General Solution to a Linear System

If $\boldsymbol{b} \in \mathcal{C}(\mathbf{A})$: $\boldsymbol{x}_g = \boldsymbol{x}_p + \boldsymbol{x}_n$ where $\boldsymbol{x}_p \in \mathbb{R}^n$ and $\boldsymbol{x}_n \in \mathcal{N}(\mathbf{A})$.

Minimum Norm Solution

$\boldsymbol{x}_r \in \mathcal{C}(\mathbf{A}^\top)$ where $\boldsymbol{x}_r = \text{proj}_{\mathcal{C}(\mathbf{A}^\top)}(\boldsymbol{x}_g)$.

Least Squares (LS)

If $\boldsymbol{b} \notin \mathcal{C}(\mathbf{A})$: $\boldsymbol{x} = \arg\min_{\boldsymbol{x}^* \in \mathbb{R}^n} \|\boldsymbol{b} - \mathbf{A}\boldsymbol{x}^*\|$.
$\boldsymbol{b} - \mathbf{A}\boldsymbol{x} \in \mathcal{N}(\mathbf{A}) \implies \mathbf{A}^\top(\boldsymbol{b} - \mathbf{A}\boldsymbol{x}) = \boldsymbol{0}$.
$\quad \mathbf{A}^\top \mathbf{A}\boldsymbol{x} = \mathbf{A}^\top \boldsymbol{b}$ (Normal Equations)

Orthogonal Projection

$$\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\top \mathbf{A}\right)^{-1} \mathbf{A}^\top$$
$$\mathbf{P}\boldsymbol{b} = \text{proj}_{\mathcal{C}(\mathbf{A})}(\boldsymbol{b}) = \mathbf{A}\boldsymbol{x}$$

$\mathbf{P}$ is idempotent ($\mathbf{P}^2 = \mathbf{P}$) and $\mathbf{P}^\top = \mathbf{P}$.

Dependent Columns

If $\text{nullity}(\mathbf{A}) > 0$, NE yields infinitely many solutions as $\mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A}^\top \mathbf{A})$.

Orthogonal Complement Projections

Given $\mathbf{P} = \text{proj}_V$: $\mathbf{Q} = \text{proj}_{V^\perp} = \mathbf{I} - \mathbf{P}$
$\quad \boldsymbol{b} = \text{proj}_V(\boldsymbol{b}) + \text{proj}_{V^\perp}(\boldsymbol{b}) = \mathbf{P}\boldsymbol{b} + \mathbf{Q}\boldsymbol{b}$
$\quad (\mathbf{P}\boldsymbol{b})^\top \mathbf{Q}\boldsymbol{b} = 0$
$\quad\quad \mathbf{PQ} = \mathbf{0}$ (zero matrix)

Change of Basis

Given the basis $W = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$
$$\boldsymbol{b} = c_1 \boldsymbol{w}_1 + \cdots + c_n \boldsymbol{w}_n$$
$$\boldsymbol{b} = \mathbf{W}\boldsymbol{c} \iff (\boldsymbol{b})_W = \boldsymbol{c}.$$

Orthonormal Basis

Normalised and orthogonal basis vectors.
For $Q = \{\boldsymbol{q}_1, \dots, \boldsymbol{q}_n\}$, $\boldsymbol{q}_i^\top \boldsymbol{q}_j = \delta_{ij}$, where
$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$
$$\mathbf{Q}\boldsymbol{c} = \boldsymbol{b} \iff \mathbf{Q}^\top \boldsymbol{b} = \boldsymbol{c} = (\boldsymbol{b})_Q$$

Orthogonal Matrices

$$\mathbf{Q}^\top = \mathbf{Q}^{-1} \iff \mathbf{Q}^\top \mathbf{Q} = \mathbf{Q}\mathbf{Q}^\top = \mathbf{I}.$$

