Matrix Theory

Daniel Mao

Contents

1

Fundamentals

1.1 Definitions

Definition (Column Space)**.** *Let A be an $m \times n$ matrix. We define the* ***column space*** *of A, denoted by* $\mathrm{col}(A)$*, to be the set given by*

$$\mathrm{col}(A) := \{Av : v \in \mathbb{R}^n\}.$$

Definition (Row Space)**.** *Let A be an $m \times n$ matrix. We define the* ***row space*** *of A, denoted by* $\mathrm{row}(A)$*, to be the set given by*

$$\mathrm{row}(A) := \{A^\top v : v \in \mathbb{R}^m\}.$$

Definition (Nullspace)**.** *Let A be an $m \times n$ matrix. We define the* ***nullspace*** *of A, denoted by* $\mathrm{null}(A)$*, to be the set given by*

$$\mathrm{null}(A) := \{v \in \mathbb{R}^n : Av = \mathbf{0}\}.$$

Definition (Left Nullspace)**.** *Let A be an $m \times n$ matrix. We define the* ***left nullspace*** *of A, denoted by* $\mathrm{null}(A^\top)$*, to be the set given by*

$$\mathrm{null}(A^\top) := \{v \in \mathbb{R}^m : A^\top v = \mathbf{0}\}.$$

1.2 Main Results

Theorem 1 (The Fundamental Theorem of Linear Algebra)**.** *Let A be an $m \times n$ matrix. Then* $\mathrm{col}(A)^\perp = \mathrm{null}(A^\top)$ *and* $\mathrm{row}(A)^\perp = \mathrm{null}(A)$*.*

2

Rank

2.1 Definitions

Definition (Column Rank)**.** *Let A be a matrix. We define the **column rank** of A to be the dimension of the column space of A. i.e.*

$$\operatorname{colrank}(A) := \dim(\operatorname{col}(A)).$$

Definition (Row Rank)**.** *Let A be a matrix. We define the **row rank** of A to be the dimension of the row space of A. i.e.*

$$\operatorname{rowrank}(A) := \dim(\operatorname{row}(A)).$$

Definition (Rank)**.** *Let A be a matrix. Then the column rank and the row rank are the same. We define the **rank** of A to be this common number.*

Definition (Full Rank)**.** *Let A be an $m \times n$ matrix. We say that A has **full rank** if* $\operatorname{rank}(A) = \min\{m, n\}$.

2.2 Properties

Proposition 2.2.1. *Let A be an $m \times n$ matrix. Then*

- *A is injective if and only if A has full column rank. i.e.* $\operatorname{rank}(A) = n$*, and*
- *A is surjective if and only if A has full row rank. i.e.* $\operatorname{rank}(A) = m$.

Proposition 2.2.2. *Let A and B be matrices with appropriate dimensions. Then*

$$\operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}.$$

Proposition 2.2.3. *Let A, B, and C be matrices with appropriate dimensions. Then*

- *If B has full row rank, then* $\operatorname{rank}(AB) = \operatorname{rank}(A)$*, and*
- *If C has full column rank, then* $\operatorname{rank}(CA) = \operatorname{rank}(A)$.

Proposition 2.2.4 (Subadditivity)**.** *Let A and B be matrices with appropriate dimensions. Then*

$$\operatorname{rank}(A+B) \leq \operatorname{rank}(A) + \operatorname{rank}(B).$$

Proposition 2.2.5. *Let A be a matrix over $\mathbb{C}$. Let A^- denote the complex conjugate of A. Let $A^\top$ denote the transpose of A. Let A^* denote the conjugate transpose of A. Then*

$$\operatorname{rank}(A) = \operatorname{rank}(A^-) = \operatorname{rank}(A^\top) = \operatorname{rank}(A^*) = \operatorname{rank}(AA^*) = \operatorname{rank}(A^*A).$$

3

Matrix Inverse

3.1 Definitions

Definition (Invertible). *Let A be an $n \times n$ matrix over $\mathbb{C}$. We say that A is **invertible** if there exists another $n \times n$ matrix B over $\mathbb{C}$ such that $AB = BA = I_n$.*

