Recall, that for a function $f : X \to Y$, the image is defined as

$$Im(f) = \{ b \in Y : b = f(x) \text{ for some } x \in X \}$$

Now, if a function is a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$

$$Im(T) = \{ \vec{b} \in \mathbb{R}^n : \text{system } A_T x = \vec{b} \text{ has a solution} \}$$

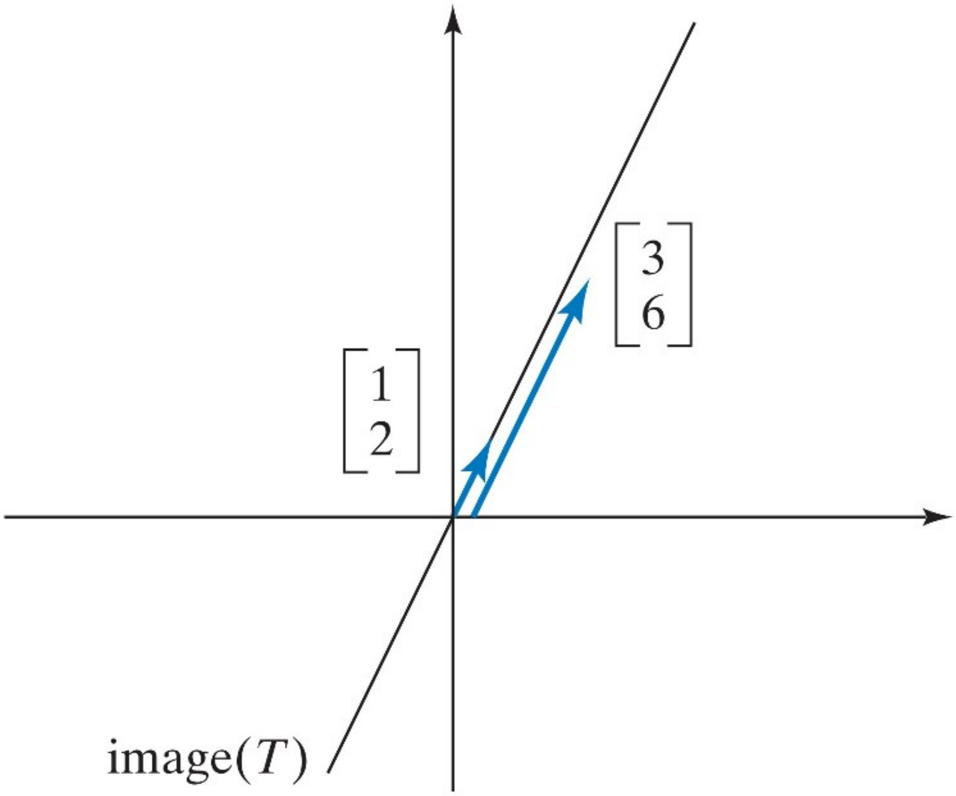
Example: 1) Say $T(\vec{x}) = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $T : \mathbb{R}^2 \to \mathbb{R}^2$

Image consists of all values of T, that is, all vectors of the form

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ 6 \end{pmatrix} =$$

$$= x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3x_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = (x_1 + 3x_2) \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad x_i \in \mathbb{R}.$$

