Chapter 1

CAAM 335 • Matrix Analysis • Spring 2016

1 Introduction to Vectors and Matrices

In this class, we will look a variety of interesting problems drawn from the (nearly) real-world. What they all have in common is that they involve mathematical problems we can solve using *matrix* techniques. *Linear algebra* is the branch of math which studies matrices and vectors and is our main focus (although we will also use other areas of math as well.)

2 Vectors

Our basic objects in linear algebra are *vectors*.

Vectors

Vectors are lists of numbers:

$$\vec{v} = \begin{bmatrix} 1 \\ -3.5 \\ 0 \end{bmatrix} \qquad \vec{w} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Notation: Vectors will be denoted by lower-case letters with an arrow atop: $\vec{v}$, $\vec{w}$.

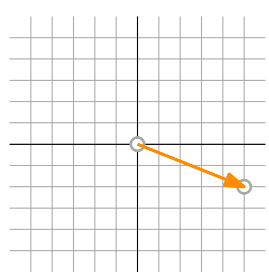
In our first examples, $\vec{v}$ is a 3-element vector of real numbers. The number of elements (or *dimension*) of a vector is important, so we introduce some further notation:

Notation: If n is a nonnegative integer, $\mathbb{R}^n$ is the set of all n-element vectors of real numbers.

Tip: n (and later m and p) will typically represent dimensions.

Looking at our example vectors, we have $\vec{v} \in \mathbb{R}^3$ (where $\in$ means "in"), and $\vec{w} \in \mathbb{R}^2$.

If you have seen vectors before, you have probably seen them illustrated as little arrows, with a direction and magnitude. For instance, $\vec{w}$ would look like this:



Some of the most familiar applications of vectors are in physics, where vectors represent velocity, momentum, force— or anything that has both a direction and magnitude.

However, the elements in a vector can be any selection of quantities we want to consider together. In this class, we will see vectors whose elements represent currents across different elements of a circuit, or the amounts different springs are stretched, or perhaps the heights of a wave at different points (if we get that far). Vectors are our basic objects. We give a special name, *scalars*, for things that are plain numbers rather than vectors.

2.1 Basic Operations

The basic operations we can perform on vectors are *addition* and *scalar multiplication*. Two vectors of the same length can be added by adding the correpsonding elements together:

$$\begin{bmatrix} 1 \\ -3.5 \\ 0 \end{bmatrix} + \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ -0.5 \\ 2 \end{bmatrix}.$$

Scalar multiplication, as its name clues, involves multiplying a vector by a scalar. Each element of the vector is multiplied by the scalar:

$$2 \cdot \begin{bmatrix} 1 \\ -3.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -7 \\ 0 \end{bmatrix}.$$

Vector Basic Operations

Notation 3: If $\vec{v}$ is a vector, v_j represents its j^{th} element.

If $\vec{v}, \vec{w} \in \mathbb{R}^n$, and $c \in \mathbb{R}$, then

$$\begin{aligned} v + w &= u \in \mathbb{R}^n, & \text{where } u_j &= v_j + w_j, \\ cv &= z \in \mathbb{R}^n, & \text{where } z_j &= cv_j. \end{aligned}$$

What about multiplying two vectors together? Just like vector addition, if we have two vectors $\vec{u}$, $\vec{v}$ of the same dimension, we could multiply their corresponding elements together. It turns out, though, this is not an operation we will often need. However, if we multiply corresponding elements, then sum, we get a very useful operation, the *inner product* or *dot product*. This is written as $\vec{v} \cdot \vec{w}$ or $\vec{v}^\top \vec{w}$ (the reason for the $\top$ notation will appear later).

Inner Product

If $\vec{v}, \vec{w} \in \mathbb{R}^n$, then

$$\vec{v} \cdot \vec{w} = \vec{v}^\top \vec{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = \sum_{i=1}^{n} v_i w_i.$$

One of the uses of the inner product is for finding the length, or *norm*, of a vector. A norm is a way of measuring how large a vector is. Depending on the application, we may be interested in different norms. The inner product provides us with one way of measuring vectors: the *2-norm* or *Euclidean norm*.

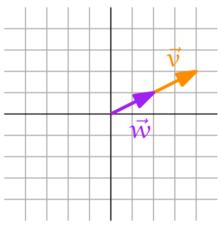
Euclidean Norm

The *Euclidean norm* of $\vec{v} \in \mathbb{R}^n$ is

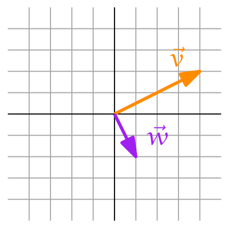
$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

As an example, the norm of $\vec{w} = \left[\begin{smallmatrix} 5 \\ -2 \end{smallmatrix}\right]$ works out to be $\sqrt{5^2 + (-2)^2} = \sqrt{29}$.

