Numerical Linear Algebra for Computational Science and Information Engineering

Essentials of Linear Algebra

Nicolas Venkovic
nicolas.venkovic@tum.de

Chair of Computational Mathematics
School of Computation, Information and Technology
Technical University of Munich

Outline I

Vector spaces

Section 2.1 in Darve & Wootters (2021)

Vectors

- We are interested in vectors in vector spaces $\mathbb{F}^n$ with real ($\mathbb{F} := \mathbb{R}$) and complex ($\mathbb{F} := \mathbb{C}$) scalar coefficients.

  A vector $\mathbf{x} \in \mathbb{F}^n$ is an n-tuple given by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ with scalar coefficients } x_1, x_2, \ldots, x_n \in \mathbb{F}.$$

- The vector space $\mathbb{F}^n$ is said to support addition, and scalar multiplication. That is, for every pair $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ and $\alpha \in \mathbb{F}$, we have

$$\mathbf{x} + \alpha \mathbf{y} = \begin{bmatrix} x_1 + \alpha y_1 \\ x_2 + \alpha y_2 \\ \vdots \\ x_n + \alpha y_n \end{bmatrix} \in \mathbb{F}^n.$$

Vector spaces

- Vector spaces are fundamental structures of (numerical) linear algebra.

Definition (Vector space)

A vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$ is a non-empty set which supports addition and multiplication with the following axioms:

1. Associative addition: $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z} \;\forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$
2. Commutative addition: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \;\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{V}$
3. Additive identity: $\exists\, \mathbf{0} \in \mathcal{V}$ s.t. $\mathbf{x} + \mathbf{0} = \mathbf{x} \;\forall\, \mathbf{x} \in \mathcal{V}$
4. Additive inverse: $\forall\, \mathbf{x} \in \mathcal{V}$, $\exists\, -\mathbf{x} \in \mathcal{V}$ s.t. $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$
5. Field-multiplication compatibility: $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x} \;\forall\, \mathbf{x} \in \mathcal{V}, \alpha, \beta \in \mathbb{F}$
6. Field-multiplicative identity: $\exists\, 1 \in \mathbb{F}$ s.t. $1\mathbf{x} = \mathbf{x} \;\forall\, \mathbf{x} \in \mathcal{V}$
7. Distributive field multiplication w.r.t. vector addition:
$$\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y} \;\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{V}, \alpha \in \mathbb{F}$$
8. Distributive field multiplication w.r.t. field addition:
$$(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x} \;\forall\, \mathbf{x} \in \mathcal{V}, \alpha, \beta \in \mathbb{F}$$

Vector spaces, cont'd

- In practice, most axioms are trivially satisfied, and verifying whether $\mathcal{V}$ is a vector space boils down to checking for **closure under addition** and **scalar multiplication**, i.e., whether

$$\mathbf{x} + \mathbf{y} \in \mathcal{V} \text{ and } \alpha\mathbf{x} \in \mathcal{V} \;\forall\; \mathbf{x}, \mathbf{y} \in \mathcal{V}, \alpha \in \mathbb{F}.$$

Example and counter-examples

- $\mathbb{R}^n$ **is** a vector space over $\mathbb{R}$.
- $\mathcal{V} := \{\mathbf{x} \in \mathbb{R}^n \text{ s.t. } x_i > 0,\ i = 1, \dots, n\}$ **is not** a vector space.
- The set of floating-point numbers **does not** form a field, and cannot serve as the scalar field for a vector space.

Disclaimer

- The elements of a vector space $\mathcal{V}$ may not always be finite-dimensional vectors; they can also be **functions**, **even when $\mathcal{V}$ is finite-dimensional**.
- In this course, we generally assume $\mathcal{V} \subseteq \mathbb{F}^n$, where the field $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.
- Note that much of what we cover here also applies to function spaces, which can be useful when developing methods to solve PDEs, ODEs, when processing spatio-temporal data, ...

Linear combinations

- A linear combination of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{F}^n$ is prescribed by some scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ and given by

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k \in \mathbb{F}^n.$$

- The span of $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{F}^n$ is a subspace of $\mathbb{F}^n$ which consists of all the linear combinations of $\mathbf{x}_1, \ldots, \mathbf{x}_n$, i.e.,

$$\text{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} := \{\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k \text{ s.t. } \alpha_1, \ldots, \alpha_k \in \mathbb{F}\}$$

- The information that can be captured by linear combination depends on the **linear independence**, or lack thereof, of the spanning vectors.

Definition (Linear independence)

$\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{F}^n$ are linearly independent if no $\mathbf{x}_i$ can be written as a linear combination of the other vectors. Or, equivalently, if

$$\sum_{i=1}^{k} \alpha_i \mathbf{x}_i = 0 \implies \alpha_1, \ldots, \alpha_k = 0.$$

Bases, dimension and subspaces of vector spaces

- The elements of a vector space can all be represented using a basis.

Definition (Basis & dimension of vector space)

- The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathcal{V}$ form a basis of the vector space $\mathcal{V}$, if they are linearly independent, and span$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \mathcal{V}$.
- While $\mathcal{V}$ admits infinitely many different bases, all of these consist of k linearly independent vectors. We call k the dimension of $\mathcal{V}$, and we write $\dim(\mathcal{V}) = k$.

For a basis $\mathbf{v_1}, \ldots, \mathbf{v}_k \in \mathcal{V}$ of $\mathcal{V}$, we define $\mathbf{V} := [\mathbf{v_1}, \ldots, \mathbf{v}_k]$, so that for each $\mathbf{x} \in \mathcal{V}$, there is a unique $\boldsymbol{\alpha} \in \mathbb{F}^k$ such that $\mathbf{x} = \mathbf{V}\boldsymbol{\alpha} = \sum_{i=1}^k \alpha_i \mathbf{v}_i$.

- In practice, linear subspaces, i.e., lower-dimensional vector spaces within vector spaces, are often used in place of high-dimensional vector spaces.

Definition (Linear subspace)

A linear subspace $\mathcal{S} \subset \mathcal{V}$ of a vector space $\mathcal{V}$ is a non-empty subset which is **closed under addition** and **scalar multiplication**, i.e.,

$$\mathbf{x} + \mathbf{y} \in \mathcal{S} \text{ and } \alpha\mathbf{x} \in \mathcal{S} \,\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{S}, \alpha \in \mathbb{F}.$$

Inner products and norms

Section 2.2 in Darve & Wootters (2021)

Vector inner products

- The abstract definition of length of a vector, angles and orthogonality between vectors, which are important notions in numerical linear algebra, is made possible with the use of inner products.

Definition (Inner product)

An inner product $(\cdot,\cdot)$ on a vector space $\mathcal{V}$ over $\mathbb{F}$ is a mapping $\mathcal{V}\times\mathcal{V}\rightarrow\mathbb{F}$ s.t.

1. $(\cdot,\cdot)$ is linear w.r.t. to its first argument:
$$(\alpha\mathbf{x}_1+\beta\mathbf{x}_2,\mathbf{y})=\alpha(\mathbf{x}_1,\mathbf{y})+\beta(\mathbf{x}_2,\mathbf{y})\ \forall\,\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}\in\mathcal{V},\alpha,\beta\in\mathbb{F}$$
2. $(\cdot,\cdot)$ is Hermitian: $(\mathbf{y},\mathbf{x})=\overline{(\mathbf{x},\mathbf{y})}\ \forall\,\mathbf{x},\mathbf{y}\in\mathcal{V}$
3. $(\cdot,\cdot)$ is positive-definite: $(\mathbf{x},\mathbf{x})\geq 0\ \forall\,\mathbf{x}\in\mathcal{V}$ and $(\mathbf{x},\mathbf{x})=0\iff\mathbf{x}=\mathbf{0}$

$\overline{\alpha}:=\Re\{\alpha\}-i\,\Im\{\alpha\}$ is the complex conjugate of $\alpha\in\mathbb{F}$.

- In particular, **dot products**, which are often used on $\mathcal{V}\subseteq\mathbb{F}^n$, are given by:
  - $(\mathbf{x},\mathbf{y})=\mathbf{x}\cdot\mathbf{y}=\mathbf{x}^T\mathbf{y}=\sum_{i=1}^n x_iy_i\ \forall\,\mathbf{x},\mathbf{y}\in\mathcal{V}\subseteq\mathbb{R}^n$,
  - $(\mathbf{x},\mathbf{y})=\mathbf{x}\cdot\mathbf{y}=\mathbf{x}^H\mathbf{y}=\sum_{i=1}^n \overline{x_i}y_i\ \forall\,\mathbf{x},\mathbf{y}\in\mathcal{V}\subseteq\mathbb{C}^n$.

$\mathbf{x}^H:=\overline{\mathbf{x}}^T$ denotes the conjugate transpose of $\mathbf{x}\in\mathbb{F}^n$.

Properties of inner products

- **By** the **Hermitian property**, every inner product $(\cdot,\cdot)$ on a vector space $\mathcal{V}$ is such that, for all $\mathbf{x}\in\mathcal{V}$, $(\mathbf{x},\mathbf{x})$ is **real**, even if $\mathcal{V}$ is defined over $\mathbb{C}$.

Theorem (Cauchy-Schwarz inequality)

Inner products $(\cdot,\cdot):\mathcal{V}\times\mathcal{V}\to\mathbb{F}$ *are s.t.* $|(\mathbf{x},\mathbf{y})|^2\leq(\mathbf{x},\mathbf{x})(\mathbf{y},\mathbf{y})\;\forall\,\mathbf{x},\mathbf{y}\in\mathcal{V}$.

Proof.

Let $f(\mathbf{x},\mathbf{y},\alpha):=(\mathbf{x}-\alpha\mathbf{y},\mathbf{x}-\alpha\mathbf{y})$ for $\mathbf{x},\mathbf{y}\in\mathcal{V}\setminus\{\mathbf{0}\},\alpha\in\mathbb{F}$.

