function

("linear function")

3.6 Linear maps

Each map that conserves the structure of our vector space $\mathbb{R}^n$ is called a *linear map*. We already know that the only structure we have is the vector addition and the scalar multiplication.

Definition 3.13. Linearity of maps

A map $f : \mathbb{R}^n \to \mathbb{R}^m$ is called linear if for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, we have:

Addition in $\mathbb{R}^n$ Addition in $\mathbb{R}^m$

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \qquad (+)$$

$$f(\lambda \mathbf{x}) = \lambda f(\mathbf{x}) \qquad (\cdot)$$

Rule of thumb:

Equation $(+)$ means: First adding, then mapping = First mapping, then adding
Equation $(\cdot)$ means: First scaling, then mapping = First mapping, then scaling

We already know that for each matrix $A \in \mathbb{R}^{m\times n}$ there is an associated map f_A. This map is indeed a linear map.

$f_A(x) = Ax$

Proposition 3.14. f_A is linear

Let $m, n \in \mathbb{N}$, $A \in \mathbb{R}^{m\times n}$ and $f_A : \mathbb{R}^n \to \mathbb{R}^m$ with $\mathbf{x} \overset{f_A}{\mapsto} A\mathbf{x}$. Then the following holds:

distributive

(a) For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we have

$$f_A(\mathbf{x}+\mathbf{y}) = f_A(\mathbf{x}) + f_A(\mathbf{y}), \quad \text{i.e.} \quad A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}. \qquad (+)$$

compatible rule

(b) For all $\lambda \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$ one has

$$f_A(\lambda \mathbf{x}) = \lambda f_A(\mathbf{x}), \quad \text{i.e.} \quad A(\lambda \mathbf{x}) = \lambda A\mathbf{x}. \qquad (\cdot)$$

Proof. This follows immediately from the properties of the matrix product in Proposition 3.12. However, it may be helpful to write down a direct proof for the case $n = 2$.

(a) Let $\mathbf{x} = \binom{x_1}{x_2}$ and $\mathbf{y} = \binom{y_1}{y_2}$ be vectors in $\mathbb{R}^2$. Then we have:

$$\begin{aligned} f_A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) &= \begin{pmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{pmatrix} \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \overset{(3.2)}{=} (x_1+y_1)\mathbf{a}_1 + (x_2+y_2)\mathbf{a}_2 \\ &= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + y_1\mathbf{a}_1 + y_2\mathbf{a}_2 \overset{(3.2)}{=} \begin{pmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= A\mathbf{x} + A\mathbf{y} = f_A(\mathbf{x}) + f_A(\mathbf{y}). \end{aligned}$$

Matrix vector multiplication

(b) Let $\lambda \in \mathbb{R}$ and $\mathbf{x} = \binom{x_1}{x_2} \in \mathbb{R}^2$. Then:

$$f_A(\lambda\mathbf{x}) = A(\lambda\mathbf{x}) = \begin{pmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{pmatrix} \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix} \stackrel{(3.2)}{=} (\lambda x_1)\mathbf{a}_1 + (\lambda x_2)\mathbf{a}_2$$

$$= \lambda(x_1\mathbf{a}_1 + x_2\mathbf{a}_2) \stackrel{(3.2)}{=} \lambda \begin{pmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda A\mathbf{x} = \lambda f_A(\mathbf{x}). \qquad \square$$

$$\left.\begin{array}{l} A \in \mathbb{R}^{m\times k} \\ B \in \mathbb{R}^{k\times n} \end{array}\right\} AB \in \mathbb{R}^{m\times n} \qquad f_A: \mathbb{R}^k \to \mathbb{R}^m, \quad f_B: \mathbb{R}^n \to \mathbb{R}^k, \quad x \in \mathbb{R}^n$$

$$f_A \circ f_B : \mathbb{R}^n \to \mathbb{R}^m : \quad (f_A \circ f_B)(x) = f_A(f_B(x)) = f_A(\underbrace{Bx}_{y})$$

$$\overset{\shortparallel}{f_{AB}} \text{ linear map} \qquad = A(Bx) = (A\cdot B)\cdot x = f_{AB}(x)$$