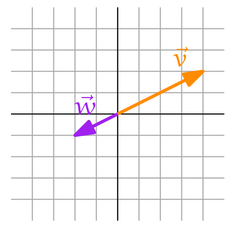
Besides providing information on a vector's length, the inner product tells us how much two vectors point in the same direction. For some rough intuition, here are three examples of $\vec{v} \cdot \vec{w}$ with the same $\vec{v}$ but different choices of $\vec{w}$:



$\vec{v} \cdot \vec{w} = +10$



$\vec{v} \cdot \vec{w} = +0$



$\vec{v} \cdot \vec{w} = -10$

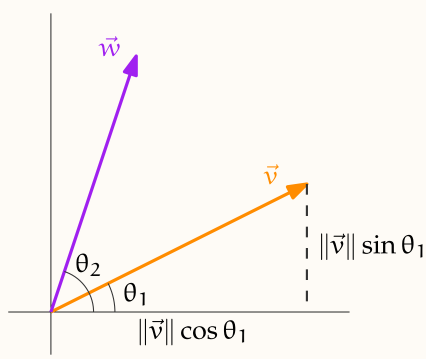
Here is the precise statement relating two vectors' directions and their inner product:

Inner Product Geometry

If $\vec{v}, \vec{w} \in \mathbb{R}^n$, $n \geq 2$, and the angle between them is θ, then

$$\vec{v} \cdot \vec{w} = \|v\| \, \|w\| \cos \theta.$$

Proof for $n = 2$*:*



In the picture, $\vec{v}, \vec{w}$ are in the first quadrant, but this isn't necessary for what follows. From trigonometry, we know.

$$\begin{aligned} v_1 &= \|v\| \cos \theta_1 & \qquad w_1 &= \|w\| \cos \theta_2 \\ v_2 &= \|v\| \sin \theta_1 & \qquad w_2 &= \|w\| \sin \theta_2 \end{aligned}$$

From the definition of inner products,

$$\begin{aligned}\vec{v}\cdot\vec{w} = v_1w_1 + v_2w_2 &= \|v\|\,\|w\|\,(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2)\\ &= \|v\|\,\|w\|\cos(\theta_1-\theta_2)\\ &= \|v\|\,\|w\|\cos\theta.\end{aligned}$$

One consequence of this geometric result is the *Cauchy–Schwarz inequality*. Since $|\cos\theta|$ is at most 1, we can conclude the following:

Cauchy–Schwarz

For any $\vec{v},\vec{w}\in\mathbb{R}^n$,

$$|\vec{v}\cdot\vec{w}| \le \|v\|\,\|w\|\,.$$

Before leaving inner products, we make a final note. In linear algebra, the property of two vectors having an inner product of zero comes up over and over, so it deserves a name of its own: *orthogonality*.

Orthogonal Vectors

Vectors $\vec{v},\vec{w}$ are *orthogonal* if $\vec{v}\cdot\vec{w}=0$.

3 Matrices

In this class, vectors will typically represent the *state* of some system of interest in our problem. Often, we will want to transform the system's current state in some way, either to obtain a new state, or to derive some new quantities of interest that depend on the current state. In mathematical terms, such a transformation (or *operator*) takes a vector $\vec{u}\in\mathbb{R}^n$ and in return gives us a new vector $\vec{v}\in\mathbb{R}^m$. If our problems has constraints, those can also be rephrased in terms of operators.

It turns out that many operators that arise in applications have a special property, *linearity*. (And, people often try to approximate non-linear operators by linear operators!) Linear operators can be boiled down to a two-dimensional array of numbers: a *matrix*. So, we will start by considering matrices and how they operate on vectors.

Matrices

A *matrix* is a rectangular array of numbers:

$$A=\begin{bmatrix}2 & -2\\ 0 & 1\\ 3 & 5\end{bmatrix}\qquad B=\begin{bmatrix}0 & 0 & -1 & 1\\ 0 & -1 & 1 & 0\\ -1 & 1 & 0 & 0\end{bmatrix}\qquad C=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$$

Notation: A matrix with m rows and n columns is said to be m-*by*-n. The set of all m-by-n matrices with real entries is denoted $\mathbb{R}^{m\times n}$.

So, for the example above, $A \in \mathbb{R}^{3\times 2}$, $B \in \mathbb{R}^{3\times 4}$.

When we need to pinpoint an element from a matrix, we give its row number, then the column number. Conventionally, we use uppercase letters for matrices and the corresponding lowercase letters for their entries.

