Vector spaces

Def A set G of elements with an operation $+: G \times G \to G$ is called a group if the following properties hold:

(G1) Associativity: $\forall a, b, c \in G: \quad (a+b)+c = a+(b+c)$

(G2) Identity element: $\exists e \in G \ \forall g \in G: \quad e+g = g+e = g$

(G3) Inverse elements: $\forall a \in G \ \exists b \in G: \quad a+b = b+a = e$

The group is called a commutative group (Abelian group) if we have additionally that

(G4) $\forall a, b \in G: \quad a+b = b+a$

Examples:
* $(\mathbb{R}^n, +)$ is a group
* $(\mathbb{R}^+, \cdot)$ "
* $(\mathbb{R}^-, \cdot)$ is not a group.
* $S_n := \{\pi : \{1, \dots, n\} \to \{1, \dots, n\} \mid \pi \text{ is bijective}\}$

  $\circ : S_n \times S_n \to S_n, \quad \pi_1 \circ \pi_2 (i) = \pi_1 (\pi_2 (i))$

  $(S_n, \circ)$ is a group.

Def A set F with two operations $+, \cdot : F \times F \to F$ is called a field if the following properties hold:

(F1) $(F, +)$ is a commutative group, with identity element 0.

(F2) $(F \setminus \{0\}, \cdot)$ is a commutative group with id. el. 1

(F3) Distributivity: $\forall a, b, c \in F: \quad a \cdot (b+c) = a \cdot b + a \cdot c$

Examples:
* $(\mathbb{R}, +, \cdot)$
* $(\mathbb{C}, +, \cdot)$

- $n \in \mathbb{Z}$, consider $\mathbb{Z}_n := \{0, 1, \dots, n-1\}$

  $a +_n b := (a+b) \bmod n$

  $a \cdot_n b := (a \cdot b) \bmod n$

  Then $(\mathbb{Z}_n, +_n, \cdot_n)$ is a field if and only if n is prime.

<u>Def</u> Let F be a field with id. elements 0 and 1.
A vector space over the field F is a ~~the~~ set V with a mapping $+ : V \times V \to V$ ("vector addition") and a mapping $\cdot : F \times V \to V$ ("scalar multiplication") such that:

(V1) $(V, +)$ is a commutative group.

(V2) Multiplicative identity: $\forall v \in V: \ 1 \cdot v = v$

(V3) Distributive properties: $\forall a, b \in F \quad \forall u, v \in V$

$$a \cdot (u + v) = a \cdot u + a \cdot v$$
$$(a+b)\, u = a \cdot u + b \cdot u$$

Elements of V are called vectors, elements of F are called scalars.

<u>Examples</u>:

- $\mathbb{R}^n$ with the standard operations.
- Function spaces:
  - $\mathbb{R}^{\mathcal{X}} := \{ f : \mathcal{X} \to \mathbb{R} \}$ the space of all real valued fcts on a set $\mathcal{X}$. Define:

    $+ : \mathbb{R}^{\mathcal{X}} \times \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}}, \quad (f+g)(x) := f(x) + g(x)$

$\cdot : \mathbb{R} \times \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}}$, $(\lambda \cdot f)(x) := \lambda \cdot (f(x))$

Then $(\mathbb{R}^{\mathcal{X}}, +, \cdot)$ is a real vector space.

- $\mathcal{C}(\mathcal{X}) := \{ f : \mathcal{X} \to \mathbb{R} \mid f \text{ is continuous} \}$
- $\mathcal{C}^r([a,b]) = \{ f : [a,b] \to \mathbb{R} \mid f \text{ is } r \text{ times cont. differentiable} \}$

Def Let V be a vector space, $U \subset V$ non-empty set. We call U a <u>subspace</u> of V if it is closed under linear combinations: $\forall \lambda, \mu \in F \ \forall u, v \in U: \ \lambda u + \mu \cdot v \in U$

Examples:
- $\mathcal{C}(X)$ is a subspace of $\mathbb{R}^X$.
- The set S of symmetric matrices of size $n \times n$ is a subspace of $\mathbb{R}^{n \times n}$.

Def V vector space over F, $u_1, \dots, u_n \in V$, $\lambda_1, \dots, \lambda_n \in F$. Then $\sum_{i=1}^n \lambda_i u_i$ is called a <u>linear combination</u>. The set of all lin. comb. of $u_1, \dots, u_n$ is called the <u>span</u> (linear hull) of $u_1, \dots, u_n$. Notation:

$$\operatorname{span}(u_1, \dots, u_n) := \left\{ \sum_{i=1}^n \lambda_i u_i \;\middle|\; \lambda_i \in F \right\}.$$

The set $U := \{u_1, \dots, u_n\}$ is the <u>generator</u> of span (U).

Def A set of vectors $v_1, \dots, v_n$ is called <u>linearly independent</u> if the following holds:

$$\sum_{i=1}^n \lambda_i v_i = 0 \Rightarrow \lambda_1 = \dots = \lambda_n = 0.$$

Examples:
- The vectors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} \in \mathbb{R}^3$ are lin. indep.
- The functions $\sin(x)$ and $\cos(x) \in \mathbb{R}^{\mathbb{R}}$ are lin. ind.
- Any set of $d+1$ vectors in $\mathbb{R}^d$ is lin. dependent.

Basis and dimension

Def A subset $\mathcal{B}$ of a vectorspace V is called a (Hamel) basis if

(B1) $\mathrm{Span}(\mathcal{B}) = V$

(B2) $\mathcal{B}$ is lin. independent.

Example

- The canonical basis of $\mathbb{R}^3$ is $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$
- Another basis of $\mathbb{R}^3$ is given by $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}$ or $\begin{pmatrix}0.5\\0.8\\0.4\end{pmatrix}, \begin{pmatrix}1.8\\0.3\\0.3\end{pmatrix}, \begin{pmatrix}-2.2\\-1.3\\3.5\end{pmatrix}$

Proposition: If $U = \{u_1, \dots, u_n\}$ spans a VS V, then the set U can be reduced to a basis of V.

Proof
- If U is already lin. ind: done
- If U is dependent: there ex. $u \in U$ that is a lin. comb. of the other vectors in U. We remove it. Keep on doing until remaining set is lin. ind.

Def A VS is called finite-dim if it has a finite basis.

Prop Let $U = \{u_1, \dots, u_n\} \subset V$ be a set of lin. ind. vectors, and let V be a finite-dim VS. Then U can be extended to a basis of V.

Proof (Sketch) Let $w_1, \dots, w_m$ be a basis of V. Consider the set

$\{u_1, \dots, u_n, w_1, \dots, w_m\}$. Remove vectors "from the end" until the remaining vectors are lin. independent.

- remaining set spans V
- remaining set is lin. ind. by constr.
- " contains U. ∎

<u>Corollary</u> Let V be a finite-dim VS. Then any two bases of V have the same length.

<u>Def</u> The length of a basis of a finite-dim VS is called the <u>dimension</u> of V.

<u>Def</u> Assume that we have U_1, U_2 subspaces of V. The <u>sum</u> of the two spaces is defined as

$$U_1 + U_2 := \{ u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2 \}$$

The sum is called a <u>direct sum</u>, if each element in the sum can be written in exactly one way.

Notation: $U_1 \oplus U_2$

Prop Suppose V is finite-dim, and $U \subset V$ is a subspace. Then there exists a subspace $W \subset V$ such that $U \oplus W = V$.

Proof (sketch) Let the set $\{u_1, \dots, u_k\}$ basis of of U. Extend it to a basis of V, say the resulting set is

$$\{\underbrace{u_1, \dots, u_k}_{\rightsquigarrow U}, \underbrace{v_1, \dots, v_m}_{\rightsquigarrow W}\}.$$

Define

$$W = \operatorname{span}\{v_1, \dots, v_m\}.$$

Linear Mappings

Def Let U, V VS over F. A mapping $T: U \to V$ is called linear if $\forall u_1, u_2 \in U$, $\forall \lambda \in F$

$$f(u_1+u_2) = f(u_1) + f(u_2)$$

$$f(\lambda u_1) = \lambda f(u_1)$$

The set of all linear mappings from $U \to V$ is denoted $\mathcal{L}(U, V)$. If $U = V$, then we write $\mathcal{L}(U)$.

Examples:
- $T: \mathcal{C}[a,b] \to \mathbb{R}$, $f \mapsto \int_a^b f(x)\,dx$ (Integration)
- $D: \mathcal{C}^\infty[a,b] \to \mathcal{C}^\infty[a,b]$, $f \mapsto f'$ (Differentiation)

Def. $T \in \mathcal{L}(U,V)$. Then kernel of T (null space of T) is defined as

$$\ker(T) := \operatorname{null}(T) := \{u \in U \mid Tu = 0\}$$

Prop
- $\ker(T)$ is a subspace of U.
- T injective iff ($\Leftrightarrow$) $\ker T = \{0\}$.

Def The range of T (image of T) is defined as

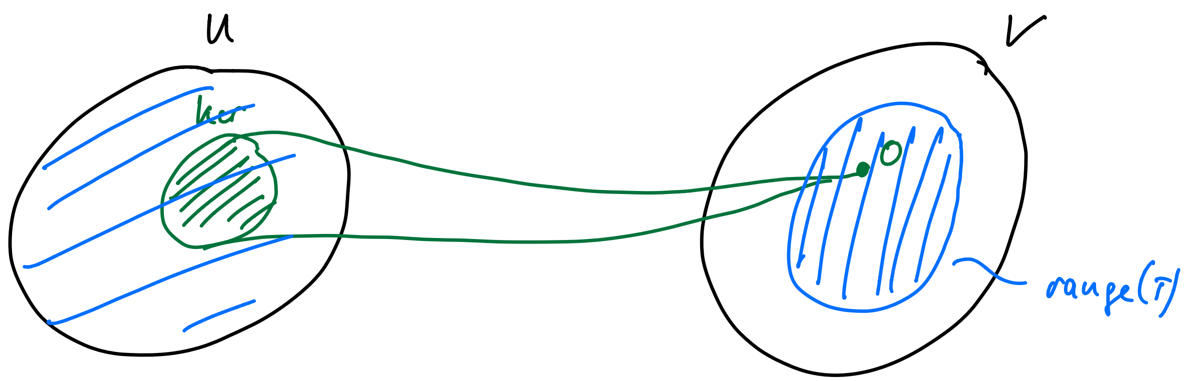
$$\operatorname{range}(T) := \operatorname{Im}(T) := \{Tu \mid u \in U\}$$

Prop
- The range is always a subspace of V.
- T is surjective iff $\operatorname{range}(T) = V$.

__Def__ : $V' \subset V$, v' any set. The __pre-image__ of V' is defined as

$$T^{-1}(V') := \{ u \in U \mid Tu \in V' \}.$$

__Prop__ : If $V' \subset V$ is a subspace of V, then $T^{-1}(V')$ is a subspace of U.



__Theorem__ : Let V be finite-dim, W any VS, $T \in \mathcal{L}(V, W)$.

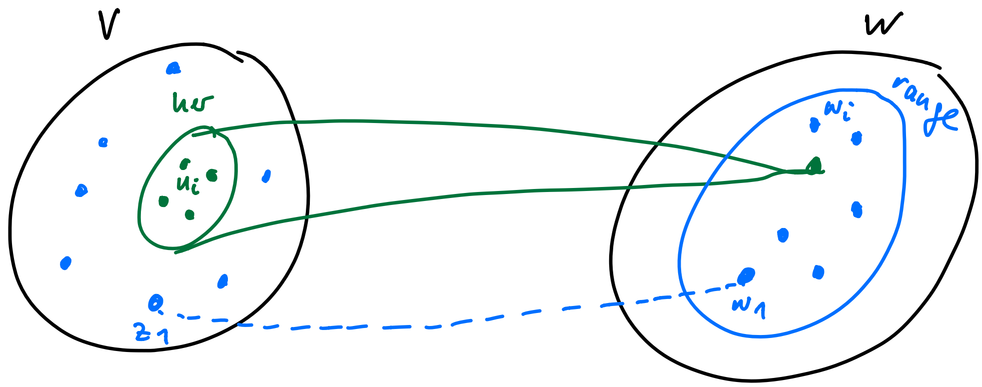
Let $u_1, \dots, u_n$ be a basis of $\ker(T) \subset V$

Let $w_1, \dots, w_m$ be a basis of range $(T) \subset W$.

Then $u_1, \dots, u_n, T^{-1}(w_1), \dots, T^{-1}(w_m) \subset V$ forms a basis of V.

In particular, $\dim(V) = \dim(\ker(T)) + \dim(\text{range}(T))$.

__Proof__ Denote $T^{-1}(w_1) =: z_1, \dots, T^{-1}(w_m) = z_m$.



Step 1: $V \subset \operatorname{span}\{u_1, \dots, u_n, z_1, \dots, z_m\}$

Let $v \in V$, consider $Tv \in \operatorname{range}(T)$.

$\Rightarrow \exists \lambda_1, \dots, \lambda_m$ s.t.

$$Tv = \lambda_1 w_1 + \dots + \lambda_m w_m$$

$$= \lambda_1 T(z_1) + \dots + \lambda_m T(z_m)$$

$$= \underline{T(\lambda_1 z_1 + \dots + \lambda_m z_m)}$$

$$\Rightarrow Tv - T(\lambda_1 z_1 + \dots + \lambda_m z_m) = 0$$

$$= T(\underbrace{v - (\lambda_1 z_1 + \dots + \lambda_m z_m)}_{\in \ker(T)})$$

$\Rightarrow \exists \mu_1, \dots, \mu_n$ s.t. $v - (\lambda_1 z_1 + \dots + \lambda_m z_m) = \mu_1 u_1 + \dots + \mu_n u_n$

$\Rightarrow v = \lambda_1 z_1 + \dots + \lambda_m z_m + \mu_1 u_1 + \dots + \mu_n u_n$

Step 2 : $u_1, \dots, u_n, z_1, \dots, z_m$ are lin. indep.

