1 Last time: Vectors

A *(column) vector*

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

is matrix with one column. A vector has the same data as a list of real numbers.

The *dimension* of a vector is the number of its rows.

Let $\mathbb{R}^n$ be the set of n-dimensional vectors.

We can add two vectors of the same dimension: $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$.

We can multiply a vector by a *scalar*: $cv = c \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}$ for $c \in \mathbb{R}$ and $v \in \mathbb{R}^n$

The word "scalar" is a synonym for real number.

It is sometimes useful to visualize vectors $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2$ as arrows in the Cartesian plane from the origin the point $(x, y) = (a_1, a_2)$:
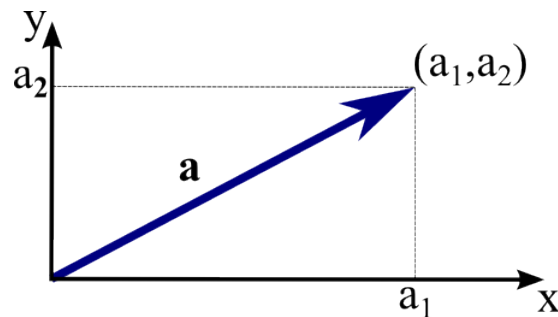
Relative to this picture, the sum $a + b$ of two vectors $a, b \in \mathbb{R}^2$ is the vector represented by the arrow from the origin to the point which is the opposite vertex of the parallelogram with sides a and b:
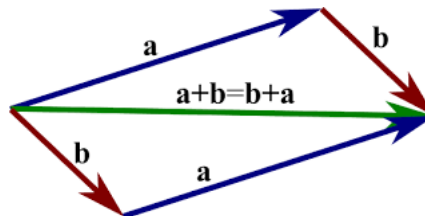
The *zero vector* $0 \in \mathbb{R}^n$ is the vector

$$0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

whose entries are all zero. Note that $0 + v = v + 0 = v$ for any vector v.

A *linear combination* of vectors $v_1, v_2, \ldots, v_p \in \mathbb{R}^n$ is any vector of the form

$$y = c_1 v_1 + c_2 v_2 + \cdots + c_p v_p \in \mathbb{R}^n$$

where $c_1, c_2, \ldots, c_p \in \mathbb{R}$.

The *span* of some vectors $v_1, v_2, \ldots, v_p \in \mathbb{R}^n$ is the set of all of their linear combinations. Denote this by

$$\mathbb{R}\text{-span}\{v_1, v_2, \ldots, v_p\} \qquad \text{or} \qquad \text{span}\{v_1, v_2, \ldots, v_p\}.$$

Proposition. If $v_1, v_2, \ldots, v_p \in \mathbb{R}^n$, then a vector $y \in \mathbb{R}^n$ belongs to $\mathbb{R}\text{-span}\{v_1, v_2, \ldots, v_p\}$ if and only if the matrix $\begin{bmatrix} a_1 & a_2 & \ldots & a_n & b \end{bmatrix}$ is the augmented matrix of a consistent linear system.

The span of vectors in $\mathbb{R}^2$ can be interpreted geometrically as either a point (at the origin), a line (through the origin), or the whole plane $\mathbb{R}^2$.

2   Multiplying vector by matrices

We have seen that one way to view a matrix is as a compact notation for representing a linear system.

Today we introduce a second, perhaps more fundamental way of viewing a matrix: namely, as an object that transforms one vector to another.

How does a matrix transform a vector?

Definition. If A is a matrix with columns $a_1, a_2, \ldots, a_n \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$, so that

$$A = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} \qquad \text{and} \qquad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

then the *matrix-vector product* Av is the vector in $\mathbb{R}^m$ given by the linear combination of the columns of A with coefficients from v:

$$Av = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1 a_1 + v_2 a_2 + \cdots + v_n a_n \in \mathbb{R}^m.$$

Example. If $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}$ and $v = \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix}$ then $a_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $a_2 = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, and $a_3 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ so

$$Av = 4a_1 + 3a_2 + 7a_3 = \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Example. If $A = \begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix}$ and $v = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$ then $a_1 = \begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix}$ and $a_2 = \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}$ so we have

$$Av = 4a_1 + 7a_2 = 4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}.$$

Note that if A is $m \times n$ then Av is only defined for $v \in \mathbb{R}^n$, and in this case we have $Av \in \mathbb{R}^m$. Thus A transforms vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$.

