Basic Linear Algebra

with Applications to Data Analysis

Carl O. R. Lutz
carl.lutz@uni.lu
carlorlutz.github.io

Université du Luxembourg
September 9, 2026

Refresher I

— Basic Notions —

- Groups, Fields, Vector Spaces
- Subspaces, Spanning Sets, Bases
- Direct Sums, Linear Complements
- Linear Functions, Isomorphisms
- Rank, Kernel, Rank-Nullity Theorem

Groups, Fields & Vector Spaces

Definition (Group)

A group is a set G with a binary operation $\cdot : G \times G \to G$ s.t.

- Associativity: $\forall a, b, c \in G : (a \cdot b) \cdot c = a \cdot (b \cdot c)$
- Identity element: $\exists! e \in G \forall a \in G : a \cdot e = a = e \cdot a$
- Inverse element: $\forall a \in G \exists! a^{-1} \in G : a \cdot a^{-1} = e = a^{-1} \cdot a$

The group is abelian/commutative if

$$a + b = b + a \qquad \forall a, b \in G.$$

Groups, Fields & Vector Spaces

Definition (Field)

A field is a set $\mathbb{K}$ with binary operations $+, \cdot \colon \mathbb{K} \times \mathbb{K} \longrightarrow \mathbb{K}$ s.t.

- $(\mathbb{K}, +)$ is an abelian group with identity element 0;
- $\mathbb{K}_* \coloneqq \mathbb{K} \setminus \{0\}$ is an abelian group with identity element 1;
- $\forall x, y, z \in \mathbb{K}$:

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Groups, Fields & Vector Spaces

Definition (Vector Space)

A vector space over a field $\mathbb{K}$ is a set V with a binary operation $+\colon V \times V \to V$ and a *scalar multiplication* $\cdot\colon \mathbb{K} \times V \to V$ s.t.

- $(V, +)$ is an abelian group;
- $+$ and $\cdot$ are compatible in the sense that for all $\lambda, \lambda_1, \lambda_2 \in \mathbb{K}$ and $v, w \in V$ holds
  - (i) $1 \cdot v = v$;
  - (ii) $(\lambda_1 + \lambda_2) \cdot v = (\lambda_1 \cdot v) + (\lambda_2 \cdot v)$.
  - (iii) $(\lambda_1 \lambda_2) \cdot v = \lambda_1 \cdot (\lambda_2 \cdot v)$;
  - (iv) $\lambda \cdot (v_1 + v_2) = (\lambda \cdot v_1) + (\lambda \cdot v_2)$;

Subspaces & Spanning Sets

Definition

A linear subspace of a vector space V over $\mathbb{K}$ is a subset $U \subset V$ that is closed under addition and scalar multiplication:

- $0 \in U$;
- if $u_1, u_2 \in U$ then $u_1 + u_2 \in U$;
- if $u \in U$ then $\lambda \cdot u \in U$ for all $\lambda \in \mathbb{K}$.

Definition

The span of a set $\mathcal{S} \subset V$ is

$$\operatorname{span}(\mathcal{S}) := \left\{ \sum_{i=1}^{N} \lambda_i \cdot s_i \mid N \in \mathbb{N}, s_i \in S, \lambda_i \in \mathbb{K} \right\}.$$

If $\operatorname{span}(\mathcal{S}) = V$, then $\mathcal{S}$ is a spanning set of V.

Bases of Vector Spaces

Definition

A set $\mathcal{B} \subset V \setminus \{0\}$ is called linear independent if for all $N \in \mathbb{N}$ and $b_1, \ldots, b_N \in \mathcal{B}$ holds:

$$\lambda_1 b_1 + \ldots + \lambda_N b_N = 0 \quad \Longleftrightarrow \quad \lambda_1 = \ldots = \lambda_N = 0.$$

If $\mathcal{B}$ is also a *spanning set* of V, then $\mathcal{B}$ is called a basis of V. The dimension of V is the cardinality $\dim_{\mathbb{K}}(V) = |\mathcal{B}|$. In particular, V is finite dimensional if $\dim_{\mathbb{K}}(V) < \infty$.

New from Old: Union ($\cup$) & Intersection ($\cap$)

Proposition

Let V_1, V_2 be subspaces of V. Then

- $V_1 \cap V_2$ is *always* a subspace of V;
- $V_1 \cup V_2$ is a linear subspace of V *if and only if* $V_1 \subset V_2$ or $V_2 \subset V_1$.

Proposition

Let $\mathcal{S} \subset V$. The span of $\mathcal{S}$ is the smallest linear subspace of V containing $\mathcal{S}$ in the sense that

$$\operatorname{span}(\mathcal{S}) = \bigcap_{\substack{\mathcal{S} \subset U \subset V \\ \text{subspace}}} U.$$

New from Old: Sum (+) & Direct Sum ($\oplus$)

Definition (Direct Sum)

Let V, W be vector spaces over $\mathbb{K}$. The direct sum $V \oplus W$ is the set $V \times W$ endowed with the

- *addition*: $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$;
- *scalar mult.*: $\lambda \cdot (u, w) = (\lambda v, \lambda w)$.

