Linear independence (reprise)

Yesterday, we defined vectors $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^m$ to be linearly independent if the homogeneous system

$$[\vec{a}_1 \ \vec{a}_2 \ \cdots \ \vec{a}_n]\vec{x} = \vec{0}$$

has only the trivial solution ($\vec{x} = \vec{0}$). Otherwise, if there is a nontrivial solution, we call the vectors linearly dependent.

The book gives a slightly different definition, but I like this one because it suggests how to determine whether vectors are independent: compute $\mathrm{rref}([\vec{a}_1 \cdots \vec{a}_n])$. Using our knowledge of homogeneous systems, we know that a pivot in every column implies independent, where a free column implies dependent.

A dependence is a nontrivial solution $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, written as

$$x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{0}.$$

That is, writing $\vec{0}$ as a linear combination of the vectors, not all of whose weights are 0.

example If $A\vec{x} = \vec{b}$ has a solution $\vec{x} \in \mathbb{R}^n$, then there is a dependence

$$x_1\vec{a}_1 + \cdots + x_n\vec{a}_n + (-1)\vec{b} = \vec{0}$$

between the vectors $\vec{a}_1, \dots, \vec{a}_n, \vec{b}$. (The last weight, -1, is not zero.)

Again, let us make a set to help capture this idea.

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

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

Then, $\vec{a}_1, \cdots, \vec{a}_n$ are independent if and only if

$$\text{null}([\vec{a}_1 \cdots \vec{a}_n]) = \{\vec{0}\}$$

(the only solution is the trivial solution).

We saw yesterday how the solution set for a homogeneous system is a span of vectors (one vector per free variable), so $\text{null}(A)$ is also a span.

example $\text{null}\left(\begin{bmatrix} 1 & 2 & 6 \\ 1 & 3 & 3 \end{bmatrix}\right)$.

$$\begin{bmatrix} 1 & 2 & 6 \\ 1 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 6 \\ 0 & 1 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 12 \\ 0 & 1 & -3 \end{bmatrix}$$

so nullspace is $\text{Span}\left\{ \begin{bmatrix} -12 \\ 3 \\ 1 \end{bmatrix} \right\}$.

Linear transformations

We have defined multiplication by a matrix. To a modern mathematician, this suggests $A\vec{x}$ is a function of $\vec{x}$. Traditional terminology is to call such a function a transformation.

def A *transformation* (or *function* or *mapping*) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule assigning a single vector $T(\vec{x})$ in $\mathbb{R}^m$ to every vector $\vec{x} \in \mathbb{R}^n$.

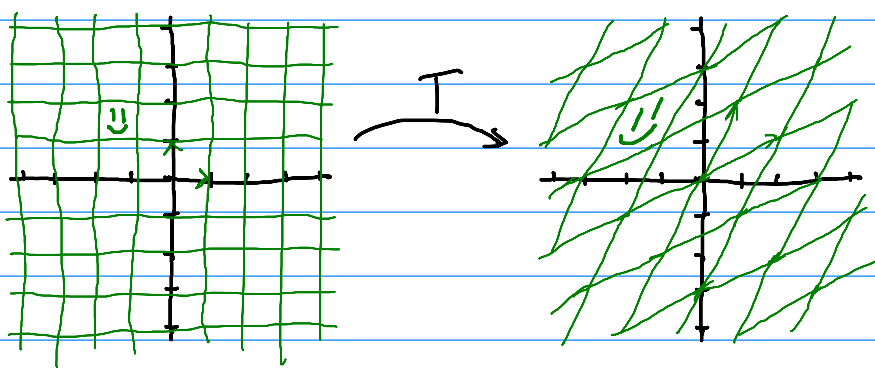
Notation is $T : \underbrace{\mathbb{R}^n}_{\text{domain}} \longrightarrow \underbrace{\mathbb{R}^m}_{\text{codomain}}$

The vector $T(\vec{x})$ is called the image of $\vec{x}$ under the action of T. The set of all images is called the range or image of T, denoted $T(\mathbb{R}^n) = \{T(\vec{x}) \mid \vec{x} \in \mathbb{R}^n\}$ or $\mathrm{im}(T)$.

