June 23

p1

Existence of Solutions

Recall that yesterday we defined $A\vec{x}$, for $A = [\vec{a}_1 \cdots \vec{a}_n]$ an $m \times n$ matrix with columns $\vec{a}_1, \cdots, \vec{a}_n \in \mathbb{R}^m$, and $\vec{x} \in \mathbb{R}^n$ to be the linear combination

$$A\vec{x} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n.$$

We described how a system $[A \,\vdots\, \vec{b}]$ being consistent is equivalent to whether $\vec{b}$ is in the span of the columns of A. That is, whether

$$\vec{b} \in \text{Span}\{\vec{a}_1, \cdots, \vec{a}_n\}.$$

One important question is whether *every* vector in $\mathbb{R}^m$ is in the span, that is, whether $\text{Span}\{\vec{a}_1, \cdots, \vec{a}_n\} = \mathbb{R}^m$. When this happens, we say the vectors *span $\mathbb{R}^m$*.

Note: This is a question of set equality. The definition for two sets being equal is that every element of one is an element of the other, and vice versa. That is, each is a subset of the other (not *proper* subset — we consider a set to be a subset of itself). Since $\text{Span}\{\vec{a}_1, \cdots, \vec{a}_n\} \subset \mathbb{R}^m$, basically by definition, the only remaining question is whether $\mathbb{R}^m \subset \text{Span}\{\vec{a}_1, \cdots, \vec{a}_n\}$

Theorem The following are equivalent: (A is $m \times n$, $\vec{a}_i$ is column i)

(a) For each $\vec{b} \in \mathbb{R}^m$, $A\vec{x} = \vec{b}$ has a solution $\vec{x} \in \mathbb{R}^n$.

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

(c) $\mathbb{R}^m = \text{Span}\{\vec{a}_1, \cdots, \vec{a}_n\}$.

(d) A has a pivot in every row.

This is a very important theorem, relating the row reduction calculation to more abstract judgments.

Warning: A is a <u>coefficient</u> matrix.

Let us see why they are equivalent. To do this, we show that they each imply one another.

(a) implies (b). Suppose $A\vec{x}=\vec{b}$ has a solution for each $\vec{b}$. Then since $A\vec{x}$ is a linear combination of the columns of A, each $\vec{b}$ is such a lin. comb.!
(b) implies (c). We need to show $\mathbb{R}^m \subset \text{Span}\{\vec{a}_1, \cdots, \vec{a}_n\}$, which we do by showing every element in $\mathbb{R}^m$ is in the span. Let $\vec{b} \in \mathbb{R}^m$, which by (b) is a lin. comb. of the columns of A, so it is in the span of $\vec{a}_1, \cdots, \vec{a}_n$.
(c) implies (a). For $\vec{b} \in \mathbb{R}^m$, by (c) it is in the span of the columns, so there are weights $x_1, \cdots, x_n$ with $\vec{b} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. So $A\vec{x}=\vec{b}$ has a sol'n.
Contrapositive (a) implies (d). If A does <u>not</u> have a pivot in every row, then we can find a $\vec{b}$ such that $[A \,\vdots\, \vec{b}\,]$ is <u>inconsistent</u>, which would mean $A\vec{x}=\vec{b}$ has <u>no</u> solution. To do this, compute $\text{rref}(A)$, then reverse each step and apply the reversed steps to $\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$ — this gives $\vec{b}$ (since $[A \,\vdots\, \vec{b}\,]$ has a pivot in the last column).
(d) implies (a). If A has a pivot in every row, then $[A \,\vdots\, \vec{b}\,]$ does not have a pivot in the last column. Hence, $A\vec{x}=\vec{b}$ has a solution for each $\vec{b} \in \mathbb{R}^m$.

Altogether,
$$\begin{tikzcd} (b) \arrow[d, Rightarrow] & & \\ (c) \arrow[r, Rightarrow] & (a) \arrow[ul, Rightarrow] \arrow[r, Rightarrow, bend left] & (d) \arrow[l, Rightarrow, bend left] \end{tikzcd}$$
, so they are logically equivalent

This theorem is a logical equivalence. Not only is it that if any one is true then so are the rest, but if any one is false, then so are the rest!

example

1) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ doesn't have a pivot in every row, so $A\vec{x} = \vec{b}$ isn't always solvable,

2) If $\vec{u}, \vec{v} \in \mathbb{R}^3$, $A = [\vec{u} \;\; \vec{v}]$ has at most two pivots (columns are limiting), so $\text{Span}\{\vec{u}, \vec{v}\} \neq \mathbb{R}^3$.

Methods for computing $A\vec{x}$

While $A\vec{x}$ is a linear combination of the columns of A, there is a dual method which is easier in practice, which comes from performing the scalar multiplication and vector addition all at once.

ex

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 6 \\ 7 \end{bmatrix} = 6 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + 7 \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 6(2) + 7(3) \\ 6(4) + 7(5) \end{bmatrix}$$

more generally,

$$\begin{bmatrix} \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \cdots & b_n \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \vdots \\ x_1 b_1 + \cdots + x_n b_n \\ \vdots \end{bmatrix}$$

Algorithm: for each row, match up corresponding entries in $\vec{x}$, and take the sum of the products.