By linearity of 1st argument and Hermitian property, we have

$$f(\mathbf{x},\mathbf{y},\alpha)=(\mathbf{x},\mathbf{x}-\alpha\mathbf{y})-\alpha(\mathbf{y},\mathbf{x}-\alpha\mathbf{y})=\overline{(\mathbf{x}-\alpha\mathbf{y},\mathbf{x})}-\alpha\overline{(\mathbf{x}-\alpha\mathbf{y},\mathbf{y})}$$

$$=(\mathbf{x},\mathbf{x})-\overline{\alpha}\cdot(\mathbf{x},\mathbf{y})-\alpha\cdot(\mathbf{y},\mathbf{x})+|\alpha|^2\cdot(\mathbf{y},\mathbf{y})$$

By positive-definiteness, we have $f(\mathbf{x},\mathbf{y},\alpha)\geq 0$ so that

$$(\mathbf{x},\mathbf{x})+|\alpha|^2\cdot(\mathbf{y},\mathbf{y})\geq\overline{\alpha}\cdot(\mathbf{x},\mathbf{y})+\alpha\cdot(\mathbf{y},\mathbf{x})$$

Then, if we let $\alpha:=(\mathbf{x},\mathbf{y})/(\mathbf{y},\mathbf{y})$, we get

$$(\mathbf{x},\mathbf{x})+\frac{|(\mathbf{x},\mathbf{y})|^2}{(\mathbf{y},\mathbf{y})}\geq\frac{|(\mathbf{x},\mathbf{y})|^2}{(\mathbf{y},\mathbf{y})}+\frac{|(\mathbf{x},\mathbf{y})|^2}{(\mathbf{y},\mathbf{y})}\;\forall\,\mathbf{x},\mathbf{y}\in\mathcal{V}\setminus\{\mathbf{0}\},\text{ then }\times\text{ by }(\mathbf{y},\mathbf{y}).$$

$\square$

Vector norms

- Vector norms are abstract measures of length in vector spaces.

Definition (Vector norm)

A vector norm $\|\cdot\|$ on a vector space $\mathcal{V}$ over $\mathbb{F}$ is any real-valued function s.t.

1. $\|\cdot\|$ is positive-definite: $\|\mathbf{x}\| \geq 0 \;\forall\, \mathbf{x} \in \mathcal{V}$ and $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$
2. $\|\cdot\|$ is homogeneous: $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\| \;\forall\, \mathbf{x} \in \mathcal{V}, \alpha \in \mathbb{F}$
3. $\|\cdot\|$ satisfies the triangular inequality: $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \;\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{V}$

- Popular vector norms on vector spaces $\mathcal{V} \subseteq \mathbb{F}^n$ are
  - 1-norm: $\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i| \;\forall\, \mathbf{x} \in \mathcal{V}$.
  - 2-norm: $\|\mathbf{x}\|_2 := \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} \;\forall\, \mathbf{x} \in \mathcal{V}$.
    - The 2-norm is induced by the dot product, i.e., $\|\mathbf{x}\|_2 = (\mathbf{x}\cdot\mathbf{x})^{1/2}$.
    - Every inner product $(\cdot,\cdot)$ on $\mathcal{V}$ induces a norm $\|\mathbf{x}\| := (\mathbf{x},\mathbf{x})^{1/2} \;\forall\, \mathbf{x} \in \mathcal{V}$.
  - p-norm: $\|\mathbf{x}\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \;\forall\, \mathbf{x} \in \mathcal{V}$.
  - ∞-norm: $\|\mathbf{x}\|_\infty := \max_{1\leq i\leq n} |x_i| \;\forall\, \mathbf{x} \in \mathcal{V}$.

Equivalence of vector norms

- A sequence of vectors $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \mathcal{V}$ **converges to a vector** $\mathbf{x} \in \mathcal{V}$ **under the norm** $\|\cdot\|$ defined on $\mathcal{V}$ if, for any real value $\epsilon > 0$, there exists K s.t. $\|\mathbf{x}_k - \mathbf{x}\| < \epsilon$ for all $k \geq K$.
- Two vector norms $\|\cdot\|$ and $\|\cdot\|'$ are **equivalent** if convergence under one norm implies convergence under the other.
  The equivalence of vector norms is usually revealed by making use of the following theorem:

Theorem (Equivalent vector norms)

Two vector norms $\|\cdot\|$ and $\|\cdot\|'$ on a vector space $\mathcal{V}$ are equivalent iff there exist real constants $C_1, C_2 > 0$ s.t. $C_1\|\mathbf{x}\| \leq \|\mathbf{x}\|' \leq C_2\|\mathbf{x}\|\ \forall\, \mathbf{x} \in \mathcal{V}$.

- In finite-dimensional vector spaces $\mathcal{V} \subseteq \mathbb{F}^n$, **all norms are equivalent**.
  In particular, we have:
  - $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n}\,\|\mathbf{x}\|_2\ \forall\, \mathbf{x} \in \mathcal{V}$,
  - $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\,\|\mathbf{x}\|_\infty\ \forall\, \mathbf{x} \in \mathcal{V}$,
  - $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq n\,\|\mathbf{x}\|_\infty\ \forall\, \mathbf{x} \in \mathcal{V}$.

Orthogonality and orthonormality

- For any inner product $(\cdot, \cdot)$ defined on a vector space $\mathcal{V}$ over $\mathbb{R}$, the notion of angle between two vectors $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ is introduced through the relation

$$(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|\|\mathbf{y}\| \cos(\angle(\mathbf{x}, \mathbf{y})) \;\forall\, \mathbf{x}, \mathbf{y} \in \mathcal{V}$$

where $\|\cdot\|$ is the induced norm $\|\mathbf{x}\| = (\mathbf{x}, \mathbf{x})^{1/2} \;\forall\, \mathbf{x} \in \mathcal{V}$.

Consequently, non-zero vectors $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ form a right angle, which is $cos(\angle(\mathbf{x}, \mathbf{y})) = 0$, iff $(\mathbf{x}, \mathbf{y}) = 0$. More generally,

Definition (Orthogonality & orthonormality)

- A set of vectors $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathcal{V}$ in a vector space $\mathcal{V}$ equipped with an inner product $(\cdot, \cdot)$ over a scalar field $\mathbb{F}$ is orthogonal if

$$i \neq j \implies (\mathbf{x}_i, \mathbf{x}_j) = 0 \;\text{ for }\; i, j = 1, \dots, k.$$

- If $(\mathbf{x}_i, \mathbf{x}_j) = \delta_{ij}$ for $i, j = 1, \dots, k$, then the vectors $\mathbf{x}_1, \dots, \mathbf{x}_k$ are orthonormal.

Orthogonality and orthonormality, cont'd

- The notion of orthogonal subspaces is useful for the definition and analysis of numerical methods in linear algebra.

Definition (Orthogonal subspace & orthogonal complement)

- Let $\mathcal{S}$ and $\mathcal{T}$ be linear subspaces of the vector space $\mathcal{V}$ equipped with an inner product $(\cdot,\cdot)$. Then, we say that $\mathcal{S}$ is orthogonal to $\mathcal{T}$, i.e., $\mathcal{S} \perp \mathcal{T}$, iff $(\mathbf{x},\mathbf{y}) = 0 \;\forall\, \mathbf{x},\mathbf{y} \in \mathcal{S} \times \mathcal{T}$.
- The orthogonal complement of $\mathcal{S}$, denoted by $\mathcal{S}^\perp$, consists of all the vectors in $\mathcal{V}$ which are orthogonal to $\mathcal{S}$, i.e.,

$$\mathcal{S}^\perp := \{\mathbf{x} \in \mathcal{V} \text{ s.t. } (\mathbf{x},\mathbf{y}) = 0 \;\forall\, \mathbf{y} \in \mathcal{S}\}.$$

Theorem (Orthogonal decomposition of vector spaces)

If $\mathcal{S}$ is a linear subspace of a vector space $\mathcal{V} \subseteq \mathbb{F}^n$, then every $\mathbf{x} \in \mathcal{V}$ admits a **unique decomposition** $\mathbf{x} = \mathbf{y} + \mathbf{z}$ where $\mathbf{y} \in \mathcal{S}$ and $\mathbf{z} \in \mathcal{S}^\perp$:

$$\mathcal{V} = \mathcal{S} \oplus \mathcal{S}^\perp \;\text{ and }\; \dim(\mathcal{V}) = \dim(\mathcal{S}) + \dim(\mathcal{S}^\perp).$$

We say that $\mathcal{V}$ is decomposed by the **direct sum** of $\mathcal{V}$ and $\mathcal{V}^\perp$.

Linear transformations and matrices

Section 2.3 in Darve & Wootters (2021)

From linear transformations between vector spaces to matrices

- Linear transformations (a.k.a. linear maps) are essential operations which are used over and over again in (numerical) linear algebra.

Definition (Linear transformation)

A linear transformation T from a vector space $\mathcal{V}$ to another vector space $\mathcal{W}$, both defined over a field $\mathbb{F}$, i.e., $T : \mathbf{x} \in \mathcal{V} \mapsto T(\mathbf{x}) \in \mathcal{W}$, is such that

$$T(\mathbf{x} + \alpha\mathbf{y}) = T(\mathbf{x}) + \alpha T(\mathbf{y}) \;\forall\; \mathbf{x}, \mathbf{y} \in \mathcal{V}, \alpha \in \mathbb{F}.$$

- Irrespective of whether a linear map operates between function or discrete vector spaces, practical problem-solving often benefits from expressing the action of such maps in a discrete, matrix form.

Proposition (Matrix representation)

The action of a linear transformation $T : \mathcal{V} \rightarrow \mathcal{W}$ between **finite-dimensional** vector spaces defined over a same field $\mathbb{F}$, can be recast into a matrix-vector product with a matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ where $n := \dim(\mathcal{V})$ and $m := \dim(\mathcal{W})$.

Remark: A **function space** can be a finite dimensional vector space.

Review of matrix arithmetic

The components of $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{m \times n}$ and $\mathbf{C} \in \mathbb{F}^{n \times p}$ are denoted by a_{ij}, b_{ij} and c_{ij}.

The vector $\mathbf{x} \in \mathbb{F}^n$ has components x_i and $\alpha \in \mathbb{F}$.

