In summary:

$$\det(A) = \det(B) = 1 \cdot \det(C) = \det(D) = 2 \cdot \det(E) = 2 \cdot 13 = \underline{\underline{26}}.$$

**Remark:**

- $\det(A^{-1}) = \frac{1}{\det(A)}$ *(if the inverse exists)*
- *If $Q$ is an orthogonal matrix ($Q^TQ = \mathbb{1}$), then* $\det(Q) = \pm 1$
- *Let $P$ be a row permutation matrix, then* $\det(P) = 1$, *if the number of row exchanges is even, and* $\det(P) = -1$ *if it is odd.*
- *If $PA = LU$, then* $\det(A) = \frac{1}{\det(P)} \det(L)\det(U) = \det(P)\det(U) = \pm\det(U)$.
- *If $A = S^{-1}BS$, then* $\det(A) = \frac{1}{\det(S)} \det(B)\det(S) = \det(B)$ *(similar matrices have the same determinant).*

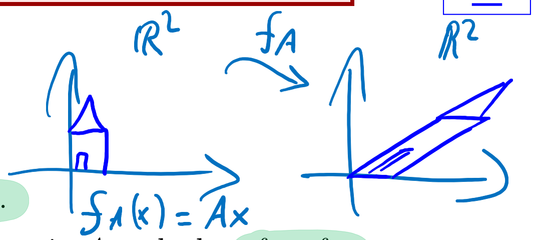
**Attention! Comparison: $n^3/3$ (Gauß) vs. $n!$ (Laplace/Leibniz formula)**

| $n$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | $\cdots$ | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n^3/3$ | 2 | 9 | 21 | 42 | 72 | 114 | 171 | 243 | 333 | $\cdots$ | 2667 |
| $n!$ | 2 | 6 | 24 | 120 | 720 | 5040 | 40320 | 362880 | 3628800 | $\cdots$ | $2.4 \cdot 10^{18}$ |

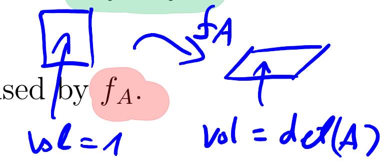
VL11 ↓

det(A) = volume

## 4.5 Determinants for linear maps




- For each matrix $A$, there is the linear map $f_A : \mathbb{R}^n \to \mathbb{R}^n$.
- For each linear map $f : \mathbb{R}^n \to \mathbb{R}^n$, there is a exactly one matrix $A$ such that $f = f_A$.
- The columns of $A$ are then the images of the unit cube under $f_A$.
- Then $\det(A)$ is the *relative change of volume* (of the unit cube) caused by $f_A$.