Definition (Sum & Linear Complement)

Let V_1, V_2 be linear subspaces of V.

- The sum of V_1 and V_2 is the subspace $V_1 + V_2 := \mathrm{span}(V_1 \cup V_2)$.
- V_1 and V_2 are linear complements (in V) if $V_1 \oplus V_2$ is isomorphic to V.

Linear Functions

Definition

Let V, W be vector spaces over $\mathbb{K}$.

- A linear function is a function $f\colon V \to W$ such that

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2)$$

  for all $\lambda_1, \lambda_2 \in \mathbb{K}$ and $v_1, v_2 \in V$.
- A linear function is called isomorphism if it admits a linear inverse $f^{-1}\colon W \to V$, *i.e.*, f^{-1} is linear and $f^{-1} \circ f = \mathrm{id}_V$.
  In this case we call V and W *(linear) isomorphic* and write $V \simeq W$.

Proposition

Let $\mathcal{B} \subset V$ be a basis. Every linear map $f\colon V \to W$ is determined by its action on $\mathcal{B}$, *i.e.*, if $w_b \in W$ for $b \in \mathcal{B}$, then

$$f(b) := w_b$$

defines f completely.

Proposition

The set $\mathrm{Lin}(V, W)$ of linear functions from V to W is a vector space. If $\dim(V), \dim(W) < \infty$, then $\dim(\mathrm{Lin}(V, W)) = \dim(V) \cdot \dim(W)$.

Proposition

The set $\mathrm{GL}(V)$ of isomorphisms $V \to V$ is a group.

Basics of Matrices

Definition

An (m,n)-matrix over a field $\mathbb{K}$ is a map $A\colon \{1,\dots,m\}\times\{1,\dots,n\}\to\mathbb{K}$. We usually write

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = (a_{ij})_{mn} = (a_{ij})$$

For $i \in \{1,\dots,m\}$ and $j \in \{1,\dots,n\}$ is $\begin{pmatrix} a_{i1} & \dots & a_{in} \end{pmatrix}$ the ith row and $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ the jth column of A.

The set of all (m,n)-matrices of $\mathbb{K}$ is denoted by $\mathbb{K}^{m,n}$.

Basics of Matrices

Definition

- The matrix product $\cdot\colon \mathbb{K}^{m,n} \times \mathbb{K}^{n,p} \longrightarrow \mathbb{K}^{m,p}$ is

$$(a_{ij})_{mn} \cdot (b_{ij})_{np} := \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)_{mp}$$

- The *transpose* of a matrix $A = (a_{ij})_{mn} \in \mathbb{K}^{m,n}$ is $A^T = (\tilde{a}_{ij})_{nm} \in \mathbb{K}^{n,m}$ with $\tilde{a}_{ij} = a_{ji}$.

Matrices & Linear Functions

Proposition

- Each $A \in \mathbb{K}^{m,n}$ defines a linear function $L_A \colon \mathbb{K}^n \to \mathbb{K}^m$ via

$$L_A(x) \coloneqq A \cdot v, \qquad v \in \mathbb{K}^n.$$

- For all $A \in \mathbb{K}^{m,n}$ and $B \in \mathbb{K}^{n,p}$ holds $L_A \circ L_B = L_{A \cdot B}$.

Matrices & Linear Functions

Proposition

Let $\mathcal{B} \subset V$, $\mathcal{B}' \subset W$ be bases with $\dim V = n$ and $\dim W = m$.
Then for $f \in \operatorname{Lin}(V, W)$ and all $j \in \{0, \ldots, n\}$ there are $a_{1j}, \ldots, a_{mj} \in \mathbb{K}$ such that

$$f(b_j) = a_{1j}b'_1 + \ldots + a_{mj}b'_m.$$

The matrix representation of f is the matrix $[f]_{\mathcal{B},\mathcal{B}'} := (a_{ij})_{ij}$.
It satisfies that for $v = \lambda_1 b_1 + \ldots + \lambda_n b_n \in \mathbb{K}^n$ and
$\mu := L_{[f]_{\mathcal{B},\mathcal{B}'}}(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^m$ holds

$$f(v) = \mu_1 b'_1 + \ldots + \mu_m b'_m.$$

Rank-Nullity of Linear Functions

Definition

Let $f: V \to W$ be a linear function and $A \in \mathbb{K}^{m,n}$.

- The kernel of f is

  $$\ker(f) := \{v \in V \mid f(v) = 0\}.$$

  If $\dim(V) < \infty$, then the nullity of f is $\dim \ker(f)$.

- The image of f is

  $$\mathrm{im}(f) := \{f(v) \in W \mid v \in V\}.$$

  If $\dim(V) < \infty$, then the rank of f is $\dim \mathrm{im}(f)$.