Proposition 3.1.1. *Let A be an $n \times n$ invertible matrix over $\mathbb{C}$. Then the $n \times n$ matrix B over $\mathbb{C}$ satisfying $AB = BA = I_n$ is unique.*

Definition (Inverse). *Let A be an $n \times n$ matrix over $\mathbb{C}$. We define the **inverse** of A, denoted by A^{-1}, to be the unique $n \times n$ matrix over $\mathbb{C}$ satisfying $AA^{-1} = A^{-1}A = I_n$.*

Definition (Left/Right Inverse). *Let A be an $m \times n$ matrix over $\mathbb{C}$. We define*

- *the **left inverse** of A, to be an $n \times m$ matrix B over $\mathbb{C}$ such that $BA = I_n$.*
- *the **right inverse** of A, to be an $n \times m$ matrix B over $\mathbb{C}$ such that $AB = I_n$.*

3.2 Characterization

Proposition 3.2.1. *Let A be an $n \times n$ matrix over field K. Then the following statements are equivalent.*

- *A is invertible.*
- $\dim(\text{row}(A)) = n$.
- $\dim(\text{col}(A)) = n$.
- $\dim(\text{null}(A)) = 0$.

Proposition 3.2.2. *Let A be an $n \times n$ matrix over field K. Then the following statements are equivalent.*

- *A is invertible.*
- *A is row-equivalent to I_n.*
- *A is column-equivalent to I_n.*
- *A can be written as a finite product of elementary matrices.*

Proposition 3.2.3. *Let A be an $n \times n$ matrix over field K. Then A is invertible if and only if* $\det(A) \neq 0$.

Proposition 3.2.4. *Let A be an $n \times n$ matrix over field K. Then A is invertible if and only if 0 is not an eigenvalue of A.*

3.3 Arithmetic Properties

Proposition 3.3.1. *Let A be an invertible matrix. Then*

- $(A^{-1})^{-1} = A$.
- $(kA)^{-1} = k^{-1}A^{-1}$.
- $(AB)^{-1} = B^{-1}A^{-1}$.
- $(A^T)^{-1} = (A^{-1})^T$.

3.4 Pseudo-Inverse

Definition (Moore-Penrose Pseudo-Inverse)**.** *Let A be an $n \times d$ matrix. We define the **Moore-Penrose pseudo-inverse** of A, denoted by $A^\dagger$, to be a $d \times n$ matrix G such that*

$$AGA = A, \quad GAG = G, \quad (AG)^\top = AG, \quad (GA)^\top = GA.$$

4

Determinant

4.1 Definitions

Definition (Cofactor). *Let M be an $n \times n$ matrix over field $\mathbb{F}$. We define the $(i,j)^{th}$* ***cofactor*** *of A, denoted by $C_{ij}(A)$, to be a number given by*

$$C_{ij}(A) := (-1)^{i+j} \det(M_{ij})$$

where M_{ij} denotes the submatrix obtained from A by removing the i^{th} row and the j^{th} column.

4.2 Properties

Proposition 4.2.1. *Let A be a matrix. Then*

$$\det(A^T) = \det(A).$$

Proposition 4.2.2. *Let A and B be positive semi-definite matrices with appropriate dimensions. Then*

$$\det(A + B) \geq \det(A) + \det(B).$$

Proposition 4.2.3. *Let A be an $n \times n$ matrix. Let c be some scalar. Then*

$$\det(cA) = c^n \det(A).$$

Proposition 4.2.4. *Let A be an invertible matrix. Then*

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

Proposition 4.2.5. *Let A and B be matrices with appropriate dimensions. Then*

$$\det(AB) = \det(A)\det(B).$$

Proposition 4.2.6. *The determinant operator is a multi-linear operator on the rows/columns.*

4.3 Adjoint of a Matrix

Definition (Adjoint)**.** *Let M be an $n \times n$ matrix. We define the* ***adjoint*** *of M, denoted by* $\operatorname{adj}(M)$*, to be an $n \times n$ matrix given by*

$$(\operatorname{adj}(M))_{ij} = C_{ji}(M),$$

for $i, j = 1, ..., n$.