Assume that $\mu_1 u_1 + \dots + \mu_n u_n + \lambda_1 z_1 + \dots + \lambda_m z_m = 0$ (*)

$$\lambda_1 w_1 + \dots + \lambda_m w_m = \lambda_1 T(z_1) + \dots + \lambda_m T(z_m)$$

$$= \lambda_1 T(z_1) + \dots + \lambda_m T(z_m) + \underbrace{\mu_1 T(u_1) + \dots + \mu_n T(u_n)}_{=0}$$

$$= T(\underbrace{\lambda_1 z_1 + \dots + \lambda_m z_m + \mu_1 u_1 + \dots + \mu_n u_n}_{=0 \text{ by } (*)}) = 0$$

$$\Rightarrow \lambda_1 w_1 + \dots + \lambda_m w_m = 0 \underset{\substack{w_1, \dots, w_m \\ \text{basis}}}{\Rightarrow} \lambda_1 = \dots = \lambda_m = 0$$

$\Rightarrow \mu_1 u_1 + \dots + \mu_n u_n = 0$ by (*)

$\Rightarrow \mu_1 = \dots = \mu_n = 0$ because $u_1, \dots, u_n$ basis.

$\blacksquare$

typo corrected

Prop $T \in \mathcal{L}(V, V)$, V finite-dim. Then the following three statements are equivalent:

(i) T injective.

(ii) T surjective.

(iii) T bijective.

Proof Direct consequence of theorem.

⚠ Does not hold in ∞-dim spaces!

Matrices and linear maps

Notation:

$$A = m \text{ rows} \left\{ \overbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}^{n \text{ col.}} \right. = (a_{ij})_{\substack{i=1\dots m \\ j=1\dots n}}$$

Consider $T \in \mathcal{L}(V, W)$, V, W finite-dim,

let $v_1, \dots, v_n$ be a basis of V

$w_1, \dots, w_m$ a basis of W

- $v = \lambda_1 v_1 + \dots + \lambda_n v_n$

  $T(v) = T(\lambda_1 v_1 + \dots + \lambda_n v_n)$

  $= \lambda_1 T(v_1) + \dots + \lambda_n T(v_n)$

- Each image vector $T(v_j)$ can be expressed in basis $w_1, \dots, w_m$:

  there exist coefficients $a_{1j}, \dots, a_{mj}$ s.t.

$$\boxed{T(v_j) = a_{1j} w_1 + \dots + a_{mj} w_m}$$

- we now stack these coefficients in a matrix:

m rows, one for each basis vector of W

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & \boxed{a_{1j}} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & \boxed{a_{mj}} & \cdots & a_{mn} \end{pmatrix}}_{\text{n cols, one for each basis vector of } V}$$

col j

= matrix of mapping T with respect to the bases $v_1, \dots, v_n$ of V, $w_1 \dots w_m$ of W.

<u>Notation</u> Let $T: V \to W$ be linear, let $\mathcal{B}$ a basis of V, $\mathcal{C}$ basis of W. We denote by

$$M(T, \mathcal{B}, \mathcal{C})$$

the matrix corresponding to T wrt bases $\mathcal{B}$ and $\mathcal{C}$.

<u>Convenient properties of matrices</u>: Let V, W vector spaces, consider the basis fixed. Let $S, T \in \mathcal{L}(V, W)$

- $M(S+T) = M(S) + M(T)$
- $M(\lambda S) = \lambda M(S)$

- For $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ we have that

$$\underbrace{T(v)}_{\text{image of } V \text{ under } T} = \underbrace{M(T) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}}_{\text{matrix-vector product}}$$

where $v_1, \dots, v_n$ is basis of V

- $T: U \to V$, $S: V \to W$ linear, then
$$M(S \circ T) = M(S) \cdot M(T)$$

Def Given a matrix $A = (a_{ij})_{ij} \in F^{m \times n}$, the <u>transpose</u> matrix is given as

$$(A^t)_{kj} = A_{jk}$$

$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, $A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

Notation: A^t, A'

If $F = \mathbb{C}$, then the <u>conjugate transpose</u> matrix is defined as

$$(A^*)_{ij} = \overline{a_{ji}}$$

$(A)_{ij} = a_{ij}$

$x = a + ib$
$\bar{x} = a - ib$

Prop Assume F is $\mathbb{R}$. Then:

$$\langle x, Ay \rangle_{\mathbb{R}^n} = \langle A^T x, y \rangle_{\mathbb{R}^n}$$

Assume $F = \mathbb{C}$. Then:

$$\langle x, Ay \rangle_{\mathbb{C}^n} = \langle A^* x, y \rangle_{\mathbb{C}^n}$$

Invertible maps and matrices

Def $T \in \mathcal{L}(V, W)$ is called invertible if there exists a linear map $S \in \mathcal{L}(W, V)$ such that

$$S \circ T = Id_V \quad \text{and} \quad T \circ S = Id_W$$

The map S is called the inverse of T, denoted by T^{-1}.

Rem Inverse maps are unique.

Prop A linear map is invertible iff it is inj. and surj.

Proof "$\Rightarrow$" Invertible $\Rightarrow$ injective:

suppose $T(u) = T(v)$. Then $u = T^{-1}(T(u))$

$= T^{-1}(Tv) = v$ $\Rightarrow$ injective

Invertible $\Rightarrow$ surjective: $w \in W$. Then

$w = T(T^{-1}(w)) \Rightarrow w \in$ range of T

$\Rightarrow$ surjective.

"$\Leftarrow$" inj & surj $\Rightarrow$ invertible.

Let $w \in W$. There exists unique $v \in V$ s.th. $T(v) = w$.

Define the mapping: $S(w) = v$. Clearly here $T \circ S = Id$.

Let $v \in V$. Then $T((S \circ T)(v)) =$

$= (T \circ S)(Tv) = Id \circ Tv = Tv$.

$\Rightarrow (S \circ T)v = v \Rightarrow S \circ T = Id$

Linearity: $T(Sw_1 + Sw_2) = TSw_1 + TSw_2 =$

$= w_1 + w_2$.

$\Rightarrow$ T maps $Sw_1 + Sw_2 \longmapsto w_1 + w_2$

$\Rightarrow$ $S(w_1 + w_2) \mapsto w_1 + w_2$

$\Rightarrow$ $S(w_1 + w_2) = Sw_1 + Sw_2$

Similarly for scalar mult. $\square$

Inverse matrix

Def A square matrix $A \in F^{n \times n}$ is invertible if there exists a square matrix $B \in F^{n \times n}$ such that

$$A \cdot B = B \cdot A = Id = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

The matrix B is called the inverse matrix, and is denoted by A^{-1}.

Prop The inverse matrix represents the inverse of the corr. lin. map, that is: $T : V \to V$

$$M(T^{-1}) = \big(M(T)\big)^{-1}$$

matrix of (inverse map) inverse matrix (of the original matrix)

In particular, a matrix is invertible iff the corr. map is invertible.

Remarks :

- The inverse matrix does not always exist.
- $(A^{-1})^{-1} = A$, $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$
- A^t invertible $\Leftrightarrow$ A invertible,

  $(A^t)^{-1} = (A^{-1})^t$
- $A \in F^{n \times n}$ invertible $\Leftrightarrow$ rank $(A) = n$
- The set of all invertible matrices is called general linear group:

  $$GL(n, F) = \{ A \in F^{n \times n} \mid A \text{ invertible} \}$$

Change of basis

Consider the identity mapping $I : V \to V,\ x \mapsto x$.

Assume we fix a basis of V (both on source and target space), then the corr. matrix looks as follows:

$$M(I, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Now consider $\mathcal{A} = \{a_1, \dots, a_n\}$ and $\mathcal{B} = \{b_1, \dots, b_n\}$ both bases on V. How does the matrix of the id. mapping

$$I : (V, \mathcal{A}) \to (V, \mathcal{B}) \quad \text{look like?}$$

Because $\mathcal{B}$ is basis, we can write each of the vectors in $\mathcal{A}$ as lin. comb:

$$a_1 = \boxed{t_{11} b_1 + t_{21} b_2 + \dots + t_{n1} b_n}$$

$$a_2 = \dots$$

Now we form the corr. matrix T

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & & \vdots \\ t_{n1} & \cdots & t_{nn} \end{pmatrix}$$

This matrix represents the Identity:

- In the basis $\mathcal{A}$, the first basis vector a_1 has the representation $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

$a_1 = 1 \cdot a_1 + 0 \cdot a_2 + 0 \cdot a_3 \cdots + 0 \cdot a_n$

- $T \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} t_{11} \\ t_{21} \\ \vdots \\ t_{n1} \end{pmatrix}$ This vector gives us Ta_1 as expressed in basis $\mathcal{B}$

- $t_{11} b_1 + \dots + t_{n1} b_n = a_1$

- $T a_1 = a_1$.

Prop Let $\mathcal{A}, \mathcal{B}$ be two bases of V. Then the matrices $M(\mathrm{Id}, \mathcal{A}, \mathcal{B})$ and $M(\mathrm{Id}, \mathcal{B}, \mathcal{A})$ are invertible and each is the inverse of each other.

Prop Let $\mathcal{A}, \mathcal{B}$ be two bases V. Consider the transformation matrix $A = M(\mathrm{Id}, \mathcal{A}, \mathcal{B})$, and $A^{-1} = M(\mathrm{Id}, \mathcal{B}, \mathcal{A})$.

Let $T: V \to V$ linear, and $X := M(T, \mathcal{A}, \mathcal{A})$. Then $\boxed{Y := A \cdot X \cdot A^{-1}}$ represents ~~then~~ T in basis $\mathcal{B}$, that is $Y = M(T, \mathcal{B}, \mathcal{B})$.

$$\begin{tikzcd}[column sep=huge, row sep=huge]
(V, \mathcal{A}) \arrow[d, "\text{mapping } T"', "\text{matrix } X"] & (V, \mathcal{B}) \arrow[l, dashed, "\text{Id map.}"', "\text{matrix } A^{-1}"] \arrow[d, "\text{mp. } T"', "\text{matrix } Y"] \\
(V, \mathcal{A}) \arrow[r, dashed, "\text{Id map}", "\text{matrix } A"'] & (V, \mathcal{B})
\end{tikzcd}$$

Rank of a matrix

Def $A \in F^{m\times n}$. The column rank of A is

$$\dim(\operatorname{span}(\text{column vectors of } A))$$

The row rank is defined accordingly.

Prop For a matrix, the row and column rank always coincide. We now call it the <u>rank</u> of the matrix.

Prop $T \in \mathcal{L}(V, W)$. Then $\operatorname{rank}(M(T)) = \dim(\operatorname{range}(T))$.

Quotient spaces

Def Consider a set S. A subset $R \subset S \times S$ is called an equivalence relation on S if $\forall x, y, z \in S$:

(E1) $(x, x) \in R$ (reflexivity)

(E2) $(x, y) \in R \Rightarrow (y, x) \in R$ (symmetry)

(E3) $(x, y) \in R$, $(y, z) \in R \Rightarrow (x, z) \in R$ (transitivity)

Notation: $(x, y) \in R \Leftrightarrow x \sim y$

Example V VS, $W \subset V$ subspace.

$$v \sim u \;:\Leftrightarrow\; v - u \in W$$

Example: Consider the space $\mathcal{L}(\mathbb{R})$ of all functions $f: \mathbb{R} \to \mathbb{R}$ that are Lebesgue integrable. Define

$$f \sim g \;:\Leftrightarrow\; f = g \text{ almost everywhere}$$

Def The equivalence class of an element $a \in S$ under equivalence relation $\sim$ is defined as

$$[a] := \{ b \in S \mid b \sim a \}$$

Prop Two equivalence classes $[a]$ and $[b]$ are either identical or disjoint.

<u>Consequence</u> : An equ. relation on S results in a disjoint partition of equivalence classes.

<u>Constructing quotient spaces:</u>

V VS, $W \subset V$ subspace, equivalence relation

$$v \sim u \Leftrightarrow v - u \in W$$

Denote the equivalence classes as $[v]$.

Observe: the equ. classes have the form

$$[v] = v + W = \{u \in S \mid \exists w \in W : u = v + w\}$$

$$= \{v + w \mid w \in W\} \subset V$$

Define the quotient "space" as

$$V/W :\Leftrightarrow \{[v] \mid v \in V\}$$

$$[v], [u] \in V/W$$

$$[v] + [u] :\Leftrightarrow [v+u]$$

$$\lambda [v] :\Leftrightarrow [\lambda v]$$

These operations are <u>well-defined</u>:

- suppose $v' \sim v$ (i.e. $v' \in [v]$, $[v] = [v']$)

$u' \sim u$

$$[v] + [u] \overset{?}{=} [v'] + [u']$$

$v \sim v' \Leftrightarrow \exists\, w \in W \quad v - v' = w$

$u \sim u' \Leftrightarrow \exists\, \tilde{w} \in W: \; u - u' = \tilde{w}$

$[v] + [u] = [v+u]$
$[v'] + [u'] = [v'+u']$? $(v+u) \sim (v'+u')$

$(v+u) - (v'+u')$

$= \underbrace{(v-v')}_{w} + \underbrace{(u-u')}_{\tilde{w}} \in W$

- similarly, for scalar mult.

$(V/W, +, \cdot)$ is a vector space: exercise.

<u>Prop</u>: Consider $g: V \to V/W$, $v \mapsto [v]$. Then:

- g is linear
- $\ker(g) = W$
- $\operatorname{range}(g) = V/W$
- If V has finite dim, then $\dim V/W = \dim V - \dim W$.

The determinant

Def Consider a linear mapping $d: F^{n\times n} \to F$. Then d is called a <u>determinant</u> if:

(D1) d is linear in each column of the matrix:

$$\begin{pmatrix} | & | & \\ a_1 & a_2 & \cdots \\ | & | & \end{pmatrix}$$

Let A be a matrix with columns $a_1, \dots, a_n$. Consider column a_i, assume $a_i = a_i' + a_i''$ for some $a_i', a_i'' \in F^{n\times 1}$. Then it holds that

- $\det((a_1, \dots, \underline{a_i}, \dots, a_n)) = \det((a_1, \dots, \underline{a_i'}, \dots, a_n)) + \det((a_1, \dots, \underline{a_i''}, \dots, a_n))$

- $\det((a_1, \dots, \underline{\lambda a_i}, \dots, a_n)) = \lambda \cdot \det((a_1, \dots, \underline{a_i}, \dots, a_n))$

(D2) d is alternating: if A has two identical columns, then $\det A = 0$.