**Definition 4.20. Determinant for $f_A$**

*For a linear map $f : \mathbb{R}^n \to \mathbb{R}^n$, we define*

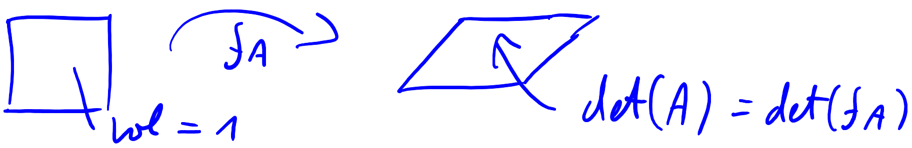
$$\det(f) := \det(A)$$

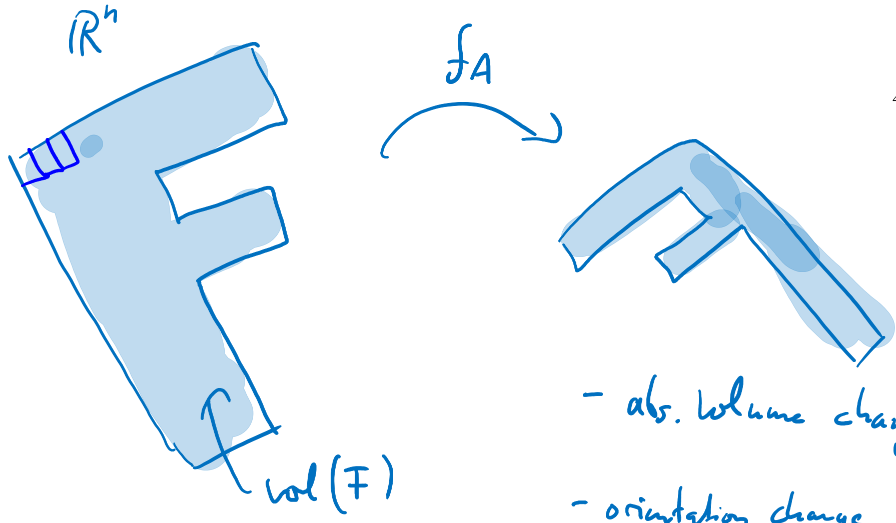
*where $A$ it the uniquely determined matrix with $f = f_A$.*

In fact $\det(f)$ is the relative change of all volumes and we remind that we have the following:

*Let $A, B \in \mathbb{R}^{n \times n}$. We have the formula:*

$$\det(f_A \circ f_B) = \det(f_A)\det(f_B)$$






- abs. volume change is given by $|\det(A)|$
- orientation change is given by the sign of $\det(A)$

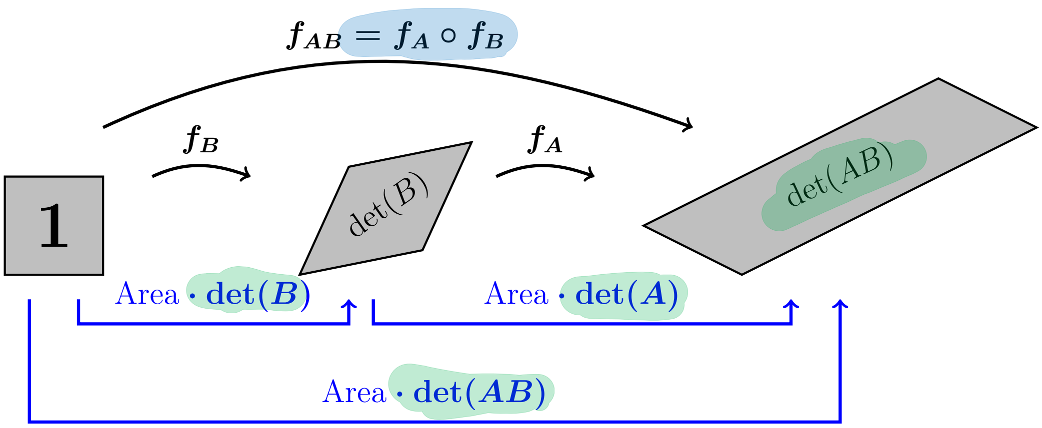
In general, $\det(A) = \det(f_A)$ describes the change of volume for every figure:

**$\det(A)$ = factor for change of volume under $f_A$**

$$\textit{figure } F \quad \overset{f_A}{\mapsto} \quad \textit{figure } F' = f_A(F)$$

$$\xrightarrow{\textit{volume} \cdot \det(A)}$$

For the composition, we get the following picture:




## 4.6 Determinants and systems of equations

Simple reasoning: if $\det(A) = 0$, then $A$ is not invertible, and vice versa. A matrix with $\det(A) = 0$ is called *singular*.

$\Leftrightarrow$ A is not invertible

$\Leftrightarrow$ LES $Ax = b$ has not a unique solution.

**Example 4.21.** $\lambda \in \mathbb{R}$

$$A(\lambda) = \begin{pmatrix} \lambda & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}, \qquad \det(A(\lambda)) = \lambda(4-3) - 1(2-2) + 1(3-4) = \lambda - 1.$$

$A(1) = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}$ columns are lin. dependent

This matrix is singular if and only if $\lambda = 1$, and in indeed, for $\lambda = 1$, we have for the column vectors $\mathbf{a}_1(\lambda) + \mathbf{a}_2 = \mathbf{a}_3$.

