Matrix-Vector Products and the Matrix Equation $A\mathbf{x} = \mathbf{b}$

A. Havens

Department of Mathematics
University of Massachusetts, Amherst

January 31, 2018

Outline

1. Matrices Acting on Vectors
   - Linear Combinations and Systems
   - Matrix-Vector Products
   - Computing Matrix-Vector Products
2. The equation $A\mathbf{x} = \mathbf{b}$
   - Returning to Systems
   - Some Examples in three dimensions
3. Geometry of Lines and Planes in $\mathbb{R}^3$
   - Vector description of a line
   - Planes, Displacement Vectors, and Normals

A Recollection

Fix a collection $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of vectors in $\mathbb{R}^m$.

1. We can connect the question of whether a vector $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ to the question of whether the system with augmented matrix

$$\left[\begin{array}{ccc|c} \mathbf{a}_1 & \ldots & \mathbf{a}_n & \mathbf{b} \end{array}\right]$$

has a solution.

2. Indeed, there exists some collection of n real numbers $x_1, \ldots x_n$ such that

$$\mathbf{b} = x_1\mathbf{a}_1 + \ldots x_n\mathbf{a}_n$$

if and only if there is a solution $(x_1, \ldots x_n)$ to the system with the above augmented matrix.

Translating from systems to vector equations

In particular, if $\mathbf{b}$ is a linear combination of the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ then it must be the case that there is some matrix A' that is row-equivalent to the matrix $\mathrm{A} = \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix}$ such that

$$\mathrm{RREF}\left(\left[\mathbf{a}_1 \quad \ldots \quad \mathbf{a}_n \mid \mathbf{b}\right] \right) = \left[\mathrm{A}' \mid \mathbf{x}\right]$$

where

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

Conversely, if you can row reduce the augmented matrix of a system to obtain a solution, then you've realized the column vector of constants $\mathbf{b}$ as a linear combination of the columns of the coefficient matrix A.

Defining Matrix-vector Multiplication

The perspective above suggests that given an $m \times n$ matrix and a vector $\mathbf{x} \in \mathbb{R}^n$, there is a natural way to create a linear combination $x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n \in \mathbb{R}^m$ using the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of A.

Defining Matrix-vector Multiplication

Thus, we make the following definition:

Definition

Given an $m \times n$ matrix $\mathrm{A} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$ and a vector $\mathbf{x} \in \mathbb{R}^n$ we define the matrix vector product $\mathrm{A}\mathbf{x}$ to be the vector giving the linear combination

$$x_1\mathbf{a}_1 + \dots x_n\mathbf{a}_n \in \mathbb{R}^m$$

of the columns fo A, where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Observation

Matrix-vector products are only defined when the sizes of the matrix and vector are compatible – the number of components of the vector $\mathbf{x}$ must equal the number of *columns* of the matrix. The result will be a vector with as many components as the number of *rows* of the matrix.

Remark

Later, we will interpret matrix-vector products as describing a special kind of transformation, called a *linear transformation*. In particular, an $m \times n$ matrix acts on an n-vector $\mathbf{x} \in \mathbb{R}^n$ to produce an m-vector $\mathrm{A}\mathbf{x} \in \mathbb{R}^m$, so we can describe a certain kind of map of vectors from $\mathbb{R}^n$ to $\mathbb{R}^m$. If $n = m$, these linear transformations allow us to describe geometric transformations of space, such as rotations and reflections, as well as other more general maps of n-vectors.

An Example

Example

Consider the matrix

$$A = \begin{bmatrix} 1 & -4 & 7 \\ -2 & 5 & -8 \\ 3 & -6 & 9 \end{bmatrix}$$

and the vector

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}.$$

Compute $A\mathbf{x}$.