(D3) d is normed: $\det \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = 1$.

<u>Theorem</u>: The mapping d exists and is unique.

Based on (D1), (D2), (D3) we can now prove many important properties of the determinant:

- The determinant of an linear mapping does <u>not</u> depend on the basis.
- $\det(c \cdot A) = c^n \det(A)$
- $\det(A \cdot B) = (\det A) \cdot (\det B)$
- $\det(A^t) = \det(A)$
- $\det(A^{-1}) = 1/\det(A)$ (if A is invertible)
- A invertible $\Leftrightarrow$ $\det(A) \neq 0$
- $\det(A+B) \neq \det(A) + \det(B)$
- If A is upper triangular, that is

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

then $\det A = \lambda_1 \cdot \ldots \cdot \lambda_n$.

Explicit formulas for the determinant:

<u>Leibniz formula</u>: Denote by S_n the set of all permutations of $\{1, \ldots, n\}$. Then

$$\det A = \sum_{\underbrace{\sigma \in S_n}_{\text{all permutations}}} \underbrace{\mathrm{sign}(\sigma)}_{\text{sign of a permutation}} \, a_{\underbrace{1\sigma(1)}_{\text{position in the matrix}}} \cdot \ldots \cdot a_{n\sigma(n)}$$

Special cases:

$\underline{n=1}$ $\quad \det(a) = a$

$\underline{n=2}$
$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}\cdot a_{22} - a_{12}\cdot a_{21}$$

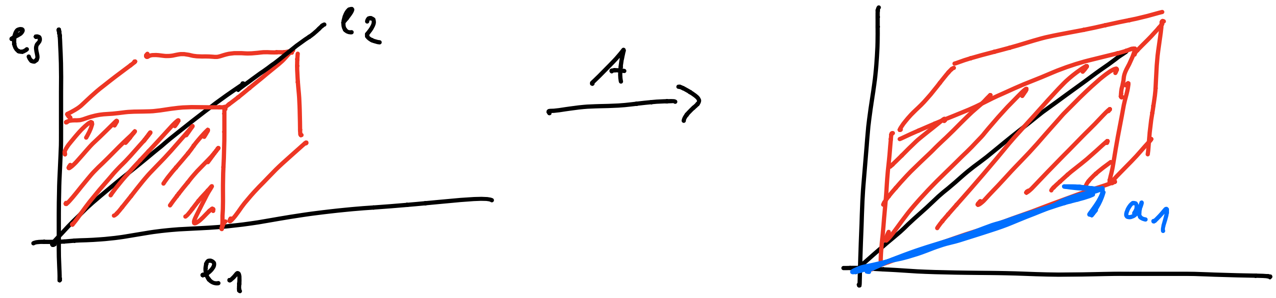
$\underline{n=3}$
$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a\cdot\det\begin{pmatrix} e & f \\ h & i \end{pmatrix} - b\cdot\det\begin{pmatrix} d & f \\ g & i \end{pmatrix} + c\cdot\det\begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

In general, there exists the <u>formula of Laplace</u> that expresses the determinant of an $n\times n$ matrix as a lin. comb. of det. of many $(n-1)\times(n-1)$-submatrices.

Geometric intuition

Consider an $n\times n$ matrix A with columns $(a_1 \,|\, a_2 \,|\, \cdots \,|\, a_n) = A$.

Consider the unit cube $U = \{c_1 e_1 + \ldots + c_n e_n \mid 0 \leq c_i \leq 1\}$



$$U \longmapsto P := \{c_1 a_1 + \ldots + c_n a_n \mid 0 \leq c_i \leq 1\}$$ parallelotope.

$\boxed{\text{Then } \det(A) \text{ gives us the (signed) volume of } P.}$

<u>intuition</u>:

vol $(U) = 1$ $\qquad\qquad$ vol$(P) = \lambda_1 \cdot \lambda_2 \cdot \lambda_3$

$\det(A)$ = product of eigenvalues $\lambda_1 \cdot \lambda_2 \cdot \lambda_3$

vol (U) changes by $\lambda_1 \lambda_2 \lambda_3$

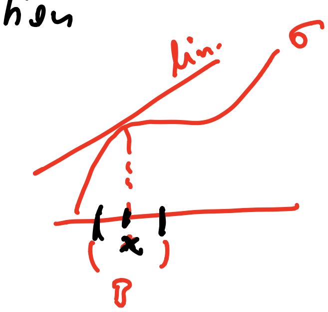
<u>Application to integrals</u>:

<u>Proposition</u>: $\Omega \subset \mathbb{R}^n$ open subset, $\sigma : \Omega \to \mathbb{R}^n$ differentiable, $f : \sigma(\Omega) \to \mathbb{R}$. Then:

derivative, linear

$$\int_{\sigma(\Omega)} f(y) \underbrace{dy}_{\text{volume element}} = \int_{\Omega} f(\sigma(x)) \underbrace{|\det(\sigma'(x))| \, dx}_{\text{volume element}}$$

<u>Intuition</u>: σ differentiable, that is we can locally (on a small ball B around x) approximate σ by a linear function



$$\sigma'(x) = \begin{pmatrix} \frac{\partial \sigma_1}{\partial x_1} & \cdots & \frac{\partial \sigma_n}{\partial x_1} \\ \vdots & & \\ \frac{\partial \sigma_n}{\partial x_1} & \cdots & \frac{\partial \sigma_n}{\partial x_n} \end{pmatrix}$$

$$\mathrm{vol}\ \sigma(B) \approx \mathrm{vol}\left((\sigma'(x)) \cdot B\right)$$

$$\approx |\det(\sigma'(x))| \cdot \mathrm{vol}(B)$$

Substitution: $y = \sigma(x)$

$$f(y) \cdot \underbrace{\mathrm{vol}(\sigma(B))}_{dy} \approx f(\overset{y}{\widetilde{\sigma(x)}}) \cdot |\det(\sigma'(x))| \cdot \underbrace{\mathrm{vol}(B)}_{dx}$$

$$\int f(y)\, dy \approx \int f(\sigma(x))\, |\det \sigma'(x)|\, dx$$

Eigenvalues

Def Let $T: V \to V$. A scalar $\lambda \in F$ is called an <u>eigenvalue</u> if there exists a $v \in V$, $v \neq 0$, such that $Tv = \lambda \cdot v$. A vector $v \neq 0$ with this property is called an <u>eigenvector</u> corresponding to eigenvalue λ. The set of all eigenvectors of λ is called the <u>eigenspace</u> $E(\lambda, T) = \ker(T - \lambda I)$.

$Tv = \lambda v$

$Tv - \lambda v = 0$

$Tv - \lambda I v = 0$

$(T - \lambda I) v = 0$

<u>Remarks</u>

- Eigenvalue/eigenvector realizes a "streching"
  $v \mapsto \lambda v$
- Many mappings do not have eigenvectors for example, a rotation.
- If λ is an eigenvalue, it has many eigenvectors!
  For example, if v is eigenvector, then also $a \cdot v$ $(a \in K)$ is an eigenvector!

$$T(\underbrace{a \cdot v}) = a \cdot T(v) = a \cdot \lambda \cdot v = \lambda \, (\underbrace{a \cdot v})$$

- Eigenvectors corr. to <u>distinct</u> eigenvalues are linearly independent.

Intuition: λ_1, λ_2 two eigenvalues, $\lambda_1 \neq \lambda_2$

Assume v_1, v_2 are eigenvectors that are not lin. independent: $v_2 = c \cdot v_1$

$$T v_1 = \lambda_1 v_1$$

$$T v_2 = \lambda_2 v_2 = \lambda_2 (c \cdot v_1) \neq$$

because $\lambda_1 \neq \lambda_2$

$$T v_2 = T(c \cdot v_1) = c \cdot T v_1 = c \cdot \lambda_1 v_1$$

$$\lightning$$

- Eigenvectors that corr. to the same eigenvalue do not need to be independent

  Simple example: v eig. $\Rightarrow$ $c \cdot v$ eig. but v and $c \cdot v$ are not lin. ind.

  They can be lin. independent:

  Easy example: $A = I$, then every vector v is an eigenvector of eigenvalue 1.

  $$I \cdot v = 1 \cdot v$$

- The eigenspace $E(\lambda, T)$ is always a lin. subspace of V.

Prop For finite-dim VS, the following statements are equivalent:

(i) λ eigenvalue of T

(ii) $T - \lambda I$ not injective

(iii) $T - \lambda I$ not surjective

(iv) not bijective.

Prop Suppose V is finite-dim, $T \in \mathcal{L}(V)$, and $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T. Then a sum of eigenspaces

$$E(\lambda_1, T) + E(\lambda_2, T) + \dots + E(\lambda_m, T)$$

is a direct sum. In particular

$$\dim(E(\lambda_1, T)) + \dots + \dim(E(\lambda_m, T)) \leq \dim V$$

Theorem: Every operator $T: V \to V$ on a finite-dim, complex vs V has at least one eigenvalue.

Proof Let $n = \dim V$. Choose a vector $v \in V$, $v \neq 0$. Then the set

$$v, Tv, T^2 v, \dots, T^n v$$

has to be linearly dependent (it consists of $n+1$ vectors in an n-dim space). Find coefficients $a_0, a_1, \dots, a_n$ such that

$$a_0 v + a_1 Tv + \dots + a_n T^n v = 0.$$

Now consider a polynomial on $\mathbb{C}$ with these coefficients:

$$p(z) := a_0 + a_1 \cdot z + \dots + a_n z^n$$

Over $\mathbb{C}$, we can factorize it:

n and m can be different

$$p(z) = c \cdot (z - \lambda_1)(z - \lambda_2) \dots (z - \lambda_m)$$

Hence, $0 = a_0 v + a_1 Tv + \dots + a_n T^n v =$

$$= \underbrace{(a_0 + a_1 T + \dots + a_n T^n)}_{c \cdot (T - \lambda_1 I)(T - \lambda_2 I) \dots (T - \lambda_m I)} v$$

$$= c(T - \lambda_1 I)(T - \lambda_2 I) \dots (T - \lambda_m I) \cdot v$$

$\Rightarrow v \in \ker($ big operator $)$

$\Rightarrow$ there must exist $i \in \{0, \dots, m\}$ such that

$(T - \lambda_i I)$ not injective

$\Rightarrow \lambda_i$ is an eigenvalue of T !

$\blacksquare$

Characteristic polynomial

Motivation: $Av = \lambda v$ 　 A $n \times n$-matrix, $v \neq 0$

$$\Leftrightarrow (A - \lambda I) v = 0$$

$$\Leftrightarrow v \in \ker(A - \lambda I)$$

$$\Leftrightarrow \operatorname{rank}(A - \lambda I) < n$$

$$\Leftrightarrow \det(A - \underline{\underline{\lambda}} I) = 0$$

Def The characteristic polynomial of an $n \times n$-matrix A is defined as

$$p_A(t) := \det(A - t \cdot I)$$

Example: $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

$$\det(A - t \cdot I) = \det\left(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - t \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

$$= \det \begin{pmatrix} a_{11} - t & a_{12} \\ a_{21} & a_{22} - t \end{pmatrix}$$

$$= (a_{11} - t)(a_{22} - t) - a_{12} \cdot a_{21}$$

$$= t^2 + t(-a_{11} - a_{22}) - a_{12} \cdot a_{21} + a_{11} \cdot a_{22}$$

Observations

- $p_A(t)$ is a polynomial with degree n

- Char. pol. does <u>not</u> depend on the basis:

  <u>Proof</u> Consider A, basis transformation matrix U.
  Want to look at char. pol. of UAU^{-1}.

$$\det(UAU^{-1} - t\cdot I)$$
$$= \det(UAU^{-1} - t\cdot U\cdot U^{-1})$$
$$= \det(U(A - t\cdot I)U^{-1})$$
$$= \det(U)\cdot\det(A - t\cdot I)\cdot\det(U^{-1})$$
$$= \det(A - tI)$$

- The roots of the characteristic poly. correspond exactly to the eigenvalues of A.

- Over $\mathbb{C}$, the char. poly. always has n roots, so the matrix has "n eigenvalues" (not nec. distinct).

- A is invertible $\Leftrightarrow$ 0 is <u>not</u> an eigenvalue.

  If 0 is an eigenvalue, ~~e~~ e.v v with
  $Av = 0\cdot v = 0$
  $\Leftrightarrow$ $\ker(A)$ ~~n~~non-trivial $\Leftrightarrow$ A not invertible

- Let $A \in \mathcal{L}(V)$, λ eig. of A. Then λ^k is an eig. of A^k.

- Let A be invertible; λ eig of A. Then $1/\lambda$ is an eig. of A^{-1}.

__Def__ For an operator A with eigenvalue λ, we define its __geometric multiplicity__ as the dimension of the corr. eigenspace $E(\lambda, A)$.
The __algebraic multiplicity__ is the multiplicity of the root λ in the char. poly.

In general, the two notions __do not__ coincide.

Computing eigs

- Write down the char. pol., find the roots.
  $\rightsquigarrow$ eigenvalues

- To compute the eigenvectors, solve the lin. system
  $$Ax = \lambda x$$

Trace of a matrix

<u>Def</u> The trace of a square matrix $A \in F^{n \times n}$ is the sum of its diagonal elements:

$$tr(A) = \sum_{i=1}^{n} a_{ii} .$$

<u>Remarks</u>:

- $tr: \mathbb{R}^{n \times n} \to \mathbb{R}$ is a linear operator

  In particular, $tr(A+B) = tr(A) + tr(B)$.

- $tr(A \cdot B) = tr(B \cdot A)$

  ⚠ $tr(A \cdot B) \neq tr(A) \cdot tr(B)$

- trace does <u>not</u> depend on the basis:

  Let $T \in \mathcal{L}(V)$, and U and W two bases of V. Then:

  $tr(M(T, U)) = tr(M(T, W))$.