If we have a linear map $f : \mathbb{R}^n \to \mathbb{R}^m$, we can write it as

$x \in \mathbb{R}^n, \; x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

arbitrary

$$f(\mathbf{x}) = f(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1 f(\mathbf{e}_1) + \cdots + x_n f(\mathbf{e}_n)$$

$(+), (\cdot)$

and immediately find:

Remark: Linear maps induce matrices

For each linear map $f : \mathbb{R}^n \to \mathbb{R}^m$, there is exactly one matrix $A \in \mathbb{R}^{m\times n}$ with $f = f_A$. In the columns of A, one finds the images of the canonical unit vectors:

$$A := \begin{pmatrix} | & & | \\ f(\mathbf{e}_1) & \cdots & f(\mathbf{e}_n) \\ | & & | \end{pmatrix} \in \mathbb{R}^{m\times n} \qquad f(e_i) \in \mathbb{R}^m \qquad (3.6)$$

A is often called the transformation matrix *of f.*

$$f_A(x) = f_A\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = A\cdot\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} | & & | \\ f(e_1) & \cdots & f(e_n) \\ | & & | \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= x_1 f(e_1) + x_2 f(e_2) + \cdots + x_n f(e_n) = f(x) \quad \text{for all } x \in \mathbb{R}^n$$

Assume: $\left.\begin{array}{l} f(x) = f_A(x) \\ f(x) = f_B(x) \\ \text{for all } x \end{array}\right\} \Rightarrow f_A(x) = f_B(x) \text{ for all } x \Rightarrow Ax = Bx \text{ for all } x$

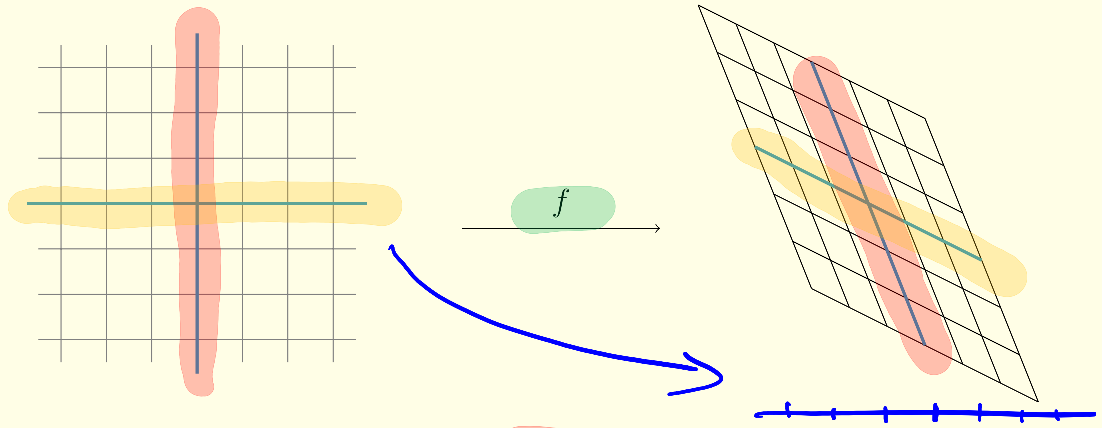
$$\Rightarrow (A-B)x = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \text{ for all } x$$

$\Rightarrow$ All columns of $A-B$ are zero $\Rightarrow A = B$ ✓

Matrix $\rightleftarrows$ Linear map

Rule of thumb: Linear map = lines stay lines

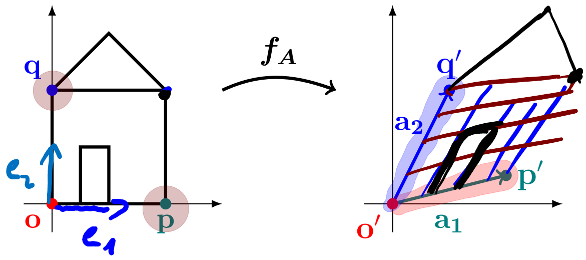
A linear map $f : \mathbb{R}^n \to \mathbb{R}^m$ conserves the linear structure: If $U \subset \mathbb{R}^n$ is a linear subspace then also the image $f(U) \subset \mathbb{R}^m$. Or in other words: Lines on the left stay lines on the right:



(However, lines could shrink down to the origin.)