An Example

Example

The result of the matrix vector product $\mathrm{A}\mathbf{x}$ is the linear combination

$$\mathrm{A}\mathbf{x} = (2)\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} + (1)\begin{bmatrix} -4 \\ 5 \\ -6 \end{bmatrix} + (-1)\begin{bmatrix} 7 \\ -8 \\ 9 \end{bmatrix}$$

An Example

Example

By properties of scaling and vector addition:

$$A\mathbf{x} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix} + \begin{bmatrix} -4 \\ 5 \\ -6 \end{bmatrix} + \begin{bmatrix} -7 \\ 8 \\ -9 \end{bmatrix}$$
$$= \begin{bmatrix} 2-4-7 \\ -4+5+8 \\ 6-6-9 \end{bmatrix} = \begin{bmatrix} -9 \\ 9 \\ -9 \end{bmatrix}.$$

An Example

Example

We can write the computation slightly differently, to see a pattern that will make computation easier:

$$A\mathbf{x} = \begin{bmatrix} (1)(2) + (-4)(1) + (7)(-1) \\ (-2)(2) + (5)(1) + (-8)(-1) \\ (3)(2) + (-6)(1) + (9)(-1) \end{bmatrix}.$$

Observe that each entry is the result of taking the *row entries* and pairing them with the entries of $\mathbf{x}$ component-wise to make products, and then summing these products.

Dot Products

This construction will be familiar to anyone who has encountered the *dot product* of vectors:

Definition

Given two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ the *dot proodoct* $\mathbf{u} \cdot \mathbf{v}$ of $\mathbf{u}$ and $\mathbf{v}$ is the scalar quantity defined by the formula

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n \,.$$

We won't study this product in depth at the moment, but we will examine it in greater detail at the end of the course. It's also commonly encountered and studied in multivariable/vector calculus courses, and in physics courses.

Dot Product Properties

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be arbitrary vectors in $\mathbb{R}^n$, and $s \in \mathbb{R}$ any real scalar. Here's just a few properties of dot products, which I will not prove at this time:

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $(s\mathbf{u}) \cdot \mathbf{v} = s(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (s\mathbf{v})$
- $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
- $\mathbf{0} \cdot \mathbf{u} = 0$
- $\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2$, where $\|\mathbf{u}\| = \sqrt{u_1^2 + \ldots u_n^2}$ is the *magnitude* of $\mathbf{u}$.

A Geometric Interpretation

Dot products are not just a neat algebraic trick for computing matrix vector products; there's a handy geometric meaning as well.

Proposition

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ be two vectors separated by an angle of $\theta \in [0, \pi]$. Then the dot product $\mathbf{u} \cdot \mathbf{v}$ is the scalar quantity

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta \, .$$

We'll come back to this interpretation eventually, as it allows us to better understand the geometry of planes.

The Matrix-Vector Product in terms of Dot Products

Let $\mathbf{r}_1, \ldots, \mathbf{r}_m$ be vectors whose entries correspond to the *rows* of an $m \times n$ matrix A. Note that each $\mathbf{r}_i \in \mathbb{R}^n$. Then for any $\mathbf{x} \in \mathbb{R}^n$

$$\mathrm{A}\mathbf{x} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \mathbf{r}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{x} \end{bmatrix} \in \mathbb{R}^m .$$

Properties of the Matrix-Vector Product

Proposition

Let $\mathbf{u}$ *and* $\mathbf{v}$ *be arbitrary vectors in* $\mathbb{R}^n$*, let* $s \in \mathbb{R}$ *be any real scalar, and let* A *be any* $m \times n$ *matrix. Then the matrix vector product satisfies the following two properties:*

(i.) $\mathrm{A}(\mathbf{u} + \mathbf{v}) = \mathrm{A}\mathbf{u} + \mathrm{A}\mathbf{v}$,

(ii.) $\mathrm{A}(s\mathbf{u}) = s(\mathrm{A}\mathbf{u})$.

Are these properties familiar?
These will be the two properties we *demand* of *linear transformations*. That is, to be called linear, a function, map transformation, or whatever from one vector space to another must meet conditions like those above.