- The trace of an operator equals the sum of its complex eigenvalues, counted according to multiplicity:

$$\tilde{A} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$ wrt some basis $v_1, \dots, v_n$

$$\Rightarrow tr(\tilde{A}) = \sum_{i=1}^{n} \lambda_i$$

Curious little fact: Over $\mathbb{C}$, we can always find basis of eigenvectors: $A \in \mathbb{R}^{n \times n}$, over $\mathbb{C}$ I can find representation

$$\tilde{A} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad \lambda_i \in \mathbb{C}$$

$$tr(\tilde{A}) = \underbrace{\sum \lambda_i}_{\in \mathbb{C}} = \sum_{i=1}^{n} \underbrace{a_{ii}}_{\in \mathbb{R}} = \underbrace{tr(A)}_{\in \mathbb{R}}$$

indep. of base

$$\Rightarrow \sum_i \lambda_i \in \mathbb{R}$$

- trace equals the negative of the coefficient in front of t^{n-1} in the char. polynomial

$$p_A(t) = t^n + a_{n-1} t^{n-1} + \dots +$$

- $tr(A) =$ sum of eigenvalues (if exist)

  $\det(A) =$ product of eigenvalues (if exist)

Example: Consider a rotation matrix

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

- A does not have any real eigenvalues.

- The trace is given as $2 \cdot \cos\theta$.
- The char. polyn. of A is

$$p(t) = \det(A - tI) = \det \begin{pmatrix} (\cos\theta) - t & -\sin\theta \\ \sin\theta & (\cos\theta) - t \end{pmatrix}$$

$$= (\cos\theta - t)^2 + \sin^2\theta$$

$$= t^2 - 2\cos\theta \cdot t + \underbrace{\cos^2\theta + \sin^2\theta}_{=1}$$

$$= t^2 - (2\cos\theta) \cdot t + 1$$

$$4\cos^2\theta - 4$$
$$= 4(\cos^2\theta - 1)$$
$$= 4(-\sin\theta)$$

- The roots of the char. pol.

$$\lambda_{1/2} = \frac{2\cos\theta \pm \sqrt{(2\cos\theta)^2 - 4}}{2}$$

$$= \cos\theta \pm i \cdot \sin\theta$$

- The matrix has a diagonal representation

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$\mathrm{tr}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \cos\theta + i\sin\theta + \cos\theta - i\sin\theta$$

$$= 2\cos\theta$$

Diagonalization

Def An operator $T \in \mathcal{L}(V)$ is <u>diagonalizable</u> if there exists a basis of V such that the corr. matrix is diagonal:

$$M(T) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Prop V finite-dim, $A \in \mathcal{L}(V)$. Then the following statements are equivalent:

(i) A is diagonalizable.

(ii)
- The char. pol. p_A can be decomposed into linear factors

  AND

- The algebraic multiplicity of the roots of p_A are equal to the geometric multiplicities.

(iii) If $\lambda_1, \dots, \lambda_k$ are the pairwise distinct eigenvalues of A, then

$$V = \operatorname{eig}(A, \lambda_1) \oplus \dots \oplus \operatorname{eig}(A, \lambda_k).$$

Triangular matrices

A matrix is called upper triangular, if it has the form

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Prop $T \in \mathcal{L}(V)$, $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ a basis.
Then equivalent:

(a) $M(T, \mathcal{B})$ is upper triangular.

(b) $T v_j \in \text{span}\{v_1, \dots, v_j\} \quad \forall j = 1, \dots, n$

$$Tv_1 = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} \\ & \lambda_2 & a_{23} \\ 0 & & \lambda_3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix} = \lambda_1 \cdot v_1$$

$$Tv_2 = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} \\ & \lambda_2 & a_{23} \\ 0 & & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{12} \\ \lambda_2 \\ 0 \end{pmatrix} = a_{12} \overset{v_1}{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}} + \lambda_2 \overset{v_2}{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}$$

$\in \text{span}(v_1, v_2)$

Prop V complex, finite-dim VS, $T \in \mathcal{L}(V)$. Then
$M(T)$ has an upper triangular form for some basis.

Prop Suppose $T \in \mathcal{L}(V)$, V any finite-dim VS, has an upper triangular form. Then the entries on the diagonal are precisely the eigenvalues of T.

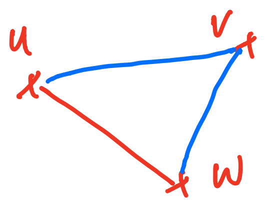
Metric space

<u>Definition</u>: Let X be a set. A function $d: X \times X \to \mathbb{R}$ is called a metric if the following conditions hold:
$\forall u, v, w \in X$

(1) $d(x,y) > 0$ if $x \neq y$ and
$d(x,x) = 0$

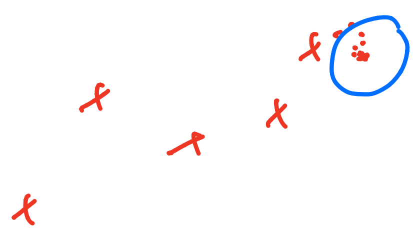
(2) $d(x,y) = d(y,x)$ (symmetry)

(3) $d(u,v) + d(v,w) \geq d(u,w)$



<u>Def</u> A sequence $(x_n)_{n \in \mathbb{N}}$ in a metric space (X, d) is called a Cauchy sequence if

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n, m > N \quad d(x_n, x_m) < \varepsilon$$



A sequence $(x_n)_n$ converges to $x \in X$ if

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n > N, \ d(x_n, x) < \varepsilon$$

Notation: $x_n \to x$, $\lim_{n\to\infty} x_n = x$

Sequence $(x_n)_n$, $x_n = \frac{1}{n}$ on $X =]0,1[$

Here $(x_n)_n$ is a Cauchy sequence, but does not converge.

Sequence $(x_n)_n$, $x_n = \frac{1}{n}$ on $\tilde{X} = [0,1]$.

Here, (x_n) is a Cauchy sequence that converges to 0.

Def A metric space is called complete if every Cauchy-sequence converges.

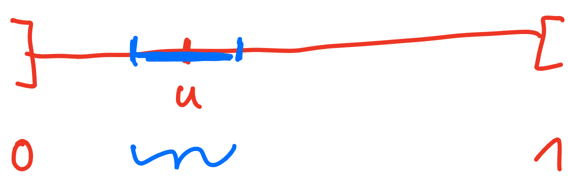
Notation: $B_\varepsilon(u) := \{ x \in X \mid d(x,u) < \varepsilon \}$ ball

Def A set $U \subset X$ is called closed if all Cauchy-sequences converge and have their limit point in U.

A set $U \subset X$ is called open if

$\forall u \in U \ \exists \varepsilon > 0 : \ B_\varepsilon(u) \subset U$.

- Set $[0,1]$ is closed
- set $]0,1[$ is open:



$B =]u-\varepsilon, u+\varepsilon[$

- A set U can be neither open nor closed:

$[0,1[$

Def A point $u \in U$ is an *interior point* of U if there exists a $\varepsilon > 0$ s.th. $B_\varepsilon(u) \subset U$.

$U = [0,1]$, then $x \in]0,1[$ are interior pts

The (topological) *closure* of a set U is defined as the set of points that can be approximated by Cauchy sequences in U:

$$w \in \overline{U} \iff \forall \varepsilon > 0 \; \exists z \in U : d(w,z) < \varepsilon$$

Notation: $\overline{U}$ is the closure of U.

The (topological) *interior* of a set U is defined as the set of interior points of U.

Notation: U°

The (topological) *boundary* of a set U is defined as the set $\overline{U} \setminus U^\circ$

literature not always consistent here ... sometimes one also reads $U \setminus U^\circ$ instead of $\overline{U} \setminus U^\circ$.

$X = [0,1[$

$\overline{X} = [0,1]$

$X^\circ =]0,1[$

$\Rightarrow \text{boundary}_1(X) = \overline{X} \setminus X^\circ = \{0,1\}$

$(\text{boundary}_2(X) = X \setminus X^\circ = \{0\})$

Def A set U is dense in X if we can approximate every $x \in X$ by a sequence in U. Formally,

$$\forall x \in X \quad \forall \varepsilon > 0 \quad B_\varepsilon(x) \cap U \neq \emptyset$$

Example: $\mathbb{Q} \subset \mathbb{R}$ is dense.

Def A set $U \subset X$ is bounded if there exists $D > 0$ such that $\forall u, v \in U$, $d(u,v) < D$

Normed spaces

Def Let V be a vector space. A norm on V is a function $\|\cdot\| : V \to \mathbb{R}$ such that $\forall x, y \in V$, $\lambda \in F$ the conditions are true:

(N1) $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$ (homogeneous)

(N2) $\|x+y\| \leq \|x\| + \|y\|$ (triangle inequality)

(N3) $\|x\| = 0 \Leftarrow x = 0$

(N4) $\|x\| = 0 \Rightarrow x = 0$

$\|\cdot\|$ is a semi-norm if (N1) - (N3) are satisfied.

Intuition norm (x) = "length of x"
= distance $(x, 0)$

Examples Euclidean norm on $\mathbb{R}^d$: $\|x\| = \left(\sum_{i=1}^{d} x_i^2\right)^{1/2}$

p - Norms

Consider $V = \mathbb{R}^d$. Define $\|\cdot\|_p : \mathbb{R}^d \to \mathbb{R}$,

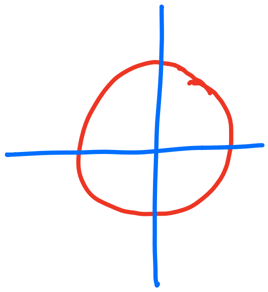
$$\|x\|_p := \left(\sum_{i=1}^{d} |x_i|^p\right)^{1/p} \quad \text{for } 0 < p < \infty$$

- $\|\cdot\|_p$ is a norm if $p \geq 1$

- Unit balls: the unit ball of a norm is the set of points such that norm ≤ 1:

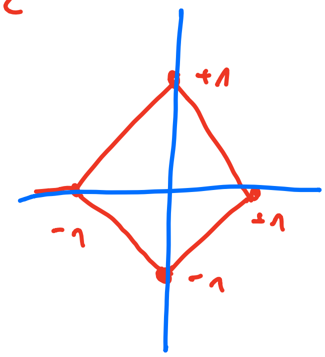
$$B_p := \{ x \in \mathbb{R}^d \mid \|x\|_p \leq 1 \}$$

Example: $\mathbb{R}^2$



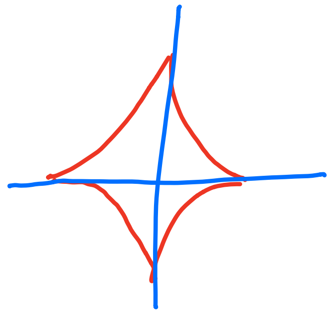
$p = 2$

(ball convex)



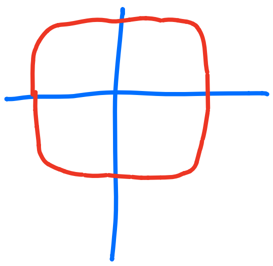
$p = 1$

(convex)

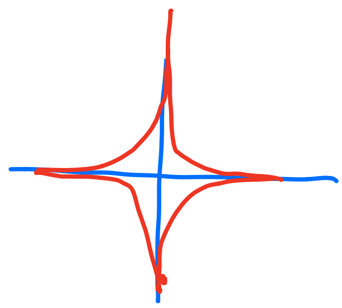


$p = 0.5$

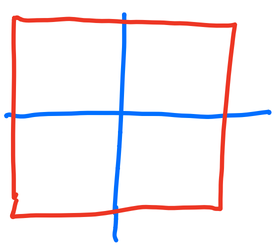
(not convex)



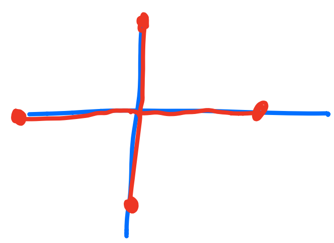
$p = 5$



$p = 0.1$



$p = \infty$



$p = 0$

Def : $\|x\|_\infty := \max |x_i|$ (is a norm)

$\|x\|_0 :=$ number of non-zero coordinates

$= \sum_{i=1}^{d} \mathbb{1}\{x_i \neq 0\}$

⚠ $\|x\|_0$ is not a norm

Equivalent norms

Theorem All norms on $\mathbb{R}^n$ are (topologically) equivalent:

If $\|\cdot\|_a$ and $\|\cdot\|_b : \mathbb{R}^n \to \mathbb{R}$ are two norms on $\mathbb{R}^n$,

then there exist constants $\alpha, \beta > 0$ such that

$$\forall x \in \mathbb{R}^n : \quad \alpha \|x\|_a \leq \|x\|_b \leq \beta \|x\|_a \qquad (*)$$

Proof: W.l.o.g. we prove that if $\|\cdot\|$ is any norm on $\mathbb{R}$, then it is equivalent to $\|\cdot\|_\infty$ on $\mathbb{R}^d$.

$$\boxed{\text{First inequality: } \exists c_1 > 0 : \forall x \quad \|x\| \leq c_1 \|x\|_\infty}$$

Let $x = \sum x_i e_i$ the representation of x in the standard basis of $\mathbb{R}^d$.

$$\|x\| = \Big\| \sum_{i=1}^d x_i e_i \Big\|$$

$$\leq \sum_i \| x_i e_i \|$$

$$= \sum_i \underbrace{|x_i|}_{\leq \|x\|_\infty} \|e_i\|$$

$$\leq \sum_i \|x\|_\infty \cdot \|e_i\|$$

$$= \|x\|_\infty \cdot \underbrace{\sum_i \|e_i\|}_{=: c_1}$$

$$\boxed{\underline{\text{Second inequality}}: \ \exists c_2 > 0 \ \forall x \ \|x\|_\infty \leq c_2 \cdot \|x\|}$$

Let $S := \{x \in \mathbb{R}^d \mid \|x\|_\infty = 1\}$ be the unit sphere wrt $\|\cdot\|_\infty$

Consider $f: S \to \mathbb{R}$, $x \mapsto \|x\|$.