Notation: If A is a matrix, a_{ij} is the entry in the i^{th} row and j^{th} column.

For instance, $a_{32} = 5$ and $b_{14} = 1$.

3.1 Basic Matrix Operations

Like vectors, two matrices of the same size can be added by adding corresponding entries. Likewise, multiplying a matrix by a scalar is defined by multiplying each entry by that scalar:

$$\begin{bmatrix} 2 & -2 \\ 0 & 1 \\ 3 & 5 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 1 & 1 \\ 4 & 5 \end{bmatrix}, \qquad 3 \cdot \begin{bmatrix} 2 & -2 \\ 0 & 1 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 6 & -6 \\ 0 & 3 \\ 9 & 15 \end{bmatrix}.$$

3.2 Matrix-Vector Multiplication

So far, matrices are just inert piles of numbers. However, as promised, they are really operators on vectors in disguise. Starting from an $m \times n$ matrix A can be multiplied with a vector in $\mathbb{R}^n$, we can generate a new vector in $\mathbb{R}^m$. The new vector's entries come from taking the inner product of each row of the matrix with the vector. This process is called *matrix-vector multiplication*, and written like multiplication.

For instance, we can multiply $A = \left[\begin{smallmatrix} 2 & -2 \\ 0 & 1 \\ 3 & 5 \end{smallmatrix}\right]$ and $\vec{w} = \left[\begin{smallmatrix} 5 \\ -2 \end{smallmatrix}\right]$. Going row-by-row through A, we form three inner products:

$$Aw = \begin{bmatrix} 2 \cdot 5 + (-2) \cdot (-2) \\ 0 \cdot 5 + 1 \cdot (-2) \\ 3 \cdot 5 + 5 \cdot (-2) \end{bmatrix} = \begin{bmatrix} 14 \\ -2 \\ 5 \end{bmatrix}.$$

Here is the operation codified in symbols:

Matrix Multiplication

Let $A \in \mathbb{R}^{m\times n}$ and $\vec{v} \in \mathbb{R}^n$. Then

$$Av = w \in \mathbb{R}^m, \quad \text{where } w_i = \sum_{j=1}^{n} a_{ij} v_j.$$

3.3 Example of a Matrix's Operation

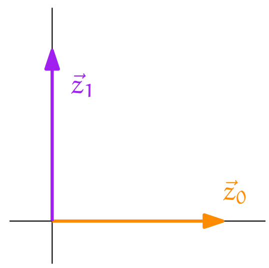
Let's consider a more useful example of a matrix acting as an operator:

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Applying A to the vector $\vec{z}_0 = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$:

$$A\vec{z}_0 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 - 1 \cdot 0 \\ 1 \cdot 1 + 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

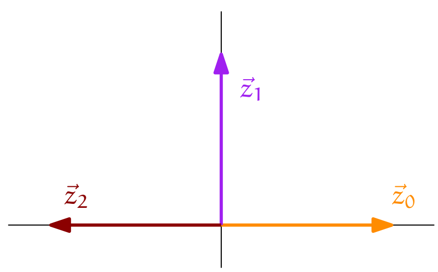
If we call this new vector $\vec{z}_1$, and plot it with $\vec{z}_0$, we get:



Applying A again, we have

$$A\vec{z}_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

Let $\vec{z}_2$ be this new vector. Plotting it too, we find a pattern emerging:



It appears that A is rotating vectors $\frac{\pi}{2}$ radians counterclockwise, and that's true.

3.4 Matrix Multiplication

If applying one matrix is useful, maybe we want to apply more. Suppose then we have two matrices $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$. Given a vector $\vec{v} \in \mathbb{R}^p$, we can multiply it by B to it to get a vector in $\mathbb{R}^n$, then multiply it by A to get a new vector $\vec{w} \in \mathbb{R}^m$. Is there one matrix that goes straight from $\vec{v}$ to $\vec{w}$? The answer is yes.

Matrix-Matrix Multiplication

Let $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$. Then for all $\vec{v} \in \mathbb{R}^p$,

$$A(B\vec{v}) = C\vec{v}, \qquad \text{where } c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \qquad (C \in \mathbb{R}^{m\times p}).$$

The matrix C above is the *product* of A and B, and we write $C = AB$.

Looking carefully at this result, note that c_{ij}— the entry in the i^{th} row and j^{th} column— is the inner product of the i^{th} row of A and the j^{th} column of B.