Conclusion: singular matrices do not appear very often. Whatever this means.

Warning: this is only good for pen-and-paper computations. In numerical computations, $\det(A + \text{round off})$ says *nothing* about invertibility of $A$, only about change of volume:

$$\det \begin{pmatrix} \varepsilon & 0 \\ 0 & 1/\varepsilon \end{pmatrix} = 1.$$

We summarise our knowledge:

**Proposition 4.22. Nonsingular matrices and LES**

*Let $A \in \mathbb{R}^{n \times n}$. Then the following is equivalent*

*(i)* $\det(A) \neq 0$,

*(ii) the columns of $A$ are linearly independent,*

*(iii) the rows of $A$ are linearly independent,* $\updownarrow$ $A^T = A$

*(iv)* $\operatorname{rank}(A) = n$,

*(v) $A$ is invertible,*

*(vi) $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbb{R}^n$,*

*(vii)* $\operatorname{Ker}(A) = \{\mathbf{o}\}$.

*Proof.* Exercise! □

## 4.7 Cramer's rule

Consider the linear system of equations, with full rank matrix $A$:

$$A\mathbf{x} = \mathbf{b}\,.$$

Then by our formula for the inverse we get:

$$\boxed{A^{-1} = \frac{C^T}{\det(A)}}$$

$$A\mathbf{x} = \mathbf{b} \Rightarrow \mathbf{x} = A^{-1}\mathbf{b} = \frac{C^T\mathbf{b}}{\det(A)}\,.$$

This let us say the following about the components of a solution:

**Proposition 4.23. Cramer's rule**

*Let $A \in \mathbb{R}^{n \times n}$ invertible and $\mathbf{b} \in \mathbb{R}^n$. Then the unique solution $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ of the LES $A\mathbf{x} = \mathbf{b}$ is given by:*

$$x_i = \frac{\det\begin{pmatrix} | & & | & \color{red}{|} & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \color{red}{\mathbf{b}} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & \color{red}{|} & | & & | \end{pmatrix}}{\underbrace{\det\begin{pmatrix} | & & | & \color{red}{|} & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \color{red}{\mathbf{a}_i} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & \color{red}{|} & | & & | \end{pmatrix}}_{A}} \qquad \text{for } i = 1, \ldots, n.$$

*Proof.* Having the cofactor matrix $C$, we already know that the solution is given by

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{C^T\mathbf{b}}{\det(A)}$$

Therefore, we just have to look at the $i$th row of the matrix $C^T\mathbf{b}$ which is given by:

co factor matrix entries

$$(C^T\mathbf{b})_i = \sum_{k=1}^{n} c_{ki} b_k = \sum_{k=1}^{n} \underbrace{\det\begin{pmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{e}_k & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{pmatrix}}_{C_{ki}} b_k$$

Linearity of det

$$= \det\begin{pmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{pmatrix}$$ □

**Attention! Do not use Cramer's rule to solve a system $A\mathbf{x}=\mathbf{b}$!**

*Cramer's rule is less efficient than Gaussian elimination. That is noticeable for large matrices.*

For computational reasons the Cramer's rule can only be used for small matrices, but the real advantage is the theoretical interest. You can use Cramer's rule in proofs if you need claims about a single component $x_i$ of the solution $\mathbf{x}$.

## Summary

- The determinant is the volume form.
- The determinant fulfils three defining properties:
  (1) Linear in each column.
  (2) Alternating when exchanging columns.
  (3) The identity matrix has determinant 1.
- To calculate a determinant, you have the Leibniz formula, the Laplace expansion or Gaussian elimination (without scaling!).

