear combination of the cols. of A.*
3. $\operatorname{Col} A = \mathbb{R}^m$.
4. *A has a pivot in every row.*

Examples

Example 1

Do the columns of

$$A = \begin{pmatrix} -2 & -3 & 4 & 3 \\ 4 & -2 & 3 & 1 \\ 2 & 3 & -5 & 4 \end{pmatrix}$$

span $\mathbb{R}^3$?

Solution. We (partially) row reduce:

$$A \xrightarrow[R_1+R_3\to R_3]{2R_1+R_2\to R_2} \begin{pmatrix} -2 & -3 & 4 & 3 \\ 0 & -8 & 11 & 7 \\ 0 & 0 & -1 & 7 \end{pmatrix} = B.$$

Because B is in echelon form (although not reduced), we see that A does indeed have a pivot in every row.

Therefore,

$$\boxed{\text{Yes, the columns of } A \text{ span } \mathbb{R}^3.}$$

□

Example 2

Let A be an $m \times n$ matrix. If $m > n$, can the columns of A span $\mathbb{R}^m$?

Solution. Because there cannot be more than one pivot per row or column,

$$\text{total \# pivots of } A \leq \min\{m, n\} = n < m.$$

That is, there *cannot* be one pivot in each row of A.

Thus the columns of A *do not* span $\mathbb{R}^m$.

$\square$

Put another way, if we have more equations than variables in a linear system (the system is overdetermined), it is always possible to choose the RHS constants so that the system is inconsistent.

Properties of the Matrix-Vector Product

Given a matrix A, so far we have been concerned with whether or not we can solve $A\mathbf{x} = \mathbf{b}$ for a given $\mathbf{b}$.

That is, if $A\mathbf{x} = \mathbf{b}$, what does $\mathbf{b}$ tell us about $\mathbf{x}$?

Let's turn this question around. If we change $\mathbf{x}$, what happens to $A\mathbf{x}$?

Specifically, what happens if we apply our vector operations (addition and scalar multiplication) to $\mathbf{x}$?

Let $A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}$ be $m \times n$ and let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Using the algebraic properties of vector arithmetic, we then have

$$
\begin{aligned}
A(\mathbf{x} + \mathbf{y}) &= A \left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right) = A \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \\
&= (x_1 + y_1)\mathbf{a}_1 + (x_2 + y_2)\mathbf{a}_2 + \cdots + (x_n + y_n)\mathbf{a}_n \\
&= x_1\mathbf{a}_1 + y_1\mathbf{a}_1 + x_2\mathbf{a}_2 + y_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n + y_n\mathbf{a}_n \\
&= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n + y_1\mathbf{a}_1 + y_2\mathbf{a}_2 + \cdots + y_n\mathbf{a}_n \\
&= A\mathbf{x} + A\mathbf{y}.
\end{aligned}
$$

That is, the matrix-vector product distributes over vector addition!

Likewise, if $c \in \mathbb{R}$, we have

$$\begin{aligned} A(c\mathbf{x}) &= A \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix} \\ &= (cx_1)\mathbf{a}_1 + (cx_2)\mathbf{a}_2 + \cdots + (cx_n)\mathbf{a}_n \\ &= c(x_1\mathbf{a}_1) + c(x_2\mathbf{a}_2) + \cdots + c(x_n\mathbf{a}_n) \\ &= c\left(x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n\right) \\ &= c(A\mathbf{x}). \end{aligned}$$

In words, the matrix-vector product commutes with scalar multiplication!

Let's record these observations.

Theorem 2 (Properties of Matrix-Vector Multiplication)

Let A be an $m \times n$ matrix, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then:

1. $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$
2. $A(c\mathbf{x}) = c(A\mathbf{x})$

It is because of these properties that we call the matrix-vector operation $A\mathbf{x}$ "mutliplication."

Remark. Given a matrix A, the rule $\mathbf{x} \mapsto A\mathbf{x}$ defines a function

$$\mathbb{R}^n \to \mathbb{R}^m.$$

The properties of matrix-vector multiplication given above show that this function is *linear*.