Projection onto a Vector

$$\text{proj}_{\boldsymbol{a}}(\boldsymbol{b}) = \boldsymbol{a}\left(\boldsymbol{a}^\top \boldsymbol{a}\right)^{-1} \boldsymbol{a}^\top \boldsymbol{b}$$
$$= \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|^2} \boldsymbol{a} \cdot \boldsymbol{b}$$

Using a unit vector $\boldsymbol{q}$:
$$\text{proj}_{\boldsymbol{q}}(\boldsymbol{b}) = \boldsymbol{q}(\boldsymbol{q} \cdot \boldsymbol{b})$$

Gram-Schmidt Process

Converts the basis W that spans $\mathcal{C}(\mathbf{A})$ to an orthonormal basis Q.

$$\boldsymbol{v}_i = \boldsymbol{w}_i - \sum_{j=1}^{i-1} \boldsymbol{q}_j \langle \boldsymbol{q}_j, \boldsymbol{w}_i \rangle \quad \boldsymbol{q}_i = \boldsymbol{v}_i / \|\boldsymbol{v}_i\|$$

V and Q span W, and V is orthogonal.

QR Decomposition

$\mathbf{A} = \mathbf{QR}$

$$\mathbf{R} = \begin{bmatrix} \|\boldsymbol{v}_1\| & \langle \boldsymbol{q}_1, \boldsymbol{w}_2 \rangle & \cdots & \langle \boldsymbol{q}_1, \boldsymbol{w}_n \rangle \\ 0 & \|\boldsymbol{v}_2\| & \ddots & \vdots \\ \vdots & \ddots & \ddots & \langle \boldsymbol{q}_{n-1}, \boldsymbol{w}_n \rangle \\ 0 & \cdots & 0 & \|\boldsymbol{v}_n\| \end{bmatrix}$$

where $\mathbf{Q}$ is found by applying the Gram-Schmidt process and $\mathbf{R}$ is upper triangular. $\mathbf{R}\boldsymbol{x} = \mathbf{Q}^\top \boldsymbol{b}$ solves LS.

Eigenvalues and Eigenvectors

$$\mathbf{A}\boldsymbol{v} = \lambda \boldsymbol{v} \iff (\lambda \mathbf{I} - \mathbf{A})\boldsymbol{v} = \boldsymbol{0} : \boldsymbol{v} \neq \boldsymbol{0}$$

Characteristic Polynomial

$$P(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}) = 0.$$

Eigen Decomposition

$$\mathbf{AV} = \mathbf{VD} \iff \mathbf{A} = \mathbf{VDV}^{-1}$$
$$\mathbf{V} = \begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{bmatrix}$$
$$\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Algebraic Multiplicity $\mu(\lambda_i)$

Multiplicity of λ_i in $P(\lambda)$, for $d \leq n$ distinct eigenvalues,
$$P(\lambda) = (\lambda - \lambda_1)^{\mu(\lambda_1)} \cdots (\lambda - \lambda_d)^{\mu(\lambda_d)}.$$
In general
$$1 \leq \mu(\lambda_i) \leq n$$
$$\sum_{i=1}^{d} \mu(\lambda_i) = n$$
If $\text{nullity}(\mathbf{A}) > 0$
$$\exists k : \lambda_k = 0 : \mu(\lambda_k) = \text{nullity}(\mathbf{A})$$

Geometric Multiplicity $\gamma(\lambda_i)$

Dimension of the eigenspace associated with λ_i.
$$\gamma(\lambda_i) = \text{nullity}(\lambda_i \mathbf{I} - \mathbf{A}).$$
Given $d \leq n$ distinct eigenvalues,
$$1 \leq \gamma(\lambda_i) \leq \mu(\lambda_i) \leq n$$
$$d \leq \sum_{i=1}^{d} \gamma(\lambda_i) \leq n.$$

Eigenvectors corresponding to distinct eigenvalues are linearly dependent.

