Chapter 2 : Matrix Algebra

$$\left| \begin{matrix} ax_1 + bx_2 = c \\ ex_1 + fx_2 = g \end{matrix} \right| \longrightarrow \begin{pmatrix} a & b \\ e & f \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c \\ g \end{pmatrix}$$

$2\times 2 — 2\times 1$

Matrix Multiplication

$$\underbrace{\begin{pmatrix} a & b \\ e & f \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{x} = \begin{pmatrix} ax_1 + bx_2 \\ ex_1 + fx_2 \end{pmatrix} = \underbrace{\begin{pmatrix} c \\ g \end{pmatrix}}_{b}$$

Matrix representation $Ax = b$

<u>Matrix</u> $Mat(n \times m, \mathbb{R})$ as a „vector of vectors"

$$A = \begin{pmatrix} a_{11} & - - & - - & a_{1m} \\ \vdots & & & \vdots \\ a_{n1} & - - & - - & a_{nm} \end{pmatrix}$$

$a_i^r = (a_{i1}, \dots a_{im})$ i'th row vector

$a_j^c = (a_{1j}, \dots, a_{nj})$ is the j'th column vector

$\Rightarrow$ Vector space of real matrices $\underline{Mat(n \times m, \mathbb{R})}$

$\hookrightarrow$ Vectors of real vectors

Some rules :

- Addition $\forall n, m \in \mathbb{N}$, $A, B \in M_{n\times m}$

  $A + B = (a_{ij} + b_{ij})_{i \in \{1, \dots n\}\, j \in \{1, \dots m\}}$

  Example:

$$\underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{\in M_{2\times 2}} + \underbrace{\begin{pmatrix} e & f \\ g & h \end{pmatrix}}_{\in M_{2\times 2}} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix} \in M_{2\times 2}$$

- Scalar multiplication $\lambda \in \mathbb{R}$ $\quad A \in M_{n \times m}$

$$\lambda A = (\lambda \, a_{ij})_{i \in \{1, \dots, n\}, j \in \{1, \dots, m\}}$$

$$\lambda \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$$

- $A = B \iff \forall \, i, j \quad a_{ij} = b_{ij}$

- transposed Matrix A', A^t : swap row and column index

Example $\begin{pmatrix} 1 & 0 & 4 \\ 3 & 1 & 2 \end{pmatrix}' = \begin{pmatrix} 1 & 3 \\ 0 & 1 \\ 4 & 2 \end{pmatrix}$

- Zero Matrix $0_{n \times m} = (a_{ij})_{i \in (1, \dots n), j = (1, \dots, n\}}$ for which $\forall i, j \; a_{ij} = 0$

- Square Matrix $A \in M_{n \times n}$ for any $n \in \mathbb{N}$
- symmetric Matrix $A^t = A$
- Diagonal Matrix $A = (a_{ij})_{ij \in \{1, \dots n\}}$ (is a square matrix)

where $\forall \; i \neq j : a_{ij} = 0$

$$\text{e.g.} \quad \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \mathrm{diag}(a, b)$$

- Identity Matrix $I_n = \mathbb{1}_n = \mathrm{diag}(1, \dots, 1)$

$\forall \; A \in M_{n \times m} \quad B \in M_{m' \times n}$

$$\mathbb{1}_n \cdot A = A \wedge B \cdot \mathbb{1}_n = B$$

- upper (lower) triangular Matrix

$A_{n \times n}$ is upper (lower) triangular if

$i > j : a_{ij} = 0 \quad (i < j : a_{ij} = 0)$

$A_{n \times n}$ is upper (lower) triangular if

$i > j : a_{ij} = 0 \quad (i < j : a_{ij} = 0)$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \leftarrow \text{upper triangular}$$

- Matrix Product: $A \in \mathbb{R}^{n \times m}$ $\quad B \in \mathbb{R}^{m \times k}$

(column dimension of A = row dimension of B)