Basic Operations:

- Addition: $(\mathbf{A}+\mathbf{B})_{ij} = a_{ij} + b_{ij}$
- Scalar mult.: $(\alpha \mathbf{A})_{ij} = \alpha a_{ij}$
- Matrix mult.: $(\mathbf{AC})_{ij} = \sum_k a_{ik} c_{kj}$
- Matrix-vector: $(\mathbf{Ax})_i = \sum_j a_{ij} x_j$

Transposition and Conjugation:

- $(\mathbf{A}^T)_{ij} = a_{ji}$, $(\mathbf{A}^H)_{ij} = \overline{a_{ji}}$
- $(\mathbf{A}+\mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H$
- $(\mathbf{AC})^H = \mathbf{C}^H \mathbf{A}^H$
- $(\mathbf{A}^H)^H = \mathbf{A}$

Inverse $(m = n,\ \mathbf{A}, \mathbf{B} \in \mathrm{GL}(n, \mathbb{F}))$:

- $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- $(\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H$

Trace:

- $\mathrm{tr}(\mathbf{A}) = \sum_i a_{ii}$
- $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
- $\mathrm{tr}(\mathbf{AC}) = \mathrm{tr}(\mathbf{CA})$
- $\mathrm{tr}(\mathbf{A}^H) = \overline{\mathrm{tr}(\mathbf{A})}$

Determinant $(m = n)$:

- $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$
- $\det(\mathbf{A}^H) = \overline{\det(\mathbf{A})}$
- $\det(\mathbf{A}^{-1}) = (\det(\mathbf{A}))^{-1} \quad (\mathbf{A} \in \mathrm{GL}(n, \mathbb{F}))$
- $\det(\alpha \mathbf{A}) = \alpha^n \det(\mathbf{A})$

Other Identities:

- $\mathrm{tr}(\mathbf{A}^H\mathbf{A}) = \|\mathbf{A}\|_F^2 = \sum_{i,j} |a_{ij}|^2$

Other Identities $(m = n)$:

- $\mathrm{tr}(\mathbf{A}) = \sum_i \lambda_i$ (λ_i are eigenvalues)
- $\det(\mathbf{A}) = \prod_i \lambda_i$

Fundamental subspaces associated with a matrix

- For every matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$, four linear subspaces of $\mathbb{F}^n$ and $\mathbb{F}^m$ are defined, whose characterization provides insight into the structure of linear systems, revealing key properties like solvability and the geometry of solutions.

Definition (Range of $\mathbf{A}$)

The range (or column space) of a matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is a linear subspace of $\mathbb{F}^m$ given by

$$\operatorname{range}(\mathbf{A}) := \{\mathbf{A}\mathbf{x} \ \text{ s.t. } \mathbf{x} \in \mathbb{F}^n\}.$$

Definition (Null space of $\mathbf{A}$)

The null space (or kernel) of a matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is a linear subspace of $\mathbb{F}^n$ given by

$$\operatorname{null}(\mathbf{A}) := \{\mathbf{x} \in \mathbb{F}^n \ \text{ s.t. } \mathbf{A}\mathbf{x} = \mathbf{0}\}.$$

Definition (Row space of $\mathbf{A}$)

The row space of a matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is a linear subspace of $\mathbb{F}^n$ given by

$$\operatorname{range}(\mathbf{A}^H) := \{\mathbf{A}^H\mathbf{y} \ \text{ s.t. } \mathbf{y} \in \mathbb{F}^m\}.$$

Fundamental subspaces associated with a matrix, cont'd

Definition (Left null space of $\mathbf{A}$)

The left null space of a matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is a linear subspace of $\mathbb{F}^m$ given by

$$\text{null}(\mathbf{A}^H) := \{\mathbf{y} \in \mathbb{F}^m \text{ s.t. } \mathbf{A}^H\mathbf{y} = \mathbf{0}\}.$$

Definition / Theorem (Rank of $\mathbf{A}$)

- The **column rank** of a matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is the dimension of the column space of $\mathbf{A}$, i.e., $\dim(\text{range}(\mathbf{A}))$.
- The **row rank** of a matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is the dimension of the row space of $\mathbf{A}$, i.e., $\dim(\text{range}(\mathbf{A}^H))$.
- The column rank and row ranks of $\mathbf{A} \in \mathbb{F}^{m \times n}$ are always equal. This common value is the **rank** of $\mathbf{A}$, and is denoted by $\text{rank}(\mathbf{A})$, with $\text{rank}(\mathbf{A}) \leq \min(m, n)$.

Definition (Full rank)

- A matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is **full-column-rank** if $\dim(\text{range}(\mathbf{A})) = n$.
- A matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is **full-row-rank** if $\dim(\text{range}(\mathbf{A}^H)) = m$.
- A matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ is **full-rank** if $\text{rank}(\mathbf{A}) = \min(m, n)$.

Equivalence of column and row ranks

- The equivalence between the row rank and the column rank of a matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ can be shown as follows.

Proof.

- Let $\mathbf{A}$ have row rank r, and $\mathbf{r}_1 \dots, \mathbf{r}_r \in \mathbb{F}^n$ be a basis of $\mathrm{range}(\mathbf{A}^H)$.
- Then, the vectors $\mathbf{A}\mathbf{r}_1, \dots, \mathbf{A}\mathbf{r}_r$ are linearly independent:
  - Consider $\alpha_1, \dots, \alpha_r \in \mathbb{F}$ such that $\alpha_1 \mathbf{A}\mathbf{r}_1 + \dots + \alpha_r \mathbf{A}\mathbf{r}_r = \mathbf{0}$.
  - Then, $\mathbf{x} := \alpha_1 \mathbf{r}_1 + \dots + \alpha_r \mathbf{r}_r \in \mathrm{range}(\mathbf{A}^H)$ is such that $\mathbf{A}\mathbf{x} = \mathbf{0}$.
  - For all $\mathbf{y} \in \mathbb{F}^m$, we then have $\mathbf{y}^H \mathbf{A}\mathbf{x} = \mathbf{x}^H \mathbf{A}^H \mathbf{y} = \mathbf{0}$, which implies $\mathbf{x} \perp \mathrm{range}(\mathbf{A}^H)$.
  - $\mathbf{x} \in \mathrm{range}(\mathbf{A}^H)$ and $\mathbf{x} \perp \mathrm{range}(\mathbf{A}^H)$ imply $\mathbf{x} = \mathbf{0}$. $\square$
- Then, since $\mathbf{A}\mathbf{r}_i \in \mathrm{range}(\mathbf{A})$ for $i = 1, \dots, r$, we have $c := \dim(\mathrm{range}(\mathbf{A})) \geq r$.
- Let $\mathbf{c}_1 \dots, \mathbf{c}_c \in \mathbb{F}^m$ be a basis of $\mathrm{range}(\mathbf{A})$.
- Similarly, we can show that the vectors $\mathbf{A}^H \mathbf{c}_1, \dots, \mathbf{A}^H \mathbf{c}_c$ are linearly independent.
- Since $\mathbf{A}^H \mathbf{c}_i \in \mathrm{range}(\mathbf{A}^H)$ for $i = 1, \dots, c$, we have $r = \dim(\mathrm{range}(\mathbf{A}^H)) \geq c$.
- From $c \geq r$ and $r \geq c$, we have $r = c$.

$\square$

Characterization of linear systems

- A key goal of numerical linear algebra is to develop and analyze methods to find an unknown vector $\mathbf{x} \in \mathbb{F}^n$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{F}^{m \times n}$ is a matrix and $\mathbf{b} \in \mathbb{F}^m$ is a given right-hand side.
- Each such matrix equation corresponds to a system of m linear equations with n scalar unknowns $x_1, \ldots, x_n \in \mathbb{F}$, of the form

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

- A linear system is either:
  - **over-determined**: the number m of equations is larger than the number n of unknowns, i.e., $m > n$. $\mathbf{A}$ is a "tall"/"skinny" matrix.
  - **under-determined**: the number m of equations is smaller than the number n of unknowns, i.e., $m < n$. $\mathbf{A}$ is a "short"/"fat" matrix.
  - **square**: the number of equations m is equal to the number n of unknowns, i.e., $m = n$. $\mathbf{A}$ is a square matrix.

Characterization of linear systems, cont'd

- A linear system is characterized by either of 3 situations:
  - The system admits **no solution**.
  - The system has a **unique solution**.
  - The system admits **infinitely many solutions**.
- A proper characterization can be made using two notions:
  1. **Consistency**: A linear system is consistent if $\mathbf{b} \in \mathrm{range}(\mathbf{A})$, which means there exists at least one vector $\mathbf{x} \in \mathbb{F}^n$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$.

     An equivalent statement of consistency is $\mathrm{rank}([\mathbf{A}, \mathbf{b}]) = \mathrm{rank}(\mathbf{A})$, which implies that $\mathbf{b}$ can be formed by linear combination of the columns of $\mathbf{A}$.
  2. **Full column rank**: A **consistent** linear system has a **unique solution** iff the column rank of $\mathbf{A}$ equals the number of unknowns, i.e., $\dim(\mathrm{range}(\mathbf{A})) = n$.

     On the other hand, if $\dim(\mathrm{range}(\mathbf{A})) < n$, and the linear system is consistent, then there exist infinitely many vectors $\mathbf{x} \in \mathbb{F}^m$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$.
- In conclusion:
  - A linear system has **solution(s)** iff it is **consistent**.
  - A linear system has a **unique solution** iff it is **consistent** and it has **full column rank**.
  - An **under-determined** linear system **cannot have** a **unique solution**.