A matrix transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is a transformation of the form $T(\vec{x}) = A\vec{x}$, for some $m \times n$ matrix A. Another notation: $T = (\vec{x} \mapsto A\vec{x})$ ("T equals the rule $\vec{x}$ maps to $A\vec{x}$").

Now, the question of whether $[A \,\vdots\, \vec{b}]$ is consistent becomes the question of whether $\vec{b}$ is an image of $T(\vec{x}) = A\vec{x}$. Also, whether the image of T equals $\mathbb{R}^m$ is whether there is a pivot in every row of A!

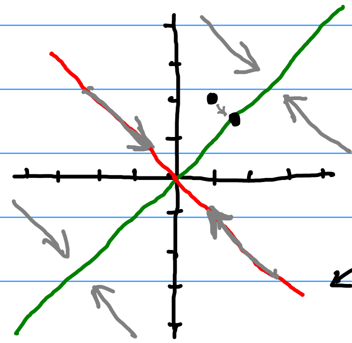
example $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\vec{x} \mapsto \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \vec{x}$.
We can think of this transforming the plane:



$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = A\begin{bmatrix} x \\ y \end{bmatrix} = A\left(x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = xA\begin{bmatrix} 1 \\ 0 \end{bmatrix} + yA\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$= x\begin{bmatrix} 2 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$ (a linear combination of columns, hence the grid diagram).

example $T(\vec{x}) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \vec{x}$

$T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 3/2 \\ 3/2 \end{bmatrix}$



← image of $T = \operatorname{Span}\left\{\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}\right\}$

← solutions to $T(\vec{x}) = \vec{0}$ $= \operatorname{null}\left(\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}\right) = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$

projection onto the line $y = x$.

More generally, a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is a transformation with the following two additional properties: for $\vec{u}, \vec{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$

(i) $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$

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

Also known as a linear map

Notice that matrix transformations satisfy these because $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$ and $A(c\vec{u}) = c(A\vec{u})$, so we know a bunch of linear transformations!

These two properties conspire to produce a much more useful and important property:

$$T(c_1\vec{u}_1 + \cdots + c_k\vec{v}_k) = c_1T(\vec{u}_1) + \cdots + c_kT(\vec{u}_k)$$

* May as well be the definition though (i) and (ii) can be easier to verify for a given T.

That is: T of a linear combination of vectors is a linear combination of T of the vectors.

We can show certain facts about T without knowing anything other than it is linear.

For instance, what is the image of $\vec{0}$ under the action of T?

$$T(\vec{0}) = T(0\cdot\vec{0}) \overset{(ii)}{=} 0\cdot T(\vec{0}) = \vec{0}$$

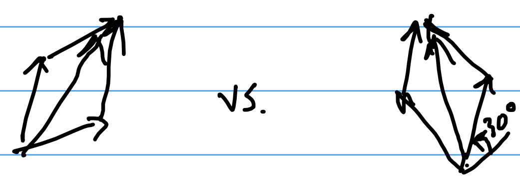
So, if T is some mapping with $T(\vec{0}) \neq \vec{0}$, it is <u>not</u> a linear mapping. <u>ex</u> translation of $\mathbb{R}^2$ has $T(\vec{0}) \neq \vec{0}$.

<u>example</u> $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\vec{x}) = r\vec{x}$, with $r \in \mathbb{R}$ a constant, is a linear map.

(i) $T(\vec{u}+\vec{v}) = r(\vec{u}+\vec{v}) = r\vec{u} + r\vec{v} = T(\vec{u}) + T(\vec{v})$
(ii) $T(c\vec{u}) = r(c\vec{u}) = (rc)\vec{u} = (cr)\vec{u} = c(r\vec{u}) = cT(\vec{u})$.
It satisfies both properties, so it is linear.

T scales the plane by r.