- We define $\ker(A) := \ker(L_A)$ and $\mathrm{im}(A) := \mathrm{im}(L_A)$.

Rank-Nullity of Linear Functions

Theorem (Rank-Nullity)

Let $f : V \to W$ be a linear function and $A \in \mathbb{K}^{m,n}$. Then

- $V \simeq \ker(f) \oplus \mathrm{im}(f)$
- *If* $\dim(V) < \infty$, *then* $\dim(V) = \dim\ker(f) + \dim\mathrm{im}(f)$
- $n = \dim\ker(L_A) + \dim\mathrm{im}(L_A)$

The Determinant

Definition

The determinant $\det \colon \mathbb{K}^{n,n} \to \mathbb{K}$ is

$$\det A \coloneqq \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots a_{n(\sigma n)}.$$

Proposition

Some of the properties of the determinant are:

- $\forall A, B \in \mathbb{K}^{n,n}\colon \det(A \cdot B) = \det(A) \det(B)$
- $\det(A) = 0 \Leftrightarrow A \in \mathrm{GL}(\mathbb{K}; n)$
- $\det$ is multi-linear, but <u>not</u> linear!

The Determinant: Adjugate Formula

Definition

- Let $A \in \mathbb{K}^{n,n}$. The ij-minor matrix $A^{ij} \in \mathbb{K}^{n-1,n-1}$ is obtained by deleting the ith row and jth column of A.
- The adjugate $\mathrm{adj}(A) \in \mathbb{K}^{n,n}$ of $A \in \mathbb{K}^{n,n}$ is given by

$$\mathrm{adj}(A)_{ij} := (-1)^{i+j} \det(M^{ji}).$$

Proposition

For all $A \in \mathbb{K}^{n,n}$ holds

$$A \,\mathrm{adj}(A) = \mathrm{adj}(A) A = \det(A) I_n.$$

In particular, for $A \in \mathrm{GL}(\mathbb{K}; n)$ we have

$$A^{-1} = \frac{1}{\det(A)} \,\mathrm{adj}(A).$$

The Determinant: Isomorphisms

Proposition

A map $f \in \mathrm{Lin}(V, W)$ between finite dimensional vector spaces is an isomorphism *if and only if* there exists a representation as a square matrix $[f]_{\mathcal{B}_V, \mathcal{B}_W} \in \mathbb{K}^{n,n}$ with $\det([f]_{\mathcal{B}_V, \mathcal{B}_W}) \neq 0$.
Moreover, the value of $\det([f]_{\mathcal{B}_V, \mathcal{B}_W})$ is independent of the choice of bases $\mathcal{B}_V$ and $\mathcal{B}_W$.

Refresher II

— Eigenvalue Problems —

- Eigenvalues and Eigenvectors
- Characteristic Polynomial
- Criteria for Diagonalizability
- Jordan Decomposition

Eigenvalues & Eigenvectors

Definition

- An eigenvalue-eigenvector pair of $f \in \mathrm{Lin}(V, V)$ is are $\lambda \in \mathbb{K}$ and $v \in V \setminus \{0\}$ s.t.
$$f(v) = \lambda v.$$
- The eigenpairs of $A \in \mathbb{K}^{n,n}$ are the eigenpairs of $L_A \in \mathrm{Lin}(\mathbb{K}^n, \mathbb{K}^n)$.

Proposition

The eigenspace $\mathrm{Eig}(f; \lambda) \subset V$ of f for $\lambda \in \mathbb{K}$ is the vector space of eigenvectors of f for λ. In particular,

$$\mathrm{Eig}(f; \lambda) = \ker(\lambda \, \mathrm{id}_V - f).$$

Characteristic Polynomial

Definition

Let $A \in \mathbb{K}^{n,n}$. The characteristic polynomial of A is

$$\chi_A(t) := \det(tI_n - A) \in \mathbb{K}[t].$$

Proposition

- $\chi_A(t)$ is monic of degree n. In particular,

$$\chi_A(t) = t^n - \operatorname{tr}(A) + \cdots + (-1)^n \det(A).$$

- $\lambda \in \mathbb{K}$ is an eigenvalue of A *if and only if* λ is a root of $\chi_A(t)$.

Criteria for Diagonalizability

Definition

A square matrix $A \in \mathbb{K}^{n,n}$ is called diagonalizable if there exists $S \in \mathrm{GL}(\mathbb{K}; n)$ and $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ such that

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Proposition

Let $A \in \mathbb{K}^{n,n}$.

- A is diagonalizable *if and only if* $\mathbb{K}^n$ possesses a basis of eigenvectors of A.
- If $\chi_A(t)$ has n distinct roots, then A is diagonalizable.

Minimal Polynomial

Definition

Let $A \in \mathbb{K}^{n,n}$. The minimal polynomial $\mu_A(t) \in \mathbb{K}[t]$ of A is the monic polynomial of smallest degree that satisfies $\mu_A(A) = 0$.