Now, in general $AB \neq BA$. For one thing, we can't multiply just any two matrices together; their sizes must match! Matrix multiplication is only defined when the number of columns in the first matrix equals the number of rows in the second. But even if AB and BA are both defined, they aren't usually the same. For instance, take $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (our rotation matrix), and $B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix}$$
$$BA = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix}$$

Geometrically, B stretches a vector in the x-direction by 2. The fact that $AB \neq BA$ indicates that rotating then stretching is different from the opposite order, stretching then rotating— is that a surprise?

3.5 More Operations and Facts

3.5.1 Transpose

In this section we will cover a few more handy operations on matrices. The first is the *transpose*, which flips rows and columns. It is denoted by a superscript $^\top$:

$$A = \begin{bmatrix} 2 & -2 \\ 0 & 1 \\ 3 & 5 \end{bmatrix} \qquad A^\top = \begin{bmatrix} 2 & 0 & 3 \\ -2 & 1 & 5 \end{bmatrix}.$$

Transpose

The transpose of $A \in \mathbb{R}^{m\times n}$ is the matrix $A^\top \in \mathbb{R}^{n\times m}$ with entries

$$(a^\top)_{ij} = a_{ji}.$$

The transpose, despite its odd appearance, pops up everywhere in linear algebra. Here's one reason: transposes let us move matrices around in expressions involving inner products.

Transpose and Inner Product

If $\vec{v} \in \mathbb{R}^n$, $\vec{w} \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$, then

$$(A\vec{v}) \cdot \vec{w} = \vec{v} \cdot (A^\top \vec{w}).$$

Proof: Let's compare both sides of the equation:

$$(A\vec{v}) \cdot \vec{w} = \sum_{i=1}^{m} (Av)_i w_i = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} v_j w_i \tag{$\dagger$}$$

$$\vec{v} \cdot (A^\top \vec{w}) = \sum_{j=1}^{n} v_j (A^\top w)_j = \sum_{j=1}^{n} \sum_{i=1}^{m} v_j (a^\top)_{ji} w_i.$$

Since $(a^\top)_{ij} = a_{ji}$, the two sides are equal.

This fact also lets us transpose products.

Transposes of Products

The transpose of the product is the reversed product of the transposes. If $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$.

$$(AB)^\top = B^\top A^\top.$$

Proof: Using the inner product property we just derived,

$$(AB\vec{v}) \cdot \vec{w} = (B\vec{v}) \cdot (A^\top \vec{w}) = \vec{v} \cdot (B^\top A^\top \vec{w}).$$

But on the other hand, the property also says

$$(AB\vec{v}) \cdot \vec{w} = \vec{v} \cdot [(AB)^\top \vec{w}]$$

for all $\vec{v}$ and $\vec{w}$. It looks like $B^\top A^\top$ might indeed be $(AB)^\top$.

You can now check that if M and N are two matrices of the same size, and $\vec{v} \cdot (M\vec{w}) = \vec{v} \cdot (N\vec{w})$ for all $\vec{v}, \vec{w}$, then $M = N$. (*Hint:* Look at equation ($\dagger$) above, and then show for any given i, j you can find $\vec{v}$ and $\vec{w}$ so that $\vec{v} \cdot (M\vec{w}) = m_{ij}$.)

3.5.2 Special Classes of Matrices

In linear algebra, there are a number of special types of matrices that have nice properties, whether for applications, for theory, or for computations. These are best illustrated by picture (nonzero elements are marked by $*$).

$$\begin{bmatrix} * & & \\ & * & \\ & & * \end{bmatrix} \quad \begin{bmatrix} * & * & * \\ & * & * \\ & & * \end{bmatrix} \quad \begin{bmatrix} * & & \\ * & * & \\ * & * & * \end{bmatrix} \quad \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ & * & * & * \\ & & * & * \end{bmatrix}$$

Diagonal Upper Triangular Lower Triangular Upper Hessenberg

For short, we use **diag** to abbreviate diagonal matrices:

$$\operatorname{diag}(1,2,3) = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}.$$

3.5.3 Trace

The next operation is the *trace*, written tr. It only works on square matrices, and is just the sum of the *diagonal* entries of the matrix:

$$\operatorname{tr} \begin{bmatrix} 3 & 0 & 2 \\ -1 & 5 & 7 \\ 2 & 4 & 0 \end{bmatrix} = 8.$$

Trace

If $A \in \mathbb{R}^{n\times n}$ is a square matrix,

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}.$$

3.5.4 Norms

As with vectors, there are many different ways of measuring how "big" a matrix is, or in other words, its norm. Having seen the 2-norm for vectors, the natural extension for matrices is the *Frobenius norm*:

Frobenius Norm

The Frobenius norm of $A \in \mathbb{R}^{m\times n}$ is

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}.$$