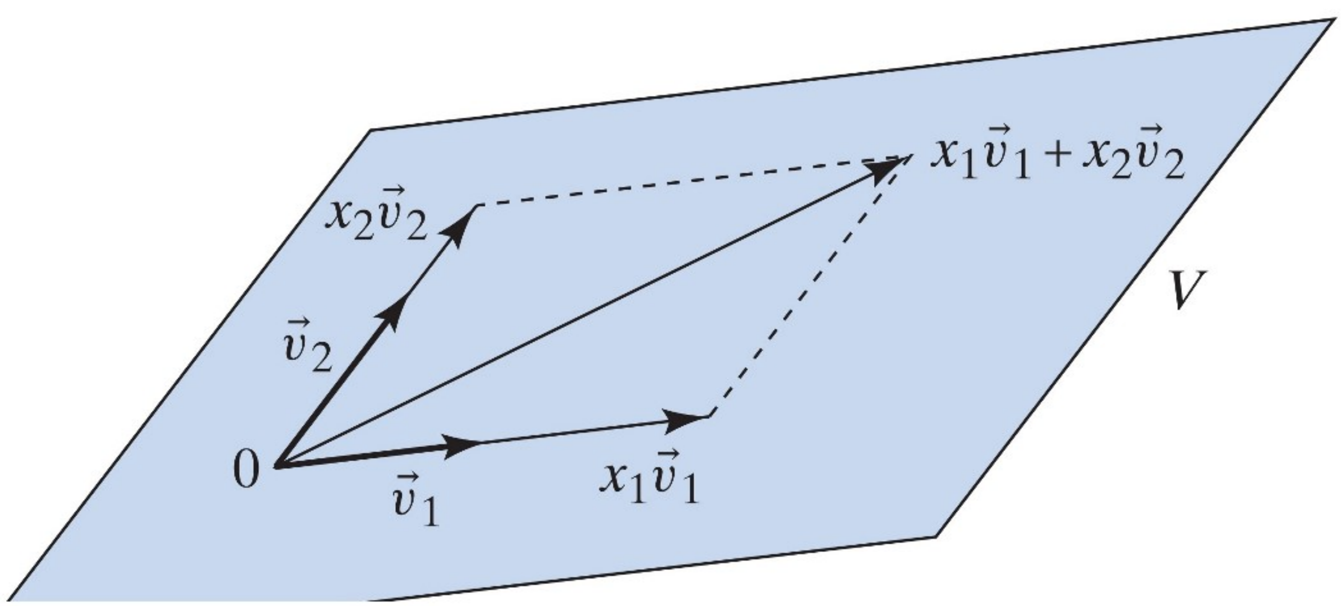
So, the image is just scalar multiples of vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

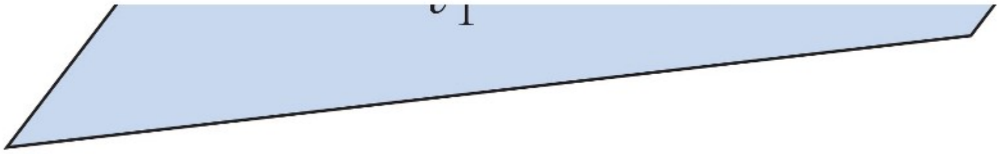


$\begin{pmatrix}3\\6\end{pmatrix}$ and $\begin{pmatrix}1\\2\end{pmatrix}$ are parallel.

2) Say $T: \mathbb{R}^2 \to \mathbb{R}^3$ is given by $T\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}1&1\\1&2\\1&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$.

$T\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}1&1\\1&2\\1&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1\underbrace{\begin{pmatrix}1\\1\\1\end{pmatrix}}_{v_1} + x_2\underbrace{\begin{pmatrix}1\\2\\3\end{pmatrix}}_{v_2}$. So, the image V of T consists of all linear combinations of column vectors of a matrix of T.





This motivates the definition

Span

Consider the vectors $\vec{v}_1, \ldots, \vec{v}_m$ in $\mathbb{R}^n$. The set of all linear combinations $c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$ of the vectors $\vec{v}_1, \ldots, \vec{v}_m$ is called their *span*:

$$\operatorname{span}(\vec{v}_1, \ldots, \vec{v}_m) = \{c_1\vec{v}_1 + \cdots + c_m\vec{v}_m : c_1, \ldots, c_m \text{ in } \mathbb{R}\}.$$

Theorem: The image of the linear transformation $T(\vec{x}) = A\vec{x}$ is the span of column vectors of A. (sometimes called a „column space" of A)

Proof:

$$T(\vec{x}) = A\vec{x} = \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = x_1\vec{v}_1 + \cdots + x_m\vec{v}_m$$

$\square$

Some properties of the image

The image of a linear transformation T (from $\mathbb{R}^m$ to $\mathbb{R}^n$) has the following properties:

- **a.** The zero vector $\vec{0}$ in $\mathbb{R}^n$ is in the image of T.
- **b.** The image of T is *closed under addition*: If $\vec{v}_1$ and $\vec{v}_2$ are in the image of T, then so is $\vec{v}_1 + \vec{v}_2$.
- **c.** The image of T is *closed under scalar multiplication*: If $\vec{v}$ is in the image of T and k is an arbitrary scalar, then $k\vec{v}$ is in the image of T as well.