Defective Matrix

$\mathbf{A}$ lacks a complete eigenbasis:
$$\exists k : \gamma(\lambda_k) < \mu(\lambda_k)$$

Matrix Similarity

$\mathbf{A}$ and $\mathbf{B}$ are similar if
$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}.$$
They share $P(\lambda)$, ranks, determinants, traces, and eigenvalues (also μ and γ).

Symmetric Matrices $\mathbf{S}^\top = \mathbf{S}$

$\mathbf{S}$ is always diagonalisable and has real eigenvalues with real orthogonal eigenvectors: $\mathbf{S} = \mathbf{QDQ}^\top$, where $\mathbf{Q}$ is found through QR: $\mathbf{V} = \mathbf{QR}$.

Skew-Symmetric Matrices $\mathbf{K}^\top = -\mathbf{K}$

Eigenvalues are always purely imaginary.

Positive-Definite Matrices

$\mathbf{S}$ is (symmetric) positive definite (SPD) if all its eigenvalues are positive, likewise
$$\boldsymbol{x}^\top \mathbf{S}\boldsymbol{x} > 0 : \forall \boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$$

Matrix Functions

Given a nondefective matrix:
$f(\mathbf{A}) = \mathbf{V}f(\mathbf{D})\mathbf{V}^{-1}$
$$= \mathbf{V} \text{diag}(f(\lambda_1), \dots, f(\lambda_n))\mathbf{V}^{-1}.$$
for an analytic function f.

Cayley-Hamilton Theorem

$$\forall \mathbf{A} : P(\mathbf{A}) = \mathbf{0} \quad \text{(zero matrix)}$$

Singular Value Decomposition

$$\mathbf{AV} = \mathbf{U}\boldsymbol{\Sigma} \iff \mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$$
$$\mathbf{V}^\top = \mathbf{V}^{-1}, \quad \mathbf{U}^\top = \mathbf{U}^{-1}$$
$$\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0).$$
Left singular vectors $\boldsymbol{u}$: $\mathbf{U} \in \mathbb{R}^{m \times m}$
$$\mathcal{C}(\mathbf{A}) = \text{span}(\{\boldsymbol{u}_{i \leq r}\})$$
$$\mathcal{N}(\mathbf{A}^\top) = \text{span}(\{\boldsymbol{u}_{r < i \leq m}\})$$
Right singular vectors $\boldsymbol{v}$: $\mathbf{V} \in \mathbb{R}^{n \times n}$
$$\mathcal{C}(\mathbf{A}^\top) = \text{span}(\{\boldsymbol{v}_{i \leq r}\})$$
$$\mathcal{N}(\mathbf{A}) = \text{span}(\{\boldsymbol{v}_{r < i \leq n}\})$$
Singular values σ_i: $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times n}$
The eigenvalues of $\mathbf{A}^\top \mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ are equal, $\boldsymbol{\Sigma}^\top \boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$ have the same diagonal entries, and when $m = n$, $\boldsymbol{\Sigma}^\top \boldsymbol{\Sigma} = \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top = \boldsymbol{\Sigma}^2$. To find σ_i compute:
$$\mathbf{A}^\top \mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}^\top \boldsymbol{\Sigma}\mathbf{V}^\top$$
$$\mathbf{A}\mathbf{A}^\top = \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \mathbf{U}^\top$$
so that $\sigma_i = \sqrt{\lambda_i}$ where $\sigma_1 \geq \cdots \geq \sigma_r > 0$.

Reduced SVD

Ignores $m - n$ "0" rows in $\boldsymbol{\Sigma}$ so that $\mathbf{U} \in \mathbb{R}^{m \times n}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$, and $\mathbf{V} \in \mathbb{R}^{n \times n}$.

