General Solution to a Linear System

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

Minimum Norm Solution

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

Least Squares (LS)

If $\mathbf{b} \notin \mathcal{C}(\mathbf{A})$: $\mathbf{x} = \underset{\mathbf{x}^* \in \mathbb{R}^n}{\arg\min} \|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|$.

$\mathbf{b} - \mathbf{A}\mathbf{x} \in \mathcal{N}(\mathbf{A}) \implies \mathbf{A}^\top(\mathbf{b} - \mathbf{A}\mathbf{x}) = \mathbf{0}$.

Orthogonal Projection

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

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

Dependent Columns

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

Orthogonal Complement Projections

Given $\mathbf{P} = \operatorname{proj}_V$: $\mathbf{Q} = \operatorname{proj}_{V^\perp} = \mathbf{I} - \mathbf{P}$

$$\mathbf{b} = \operatorname{proj}_V(\mathbf{b}) + \operatorname{proj}_{V^\perp}(\mathbf{b}) = \mathbf{P}\mathbf{b} + \mathbf{Q}\mathbf{b}$$
$$(\mathbf{P}\mathbf{b})^\top \mathbf{Q}\mathbf{b} = 0$$
$$\mathbf{P}\mathbf{Q} = \mathbf{0} \qquad \text{(zero matrix)}$$

Change of Basis

Given the basis $W = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$

$$\mathbf{b} = c_1\mathbf{w}_1 + \dots + c_n\mathbf{w}_n$$
$$\mathbf{b} = \mathbf{W}\mathbf{c} \iff (\mathbf{b})_W = \mathbf{c}.$$

Orthonormal Basis

Normalised and orthogonal basis vectors. For $Q = \{\mathbf{q}_1, \dots, \mathbf{q}_n\}$, $\mathbf{q}_i^\top\mathbf{q}_j = \delta_{ij}$, where

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$
$$\mathbf{Q}\mathbf{c} = \mathbf{b} \iff \mathbf{Q}^\top\mathbf{b} = \mathbf{c} = (\mathbf{b})_Q$$

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

$$\operatorname{proj}_{\mathbf{a}}(\mathbf{b}) = \mathbf{a}\left(\mathbf{a}^\top\mathbf{a}\right)^{-1}\mathbf{a}^\top\mathbf{b} = \frac{\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \cdot \mathbf{b}$$
$$\operatorname{proj}_{\mathbf{q}}(\mathbf{b}) = \mathbf{q}(\mathbf{q} \cdot \mathbf{b}) \qquad \text{(unit vector)}$$

Gram-Schmidt Process

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

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

V and Q span W, and V is orthogonal.

QR Decomposition

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

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

Eigenvalues and Eigenvectors

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

Characteristic Polynomial

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

Eigen Decomposition

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

Eigenspace

The eigenspace associated with λ_i is the span of eigenvectors: $\mathcal{N}(\lambda_i\mathbf{I} - \mathbf{A})$.

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 $\operatorname{nullity}(\mathbf{A}) > 0$

$$\exists k : \lambda_k = 0 : \mu(\lambda_k) = \operatorname{nullity}(\mathbf{A})$$

Geometric Multiplicity $\gamma(\lambda_i)$

The dimension of each eigenspace λ_i

$$\gamma(\lambda_i) = \operatorname{nullity}(\lambda_i\mathbf{I} - \mathbf{A}).$$

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 $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. 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 eigenspaces: $\mathbf{S} = \mathbf{Q}\mathbf{D}\mathbf{Q}^\top$, where $\mathbf{Q}$ is found through QR: $\mathbf{V} = \mathbf{Q}\mathbf{R}$.

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

$$\mathbf{x}^\top\mathbf{S}\mathbf{x} > 0 : \forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$$

Matrix Functions

Given a nondefective matrix:

$$f(\mathbf{A}) = \mathbf{V} f(\mathbf{D}) \mathbf{V}^{-1} = \mathbf{V} \operatorname{diag}(f(\lambda_1), \dots, f(\lambda_n)) \mathbf{V}^{-1}$$

for an analytic function f.

Cayley-Hamilton Theorem

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

Singular Value Decomposition

$$\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma} \iff \mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$$
$$\mathbf{V}^\top = \mathbf{V}^{-1}, \quad \mathbf{U}^\top = \mathbf{U}^{-1}$$
$$\mathbf{\Sigma} = \operatorname{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0).$$

Left singular vectors $\mathbf{u}$: $\mathbf{U} \in \mathbb{R}^{m \times m}$

$$\mathcal{C}(\mathbf{A}) = \operatorname{span}(\{\mathbf{u}_{i \leq r}\})$$
$$\mathcal{N}(\mathbf{A}^\top) = \operatorname{span}(\{\mathbf{u}_{r < i \leq m}\})$$

Right singular vectors $\mathbf{v}$: $\mathbf{V} \in \mathbb{R}^{n \times n}$

$$\mathcal{C}(\mathbf{A}^\top) = \operatorname{span}(\{\mathbf{v}_{i \leq r}\})$$
$$\mathcal{N}(\mathbf{A}) = \operatorname{span}(\{\mathbf{v}_{r < i \leq n}\})$$

Singular values σ_i: $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$

The eigenvalues of $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ are equal, $\mathbf{\Sigma}^\top\mathbf{\Sigma}$ and $\mathbf{\Sigma}\mathbf{\Sigma}^\top$ have the same diagonal entries, and when $m = n$, $\mathbf{\Sigma}^\top\mathbf{\Sigma} = \mathbf{\Sigma}\mathbf{\Sigma}^\top = \mathbf{\Sigma}^2$. To find σ_i compute:

$$\mathbf{A}^\top\mathbf{A} = \mathbf{V}\mathbf{\Sigma}^\top\mathbf{\Sigma}\mathbf{V}^\top$$
$$\mathbf{A}\mathbf{A}^\top = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^\top\mathbf{U}^\top$$

so that $\sigma_i = \sqrt{\lambda_i}$ where $\sigma_1 \geq \dots \geq \sigma_r > 0$.