Proposition

Let $A \in \mathbb{K}^{n,n}$.

- The characteristic polynomial $\chi_A(t)$ is a multiple of $\mu_A(t)$, *i.e.*, $\exists q(t) \in \mathbb{K}[t]$ with $\chi_A(t) = q(t)\mu_A(t)$.
- A is diagonalizable *if and only if* $\mu_A(t)$ splits over $\mathbb{K}$ and has distinct roots, *i.e.*, $\exists m < n, \lambda_i \in \mathbb{K}$ with $\lambda_i \neq \lambda_j$ s.t. $\mu_A(t) = \prod_{i=1}^{m}(t - \lambda_i)$.

Corollary

Every projector $P \in \mathrm{Lin}(V, V)$, $\dim(V) < \infty$, is diagonalizable with eigenvalues in $\{0, 1\}$.

Jordan Decomposition

Definition

The $n \times n$ Jordan block for $\lambda \in \mathbb{K}$ is $J_{\lambda,n} := \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathbb{K}^{n,n}$.

Theorem (Jordan Decomposition)

Let $A \in \mathbb{K}^{n,n}$ and assume that $\chi_A(t)$ splits over $\mathbb{K}$. Then there is a $k \in \mathbb{N}$, $(\lambda_i, n_i) \in \mathbb{C} \times \mathbb{N}$ and $S \in \mathrm{GL}(\mathbb{K}; n)$ s.t.

$$S^{-1}AS = \begin{pmatrix} J_{\lambda_1,n_1} & & 0 \\ & \ddots & \\ 0 & & J_{\lambda_k,\lambda_k} \end{pmatrix}.$$

Jordan Decomposition: Some Corollaries

Theorem (Fundamental Theorem of Algebra)

Every polynomial with coefficients in $\mathbb{C}$ splits, i.e., for every $p(t) \in \mathbb{C}[t]$ there exist $c, \lambda_1, \cdot, \lambda_n \in \mathbb{C}$ with $p(t) = c(t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)$.

Corollaries

- Every $A \in \mathbb{C}^{n,n}$ possesses a Jordan decomposition.
- For $A \in \mathbb{C}^{n,n}$ holds $\det(A) = \prod_{i=1}^{k} \lambda_i^{n_i}$ and $\mathrm{tr}(A) = \sum_{i=1}^{k} n_i \lambda_i$.

Jordan Decomposition: Computation

Let $A \in \mathbb{K}^{n,n}$.

(i) Factorize $\chi_A(t)$: find $m, n_i \in \mathbb{N}$ and $\lambda_i \in \mathbb{K}$ s.t. $\chi_A(t) = \prod_{i=1}^{m}(t - \lambda_i)^{n_i}$.

(ii) $\forall i$: find $k_i \in \mathbb{N}$ s.t $(I_n - \lambda_i A)^{k_i - 1} \neq 0$ and $(I_n - \lambda_i A)^{k_i} = 0$.

(iii) $\forall i$: Compute a linear complement V_i of

$$\ker(I_n - \lambda_i A)^{k_i - 1} \subset \ker(I_n - \lambda_i A)^{k_i}.$$

(iv) $\forall i$: Compute a basis $v_1^{(i)}, \ldots, v_{l_i}^{(i)}$ of V_i.

(v) Construct the $S \in \mathrm{GL}(\mathbb{K}; n)$ with column vectors

$$v_j^{(i)}, \quad (\lambda_i I_n - A)v_j^{(i)}, \quad \ldots \quad , (\lambda_i I_n - A)^{k-1} v_j^{(i)}.$$

Refresher III
— Inner Product Spaces —

- Inner Products
- Orthogonal Complements, Orthonormal Basis
- Grahm–Schmidt Procedure
- QR decomposition, Least Squares Approximation
- Isometries, Self-Adjoint maps
- Spectral theorem
- Singular Value Decomposition

Inner Product Spaces

Definition

An inner product space is a vector space V over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ endowed with an inner product $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{K}$ that satisfies $\forall u, v, w \in V$ and $\lambda \in \mathbb{K}$:

- (symmetry) $\langle v, w \rangle = \overline{\langle w, v \rangle}$
- (linearity) $\langle v + \lambda w, u \rangle = \langle v, u \rangle + \lambda \langle w, u \rangle$
- (definiteness) $\langle v, v \rangle > 0$ if $v \neq 0$.

If $\mathbb{K} = \mathbb{R}$, V is also called an Euclidean space, and if $\mathbb{K} = \mathbb{C}$, V is an unitary space.

Note: For the rest of *Refresher III* we **always** assume $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$!

Inner Product Spaces: Examples

Example

The standard inner product on $\mathbb{K}^n \simeq \mathbb{K}^{n,1}$ is $\langle v, w \rangle = \bar{w}^T v = \sum_{i=1}^n \bar{v}_i w_i$.