The mapping f is continuous wrt $\|\cdot\|_\infty$:

(this follows directly from the fact that

$$\begin{aligned} |f(x) - f(y)| &= \big| \|x\| - \|y\| \big| \\ &\leq \|x - y\| \leq c_1 \cdot \|x - y\|_\infty \end{aligned}$$).

The S is closed and bounded, thus by Theorem of Heine-Borel, S compact. Any continuous mapping on a compact set takes its min and max.

$$\tilde{c}_2 := \min\{f(x) \mid x \in S\}$$

$x \in S$: $\quad \underbrace{\|x\|}_{\geq \tilde{c}_2} = \left\| \frac{x}{1} \right\| = \left\| \frac{x}{\|x\|_\infty} \right\| = \frac{\|x\|}{\|x\|_\infty}$

$$\Rightarrow \quad \tilde{c}_2 \leq \frac{\|x\|}{\|x\|_\infty}$$

$$\Rightarrow \quad \|x\|_\infty \leq \underbrace{\frac{1}{\tilde{c}_2}}_{c_2 := \frac{1}{\tilde{c}_2}} \|x\|$$

$$\|x\|_\infty \leq c_2 \cdot \|x\| .$$

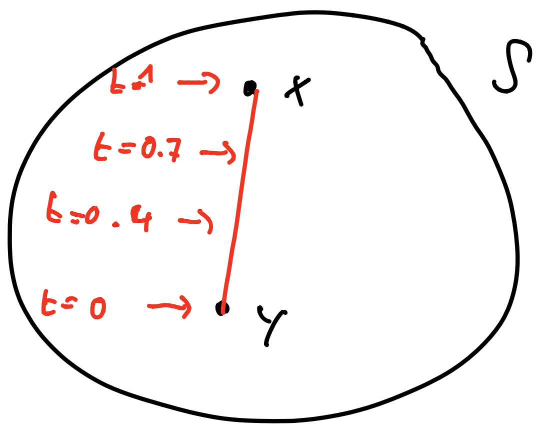
$\blacksquare$

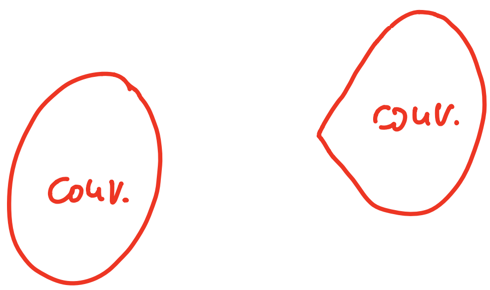
Convex sets – unit balls of norms

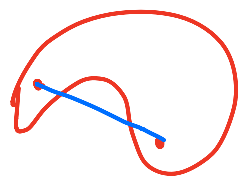
Def Consider a real VS V, $S \subset V$. S is called <u>convex</u> if $\forall t$, $0 \le t \le 1$ and $\forall x, y \in S$,

$$t \cdot x + (1-t) y \in S$$

<u>Intuition</u>

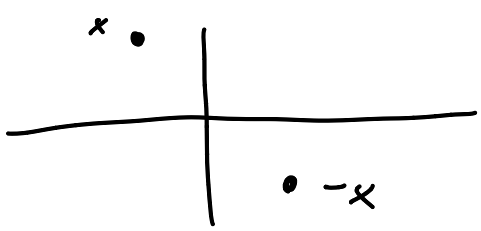


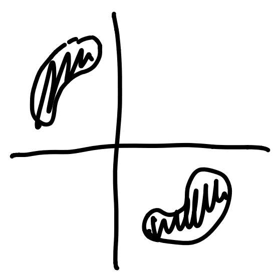




Def A set $C \subset V$ is called symmetric if

$$x \in C \Rightarrow -x \in C$$





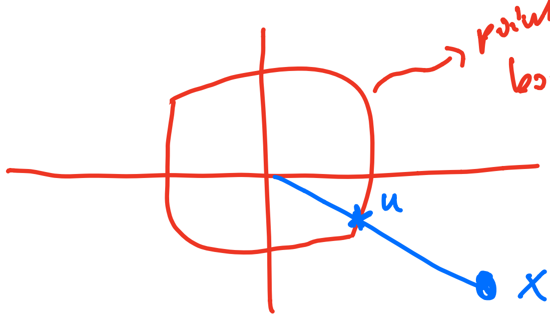
Theorem: (1) Let $C \subset \mathbb{R}^d$ closed, convex, symmetric and has non-empty interior. Define

$$p(x) := \inf \left\{ t > 0 \;\middle|\; \frac{x}{t} \in C \right\}. \text{ Then}$$

$= \inf \{ t > 0 \mid x \in t \cdot C \}$, this might be more intuitive.

p is a semi-norm. If C is bounded, then p is a norm, and its unit ball coincides with C (that is, $C = \{ x \in \mathbb{R}^d \mid p(x) \leq 1 \}$

(2) For any norm $\|\cdot\|$ on $\mathbb{R}^d$, the set $C := \{ x \in \mathbb{R}^d \mid \|x\| \leq 1 \}$ is bounded, symmetric, closed, convex, and has non-empty interior.



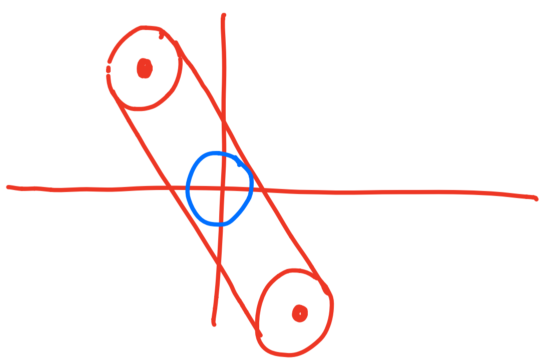
Proof $\boxed{p(x) \text{ is well defined}}$

Want to prove: given $x \in \mathbb{R}^d$, the set $\{ t > 0 \mid \frac{x}{t} \in C \} \neq \emptyset$

We are going to prove: $\exists\, \varepsilon > 0$ such that

$$B_\varepsilon(0) = \{ e \in \mathbb{R}^n \mid \|e\|_2 < \varepsilon \} \subset C.$$

Intuition:



- By ass, C has at least one interior point
  $v \in C^{\circ} \Rightarrow \exists\, \varepsilon$ such that

  $$B_{\varepsilon}(v) \subset C$$
  $$\overset{\shortparallel}{v + B_{\varepsilon}(0)} = \{ v + e \mid e \in B_{\varepsilon}(0) \}$$

- By symmetry, $v + e \in C \Rightarrow -(v+e) \in C$
- By convexity, $\frac{1}{2}(v+e) + \frac{1}{2}(-v-e) = e \in C$

So $B_{\varepsilon}(0) \subset C$, so the set $\{ t > 0 \mid \frac{x}{t} \in C \}$ is non-empty.

The infimum of $\inf \overbrace{\{ t > 0 \mid \frac{x}{t} \in C \}}^{S_x}$ exists because $S_x \subset \mathbb{R}$, 0 is a lower bound.

Now we need to prove all axioms of a norm:

$$\boxed{p(0) = 0}$$

- Have seen: $0 \in C$
- $\forall t > 0 : \frac{0}{t} = 0 \in C$
- $\inf \{ t \mid \frac{0}{t} \in C \} = 0.$

$$\Rightarrow p(0) = 0$$

$$\boxed{p(\alpha x) = |\alpha|\, p(x)}$$

- For all $\alpha > 0$ we have

$$p(\alpha \cdot x) = \inf\left\{ t > 0 \;\middle|\; \frac{\alpha x}{t} \in C \right\} \overset{s := \frac{t}{\alpha}}{=}$$

$$= \inf\left\{ \alpha \cdot s > 0 \;\middle|\; \frac{x}{s} \in C \right\}$$

$$= \alpha \cdot \underbrace{\inf\left\{ s > 0 \;\middle|\; \frac{x}{s} \in C \right\}}_{p(x)}$$

$\Rightarrow p(\alpha x) = \alpha \cdot p(x)$

- By symmetry we also get

$$p(-x) = \inf\left\{ t > 0 \;\middle|\; \frac{-x}{t} \in C \right\} = \quad -\frac{x}{t} \in C \Rightarrow \frac{x}{t} \in C$$

$$= \inf\left\{ t > 0 \;\middle|\; \frac{x}{t} \in C \right\} = p(x)$$

- Combining the two statements gives homogeneity.

Δ-inequality Consider $x, y \in \mathbb{R}^d$, $s, t > 0$ such that

$\frac{x}{s} \in C$, $\frac{y}{t} \in C$.

Observe: $\frac{s}{s+t} + \frac{t}{s+t} = 1$. Thus, by convexity,

$$\underbrace{\frac{s}{s+t}}\cdot \underbrace{\frac{x}{s}}_{\in C} + \underbrace{\frac{t}{s+t}}\cdot \underbrace{\frac{y}{t}}_{C} \in C \quad (*)$$

two scalars that sum up to 1

$\boxed{\text{Want to prove:}}$

$$p(x+y) = \inf\left\{u > 0 \;\middle|\; \frac{x+y}{u} \in C\right\}$$

$$\leq \underbrace{\inf\left\{s > 0 \;\middle|\; \frac{x}{s} \in C\right\}}_{p(x)} + \underbrace{\inf\left\{t > 0 \;\middle|\; \frac{y}{t} \in C\right\}}_{p(y)}$$

$$\frac{s}{s+t}\,\frac{x}{s} + \frac{t}{s+t}\,\frac{y}{t} \in C$$

$$\frac{x+y}{\underbrace{s+t}_{=u_0}} \in C$$

$$p(x+y) = \inf\left\{u>0 \;\middle|\; \frac{x+y}{u} \in C\right\} \leq u_0$$

$$= \underbrace{s}_{p(x)} + \underbrace{t}_{p(y)}$$

s was chosen such that $\frac{x}{s} \in C$

t $\frac{y}{t} \in C$

Consider a sequence $(s_i)_{i\in\mathbb{N}}$ such that

$$\frac{x}{s_i} \in C \quad \text{and} \quad s_i \to p(x)$$

Similarly $(t_i)_{i\in\mathbb{N}}$ such that $\frac{y}{t_i} \in C$ and $t_i \to p(y)$.

By the argument above, we know that

$$\forall i: \quad p(x+y) \leq \underbrace{s_i}_{\to p(x)} + \underbrace{t_i}_{\to p(y)}$$

$$\Rightarrow \quad p(x+y) \leq p(x) + p(y).$$

$$\boxed{p(x) = 0 \Rightarrow x = 0}$$

$$p(x) = 0 \Leftrightarrow \inf \left\{ t > 0 \;\middle|\; \frac{x}{t} \in C \right\} = 0$$

$\Rightarrow$ There exists a sequence $(t_k)_{k \in \mathbb{N}}$ such that

$t_k \to 0$ and $\frac{x}{t_k} \in C \quad \forall k.$

Now assume that $x \neq 0$. Then the sequence

$\left(\frac{x}{t_k}\right)_{k \in \mathbb{N}}$ is unbounded. $\lightning$ Contradiction because C is bounded.

$\blacksquare$

Examples of normed function spaces

Space of continuous fcts:

Let T be a metric space,

$$\mathcal{C}^b(T) := \{ f : T \to \mathbb{R} \mid f \text{ is continuous and bounded} \}$$

$\exists c \in \mathbb{R} : \forall t \in T : |f(t)| < c$

As norm on $\mathcal{C}^b(T)$ we now use

$$\|f\|_\infty := \sup_{t \in T} |f(t)|$$

Then the space $\mathcal{C}^b(T)$ with norm $\|\cdot\|_\infty$ is a Banach space.

Proof outline:

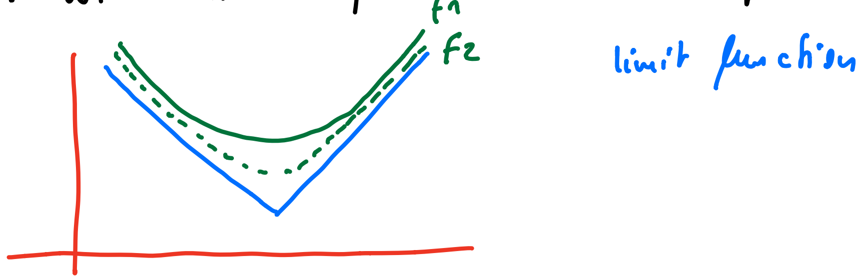
- need to check vector space axioms
- norm axioms
- completeness: follows from the fact that $\|\cdot\|_\infty$ induces uniform convergence

Space of differentiable functions:

Let $[a,b] \subset \mathbb{R}$, $\mathcal{C}^1([a,b]) = \{ f : [a,b] \to \mathbb{R} \mid f \text{ is cont. differentiable} \}$

Which norm?

- Consider $\|\cdot\|_\infty$. With this norm, $\mathcal{C}^1$ is not complete.



- Consider $\|f\| := \sup_{t\in[a,b]} \max\{ |f(t)|, |f'(t)| \}$

$$|||f||| := \|f\|_\infty + \|f'\|_\infty$$

$\mathcal{C}^1([a,b])$ with any of these two norms is a Banach space.

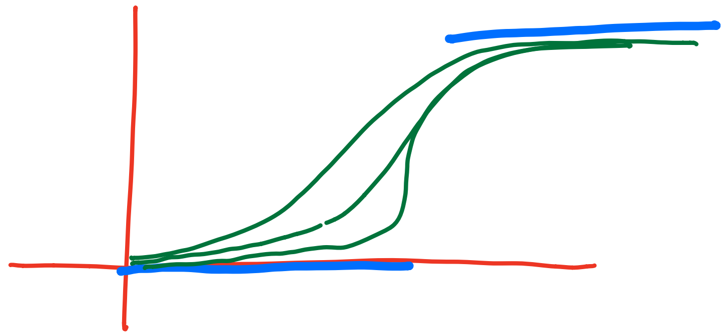
Constructing L_p-spaces

Consider $\mathcal{C}^b([a,b])$ with the norm

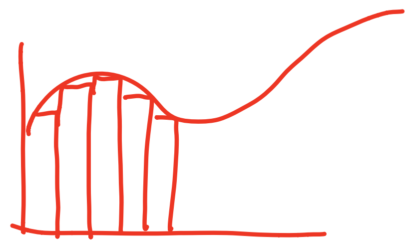
$$\|f\|_1 := \int_a^b |f(t)|\, dt$$

Can see: $\|\cdot\|_1$ is a norm,

but the space is not complete.