A linear map is completely determined when one knows how it acts on the canonical unit vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Therefore, in $\mathbb{R}^2$, a good visualisation is to look at "houses": A house H is given by two points. Now what happens under a linear map f_A associated to a matrix A? One just have to look at the corners:

$$\mathbf{o}' := f_A(\mathbf{o}) = A\binom{0}{0} = 0\mathbf{a_1} + 0\mathbf{a_2} = \mathbf{o}$$
$$\mathbf{p}' := f_A(\mathbf{p}) = A\binom{1}{0} = 1\mathbf{a_1} + 0\mathbf{a_2} = \mathbf{a_1}$$
$$\mathbf{q}' := f_A(\mathbf{q}) = A\binom{0}{1} = 0\mathbf{a_1} + 1\mathbf{a_2} = \mathbf{a_2}$$

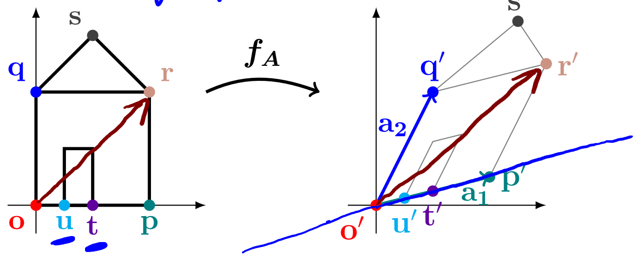


With the help of the linearity, we also know what happens with the other parts of the house, for example the corners of the door:

$r = p + q \quad , \quad f_A(r) = f_A(p) + f_A(q)$

Since $\mathbf{t} = \frac{1}{2}\mathbf{p}$ and $\mathbf{u} = \frac{1}{4}\mathbf{p}$, we have:

$$f_A(\mathbf{t}) = f_A(\tfrac{1}{2}\mathbf{p}) \overset{(\cdot)}{=} \tfrac{1}{2} f_A(\mathbf{p}) = \tfrac{1}{2}\mathbf{p}'$$
$$f_A(\mathbf{u}) = f_A(\tfrac{1}{4}\mathbf{p}) \overset{(\cdot)}{=} \tfrac{1}{4} f_A(\mathbf{p}) = \tfrac{1}{4}\mathbf{p}'$$

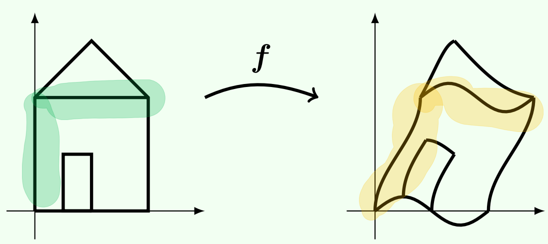


Example 3.15. A non-linear map

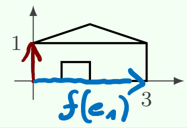
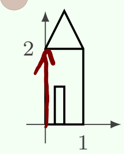
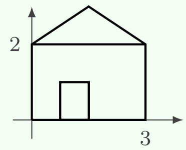
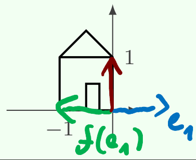
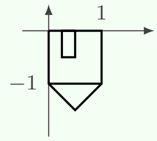
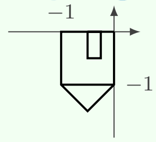
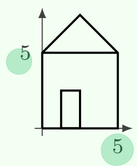
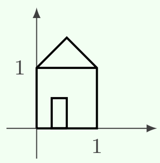
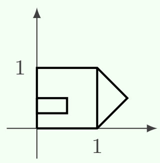
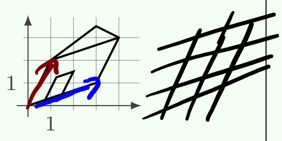
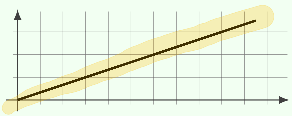
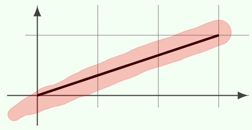
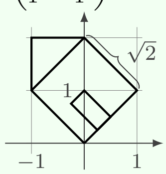
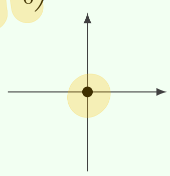
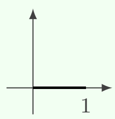
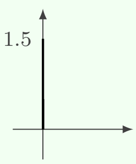
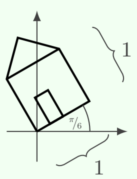
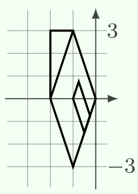
A map $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x - \frac{1}{5}(\cos(\pi y) - 1) \\ y + \frac{1}{8}\sin(2\pi x) \end{pmatrix}.$$

