Linear Algebra

Vector spaces

A vector space over a field F $(F, V, V\times V \to V, F\times V \to V)$ is a set V together with a function $V\times V \to V,\ (\boldsymbol{x},\boldsymbol{y}) \mapsto \boldsymbol{x}+\boldsymbol{y}$ (called **addition**) and a function $F\times V \to V,\ (c,\boldsymbol{x}) \mapsto c\boldsymbol{x}$ (called **scalar multiplication**) for which

$\forall \boldsymbol{x},\boldsymbol{y}\in V$	:	$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$	commutativity
$\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in V$	:	$(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$	associativity
$\exists \boldsymbol{0}\in V \forall \boldsymbol{x}\in V$	:	$\boldsymbol{0}+\boldsymbol{x}=\boldsymbol{x}$	additive identity
$\forall \boldsymbol{x}\in V \exists -\boldsymbol{x}\in V$	:	$(-\boldsymbol{x})+\boldsymbol{x}=\boldsymbol{0}$	additive inverse
$\forall c\in F \forall \boldsymbol{x},\boldsymbol{y}\in V$	:	$c(\boldsymbol{x}+\boldsymbol{y})=c\boldsymbol{x}+c\boldsymbol{y}$	distributivity
$\forall c,d\in F \forall \boldsymbol{x}\in V$	:	$(c+d)\boldsymbol{x}=c\boldsymbol{x}+d\boldsymbol{x}$	distributivity
$\forall c,d\in F \forall \boldsymbol{x}\in V$	:	$(cd)\boldsymbol{x}=c(d\boldsymbol{x})$	associativity
$\forall \boldsymbol{x}\in V$	:	$1\boldsymbol{x}=\boldsymbol{x}$	scalar identity

Closure is implied by this axioms.

A subset W of a vector space V is a **subspace** iff it is a vector space with the same vector space operations as V.

$W\subset V$ is a **subspace** iff
$\boldsymbol{0}\in W$
$\forall \boldsymbol{x},\boldsymbol{y}\in W \forall c\in F\ :\ \boldsymbol{x}+\boldsymbol{y}\in W\ \wedge\ c\boldsymbol{x}\in W$ closure

Linear combination & Bases

A **linear combination** is a finite sum (even in infinite dimensional vector spaces) given by

$$\boldsymbol{Lin}(\mathcal{S},c_1,\dots,c_n) = \sum_{\boldsymbol{v}\in\mathcal{S}} c_i\boldsymbol{v} \qquad \mathcal{S}\subset V,\ c_i\in F$$
$$\boldsymbol{Lin}(\emptyset) = \boldsymbol{0}$$

The **span** of vectors in $\mathcal{S}\subset V$ is the set of all linear combinations
$\mathrm{Span}(\mathcal{S}) = \{\boldsymbol{Lin}(\mathcal{S},c_1,\dots,c_n)\mid c_1,\dots,c_n\in F\}$

A set of vectors $\mathcal{S}$ is **linearly independent** iff
$\boldsymbol{Lin}(\mathcal{S},c_1,\dots,c_n)=\boldsymbol{0} \iff c_1,\dots,c_n=0$

A set of vectors $\mathcal{B}\subset V$ in a vector space V is called a **basis** iff

- $\mathcal{B}$ is linearly independent
- $\mathrm{Span}(\mathcal{B})=V$

Every vector space has a basis.

We call $(\boldsymbol{e}_i)_j=\delta_{ij},\ i=1,\dots,n$ the standard basis of F^n.

We call $\dim(V)=\#(\mathcal{B})$ the **dimension** of V. The dimension (and thus the number of vectors in a basis) is unique for every vector space.

There are at most $\dim(V)$ linearily independent vectors in V.

Linear functions

A **linear function** (or linear transformation) is a function $T:V\to W$ such that

$$T(\boldsymbol{x}+\boldsymbol{y}) = T(\boldsymbol{x})+T(\boldsymbol{y})$$
$$T(c\boldsymbol{x}) = cT(\boldsymbol{x})$$

The set of all linear functions $T:V\to W$ (denoted by $\mathrm{L}(V,W)$) is a vector space.