Example

For $M \in \mathbb{K}^{n,n}$ let $M^\dagger := M^* := \bar{M}^T$. Then $\langle M, N \rangle = \mathrm{tr}(N^H M)$ is an inner product on $\mathbb{K}^{n,n}$.

Example

The L^2 product on the space of smooth functions $[0,1] \to \mathbb{C}$ is

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\, dt.$$

Orthogonality: Definitions

Definition

- $v, w \in V$ are orthogonal if $\langle v, w \rangle = 0$.
- A basis $\mathcal{B} \subset V$ is orthogonal if all pairs $v, w \in \mathcal{B}$, $v \neq w$, are orthogonal.
- This basis is orthonormal if moreover $1 = \|v\| := \sqrt{\langle v, v \rangle}$ for all $v \in \mathcal{B}$.
- The orthogonal complement of a subspace $U \subset V$ is

$$U^\perp := \{v \in V \mid \langle u, v \rangle = 0 \; \forall u \in U\}.$$

Orthogonality: Basic Properties

Proposition

- The orthogonal complement $U^\perp$ is always a linear subspace, and $U \cap U^\perp = \{0\}$.
- All pairs $v, w \in V$ satisfy the Cauchy-Schwarz inequality

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

- If $v, w \in V$ are orthogonal, then the Pythagorean theorem holds:

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Grahm–Schmidt Procedure

Theorem (Grahm–Schmidt)

Let $(V, \langle \cdot, \cdot \rangle)$ *be a finite dimensional inner product space with basis* $\{u_1, \ldots, u_n\}$*. We can inductively build an orthonormal basis* $\{q_1, \ldots, q_n\}$*:*

$$q_1 := {}^{u_1}/_{\|u_1\|},$$

$$q_i' := u_i - \sum_{k=1}^{i} \langle u_i, q_k \rangle \, q_k, \qquad q_i := {}^{q_i'}/_{\|q_i'\|}.$$

Orthogonal Projection

Proposition

Let $(V, \langle \cdot, \cdot \rangle)$ a finite dimensional inner product space. Every linear subspace $U \neq \{0\}$ possesses an orthonormal basis $\{q_1, \ldots, q_m\}$. The orthogonal projection onto U is given by

$$P_U(v) \coloneqq \sum_{i=1}^{m} \langle v, q_i \rangle \, q_i \qquad \forall v \in V.$$

It satisfies $P_U^2 = P_U$, $\mathrm{im}(P_U) = U$, and $\ker(P_U) = U^\perp$.

Orthogonal Projection: Examples

Exercise

Let $\lambda_1, \ldots, \lambda_m \in \mathbb{R}^m_+$ and $D \in \mathbb{R}^{m,m}$ be the diagonal matrix with $D_{ii} = \lambda_i$. Prove that $\langle v, w \rangle_D := w^* D v$ is an inner product on $\mathbb{K}^m$.

Proposition

Let $U \neq \{0\}$ be a subspace of $\mathbb{K}^m$ with $\langle \cdot, \cdot \rangle_D$-orthonormal basis $\{u_1, \ldots, u_n\}$. Then the matrix representation of P_U w.r.t. the standard basis of $\mathbb{K}^n$ is

$$[P_U] = \underbrace{\begin{pmatrix} | & & | \\ u_1 & \ldots & u_n \\ | & & | \end{pmatrix}}_{:=A} \begin{pmatrix} — & u_1^* & — \\ & \vdots & \\ — & u_n^* & — \end{pmatrix} D.$$

If $\{u_1, \ldots, u_n\}$ is a general basis, this becomes $[P_U] = A(A^* D A)^{-1} A^* D$.

Orthogonal Projection: Corollaries

Corollary

Let $(V, \langle \cdot, \cdot \rangle)$ be a finite dimensional inner product space and $U \subset V$ a linear subspace. Then $V = U \oplus U^{\perp}$.

Proposition (Distance Minimization)

Let $(V, \langle \cdot, \cdot \rangle)$ a finite dimensional inner product space and $U \neq \{0\}$ a subspace. For every $v \in V$ holds

$$\operatorname{argmin}_{u \in U} \|v - u\| = P_U(v).$$

Orthogonal & Unitary Matrices

Definition

- $Q \in \mathbb{R}^{n,n}$ is called orthogonal if $Q^T Q = I_n$. The set of orthogonal matrices is $\mathrm{O}(n)$ and the subset of matrices with $\det(Q) = 1$ is $\mathrm{SO}(n)$.
- $U \in \mathbb{C}^{n,n}$ is called unitary if $U^* U = I_n$. The set of unitary matrices is $\mathrm{U}(n)$ and the subset of matrices with $\det(U) = 1$ is $\mathrm{SU}(n)$.

Proposition

- Every orthogonal Q is unitary.
- A matrix $\mathbb{K}^{n,n}$ is unitary *if and only if* its columns are orthonormal w.r.t. the standard inner product on $\mathbb{K}^n$.
- $\mathrm{O}(n)$, $\mathrm{SO}(n)$, $\mathrm{U}(n)$, and $\mathrm{SU}(n)$ are subgroups of $\mathrm{GL}(n)$.

