1 Last time: row reduction to (reduced) echelon form

The *leading entry* in a nonzero row of a matrix is the first nonzero entry from left going right. For example, the row $\begin{bmatrix} 0 & 0 & 7 & 0 & 5 \end{bmatrix}$ has leading entry 7, which occurs in the 3rd column.

Definition. A matrix with m rows and n columns is in *echelon form* if it has the following properties:

1. If a row is nonzero, then every row above it is also nonzero.
2. The leading entry in a nonzero row is in a column to the right of the leading entry in the row above.
3. If a row is nonzero, then every entry below its leading entry in the same column is zero.

For example,

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{*}$$

is in echelon form, but none of

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 5 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 2 & 0 & 0 \\ 1 & 3 & 5 & 0 \\ 0 & 0 & 4 & 5 \end{bmatrix}$$

is in echelon form.

Definition. A matrix is in *reduced echelon form* if

1. The matrix is in echelon form.
2. Each nonzero row has leading entry 1.
3. The leading 1 in each nonzero row is the only nonzero number in its column.

The matrix

$$\begin{bmatrix} 1 & 0 & -10/3 & 0 \\ 0 & 1 & 5/3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is in reduced echelon form and is row equivalent to the matrix (*).

Theorem. Each matrix A is row equivalent to exactly one matrix in reduced echelon form, which we denote $\texttt{RREF(A)}$.

The *row reduction algorithm* is a way of constructing $\texttt{RREF(A)}$ given A. This algorithm is something you should memorize and be able to perform quickly. We won't review the full definition again in this lecture, but let's do an example.

Example. Writing $\to$ to indicate a sequence of row operations, we have

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 1 \\ 1 & 3 & 9 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 2 & 8 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and the last matrix is the reduced echelon form of the first matrix.

We could actually have figured this out with less work: note that the first matrix is the augmented matrix of the linear system

$$a + b + c = 0$$
$$a + 2b + 2^2c = 1$$
$$a + 3b + 3^2c = 2$$

where I've used a, b, c as variables rather than x_1, x_2, x_3 as usual. Numbers a, b, c satisfying this system are the coefficients of a polynomial $f(x) = a + bx + cx^2$ with $f(1) = 0$, $f(2) = 1$, and $f(3) = 2$, i.e., a quadratic function whose graph is a parabola passing through the points $(x, y) = (1, 0), (2, 1), (2, 2)$. But these three points are all on the line $y = x - 1$, so we must have $f(x) = x - 1$ and $(a, b, c) = (-1, 1, 0)$ must be the unique solution to our system. This forces the reduced echelon form of our augmented matrix to be what we computed.

A *pivot column* of a matrix A is a column containing a leading 1 in `RREF(A)`.

If A is the augmented matrix of a linear system in variables $x_1, x_2, \dots, x_n$, then we say that x_i is a *basic variable* if i is a pivot column and that x_i is a *free variable* if i is not a pivot column.

To determine the basic and free variables of the system, we have to perform the row reduction algorithm to figure out what `RREF(A)` is first. Once we have this we can conclude that:

- The system has 0 solutions if the last column is a pivot column of A.
- The system has ∞ solutions if the last column is not a pivot column but there is ≥ 1 free variable.
- The system has 1 solution if there are no free variables, and the last column is not a pivot column.

Moreover, here's how you find all the solutions to the system: choose any values for the free variables, then solve the basic variables in terms of the free variables via the equations which make up the linear system corresponding to `RREF(A)`.

2 Vectors

Until we see vector spaces later in this course, the term *vector* will always refer to an ordered list of numbers in $\mathbb{R}$. A *vector* (sometimes to be called a *column vector*) is such a list oriented vertically; in other words, a matrix with one column:

$$\begin{bmatrix} 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 5 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \sqrt{7} \\ \sqrt{6} \end{bmatrix}.$$

We write a general column vector as

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

where each v_i is a real number. Two vectors u and v are equal if they have the same number of rows and the same entries in each row.

The sum of two vectors is

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}.$$

Note: $u + v = v + u$, but we can only add together vectors with the same number of rows.

If v is a vector and $c \in \mathbb{R}$ is a *scalar*, i.e., a real number, then we define

$$cv = c \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}.$$

We call the new vector cv the *scalar multiple* of v by c.