Invertible and singular matrices

- The **identity matrix**, which we denote by $\mathbf{I}_n \in \mathbb{R}^{n\times n}$, is the matrix with ones on the diagonal, and zeros everywhere else.
- A **square** matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ is **invertible** if there exists a matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$ such that

$$\mathbf{BA} = \mathbf{AB} = \mathbf{I}_n.$$

  If such a matrix exists, it is unique, denoted by $\mathbf{A}^{-1}$, and referred to as the inverse of $\mathbf{A}$.
- A matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ has an inverse $\mathbf{A}^{-1} \in \mathbb{F}^{n\times n}$ iff
  - $\mathbf{Ax} = \mathbf{b}$ has a unique solution $\mathbf{x} \in \mathbb{F}^n$ for every $\mathbf{b} \in \mathbb{F}^n$.
  - $\mathbf{A}$ is full-rank, i.e., $\mathrm{rank}(\mathbf{A}) = n$.
  - The columns of $\mathbf{A}$ form a basis of $\mathbb{F}^n$, i.e., $\mathrm{range}(\mathbf{A}) = \mathbb{F}^n$.
  - The null space of $\mathbf{A}$ is trivial, i.e., $\mathrm{null}(\mathbf{A}) = \{\mathbf{0}\}$.
  - 0 is neither an eigenvalue nor a singular value of $\mathbf{A}$.
  - $\det(\mathbf{A}) \neq 0$.
- The set of invertible matrices is the general linear group of degree n over $\mathbb{F}$, denoted by $\mathrm{GL}(n, \mathbb{F})$.

Moore-Penrose inverse

- The **pseudo-inverse** (or **Moore-Penrose inverse**) of a matrix generalizes the concept of matrix inverse to rectangular and singular matrices.

Definition (Moore-Penrose inverse)

For a given matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$, the pseudo-inverse $\mathbf{A}^\dagger \in \mathbb{F}^{n \times m}$ is **unique** and such that

- $\mathbf{A}^\dagger$ is consistent with $\mathbf{A}$: $\quad \mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}$
- $\mathbf{A}^\dagger$ is consistent with $\mathbf{A}^\dagger$: $\quad \mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger$
- $\mathbf{A}^\dagger\mathbf{A}$ is Hermitian: $\quad (\mathbf{A}\mathbf{A}^\dagger)^H = \mathbf{A}\mathbf{A}^\dagger$
- $\mathbf{A}\mathbf{A}^\dagger$ is Hermitian: $\quad (\mathbf{A}^\dagger\mathbf{A})^H = \mathbf{A}^\dagger\mathbf{A}$

- If a is invertible, then $\mathbf{A}^\dagger = \mathbf{A}^{-1}$.
- Pseudo-inverses are used to provide **alternative solutions** to **linear systems which admit no standard solution**, or **infinitely many solutions**.

Different methods for different types of linear systems

Methods for solving linear systems depend on the system's characteristics:

- Systems with **square matrices**:
  - **Full-rank** (**invertible**) matrices:
    - Direct methods: Gaussian elimination, LU decomposition.
    - Iterative methods: Fixed-point methods, Krylov subspace methods.
  - **Singular** matrices but **consistent** systems:
    - Low-rank approximation (SVD, pseudoinverse).
    - Specialized iterative methods that exploit the structure of fundamental subspaces.
- **Over-determined** systems (**tall-and-skinny matrices**):
  - **Full column-rank** matrices:
    - **Consistent** systems: QR decomposition or transformation to normal equations.
    - **Inconsistent** systems: Least squares problems.
  - **Rank-deficient** matrices: Least squares problems with regularization.
- **Under-determined** systems (**short-and-fat matrices**):
  - **Consistent** systems: Low-rank approximation (SVD, pseudoinverse).
  - **Inconsistent** systems: Least squares problems with regularization.
- Least squares problems:
  - direct methods (through QR factorization), or iterative methods (LSQR, LSMR).

Types of matrices

- Normal matrices, i.e., $\mathbf{A} \in \mathbb{F}^{n\times n}$ s.t. $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$:
  - Diagonal matrices, i.e., $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_n)$.
  - Hermitian matrices, i.e., $\mathbf{A} \in \mathbb{F}^{n\times n}$ s.t. $\mathbf{A}^H = \mathbf{A}$.
  - Symmetric matrices, i.e., $\mathbf{A} \in \mathbb{R}^{n\times n}$ s.t. $\mathbf{A}^T = \mathbf{A}$.
  - Unitary matrices, i.e., $\mathbf{U} \in \mathbb{F}^{n\times n}$ s.t. $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}_n$, i.e., $\mathbf{U}^{-1} = \mathbf{U}^H$.
  - Skew-Hermitian matrices, i.e., $\mathbf{A} \in \mathbb{F}^{n\times n}$ s.t. $\mathbf{A}^H = -\mathbf{A}$.
  - Skew-symmetric matrices, i.e., $\mathbf{A} \in \mathbb{R}^{n\times n}$ s.t. $\mathbf{A}^T = -\mathbf{A}$.
- Orthogonal matrices: $\mathbf{Q} \in \mathbb{R}^{m\times n}$, $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$.
- Tridiagonal matrices: 
- Triangular matrices:
  - Lower-triangular matrices: 
  - Upper-triangular matrices: 
- Hessenberg matrices:
  - Lower Hessenberg matrices: 
  - Upper Hessenberg matrices: 
- Block diagonal matrices, i.e., $\mathbf{A} \in \mathbb{F}^{n\times n}$ s.t. $\mathbf{A} = \mathrm{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_{n_b})$.
- ...

Projections in $\mathbb{F}^n$

- Projections, and their abstract representation, play an important role in formulating and understanding methods in NLA.

Definition (Projection in $\mathbb{F}^n$ & projector)

- A projection in $\mathbb{F}^n$ is an **idempotent linear map** $\mathbf{x} \in \mathbb{F}^n \mapsto \mathbf{P}\mathbf{x} \in \mathbb{F}^n$, i.e., such that $\mathbf{P}^2\mathbf{x} = \mathbf{P}\mathbf{x} \;\forall\, \mathbf{x} \in \mathbb{F}^n$. The matrix $\mathbf{P} \in \mathbb{F}^{n\times n}$ is called a **projector**.
- The range of a non-trivial projector is a proper subset of $\mathbb{F}^n$, i.e., $\mathrm{rank}(\mathbf{P}) < n$.

Proposition (Complementary projector)

If $\mathbf{P} \in \mathbb{F}^{n\times n}$ is a projector, then $\mathbf{I}_n - \mathbf{P}$ is a projector onto $\mathrm{null}(\mathbf{P})$.

Proof.

- $\mathbf{I}_n - \mathbf{P}$ is idempotent, i.e., $(\mathbf{I}_n - \mathbf{P})^2 = \mathbf{I}_n - \mathbf{P} - \mathbf{P} + \mathbf{P}^2 = \mathbf{I}_n - 2\mathbf{P} + \mathbf{P} = \mathbf{I}_n - \mathbf{P}$.
- $\mathrm{range}(\mathbf{I}_n - \mathbf{P}) = \mathrm{null}(\mathbf{P})$:
  - $\mathbf{x} = \mathbf{P}\mathbf{x} + (\mathbf{I}_n - \mathbf{P})\mathbf{x} \;\forall\, \mathbf{x} \in \mathbb{F}^n$, so that $\mathbf{P}\mathbf{x} = \mathbf{0} \implies \mathbf{x} = (\mathbf{I}_n - \mathbf{P})\mathbf{x}$
    $\implies \mathrm{null}(\mathbf{P}) \subseteq \mathrm{range}(\mathbf{I}_n - \mathbf{P})$.
  - $\mathbf{P}(\mathbf{I}_n - \mathbf{P})\mathbf{x} = (\mathbf{P} - \mathbf{P}^2)\mathbf{x} = (\mathbf{P} - \mathbf{P})\mathbf{x} = \mathbf{0} \;\forall\, \mathbf{x} \in \mathbb{F}^n \implies \mathrm{range}(\mathbf{I}_n - \mathbf{P}) \subseteq \mathrm{null}(\mathbf{P})$.

$\square$

Projections in $\mathbb{F}^n$, cont'd

Theorem (Decomposition of $\mathbb{F}^n$ via projection)

Given a projector $\mathbf{P} \in \mathbb{F}^{n\times n}$, every $\mathbf{x} \in \mathbb{F}^n$ is uniquely decomposed into a sum $\mathbf{x} = \mathbf{P}\mathbf{x} + (\mathbf{I}_n - \mathbf{P})\mathbf{x}$ with $\mathbf{P}\mathbf{x} \in \mathrm{range}(\mathbf{P})$ and $(\mathbf{I}_n - \mathbf{P})\mathbf{x} \in \mathrm{null}(\mathbf{P})$:

$$\begin{aligned}\mathbb{F}^n &= \mathrm{range}(\mathbf{P}) \oplus \mathrm{null}(\mathbf{P}) \\ n &= \mathrm{rank}(\mathbf{P}) + \mathrm{rank}(\mathbf{I}_n - \mathbf{P}).\end{aligned}$$

Proof.

- $\mathrm{range}(\mathbf{P}) \cap \mathrm{null}(\mathbf{P}) = \{\mathbf{0}\}$.
  Let $\mathbf{y} \in \mathrm{range}(\mathbf{P}) \cap \mathrm{null}(\mathbf{P})$, then
  - Since $\mathbf{P}$ is idempotent and $\mathbf{y} \in \mathrm{range}(\mathbf{P})$, we have $\mathbf{P}\mathbf{y} = \mathbf{y}$.
  - From $\mathbf{y} \in \mathrm{null}(\mathbf{P})$, we have $\mathbf{P}\mathbf{y} = \mathbf{0}$.

  $\implies \mathbf{y} = \mathbf{0}$. $\square$
- Since $\mathbf{x} = \mathbf{P}\mathbf{x} + (\mathbf{I}_n - \mathbf{P})\mathbf{x} \ \forall\, \mathbf{x} \in \mathbb{F}^n$, in which $\mathbf{x} \mapsto \mathbf{P}\mathbf{x}$ is onto $\mathrm{range}(\mathbf{P})$, while $\mathbf{x} \mapsto (\mathbf{I}_n - \mathbf{P})\mathbf{x}$ is onto $\mathrm{null}(\mathbf{P})$, and $\mathrm{range}(\mathbf{P}) \cap \mathrm{null}(\mathbf{P}) = \{\mathbf{0}\}$, we have that the decomposition of $\mathbf{x}$ via projection is unique.