$\dim V=n,\ \dim W=m$ and $\mathcal{B}_V,\mathcal{B}_W$ is a basis of V,W then there exists a unique matrix $\boldsymbol{A}\in F^{m\times n}$ such that $T(\boldsymbol{v})=\boldsymbol{A}\boldsymbol{v}$.

Properties that are independent of the chosen basis (such as trace, determinant or eigenvalues) are assigned to the function represented by the matrices.

Two matrices $\boldsymbol{A},\boldsymbol{B}\in F^{m\times n}$ are **equivalent** iff
$\exists \boldsymbol{P}\in \mathrm{GL}_m(F),\ \boldsymbol{Q}\in\mathrm{GL}_n(K)\ :\ \boldsymbol{B}=\boldsymbol{PAQ}$ or
$\mathrm{rg}(\boldsymbol{A})=\mathrm{rg}(\boldsymbol{B})$ or
iff $\boldsymbol{A},\boldsymbol{B}$ represent the same functions $T:V\to W$ regarding different bases.

Two matrices $\boldsymbol{A},\boldsymbol{B}\in F^{n\times n}$ are **similar** iff
$\exists \boldsymbol{T}\in\mathrm{GL}_n(K)\ :\ \boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{AT}$ or
iff $\boldsymbol{A},\boldsymbol{B}$ represent the same function $f:V\to V$ regarding different bases.

The **kernel** of a linear function is defined by $\ker(T)=\{\boldsymbol{x}\in V\mid f(\boldsymbol{x})=0_W\}$.

The **image/range** of a linear function is defined by
$\mathrm{im}(T)=\mathrm{range}(T)=\{f(\boldsymbol{x})\mid \boldsymbol{x}\in V\}$.

Matrices

A $m\times n$ **matrix** $\boldsymbol{A}\in F^{m\times n}$ over F is a function (usually written as a rectangular grid)

$$\boldsymbol{A}\ :\ \{1,\dots,m\}\times\{1,\dots,n\}\to F,\ (i,j)\mapsto \boldsymbol{A}_{ij}$$

A matrix $\boldsymbol{A}$ is square iff $m=n$. Then we call n the size of $\boldsymbol{A}$.

Equality, addition, and scalar multiplication is defined component-wise.

The vectors in F^n are usually identified with the matrices in $F^{n\times 1}$(thus are column vectors).

Identity matrix $(\boldsymbol{I}_n)_{ij}=\delta_{ij}$. It holds that $\boldsymbol{I}_n\boldsymbol{A}=\boldsymbol{A}\boldsymbol{I}_n=\boldsymbol{A}$.

Special matrices

$\forall i\neq j$	:	$\boldsymbol{A}_{ij}=0$	diagonal matrix
$\forall i>j$	:	$\boldsymbol{A}_{ij}=0$	upper triangular matrix
$\forall i<j$	:	$\boldsymbol{A}_{ij}=0$	lower triangular matrix

$\boldsymbol{A},\boldsymbol{B}\in F^{n\times n}$ and upper (lower) triangular
$\implies \boldsymbol{AB}$ is upper (lower) triangular

An **elementary matrix** is an invertible square matrix obtained by

- A multiple of one row of $\boldsymbol{I}_n$ is added to a different row
- Two different rows of $\boldsymbol{I}_n$ are exchanged
- One row of $\boldsymbol{I}_n$ is multiplied by a nonzero scalar

Matrix multiplication

$$\boldsymbol{A}\in F^{m\times n},\boldsymbol{B}\in F^{n\times p}$$
$$(\boldsymbol{AB})_{ij}\in F^{m\times p} = \textstyle\sum_{k=1}^n \boldsymbol{A}_{ik}\boldsymbol{B}_{kj}$$

Matrix multiplication is associative and distributive over matrix addition (from the left and from the right) but **not** commutative.

Transpose of a matrix

$$\boldsymbol{A}\in F^{m\times n}$$
$$(\boldsymbol{A}^T)_{ij}\in F^{n\times m}=\boldsymbol{A}_{ji}$$

$$(\boldsymbol{A}^T)^T = \boldsymbol{A}$$
$$(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T+\boldsymbol{B}^T$$
$$(c\boldsymbol{A})^T = c\boldsymbol{A}^T$$
$$(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$$

Rank of a matrix

The **range** of a matrix is the vector space spanned by the columns (also called the **column space**).