Reduced SVD

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

Pseudoinverse

Consider the inverse mapping $\mathbf{u}_i \mapsto \frac{1}{\sigma_i}\mathbf{v}_i$

$$\mathbf{A}^\dagger\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{v}_i \iff \mathbf{A}^\dagger\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{v}_i\mathbf{u}_i^\top\mathbf{u}_i$$
$$\mathbf{A}^\dagger = \sum_{i=1}^{r} \frac{1}{\sigma_i}\mathbf{v}_i\mathbf{u}_i^\top \iff \mathbf{A}^\dagger = \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^\top$$

where $\mathbf{\Sigma}^\dagger = \operatorname{diag}\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0\right)$.

$\mathbf{x} = \mathbf{A}^\dagger\mathbf{b}$ also solves LS.

Truncated SVD

Express $\mathbf{A}$ as the sum of rank-1 matrices:

$$\mathbf{A} = \sum_{i=1}^{n} \sigma_i\mathbf{u}_i\mathbf{v}_i^\top \approx \tilde{\mathbf{A}} = \sum_{i=1}^{\nu} \sigma_i\mathbf{u}_i\mathbf{v}_i^\top$$

for the rank-ν approximation of $\mathbf{A}$. Using the SVD:

$$\tilde{\mathbf{A}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$$

$\mathbf{U} \in \mathbb{R}^{m \times \nu}$, $\mathbf{\Sigma} \in \mathbb{R}^{\nu \times \nu}$, 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 $\mathbf{v} \in V$ if the following 10 axioms are satisfied for $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\forall k, m \in \mathbb{F}$, given an addition and scalar multiplication operation.

For the addition operation:

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

For the scalar multiplication operation:

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

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

- The linear combination of S is a vector of the form $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$.
- S is linearly independent iff $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ 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

$$T(k\mathbf{u}) = kT(\mathbf{u})$$
$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$$

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

Consider the map of $(\mathbf{v})_B = \mathbf{x}$ of $\mathbf{v} \in V$ to $(\mathbf{w})_{B'} = \mathbf{b}$ of $\mathbf{w} \in W$, where $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $B' = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$.

$$T(\mathbf{v}) = \mathbf{w}$$
$$\begin{bmatrix}T(\mathbf{v}_1) & \cdots & T(\mathbf{v}_n)\end{bmatrix}\mathbf{x} = \mathbf{W}\mathbf{b}$$
$$\begin{bmatrix}(T(\mathbf{v}_1))_{B'} & \cdots & (T(\mathbf{v}_n))_{B'}\end{bmatrix}\mathbf{x} = \mathbf{b}$$
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

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 V that T maps to $\mathbf{0}_W$ is the **kernel** of T, denoted $\ker(T)$.

If finite, $\dim(\operatorname{im}(T)) = \operatorname{rank}(T)$ 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 $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $k \in \mathbb{R}$:

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

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

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

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

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

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

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 $\mathbf{P} = \mathbf{I}$ otherwise $\mathbf{P} = \mathbf{Q}\mathbf{Q}^\top$ using QR.
- $\mathbf{P}^2 = \mathbf{P} \wedge \mathbf{P}^\top = \mathbf{P} \iff \mathbf{P} = \operatorname{proj}_{\mathcal{C}(\mathbf{P})}$. $\mathbf{P}\mathbf{v} = \mathbf{P}^2\mathbf{v} \iff \lambda\mathbf{v} = \lambda^2\mathbf{v}$ implies $\lambda = 0, 1$.
- $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ share eigenvalues,
$$\mathbf{A}^\top\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$
$$(\mathbf{A}\mathbf{A}^\top)(\mathbf{A}\mathbf{v}) = \lambda(\mathbf{A}\mathbf{v}).$$
$\mathbf{A}\mathbf{v} = \mathbf{0} \implies \lambda = 0$, else $\mathbf{w} = \mathbf{A}\mathbf{v}$ is an eigenvector of $\mathbf{A}\mathbf{A}^\top$.
- For symmetric $\mathbf{S} \in \mathbb{R}^{n\times n}$:
$$\mathbf{S} = \sum_{i=1}^n \lambda_i \mathbf{q}_i\mathbf{q}_i^\top = \sum_{i=1}^n \lambda_i \operatorname{proj}_{\mathbf{q}_i}$$
- For $\mathbf{W} = \mathbf{w} \in \mathbb{R}^{n\times 1}$:
$$\mathbf{W} = [\hat{\mathbf{w}}]\,[\|\mathbf{w}\|]\,[1]$$
$$\mathbf{W}^\dagger = \hat{\mathbf{w}}^\top / \|\mathbf{w}\|$$

Identities

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