is not linear!

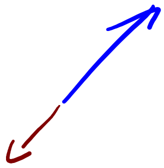


Example 3.16. Some linear house transformations

$A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$ 	$B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ 	$C = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ 
$D = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ 	$E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ 	$F = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ 
$G = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$ $H' = 5H$ 	$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 𝟙 $H' = H$ $f_I = id$ 	$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 
$K = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ 	$L = \begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix}$ 	$M = \begin{pmatrix} 3 & 0 \\ 1 & 0 \end{pmatrix}$ 
$N = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ 	$O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ 	$P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 
$Q = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ 	$R = \begin{pmatrix} \cos(\frac{\pi}{6}) & -\sin(\frac{\pi}{6}) \\ \sin(\frac{\pi}{6}) & \cos(\frac{\pi}{6}) \end{pmatrix}$ 	$S = \begin{pmatrix} -1 & -1 \\ -3 & 3 \end{pmatrix}$ 

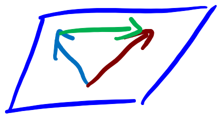
3.7 Linear dependence, linear independence, basis and dimension

We have seen that in $\mathbb{R}^2$ two vectors can be parallel (*colinear*):



$$\text{There is a } \lambda \in \mathbb{R} \text{ with } \mathbf{a} = \lambda \mathbf{b}\,.$$

Similarly, in $\mathbb{R}^3$ three vectors can be in the same plane (*coplanar*) or not:

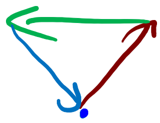


$$\text{There are } \lambda, \mu \in \mathbb{R} \text{ with } \mathbf{a} = \lambda \mathbf{b} + \mu \mathbf{c}\,.$$

If this is the case, we can build a loop of vectors, starting at $\mathbf{o}$ and ending at $\mathbf{o}$ again:

$$\mathbf{o} = (-1)\mathbf{a} + \lambda \mathbf{b} + \mu \mathbf{c}\,.$$

non-trivial linear combination



Let us generalise this:

Definition 3.17. Linear dependence and indepedence

A family $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ of k vectors from $\mathbb{R}^n$ is called linearly dependent *if we find a non-trivial linear combination for $\mathbf{o}$. This means that we can find $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ that are not all equal zero such that*

$$\sum_{j=1}^{k} \lambda_j \mathbf{v}_j = \mathbf{o}\,.$$

If this is not possible, we call the family $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ linearly independent*. This means that*

$$\sum_{j=1}^{k} \lambda_j \mathbf{v}_j = \mathbf{o} \quad \Rightarrow \quad \lambda_1, \dots, \lambda_k = 0$$

holds.

Example 3.18. Let us look at examples:

(a) The family $(\binom{1}{0}, \binom{1}{1}, \binom{0}{1})$ is linearly dependent since

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} + (-1)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

(b) $\left(\binom{1}{0}, \binom{1}{1}, \binom{2}{2}\right)$ is linearly dependent.

$$\begin{pmatrix}0\\0\end{pmatrix} = 0\cdot\begin{pmatrix}1\\0\end{pmatrix} + 2\cdot\begin{pmatrix}1\\1\end{pmatrix} + (-1)\begin{pmatrix}2\\2\end{pmatrix}$$