Bonus: $\vec{y} = A\vec{x}$ has $y_i = \sum_{j=1}^{n} A_{ij} x_j$, A_{ij} is entry in row i, column j.

Algebraic properties of matrices

Another step of our journey for algebrization of systems of equations is the properties of matrix-vector products. For A an $m \times n$ matrix and $\vec{u}, \vec{v} \in \mathbb{R}^n$, $c \in \mathbb{R}$, we have

(i) $A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$; and

(ii) $A(c\vec{u}) = c(A\vec{u})$.

These are <u>very</u> important. They say addition and scalar products commute with A — together, these imply the following for linear combinations:

$$A(c_1\vec{v_1} + \cdots + c_n\vec{v_n}) = c_1(A\vec{v_1}) + \cdots + c_n(A\vec{v_n})$$

The linear combination of $\vec{v_1}, \cdots, \vec{v_n} \in \mathbb{R}^n$ gets tranformed into one of $A\vec{v_1}, \cdots, A\vec{v_n} \in \mathbb{R}^m$.

Proof of (i): Let's just do it for $\mathbb{R}^2$.

$$\begin{aligned} A(\vec{u}+\vec{v}) = A\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix} &= (u_1+v_1)\vec{a_1} + (u_2+v_2)\vec{a_2} \\ &= u_1\vec{a_1} + v_1\vec{a_1} \\ &\quad + u_2\vec{a_2} + v_2\vec{a_2} \\ &= A\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + A\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = A\vec{u} + A\vec{v}. \end{aligned}$$

Proof of (ii): Again, only for $\mathbb{R}^2$.

$$\begin{aligned} A(c\vec{u}) = A\begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix} &= (cu_1)\vec{a_1} + (cu_2)\vec{a_2} \\ &= c(u_1\vec{a_1} + u_2\vec{a_2}) \\ &= c(A\vec{u}). \end{aligned}$$

Homogeneous linear systems

One consequence of the previous fact is the following:

> If $\vec{x}_h, \vec{x}_p \in \mathbb{R}^n$ are solutions to $A\vec{x}_h = \vec{0}$ and $A\vec{x}_p = \vec{b}$, then $A(\vec{x}_p + \vec{x}_h) = A\vec{x}_p + A\vec{x}_h = \vec{b} + \vec{0} = \vec{b}$.
> So $A\vec{x} = \vec{b}$ has $\vec{x}_p + \vec{x}_h$ as a solution.
>
> Furthermore, if $\vec{x}_1$ and $\vec{x}_2$ are two solutions to $A\vec{x} = \vec{b}$, then $A(\vec{x}_1 - \vec{x}_2) = A\vec{x}_1 - A\vec{x}_2 = \vec{b} - \vec{b} = \vec{0}$, so $\vec{x}_1 - \vec{x}_2$ is a solution to $A\vec{x} = \vec{0}$.

These suggest that the set of solutions to $A\vec{x} = \vec{b}$ (if it is nonempty) is completely governed by the solutions to $A\vec{x} = \vec{0}$!

A homogeneous system is one which is of the form $A\vec{x} = \vec{0}$. (With A $m \times n$ and $\vec{0} \in \mathbb{R}^m$).

A homogeneous system always has the solution $\vec{x} = \vec{0}$, since $A\vec{0} = \vec{0}$. This is the trivial solution. Other solutions (if they exist) are nontrivial solutions.

Thm $A\vec{x} = \vec{0}$ has nontrivial solutions if and only if A has a free (i.e., non-pivot) column.

Proof If A has no free columns, $A\vec{x} = \vec{0}$ has at most one solution, which we know $\vec{0}$ to be one.
If A has a free column, say column i, take a solution with $x_i = 1$. Then $A\vec{x} = \vec{0}$ and $\vec{x} \neq \vec{0}$.

"P if and only if Q" means both "if P then Q" and "if Q then P"
We show not Q ⇒ not P and Q ⇒ P
(logical equivalence)

example If A is 2×3, $A\vec{x} = \vec{0}$ has a nontrivial solution. (At most 2 pivots < 3 columns) example $A=\begin{bmatrix}1&0&2&0\\0&1&1&0\end{bmatrix}$

$$\begin{array}{cccc} & & \uparrow & \uparrow \end{array}$$

free columns

so $\begin{cases} x_1 = -2x_3 \\ x_2 = -x_3 \\ x_3 \text{ free} \\ x_4 \text{ free} \end{cases}$

Let us write this as a vector:

$$\vec{x} = \begin{bmatrix}-2x_3\\-x_3\\x_3\\x_4\end{bmatrix} = x_3\begin{bmatrix}-2\\-1\\1\\0\end{bmatrix} + x_4\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

"parametric vector form"

So the solution set is

$$\text{Span}\left\{\begin{bmatrix}-2\\-1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}$$

In fact, solution sets for homogeneous systems are <u>always</u> the span of some number of vectors, one per free variable.

and vice versa!