The **rank** of a matrix (denoted by $\mathrm{rank}\boldsymbol{A}$) is the number of independent columns of $\boldsymbol{A}$ (This is equivalent to the number of linearly independent rows of $\boldsymbol{A}$.
Equivalently: $\mathrm{rank}\boldsymbol{A}=\dim(\mathrm{range}(\boldsymbol{A}))$

$$\mathrm{rank}\boldsymbol{A}=0 \iff \boldsymbol{A}=\boldsymbol{0}$$
$$\mathrm{rank}(\boldsymbol{A}+\boldsymbol{B}) \le \mathrm{rank}\boldsymbol{A}+\mathrm{rank}\boldsymbol{B}$$
$$\mathrm{rank}(\boldsymbol{AB}) \le \min(\mathrm{rank}\boldsymbol{A},\mathrm{rank}\boldsymbol{B})$$
$$\mathrm{rank}(\boldsymbol{A}^T\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}\boldsymbol{A}^T)=\mathrm{rank}(\boldsymbol{A})$$

Iff $F=\mathbb{C}$ we call $\boldsymbol{A}^*=\overline{\boldsymbol{A}}^T$ the **Hermitian adjoint.**

We call a square matrix **symmetric** iff $\boldsymbol{A}^T=\boldsymbol{A}$ and **skew-symmetric** iff $\boldsymbol{A}^T=-\boldsymbol{A}$.

We call a square matrix **hermitian** iff $\boldsymbol{A}^*=\boldsymbol{A}$ and **skew-hermitian** iff $\boldsymbol{A}^*=-\boldsymbol{A}$.

Permutations

A **permutation** is a bijective function $\sigma:S\to S$ where S is a finite set. We write e.g. $\begin{pmatrix} 1 & 2 & \dots & n\\ \sigma(1) & \sigma(2) & \dots & \sigma(n)\end{pmatrix}$

A **cycle** is a permutation such that $\exists a_1,\dots,a_k\in S: f(a_i)=a_{i+1}\wedge f(a_k)=a_1$. Every permutation can be written as a product of cycles.
A cycle of length two is called a **transposition**. Every cycle can be written as a product (combination) of transpositions.

The **sign** of a permutation is defined by $\mathrm{sgn}(\sigma)=\begin{cases}1 & n \text{ even}\\ -1 & n \text{ odd}\end{cases}$ (where n is the number of transposition in the decomposition of the permutation)

Determinant of a (square) matrix

$$\det \boldsymbol{A} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \boldsymbol{A}_{i,\sigma(i)}$$

Rule of Sarrus

$$\det \boldsymbol{A} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$\det \boldsymbol{A} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh$$

Laplace expansion

$\det \boldsymbol{A} = \sum_{j=1}^{n} \boldsymbol{A}_{ij}(-1)^{i+j}\boldsymbol{M}_{ij}$

$\boldsymbol{M}_{ij}$ is the ij–th minor (row i and column j are removed from $\boldsymbol{A}$).

Determinant by Gaussian Elimination

- Exchange of a row/column changes the sign of the determinant
- Multiplying a row/column by c multiplies the determinant by c
- Adding a multiple of a row/column to another leaves the determinant unchangend

then use: If $\boldsymbol{A}$ is a **triangular matrix** then $\det \boldsymbol{A} = \sum_{i=1}^{n} a_{ii}$.

$$\begin{aligned} \det(\boldsymbol{AB}) &= \det \boldsymbol{A} \det \boldsymbol{B} \\ \det(c\boldsymbol{A}) &= c^n \det(\boldsymbol{A}) \\ \det(\boldsymbol{A}^{-1}) &= \det(\boldsymbol{A})^{-1} \\ \det(\boldsymbol{A}^T) &= \underline{\det(\boldsymbol{A})} \\ \det(\boldsymbol{A}^*) &= \overline{\det(\boldsymbol{A})} \\ \det \boldsymbol{A} &= \lambda_1 \cdot \ldots \cdot \lambda_n \quad n \times n \text{ matrices} \end{aligned}$$