$\square$

Abstract and geometric formulation of projections in $\mathbb{F}^n$

- For every pair $(\mathcal{M}, \mathcal{S})$ of linear subspaces of $\mathbb{F}^n$ such that $\mathbb{F}^n = \mathcal{M} \oplus \mathcal{S}$, there exists a unique projector $\mathbf{P} \in \mathbb{F}^{n \times n}$ such that

$$\mathcal{M} = \mathrm{range}(\mathbf{P}) \text{ and } \mathcal{S} = \mathrm{null}(\mathbf{P}).$$

- In general, $\mathbf{y} \in \mathcal{S} = \mathrm{null}(\mathbf{P}) \not\Leftrightarrow \mathbf{y} \perp \mathcal{M} = \mathrm{range}(\mathbf{P})$.
- For every such pair $(\mathcal{M}, \mathcal{S})$, the linear map $\mathbf{x} \mapsto \mathbf{u} := \mathbf{Px}$ is recast into:

$$\begin{aligned} \text{Find} \quad & \mathbf{u} \in \mathcal{M} \\ \text{s.t.} \quad & \mathbf{x} - \mathbf{u} \in \mathcal{S}. \end{aligned}$$

Since $\mathbf{y} \in \mathcal{S} \iff \mathbf{y} \perp \mathcal{L}$ where $\mathcal{L} := \mathcal{S}^\perp$, the projection is recast into

$$\boxed{\begin{aligned} \text{Find} \quad & \mathbf{u} \in \mathcal{M} & \text{where} \quad & \mathcal{M} = \mathrm{range}(\mathbf{P}) \\ \text{s.t.} \quad & \mathbf{x} - \mathbf{u} \perp \mathcal{L} & & \mathcal{L} = \mathrm{null}(\mathbf{P})^\perp \end{aligned}} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

where $\mathcal{L}$ is the **orthogonality subspace** (or **space of constraints**). We say that Eqs. (1)-(2) define a projector $\mathbf{P}$ **onto** $\mathcal{M}$ **perpendicular to** $\mathcal{L}$ (or **along** $\mathcal{L}$). An advantage of this reformulation is that $\dim(\mathcal{L}) = \dim(\mathcal{M})$.

Orthogonal and oblique projections in $\mathbb{F}^n$

- Every projection induced by the decomposition of $\mathbb{F}^n$ into complementary subspaces $\mathcal{M}, \mathcal{L}^\perp \subset \mathbb{F}^n$ is either orthogonal, or oblique.

Definition (Orthogonal projections & oblique projections)

- The projection in $\mathbb{F}^n$ induced by the linear subspaces $\mathcal{M}, \mathcal{L} \subset \mathbb{F}^n$ such that $\mathbb{F}^n = \mathcal{M} \oplus \mathcal{L}^\perp$, is **orthogonal** if the space of constraints $\mathcal{L}$, is the approximation space $\mathcal{M}$, i.e., $\mathcal{L} = \mathcal{M}$. Then, for every $\mathbf{x} \in \mathbb{F}^n$, the projection $\mathbf{x} \mapsto \mathbf{u} := \mathbf{P}\mathbf{x}$ is the solution of

$$\begin{aligned} \text{Find} \quad & \mathbf{u} \in \mathcal{M} \\ \text{s.t.} \quad & \mathbf{x} - \mathbf{u} \perp \mathcal{M}. \end{aligned}$$

- A projection in $\mathbb{F}^n$ that is not orthogonal, is **oblique**. Then, for every $\mathbf{x} \in \mathbb{F}^n$, the projection $\mathbf{x} \mapsto \mathbf{u} := \mathbf{P}\mathbf{x}$ is the solution of

$$\begin{aligned} \text{Find} \quad & \mathbf{u} \in \mathcal{M} \\ \text{s.t.} \quad & \mathbf{x} - \mathbf{u} \perp \mathcal{L} \end{aligned}$$

where the space of constraints $\mathcal{L}$ is different from the approximation space $\mathcal{M}$, i.e., $\mathcal{L} \neq \mathcal{M}$.

Optimality of orthogonal projections in $\mathbb{F}^n$

- Orthogonal projections have important approximation properties.

Theorem (Optimality of orthogonal projections in $\mathbb{F}^n$)

Let $\mathbf{x} \mapsto \mathbf{P}\mathbf{x}$ be the orthogonal projection induced by the linear subspace $\mathcal{M} \subset \mathbb{F}^n$. Then, for every $\mathbf{x} \in \mathbb{F}^n$, $\mathbf{P}\mathbf{x}$ is the best approximation of $\mathbf{x}$ in $\mathcal{M}$:

$$\|\mathbf{x} - \mathbf{P}\mathbf{x}\| = \min_{\mathbf{u} \in \mathcal{M}} \|\mathbf{x} - \mathbf{u}\| \; \forall \; \mathbf{x} \in \mathbb{F}^n.$$

Proof.

- Every orthogonal projection $\mathbf{x} \mapsto \mathbf{P}\mathbf{x}$ is such that $\mathrm{range}(\mathbf{P})^\perp = \mathrm{null}(\mathbf{P})$.
- Then, for all $(\mathbf{x}, \mathbf{u}) \in \mathbb{F}^n \times \mathcal{M}$, we have:

$$\|\mathbf{x} - \mathbf{u}\| = \|(\mathbf{I}_n - \mathbf{P})\mathbf{x} + \mathbf{P}\mathbf{x} - \mathbf{u}\| = \|(\mathbf{I}_n - \mathbf{P})\mathbf{x}\| + \|\mathbf{P}\mathbf{x} - \mathbf{u}\|$$

  where we use the fact that $\mathrm{range}(\mathbf{I}_n - \mathbf{P}) \perp \mathcal{M}$ and $\mathbf{u} \in \mathcal{M} = \mathrm{range}(\mathbf{P})$, and we assume $\|\cdot\|$ is induced by the inner product from the definition of orthogonality.
- Then, we have $\|\mathbf{x} - \mathbf{u}\| \geq \|\mathbf{x} - \mathbf{P}\mathbf{x}\| \; \forall \; \mathbf{u} \in \mathcal{M}$.
- Moreover, $\|\mathbf{x} - \mathbf{u}\|$ is minimized when $\mathbf{u} = \mathbf{P}\mathbf{x}$.

$\square$

Matrix form of orthogonal projections in $\mathbb{F}^n$

- Let the linear subspace $\mathcal{M} \subset \mathbb{F}^n$ of an orthogonal projection $\mathbf{x} \mapsto \mathbf{P}\mathbf{x}$ have dimension m, i.e., $\dim(\mathcal{M}) = m < n$.
- Then, there exists a basis $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathcal{M}$ such that, for every $\mathbf{u} \in \mathcal{M}$, there is a unique $\hat{\mathbf{u}} \in \mathbb{F}^m$ such that $\mathbf{u} = \mathbf{V}\hat{\mathbf{u}}$, where $\mathbf{V} := [\mathbf{v}_1, \dots, \mathbf{v}_m]$.
- Then, from the geometric definition of an orthogonal projection onto $\mathcal{M}$ given by $\mathbf{x} \mapsto \mathbf{u} := \mathbf{P}\mathbf{x}$, for every $\mathbf{x} \in \mathbb{F}^n$, there exists $\hat{\mathbf{u}} \in \mathbb{F}^m$ such that

$$\mathbf{u} \in \mathcal{M},\ \mathbf{x} - \mathbf{u} \perp \mathcal{M} \iff \mathbf{u} = \mathbf{V}\hat{\mathbf{u}},\ (\mathbf{v}_i, \mathbf{x} - \mathbf{V}\hat{\mathbf{u}}) = 0 \text{ for } i = 1, \dots, m.$$

- Let the inner product $(\cdot,\cdot)$ be a dot product, in which case we have

$$\mathbf{V}^H(\mathbf{x} - \mathbf{V}\hat{\mathbf{u}}) = \mathbf{0}$$
$$\mathbf{V}^H\mathbf{x} = \mathbf{V}^H\mathbf{V}\hat{\mathbf{u}}$$

so that $\hat{\mathbf{u}} = (\mathbf{V}^H\mathbf{V})^{-1}\mathbf{V}^H\mathbf{x}$, and $\mathbf{u} = \mathbf{P}\mathbf{x}$, where

$$\boxed{\mathbf{P} = \mathbf{V}(\mathbf{V}^H\mathbf{V})^{-1}\mathbf{V}^H}.$$

- Note that the basis $\mathbf{v}_1, \dots, \mathbf{v}_m$ does not need to be orthogonal for $\mathbf{x} \mapsto \mathbf{P}\mathbf{x}$ to be an orthogonal projection. But if it is, then $\mathbf{P} = \mathbf{V}\mathbf{V}^H$.

Matrix form of oblique projections in $\mathbb{F}^n$

- Let the linear subspaces $\mathcal{M}, \mathcal{L} \subset \mathbb{F}^n$ of an oblique projection $\mathbf{x} \mapsto \mathbf{P}\mathbf{x}$ have dimension m, i.e., $\dim(\mathcal{M}) = \dim(\mathcal{L}) = m < n$ and $\mathcal{M} \oplus \mathcal{L}^\perp = \mathbb{F}^n$.
- Then, there exist two bases $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathcal{M}$ and $\mathbf{w}_1, \ldots, \mathbf{w}_m \in \mathcal{L}$. The first one is such that, for every $\mathbf{u} \in \mathcal{M}$, there is a unique $\hat{\mathbf{u}} \in \mathbb{F}^m$ such that $\mathbf{u} = \mathbf{V}\hat{\mathbf{u}}$, where $\mathbf{V} := [\mathbf{v}_1, \ldots, \mathbf{v}_m]$.
- From the definition of an oblique projection onto $\mathcal{M}$ perpendicular to $\mathcal{L}$ given by $\mathbf{x} \mapsto \mathbf{u} := \mathbf{P}\mathbf{x}$, for every $\mathbf{x} \in \mathbb{F}^n$, there exists $\hat{\mathbf{u}} \in \mathbb{F}^m$ such that

$$\begin{aligned}
\mathbf{u} \in \mathcal{M},\ \mathbf{x} - \mathbf{u} \perp \mathcal{L} \iff \mathbf{u} = \mathbf{V}\hat{\mathbf{u}},\ (\mathbf{w}_i, \mathbf{x} - \mathbf{V}\hat{\mathbf{u}}) = 0 \ \text{ for } \ i = 1, \ldots, m \\
\mathbf{W}^H(\mathbf{x} - \mathbf{V}\hat{\mathbf{u}}) = \mathbf{0} \\
\mathbf{W}^H\mathbf{x} = \mathbf{W}^H\mathbf{V}\hat{\mathbf{u}}.
\end{aligned}$$

where the inner product is a dot product, so that $\hat{\mathbf{u}} = (\mathbf{W}^H\mathbf{V})^{-1}\mathbf{W}^H\mathbf{x}$, and $\mathbf{u} = \mathbf{P}\mathbf{x}$, where

$$\boxed{\mathbf{P} = \mathbf{V}(\mathbf{W}^H\mathbf{V})^{-1}\mathbf{W}^H}$$

where $\mathbf{W}^H\mathbf{V}$ is not singular because, by definition, we have $\mathcal{M} \cap \mathcal{L}^\perp = \{\mathbf{0}\}$.