QR Decomposition

Theorem (QR Decomposition)

Let $A \in \mathbb{K}^{m,n}$.

- *There always exists a unitary matrix $Q \in \mathrm{U}(m)$ and an upper triangular matrix $R \in \mathbb{K}^{m,n}$, i.e., $R_{ij} = 0$ if $i > j$, such that $A = QR$.*
- *If A has full rank, i.e., $\mathrm{rank}(A) = \min\{m, n\}$, then this QR decomposition is unique.*
- *If A has full rank and $m \geq n$, then there is $Q \in \mathbb{K}^{m,n}$ and $R \in \mathrm{GL}(\mathbb{K}; n)$ such that $A = QR$ and the columns of Q are an orthonormal basis for $\mathrm{im}(Q)$.*
  This is called the reduced QR decomposition.

Least Squares Approximation

Proposition

Let $m \geq n$, $A \in \mathbb{R}^{m,n}$, and $b \in \mathbb{R}^m$. Denote by $Q \in \mathbb{R}^{m,n}$ and $R \in \mathrm{GL}(\mathbb{K}; n)$ the reduced QR decomposition of A. Then the least squares approximation of the linear system "$Ax = b$" is

$$\operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\| = R^{-1} Q^* b.$$

Note: Computing an inverse is usually numerically unstable. It should be avoided if possible. Luckily R is an upper triangle matrix. Thus, we can find $x \in \mathbb{R}^n$ with $Rx = Q^* b$ by backward substitution.

Polynomial Regression

Definition

Let $n, m \in \mathbb{N}$ and $x_0 < x_1 < \ldots < x_m$. The Vandermonde matrix of these points is

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \ldots & x_m^n \end{pmatrix} \in \mathbb{R}^{(m+1),(n+1)}$$

Proposition

The Vandermonde matrix has full rank.

Polynomial Regression

Theorem

Let $m \geq n$, $x_0 < x_1 < \ldots < x_m$, *and* $y \in \mathbb{R}^{m+1}$. *For* $\alpha \in \mathbb{R}^{n+1}$ *let* $p_\alpha(x) = \sum_{i=0}^{n} \alpha_i x^i$. *Let* $Q \in \mathbb{R}^{m+1,n+1}$ *and* $R \in \mathrm{GL}(\mathbb{R}; n+1)$ *be the reduced QR decomposition of the Vandermonde matrix of the* x_i. *Then*

$$\operatorname{argmin}_{\alpha \in \mathbb{R}^{n+1}} \sum_{i=0}^{m} |y_i - p_\alpha(x_i)|^2 = R^{-1} Q^T y.$$

Isometries

Definition

A linear map $f \in \mathrm{Lin}(V, W)$ between inner product spaces $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ is an isometry if

$$\langle f(v_1), f(v_2) \rangle_W = \langle v_1, v_2 \rangle_V \qquad \forall v_1, v_2 \in V.$$

Isometries: Properties

Proposition

Let $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ be inner product spaces, and $f \in \mathrm{Lin}(V, W)$. The following are equivalent:

- f is an isometry.
- For all $v \in V$: $\|v\|_V = 1 \Rightarrow \|f(v)\|_W = 1$.
- For all $v \in V$: $\|v\|_V = \|f(v)\|_W$.
- If $\{b_1, \ldots, b_n\}$ is orthonormal in V, then $\{f(b_1), \ldots, f(b_n)\}$ in W orthonormal.

Isometries: Corollaries

Corollary

Every isometry $f \in \mathrm{Lin}(V, W)$ is injective.

Corollary

Let $f \in \mathrm{Lin}(V, V)$ be an isometry and $\mathcal{B} = \{b_1, \dots, b_n\}$ be an orthonormal basis of V. Then

- $\Phi_V : V \longrightarrow \mathbb{K}^n$ given by $b_i \mapsto e_i$ is an isometry, where e_i are the standard basis vectors of $\mathbb{K}^n$ which we equip with the standard dot product.
- The matrix representation $[f]_{\mathcal{B}\mathcal{B}}$ of f is unitary, *i.e.*, $[f]_{\mathcal{B}\mathcal{B}} \in \mathrm{U}(n)$.
- f is invertible with $[f^{-1}]_{\mathcal{B}\mathcal{B}} = [f]_{\mathcal{B}\mathcal{B}}^{-1} = [f]_{\mathcal{B}\mathcal{B}}^*$.

The Adjoint Map

Definition

Let $f \in \mathrm{Lin}(V, W)$ be a map between inner product spaces.

- A map $f^* \in \mathrm{Lin}(W, V)$ is called adjoint to f if

  $$\langle f(v), w \rangle_W = \langle v, f^*(w) \rangle_V$$

  for all $v \in V$ and $w \in W$.
- If $W = V$ and $f = f^*$, then we call f self-adjoint.