One Linear Equation to Capture them All

We now return to thinking about systems. Our discussion of linear combinations allowed us to conclude that numbers $x_1, \ldots, x_n$ solve a system of m equations with coefficient matrix A and column vector of constants $\mathbf{b}$ if and only if $\mathbf{b}$ was a linear combination of the columns of A, with $x_1, \ldots, x_n$ as the weights.

We can now rephrase this as follows: the vector $\mathbf{x}$ whose components are $x_1, \ldots, x_n$ solves the matrix-vector equation

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

if and only if $x_1, \ldots, x_n$ solve the system with augmented matrix $\left[\mathrm{A} \mid \mathbf{b}\right]$.

In particular, any linear system is captured by an equation of the form $\mathrm{A}\mathbf{x} = \mathbf{b}$.

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

Given a matrix A and a vector $\mathbf{b}$, solving $\mathrm{A}\mathbf{x} = \mathbf{b}$ amounts to expressing $\mathbf{b}$ as a linear combination of the columns of A, which one can do by solving the corresponding linear system.

Thus, to find a solution, one can row reduce the augmented matrix $\left[\mathrm{A} \mid \mathbf{b}\right]$.

A Proposition on Existence of Solutions

Proposition

Let A be an $m \times n$ matrix. Then the following statements are equivalent:

- *For every $\mathbf{b} \in \mathbb{R}^m$, the system $\mathrm{A}\mathbf{x} = \mathbf{b}$ has a solution,*
- *Each $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns of A,*
- *The span of the columns of A is all of $\mathbb{R}^m$,*
- *There is a pivot position in each row of A.*

Example: Is $\mathbf{b}$ in $\mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$?

Consider the vectors

$$\mathbf{a}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \qquad \mathbf{a}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 8 \\ 3 \\ -5 \end{bmatrix}.$$

Is $\mathbf{b} \in \mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$, i.e., does the vector **b** lie in the plane spanned by $\mathbf{a}_1$ and $\mathbf{a}_2$?

Let's reframe this as a matrix-vector equation.

Example: Is $\mathbf{b}$ in $\mathrm{Span}\,\{\mathbf{a}_1, \mathbf{a}_2\}$?

Thus, we want to know, is there a vector $\mathbf{x} \in \mathbb{R}^2$ such that

$$\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \begin{bmatrix} 2 & -1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

The last expression comes from the definition of the matrix vector product. If we compute the product, we recover a system in a vector form:

$$\begin{bmatrix} 2 & -1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_2 \\ x_1 \\ -x_1 + x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \\ -5 \end{bmatrix}.$$

Example: Is $\mathbf{b}$ in $\mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$?

We can thus row reduce the corresponding augmented matrix

$$\left[\begin{array}{cc|c} 2 & -1 & 8 \\ 1 & 0 & 3 \\ -1 & 1 & -5 \end{array}\right].$$

Row reducing, we have

$$\left[\begin{array}{cc|c} 2 & -1 & 8 \\ 1 & 0 & 3 \\ -1 & 1 & -5 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{array}\right].$$

Thus $\mathbf{b} = 3\mathbf{a}_1 - 2\mathbf{a}_2 \in \mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$.

Existence Problems for Plane Intersections

Here I pose two questions that we are now better equipped to answer and understand:

- Given two plane equations, when do they intersect in a single line? How can we describe this solution using vectors?
- Given three plane equations, what geometric conditions correspond to the *non-existence* of a solution?

Perhaps the nicest, and most geometric answers to these questions involve us thinking a bit about the dot product introduced above.

As far as describing *solution sets*, the idea is to use linear combinations to express the solutions in terms of some *parameters*. This will be the topic of discussion for the next lecture.