Matrix norms

- Matrix norms are abstract measures of the strength of a transformation.

Definition (Matrix norm)

A matrix norm is a function $\mathbb{F}^{m\times n} \rightarrow \mathbb{R}$ such that

1. $\|\cdot\|$ is positive-definite: $\|\mathbf{A}\| \geq 0$ and $\|\mathbf{A}\| = 0 \iff \mathbf{A} = \mathbf{0}$
2. $\|\cdot\|$ is homogeneous: $\|\alpha\mathbf{A}\| = |\alpha|\|\mathbf{A}\| \; \forall \; \mathbf{A} \in \mathbb{F}^{m\times n}, \alpha \in \mathbb{F}$
3. $\|\cdot\|$ satisfies the triangular inequality: $\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\| \; \forall \; \mathbf{A}, \mathbf{B} \in \mathbb{F}^{m\times n}$

- Some matrix norms are naturally induced by vector norms.

Definition (Induced norms)

Let $\|\cdot\|_\beta : \mathbb{F}^m \rightarrow \mathbb{R}$ and $\|\cdot\|_\alpha : \mathbb{F}^n \rightarrow \mathbb{R}$ be vector norms. The **induced norm** (or subordinate norm, or operator norm) $\|\cdot\| : \mathbb{F}^{m\times n} \rightarrow \mathbb{R}$ is the matrix norm defined as

$$\|\mathbf{A}\| = \sup_{\mathbf{x}\in\mathbb{F}^n\setminus\{\mathbf{0}\}} \frac{\|\mathbf{A}\mathbf{x}\|_\beta}{\|\mathbf{x}\|_\alpha} \; \forall \; \mathbf{A} \in \mathbb{F}^{m\times n}.$$

Properties of matrix norms

- A number of properties of matrix norms can come in handy, in particular

Definition (Consistency of matrix norms)

- The matrix norms $\|\cdot\|_\alpha : \mathbb{F}^{m\times n} \to \mathbb{R}$, $\|\cdot\|_\beta : \mathbb{F}^{n\times p} \to \mathbb{R}$ and $\|\cdot\|_\gamma : \mathbb{F}^{m\times p} \to \mathbb{R}$ are mutually **consistent** if

$$\|\mathbf{AB}\|_\gamma \leq \|\mathbf{A}\|_\alpha \|\mathbf{B}\|_\beta \;\forall\; \mathbf{A} \in \mathbb{F}^{m\times n}, \mathbf{B} \in \mathbb{F}^{n\times p}.$$

- In case $m = n = p$ and $\gamma = \beta = \alpha$, we say that the matrix norm $\|\cdot\|_\alpha$ is **sub-multiplicative**.

- All **induced norms are consistent**, i.e., mutually consistent with themselves.
- The **Frobenius norm**, which is **not induced by any vector norm**, and is defined as

$$\|\mathbf{A}\|_F := \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}$$

is consistent.

Properties of matrix norms, cont'd

- A number of properties of matrix norms can come in handy, in particular

Definition (Consistency of matrix and vector norms)

- A matrix norm $\|\cdot\| : \mathbb{F}^{m\times n} \rightarrow \mathbb{R}$ is **consistent with** the vector norms $\|\cdot\|_\beta : \mathbb{F}^m \rightarrow \mathbb{R}$ and $\|\cdot\|_\alpha : \mathbb{F}^n \rightarrow \mathbb{R}$ if

$$\|\mathbf{A}\mathbf{x}\|_\beta \leq \|\mathbf{A}\|\|\mathbf{x}\|_\alpha \;\forall\; \mathbf{A} \in \mathbb{F}^{m\times n}, \mathbf{x} \in \mathbb{F}^n.$$

- In case $m = n$ and $\beta = \alpha$, we say the matrix norm $\|\cdot\|$ is **compatible with** $\|\cdot\|_\alpha$.
- All **induced norms** are **consistent** with their underlying vector norms **by definition**.

- The **Frobenius norm** is **consistent with vector 2-norms**.
- For every matrix norm $\|\cdot\| : \mathbb{F}^{m\times n} \rightarrow \mathbb{R}$ induced by the vector norms $\|\cdot\|_\beta : \mathbb{F}^m \rightarrow \mathbb{R}$ and $\|\cdot\|_\alpha : \mathbb{F}^m \rightarrow \mathbb{R}$, and every matrix $\mathbf{A} \in \mathbb{F}^{m\times n}$, there exists $\mathbf{x} \in \mathbb{F}^n$ s.t. $\|\mathbf{A}\mathbf{x}\|_\beta = \|\mathbf{A}\|\|\mathbf{x}\|_\alpha$.

Equivalence of matrix norms

- A sequence of matrices $\{\mathbf{A}_k\}_{k\in\mathbb{N}} \subset \mathbb{F}^{m\times n}$ **converges to a matrix** $\mathbf{A} \in \mathbb{F}^{m\times n}$ **under the norm** $\|\cdot\|$ defined on $\mathbb{F}^{m\times n}$ if, for any real $\epsilon > 0$, there exists K s.t. $\|\mathbf{A}_k - \mathbf{A}\| < \epsilon$ for all $k > K$.
- Two matrix norms $\|\cdot\|$ and $\|\cdot\|'$ are **equivalent** if convergence under one norm implies convergence under the other. This can be checked by

Theorem (Equivalent matrix norms)

Two matrix norms $\|\cdot\|$ and $\|\cdot\|'$ on $\mathbb{F}^{m\times n}$ are equivalent iff there exist real constants $C_1, C_2 > 0$ s.t. $C_1\|\mathbf{A}\| \leq \|\mathbf{A}\|' \leq C_2\|\mathbf{A}\| \;\forall\; \mathbf{A} \in \mathbb{F}^{m\times n}$.

- **All matrix norms** defined **on** $\mathbb{F}^{m\times n}$ **are equivalent**. In particular,
  - $\frac{1}{\sqrt{n}}\|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_2 \leq \sqrt{m}\|\mathbf{A}\|_\infty \;\forall\; \mathbf{A} \in \mathbb{F}^{m\times n}$,
  - $\frac{1}{\sqrt{m}}\|\mathbf{A}\|_1 \leq \|\mathbf{A}\|_2 \leq \sqrt{n}\|\mathbf{A}\|_1 \;\forall\; \mathbf{A} \in \mathbb{F}^{m\times n}$,
  - $\frac{1}{\sqrt{\min(m,n)}}\|\mathbf{A}\|_F \leq \|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F \;\forall\; \mathbf{A} \in \mathbb{F}^{m\times n}$,
  - $\max_{i,j}|a_{ij}| \leq \|\mathbf{A}\|_2 \leq \sqrt{mn}\max_{i,j}|a_{ij}| \;\forall\; \mathbf{A} \in \mathbb{F}^{m\times n}$.

Eigenvalues and eigenvectors

Section 2.4 in Darve & Wootters (2021)

Eigenvalue and eigenvectors

- **Eigenvalues and eigenvectors** reveal how a square matrix transforms space by identifying invariant directions (along eigenvectors) and their scaling factors (eigenvalues).

Definition (Eigenvalues, eigenvectors & spectrum)

- A scalar $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathbf{A} \in \mathbb{F}^{n \times n}$ if there is a non-zero vector $\mathbf{u} \in \mathbb{C}^n$ such that $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$.
- $\mathbf{u}$ is called an **eigenvector** of $\mathbf{A}$.
- The set of all the eigenvalues of $\mathbf{A}$, denoted by $\mathrm{Sp}(\mathbf{A})$, is the **spectrum** of $\mathbf{A}$.

- Equivalently, $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathbf{A} \in \mathbb{F}^{n \times n}$ if it is a root of the characteristic polynomial:

$$p_{\mathbf{A}} : \lambda \in \mathbb{C} \mapsto \det(\mathbf{A} - \lambda \mathbf{I}_n) \in \mathbb{F}.$$

Definition (Left eigenpair)

- A scalar $\vartheta \in \mathbb{C}$ and a non-zero vector $\mathbf{v} \in \mathbb{C}^n$ such that $\mathbf{v}^H\mathbf{A} = \vartheta\mathbf{v}^H$ is called a **left eigenpair** of $\mathbf{A}$.