The Adjoint Map: Properties

Proposition

Let $f \in \mathrm{Lin}(V, W)$.

- f possesses at most one adjoint.
- If f^* is the adjoint of f, then f is the adjoint of f^*, *i.e.*, $(f^*)^* = f$.
- If f is an isometry and $\dim(V) = \dim(W) < \infty$, then f^{-1} is the adjoint of f.

The Adjoint Map: Properties

Proposition

Let $f \in \mathrm{Lin}(V, W)$.

- If $\{b_1, \ldots, b_n\}$ is an orthonormal basis of V, then the adjoint of f is given by
$$f^*(w) := \sum_{i=1}^{n} \langle w, f(b_i) \rangle_W \, b_i \qquad \forall w \in W.$$
- The matrix representation of the adjoint w.r.t. (finite) orthonormal bases $\mathcal{B}_V$ and $\mathcal{B}_W$ satisfies $[f^*]_{\mathcal{B}_W \mathcal{B}_V} = [f]^*_{\mathcal{B}_V \mathcal{B}_W}$.
- If $\dim(V) < \infty$, f is self-adjoint if and only if $[f]_{\mathcal{B}_V \mathcal{B}_V}$ is hermitian, *i.e.*, $[f]_{\mathcal{B}_V \mathcal{B}_V} = [f]^*_{\mathcal{B}_V \mathcal{B}_V}$. If $\mathbb{K} = \mathbb{R}$, this implies that $[f]_{\mathcal{B}_V \mathcal{B}_V}$ is symmetric, *i.e.*, $[f]_{\mathcal{B}_V \mathcal{B}_V} = [f]^T_{\mathcal{B}_V \mathcal{B}_V}$.

Spectral Theorem

Theorem (Spectral Theorem)

Let V *be an inner product space over* $\mathbb{C}$ *with* $\dim(V) < \infty$*, and* $f \in \mathrm{Lin}(V, V)$*. The following are equivalent*

- $f \circ f^* = f^* \circ f$*.*
- *There exists an orthonormal basis of* V *that only consists of eigenvectors of* f*.*

Spectral Theorem: Corollaries

Proposition

If $f \in \mathrm{Lin}(V, V)$ is self-adjoint or an isometry of the finite dimensional vector space V over $\mathbb{C}$, then V possess an orthonormal basis that only consists of eigenvectors of f.

Proposition

- For every hermitian matrix $A \in \mathbb{C}^{n,n}$, *i.e.*, $A = A^*$, there exists a diagonal matrix $D \in \mathbb{R}^{n,n}$ and a unitary matrix $U \in \mathrm{U}(n)$ such that $A = UDU^*$.
- For every symmetric matrix $A \in \mathbb{R}^{n,n}$, *i.e.*, $A = A^T$, there exists a diagonal matrix $D \in \mathbb{R}^{n,n}$ and an orthogonal matrix $Q \in \mathrm{O}(n)$ such that $A = QDQ^T$.

Spectral Theorem: Corollaries

Proposition

Let $A \in \mathbb{K}^{n,n}$ be a hermitian matrix with all eigenvalues $\lambda_i \in \mathbb{R}_+$. Then

$$\langle v, w \rangle_A \coloneqq w^* A v, \qquad v, w \in \mathbb{K}^n$$

is an inner product of $\mathbb{K}^n$. Moreover, all inner products are of this form.

Spectral Theorem: Corollaries

Proposition

Let $A \in \mathbb{K}^{n,n}$ be hermitian and let $\lambda_1 \geq \ldots \geq \lambda_n$ be its eigenvalues.

- There exits $U \in \mathrm{U}(n)$ such that the kth power A^k of A is given by

$$A^k = U \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} U^*.$$

- If $0 < \lambda_n$, then the square root of A, *i.e.*, the matrix $B \in \mathbb{C}^{n,n}$ with $B^2 = A$, is given by

$$B = U \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} U^*.$$

Singular Value Decomposition

Theorem

Let $A \in \mathbb{C}^{m,n}$. Then there exit $U \in \mathrm{U}(m)$, $V \in \mathrm{U}(n)$, and a diagonal matrix $\Sigma \in \mathbb{C}^{m,n}$ such that

$$A = U\Sigma V^*.$$

If $A \in \mathbb{R}^{m,n}$ we can replace unitary by orthogonal matrices.
The diagonal values $\sigma \in \mathbb{R}_+$ of Σ are called singular values of A.

Refresher IV

— Normed Vector Spaces —

- Norms on Vector Spaces
- p-Norms, Operator Norm, Matrix Norms

Norms on Vector Spaces

Definition

A normed vector space V is a vector space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ endowed with a map $\|\cdot\|: V \longrightarrow \mathbb{R}$ called a norm, which satisfies $\forall\ u, v \in V$ and $\lambda \in \mathbb{K}$:

- (Δ-inequality) $\|u + v\| \leq \|u\| + \|v\|$
- (homogeneity) $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$
- (definiteness) $\|v\| \geq 0$ with $\|v\| = 0$ if and only if $v = 0$.