Proof: a) $T(\vec{0}) = A\vec{0} = \vec{0}$

b) Say $\vec{w}_1, \vec{w}_2 \in \operatorname{Im}(T)$. This means, $\vec{w}_1 = T(\vec{v}_1)$, $\vec{w}_2 = T(\vec{v}_2)$

b) Say $\vec{w}_1, \vec{w}_2 \in \mathrm{Im}(T)$. This means, $\vec{w}_1 = T(\vec{v}_1)$, $\vec{w}_2 = T(\vec{v}_2)$ for some $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^m$. Then, $\vec{w}_1 + \vec{w}_2 = T(\vec{v}_1) + T(\vec{v}_2) =$

$= T(\vec{v}_1 + \vec{v}_2) \in \mathrm{Im}(T)$

c) Say $\vec{w} \in \mathrm{Im}(T) \Rightarrow \vec{w} = T(\vec{v})$ for some $\vec{v} \in \mathbb{R}^m \Rightarrow$

$k\vec{w} = kT(\vec{v}) = T(k\vec{v}) \in \mathrm{Im}(T)$. □

It follows that $\mathrm{Im}(T)$ is closed under all linear combinations.

Kernel

For a linear transformation $T: \mathbb{R}^m \to \mathbb{R}^n$, vectors mapped to zero are of special interest.

Kernel (also called a null space)

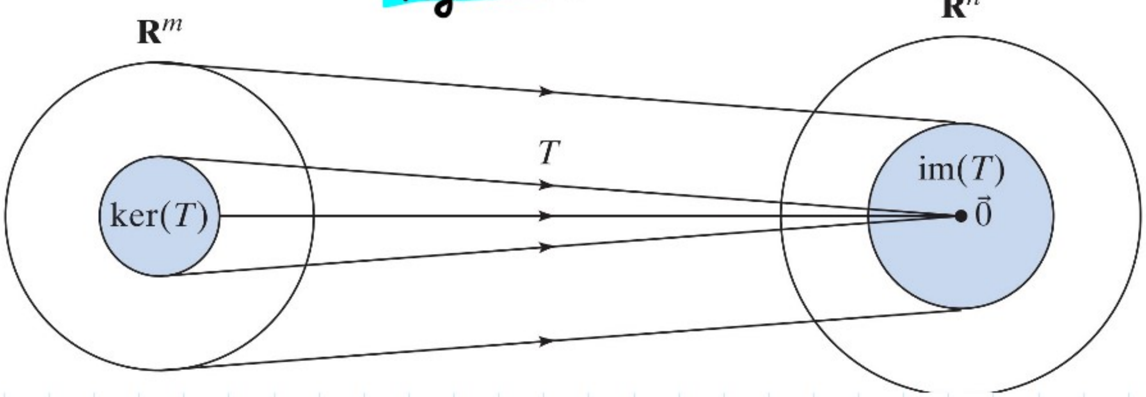
The *kernel* of a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^m$ to $\mathbb{R}^n$ consists of all zeros of the transformation, that is, the solutions of the equation $T(\vec{x}) = A\vec{x} = \vec{0}$. See Figure 10, where we show the kernel along with the image.