Consider $R([a,b])$ of all Riemann-integrable functions on $[a,b] \subset \mathbb{R}$, together with $\|\cdot\|_1$.



However, on $R([a,b])$, $\|\cdot\|_1$ is not a norm: it is not true that

$$\|f\| = 0 \Rightarrow f = 0$$

$\int f\, dt = 0$

but $f \not\equiv 0$

$\mathcal{L}_p$ - space

$$\mathcal{L}_p([a,b]) := \Big\{ f : [a,b] \to \mathbb{R},\ f \text{ measurable},\ \int |f|^p \, d\lambda < \infty \Big\}$$

$\int |f(t)|^p \, dt$

for $1 \leq p < \infty$

$$\|f\|_p := \left(\int |f|^p \, d\lambda \right)^{1/p}$$

<u>Proposition 1</u> : $\|f\|_p$ is a semi-norm on $\mathcal{L}_p$.

Proof : • Vector space (clear)

• semi-norm : observe: $\|f\|_p = 0 \Rightarrow f = 0$ almost everywhere

so we do <u>not</u> have that $\|f\|_p = 0 \Rightarrow f = 0$. a.e.

<u>Proposition 2</u> $\mathcal{L}_p$ is complete under $\|\cdot\|_p$.

<u>Proof</u> : If $(f_i)_{i \in \mathbb{N}}$ is a Cauchy-sequence in $\mathcal{L}_p$, then want to prove that $\lim_{i \to \infty} f_i \in \mathcal{L}_p$.

This is equivalent to proving the following:

Let $(f_i)_i$ be a sequence such that

$$a := \sum_{i=1}^{\infty} \|f_i\|_p < \infty \qquad (*)$$

then there exists $f \in \mathcal{L}_p$ such that $f_i \to f$ (in $\|\cdot\|_p$).

Define

$$\hat{g} := \sum_{i=1}^{\infty} |f_i|$$

Note: this might not yet be a well-defined fct from $[a,b]$ to $\mathbb{R}$, might be ∞ at certain points.

$$\hat{g}_n := \sum_{i=1}^{n} |f_i| \in \mathcal{L}_p$$

By Minkowski,

$$\|\hat{g}_n\|_p \overset{\text{def}}{=} \Big\| \sum_{i=1}^{n} |f_i| \Big\|_p \overset{\text{Mink.}}{\leq} \sum_{i=1}^{n} \|f_i\|_p \overset{\text{ass. *}}{<} a$$

$\hat{g}_n \to g$ monotonously

By theorem of monotonic convergence, g is measurable and we have

$$\lim_{n\to\infty} \int \hat{g}_n^{\,p} \, d\lambda = \int \lim \hat{g}_n^{\,p} \, d\lambda$$

$$\overset{\text{by def}}{=} \int \hat{g}^p \, d\lambda$$

$$\leq a^p .$$

$\Rightarrow$ $\hat{g} < \infty$ a.e., that is there exists a set N of measure 0 such that on $[a,b] \setminus N$, $\hat{g}$ is finite.

Now we can define

$$g(t) = \begin{cases} \hat{g}(t) & t \in [a,b] \setminus N \\ 0 & t \in N \end{cases}$$

$$\in \mathcal{L}_p$$

From this it now follows that $f(t) = \sum_{i=1}^{\infty} f_i(t)$, $t \notin N$ exists. For $t \in N$, we set $f(t) = 0$.

Now f is measurable, and in $\mathcal{L}_p$

because $\int |f|^p \, d\lambda \leq \int \hat{g}^p \, d\lambda < \infty$

Finally, $\sum^{\infty} f_n$ converges to f in $\|\cdot\|_p$ because of the theorem of dominated convergence. ∎

From $\mathcal{L}_p$ to L_p

We constructed a space $\mathcal{L}_p$ with the Lebesgue integral as a semi-norm. This means,

given $f \in \mathcal{L}_p$, we can change the values of f in a set of measure 0, resulting in $\tilde{f}$, but the norm "does not see a difference":

$$\| f - \tilde{f} \| = 0$$

Define equivalence relation

$$f \sim \tilde{f} \;:\Longleftrightarrow\; f = \tilde{f} \text{ a.e.}$$

Formally: $N := \ker(\|\cdot\|_p) := \{ f \in \mathcal{L}_p \mid \|f\|_p = 0 \}$ is a subspace of $\mathcal{L}_p$.

$$L_p([a,b]) := \mathcal{L}_p([a,b]) / N$$

Elements ("functions") in L_p are equivalence classes $[f]$ consisting of all functions that coincide a.e.

⚠ It does not make sense to evaluate $f(0)$ because $\{0\}$ has Lebesgue measure 0.

Define a norm on L_p by

$$\underbrace{\| [f] \|_p}_{\text{new}} = \underbrace{\| f \|_p}_{\text{old}}$$

This norm is well-defined: if $f, \tilde{f} \in [f]$, then $\| f \|_p = \| \tilde{f} \|_p$.

This "norm" is a norm, because

$$\| [f] \|_p = 0 \Rightarrow [f] = [0].$$

Conclusion: L_p with $\| \cdot \|_p$ is a Banach space!

For simplicity, in future we write $\| f \|_p$ for $\| [f] \|_p$.

Scalar product

<u>Def</u> Consider vector space V. A mapping $\langle \cdot , \cdot \rangle : V \times V \to \mathbb{R}$ is called a scalar product if

linearity
(S1) $\langle x_1 + x_2 , y \rangle = \langle x_1 , y \rangle + \langle x_2 , y \rangle$
(S2) $\langle \lambda x , y \rangle = \lambda \langle x , y \rangle$

symmetry
(S3) $\langle x , y \rangle = \langle y , x \rangle$ (if on $\mathbb{R}$)
$\langle x , y \rangle = \overline{\langle y , x \rangle}$ (if on $\mathbb{C}$)

complex conjugate

pos. def.
(S4) $\langle x , x \rangle \geq 0$
(S5) $\langle x , x \rangle = 0 \iff x = 0$

<u>Examples</u>:
- Euclidean scalar product on $\mathbb{R}^n$; $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$$\langle x , y \rangle = \sum_{i=1}^{n} x_i y_i$$

- On $\mathbb{C}^n$, $\langle x , y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$

- $\mathcal{C}([a,b])$: $\langle f , g \rangle = \int_a^b f(t)\, g(t)\, dt$

is a scalar product (but space would not be complete).

Def A vector space with a norm is called a normed space. If a normed space is complete (each Cauchy sequence converges), then V is called a Banach space. A VS with a scalar product is called a pre-Hilbert-space. If it is additionally complete, then it is called Hilbert space.

$$\boxed{\text{Scalar product} \Rightarrow \text{norm} \quad \not\Leftarrow}$$

Consider a VS with a scalar product $\langle \cdot, \cdot \rangle$. Define $\| \cdot \| : V \to \mathbb{R}$ as $\|x\| := \sqrt{\langle x, x \rangle}$. Then $\| \cdot \|$ is a norm on V, the norm induced by $\langle \cdot, \cdot \rangle$.

The other way round does not work in general.

$$\boxed{\text{norm} \Rightarrow \text{metric} \quad \not\Leftarrow}$$

Consider a VS V with norm $\| \cdot \|$. Then

$$d : V \times V \to \mathbb{R}, \quad d(x,y) := \|x - y\|$$

is a metric on V, the metric induced by the norm.

The other direction does not work in general.

Orthogonal basis and projections

Def Consider a pre-Hilbert space V. Two vectors $v_1, v_2 \in V$ are called <u>orthogonal</u> if $\langle v_1, v_2 \rangle = 0$.

Notation: $v_1 \perp v_2$

Two sets $V_1, V_2 \subset V$ are called orthogonal if

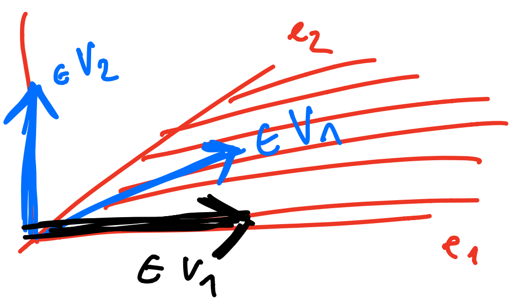
$$\forall v_1 \in V_1 \ \forall v_2 \in V_2 : \ \langle v_1, v_2 \rangle = 0$$

$V_1 = \mathrm{span}\{e_1, e_2\}$

$V_2 = \mathrm{span}\{e_3\}$

standard Euclidean scalar product. Then:

$V_1 \perp V_2$



Vectors are called <u>orthonormal</u> if additionally, the two vectors have norm 1:

- $\langle v_1, v_2 \rangle = 0$
- $\|v_1\| = 1, \quad \|v_2\| = 1$

A set of vectors $v_1, v_2, \dots, v_n$ is called orthonormal if any two vectors are orthonormal.

For a set $S \subset V$ we define its orthogonal complement $S^\perp$ as follows:

$$S^\perp := \{ v \in V \mid v \perp s \ \forall s \in S \}$$

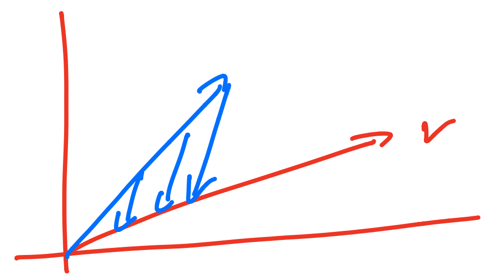
Remark: We are particularly interested in orthogonal/orthonormal bases of a space. In an orthonormal basis $u_1, \dots, u_n$, the representation of a vector v is given as

$$v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$$

Orthogonal projections

Def $A \in \mathcal{L}(V)$ is called a projection if $A^2 = A$.

Blue vector gets projected on red vector (not orthogonal)

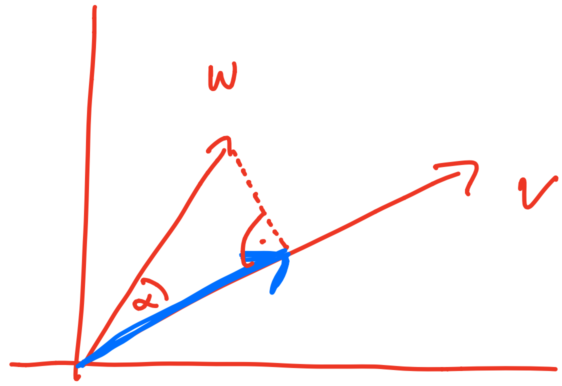


Theorem & Def: Let U be a finite-dim subspace of a pre-Hilbert space H. Then there exists a linear projection $P_U : H \to U$, and $\ker(P_U) = U^\perp$. P_U is then called the <u>orthogonal projection</u> of H on U.

Construction: Let $v_1, \dots, v_n$ be an orthogonal basis of U.

Define $P_U : V \to U$ by $P_U(w) = \sum_{i=1}^{n} \frac{\langle w, v_i \rangle}{\|v_i\|} v_i$

Intuition:



If $\|v\| = 1$, then $\| P_v(w) \| = | \langle v, w \rangle |$

$= | \cos \alpha |$

In particular, $\langle v, w \rangle = \cos \alpha$

Remark In an orthonormal basis $u_1, \dots, u_n$ the representation of a vector v is given as

$$v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$$

Gram-Schmidt orthogonalization

Is a procedure that takes any basis $v_1, \dots, v_n$ of a finite-dim VS and transforms it into another basis $u_1, \dots, u_n$ that is orthogonal:

Intuition: iterative procedure

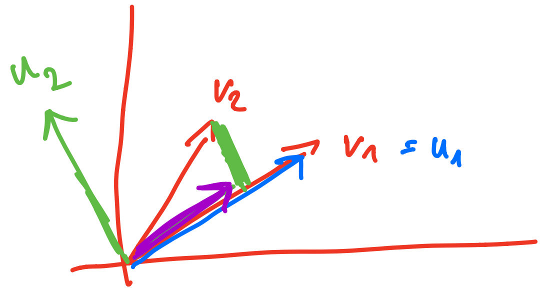
Step 1: $u_1 := \frac{v_1}{\|v_1\|}$

$U_1 := \operatorname{span}\{u_1\}$

Step k: Assume that we already identified $u_1, \dots, u_{k-1}$.

- Project v_k on U_{k-1}, and keep "the rest":

$$\tilde{u}_k := v_k - P_{U_{k-1}}(v_k)$$



- Renormalize:

$u_k = \tilde{u}_k / \|\tilde{u}_k\|$

Works ☺ (would need to prove that, skipped)

Orthogonal matrices

updated

Def Let $Q \in \mathbb{R}^{n \times n}$ be a matrix with orthonormal (!) column vectors (wrt Euclidean scalar product). Then Q is called an orthogonal (!) matrix.

If $Q \in \mathbb{C}^{n \times n}$ and the columns are orthonormal (wrt the standard scalar product on $\mathbb{C}^n$), then it is called unitary.

⚠ The literature is not completely consistent whether an orthogonal matrix needs to have rows/cols of norm 1. In any case, the definition only makes sense if the matrix is of full rank. See also (*) below.