Some analytical geometry

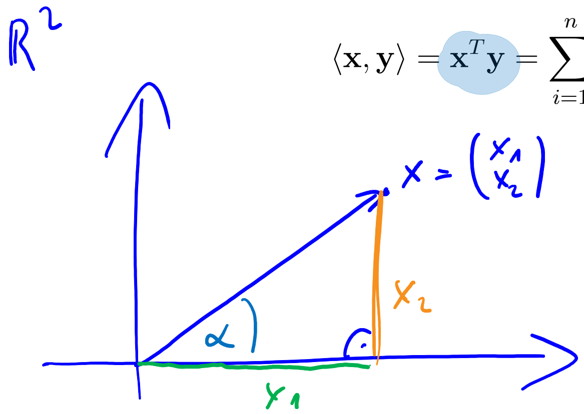
# 5 General inner products, orthogonality and distances

We have already encountered the standard inner product (also called Euclidean scalar product) in $\mathbb{R}^n$:

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

- positive definite (S1)
- linear in the first argument (S2), (S3)
- symmetric (S4)

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{n} x_i y_i \,, \quad \text{for} \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$




$$\sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + x_2^2} =: \|x\|$$

$$\cos(\alpha) = \frac{x_1}{\|x\|} = \frac{\langle x, e_1 \rangle}{\|x\| \cdot \|e_1\|}$$

With the help of this inner product, we are able to define and compute many useful things:

- length: $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$
- distances: $\operatorname{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$
- angle: $\cos(\angle(\mathbf{x}, \mathbf{y})) := \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}$
- orthogonality: $\mathbf{x} \perp \mathbf{y} :\Leftrightarrow \langle \mathbf{x}, \mathbf{y} \rangle = 0$.
- orthogonal projections, e.g. the height
- rotations about an axis by an angle
- reflections at a hyperplane

## 5.1 General inner products in $\mathbb{R}^n$

**Definition 5.1. Inner product**

*Let $V$ be $\mathbb{R}^n$ or a subspace of $\mathbb{R}^n$. We call a map of two arguments $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ an inner product if it satisfies for all $\mathbf{x}, \mathbf{y}, \mathbf{v} \in V$ and $\lambda \in \mathbb{R}$:*

*(S1) Positive definiteness:*

$$\langle \mathbf{x}, \mathbf{x} \rangle > 0 \text{ for } \mathbf{x} \neq \mathbf{o}$$

(S1') $\langle x, x \rangle \geq 0$ for all $x$ and $\langle x, x \rangle = 0 \Rightarrow x = 0$

*(S2) Additivity in the first argument:*

$$\langle \mathbf{x}+\mathbf{y}, \mathbf{v} \rangle = \langle \mathbf{x}, \mathbf{v} \rangle + \langle \mathbf{y}, \mathbf{v} \rangle$$

linear

*(S3) Homogenity in the first argument:*

$$\langle \lambda\mathbf{x}, \mathbf{v} \rangle = \lambda \langle \mathbf{x}, \mathbf{v} \rangle$$

*(S4) Symmetry:*

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$$

We usually summarise (S2) and (S3) to *linearity in the first argument*. Note that from (S3) always follows $\langle \mathbf{o}, \mathbf{o} \rangle = 0 \cdot \langle \mathbf{o}, \mathbf{o} \rangle = 0$. In combination with (S1), we get:

$$\langle \mathbf{x}, \mathbf{x} \rangle = 0 \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{o}\,. \tag{5.1}$$

Also, due to positive definiteness, we can define a norm (or length) via

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}. \in [0, \infty)$$

We call it the the associated norm with respect to $\langle\cdot,\cdot\rangle$.

- Inner products are also linear in the second argument, by symmetry.
- Later, we will define complex-valued inner products that fulfil instead of (S4): $\mathbb{C}^n$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}\,. \tag{5.2}$$

  Then the second argument actually has different properties than the first.
- In the usual real case, the binomial formulas hold:

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 \pm 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2$$
$$\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle = \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2.$$

don't use a dot!