$C \in \mathbb{R}^{n \times k}$ is called the product of A and B

$C = A \cdot B$ if $\forall i \in \{1, \dots, n\} \wedge j \in \{1, \dots k\}$

$$c_{ij} = \sum_{l=1}^{m} a_{il} b_{lj} = \sum_{l=1}^{m} a_{il} b_{lj}$$

Example

$$\underbrace{\begin{pmatrix} \boxed{1 \quad 0} \\ \boxed{3 \quad 2} \end{pmatrix}}_{2 \times 2} \cdot \underbrace{\begin{pmatrix} \boxed{\begin{matrix} -1 \\ 0 \end{matrix}} & \boxed{\begin{matrix} 2 \\ 1 \end{matrix}} & \boxed{\begin{matrix} 4 \\ 3 \end{matrix}} \end{pmatrix}}_{2 \times 3} = \begin{pmatrix} c_{11} = -1 & c_{12} = 2 & c_{13} = 4 \\ c_{21} = -3 & c_{22} = 8 & c_{23} = 18 \end{pmatrix}$$

Properties of the Matrix product

- Associativity: $(A \cdot B) \cdot C = A \cdot (B \cdot C)$
- Distributivity $(A + B) \cdot C = A \cdot C + B \cdot C$

⚠ Matrix multiplication is not commutative

$$A \cdot B \neq BA \text{ in general}$$

- $A, B \in \mathbb{R}^{n \times m}$ $\quad (A + B)' = A' + B'$

$$(A \cdot B)' = B' \cdot A'$$

⚠ Vectors can be thought of as special matrices $v \in \mathbb{R}^{n \times 1}$

A matrix representation for a linear system of equations:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 5 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Solving LS, univariate case

$$a x = b \qquad \text{if } a \neq 0$$
$$x = a^{-1} \cdot b$$

↳ for $n \times n$ Systems, we can also <u>invert</u> a matrix $A \in \mathbb{R}^{n \times n}$ under certain conditions!

given A is invertible and $A \in \mathbb{R}^{n \times n}$

$$A \vec{x} = \underline{\underline{b}}$$
$$\vec{x} = A^{-1} b$$

⚠ We can think of multiplication with a Matrix A as a linear function $A \in \mathbb{R}^{n \times n}$

$$f : \mathbb{R}^n \to \mathbb{R}^n$$
$$x \mapsto Ax$$

Inverse of A: A^{-1} inverse of $A \in \mathbb{R}^{n \times n}$ is defined as the matrix such that $A^{-1} A = \mathbb{1}_n = A A^{-1}$

Open questions: i) when does A^{-1} exist? is it unique?
ii) when is $b \in \operatorname{im}(f)$ (so we can solve the LS)

i) Is the inverse matrix unique? Yes!
(Proof by contradiction)

Consider $A, B, C \in \mathbb{R}^{n \times n}$ and assume $B \neq C$

Assume $\underbrace{A \cdot B = \mathbb{1}}{} = B \cdot A \quad \wedge \quad A \cdot C = \underbrace{\mathbb{1} = C \cdot A}{}$

$$\Rightarrow C = C \mathbb{1} = C(AB) = (CA) \cdot B = \mathbb{1} \cdot B = B$$

$$\Rightarrow B = C \qquad \lightning$$

$\Rightarrow$ The inverse of a matrix is unique!

Existence: When does A^{-1} exist (we say: A is invertible)

Let's try to construct the inverse for an arbitrary matrix $A \in \mathbb{R}^{n \times n}$

Define 3 elementary matrix operations:

E_1 : interchange rows i, j

define $A = \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & a_n & - \end{pmatrix}$ $\tilde{A} = \begin{pmatrix} - & \tilde{a}_1 & - \\ & \vdots & \\ - & \tilde{a}_n & - \end{pmatrix}$

$\tilde{a}_i = a_j \wedge \tilde{a}_j = a_i$

E_2 : scalar multiplication of row i $\lambda \neq 0$

$\tilde{a}_i = \lambda a_i$

$\tilde{a}_j = a_j \quad \forall j \neq i$

E_3 : addition of row j to row i

$\tilde{a}_i = a_i + a_j$

$\tilde{a}_j = a_j \quad \forall j \neq i$

With these 3 operations, we can define the Gauß-Jordan Algorithm

$$\left(\underbrace{\begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix}}_{A} \;\middle|\; \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{\mathbb{1}}\right)$$

Apply the 3 elementary operations until we have an identity on the left side

-> if we can construct an inverse this way, the inverse exists

Combining the elementary matrix operations corresponds to finding the inverse, as the elementary operations can be written as multiplications with specific matrices.