In other words, the kernel of T is the solution set of the linear system

$$A\vec{x} = \vec{0}.$$

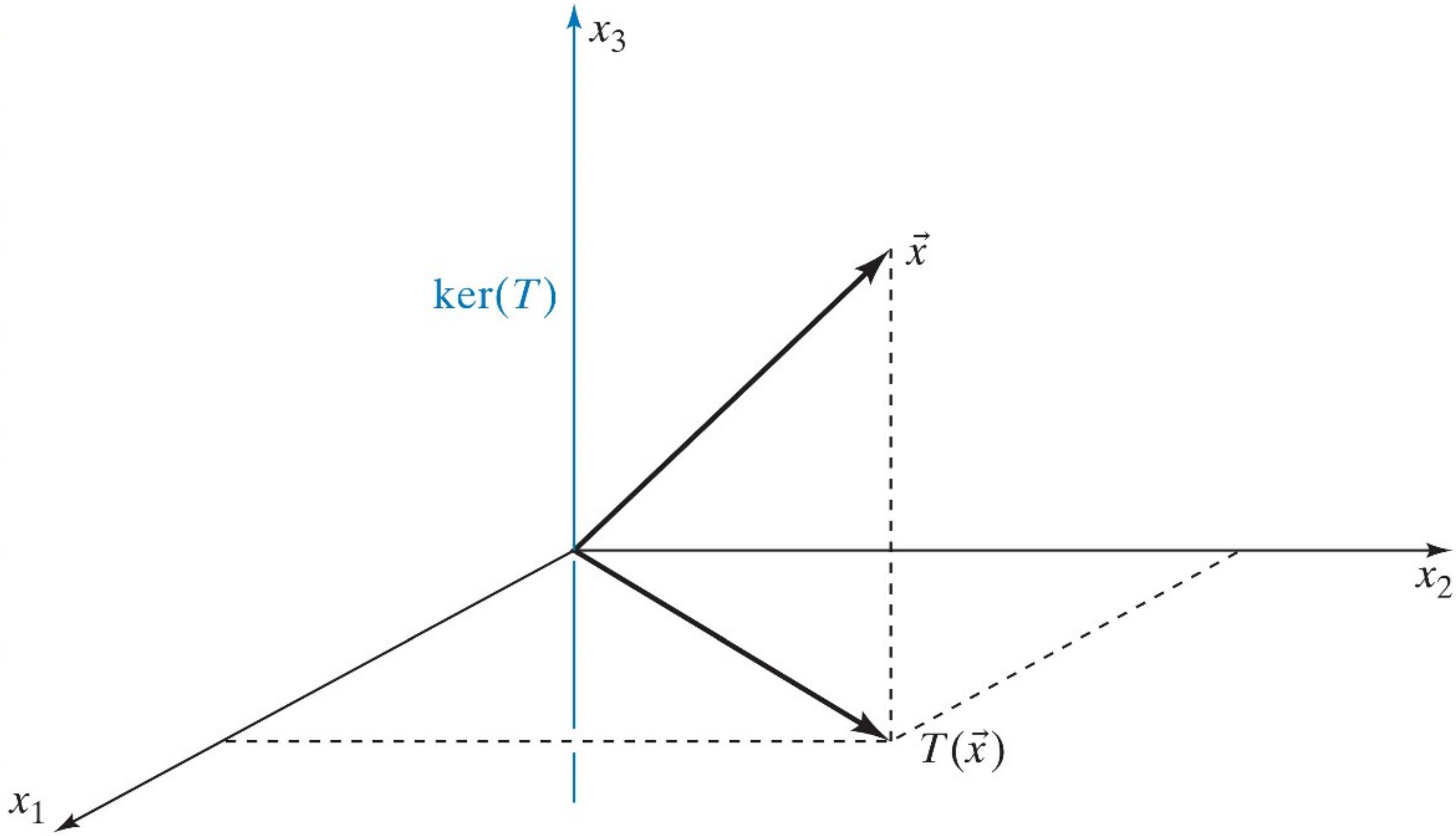
We denote the kernel of T by $\ker(T)$ or $\ker(A)$.

Figure 10



Example: 1) $T: \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation

Example: 1) $T: \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation given by orthogonal projection to x_1-x_2 plane.



$$\ker(T) = \{ \vec{x} \in \mathbb{R}^3 : \; T(\vec{x}) = 0 \}$$

which are vectors that are mapped to zero by orthogonal projection to x_1-x_2 plane. These are vectors on axis X_3, that is, span of vector $\vec{e}_3$.

Example: 1) Find $\ker(T)$ for $T(\vec{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \vec{x}$, $T: \mathbb{R}^3 \to \mathbb{R}^2$.

We need to solve $T(\vec{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$ to find vectors mapped to zero. For solving,

$$\mathrm{rref}\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \end{array} \right) = \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \end{array} \right)$$

$$\operatorname{rref}\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right)$$

$$\begin{cases} x_1 - x_3 = 0 \\ x_2 + 2x_3 = 0 \end{cases} \Rightarrow \begin{cases} x_1 = x_3 \\ x_2 = -2x_3 \end{cases}, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} t \\ -2t \\ t \end{pmatrix} = t\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

with $t \in \mathbb{R}$. So $\ker(T) = \operatorname{span}\left(\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\right)$.

In general, $T: \mathbb{R}^m \to \mathbb{R}^n$, $\ker(T)$ is infinite if $m > n$.

2) $T(\vec{x}) = A\vec{x}$ where $A = \begin{pmatrix} 1 & 2 & 2 & -5 & -6 \\ -1 & -2 & -1 & 1 & -1 \\ 4 & 8 & 5 & -8 & 9 \\ 3 & 6 & 1 & 5 & -7 \end{pmatrix}$.

To find $\ker(T)$, we need to solve $A\vec{x} = 0$.

$$\operatorname{rref}(A|0) = \left(\begin{array}{ccccc|c} 1 & 2 & 0 & 3 & -4 & 0 \\ 0 & 0 & 1 & -4 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right), \text{ that is,}$$

$$\begin{cases} x_1 = -2x_2 - 3x_4 + 4x_5 \\ x_3 = 4x_4 - 5x_5 \end{cases}$$ or in a vector notation

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2s - 3t + 4r \\ s \\ 4t - 5r \\ t \\ r \end{pmatrix} = s\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{pmatrix} + r\begin{pmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{pmatrix}$$

So, $\ker(T) = \operatorname{span}\left(\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{pmatrix}\right)$.