Solutions of Linear Systems (Again)

We are now in a position to explain the parametric structure of solutions to linear systems more precisely.

A linear system of the form $A\mathbf{x} = \mathbf{0}$ is called *homogeneous*. A linear system of the form $A\mathbf{x} = \mathbf{b}$ with $\mathbf{b} \neq \mathbf{0}$ is called *inhomogeneous*.

Definition

Given a matrix A, the set of solutions to the homogeneous equation $A\mathbf{x} = \mathbf{0}$ is called the *null space* of A:

$$\operatorname{Null} A = \{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} = \mathbf{0}\}.$$

It turns out that the solutions to the general system $A\mathbf{x} = \mathbf{b}$ are related to the null space of A.

Fix $\mathbf{b} \in \mathbb{R}^m$ and consider the matrix equation $A\mathbf{x} = \mathbf{b}$.

Suppose that we have at least one solution $\mathbf{x}_0 \in \mathbb{R}^n$.

Consider any other solution $\mathbf{y} \in \mathbb{R}^n$.

Using the properties of the matrix-vector product we find that

$$\mathbf{0} = A\mathbf{y} - A\mathbf{x}_0 = A\mathbf{y} + (-1)(A\mathbf{x}_0) = A\mathbf{y} + A(-\mathbf{x}_0) = A(\mathbf{y} - \mathbf{x}_0).$$

Hence, $\mathbf{y} - \mathbf{x}_0 = \mathbf{z} \in \operatorname{Null} A$,or $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$, with $\mathbf{z} \in \operatorname{Null} A$.

We write this as $\mathbf{y} \in \mathbf{x}_0 + \operatorname{Null} A$.

On the other hand, suppose we choose any $\mathbf{y} \in \mathbf{x}_0 + \operatorname{Null} A$.

This means that $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$ for some $\mathbf{z}$ which satisfies $A\mathbf{z} = \mathbf{0}$. Hence

$$A\mathbf{y} = A(\mathbf{x}_0 + \mathbf{z}) = A\mathbf{x}_0 + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

This shows that the *complete* set of solutions to $A\mathbf{x} = \mathbf{b}$ is given by

$$\mathbf{x}_0 + \operatorname{Null} A,$$

where $\mathbf{x}_0$ is any *particular solution*.

Theorem 3

Let A be an $m \times n$ matrix. If the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution $\mathbf{x}_0$, then the complete set of solutions to $A\mathbf{x} = \mathbf{b}$ is described by

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{z},$$

where $\mathbf{z}$ is any vector satisfying $A\mathbf{z} = \mathbf{0}$.

This says that solution sets to equations of the form $A\mathbf{x} = \mathbf{b}$ are "structurally" the same: they are just translates of the null space of A (if they are nonempty).

In particular, if we can describe $\operatorname{Null} A$ in some "nice" way, then we get a "nice" description of the set of solutions to $A\mathbf{x} = \mathbf{b}$ in general.

Homogeneous equations are somewhat easier to solve than inhomogeneous equations. First of all, they *always* have solutions, since $A\mathbf{0} = \mathbf{0}$ (we always have $\mathbf{0} \in \operatorname{Null} A$).

To completely solve $A\mathbf{x} = \mathbf{0}$ we must row reduce $\begin{pmatrix} A & \mathbf{0} \end{pmatrix}$.

However, applying any sequence of elementary row operations to $\begin{pmatrix} A & \mathbf{0} \end{pmatrix}$ will yield a matrix of the form $\begin{pmatrix} B & \mathbf{0} \end{pmatrix}$.

So the final column is redundant (it never changes) and can be ignored. That is, we can simply row reduce the coefficient matrix A and just remember that there is an "invisible" final column of zeros.