Example

E_1 to interchange row 1 and 3

$$E_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ \end{pmatrix} \cdot \begin{pmatrix} a & b & c \\ d & e & f \\ \end{pmatrix} = \begin{pmatrix} g & h & i \\ d & e & f \\ \end{pmatrix}$$

$$E_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}}_{A} = \begin{pmatrix} g & h & i \\ d & e & f \\ a & b & c \end{pmatrix}$$

E_2 to multiply row 2 by 2

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad E_2 \cdot A = \begin{pmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{pmatrix}$$

E_3: to add row 3 to row 1

$$E_3 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A^{-1} = E_k \cdot E_{k-1} \cdots E_1 \cdot \mathbb{1}$$

<u>Thm: Triangularizability of a Matrix</u>

consider $A \in \mathbb{R}^{n \times m}$ then, if $n = m$, A can be brought to an upper triangular matrix using elementary operations.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 10 & 0 & 0 & 1 \end{array}\right)$$

$$\begin{array}{c} \\ -4 \cdot I + II \\ -7 \cdot I + III \end{array} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -6 & -11 & -7 & 0 & 1 \end{array}\right)$$

$$\begin{array}{c} \\ -2 II + III \\ \\ \end{array} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

get those to zero in order to get an inverse

↑ upper triangular matrix

$$\mathrm{II} + 6 \cdot \mathrm{III} \quad \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & 0 & 2 & -11 & 6 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

$$\mathrm{II} \cdot \left(-\frac{1}{3}\right) \quad \left(\begin{array}{ccc|ccc} 1 & \boxed{2 \quad 3} & & 1 & 0 & 0 \\ 0 & 1 & 0 & -2/3 & 11/3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

$$\mathrm{I} - 3\,\mathrm{III} \quad \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & 6 & -3 \\ 0 & 1 & 0 & -2/3 & 11/3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

$$\mathrm{I} - 2\,\mathrm{II} \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 0 & \underbrace{\begin{matrix} -2+\frac{4}{3} & 6-\frac{22}{3} & -3+4 \\ -2/3 & 11/3 & -2 \\ 1 & -2 & 1 \end{matrix}}_{A^{-1}} \\ 0 & 0 & 1 & \end{array}\right)$$

Proposition: Invertibility and identity transformation

If we can use elementary matrix operations $E_1, \dots, E_k$ to bring a matrix A to the identity matrix, then A is invertible and

$$A^{-1} = E := E_k \cdot \dots \cdot E_1$$

Def: Determinant of a matrix

Let $A \in \mathbb{R}^{n \times n}$. Define the determinant of A as

i) if $n = 1$ an A is a scalar

$\det(A) = a$

ii) if $n \in \mathbb{N}$ and $n \geq 2$, with $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot \det(A_{-ij}) \qquad \text{with } i = 1$$

$$A_{-ij} = \begin{pmatrix} a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

Thm: Laplace-Expansion

For any i^*, j^* it holds that

$$\det(A) = \sum_{j=1}^{n} (-1)^{i^*+j} a_{i^*j} \det(A_{-i^*j}) = \sum_{i=1}^{n} (-1)^{i+j^*} a_{ij^*} \det(A_{-ij^*})$$

Determinant of small matrices

i) if $n=2$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow \det(A) = ad - bc$

ii) if $n=3$ and

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

$$\det A = aei + bfg + cdh - ceg - bdi - afh$$

Determinant of triangular matrices:

$$A = \begin{pmatrix} a_{11} & & & a_{14} \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & a_{44} \end{pmatrix}$$

$\det A = \text{trace}(A)$ for upper + lower triangular matrices

trace of a matrix

$$\text{trace}(B) = \prod_{i=1}^{n} b_{ii}$$

From Gauss-Jordan, we know that for a triangular matrix to be invertible we need the diagonal elements to be non-zero

$\Rightarrow \text{trace}(A) \neq 0$

$\Rightarrow \det(A) \neq 0$

$\Rightarrow$ Triangular matrices are invertible if $\det(A) \neq 0$

$\Rightarrow$ This holds more generally: A Matrix A is invertible $\Leftrightarrow \det(A) \neq 0$