Some properties of the kernel As for image.

Consider a linear transformation T from $\mathbb{R}^m$ to $\mathbb{R}^n$.

a. The zero vector $\vec{0}$ in $\mathbb{R}^m$ is in the kernel of T.

b. The kernel is closed under addition.

c. The kernel is closed under scalar multiplication.

Proof: exercize.

Say now A is $n \times n$ matrix and A is invertible. Then $A\vec{x} = 0$ has a unique solution $x = 0$. Thus $\ker(A) = \{0\}$.

Various characterizations of invertible matrices

For an $n \times n$ matrix A, the following statements are equivalent; that is, for a given A, they are either all true or all false.

i. A is invertible.

ii. The linear system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$, for all $\vec{b}$ in $\mathbb{R}^n$.

iii. $\text{rref}(A) = I_n$.

iv. $\text{rank}(A) = n$.

v. $\text{im}(A) = \mathbb{R}^n$.

vi. $\ker(A) = \{\vec{0}\}$.

vii. $T(\vec{x}) = Ax$ is a bijection.

Subspaces of $\mathbb{R}^n$

A subset W of the vector space $\mathbb{R}^n$ is called a (linear) *subspace of* $\mathbb{R}^n$ if it has the following three properties:

a. W contains the zero vector in $\mathbb{R}^n$.

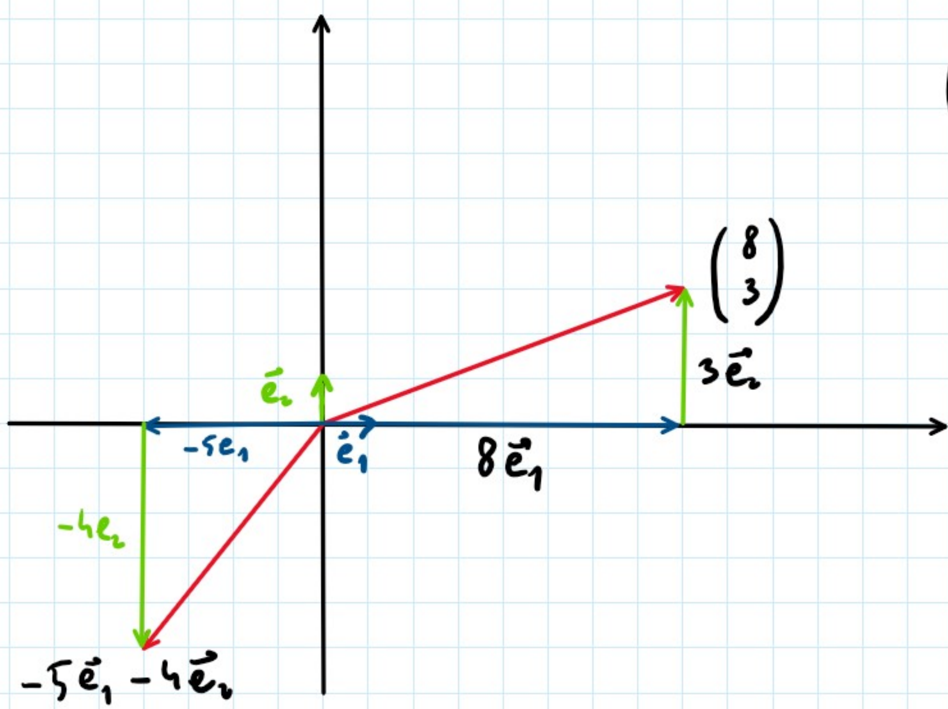
b. W is closed under addition: If $\vec{w}_1$ and $\vec{w}_2$ are both in W, then so is $\vec{w}_1 + \vec{w}_2$.

c. W is closed under scalar multiplication: If $\vec{w}$ is in W and k is an arbitrary scalar, then $k\vec{w}$ is in W.

Properties (b) and (c) together mean that W is *closed under linear combinations*: If vectors $\vec{w}_1, \ldots, \vec{w}_m$ are in W and $k_1, \ldots, k_m$ are scalars, then the linear combination $k_1\vec{w}_1 + \cdots + k_m\vec{w}_m$ is in W as well.