Examples:

- Identity: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
- Reflection: $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
- Permutation of coordinates: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
- Rotation: $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$
- Rotation in $\mathbb{R}^3$:
  - Rotation about one of the axes:

$$R_{\theta,1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

- General rotation: can be written as a product of "elementary" rotations

<u>Properties</u> of orthonormal matrices Q:

- columns are orthogonal $\Leftrightarrow$ rows are orthogonal
- Q is always invertible, and $Q^{-1} = Q^t$
- Q realizes an <u>isometry</u>: $\forall v \in V$: $\|Qv\| = \|v\|$
- Q preserves angles: $\langle Qu, Qv \rangle = \langle u, v \rangle \quad \forall u, v \in V$
- $|\det Q| = 1$

The respective properties also hold for unitary matrices U:

- $U^{-1} = \overline{U}^t$

<u>Theorem</u> Let $S \in \mathcal{L}(V)$ for a real VS V. Then equivalent:

(a) S is an isometry: $\|Sv\| = \|v\| \quad \forall v \in V$

(b) there exists an orthonormal basis of V such that the matrix of S has the following form:

$$M = \begin{pmatrix} \square & & & & & & 0 \\ & \square & & & & & \\ & & \square & & & & \\ & & & \square & & & \\ & & & & \square & & \\ & & & & & \ddots & \\ 0 & & & & & & \square \end{pmatrix}$$

where each of the little blocks

- either of a 1×1 matrix ($\hat{=}$ a real number) being 1 or -1

- or of a 2×2 rotation matrix.

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

(*) Orthonormal vs. orthogonal:

Consider the projection matrix $A = \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix}$. The columns are obviously not orthogonal. The rows formally satisfy that $\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rangle = 0$. The property that "rows orthogonal $\Leftrightarrow$ cols orthogonal" does not hold here. But note that A is not an orthogonal matrix because the latter requires all rows / cols to have norm 1 (in particular, also full rank).

Symmetric matrices

Def A matrix $A \in \mathbb{R}^{n\times n}$ is called <u>symmetric</u> if $A = A^t$.

A matrix $A \in \mathbb{C}^{n\times n}$ is called <u>hermitean</u> if $A = \bar{A}^t$.

Prop Let $A \in \mathbb{C}^{n\times n}$ be hermitean. Then all eigenvalues of A are real-valued. Eigenvectors that correspond to different eigenvalues are orthogonal.

Proof • λ eig.value of A with eigenvector x. Then

$$\underline{\lambda \langle x, x\rangle} = \langle \lambda x, x \rangle = \langle Ax, x\rangle \overset{\text{hermitean}}{=}$$

$$= \langle x, Ax\rangle = \langle x, \lambda x\rangle = \underline{\bar{\lambda}\langle x, x\rangle}$$

$$\Rightarrow \lambda = \bar{\lambda} \in \mathbb{R}$$

• (λ_1, x_1), (λ_2, x_2) are eigs of A. Then

$$\underline{\lambda_1 \langle x_1, x_2\rangle} = \text{.... as above ...} = \underline{\lambda_2 \langle x_1, x_2\rangle}$$

$$0 = \lambda_1 \langle x_1, x_2\rangle - \lambda_2 \langle x_1, x_2\rangle =$$

$$= (\lambda_1 - \lambda_2)\langle x_1, x_2\rangle$$

$\Rightarrow$ either $\lambda_1 = \lambda_2$

or if $\lambda_1 \neq \lambda_2$, then $\langle x_1, x_2\rangle = 0$

$\Rightarrow x_1 \perp x_2$. ∎

Def An operator $T \in \mathcal{L}(V)$ on a pre-Hilbert space V is called <u>self-adjoint</u> if

$$\langle Tv, w \rangle = \langle v, Tw \rangle .$$

Sometimes it is called a Hermitean operator (on $\mathbb{C}^n$)

symmetric operator (on $\mathbb{R}^n$).

Remarks Over $\mathbb{C}^n$, self-adjoint operators are represented by hermitean matrices. On $\mathbb{R}^n$, self-adjoint op. are represented by symmetric matrices.

Prop $T \in \mathcal{L}(V)$ self-adjoint. Then T has at least one eigenvalue, and it is real-valued.
(holds both on $\mathbb{C}^n$ and $\mathbb{R}^n$).

Proof (sketch) $n := \dim V$. Chose $v \neq 0$, and consider

$$v, Tv, T^2 v, \dots, T^n v.$$

These vectors have to be lin. dependent ($n+1$ vectors, $\dim = n$).

There exist $a_0, \dots, a_n$: $a_0 v + a_1 Tv + \dots + a_n T^n v = 0$.

Consider polynomial with these coefficients:

$$a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n =$$

$$= c\,\underbrace{(x^2+b_1x+c_1)\cdots(x^2+b_Mx+c_M)}_{\text{quadratic terms}}\cdot\underbrace{(x-\lambda_1)\cdots(x-\lambda_m)}_{\text{linear}}$$

Replace the x by T:

$$0 = (a_0 + a_1T + \dots + a_nT^n)v = \Big(c\,\underbrace{(\dots)}_{\text{quadr.}}\ \underbrace{(\dots)}_{\text{lin. terms}}\Big)\cdot v$$

Now can prove: the quadratic terms are invertible, and we are left with (at least one) linear factor:

$$0 = (T-\lambda_1 I)\cdots(T-\lambda_n I)\,v$$

There needs to exist at least one i such that $(T-\lambda_i I)$ is not invertible. Thus λ_i is an eigenvalue of T.

∎

Spectral theorems for symmetric / hermitian matrices

Theorem: A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is orthogonally diagonalizable: there exists an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ and a diagonal matrix $D \in \mathbb{R}^{n \times n}$ s.t.

$$A = Q D Q^t$$
$$= \sum_{i=1}^{n} \lambda_i q_i q_i^t .$$

Proof (Sketch)

By induction on $n := \dim V$

Base case $n = 1$:

Inductive step $n-1 \rightsquigarrow n$

- A symmetric $\Rightarrow$ A has at least one eigenvector u.
- $U := \operatorname{span}\{u\}$. U is invariant under A.
- Consider $U^\perp$ and the restriction of A to $U^\perp$.

  On $U^\perp$, A is again a symmetric operator

  and $\dim(U^\perp) = n-1$.
- Apply the induction hypothesis on this space of dim $n-1$.

Does the job!

$\blacksquare$

Complex version

Theorem A hermitian matrix $A \in \mathbb{C}^{n\times n}$ is unitarily diagonalizable: there exists a unitary matrix U and a diagonal matrix D s.th.

$$A = U D \bar{U}^t$$

In particular, the entries of D are real-valued.

Positive definite matrices

Def A matrix $A \in \mathbb{R}^{n\times n}$ is called

positive ~~definite~~ semi-definite (pd) if $\forall x \in \mathbb{R}^n,\ x \neq 0$:

$$x^t A x > 0 .$$

$$\geq$$

Def A matrix $A \in \mathbb{C}^{n \times n}$ is called a Gram matrix if there exists a set of vectors $v_1, \dots, v_n \in \mathbb{C}^n$ s.th. $a_{ij} = \langle x_i, x_j \rangle$. Note: Gram matrices are hermitian (similarly, on $\mathbb{R}^{n\times n}$, then Gram matrices are symmetric).

⚠ Over $\mathbb{C}$, we have that pd $\Rightarrow$ self-adjoint.

Over $\mathbb{R}$, this is <u>not</u> true!

Example:

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

on $\mathbb{R}$, if $x \neq 0$

$$x^t A x = x_1^2 + x_2^2 > 0$$

- So A is pd but not symmetric.
- Over $\mathbb{C}$, the same matrix is not pd because $x_1^2 + x_2^2$ can be negative!

<u>Theorem</u>: $A \in \mathbb{C}^{n\times n}$ hermitean. Then equivalent:

(i) A is psd (pd)

(ii) All eigenvalues of A are ≥ 0 (>0)

(iii) The mapping $\langle\cdot,\cdot\rangle_A : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ with

$$\langle x, y\rangle_A := \bar{y}^t A x$$

satisfies all properties of a scalar product except one: if $\langle x,x\rangle_A = 0$ this does not imply $x=0$.

(This mapping is a scalar product.)

(iv) A is a Gram matrix of n vectors which are not necessarily lin. independent (which are lin. independent).

$$a_{ij} = \langle x_i, x_j\rangle$$

Roots of psd matrices

<u>Theorem</u>: Let $A \in \mathbb{R}^{n\times n}$ be symmetric, psd. Then there exists a matrix $B \in \mathbb{R}^{n\times n}$, B psd such that $A = B^2$.

Sometimes B is called the square root of A, sometimes denoted as $B = (A)^{1/2}$.

<u>Proof</u> • Spectral theorem $\Rightarrow$

$$A = U D U^t \quad , \quad D \text{ diagonal.}$$

• psd $\Rightarrow$ eigenvalues ≥ 0

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad , \quad \lambda_i \geq 0$$

Define $\sqrt{D} := \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}$ and set

$B := U \sqrt{D}\, U^*$. Does the job. ∎

Variational characterization of eigenvalues

<u>Def</u> Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix.

$$R_A : \mathbb{R}^n \setminus \{0\} \to \mathbb{R} \,, \quad x \mapsto \frac{x^t A x}{x^t x}$$

is called the Rayleigh coefficient.

<u>Prop</u> Let A be symmetric, let $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$ be the eigenvalues and $v_1, \dots, v_n$ the eigenvectors of A.

Then:

$$\min_{x \in \mathbb{R}^n} R_A(x) = \min_{\|x\|=1} x^t A x = \lambda_1 \,, \quad \text{attained at } x = v_1$$

$$\max_{x \in \mathbb{R}^n} R_A(x) = \max_{\|x\|=1} x^t A x = \lambda_n \,, \quad \text{attained at } v_n.$$

<u>Intuition</u>: Assume A is expressed in terms of the basis $v_1, \dots, v_n$

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Let y be a vector, also represented in this basis

$(y = y_1 v_1 + y_2 v_2 + \dots + y_n v_n)$

(*) $y^t A y = \underline{\lambda_1} y_1^2 + \dots + \lambda_n y_n^2$

Among the vectors $\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\\vdots\end{pmatrix}, \dots, \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$

the smallest result of $y^t A y$ would be given by the vector $\underbrace{\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}}_{= v_1}$, and the value would be λ_1

__More general proof sketch:__ Assume we start with the standard basis. Let $Q = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$ be the basis transformation.

Observe: Q orthogonal, we have

$$A = Q^t \Lambda Q \quad \text{with } \Lambda \text{ diagonal.}$$

For a vector $x = \begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}$ in the original basis, we now consider

$y := Q^t x$.

$$R_A(y) = \frac{\overbrace{(Q^t x)^t}^{y^t} \overbrace{(Q^t \Lambda Q)}^{A} \overbrace{(Q^t x)}^{y}}{\underbrace{(Q^t x)}_{y^t}{}^t \underbrace{(Q^t x)}_{y}} \qquad (Q^t x)^t = x^t Q$$

$$= \frac{x^t \cancel{Q Q^t} \Lambda \cancel{Q Q^t} x}{x^t \cancel{Q Q^t} x} = \frac{x^t \Lambda x}{x^t x} = \frac{\lambda_1 x_1^2 + \dots + \lambda_n x_n^2}{\|x\|}$$

$$\min_{\|y\|=1} R(y) = \min_{\|x\|=1} \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$$

This min. is attained for $x = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, that is

$y = Q^t x = v_1$, with value $R(y) = \lambda_1$.

Prop Consider the problem

$$\min_{\substack{\|x\|=1 \\ x \perp v_1}} R(x).$$

This problem is solved with $x = v_2$, $R(v_2) = \lambda_2$.

Intuition Consider operator A restricted to the space

$V_1^\perp := (\operatorname{span}\{v_1\})^\perp$. We know that on this space, A is invariant and symmetric, so we can apply Rayleigh to this "smaller" space.

$V_1^\perp = \operatorname{span}\{v_2, \dots, v_n\}$

If we apply Rayleigh to $V_1^\perp$, then we get the solutions λ_2, v_2.

Theorem (Min-max-theorem, Courant-Fischer-Weyl-theorem)

$A \in \mathbb{R}^{n \times n}$ symmetric, eigenvalues $\lambda_1 \leq \dots \leq \lambda_n$. Then:

$$\lambda_k = \min_{\substack{U \text{ subspace} \\ \dim U = k}} \max_{x \in U \setminus \{0\}} R_A(x)$$

$$= \max_{\substack{U \text{ subspace}, \\ \dim U = n-k+1}} \min_{x \in U \setminus \{0\}} R_A(x)$$

Intuition: for case of $k=3$

- Consider the subspace U spanned by v_1, v_2, v_3.