Thm: Determinant and Elementary operations

$A \in \mathbb{R}^{n \times n}$ and $\hat{A}$ as the resulting matrix after elementary operation

i) for E_1 (interchange of rows) $\det(\hat{A}) = -\det(A)$

E_2 (multiplication with scalar) $\det(\hat{A}) = \lambda \det(A)$

E_3 (addition of rows) $\det(\hat{A}) = \det A$

$\Rightarrow$ EO do not affect property that $\det \neq 0$"

E_3 (addition of rows) $\det(\tilde{A}) = \det A$

$\Rightarrow$ EO do not affect property that „$\det \neq 0$"

For any square Matrix A

- We can use EO to bring $\tilde{A}$ to $\mathbb{1}$ if A is invertible

  $\det(\mathbb{1}) = 1 \neq 0$

  $\Rightarrow \det(\tilde{A}) \neq 0$ if and only if $\det(A) \neq 0$

  $\boxed{\Rightarrow A \text{ is invertible if and only if } \det(A) \neq 0}$

So far: A is invertible $\Rightarrow \det(A) \neq 0$

Now need to show $\det(A) \neq 0 \Rightarrow A$ is invertible

- Product rule for two square matrices $\det(A \cdot B) = \det(A) \cdot \det(B)$

  $(1) = \det(\mathbb{1}) = \det(A A^{-1}) = \det(A) \cdot \det(A^{-1})$

  $\Rightarrow \det(A) \neq 0$

Some properties

- $\det(A \cdot B) = \det(A) \cdot \det(B)$
- $\det(A^{-1}) = 1/\det(A)$ $\quad$ $(\det(\mathbb{1}) = \det(A) \cdot \det(A^{-1}) = \det(AA^{-1})$
- $A \in \mathbb{R}^{n \times n}, \lambda \in \mathbb{R}$ $\quad \det(\lambda A) = \lambda^n \det(A)$

One criterion to show that a unique solution to $Ax = b$ exists:

$\det(A) \neq 0$

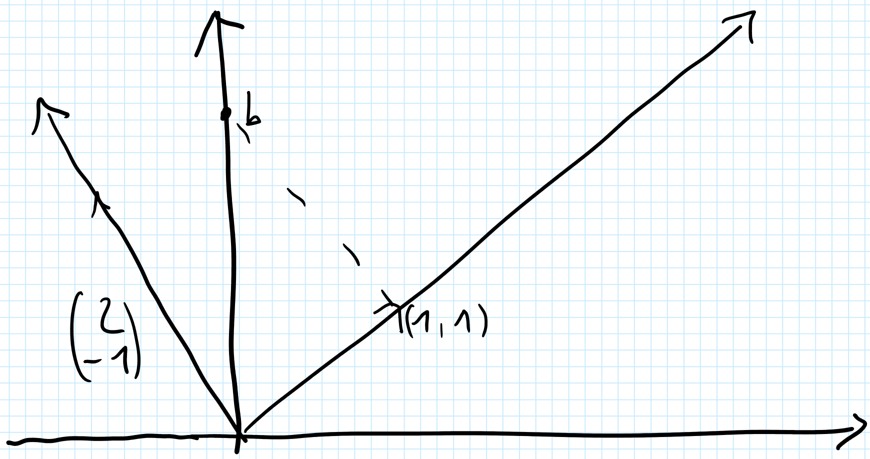
The Rank - Criterion

Column space of a Matrix

$\rightarrow$ the spaces spanned by the column vectors of a matrix

Example $A = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right)$

Column space of A = span$\left\{ \begin{pmatrix}1\\1\end{pmatrix} \begin{pmatrix}2\\-1\end{pmatrix} \right\} = \mathbb{R}^2 = Co(A)$



We can reach b through a linear combination of the two vectors!

Dimension of the column space: the <u>column rank</u>

↳ given by the number of linearly independent columns in A

<u>Row Rank:</u> Number of linearly independent rows in A

<u>Thm:</u> Column rank = row rank =: Rank of the matrix A: $rk(A)$

↳ $rk(A) \leq \min\{m, n\}$ for $A \in \mathbb{R}^{m \times n}$

<u>Full rank:</u> if $m = n = rk(A)$, then A has full rank

<u>Thm:</u> The rank of a matrix is unchanged by the elementary operations $E1$ to E_3

<u>Prop:</u> Elementary operations do not change the set of solutions

consider $Ax = b$, a system of linear equations

Then, for an elementary operation $\tilde{A} = EA$ with operation Matrix E the system $\tilde{A}x = \tilde{b} = Eb$ is equivalent to $Ax = b$.