Basis

Lets think of the coordinates for a vector in a different way.



$\begin{pmatrix} 8 \\ 3 \end{pmatrix} = 8 \cdot \vec{e}_1 + 3\vec{e}_2$. So, coordinates are just transformations that scale vectors $\vec{e}_1, \vec{e}_2$. When we think about vector coordinates as scalars, vectors $\{\vec{e}_1, \vec{e}_2\}$ are the ones they scale.

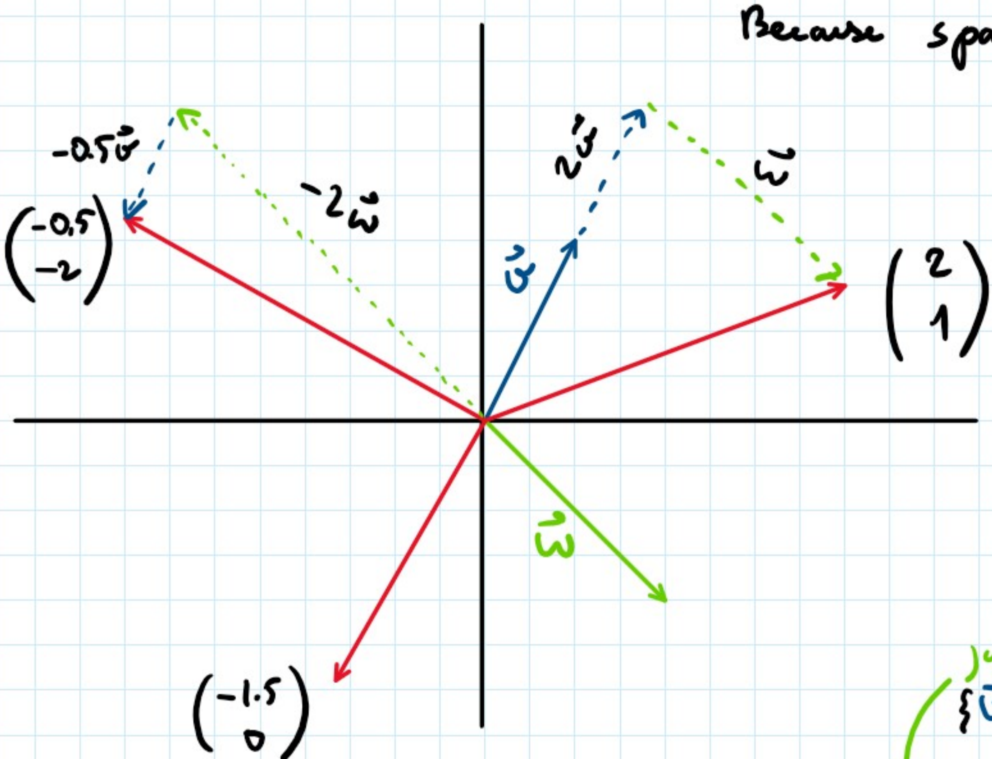
So, we are basically framing our coordinate system in terms of vectors $\vec{e}_1$ and $\vec{e}_2$.

Question: What if we choose different vectors instead of $\vec{e}_1, \vec{e}_2$?

of vectors $\vec{e}_1$ and $\vec{e}_2$.

Question: What if we choose different vectors instead of $\vec{e}_1, \vec{e}_2$?
Say, choose $\vec{v}, \vec{w}$. We get a different, „legitimate" coordinate system:
Because $\text{span}(\vec{v}, \vec{w}) = \mathbb{R}^2$.

why? exercise.



just a name for $\{\vec{v}, \vec{w}\}$ at the moment.

Take home: Every time we describe a vector numerically, it depends on an implicit choice of „basis" vector we are using.

For „most" pairs of vectors $\vec{u}_1, \vec{u}_2$, $\text{span}(\vec{u}_1, \vec{u}_2) = \mathbb{R}^2$, in other words, we can „reach" every vector in $\mathbb{R}^2$ with linear combinations of $\vec{u}_1, \vec{u}_2$. But if vectors $\vec{u}_1$ and $\vec{u}_2$ happen to line up, then $k\vec{u}_1 = \vec{u}_2$ and for any linear combination