(c) Each family which includes $\mathbf{o}$ is linearly dependent. Also each family that has the same vector twice or more is linearly dependent.

$$\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right) \qquad \begin{pmatrix}0\\0\end{pmatrix} = 2\cdot\begin{pmatrix}0\\0\end{pmatrix} + 0\begin{pmatrix}1\\1\end{pmatrix} + 0\begin{pmatrix}0\\1\end{pmatrix}$$

(d)

$$\mathbf{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix} \qquad \mathbf{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix} \qquad \mathbf{e}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix} \qquad \text{in } \mathbb{R}^3$$

These are linearly independent vectors, because

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = \lambda_1\begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda_2\begin{pmatrix}0\\1\\0\end{pmatrix} + \lambda_3\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix}$$

yields $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

If we add an arbitrary additional vector

$$\mathbf{a} = \begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}, \qquad (a, e_1, e_2, e_3)$$

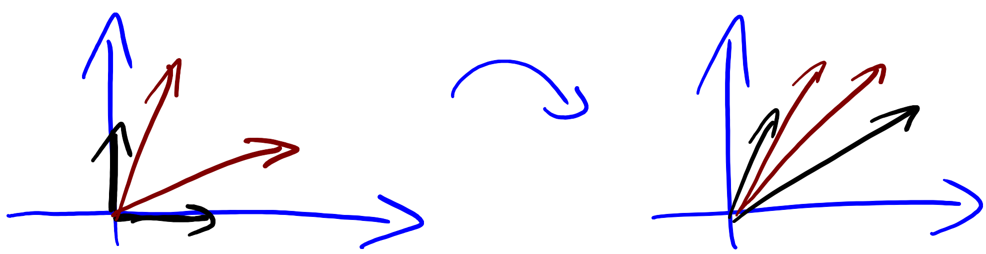
lin. dep.?

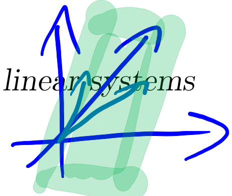
we can combine it from the other three by setting $\lambda_i = a_i$, which means:

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3$$

So the resulting set of vectors is linearly dependent.

$$0 = a_1e_1 + a_2e_2 + a_3e_3 + (-1)a$$





Proposition 3.19. Linear dependence

For a family $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *of vectors from* $\mathbb{R}^n$ *the following claims are equivalent:*

(i) $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *is linearly dependent.*

linear subspace

(ii) *There is a vector in* $\operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *that has two or more representations as a linear combination.*

(iii) *At least one of the vectors in* $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *is a linear combination of the others.*

(iv) *There is an* $i \in \{1, \ldots, k\}$ *such that we have*

omit v_i

$$\operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k).$$

(v) *There is an* $i \in \{1, \ldots, k\}$ *with* $\mathbf{v}_i \in \operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k)$.

You have to prove: $(i) \Rightarrow (ii)$, $(ii) \Rightarrow (iii)$ ... $(iii) \Rightarrow (ii)$ ...

Proof. Exercise! □

$((i) \Rightarrow (ii))$: [Show: $\neg(ii) \Rightarrow \neg(i)$]

$\neg(ii) \Rightarrow (\lambda_1 v_1 + \ldots + \lambda_k v_k = 0 \Rightarrow \lambda_1 = \cdots = \lambda_k = 0)$

$\Rightarrow (v_1, \ldots, v_k)$ lin. independent $\Rightarrow \neg(i)$ ✓

$((ii) \Rightarrow (iii))$: Assume $(ii) \Rightarrow w = \lambda_1 v_1 + \cdots + \lambda_k v_k$ and

$w = \mu_1 v_1 + \cdots + \mu_k v_k$ not all $\lambda_i = \mu_i$

$\Rightarrow (\lambda_1 - \mu_1) v_1 + \cdots + (\lambda_k - \mu_k) v_k = 0$

$\Rightarrow v_i = \ldots$

Proposition 3.20. Linear independence

For a family $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *of vectors from* $\mathbb{R}^n$ *the following are equivalent:*

(i) $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *is linearly independent.*