**Example 5.2.** The standard inner product on $\mathbb{R}^n$:

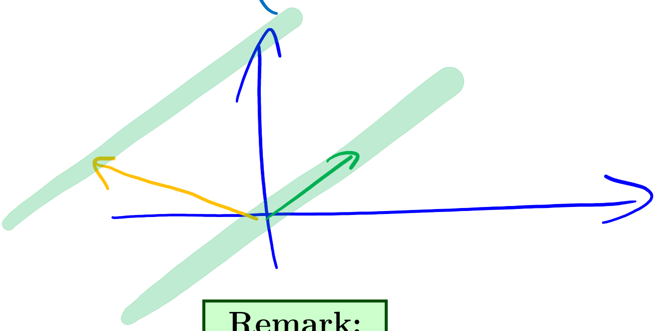
$$\langle \mathbf{x}, \mathbf{y} \rangle_{euklid} := \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i\,.$$

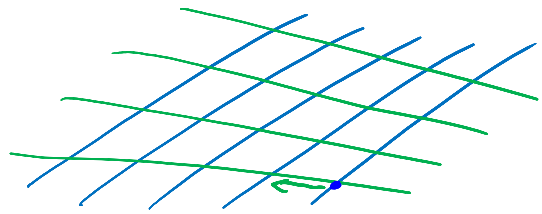
Another example: (in $\mathbb{R}^2$): $\langle x, y \rangle = 2x_1 \cdot y_1 - x_1 \cdot y_2 - x_2 \cdot y_1 + 5x_2 y_2$

↳ Check the rules (S1)-(S4).



$\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-4\\1\end{pmatrix}$ are orthogonal: $\left\langle \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-4\\1\end{pmatrix}\right\rangle = 2\cdot(-4) - 1 + 4 + 5$
$= 0$





**Remark:**

*Due to its simplicity, this inner product is prominent in theory and practice. However, in particular for very large scale problems with special structure other "specially tailored" inner products play a major role.*

**Definition 5.3. Positive definite matrix**

*A matrix $A \in \mathbb{R}^{n\times n}$ is called* positive definite *if it is symmetric ($A^T = A$) and satisfies*

$$\langle \mathbf{x}, A\mathbf{x}\rangle_{euklid} = \mathbf{x}^T A\mathbf{x} > 0$$

*for all $\mathbf{x} \neq \mathbf{o}$.*

$\left( \begin{array}{l} x^TAx \geq 0 \text{ for all } x \\ x^TAx = 0 \Rightarrow x = o \end{array} \right)$

**Example 5.4.** Each diagonal matrix $D \in \mathbb{R}^{n\times n}$ with positive entries on the diagonal is a positive definite matrix.

$D = \begin{pmatrix} 2 & & \\ & 3 & \\ & & 4 \end{pmatrix}$, $\quad \langle x, Dx\rangle_{eukl.} = 2x_1^2 + 3x_2^2 + 4x_3^2 \geq 0$

$\langle x, Dx\rangle = 0 \Rightarrow x_i^2 = 0$ for all $i$
$\Rightarrow x = 0$

$(D^T = D)$

no zero on the diagonal

Let $A \in \mathbb{R}^{n\times n}$ be a positive definite matrix. Then the following defines an inner product on $\mathbb{R}^n$:

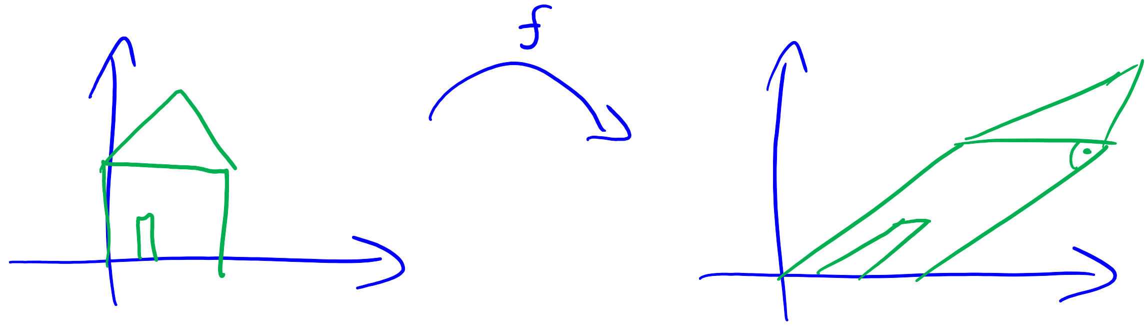
$$\langle \mathbf{x}, \mathbf{y}\rangle_A := \langle \mathbf{x}, A\mathbf{y}\rangle_{euklid} = \mathbf{x}^T A\mathbf{y}$$