Proposition 4.3.1. *Let M be an $n \times n$ matrix. Then*

$$M \operatorname{adj}(M) = \operatorname{adj}(M) M = \det(M) I_n.$$

5

Trace

Definition. *Let A be a square matrix. We define the trace of A, denoted by* $\mathrm{tr}(A)$*, to be the sum of the entries on the main diagonal of A.*

5.1 Properties

Proposition 5.1.1. *Trace is a linear operator.*

Proposition 5.1.2. *The trace of the transpose of a matrix equals the trace of the matrix itself. i.e. if M is a square matrix, then*

$$\mathrm{tr}(M) = \mathrm{tr}(M^\top).$$

Proposition 5.1.3. *If $A \in M_{m\times n}$ and $B \in M_{n\times m}$, then*

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Proposition 5.1.4. *Trace is similarity-invariant. i.e., if A is similar to B, then* $\mathrm{tr}(A) = \mathrm{tr}(B)$.

Proposition 5.1.5. *The trace of an idempotent matrix is equal to its rank.*

Proposition 5.1.6. *The trace of a matrix equals the sum of its eigenvalues.*

6

Matrix Norm

Definition. $\|A\| := \sup_{\|x\|=1} \|Ax\|$

6.1 Properties

Proposition 6.1.1. *Let A be an $n \times n$ matrix. Then if A is symmetric, we have*

$$\|A\| = \max\{\lambda_i\}_{i=1}^n$$

where $\lambda_1, ..., \lambda_n$ are the eigenvalues of A.

7

Eigenvalues and Eigenvectors

7.1 Definitions

Definition (Eigenvalue and Eigenvector)**.** *Let A be a matrix. Let x be a vector. Let λ be a scalar. We say that x is an **eigenvector** of A and that λ is an **eigenvalue** of A if $x \neq 0$ and*

$$Ax = \lambda x.$$

7.2 Properties

Proposition 7.2.1. *Let A be an invertible matrix. Let $\{\lambda_i\}_{i=1}^n$ be the eigenvalues of A. Then the eigenvalues of A^{-1} are $\{\lambda_i^{-1}\}_{i=1}^n$.*

Proof.

$$\begin{aligned} & Av = \lambda v \\ \iff & A^{-1}Av = A^{-1}\lambda v \\ \iff & v = \lambda A^{-1} v \\ \iff & A^{-1}v = \lambda^{-1} v. \end{aligned}$$

∎

Proposition 7.2.2. *Let A be an invertible matrix. Let $\{x_i\}_{i=1}^n$ be the eigenvectors of A. Then the eigenvectors of A^{-1} are also $\{x_i\}_{i=1}^n$.*

Proposition 7.2.3. *Let A be a matrix. Let n be a positive integer. Let (x, λ) be an eigenpair of A. Then*

$$A^n x = \lambda^n x.$$

Proof. I will prove by induction on n.
Base Case: $n = 1$.
This is to prove that $Ax = \lambda x$. This holds since (x, λ) is an eigenpair of A.
Inductive Step:
Assume that $A^n x = \lambda^n x$ for some $n \in \mathbb{N}$. We are to prove that $A^{n+1}x = \lambda^{n+1}x$.

$$\begin{aligned} A^{n+1}x &= A^n A x \\ &= A^n \lambda x \\ &= \lambda A^n x \\ &= \lambda \lambda^n x \text{ by the inductive hypothesis} \\ &= \lambda^{n+1} x. \end{aligned}$$

That is,

$$A^{n+1}x = \lambda^{n+1}x.$$

Summary:
By the principle of mathematical induction,

$$\forall n \in \mathbb{N}, \quad A^n x = \lambda^n x.$$

■

Proposition 7.2.4. *If a square matrix is idempotent, then its eigenvalues are either 0 or 1.*

Proof. Since A is idempotent, by definition, $A^2 = A$. Let (x, λ) be an arbitrary eigenpair of A. Then

$$Ax = \lambda x \text{ and } A^2 x = \lambda^2 x.$$

Since $A^2 = A$ and $A^2 x = \lambda^2 x$, we get $Ax = \lambda^2 x$. Since $Ax = \lambda x$ and $Ax = \lambda^2 x$, we get $\lambda x = \lambda^2 x$. Since x is an eigenvector of A, $x \neq 0$. Since $\lambda x = \lambda^2 x$ and $x \neq 0$, we get $\lambda \in \{0, 1\}$. ■

7.3 Eigenspace

Definition (Eigenspace)**.** *Let A be an $m \times n$ matrix over field $\mathbb{F}$. Let λ be an eigenvalue of A. We define the* ***eigenspace*** *of A, associated with λ, denoted by E_λ, to be a set given by*

$$E_\lambda := \{v \in \mathbb{F}^n : Av = \lambda v\}.$$

i.e., E_λ is the set of all eigenvectors of A with eigenvalue λ and the zero vector.