Trace of a (square) matrix

$\operatorname{tr}\boldsymbol{A} \in F = \sum_{k=1}^{n} \boldsymbol{A}_{kk}$

$$\begin{aligned} \operatorname{tr}(\boldsymbol{A}+\boldsymbol{B}) &= \operatorname{tr}\boldsymbol{A} + \operatorname{tr}\boldsymbol{B} \\ \operatorname{tr}(r\boldsymbol{B}) &= r \cdot \operatorname{tr}\boldsymbol{A} \\ \operatorname{tr}(\boldsymbol{A}) &= \operatorname{tr}(\boldsymbol{A}^T) \\ \operatorname{tr}(\boldsymbol{AB}) &= \operatorname{tr}(\boldsymbol{BA}) \\ \operatorname{tr}(\boldsymbol{ABC}) &= \operatorname{tr}(\boldsymbol{CAB}) = \operatorname{tr}(\boldsymbol{BCA}) \quad \text{cyclic permutation} \\ \operatorname{tr}(\boldsymbol{A}) &= \lambda_1 + \cdots + \lambda_n \quad n \times n \text{ matrix} \end{aligned}$$

System of linear equations

A matrix $\boldsymbol{A}$ is in **row echelon form (REF)** iff

- All nonzero rows are above any rows of all zeroes
- The leading coefficient (called pivot) of a row is always strictly to the right of the leading coefficient of the row above it.

A matrix $\boldsymbol{A}$ is in **reduced row echelon form (RREF)** iff additionally

- The leading entry in every row is 1 and every other entry of that column is 0

Elementary row operation on a matrix are left multiplications with elementary matrices. Thus

- Add a multiple of one row to a different row
- Exchange two different rows
- Multiply one row by a nonzero scalar

Gaussian Elimination is the algorithm that transforms a matrix by applying elementary row operations to a matrix in **REF**.
Gauss-Jordan Elimination is the algorithm that transforms a matrix by applying elementary row operations to a matrix in **RREF**.

Every **system of linear equations** can be converted to **matrix-vector** form

$$\begin{matrix} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{matrix} \longrightarrow \begin{matrix} \boldsymbol{Ax} = \boldsymbol{b} \\ \boldsymbol{A} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \\ \boldsymbol{b} = [b_1 \ \ldots \ b_m]^T \\ \boldsymbol{x} = [x_1 \ \ldots \ x_n]^T \end{matrix}$$

The system of linear equations can be solved by applying **Gauss-Jordan elimination** to the **augmented matrix** $\boldsymbol{A}|\boldsymbol{b} \in F^{m \times (n+1)}$.

The solution of a system of linear equations forms an **affine space** which can be written as $\boldsymbol{z} + \operatorname{span}(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_k)$ where $\boldsymbol{z}$ is an arbitrary solution and $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_k$ is a basis of the homogenous solution space where $k = \operatorname{rank}(\boldsymbol{A})$.

$\operatorname{span}(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_k) = \{\boldsymbol{x} | \ \boldsymbol{Ax} = \boldsymbol{0}\}$

Inverse matrix

A matrix $\boldsymbol{A} \in F^{n \times n}$ is called **invertible** or **nonsingular** iff
$\exists \boldsymbol{B} \in F^{n \times n} \ : \ \boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}_n$ We write $\boldsymbol{A}^{-1} = \boldsymbol{B}$.

A matrix $\boldsymbol{A}$ is invertible iff $\det \boldsymbol{A} \neq 0$.
A matrix $\boldsymbol{A}$ is invertible iff $0 \notin \sigma(\boldsymbol{A})$.

The group of all invertible $n \times n$ Matrices over F is denoted by $\mathrm{GL}_n(F)$ and is called the **general linear group** (for $n \geq 2$, $\mathrm{GL}_n(F)$ is **not** commutative).

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Longrightarrow \boldsymbol{A}^{-1} = \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

-

Inversion by Gauss-Jordan Elimination
Use Gauss-Jordan Elimination to transform the augmented matrix $[\boldsymbol{A}|\boldsymbol{I}_n]$ to $[\boldsymbol{I}_n\boldsymbol{B}]$ then $\boldsymbol{B} = \boldsymbol{A}^{-1}$.