Example. To be concrete, we have

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

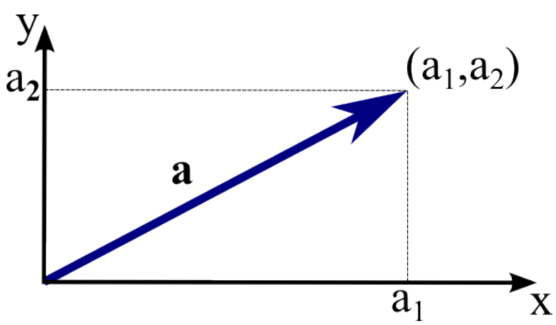
and

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

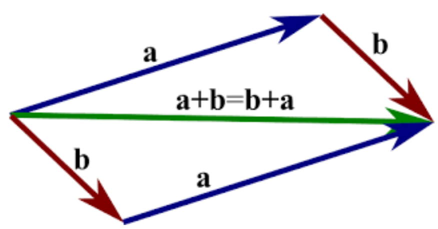
Define

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} - \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + (-1)\begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} -1 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \end{bmatrix}.$$

The *dimension* of a vector v is the number of its rows. We write $\mathbb{R}^n$ for the set of all n-dimensional vectors. Vectors in dimension 2, i.e., vectors $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2$, can be identified with arrows in the Cartesian plane from the origin the point $(x, y) = (a_1, a_2)$:



Proposition. 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:



Proof. Check that $\frac{a_2}{a_1} = \frac{(a_2+b_u)-b_2}{(a_1+b_1-b_1}$ and $\frac{b_2}{b_1} = \frac{(a_2+b_2)-a_2}{(a_1+b_1)-a_2}$. Hence the endpoint of $a+b$ is the intersection of the line through the endpoint of a parallel to b, and the line through the endpoint of b parallel to a. □

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.

Definition. Suppose $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ are distinct vectors and $c_1, c_2, \dots, c_p \in \mathbb{R}$ are *scalars*, i.e., numbers. The vector $y = c_1v_1 + c_2v_2 + \cdots + c_pv_p$ is called a *linear combination* of $v_1, v_2, \dots, v_p$. It is the linear combination with *coefficients* $c_1, c_2, \dots, c_p$.

Example. Suppose $a = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$ and $c = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$. Is c a linear combination of a and b?

If it were, we could find numbers $x_1, x_2 \in \mathbb{R}$ such that $x_1 a + x_2 b = c$, i.e., such that

$$\begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= -3. \end{aligned}$$

So to answer our question we need to determine if this linear system has a solution. To do this, use row reduction!

$$A = \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \to \texttt{RREF(A)} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

The pivot columns of A are thus 1 and 2: the last column is *not* a pivot column. Therefore our linear system is consistent, which means that c *is* a linear combination of a and b. (With what coefficients?)

We generalize this example with the following statement.

Proposition. A vector equation of the form $x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b$ where $x_1, x_2, \ldots, x_n$ are variables and $a_1, a_2, \ldots, a_n, b \in \mathbb{R}^m$ are vectors, has the same solutions as those for the linear system with augmented matrix

$$\begin{bmatrix} a_1 & a_2 & a_3 & \ldots & a_n & b \end{bmatrix}.$$

(This notation means: the matrix whose ith column is a_i and last column is b.) In particular, b is a linear combination of $a_1, a_2, \ldots, a_n$ if and only if this linear system is consistent.

Definition. The *span* of a vectors $v_1, v_2, \ldots, v_p \in \mathbb{R}^n$ is the set of all vectors $y \in \mathbb{R}^n$ which are linear combinations of $v_1, v_2, \ldots, v_p$. We typically denote the span of some set of vectors by

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

What does $\mathbb{R}$-span$\{v_1, v_2, \ldots, v_p\}$ look like?

We can visualize the span of the 0 vector as the single point consisting of just the origin. We imagine the span of a collection of vectors that all belong to the same line through the origin as that line.

In $\mathbb{R}^2$, if the span of $v_1, v_2, \ldots, v_p$ does not consist of a line, then the span is the whole plane. To see this, imagine we have two vectors $u, v \in \mathbb{R}^2$ which are not parallel. We can then get to any point in the plane by travelling some distance in the u direction, then some distance in the v direction. In other words, we can get any vector in $\mathbb{R}^2$ as the linear combination $au + bv$ for some scalars $a, b \in \mathbb{R}$. Draw a picture to illustrate this to yourself:

(Can you also explain this algebraically, in terms of the reduced echelon form of a matrix like the one in our last example?)