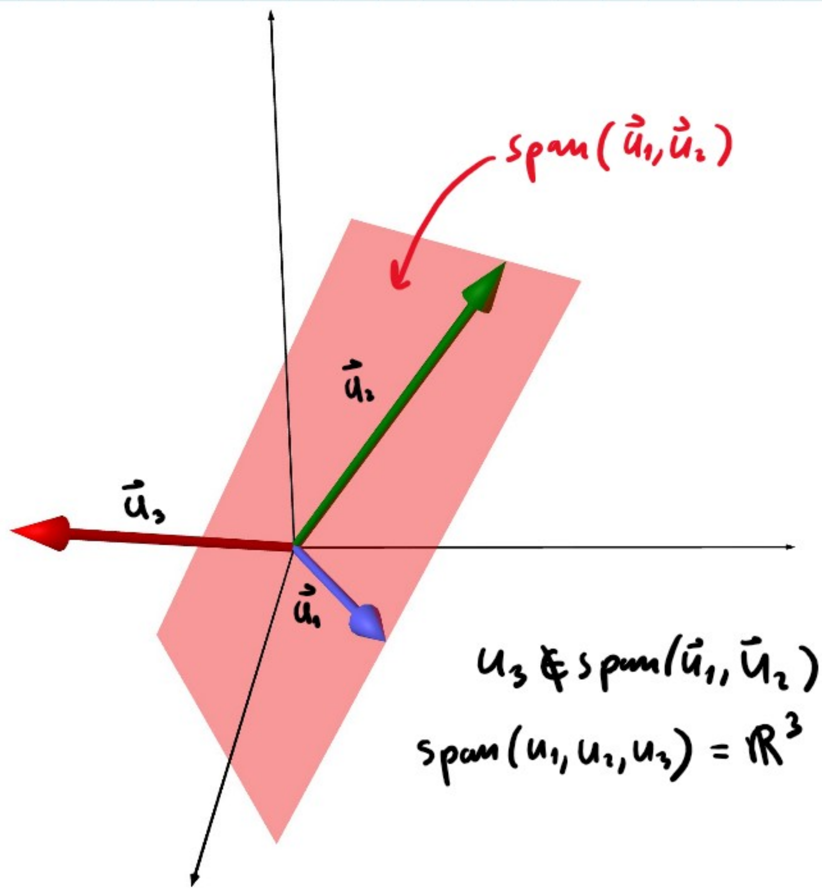
$$x\vec{u}_1 + y\vec{u}_2 = x\cdot\vec{u}_1 + y\cdot k\vec{u}_1 = (x + yk)\,\vec{u}_1$$

is a vector on the same line as $\vec{u}_1$ and $\vec{u}_2$;

is a vector on the same line as u_1 and u_2;

So, $\mathrm{span}(u_1, u_2) \neq \mathbb{R}^2$. (What happens if $\vec{u}_1 = \vec{u}_2 = 0$?)

Now in $\mathbb{R}^3$, choosing any two vectors $\vec{u}_1, \vec{u}_2 \in \mathbb{R}^3$ the $\mathrm{span}(\vec{u}_1, \vec{u}_2)$ is again a two-dimensional plane passing through the origin ($0 \in \mathrm{span}(\vec{u}_1, \vec{u}_2)$) in 3D space. Now, lets pick a third vector $\vec{u}_3 \in \mathbb{R}^3$.



if $\vec{u}_3 \in \mathrm{span}(\vec{u}_1, \vec{u}_2)$ then $\vec{u}_3 = x\vec{u}_1 + y\vec{u}_2$ and thus

$$\mathrm{span}(\vec{u}_1, \vec{u}_2) = \mathrm{span}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$$

So, in this case $\vec{u}_3$ is "redundant".

If $\vec{u}_3 \notin \mathrm{span}(\vec{u}_1, \vec{u}_2)$ (See Figure). Then any vector $\vec{b}$ can be written as linear combination $\vec{b} = x\vec{u}_1 + y\vec{u}_2 + z\vec{u}_3 = \begin{pmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, in other words $\begin{pmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{pmatrix} x = \vec{b}$ has a solution for all $\vec{b}$. ($\mathrm{rank}\begin{pmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{pmatrix} = 3$), and $\mathrm{span}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = \mathbb{R}^3$.

... $\begin{pmatrix} u_1 & u_2 & u_3 \\ | & | & | \end{pmatrix}$ ...), ... span(u_1, u_2, u_3) ...

All this motivates a

Definition: Vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ are called **linearly independent** if none of them is a linear combination of other vectors. (Otherwise, we call them **linearly dependent**)

Definition: We say that vectors $\vec{v}_1, \dots, \vec{v}_m$ in a subspace V of $\mathbb{R}^n$ form a **basis** of V if $\mathrm{span}(\vec{v}_1, \dots, \vec{v}_m) = V$ and they are linearly independent.