Pseudoinverse

Consider the inverse mapping $\boldsymbol{u}_i \mapsto \frac{1}{\sigma_i}\boldsymbol{v}_i$
$$\mathbf{A}^\dagger \boldsymbol{u}_i = \frac{1}{\sigma_i}\boldsymbol{v}_i \iff \mathbf{A}^\dagger \boldsymbol{u}_i = \frac{1}{\sigma_i}\boldsymbol{v}_i \boldsymbol{u}_i^\top \boldsymbol{u}_i$$
$$\mathbf{A}^\dagger = \sum_{i=1}^{r} \frac{1}{\sigma_i}\boldsymbol{v}_i \boldsymbol{u}_i^\top \iff \mathbf{A}^\dagger = \mathbf{V}\boldsymbol{\Sigma}^\dagger \mathbf{U}^\top$$
where $\boldsymbol{\Sigma}^\dagger = \text{diag}\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0\right)$.
$\boldsymbol{x} = \mathbf{A}^\dagger \boldsymbol{b}$ solves LS.

Truncated SVD

Express $\mathbf{A}$ as the sum of rank-1 matrices:
$$\mathbf{A} = \sum_{i=1}^{n} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top \approx \tilde{\mathbf{A}} = \sum_{i=1}^{\nu} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top.$$
for the rank-ν approximation of $\mathbf{A}$.
Using the SVD:
$$\tilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$$
$\mathbf{U} \in \mathbb{R}^{m \times \nu}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{\nu \times \nu}$, and $\mathbf{V} \in \mathbb{R}^{n \times \nu}$.
When $\nu \geq r$, $\tilde{\mathbf{A}} = \mathbf{A}$ as $\sigma_{i > r} = 0$.

General Vector Spaces

V is a vector space with vectors $\boldsymbol{v} \in V$ if the following 10 axioms are satisfied for $\forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ and $\forall k, m \in \mathbb{F}$, given an addition and scalar multiplication operation.

For the addition operation:

- Closure: $\boldsymbol{u} + \boldsymbol{v} \in V$
- Commutativity: $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u} \in V$
- Associativity:
  $$\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$$
- Identity: $\exists \boldsymbol{0} \in V : \boldsymbol{u} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{u} = \boldsymbol{u}$
- Inverse: $\exists(-\boldsymbol{u}) \in V : \boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$

For the scalar multiplication operation:

- Closure: $k\boldsymbol{u} \in V$
- Distributivity: $k(\boldsymbol{u} + \boldsymbol{v}) = k\boldsymbol{u} + k\boldsymbol{v}$
- Distributivity: $(k + m)\boldsymbol{u} = k\boldsymbol{u} + m\boldsymbol{u}$
- Associativity: $k(m\boldsymbol{u}) = (km)\boldsymbol{u}$
- Identity: $\exists 1 \in \mathbb{F} : 1\boldsymbol{u} = \boldsymbol{u}$

Examples of Vector Spaces

The set of all $m \times n$ matrices $\mathscr{M}_{mn}$ with matrix addition and scalar matrix multiplication.

The set of all functions $\mathscr{F}(\Omega) : \Omega \to \mathbb{R}$ with addition and scalar multiplication defined pointwise.

Subspaces

The subset $W \subset V$ is itself a vector space if it is closed under addition and scalar multiplication.

Examples of Subspaces

Subspaces of $\mathbb{R}^n$:

- Lines, planes and higher-dimensional analogues in $\mathbb{R}^n$ *passing through the origin.*

Subspaces of $\mathscr{M}_{nn}$:

- The set of all *symmetric* $n \times n$ matrices, denoted $\mathscr{S}_n \subset \mathscr{M}_{nn}$.
- The set of all *skew symmetric* $n \times n$ matrices, denoted $\mathscr{K}_n \subset \mathscr{M}_{nn}$.

Subspaces of $\mathscr{F}$:

- The set of all *polynomials* of degree n or less, denoted $\mathscr{P}_n(\Omega) \subset \mathscr{F}(\Omega)$.
- The set of all *continuous functions,* denoted $C(\Omega) \subset \mathscr{F}(\Omega)$.
- The set of all continuous functions with *continuous nth derivatives,* denoted $C^n(\Omega) \subset C(\Omega)$.
- The set of all functions f defined on $[0,1]$ satisfying $f(0) = f(1)$.