Proposition 7.3.1. *Eigenspaces are linear subspaces.*

8

Singular Values and Singular Vectors

Definition (Singular Value, Singular Vector). *Let M be an $m \times n$ real or complex matrix. We define a* ***singular value*** *for M to be a non-negative real number σ such that there exist unit vectors $u \in \mathbb{F}^m$ and $v \in \mathbb{F}^n$ such that $Mv = \sigma u$ and $M^*u = \sigma v$. We call u the* ***left-singular vector*** *for σ and v the* ***right-singular vector*** *for σ.*

8.1 Singular Value Decomposition

Definition (Singular Value Decomposition). *Let M be an $m \times n$ real or complex matrix. We define a* ***singular value decomposition*** *to be a factorization of the form $M = U\Sigma V^*$ where U is an $m \times m$ unitary matrix, the columns of U are the left-singular vectors of M; V is an $n \times n$ unitary matrix, the columns of V are the right-singular vectors of M; Σ is an $m \times n$ rectangular diagonal matrix, the diagonal entries of Σ are the singular values of M.*

9

Types of Matrices

9.1 Elementary Matrices

Proposition 9.1.1. *The inverse of an elementary matrix can be obtained by multiplying its off-diagonal entries by* -1.

Unconfirmed...

9.2 Definite Symmetric Matrices

9.2.1 Definitions

Definition (Definite Symmetric Matrices)**.** *Let* M *be an* $n \times n$ *Hermitian complex. We say that*

- M *is* ***positive definite****, denoted by* $M \succ 0$*, if*

$$\forall x \in \mathbb{C}^n \setminus \{\mathbf{0}\}, \quad x^* M x > 0.$$

- M *is* ***negative definite****, denoted by* $M \prec 0$*, if*

$$\forall x \in \mathbb{C}^n \setminus \{\mathbf{0}\}, \quad x^* M x < 0.$$

- M *is* ***positive semi-definite****, denoted by* $M \succeq 0$*, if*

$$\forall x \in \mathbb{C}^n \setminus \{\mathbf{0}\}, \quad x^* M x \geq 0.$$

- M *is* ***negative semi-definite****, denoted by* $M \preceq 0$*, if*

$$\forall x \in \mathbb{C}^n \setminus \{\mathbf{0}\}, \quad x^* M x \leq 0.$$

9.2.2 Eigenvalues

Proposition 9.2.1. *Let M be an $n \times n$ Hermitian matrix. Then*

- *M is positive definite if and only if all of its eigenvalues are positive.*
- *M is negative definite if and only if all of its eigenvalues are negative.*
- *M is positive semi-definite if and only if all of its eigenvalues are non-negative.*
- *M is negative semi-definite if and only if all of its eigenvalues are non-positive.*

9.2.3 Sufficient Conditions

Proposition 9.2.2. *If A is positive definite, then A^{-1} exists and is also positive definite.*

Proof Approach 1. Let y be an arbitrary vector. Then there exists some x such that $y = Ax$ since A is invertible. Now

$$y^T A^{-1} y \tag{9.1}$$

$$= x^T A^T A^{-1} A x \tag{9.2}$$

$$= x^T A^T x = \qquad x^T A x > 0. \tag{9.3}$$

Since $\forall y, y^T A^{-1} y > 0$, we get A^{-1} is positive definite. ■

Proof Approach 2. Since A is positive definite, all its eigenvalues are positive. Eigenvalues of A^{-1} are reciprocals of eigenvalues of A. So all eigenvalues of A^{-1} are positive. So A^{-1} is positive definite. ■

9.3 Hermitian Matrix

9.3.1 Definition

Definition (Hermitian Matrix)**.** *We say that a complex square matrix is* ***Hermitian****, or* ***self-adjoint****, if it equals to its complex conjugate.*

9.3.2 Properties

Proposition 9.3.1. *The eigenvalues of a Hermitian matrix are all real.*

Proof Approach 1.
Let A be a Hermitian matrix.
Let (λ, v) be an arbitrary eigenpair of A.
Since (λ, v) is an eigenpair, $Av = \lambda v$.