This transformation is *linear*:

1. If A is an $m \times n$ matrix and $u, v \in \mathbb{R}^n$ then $A(u + v) = Au + Av$.

2. If A is an $m \times n$ matrix and $v \in \mathbb{R}^n$ and $c \in \mathbb{R}$ then $A(cv) = c(Av)$.

Proof. If the columns of A are $a_1, a_2, \ldots, a_n \in \mathbb{R}^m$ and the numbers in the ith row of u, v are u_i, v_i, then

$$A(u + v) = (u_1 + v_1)a_1 + (u_2 + v_2)a_2 + \cdots = (u_1 a_1 + u_2 a_2 + \ldots) + (v_1 a_1 + v_2 a_2 + \ldots) = Au + Av.$$

If $c \in \mathbb{R}$ then

$$A(cv) = (cv_1)a_1 + (cv_2)a_2 + \cdots = c(v_1 a_1 + v_2 a_2 + \ldots) = c(Av).$$

$\square$

Let A and v be the general $m \times n$ matrix and n-dimensional vector given by

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Quick way to compute Av: match up entries in the ith column of A with the entry in the ith row of v.

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}.$$

For example, $\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} = 1 \cdot 5 + 2 \cdot 6 + 3 \cdot 7 + 4 \cdot 8 = 5 + 12 + 21 + 32 = 70.$

3   Matrix equations

If A is an $m \times n$ matrix with columns $a_1, a_2, \ldots, a_n \in \mathbb{R}^m$ and $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a vector where each x_i is

a variable and $b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$, then we call $Ax = b$ a *matrix equation*.

Proposition. The matrix equation $Ax = b$ has the same solutions as the vector equation

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b$$

and also the same solutions as the linear system whose augmented matrix is $\begin{bmatrix} a_1 & a_2 & \ldots & a_n & b \end{bmatrix}$.

Proposition. The matrix equation $Ax = b$ has a solution if and only if b is a linear combination of the columns of A, that is, $b \in \mathbb{R}\text{-span}\{a_1, a_2, \ldots, a_n\}$.

Example. Let $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$ and $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

Does $Ax = b$ have a solution for all choices of $b_1, b_2, b_3 \in \mathbb{R}$?

Note that $Ax = b$ has a solution if and only if

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix}$$

is the augmented matrix of a consistent linear system. We can determine if this system is consistent by row reducing the matrix to echelon form:

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

The last matrix is in echelon form, so its leading entries are the pivot positions of our first matrix. The corresponding linear system is consistent if and only if the last column does not contain a pivot position. This occurs precisely when $b_1 - \frac{1}{2}b_2 + b_3 = 0$.

But we can choose numbers such that $b_1 - \frac{1}{2}b_2 + b_3 \neq 0$: take $b_1 = 1$ and $b_2 = b_3 = 0$. Therefore our original matrix equation $Ax = b$ does not always have a solution.

We can generalize this example:

Theorem. Let A be an $m \times n$ matrix. If one of the following holds, then all of the statements hold. If one of the following fails, then all of the statements fail:

1. For each vector $b \in \mathbb{R}^m$, the matrix equation $Ax = b$ has a solution.

2. Each vector $b \in \mathbb{R}^m$ is a linear combination of the columns of A.

3. The span of the columns of A is all of $\mathbb{R}^m$ (say this as: "the columns of A span $\mathbb{R}^m$").

4. A has a pivot position in every row.

Proof. (1)-(3) are different ways of saying the same thing.

We must check that (1)-(3) are equivalent to (4), which is less obvious.