Method:
- Compute rref(A)
- write a vector down for each free variable
- put a 1 in entry i for the vector associated with free variable i (0 for other free variables)
- solve for the remaining entries (they are just negatives of entries in column i for free var, i)

example $A = \begin{bmatrix} 1 & 2 & 0 & 0 & 3 \\ & & 1 & 0 & 4 \\ & & & 1 & 5 \end{bmatrix}$

solution set $= \text{Span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ -4 \\ -5 \\ 1 \end{bmatrix} \right\}$ for free variables

A **nonhomogeneous system** is just a system $A\vec{x} = \vec{b}$. It's possible $\vec{b} = \vec{0}$, so this is a useless word. Suppose we have a solution $\vec{p} \in \mathbb{R}^n$. Adding this to homogeneous solutions yields all solutions to $A\vec{x} = \vec{b}$, as previously discussed.

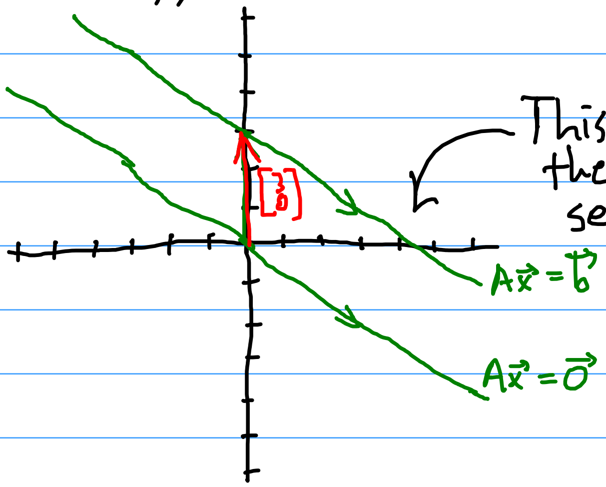
example $A = \begin{bmatrix} 2 & 3 \end{bmatrix}$ $\quad \vec{b} = \begin{bmatrix} 6 \end{bmatrix}$

solution set to $A\vec{x} = \vec{0}$ is $\text{Span}\left\{ \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$

and $A \begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \end{bmatrix}$, so the solutions to this are of the form

$$\vec{x} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \text{ for } t \in \mathbb{R}$$

Geometrically,



This is a translation of the homogeneous solution set by a particular solution.

In practice, there is no need to use the homogeneous system if $\vec{b}$ is fixed, if the goal is a parametric vector form of the solution, especially since you need a particular solution anyway.

example For above,

$$\left[\begin{array}{cc|c} 2 & 3 & 6 \end{array}\right]$$

$$\begin{array}{c}\uparrow\\ \text{free}\end{array}$$

so $\vec{x} = \begin{bmatrix} 3 - \frac{3}{2}x_2 \\ x_2 \end{bmatrix}$

$$= \begin{bmatrix} 3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3/2 \\ 1 \end{bmatrix}.$$

def The nullspace of A is the set of solutions to $A\vec{x} = \vec{0}$, denoted null(A). That is,

$$\text{null}(A) = \{ \vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0} \}$$

where A is $m \times n$.

A collection of vectors $\vec{a}_1, \ldots, \vec{a}_n \in \mathbb{R}^m$ are independent if $\text{null}([\vec{a}_1 \cdots \vec{a}_n]) = \{\vec{0}\}$. That is, if the only time

$$c_1\vec{a}_1 + \cdots + c_n\vec{a}_n = \vec{0}$$

is when $c_1 = c_2 = \cdots = c_n = 0$.

Otherwise, the vectors are dependent. The linear combination which demonstrates that the vectors are dependent is a dependence (that is, some scalars $c_1, \ldots, c_n$, not all zero, with $c_1\vec{a}_1 + \cdots + c_n\vec{a}_n = \vec{0}$. Or: a nontrivial linear combination for $\vec{0}$).

example $\vec{u},\vec{v},\vec{w},\vec{0}$ is linearly dependent since $0\vec{u}+0\vec{v}+0\vec{w}+1\cdot\vec{0}=\vec{0}$

example If $c_1\vec{u}_1+c_2\vec{u}_2+c_3\vec{u}_3=\vec{0}$, with $c_1\neq 0$,

$$\vec{u}_1=\frac{-c_2}{c_1}\vec{u}_2-\frac{c_3}{u_1}\vec{u}_3$$

That is, $\vec{u}_1$ is a linear combination of $\vec{u}_2$ and $\vec{u}_3$.

Important principle!

example For two vectors, $c_1\vec{u}_1+c_2\vec{u}_2=0$, $c_1\neq 0$, then $\vec{u}_1=\frac{-c_2}{c_1}\vec{u}_2$. One is multiple of other.

example If $\vec{a}_1,\dots,\vec{a}_n\in\mathbb{R}^m$, with $m<n$, then they are dependent.

$\left[\vec{a}_1 \cdots \vec{a}_n\right]$ has more columns than rows, hence has free variables.

Thus: independent implies $m\geq n$! (though $m\geq n$ might still be dependent, for instance $\begin{bmatrix}1\\2\\3\end{bmatrix},\begin{bmatrix}2\\4\\6\end{bmatrix}$.)