From Spans to Affine Lines

- Recall, we saw that for any nonzero vector $\mathbf{v} \in \mathbb{R}^n$, $\mathrm{Span}\,\{\mathbf{v}\} \subset \mathbb{R}^n$ describes a line through the origin $\mathbf{0}$.
- If we want an *affine line*, i.e., one not necessarily through the origin, we can displace the set of points of some $\mathrm{Span}\,\mathbf{v}$ by a constant vector, $\mathbf{b}$.
- Thus, the set of vectors $\mathbf{r}(t) = \mathbf{b} + t\mathbf{v}$ describes a line through the point with position vector $\mathbf{b}$ and parallel to the line $\mathrm{Span}\,\{\mathbf{v}\} = \{t\mathbf{v} \,|\, t \in \mathbb{R}\}$.

Three planes meeting in a line

Recall the system

$$\begin{cases} x + 2y + 3z &= -1 \\ 4x + 5y + 6z &= 0 \\ 7x + 8y + 9z &= 1 \end{cases}$$

which we had solved in the lecture on Gauss-Jordan elimination.

We found that the solution was of the form $(5/3 + z, -4/3 - 2z, z)$ for a free variable z. To emphasize that this is just a scalar parameter, I will re-label it $t = z$.

Three planes meeting in a line

We can rephrase this as seeking a vector $\mathbf{r} \in \mathbb{R}^3$ with components x, y, and z solving the matrix-vector equation

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Then observe that we can view the solution as the vector

$$\mathbf{r}(t) = \begin{bmatrix} 5/3 \\ -4/3 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},$$

for any real scalar t. This is thus a line through $(5/3, -4/3, 0)$. What's its direction?

In the preceding example, one of the plane equations was redundant, in the sense that the solution set would still be the aforementioned line if we just took two of the planes.

Can you express the last plane's equation as a linear combination of the other two equations?

What happens if we change the constant vector **b** so that its z-component is 2 instead of 1?

The Plane Equation as a Vector Equation

Recall, a plane in $\mathbb{R}^3$ has a description as the set of points (x, y, z) satisfying an equation of the form $ax + by + cz = d$, for some constants $a, b, c, d \in \mathbb{R}$.

We can rephrase this as a matrix-vector equation

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = [d] \,,$$

where the 1×1 matrix $[d]$ can be thought of as a 1-vector, i.e., a real number or simple scalar, but we included the brackets to emphasize that it is the vector result of a matrix-vector product of a 1×3 matrix with a 3-vector $\mathbf{x} \in \mathbb{R}^3$.

The Plane Equation via the Dot Product

An alternate description is via the dot product. Let

$$\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Then we can think of the above matrix-vector equation of a plane as

$$\mathbf{n}^{\mathrm{t}}\mathbf{x} = [\mathbf{n}\cdot\mathbf{x}] = [d]$$

which, removing the brackets, is just $\mathbf{n}\cdot\mathbf{x} = d$. Here, $\mathbf{n}^{\mathrm{t}}$ is the 1×3 matrix whose entries are those of the column vector $\mathbf{n}$ (this is an example of a construction called *transposing a matrix*, which turns an $m \times n$ matrix into an $n \times m$ matrix by reversing the roles of rows and columns).

Displacement Vectors

The conclusion is that the position vectors for points in the plane have constant dot product with $\mathbf{n}$!

We can take this conclusion further. Pick some initial position $\mathbf{x}_0$ known to satisfy the plane equation. Thus, $\mathbf{n} \cdot \mathbf{x}_0 = d$. Let $\mathbf{x}$ be another point on the plane.

Consider a *displacement vector* from $\mathbf{x}_0$ to $\mathbf{x}$. Such a vector is one that, if drawn originating at the position $\mathbf{x}_0$ points to the position $\mathbf{x}$.

Convince yourself via the rules for vector addition/subtraction that the components of this displacement vector are precisely the components of $\mathbf{x} - \mathbf{x}_0$.

Normal Vectors

What can we say about $\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0)$?