<u>example</u> $R: \mathbb{R}^2 \to \mathbb{R}^2$ defined by rotating a vector $30°$ CCW about the origin is a linear transformation. The properties seem true: rotate your head $30°$ CW; head-to-tail vector addition can happen before or after rotation. This is not a real proof.



Every linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is actually a matrix transformation! (You may wonder why we bother with linear transformations at all, then. (1) Later we generalize $\mathbb{R}^n$ to a "vector space", where matrices no longer [directly] make sense; (2) Sometimes it is easier to describe a transformation without a matrix.)

The principle is the following (at least in $\mathbb{R}^2$):

$$
\begin{aligned}
T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) &= T\left(x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\
&= x\,T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + y\,T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\
&= \underbrace{\begin{bmatrix} T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) & T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \end{bmatrix}}_{\text{This is the matrix which shows } T \text{ is a matrix transformation!}} \begin{bmatrix} x \\ y \end{bmatrix}.
\end{aligned}
$$

We call this the standard matrix of T, denoted $[T]$. So, $T(\vec{x}) = [T]\vec{x}$.

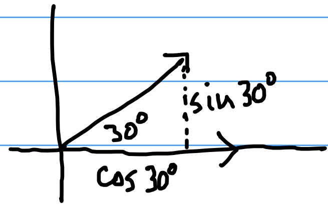
One special set of vectors in $\mathbb{R}^n$ is the standard basis $\vec{e}_1, \ldots, \vec{e}_n$ with

$$
\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \cdots \quad \vec{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}
$$

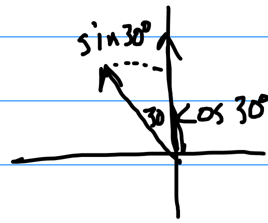
(and $[\vec{e}_1 \ \cdots \ \vec{e}_n]$ is called the identity matrix I_n)

More generally, $[T] = [T(\vec{e}_1) \ T(\vec{e}_2) \ \cdots \ T(\vec{e}_n)]$

example If T is once again rotation by $30°$ CCW,



$$T(\vec{e}_1) = T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}\sqrt{3}/2\\1/2\end{bmatrix}$$



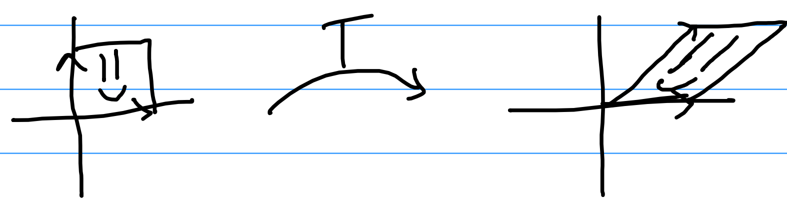
$$T(\vec{e}_2) = T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-1/2\\\sqrt{3}/2\end{bmatrix}$$

Thus, T can be implemented as

$$T(\vec{x}) = \begin{bmatrix}\sqrt{3}/2 & -1/2\\1/2 & \sqrt{3}/2\end{bmatrix}\vec{x}$$

example We can figure out what T is from its matrix (sometimes). For $T(\vec{x}) = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}\vec{x}$,

$$T(\vec{e}_1) = \begin{bmatrix}1\\0\end{bmatrix} \quad \text{and} \quad T(\vec{e}_2) = \begin{bmatrix}1\\1\end{bmatrix}$$



it is a shear transformation.

New words to learn for a $T: \mathbb{R}^n \to \mathbb{R}^m$.

- T is **onto** if $T(\mathbb{R}^n) = \mathbb{R}^m$. (each $\vec{b} \in \mathbb{R}^m$ is an image)
- T is **one-to-one** if each $\vec{b} \in \mathbb{R}^m$ is an image of at most one $\vec{x} \in \mathbb{R}^n$. ($T(\vec{x}) = \vec{b}$ has no more than one solution)

* **Thm** T one-to-one $\Leftrightarrow$ $[T]$ has pivot in every column.

Thm T onto $\Leftrightarrow$ $[T]$ has pivot in every row.