If A has a pivot position in every row, then the augmented matrix $\begin{bmatrix} A & b \end{bmatrix}$ cannot have a pivot position in the last column; saying that A has a pivot position in every row means that $\begin{bmatrix} A & b \end{bmatrix}$ has to be row equivalent to something like

$$\begin{bmatrix} 0 & 1 & * & * & * & c_1 \\ 0 & 0 & 0 & 4 & * & c_2 \\ 0 & 0 & 0 & 0 & 3 & c_3 \end{bmatrix}$$

where c_1, c_2, c_3 are numbers (i.e., 1-dimensional vectors) given by linear combinations of b_1, b_2, b_3. Regardless of what c_1, c_2, c_3 are, the given matrix has pivot columns 2, 4 and 5 but not 6.

We saw last time that not having a pivot position in the last column means that $\begin{bmatrix} A & b \end{bmatrix}$ is the augmented matrix of a consistent linear system.

On the other hand, if A doesn't have a pivot position in some row, then it is always possible to choose b such that $\begin{bmatrix} A & b \end{bmatrix}$ has a pivot position in the last column, in which case the corresponding linear system has no solution. (Think about why this is true!) ☐

This theorem is very useful and will be explored in problems on this week's homework and later on.

4   Homogeneous linear systems

Any system of linear equations

$$a_{11}x_1 + a_{12}x_2 + \cdots + x_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + x_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

can be written as the matrix equation $Ax = b$ where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad \text{and} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

A linear system is *homogeneous* if it can be written as $Ax = 0$ where $0 \in \mathbb{R}^m$ is the zero vector.

<u>Important easy fact.</u> The homogeneous equation $Ax = 0$ always has a solution given by $x = 0 \in \mathbb{R}^n$.

We call $x = 0$ the *trivial solution* to $Ax = 0$. A nonzero vector $x \in \mathbb{R}^n$ is a *nontrivial solution* if $Ax = 0$. A homogeneous matrix equation may or may not have nontrivial solutions.

Proposition. The equation $Ax = 0$ has a nontrivial solution if and only if the corresponding linear system has at least one free variable.

Proof. If the system has no free variables, then it has either zero solutions or one solution. Since $x = 0$ already gives one solution to $Ax = 0$, the system must have exactly one solution if it has no free variables, but this solution is the trivial one. □

Theorem. Suppose the matrix equation $Ax = b$ has a solution $x_0 \in \mathbb{R}^n$. Every solution of this equation has the form $x_0 + h$ where $h \in \mathbb{R}^n$ is an arbitrary solution to the homogeneous equation $Ax = 0$.

Proof. Let $x_1 \in \mathbb{R}^n$ be another solution to $Ax = b$. We want to show that $x_1 = x_0 + h$ for a vector $h \in \mathbb{R}^n$ which solves $Ax = 0$. Clearly we have to have $h = x_1 - x_0$. But this does satisfy $Ah = A(x_1 - x_0) = Ax_1 + A(-x_0) = Ax_1 - Ax_0 = b - b = 0$. □

5   Linear independence

Let $v_1, v_2, \ldots, v_p$ be vectors in $\mathbb{R}^n$.

These vectors are *linearly independent* if the homogeneous vector equation

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0$$

has only trivial solutions, i.e., if $x_1 v_1 + \cdots + x_p v_p = 0$ if and only if $x_1 = x_2 = \cdots = x_p = 0$.

The vectors $v_1, v_2, \ldots, v_p$ are *linearly dependent* otherwise, i.e., if there are some numbers $c_1, c_2, \ldots, c_p \in \mathbb{R}$, at least one of which is nonzero, such that $c_1 v_1 + c_2 v_2 + \ldots c_p v_p = 0$.

Example. If $v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ and $v_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ then

$$v_1 + v_3 = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} \quad \text{and} \quad v_2 + v_3 = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$

so $2(v_1 + v_3) - (v_2 + v_3) = 2v_1 - v_2 + v_3 = 0$. Hence v_1, v_2, v_3 are linearly dependent.