Eigenvalues & Eigenvectors

$\lambda \in F$ is an **eigenvalue** to the **eigenvector** $\boldsymbol{x} \neq \boldsymbol{0} \in F^n$ of the Matrix $\boldsymbol{A} \in F^{n \times n}$ iff $\boldsymbol{Ax} = \lambda\boldsymbol{x}$.

The set of all eigenvectors (also called the **spectrum**) is denoted by $\sigma(\boldsymbol{A}) = \{\lambda | \boldsymbol{Ax} = \lambda\boldsymbol{x}, \ \boldsymbol{x} \neq 0\}$

The **eigenspace** to the eigenvector λ is a vector space and denoted by $E_\lambda(\boldsymbol{A}) = \{\boldsymbol{x} | \ \boldsymbol{Ax} = \lambda\boldsymbol{x}, \ \boldsymbol{x} \neq 0\}$

The **characteristic polynomial** of $\boldsymbol{A}$ is given by $p_{\boldsymbol{A}}(z) = \det(z\boldsymbol{I}_n - \boldsymbol{A}) = \det(\boldsymbol{A} - z\boldsymbol{I}_n)$.

$\sigma(\boldsymbol{A}) = \{z \in F | \ p_{\boldsymbol{A}}(z) = 0\}$

The **(algebraic) multiplicity** $\alpha(\lambda)$ is the number of times the eigenvalue occurs as a root in $p_{\boldsymbol{A}}(z)$. We call the eigenvalue **simple** iff $\alpha(\lambda) = 1$.

The (geometric) multiplicity $\gamma(\lambda) = \dim(E_\lambda(\boldsymbol{A}))$

We call the matrix $\boldsymbol{A}$ nonderogatory iff $\forall \lambda \in \sigma(\boldsymbol{A}) \ : \ \gamma(\lambda) = 1$.

Properties of eigenvalues & eigenvectors

If $\boldsymbol{A}$ is a triangular matrix $\sigma(\boldsymbol{A}) = \{a_{11}, \ldots, a_{nn}\}$.

$$\begin{aligned} \sigma(\boldsymbol{A}) &= \sigma(\boldsymbol{A}^T) \quad \text{the multiplicities are also the same} \\ \lambda \in \sigma(\boldsymbol{A}) &\iff \lambda^{-1} \in \sigma(\boldsymbol{A}^{-1}) \end{aligned}$$

Diagonalization

A matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is **diagonalizable** iff it is similar to a diagonal matrix; thus, $\exists \boldsymbol{P} \in GL_n(\mathbb{C}) \ : \ \boldsymbol{P}^{-1}\boldsymbol{AP}$ is a diagonal matrix.

A matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is **diagonalizable** iff the sum of the dimensions of its eigenspaces is equal to n.

If a matrix $\boldsymbol{A}$ has n distinct eigenvalues, then $\boldsymbol{A}$ is diagonalizable.

To **find** a matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{AP}$ is a diagonal matrix we

- Find bases $\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{ir_i}$ for $E_{\lambda_i}(\boldsymbol{A})$ for each of the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ of $\boldsymbol{A}$
- $\boldsymbol{P} = [\boldsymbol{x}_{11} \ldots \boldsymbol{x}_{1r_1} \ldots \boldsymbol{x}_{k1} \ldots \boldsymbol{x}_{kr_k}]$

Properties of diagonalizability

$\boldsymbol{A}$ diagonalizable $\Longrightarrow$ $\boldsymbol{A}^T, \boldsymbol{A}^{-1}, \boldsymbol{A}^k \ k \in \mathbb{N}$ are diagonalizable

---

Cramer's rule

A system of linear equations has a **unique solution** iff $\boldsymbol{A}$ is invertible. In this case

$x_i = \frac{\det \boldsymbol{B}(i)}{\det \boldsymbol{A}}$

where $\boldsymbol{B}(i)$ is the matrix that is formed by replacing the i-th column of $\boldsymbol{A}$ by $\boldsymbol{b}$.