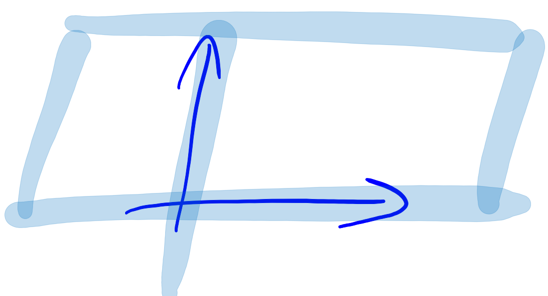
(ii) *Every vector in* $\operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *can be formed by linear combinations in exactly one way.*

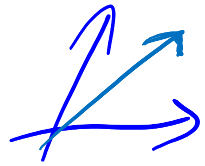
(iii) *None of the vectors in* $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *is a linear combination of the others.*

(iv) *For all* $i \in \{1, \ldots, k\}$ *we have:*

$$\operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \neq \operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k).$$

(v) *For all* $i \in \{1, \ldots, k\}$ *we have* $\mathbf{v}_i \notin \operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k)$.

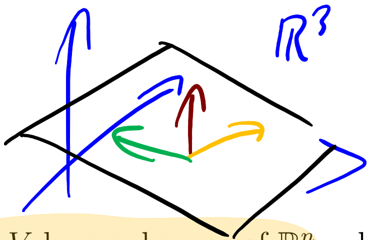


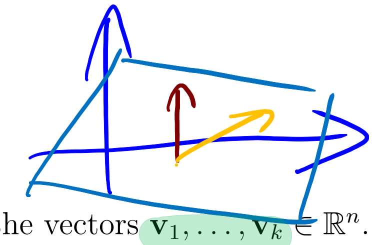


A simple consequence is:

Corollary 3.21.

If the family $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ *is linearly dependent, we can subjoin vectors and the resulting family is still linearly dependent. On the other hand, if the family* $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ *is linearly independent, we can omit vectors and the resulting family is still linearly independent.*





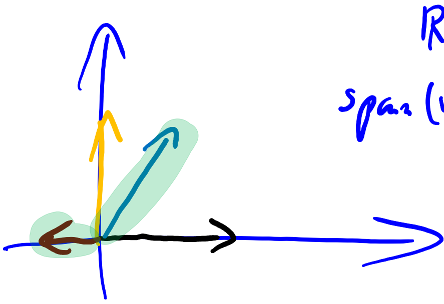
Let now V be a subspace of $\mathbb{R}^n$, which is spanned by the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$. Hence $V = \mathrm{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Question: Efficiency question:

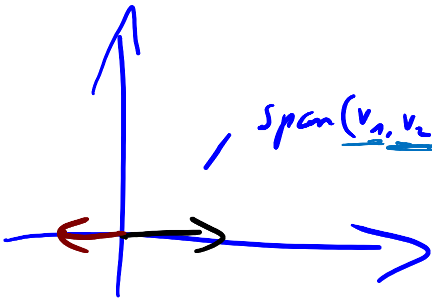
How many vectors do we actually need to span V*?*

$\mathbb{R}^2$

$\mathrm{span}(v_1, v_2, v_3, v_4) = \mathbb{R}^2$



$\mathrm{span}(v_1, v_2) \neq \mathbb{R}^2$



plane is two-dimensional

$\rightsquigarrow$ I should use two vectors

Definition 3.22. Basis, basis vectors

Let V *be a subspace of* $\mathbb{R}^n$*. A family* $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ *is called a* basis *of* V *if*

(a) $V = \mathrm{Span}(\mathcal{B})$ *and* $\rightarrow$ family generates the subspace

(b) $\mathcal{B}$ *is linearly independent.*

The elements of $\mathcal{B}$ *are called the* basis vectors *of* V.

A lot of possibilities : $-\ \mathrm{Span}(\mathcal{B}) = V$

$-$ lin. ind.

We can show that each subspace $V \subset \mathbb{R}^n$ has a basis. We define:

$\mathbb{R}^n : \quad \underbrace{(e_1, \dots, e_n)}_{\text{basis of } \mathbb{R}^n} \rightsquigarrow$ canonical basis (standard basis)