  By the argument similar as before,

$$\max_{x \in U} R_A(x) = \lambda_3 \text{, attained at } v_3$$

- Consider another subspace, for example spanned by v_8, v_9, v_{10}

$$\max_{x \in U} R_A(x) = \lambda_{10}$$

Singular value decomposition

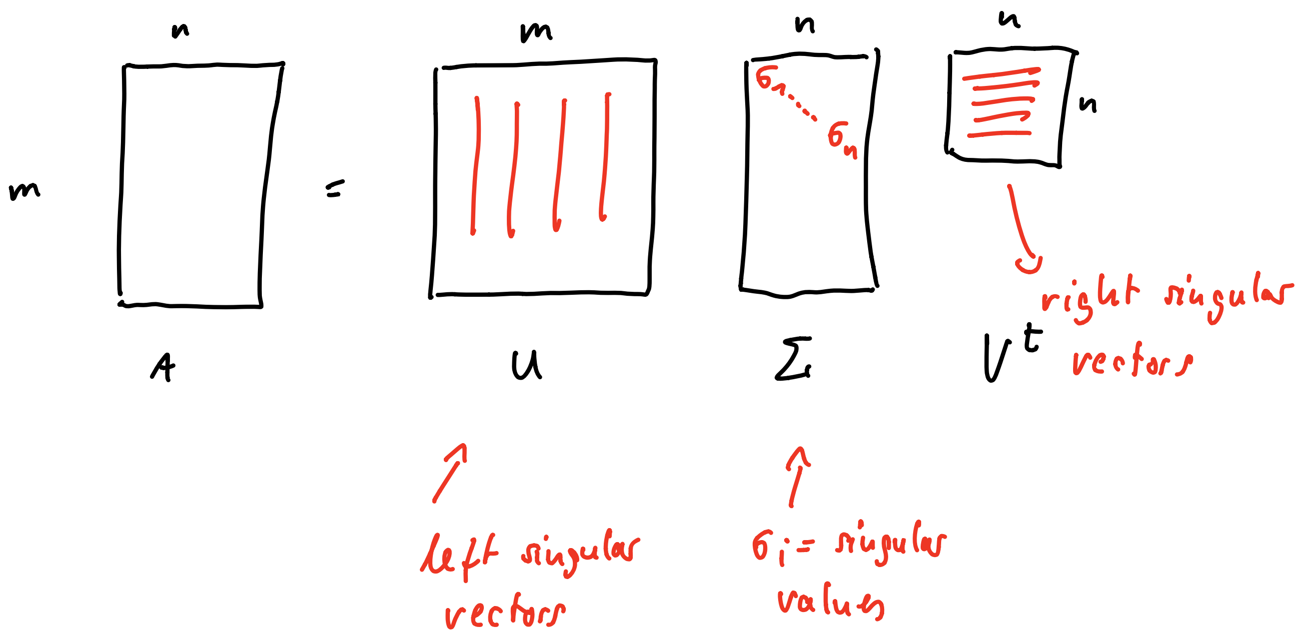
Proposition Consider $A \in \mathbb{R}^{m \times n}$ of rank r. Then we can write A in the form

$$\boxed{A = U \cdot \Sigma \cdot V^t}$$

where $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{m \times n}$ is "diagonal".

$$\begin{array}{cc} & n \\ m & \begin{pmatrix} \sigma_1 & & 0 \\ 0 & \ddots & \sigma_n \\ & 0 & \end{pmatrix} \end{array} \qquad \begin{array}{cc} & n \\ m & \begin{pmatrix} \sigma_1 & & 0 & \\ 0 & \ddots & \sigma_m & 0 \end{pmatrix} \end{array}$$

Exactly r of the diagonal values $\sigma_1, \sigma_2, \dots$ are non-zero.



Proof sketch Given $A \in \mathbb{R}^{m \times n}$, we consider

$$B := \underbrace{A^t}_{n \times m} \underbrace{A}_{m \times n} \in \mathbb{R}^{n \times n}.$$

observe:
- B is symmetric:

$$(A^t A)^t = A^t (A^t)^t = A^t A$$

- B is positive semi-definite:

$$x^t B x = \langle x, Bx \rangle = \langle x, A^t A x \rangle$$

$$= \langle Ax, Ax \rangle$$

$$= \| Ax \|^2 \geq 0$$

So there exists an orthonormal basis of eigenvectors $x_1, \dots, x_n$ with eigenvalues $\lambda_1, \dots, \lambda_n \geq 0$.

Define:

- $\Sigma =$ "diag(σ_i)" $\in \mathbb{R}^{m \times n}$

  where $\sigma_i = \sqrt{\lambda_i}$

- $U = \begin{pmatrix} | \\ r_i \\ | \end{pmatrix}$ matrix with columns

$$r_i := \frac{A x_i}{\sigma_i}$$

- $V = \begin{pmatrix} | \\ x_i \\ | \end{pmatrix}$ matrix with x_i as columns

Now we need to show that with these definitions we have $A = U \cdot \Sigma \cdot V^t$.

Sketch:

- Column of $U \cdot \Sigma$ are given as

$$\sigma_i r_i = \sigma_i \frac{A x_i}{\sigma_i} = A x_i$$

- Now multiply with V^t:
  - rows of V^t are the x_i,
  - exploit that if $i \neq j$ then $x_i \perp x_j$ and $\|x_i\| = 1$.
  - The terms consisting of i, j with $i \neq j$ cancel, the terms with $i = j$ will result in a factor of 1.

  So we will be left with matrix A.

Key differences between SVD and eigs:

- SVD always exists, no matter how A looks like!
- U, V are orthonormal! (not true for eigenvectors in general).
- singular values are always real and non-negative.
- If $A \in \mathbb{R}^{n \times n}$ is symmetric, then the SVD is "nearly the same" as the eigenvalue decomposition: (λ_i, v_i) are the eigenvalues/vectors of A, then

$(|\lambda_i|, v_i)$ are the singular values / vectors of A.

In particular, left- and right singular vectors are the same.

- Left-singular vectors of A are the eigenvectors of AA^t.
- Right- A^tA.
- $\lambda \neq 0$ is a eigenvalue of $\overset{A^tA}{AA^t}$ $\iff$ $\sqrt{\lambda} \neq 0$ is singular value of A

Matrix norms

Given a matrix $A \in \mathbb{R}^{m \times n}$. Define the following norms:

$$\|A\|_{max} = \|A\|_{\infty} = \max_{ij} |a_{ij}|$$

$$\|A\|_1 = \sum_{ij} |a_{ij}|$$

$$\|A\|_F = \sqrt{\sum_{ij} a_{ij}^2} = \sqrt{tr(A^t A)}$$

Frobenius norm

$$= \sqrt{\sum \sigma_i^2}$$ where σ_i are the singular values of A.

$\|A\|_2 = \sigma_{max}(A)$ where σ_{max} is the largest singular value

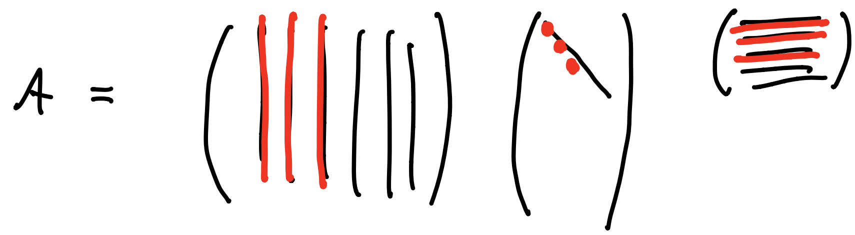
$$= \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Euclidean norm on vectors in $\mathbb{R}^m$

"Operator norm", "spectral norm"

Rank-k-approximation

Given matrix $A = U \Sigma V^t$, entries $\sigma_1, \sigma_2, \ldots$ sorted in decreasing order, $k \in \mathbb{N}$. Now we are going to define a new matrix A_k by the following procedure:



- take first k cols of U, first k entries of Σ, first k rows of V^t $\Big\}$ A_k

More formally:

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^t$$

Prop Let B be any rank-k-matrix $\in \mathbb{R}^{m \times n}$. Then:

$$\| A - A_k \|_F \leq \| A - B \|_F .$$

"A_k is the best rank-k-approximation (in Frobenius norm)."

<u>Prop</u> For any matrix B of rank k, $B \in \mathbb{R}^{m \times n}$,

$$\| A - A_k \|_2 \leq \| A - B \|_2 .$$

where $\| \cdot \|_2$ denotes the operator norm.

"A_k is the best rank-k-approximation (in operator norm)".

Pseudo-inverse

Definition For $A \in \mathbb{R}^{m \times n}$, a pseudo-inverse of A is defined as the matrix $A^\# \in \mathbb{R}^{n \times m}$ which satisfies the following conditions:

If A would be invertible
$AA^{-1} = Id \Rightarrow AA^{-1}A = A$

$$(1) \quad \underbrace{A\overbrace{A^\#}^{\neq Id \text{ in general}}}_{\neq Id \text{ in general}}A = A$$

$$(2) \quad A^\# A A^\# = A^\#$$

(1), (2): "nearly inverse"

$$(3) \quad (AA^\#)^t = AA^\#$$

$$(4) \quad (A^\# A)^t = A^\# A$$

(3), (4): symmetry

Intuition:

- A is a projection from $\mathbb{R}^3 \to \mathbb{R}^2$,
$$A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
- Cannot invert, obviously (inverting would mean to reconstruct the original point).
- But I could "invent" a reconstruction, for example: $R : \mathbb{R}^2 \to \mathbb{R}^3$,
$$R\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 17 \end{pmatrix}$$
- Now we have:
$$ARA = A$$
$$A\overset{\shortparallel}{A^\#}A = A$$

Proposition: Let $A \in \mathbb{R}^{m \times n}$, $A = U \Sigma V^t$ its SVD. Then:

$$A^{\#} := V \Sigma^{\#} U^t \quad \text{with} \quad \Sigma^{\#} \in \mathbb{R}^{m \times n} = \begin{pmatrix} \sigma_1 & \cdots & \sigma_n \\ & 0 & \end{pmatrix}$$

$$\Sigma^{\#}_{ii} = \begin{cases} 1/\Sigma_{ii} & \text{if } \Sigma_{ii} \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Intuition: Assume $A \in \mathbb{R}^{n \times n}$, invertible, assume it has eigendecomposition $A = UDU^t$. Then:

- All entries in diag (D) are $\neq 0$ (eigenvalues $\neq 0$)
- $A^{-1} = U D^{-1} U^t$ with

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

$$D^{-1} = \begin{pmatrix} 1/d_1 & & & 0 \\ & 1/d_2 & & \\ & & \ddots & \\ 0 & & & 1/d_n \end{pmatrix}$$

Proof: easy, just do it.

Operator norms

Theorem X, Y normed spaces, $T: X \to Y$ linear. Then the following statements are equivalent:

(i) T is continuous. $\longrightarrow$ $\left\{ \begin{array}{l} \forall x \in X \ \forall \varepsilon > 0 \ \exists \delta > 0 \ \forall y \in X: \\ \|x - y\| < \delta \Rightarrow \|Tx - Ty\| < \varepsilon \end{array} \right.$

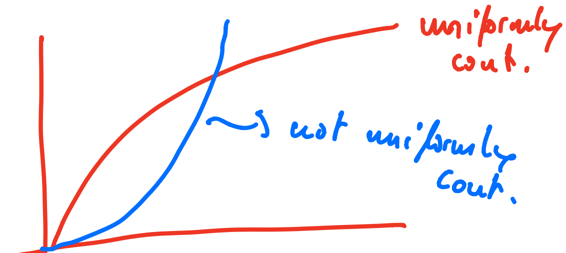
(ii) T is continuous at 0.

(iii) T is bounded:

$$\exists M > 0 \ \forall x \in X: \|Tx\| \leq M \cdot \|x\|$$

(iv) T is uniformly continuous.

$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in X \ \forall y \in X:$
$\|x - y\| < \delta \Rightarrow \|Tx - Ty\| < \varepsilon$



Def X, Y normed spaces, $T: X \to Y$ linear and continuous.

$$\|T\| := \sup_{x \in X} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x \in X \\ \|x\| \leq 1}} \|Tx\| = \sup_{\substack{x \in X \\ \|x\| = 1}} \|Tx\|$$

is called the <u>operator norm</u> of T.

Observe: coincides with the matrix norm $\|\cdot\|_2$ as we had defined it earlier.

Examples

- <u>Evaluation operator</u>: $T: \mathcal{C}[0,1] \to \mathbb{R}$, $Tf = f(0)$.

  As norms consider $\|\cdot\|_\infty$ on $\mathcal{C}[0,1]$, $|\cdot|$ on $\mathbb{R}$. Then $\|T\| = 1$.

  $$\sup_{f \in \mathcal{C}[0,1]} \frac{|Tf|}{\|f\|_\infty} = \sup \frac{|f(0)|}{\|f\|_\infty} = \underline{\text{exercise}} = 1$$

- <u>Integral operator</u>: $T: \mathcal{C}[0,1] \to \mathbb{R}$, $Tf = \int_0^1 f(t)\,dt$

  With the same norms as above, T is cont. and has $\|T\| = 1$.

- <u>Differential operator</u>: $D: \mathcal{C}^1[0,1] \to \mathcal{C}[0,1]$, $f \mapsto f'$.

  - Consider $\|\cdot\|_\infty$ on $\mathcal{C}^1$ and $\mathcal{C}$. Then D is linear, but not continuous!

  - Consider $|||f||| := \|f\|_\infty + \|f'\|_\infty$ on $\mathcal{C}^1$. With this norm, D is continuous and bounded.

Dual space

Definition V VS, $T : V \to F$ linear is called a *functional*.

Given a vector space V, the algebraic *dual space V^** consists of all linear functionals on V:

$$V^* := \mathcal{L}(V, F).$$

If V is a normed VS, then the space of all linear, continuous functionals from V to F is called the (topological) *dual space V'* of V.

Remark If V is finite dim, then $V^* = V'$ because then linear mappings are always continuous. In general, this is not true.

We endow the dual space with the operator norm

$$\|T\| := \sup_{x \in X} \frac{\|Tx\|}{\|x\|}.$$

Proposition: V' is a vector space, and the operator norm is indeed a norm on V'.

Prop: If V is a normed VS (but not necessarily complete), then V' with the operator norm is a Banach space.

Examples:

- $K \subset \mathbb{R}$ compact set, $\mathcal{C}(K)$ space of cont. fcts with $\|\cdot\|_\infty$. Then $(\mathcal{C}(K))'$ is equivalent to the space $M(K)$, the space of all (Radon) measures over K.

- $S \subset \mathbb{R}$ measurable set, $1 \leq p < \infty$, q such that $\frac{1}{p} + \frac{1}{q} = 1$. Then: the dual of $L_p(S)$ is given as $L^q(S)$.

Riesz representation theorem

Theorem: H Hilbert space, H' its dual. Then the mapping $\Phi : H \to H'$, $y \mapsto \langle \cdot , y \rangle$ is bijective, isometric, and satisfies $\Phi(\lambda x) = \bar{\lambda}\, \Phi(y)$.

Stated differently: for any mapping $x' \in H'$ there exists a unique $y \in H$ such that $x'(x) = \langle x, y \rangle$.

Adjoint operator

<u>Def</u> Let $T \in \mathcal{L}(H_1, H_2)$, H_1, H_2 Hilbert spaces. Then there exists an operator $T^*: H_2 \to H_1$ such that

$$\langle Tx, y \rangle_{H_2} = \langle x, T^* y \rangle_{H_1}.$$

for all $x \in H_1$, $y \in H_2$. T^* is called the adjoint of T.

<u>Remark</u> The existence of this operator is a consequence of the Riesz representation theorem.

<u>Def</u> An operator $T: H_1 \to H_1$ is called self-adjoint if $\langle Tx, y \rangle = \langle x, Ty \rangle$