Example 3

Compute the null space of

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

Solution. Row reduction of A yields

$$\begin{pmatrix} -2 & -3 & 4 & 3 \\ 4 & -2 & 3 & 1 \\ 2 & 3 & -5 & 4 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 0 & 1/4 \\ 0 & 1 & 0 & -21/2 \\ 0 & 0 & 1 & -7 \end{pmatrix}.$$

This tells us that the equation $A\mathbf{x} = \mathbf{0}$ is equivalent to the system

$$\begin{aligned} x_1 \qquad\qquad + \tfrac{1}{4}x_4 &= 0, \\ x_2 \qquad - \tfrac{21}{2}x_4 &= 0, \\ x_3 - 7x_4 &= 0. \end{aligned}$$

Thus, x_4 is free and

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -\frac{1}{4}x_4 \\ \frac{21}{2}x_4 \\ 7x_4 \\ x_4 \end{pmatrix} = x_4 \begin{pmatrix} -1/4 \\ 21/2 \\ 7 \\ 1 \end{pmatrix}.$$

Hence

$$\operatorname{Null} A = \operatorname{Span}\left\{ \begin{pmatrix} -1/4 \\ 21/2 \\ 7 \\ 1 \end{pmatrix} \right\}.$$

□

Example 4

Solve the matrix equation

$$\begin{pmatrix} -2 & -3 & 4 & 3 \\ 4 & -2 & 3 & 1 \\ 2 & 3 & -5 & 4 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix}.$$

Solution. We observe that

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

is a particular solution.

Therefore, by the preceding example, the general solution is given by

$$\boxed{\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + x_4 \begin{pmatrix} -1/4 \\ 21/2 \\ 7 \\ 1 \end{pmatrix}.}$$

□

Remark. Row reduction of the augmented matrix yields

$$\begin{pmatrix} -2 & -3 & 4 & 3 & 2 \\ 4 & -2 & 3 & 1 & 6 \\ 2 & 3 & -5 & 4 & 4 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 0 & 1/4 & 5/4 \\ 0 & 1 & 0 & -21/2 & -19/2 \\ 0 & 0 & 1 & -7 & -6 \end{pmatrix}.$$

This tells us that x_4 is free and

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} - \frac{1}{4}x_4 \\ -\frac{19}{2} + \frac{21}{2}x_4 \\ -6 + 7x_4 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5/4 \\ -19/2 \\ -6 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1/4 \\ 21/2 \\ 7 \\ 1 \end{pmatrix}.$$

Although this appears to be different than the solution we found above, as we vary x_4 in both cases we get exactly the same *set* of vectors.

What has happened is we have replaced the particular solution $(1, 1, 1, 1)$ with $(5/4, -19/2, -6, 0)$.

Example 5

Compute the null space of

$$A = \begin{pmatrix} 2 & -2 & -1 & 3 & 16 \\ 1 & -1 & 1 & -1 & 3 \\ 3 & -3 & 0 & -5 & -1 \\ -1 & 1 & 0 & 2 & 1 \end{pmatrix}.$$

Solution. Row reduction gives

$$\begin{pmatrix} 2 & -2 & -1 & 3 & 16 \\ 1 & -1 & 1 & -1 & 3 \\ 3 & -3 & 0 & -5 & -1 \\ -1 & 1 & 0 & 2 & 1 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore the equation $A\mathbf{x} = \mathbf{0}$ is equivalent to the system

$$\begin{aligned} x_1 \quad -x_2 \qquad\qquad\qquad +3x_5 &= 0, \\ x_3 \qquad -4x_5 &= 0, \\ x_4 \quad +2x_5 &= 0. \end{aligned}$$

The variables x_2 and x_5 are free, so that

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_2 - 3x_5 \\ x_2 \\ 4x_5 \\ -2x_5 \\ x_5 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -3 \\ 0 \\ 4 \\ -2 \\ 1 \end{pmatrix}.$$

Because x_2 and x_5 can take on any values, this means that

$$\text{Null}\, A = \text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 4 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

□

This happens in general: the solutions to $A\mathbf{x} = \mathbf{0}$ (the null space of A) can always be described as linear combinations of certain vectors using the free variables as weights.

We will discuss the details next time.