Multiplicity of eigenvalues

- Eigenvalues have two types of multiplicities: **algebraic** and **geometric**.
- **Algebraic multiplicity**: The number of times an eigenvalue appears as a root of the characteristic polynomial.
- **Geometric multiplicity**: The dimension of the eigenspace corresponding to the eigenvalue, i.e., the number of linearly independent eigenvectors associated with that eigenvalue.

$$1 \leq \text{ geometric multiplicity } \leq \text{ algebraic multiplicty}$$

- An eigenvalue is **semisimple** if its **geometric mult**. = its **algebraic mult**..
- An eigenvalue with **geometric mult**. $<$ **algebraic mult**. has fewer independent eigenvectors than the size of the eigenspace implied by the algebraic multiplicity.
- If a matrix has at least one eigenvalue with **geometric mult**. $<$ **algebraic mult**., then it is **defective**, i.e., it does not have a full set of linearly independent eigenvectors.
- $$\textbf{defective} \;\not\Leftrightarrow\; \textbf{singular}$$

Characterization of normal matrices

- A square matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ is called **normal** if it satisfies:

$$\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}.$$

- Common examples of normal matrices are
  - Diagonal matrices, i.e., $\mathbf{A} = \mathrm{diag}(a_{11}, \dots, a_{nn})$.
  - Hermitian matrices, i.e., $\mathbf{A} \in \mathbb{F}^{n \times n}$ such that $\mathbf{A}^H = \mathbf{A}$.
  - Symmetric matrices, i.e., $\mathbf{A} \in \mathbb{R}^{n \times n}$ such that $\mathbf{A}^T = \mathbf{A}$.
  - Unitary matrices, i.e., $\mathbf{A} \in \mathbb{F}^{n \times n}$ such that $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A} = \mathbf{I}_n$.
  - Skew-Hermitian matrices, i.e., $\mathbf{A} \in \mathbb{F}^{n \times n}$ such that $\mathbf{A}^H = -\mathbf{A}$.
  - Skew-symmetric matrices, i.e., $\mathbf{A} \in \mathbb{R}^{n \times n}$ such that $\mathbf{A}^T = -\mathbf{A}$.

Theorem

A normal triangular matrix must be diagonal.

Theorem (Spectral characterization of normal matrices)

A matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ is normal if and only if each of its eigenvectors is also an eigenvector of $\mathbf{A}^H$. That is, $\mathbf{A}$ is normal if and only if for every eigen-pair $(\lambda, \mathbf{u}) \in \mathbb{C} \times \mathbb{C}^n$ such that $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$, we have $\mathbf{u}^H\mathbf{A} = \lambda\mathbf{u}^H$.

Characterization of Hermitian matrices

- A square matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ is called **Hermitian** if it satisfies $\mathbf{A}^H = \mathbf{A}$.

Theorem

- *The eigenvalues of a Hermitian matrix are real.*
- *A normal matrix whose eigenvalues are real is Hermitian.*

- The ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ of a Hermitian matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ are characterized by optimality properties of the **Rayleigh quotient**

$$\mu_{\mathbf{A}} : \mathbf{x} \in \mathbb{C}^n \backslash \{\mathbf{0}\} \mapsto \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} \in \mathbb{R}.$$

Theorem (Courant-Fisher min-max principle)

The ordered eigenvalues of a Hermitian matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ are such that

$$\begin{aligned}\lambda_k &= \min_{\mathcal{S} \subseteq \mathbb{C}^n, \dim(\mathcal{S})=n-k+1} \; \max_{\mathbf{x}\in\mathcal{S}\backslash\{\mathbf{0}\}} \; \mu_{\mathbf{A}}(\mathbf{x}) \\ &= \min_{\mathcal{S} \subseteq \mathbb{C}^n, \dim(\mathcal{S})=k} \; \max_{\mathbf{x}\in\mathcal{S}\backslash\{\mathbf{0}\}} \; \mu_{\mathbf{A}}(\mathbf{x}).\end{aligned}$$

Characterization of Hermitian matrices, cont'd

- A corrolary of the (Courant-Fisher) min-max principle is that the largest and smallest eigenvalues of a Hermitian matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ are such that

$$\lambda_1 = \max_{\mathbf{x} \in \mathbb{C}^n \backslash \{\mathbf{0}\}} \mu_{\mathbf{A}}(\mathbf{x}) \quad \text{and} \quad \lambda_n = \min_{\mathbf{x} \in \mathbb{C}^n \backslash \{\mathbf{0}\}} \mu_{\mathbf{A}}(\mathbf{x}).$$

- Another way to characterize the optimality of the ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ of a Hermitian matrix is

Theorem (Courant characterization)

The largest eigenvalue λ_1 of a Hermitian matrix $\mathbf{A}$ with a corresponding eigenvector $\mathbf{u}_1$ is such that

$$\lambda_1 = \mu_{\mathbf{A}}(\mathbf{u}_1) = \max_{\mathbf{x} \in \mathbb{C}^n \backslash \{\mathbf{0}\}} \mu_{\mathbf{A}}(\mathbf{x})$$

and the other eigenvalues $\lambda_2 \geq \cdots \geq \lambda_n$, with corresponding eigenvectors, are such that

$$\lambda_k = \mu_{\mathbf{A}}(\mathbf{u}_k) = \max_{\mathbf{x} \in \mathrm{range}([\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}])^\perp \backslash \{\mathbf{0}\}} \mu_{\mathbf{A}}(\mathbf{x}).$$

Characterization of Hermitian (semi)positive-definite matrices

- A matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ is **Hermitian semipositive-definite** if $\mathbf{A}^H = \mathbf{A}$ and

$$\mathbf{x}^H\mathbf{A}\mathbf{x} \geq 0 \;\forall\, \mathbf{x} \in \mathbb{F}^n \setminus \{\mathbf{0}\}. \tag{3}$$

  The **eigenvalues** of such matrices are **non-negative**, i.e., $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$.
- If Eq. (3) is **strictly satisfied**, the matrix is **Hermitian positive-definite**, or **symmetric positive-definite**, i.e., **SPD**, if $\mathbb{F} = \mathbb{R}$.
  The **eigenvalues** of such matrices are **positive**, i.e., $\lambda_1 \geq \ldots \geq \lambda_n > 0$.
- Hermitian positive definite matrices are **invertible**.
  They admit **Cholesky decompositions** of the form $\mathbf{A} = \mathbf{L}\mathbf{L}^H$ where $\mathbf{L}$ is lower-triangular.
- Hermitian (semi)positive-definite matrices are common in practice:
  - **Low-rank approximation**: $\mathbf{A}^H\mathbf{A}$ and $\mathbf{A}\mathbf{A}^H$.
  - **Optimization**: Hessian matrices.
  - **Machine learning**: Covariance matrices, kernel matrices.
  - **Computational physics**: Discretizatized PDEs.
  - **Statistics**: Fisher information matrices.

Matrix canonical forms

Section 2.4 in Darve & Wootters (2021)

Similarity transformations

- Similarity is an equivalence relation between linear maps induced by a change of basis, that preserves key properties of the underlying linear operator.

Definition (Similar matrices)

Two square matrices $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{n \times n}$ are **similar** if there exists an invertible matrix $\mathbf{X} \in \mathbb{F}^{n \times n}$ such that $\mathbf{A} = \mathbf{X}\mathbf{B}\mathbf{X}^{-1}$.

- The mapping $\mathbf{A} \mapsto \mathbf{B}$ is a **similarity transformation**, which recasts the matrix representation of a linear map in a different basis.
- Similar matrices have the same **rank**, **characteristic polynomial** and underlying properties (i.e., **determinant**, **eigenvalues** and their **algebraic multiplicities**, **trace**, ...), **geometric multiplicities**, ...
- An eigenvector $\mathbf{v}$ of $\mathbf{B}$ is transformed into an eigenvector $\mathbf{u} := \mathbf{X}\mathbf{v}$ of the similar $\mathbf{A} = \mathbf{X}\mathbf{B}\mathbf{X}^{-1}$.
- Similarity is crucial for finding canonical forms, such as the **diagonal form**, the **Jordan form** and the **Schur form**, which provide simplified representations useful in solving and analyzing numerical linear algebraic problems.

Diagonal form and eigen-decomposition

- The simplest similar form into which a matrix may be reduced is the diagonal form, i.e., $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}$, for some diagonal matrix $\mathbf{D}$.

Theorem

A square matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ is diagonalizable iff it has n linearly independent eigenvectors or, equivalently, iff it is not defective.

- In particular, every diagonalizable matrix $\mathbf{A}$ can be recast into an **eigen-decomposition** of the form $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$, where $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues, and the columns of $\mathbf{U}$ are normalized eigenvectors.

  If $\mathbf{A}$ is **normal**, then $\mathbf{U}^{-1} = \mathbf{U}^H$. **All normal matrices are diagonalizable.**
- An **invertible** matrix is **not necessarily diagonalizable**, e.g.:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ is invertible, but defective.}$$

- A **diagonalizable** matrix is **not necessarily invertible**, e.g.:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ is diagonal, but singular.}$$

Jordan form

- For defective matrices, an alternative representation is the **Jordan form**.

Definition (Jordan block)

A Jordan block is either a scalar λ, or a matrix of the form:

$$\mathbf{J} = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{bmatrix}.$$

Theorem (Reduction to Jordan form)

- *For every matrix* $\mathbf{A} \in \mathbb{F}^{n\times n}$*, there exist* $\mathbf{X}, \mathbf{J} \in \mathbb{C}^{n\times n}$ *such that* $\mathbf{A} = \mathbf{X}\mathbf{J}\mathbf{X}^{-1}$*, where* $\mathbf{J}$ *is a block-diagonal matrix of Jordan blocks, i.e.,* $\mathbf{J} = \mathrm{diag}(\mathbf{J}_1, \dots, \mathbf{J}_k)$*.*
- *For each distinct eigenvalue* λ *of* $\mathbf{A}$*, there is a number of associated Jordan blocks in* $\mathbf{J}$ *given by the geometric multiplicity of* λ*, whereas the sizes of these blocks add up to the algebraic multiplicity of* λ*.*

- The Jordan form is a generalization of the eigen-decomposition, mostly used for the analysis of defective matrices.

Schur form/decomposition

- A Jordan form exists for every square matrix, but its computation can be challenging, and the associated basis ill-conditioned.
- An alternative to the Jordan form, which also exists for every square matrix, is the **Schur form** (or **decomposition**).