(S2), (S3) ✓

(S4) $\langle x, y\rangle_A = \langle x, Ay\rangle_{eukl} = \langle A^Tx, y\rangle_{eukl.} = \langle Ax, y\rangle_{eukl} = \langle y, Ax\rangle$
$= \langle y, x\rangle_A$

(S1) $\langle x,x\rangle_A = \langle x, Ax\rangle_{\text{eucl.}} > 0$ for all $x \neq 0$ ✓






Our abstract assumptions already yield all the useful formulas, known from our standard inner product:

**Proposition 5.5.**

*Let $\langle\cdot,\cdot\rangle$ be an inner product on a subspace $V \subset \mathbb{R}^n$ and $\|\cdot\|$ its associated norm. Then for all $\mathbf{x}, \mathbf{y} \in V$ and $\lambda \in \mathbb{R}$, we have:*

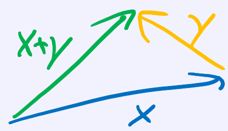
*(a)* $|\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$ *(Cauchy-Schwarz inequality). Equality holds if and only if* $\mathbf{x}$ *and* $\mathbf{y}$ *are colinear (written as* $\mathbf{x} \parallel \mathbf{y}$*).*

*(b) The norm fulfils three properties:*

*(N1)* $\|\mathbf{x}\| > 0$ *for all* $\mathbf{x} \neq \mathbf{o}$, *and* $\|\mathbf{x}\| = 0$ *only for* $\mathbf{x} = \mathbf{o}$,

*(N2)* $\|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|$,

*(N3)* $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. (triangle inequality)




*Proof.* We show the Cauchy-Schwarz inequality (CSI) in a short proof. Let $\mathbf{y} \neq \mathbf{o}$, otherwise the CSI reads $0 = 0$. Exercise 3.1 (b)

For any $\lambda \in \mathbb{R}$ the binomial formula yields:

$$0 \leq \|\mathbf{x} - \lambda\mathbf{y}\|^2\|\mathbf{y}\|^2 = \|\mathbf{x}\|^2\|\|\mathbf{y}\|^2 - 2\lambda\langle\mathbf{x},\mathbf{y}\rangle\|\mathbf{y}\|^2 + \lambda^2\|\mathbf{y}\|^4.$$

(This *is* zero, if $\mathbf{y} = \mathbf{o}$, or $\mathbf{x} = \lambda\mathbf{y}$, i.e. $\mathbf{x}$ and $\mathbf{y}$ are colinear). Setting $\lambda = \langle\mathbf{x},\mathbf{y}\rangle/\|\mathbf{y}\|^2$, we obtain

$$0 \leq \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - 2\langle\mathbf{x},\mathbf{y}\rangle^2 + \langle\mathbf{x},\mathbf{y}\rangle^2 = \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - \langle\mathbf{x},\mathbf{y}\rangle^2.$$

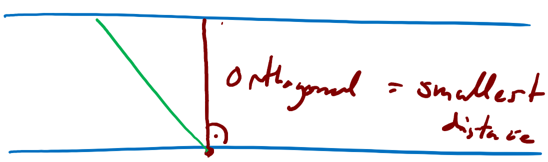
The norm properties are left as an exercise. □