Linear independence can also be stated in a different way.

Definition: Vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ are linearly independent if and only if

$$c_1\vec{v}_1 + \dots + c_m\vec{v}_m = 0 \iff c_1, c_2, \dots, c_m = 0.$$

Proof: Suppose vectors are linearly dependent. Then,

$$\vec{v}_j = \sum_{i \neq j}^{m} c_i \vec{v}_i \implies \sum_{i \neq j}^{m} c_i \vec{v}_i + (-1)\vec{v}_j = 0.$$

Conversly, suppose $\sum_{i=1}^{m} c_i \vec{v}_i = 0$, and $c_j \neq 0$. Then

$$\vec{v}_j = \sum_{i \neq j}^{m} -\frac{c_i}{c_j} \vec{v}_i. \qquad \square$$

Example: Say column vectors of $n \times m$ matrix A are linearly independent. Lets determine $\ker(A)$. We need to solve $A\vec{x} = 0$.

need to solve $A\vec{x} = 0$.

$$A\vec{x} = \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = x_1\vec{v}_1 + \dots + x_m\vec{v}_m = 0; \quad \text{So}$$

finding $\ker(A)$ amounts to finding dependencies among column vectors of A. Since they are linearly independent $\ker(A) = \{0\}$.

Kernel and relations

The vectors in the kernel of an $n \times m$ matrix A correspond to the linear relations among the column vectors $\vec{v}_1, \dots, \vec{v}_m$ of A: The equation

$$A\vec{x} = \vec{0} \quad \text{means that} \quad x_1\vec{v}_1 + \dots + x_m\vec{v}_m = \vec{0}.$$

In particular, the column vectors of A are linearly independent if (and only if) $\ker(A) = \{\vec{0}\}$, or, equivalently, if $\text{rank}(A) = m$. This condition implies that $m \leq n$.

Because we need m leading variables!

Thus, we can find at most n linearly independent vectors in $\mathbb{R}^n$.

We conclude by important, alternative characterization (We saw this in CE) of Basis.

Basis and unique representation

Consider the vectors $\vec{v}_1, \dots, \vec{v}_m$ in a subspace V of $\mathbb{R}^n$.

The vectors $\vec{v}_1, \dots, \vec{v}_m$ form a basis of V if (and only if) every vector $\vec{v}$ in V can be expressed *uniquely* as a linear combination

$$\vec{v} = c_1\vec{v}_1 + \dots + c_m\vec{v}_m.$$

(we will call the coefficients $c_1, \dots, c_m$ the *coordinates* of $\vec{v}$ with respect to the basis $\vec{v}_1, \dots, \vec{v}_m$.)

Proof: Say $\vec{v}_1, \dots, \vec{v}_m$ form a basis, that is,

Proof: Say $\vec{v}_1, \dots, \vec{v}_m$ form a basis, that is, $\text{span}(\vec{v}_1, \dots, \vec{v}_m) = V$ and $\vec{v}_1, \dots, \vec{v}_m$ are linearly indep.

Then, for any $\vec{w} \in V = \text{span}(\vec{v}_1, \dots, \vec{v}_m)$ we have

$\vec{w} = c_1 \vec{v}_1 + \dots + c_m \vec{v}_m$, if we also have some other combination, $\vec{w} = d_1 \vec{v}_1 + \dots + d_m \vec{v}_m$, then

$$0 = w - w = (c_1 - d_1) \vec{v}_1 + \dots + (c_m - d_m) \vec{v}_m$$

and by linear independence $(c_i - d_i) = 0 \iff c_i = d_i$, $\forall i$.

So, linear combination is unique.

Conversly, say every vector in V can uniquely be expressed as lin. combination of $\vec{v}_1, \dots, \vec{v}_m$. Then clearly, $V = \text{span}(\vec{v}_1, \dots, \vec{v}_m)$ and if $c_1 \vec{v}_1 + \dots + c_m \vec{v}_m = 0$

$$0 = 0 \cdot \vec{v}_1 + \dots + 0 \cdot \vec{v}_m = c_1 \vec{v}_1 + \dots + c_m \vec{v}_m$$

and by uniqueness $c_i = 0$, $i = 1, \dots, m$. Thus $\vec{v}_1, \dots, \vec{v}_m$ are lin. indep. $\square$