Thm: Rank Condition

$A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}$. Then $Ax = b$ has a unique solution if and only if $b \in Co(A)$ and $rk(A) = n$

- If $A \in \mathbb{R}^{n \times n}$, we have as many equations as unknowns
  ↳ the system is called just identified
- if $m < n$, there are less equations than unknowns
  ↳ the system is under-identified, cannot have a unique solution
- if $m > n$, there are more equations than unknowns
  ↳ the system is over-identified
    ↳ either some rows are redundant (i.e. not linearly independent from the remaining rows)
    ↳ ~~ore~~ or there is conflicting information (no solution, some rows contradict each other)

  ↳ if there is a unique solution to an over-identified system, we can reduce it to a square system by eliminating redundant rows!

Corollary: Summary on the invertibility of matrices

$A \in \mathbb{R}^{n \times n}$ (Square Matrix). The following are equivalent:

i) A is invertible
ii) $\det(A) \neq 0$
iii) $rk(A) = n$
iv) for any $b \in \mathbb{R}^n$, $Ax = b$ has a unique solution
v) Any triangular matrix $\tilde{A}$ obtained from applying elementary operations to A has only non-zero diagonal entries

diagonal entries

Eigenvalues and Definiteness of a Matrix

(Only defined for Square Matrices!)

Def: Eigenvalue of a matrix

$\lambda \in \mathbb{R}$ is called Eigenvalue of the Matrix $A \in \mathbb{R}^{n \times n}$ if

$$\exists\, x \in \mathbb{R}^n / \{0\} : Ax = \lambda x$$

Def: Eigenspace of Eigenvalue λ

the Eigenspace of an Eigenvalue λ and a matrix A is defined as

$$ES = \text{span}(\{x \in \mathbb{R}^n : Ax = \lambda x\})$$

Finding Eigenvalues

Let $A \in \mathbb{R}^{n \times n}$. we can determine the Eigenvalues of A by solving for the $\lambda \in \mathbb{R}$: $P(\lambda) = \det(A - \lambda \mathbb{1}_n) = 0$

↳ $P(\lambda)$ is called the characteristic polynomial of A

Prop: Eigenvalues and invertibility

Let $A \in \mathbb{R}^{n \times n}$. A is invertible if and only if all eigenvalues of A are non-zero.

Proof: A is invertible $\Leftrightarrow$

$\det(A) \neq 0$

$\det(A) = \det(A - 0 \cdot \mathbb{1}_n) \neq 0$

$\Rightarrow$ 0 is not an eigenvalue of A

$\Rightarrow$ if $\det(A) \neq 0 \Leftrightarrow$ the eigenvalues of A are non-zero

Def: Definiteness of a matrix

A symmetric square matrix $A \in \mathbb{R}^{n \times n}$ is called

i) positive semi-definite if $\forall x \in \mathbb{R}^n : x^t A x \geq 0$

A symmetric square matrix $A \in \mathbb{R}^{n \times n}$ is called

i) positive semi-definite if $\forall x \in \mathbb{R}^n : x'Ax \geq 0$

ii) positive definite if $\forall x \in \mathbb{R}^n : x'Ax > 0$

iii) negative semi-definite if $\forall x \in \mathbb{R}^n : x'Ax \leq 0$

iv) negative definite if $\forall x \in \mathbb{R}^n : x'Ax < 0$

otherwise, A is called indefinite.

Prop: Definiteness and Eigenvalues

A symmetric square matrix $A \in \mathbb{R}^{n \times n}$ is

i) positive (negative) definite $\iff$ all eigen values of A are strictly positive (negative)

ii) positive (negative) semi-definite $\iff$ all eigen values of A are strictly non-negative (non-positive)

Corrolary:

If $A \in \mathbb{R}^{n \times n}$ is symmetric and positive or negative definite, A is invertible.

Thm: Validity of the Gauss-Jordan Algorithm

Suppose $A \in \mathbb{R}^{n \times n}$ is invertible.

Then, we can apply elementary operations $E_1, \dots, E_k$ in ascending order of the index to A to arrive at the identity matrix $\mathbb{1}_n$.

The inverse A^{-1} can be determined as $A^{-1} = E_k \cdot \ldots \cdot E_1$

Special case: Inverse of a 2×2 - Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$