General Vector Space Terminology

Let $S = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ and $c_1, \dots, c_k \in \mathbb{F}$:

- The linear combination of S is a vector of the form $\boldsymbol{v} = c_1 \boldsymbol{v}_1 + \cdots + c_k \boldsymbol{v}_k$.
- S is linearly independent iff $c_1 \boldsymbol{v}_1 + \cdots + c_k \boldsymbol{v}_k = \boldsymbol{0}$ has the trivial solution.
- $\operatorname{span}(S)$ is the set of all linear combinations of S.

S is a *basis* for a vector space V if

- S is linearly independent.
- $\operatorname{span}(S) = V$.

The number of basis vectors denotes the dimension of V.
C is infinite dimensional.

Examples of Standard Bases

- $\mathscr{M}_{22}$:
$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$
- $\mathscr{S}_{22}$: $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$
- $\mathscr{K}_{22}$: $\left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$
- $\mathscr{P}_3$: $\{1,\, x,\, x^2,\, x^3\}$

Linear Transformations

$T : V \to W$ satisfying
$$T(k\boldsymbol{u}) = kT(\boldsymbol{u})$$
$$T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$$

Constructing $\mathbf{A} = (T)_{B',\,B}$:

Consider the map of $(\boldsymbol{v})_B = \boldsymbol{x}$ of $\boldsymbol{v} \in V$ to $(\boldsymbol{w})_{B'} = \boldsymbol{b}$ of $\boldsymbol{w} \in W$, where $B = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $B' = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$.
$$T(\boldsymbol{v}) = \boldsymbol{w}$$
$$[T(\boldsymbol{v}_1) \quad \cdots \quad T(\boldsymbol{v}_n)]\, \boldsymbol{x} = \mathbf{W}\boldsymbol{b}$$
$$\left[(T(\boldsymbol{v}_1))_{B'} \quad \cdots \quad (T(\boldsymbol{v}_n))_{B'}\right] \boldsymbol{x} = \boldsymbol{b}$$
$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$$

Isomorphism ($\cong$)

$T : V \to W$ is an isomorphism between V and W if there exists a bijection between the two vector spaces.
$\forall V : \dim(V) = n : V \cong \mathbb{R}^n$, $\mathscr{M}_{mn} \cong \mathbb{R}^{mn}$ and $\mathscr{P}_n \cong \mathbb{R}^{n+1}$.

Fundamental Subspaces of T

- The set of all vectors in V that map to W is the **image** of T, denoted $\operatorname{im}(T)$.
- The set of all vectors in W that is mapped to by a vector in V is the **range** of T, denoted $\operatorname{range}(T)$.
- The set of all vectors in V that T maps to $\boldsymbol{0}_W$ is the **kernel** of T, denoted $\ker(T)$.

If finite, $\dim(\operatorname{range}(T)) = \operatorname{rank}(T)$ and $\dim(\ker(T)) = \operatorname{nullity}(T)$.
$$\operatorname{rank}(T) + \operatorname{nullity}(T) = \dim(V).$$

Inner Product Spaces

$$\langle \cdot,\, \cdot \rangle : V \times V \to \mathbb{R}.$$

For $\boldsymbol{u},\, \boldsymbol{v},\, \boldsymbol{w} \in V$ and $k \in \mathbb{R}$:

- Symmetry: $\langle \boldsymbol{u},\, \boldsymbol{v} \rangle = \langle \boldsymbol{v},\, \boldsymbol{u} \rangle$
- Linearity:
  $$\langle \boldsymbol{u} + \boldsymbol{v},\, \boldsymbol{w} \rangle = \langle \boldsymbol{u},\, \boldsymbol{w} \rangle + \langle \boldsymbol{v},\, \boldsymbol{w} \rangle$$
- Linearity: $\langle k\boldsymbol{u},\, \boldsymbol{v} \rangle = k\langle \boldsymbol{u},\, \boldsymbol{v} \rangle$
- Positive semi-definitiveness:
  $$\langle \boldsymbol{u},\, \boldsymbol{u} \rangle \geq 0,\ \langle \boldsymbol{u},\, \boldsymbol{u} \rangle = 0 \iff \boldsymbol{u} = \boldsymbol{0}$$