It is usually not so easy to guess whether a given list of vectors is linearly independent or not. In general, to do this we have to determine whether a certain homogeneous linear system has a nontrivial solution, which involves reducing its matrix to echelon form.

Note that the columns of a matrix A are linearly independent if and only if the equation $Ax = 0$ has no nontrivial solution.

Example. Some useful observations:

1. A list containing a single v is linearly independent if and only if $v \neq 0$.

2. Two vectors $u, v \in \mathbb{R}^n$ are linear dependnet if and only if we can write $au + bv = 0$ for numbers $a, b \in \mathbb{R}$ with $a \neq 0$ or $b \neq 0$. If $a \neq 0$ then we have $u = (-b/a)v$. If $b \neq 0$ then $v = (-a/b)u$. (Note that both of these cases could occur.) Thus:

   Two vectors are linearly independent if and only if neither is a scalar multiple of the other.

3. If some $v_i = 0$ then $v_1, v_2, \ldots, v_p$ are linear dependent, since then

   $$0v_1 + \cdots + 0v_{i-1} + 5v_i + 0v_{i+1} + \cdots + 0v_p = 0.$$

   (The scalar 5 here can be replaced by any number.)

Characterization of linearly dependent vectors. The vectors $v_1, v_2, \ldots, v_p \in \mathbb{R}^n$ are linearly dependent if and only if some vector v_i is a linear combination of the other vectors $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_p$.

Proof. We first show that if the vectors are linearly dependent then some vector is a linear combination of the others. Suppose $c_1 v_1 + \cdots + c_p v_p = 0$ where $c_i \neq 0$. Then

$$v_i = (-c_1/c_i)v_1 + (-c_2/c_i)v_2 + \cdots + (-c_{i-1}/c_i)v_{i-1} + (-c_{i+1}/c_i)v_{i+1} + \cdots + (-c_p/c_i)v_p$$

so v_i is a linear combination of $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_p$.

Conversely, if we can write $v_i = c_1 v_1 + \cdots + c_{i-1}v_{i-1} + c_{i+1}v_{i+1} + \cdots + c_p v_p$ for any coefficients in $\mathbb{R}$, so that v_i is a linear combination of the remaining vectors, then

$$c_1 v_1 + \cdots + c_{i-1}v_{i-1} - v_i + c_{i+1}v_{i+1} + \cdots + c_p v_p = 0$$

which means that the vectors are linearly dependent, since the coefficient of at least v_i is nonzero. $\square$

We conclude this lecture with a useful, non-obvious fact:

Theorem. Suppose $v_1, v_2, \ldots, v_p \in \mathbb{R}^n$. If $p > n$ then these vectors are linearly dependent.

Proof. Saying these vectors are linearly dependent is the same thing as saying that the $n \times (p+1)$ matrix

$$A = \begin{bmatrix} v_1 & v_2 & \ldots & v_p & 0 \end{bmatrix}$$

is the augmented matrix of a linear system with at least one free variable. A variable x_i for $1 \leq i \leq p$ is free for this system precisely when i is a pivot column of A. There can only be 1 pivot position in each row, so there can be at most n pivot columns in A. If $p > n$, it follows that there will be at least $p - n > 0$ free variables, so our vectors must be linearly dependent. $\square$

Example. Suppose $u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $w = \begin{bmatrix} 5 \\ 60 \end{bmatrix}$. Then

$$A = \begin{bmatrix} 1 & 1 & 5 & 0 \\ 2 & 3 & 60 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 5 & 0 \\ 0 & 1 & 50 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -45 & 0 \\ 0 & 1 & 50 & 0 \end{bmatrix} = \texttt{RREF(A)}$$

so the pivot columns of A are 1 and 2, while x_3 is a free variable. Therefore u, v, w are linearly dependent.

In fact we have $x_1 u + x_2 v + x_3 w = 0$ if and only if $x_1 - 45x_3 = x_2 + 50x_3 = 0$.

Take $x_3 = 1$. Then $x_1 = 45$ and $x_2 = -50$, so $45u - 50v + w = 0$.