Norms on Vector Spaces

Proposition

- If $(V, \langle\cdot,\cdot\rangle)$ is a inner product space, then $\|v\| := \sqrt{\langle v, v\rangle}$ defines a norm on V.
- Let $(W, \|\cdot\|)$ be a normed vector space and $f \in \mathrm{Lin}(V, W)$ be injective. Then $\|v\|_f := \|f(v)\|_W$ defines a norm on V.

Example

The space of $\mathbb{K}_{\leq n}[x]$ can be equipped with the norm

$$\|a_n x^n + \cdots + a_1 x + a_0\|_\infty := \max\{|a_n|, \ldots, |a_1|, |a_0|\}.$$

p-Norms

Definition

Let $p \geq 1$ be a real number. Then the p-norm on $\mathbb{K}^n$ is defined as

$$\|x\|_p := (|x_1|^p + \cdots + |x_n|^p)^{1/p}$$

The ∞-norm (also maximum norm or supremum norm) is

$$\|x\|_\infty := \max\{|x_1|, \ldots, |x_n|\}.$$

Note:
The triangle inequality for p-norms, $p \in [1, \infty]$, is also called Minkowski inequality.

p-Norms: Hölder's inequality

Theorem (Hölder's inequality)

Let $p, q \in [1, \infty]$ *such that* ${}^1/_p + {}^1/_q = 1$. *(By convention* ${}^1/_\infty = 0$.*) Then*

$$\sum_{i=1}^{n} |x_i y_i| \leq \|x\|_p \|y\|_q$$

for all $x, y \in \mathbb{K}^n$.

Note:

- Hölder's inequality can be used to prove the Minkowski inequality.
- For $p = q = {}^1/_2$ this implies the Cauchy-Schwarz inequality.

Norms & Inner Products

Proposition

Let V be a vector space over $\mathbb{R}$. A norm $\|\cdot\|$ is induced by a inner product $\langle\cdot,\cdot\rangle$, *i.e.*, $\|v\| = \sqrt{\langle v, v\rangle}$, if and only if it satisfies the parallelogram identity

$$\|u+v\| + \|u-v\| = 2\|u\| + 2\|v\|.$$

The inner product can then be reconstructed via the polarization identity

$$\langle u, v\rangle = \frac{1}{2}\left(\|u+v\|^2 - \|u-v\|^2\right).$$

Corollary

Under all p-norms, only the 2-norm, *i.e.*, the standard norm, on $\mathbb{R}^n$ comes from a inner product.

Equivalence of Norms

Theorem

Let V be finite dimensional. Then all norms of V are equivalent, i.e., if $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are norms on V then there exist positive numbers $C, D \in \mathbb{R}_+$ such that

$$C\|v\|_\alpha \leq \|v\|_\beta \leq D\|v\|_\alpha$$

for all $v \in V$.

Operator Norm

Proposition

Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed vector spaces. Then the space $\mathrm{Lin}(V, W)$ is also a normed space with norm

$$\|f\|_{V,W} \coloneqq \sup_{v \neq 0} \frac{\|f(v)\|_W}{\|v\|_V} = \sup_{\|v\|_V = 1} \|f(v)\|_W,$$

which is called the operator norm.

proposition

The operator norm satisfies for all $v \in V$, $f \in \mathrm{Lin}(V, W)$ and $g \in \mathrm{Lin}(U, V)$:

- $\|f(v)\|_W \leq \|f\|_{V,W} \, \|v\|_V$,
- $\|f \circ g\|_{U,W} \leq \|f\|_{V,W} \, \|g\|_{U,V}$.

Norms for Matrices

Definition

A norm $\|\cdot\|$ on $\mathbb{K}^{n,n}$ is a matrix norm if

$$\|AB\| \leq \|A\| \, \|B\| \qquad \forall A, B \in \mathbb{K}^{n,n}.$$

Example

Interpreting matrices as linear maps and using the operator norm together with the p-norms, we obtain matrix norms $\|\cdot\|_p = \|\cdot\|_{p,p}$ on $\mathbb{K}^n$. For $p = 2$, it is called spectral norm.

Example

We saw that $\langle M, N \rangle = \operatorname{tr}(N^* M)$ is an inner product on $\mathbb{K}^{n,n}$. Its induced norm $\|\cdot\|_F$ is a matrix norm called Frobenius norm.

Spectral & Frobenius norm

Properties

Let $A \in \mathbb{K}^{m,n}$ with singular values $\sigma_{\max}(A) = \sigma_1 \geq \ldots \geq \sigma_r$, $r = \dim \operatorname{im}(A)$. Then

- $\|A\|_2 = \sigma_{\max}(A)$,
- $\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$,
- $\|A\|_2 \leq \|A\|_F$.