For $\boldsymbol{u},\, \boldsymbol{v} \in \mathbb{R}^n$:

- $\langle \boldsymbol{u},\, \boldsymbol{v} \rangle = \boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^\top \boldsymbol{v}$.
- $\langle \boldsymbol{u},\, \boldsymbol{v} \rangle = \boldsymbol{u}^\top \mathbf{A} \boldsymbol{v}$ where $\mathbf{A}$ is SPD.

For $\mathbf{A},\, \mathbf{B} \in \mathscr{M}_{mn}$:

- $\langle \mathbf{A},\, \mathbf{B} \rangle = \operatorname{Tr}(\mathbf{A}^\top \mathbf{B})$.

For $f,\, g \in C([a,b])$:

- $\langle f,\, g \rangle = \int_a^b f(x) g(x)\, \mathrm{d}x$.
- $\langle f,\, g \rangle = \int_a^b f(x) g(x) w(x)\, \mathrm{d}x$.

where $w(x) > 0 : \forall x \in [a,b]$.

Norms

- $\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v},\, \boldsymbol{v} \rangle}$.
- $\|\boldsymbol{v}\| \geq 0$, and $\|\boldsymbol{v}\| = 0 \iff \boldsymbol{v} = \boldsymbol{0}$.
- $\|k\boldsymbol{v}\| = |k|\|\boldsymbol{v}\| : \forall k \in \mathbb{R}$.
- $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$.

Examples:

- $\forall \mathbf{A} \in \mathscr{M}_{mn} : \|\mathbf{A}\| = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$.
- $\forall f \in C([a,b]) : \|f\| = \sqrt{\int_a^b f(x)^2\, \mathrm{d}x}$.

Orthogonality

$$\langle \boldsymbol{v},\, \boldsymbol{v} \rangle = 0.$$

Orthogonal Complements of $\mathscr{M}_n$

Given $\mathbf{P}_{\mathscr{S}_n} = \operatorname{proj}_{\mathscr{S}_n}$ and $\mathbf{P}_{\mathscr{K}_n} = \operatorname{proj}_{\mathscr{K}_n}$
$$\mathbf{P}_{\mathscr{S}_n} = \mathbf{I} - \mathbf{P}_{\mathscr{K}_n}$$
$$\mathbf{S} = \mathbf{P}_{\mathscr{S}_n} \mathbf{M} = \operatorname{proj}_{\mathbf{P}_{\mathscr{S}_n}}(\mathbf{M}) = \frac{\mathbf{M} + \mathbf{M}^\top}{2}$$
$$\mathbf{K} = \mathbf{P}_{\mathscr{K}_n} \mathbf{M} = \operatorname{proj}_{\mathbf{P}_{\mathscr{K}_n}}(\mathbf{M}) = \frac{\mathbf{M} - \mathbf{M}^\top}{2}$$
$\mathbf{S} \in \mathscr{S}_n$, $\mathbf{K} \in \mathscr{K}_n$, and $\mathbf{S} + \mathbf{K} = \mathbf{M} \in \mathscr{M}_n$.