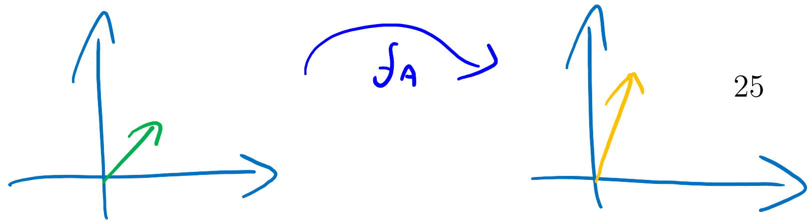
No matter which inner product we are using, we can define *orthogonality* as follows:

$$\mathbf{x} \perp \mathbf{y} \quad :\Leftrightarrow \quad \langle\mathbf{x},\mathbf{y}\rangle = 0.$$

$$|\langle x,y\rangle| \leq \|x\|\cdot\|y\|$$

$$\frac{|\langle x,y\rangle|}{\|x\|\cdot\|y\|} \in [0,1] \rightsquigarrow \text{angle definition}$$






### A norm for matrices

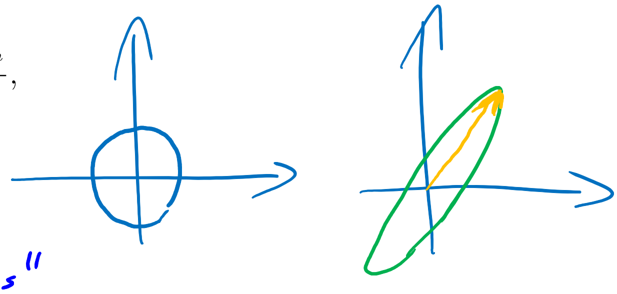
Once we can measure the size of a vector $\mathbf{v}$ by a norm $\|\mathbf{v}\|$, we may think about measuring the "size" of a linear map. Consider $A \in \mathbb{R}^{m\times n}$, and $\mathbf{w} = A\mathbf{v}$. Then the following quotient

$$\frac{\|\mathbf{w}\|_{\mathbb{R}^m}}{\|\mathbf{v}\|_{\mathbb{R}^n}} = \frac{\|A\mathbf{v}\|_{\mathbb{R}^m}}{\|\mathbf{v}\|_{\mathbb{R}^n}}$$

tells us, how much longer (or shorter) $\mathbf{w} = A\mathbf{v}$ is, compared to $\mathbf{v}$. $A$ should be "large", if it produces long vectors from short ones, and "small", if it produces short vectors from long ones. Thus, we may define

$$\|A\| := \max_{\mathbf{v}\neq 0} \frac{\|A\mathbf{v}\|_{\mathbb{R}^m}}{\|\mathbf{v}\|_{\mathbb{R}^n}},$$

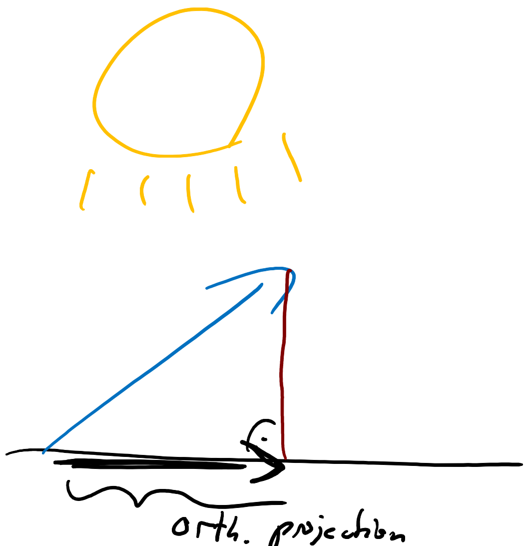
"matrix norms"



so that we have:

$$\|\mathbf{w}\|_{\mathbb{R}^m} = \|A\mathbf{v}\|_{\mathbb{R}^m} \le \|A\|\|\mathbf{v}\|_{\mathbb{R}^n}.$$

It is not easy to compute this norm. We will consider a possibility later.




## 5.2 Orthogonal projections

### 5.2.1 Orthogonal projection onto a line

Imagine you ride a rowboat on a river. You want to go in a direction $\mathbf{r} \neq \mathbf{o}$. However water flows in direction $\mathbf{x}$, which is not parallel to $\mathbf{r}$.