Computing $\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0)$, we have $\mathbf{n} \cdot \mathbf{x} - \mathbf{n} \cdot \mathbf{x}_0 = \mathbf{n} \cdot \mathbf{x} - d$. But since $\mathbf{x}$ is a position on the plane, $\mathbf{n} \cdot \mathbf{x} = d$, and so $\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0) = 0$!

The geometric interpretation of the dot product then tells us that the cosine of the angle between $\mathbf{n}$ and the displacement vector $\mathbf{x} - \mathbf{x}_0$ must be 0. This implies that $\mathbf{n}$ and $\mathbf{x} - \mathbf{x}_0$ are *perpendicular* to each other!

The vector $\mathbf{n}$ is called a *normal vector* for the plane. A normal vector $\mathbf{n}$ to a plane is any vector perpendicular to the plane.

Vectors Equation of a Plane

Thus, we have established the following vector formula for the equation of a plane: given a point $\mathbf{x}_0$ on a plane, and any normal vector $\mathbf{n}$ to the plane, the plane can be described as the set of points (x, y, z) whose position vectors $\mathbf{x}$ satisfy

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0) = 0\,.$$

The quantity $\mathbf{n} \cdot \mathbf{x}_0$ will be some constant, d, and writing out the dot product, one recovers the scalar formula:

$$0 = \mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0) = \mathbf{n} \cdot \mathbf{x} - d = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} - d$$

$$\implies 0 = ax + by + cz - d \iff ax + by + cz = d\,.$$

Here are some key observations about this result:

- The normal vector associated to a plane is not unique, since we could multiply both sides of $ax + by + cz = d$ by any nonzero scalar to get a new, equivalent equation with a scaled normal vector.
- There is however a unique *span* associated to a given plane. Given a line through the origin, and a point $\mathbf{x}_0$, we can construct a unique plane normal to that line, and containing the point $\mathbf{x}_0$.
- Existence and uniqueness questions can be rephrased using normal vectors!

Normal Vectors and two planes

Given two planes, if they possess normal vectors which are not parallel, they will meet in a line!

Indeed, if the normals aren't parallel, they are not multiples of each other.

The associated system of equations then has an augmented matrix with two rows and four columns, and the rows are not multiples of each other.

Thus there are precisely two pivot positions, and there is one free variable.

On the other hand, if two planes have parallel normals, then they are either the same plane, or disjoint parallel planes. The row operations on the augmented matrix will give you a row with zeros, except possibly the last entry. Consistency would demand that this be a zero as well!

A Challenge Problem

We know that two non-parallel vectors **u** and **v** in $\mathbb{R}^3$ span a plane in $\mathbb{R}^3$ through the origin. Can you use linear algebra to find the equation of this plane from components of **u** and **v**? Do this using linear systems, rather than using techniques you may have learned in other classes, such as computing *cross products*.

Note that this exercise is equivalent to computing some normal to the plane, using row reduction! By choosing an appropriate normal, you can actually recover the formula for the cross product of two vectors in $\mathbb{R}^3$.

Three Planes in $\mathbb{R}^3$

Here's a deeper challenge, which connects to what we'll be learning when we study matrix algebra, and linear independence:

Try to convince yourself that three planes in $\mathbb{R}^3$ have a unique intersection point precisely when their normals span $\mathbb{R}^3$.

Think about the inconsistent cases and the cases where the solution is a line, and try to see that one of the normals is a linear combination of the other two. The key is to connect the geometry of these normals to the number of pivot positions in the coefficient matrix.

Homework for Week 2

- MyMathLab for section 1.1 was due on Tuesday (last night). You can submit late for a penalty if you didn't get around to it last night.
- Section 1.2 in MyMathLab is due Thursday night.
- There will be a quiz Friday. Expect something which requires you to set up and solve a system, perhaps writing it initially in the for $A\mathbf{x} = \mathbf{b}$ for some matrix A and a given vector $\mathbf{b}$.
- Solutions to the two challenge problems regarding planes posed above can be submitted anytime before Friday, 2/9, for extra credit. The first problem is expanded upon in the next set of lecture slides.