Theorems

- $\mathbf{A}^\top \mathbf{A}$ is always positive semi-definite, and $\mathcal{N}(\mathbf{A}^\top \mathbf{A}) = \mathcal{N}(\mathbf{A})$ so that $\operatorname{rank}(\mathbf{A}^\top \mathbf{A}) = \operatorname{rank}(\mathbf{A})$. $\mathbf{A}^\top \mathbf{A}$ is positive definite and $\mathbf{A}^\top \mathbf{A}$ is invertible when $\operatorname{nullity}(\mathbf{A}) = 0$.
- When $\mathbf{A}$ is square and invertible, $(\mathbf{A}^\top \mathbf{A})^{-1} = \mathbf{A}^{-1} \mathbf{A}^{-\top}$ and $\mathbf{P} = \mathbf{I}$.
- $\mathbf{P}^2 = \mathbf{P} \wedge \mathbf{P}^\top = \mathbf{P} \iff \mathbf{P} = \operatorname{proj}_{C(\mathbf{P})}$.
- If $\operatorname{nullity}(\mathbf{A}) = 0$, $\mathbf{P} = \mathbf{Q}\mathbf{Q}^\top$ using QR.
- $\mathbf{A}^\top \mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ share eigenvalues,
  $$\mathbf{A}^\top \mathbf{A}\boldsymbol{v} = \lambda \boldsymbol{v}$$
  $$(\mathbf{A}\mathbf{A}^\top)(\mathbf{A}\boldsymbol{v}) = \lambda(\mathbf{A}\boldsymbol{v}).$$
  $\mathbf{A}\boldsymbol{v} = \boldsymbol{0} \implies \lambda = 0$, else $\boldsymbol{w} = \mathbf{A}\boldsymbol{v}$ is an eigenvector of $\mathbf{A}\mathbf{A}^\top$.
- For symmetric $\mathbf{S} \in \mathbb{R}^{n \times n}$:
  $$\mathbf{S} = \sum_{i=1}^n \lambda_i \boldsymbol{q}_i \boldsymbol{q}_i^\top = \sum_{i=1}^n \lambda_i \operatorname{proj}_{\boldsymbol{q}_i}$$
- For $\boldsymbol{w} = \mathbf{W} \in \mathbb{R}^{n \times 1}$:
  $$\mathbf{W} = [\widehat{\boldsymbol{w}}]\, [\|\boldsymbol{w}\|]\, [1]$$
  $$\mathbf{W}^\dagger = \frac{\widehat{\boldsymbol{w}}^\top}{\|\boldsymbol{w}\|}$$

Identities

- $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$.
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$ if $\mathbf{A}$, $\mathbf{B}$ invertible.
- $(\mathbf{A}^\top)^{-1} = (\mathbf{A}^{-1})^\top$ if $\mathbf{A}$ invertible:
  $$\mathbf{A}^\top (\mathbf{A}^{-1})^\top = (\mathbf{A}^{-1}\mathbf{A})^\top = \mathbf{I}$$
  $$(\mathbf{A}^{-1})^\top \mathbf{A}^\top = (\mathbf{A}\mathbf{A}^{-1})^\top = \mathbf{I}$$
- $\langle \mathbf{A}\boldsymbol{x},\, \boldsymbol{y} \rangle = \langle \boldsymbol{x},\, \mathbf{A}^\top \boldsymbol{y} \rangle$:
  $$(\mathbf{A}\boldsymbol{x})^\top \boldsymbol{y} = \boldsymbol{x}^\top (\mathbf{A}^\top \boldsymbol{y})$$
- If $\mathbf{A}$ is triangular, $\det(\mathbf{A}) = \prod_{i=1}^n a_{ii}$.
- For $\mathbf{A} \in \mathbb{R}^{n \times n}$:
  $$\operatorname{Tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i$$
  $$\det(\mathbf{A}) = \prod_{i=1}^n \lambda_i$$
  $$\det(\mathbf{A}^\top \mathbf{A}) = \det(\mathbf{A})^2 = \prod_{i=1}^n \sigma_i^2$$
- For $\mathbf{A} \in \mathbb{R}^{m \times n}$:
  $$\operatorname{Tr}(\mathbf{A}^\top \mathbf{A}) = \sum_{j=1}^m \sum_{i=1}^n a_{ij}^2$$
  $$= \sum_{i=1}^n \sigma_i^2$$
  $$\det(\mathbf{A}^\top \mathbf{A}) = \prod_{i=1}^n \sigma_i^2$$