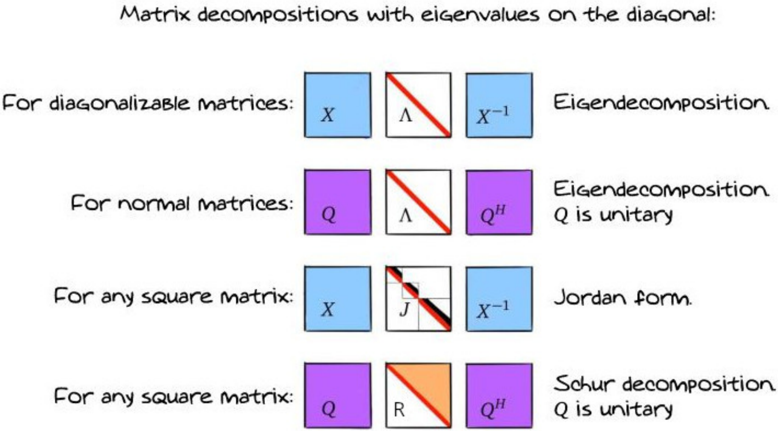
Theorem (Schur decomposition)

For every matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$, *there exist* $\mathbf{Q}, \mathbf{R} \in \mathbb{C}^{n\times n}$ *such that*

$$\mathbf{A} = \mathbf{Q}\mathbf{R}\mathbf{Q}^H$$

where $\mathbf{R}$ *is upper triangular, and* $\mathbf{Q}$ *is unitary, i.e.,* $\mathbf{Q}^{-1} = \mathbf{Q}^H$.

- Since $\mathbf{R}$ is a triangular matrix, its eigenvalues are listed on the diagonal. And since $\mathbf{R}$ and $\mathbf{A}$ are similar, the values listed on the diagonal of $\mathbf{R}$ are also the eigenvalues of $\mathbf{A}$.



Darve, E., & Wootters, M. (2021). Numerical linear algebra with Julia. Society for Industrial and Applied Mathematics.

Matrix functions

- A **matrix function** $f(\mathbf{A})$ extends scalar functions like e^x, $sin(x)$, or $\log(x)$ to matrices $\mathbf{A} \in \mathbb{F}^{n\times n}$.
- The previously introduced canonical forms can be used to define the application of matrix functions. In particular, for
  - **Diagonalizable matrices**: If $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1}$ where $\mathbf{D}$ is diagonal, then

    $$f(\mathbf{A}) = \mathbf{U}f(\mathbf{D})\mathbf{U}^{-1} \quad \text{where} \quad f(\mathbf{D}) = \mathrm{diag}(f(d_{11}), \ldots, f(d_{nn})).$$

  - **Jordan forms**: If $\mathbf{A} = \mathbf{X}\mathbf{J}\mathbf{X}^{-1}$ where $\mathbf{J}$ is a Jordan matrix, then

    $$f(\mathbf{A}) = \mathbf{X}f(\mathbf{J})\mathbf{X}^{-1}$$

    where $f(\mathbf{J})$ is block diagonal, with each block corresponding to a Jordan block. Computing functions of Jordan blocks requires handling the non-diagonal terms via derivatives of the scalar function.
  - **Schur form**: If $\mathbf{A} = \mathbf{Q}\mathbf{R}\mathbf{Q}^H$ where $\mathbf{R}$ is upper triangular, then:

    $$f(\mathbf{A}) = \mathbf{Q}f(\mathbf{R})\mathbf{Q}^H$$

    For upper triangular matrices, matrix functions can be computed recursively based on the entries of $\mathbf{R}$.

Singular value decomposition

Section 2.5 in Darve & Wootters (2021)

Singular value decomposition

- **Eigenvalues** and related **canonical forms** are **for square matrices only**. But **singular value decompositions** exist **for all matrices**.

Theorem (Singular value decomposition (SVD))

- *For every matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$, there exist decompositions of the form*

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

where $\mathbf{U} \in \mathbb{F}^{m \times m}$ and $\mathbf{V} \in \mathbb{F}^{n \times n}$ are unitary matrices, and the only non-zero entries of $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ are on the diagonal.
- *The diagonal entries $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$ of $\mathbf{\Sigma}$ are called **singular values**.*
- *The columns $\mathbf{u}_1, \ldots, \mathbf{u}_m \in \mathbb{F}^m$ of $\mathbf{U}$ are called **left singular vectors**.*
- *The columns $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^v$ of $\mathbf{V}$ are called **right singular vectors**.*

- The **singular values** $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$ are **unique**, but the matrices $\mathbf{U}$ and $\mathbf{V}$ of left and right singular vectors are not unique.
- If $\mathbf{A} \in \mathbb{F}^{m \times n}$ has rank $r < p$ where $p := \min(m, n)$, then $\sigma_k = 0$ for $k = r + 1, \ldots, p$. Thus, the number of **non-zero singular values** equals the **rank** of $\mathbf{A}$.

Relation between singular and eigenvalue problems

- If $\mathbf{A} \in \mathbb{F}^{n\times n}$ is **Hermitian** with (real) eigenvalues ordered such that $|\lambda_1| \geq \cdots \geq |\lambda_n|$, then the ordered singular values $\sigma_1 \geq \cdots \geq \sigma_n$ of $\mathbf{A}$ are given by $\sigma_i = |\lambda_i|$ for $i = 1, \ldots, n$.
- If $\mathbf{A} \in \mathbb{F}^{m\times n}$ is not Hermitian, or even square, then the singular values and singular vectors of $\mathbf{A}$ can be characterized as eigen-pairs of Gram matrices.

Theorem (Relation between the SVD and Gram matrices)

For every $\mathbf{A} \in \mathbb{F}^{m\times n}$ with $p := \min(m, n)$ singular values $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$ and a proper SVD $\mathbf{U\Sigma V}^H$ with $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ and $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$,

- *$(\sigma_1, \mathbf{u}_1), \ldots, (\sigma_p, \mathbf{u}_p) \in \mathbb{R} \times \mathbb{F}^n$ are eigen-pairs of $\mathbf{A}^H\mathbf{A} \in \mathbb{F}^{n\times n}$,*
- *$(\sigma_1, \mathbf{v}_1), \ldots, (\sigma_p, \mathbf{v}_p) \in \mathbb{R} \times \mathbb{F}^m$ are eigen-pairs of $\mathbf{A}\mathbf{A}^H \in \mathbb{F}^{m\times m}$.*

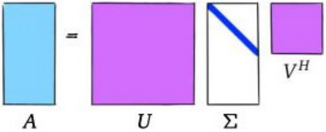
- As a corollary, we have: $\|\mathbf{A}\|_2 = \sigma_1(\mathbf{A}) = \sqrt{\lambda_{max}(\mathbf{A}^H\mathbf{A})} = \sqrt{\lambda_{max}(\mathbf{A}\mathbf{A}^H)}$

$$\text{and} \quad \|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{p} \sigma_i^2}.$$

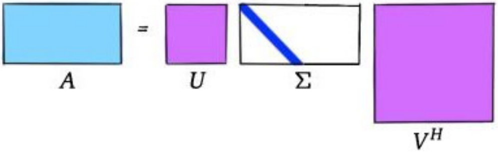
Compact singular value decomposition

- For rectangular matrices, we denote two types of structures of SVD:

Tall-and-skinny matrices ($m > n$)

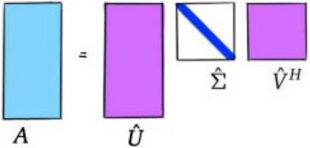


Short-and-fat matrices ($m < n$)

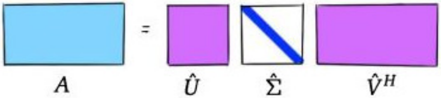


- Due to the structure of $\mathbf{\Sigma}$, a **compact SVD** $\mathbf{A} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^H$ is formed by discarding the zero block of $\mathbf{\Sigma}$, and the corresponding blocks of singular vectors in $\mathbf{U}$ or $\mathbf{V}$, leading to $\hat{\mathbf{U}} \in \mathbb{F}^{m\times p}$, $\hat{\mathbf{\Sigma}} \in \mathbb{R}^{p\times p}$ and $\hat{\mathbf{V}} \in \mathbb{F}^{n\times p}$:

Tall-and-skinny matrices ($m > n$)



Short-and-fat matrices ($m < n$)



Darve, E., & Wootters, M. (2021). Numerical linear algebra with Julia. Society for Industrial and Applied Mathematics.

Low-rank approximation

- The factors of th SVD capture essential information about the action of $\mathbf{A}$:

Theorem (Four fundamental subspaces and the SVD)

Every matrix $\mathbf{A} \in \mathbb{F}^{m\times n}$ *of rank* r *with left singular vectors* $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathbb{F}^n$*, and right singular vectors* $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{F}^m$ *is such that*

$$\mathrm{range}(\mathbf{A}) = \mathrm{span}\{\mathbf{u}_1, \dots, \mathbf{u}_r\}, \qquad \mathrm{null}(\mathbf{A}) = \mathrm{span}\{\mathbf{v}_{r+1}, \dots, \mathbf{v}_n\},$$
$$\mathrm{range}(\mathbf{A}^H) = \mathrm{span}\{\mathbf{v}_1, \dots, \mathbf{v}_r\}, \qquad \mathrm{null}(\mathbf{A}^H) = \mathrm{span}\{\mathbf{u}_{r+1}, \dots, \mathbf{u}_m\}.$$

Theorem (Eckart-Young theorem)

Consider $\mathbf{A} \in \mathbb{F}^{m\times n}$*, its* $p = \min(m, n)$ *singular values* $\sigma_1 \geq \cdots \geq \sigma_1 \geq 0$ *and corresponding left singular vectors* $\mathbf{u}_1, \dots, \mathbf{u}_{\min(p,m)}$ *and right singular vectors* $\mathbf{v}_1, \dots, \mathbf{v}_{\min(p,n)}$*. Then, for* $r < p$*, the matrix* $\mathbf{B} := \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^H$ *is such that*

$$\|\mathbf{A} - \mathbf{B}\|_2 = \sigma_{r+1} \;\; \textit{and} \;\; \|\mathbf{A} - \mathbf{B}\|_F = \sum_{i=r+1}^{p} \sigma_i^2.$$

Moreover, $\mathbf{B}$ *minimizes* $\|\mathbf{A} - \mathbf{B}\|_2$ *and* $\|\mathbf{A} - \mathbf{B}\|_F$ *among matrices of rank* r*.*