Since $Av = \lambda v$, $v^*Av = v^*\lambda v = \lambda v^*v$.
Since $(v^*Av)^* = v^*A^*v^{**} = v^*Av$, v^*Av is Hermitian.
Since $(v^*v)^* = v^*v^{**} = v^*v$, v^*v is Hermitian.
Say $v^*Av = [a]$ and $v^*v = [b]$.
Since $v^*Av = \lambda v^*v$ and $v^*Av = [a]$ and $v^*v = [b]$, $a = \lambda b$.
Since v^*Av is Hermitian, $a = \overline{a}$.
Since $a = \overline{a}$, a is real.
Since v^*v is Hermitian, $b = \overline{b}$.
Since $b = \overline{b}$, b is real.
Since $a = \lambda b$ and both a and b are real, λ is real.

■

Proof Approach 2.

$$
\begin{aligned}
&\lambda\langle v, v\rangle \\
&= \langle \lambda v, v\rangle \\
&= \langle Av, v\rangle \\
&= \langle v, A^*v\rangle \\
&= \langle v, Av\rangle \\
&= \langle v, \lambda v\rangle \\
&= \overline{\lambda}\langle v, v\rangle.
\end{aligned}
$$

That is, $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$. Since v is an eigenvector, $v \neq \vec{0}$. Since $v \neq \vec{0}$, $\langle v, v\rangle \neq 0$. Since $\langle v, v\rangle \neq 0$ and $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$, $\lambda = \overline{\lambda}$. Since $\lambda = \overline{\lambda}$, λ is real.

■

9.4 Triangular Matrix

Definition (Upper Triangular Matrix)**.**

Definition (Lower Triangular Matrix)**.**

9.4.1 Properties

Proposition 9.4.1. *The product of two upper triangular matrices is also upper triangular. i.e. if U_1 and U_2 are upper triangular matrices with appropriate dimensions, then $U := U_1U_2$ is also upper triangular.*

Proposition 9.4.2. *The inverse of an upper triangular matrix is also upper triangular, if it exists. i.e. if U is an invertible upper triangular matrix, then U^{-1} is also upper triangular.*

9.5 Unitary Matrices

9.5.1 Definition

Definition (Unitary)**.** *Let M be a complex square matrix. We say that M is* ***unitary*** *if*

$$M^*M = I$$

where M^* is the complex conjugate of M and I is the identity matrix.

9.5.2 Sufficient Conditions

Proposition 9.5.1. *The product of two unitary matrices is still unitary.*

10

Matrix Diagonalization

10.1 Unitary Diagonalization

10.1.1 Definitions

Definition (Unitarily Similar). *Let A and B be complex square matrices of the same dimension. We say that A and B are **unitarily similar** if there exists a unitary matrix U such that*

$$U^*AU = B.$$

Theorem 2 (Schur). *Any matrix is unitarily similar to an upper triangular matrix.*

Definition (Unitarily Diagonalizable). *Let M be a complex square matrix. We say that M is **unitarily diagonalizable** if M is unitarily similar to a diagonal matrix.*

Definition (Normal). *Let M be a complex square matrix. We say that M is **normal** if*

$$M^*M = MM^*M.$$

10.1.2 Properties

Proposition 10.1.1. *Unitarily diagonalizable matrices are normal.*

10.2 Sufficient Conditions

Proposition 10.2.1. *Hermitian matrices are unitarily diagonalizable.*

Proposition 10.2.2. *Normal matrices are unitarily diagonalizable.*

11

Matrix Decomposition

11.1 LU Decomposition

Theorem 3. *Let A be an $n \times n$ matrix with $\det(A) \neq 0$. Then there exists a permutation matrix P, a lower triangular matrix L, and an upper triangular matrix U.*

11.2 Eigenvalue Decomposition

Definition (Eigenvalue Decomposition)**.** *Let A be an $n \times n$ matrix where $n \in \mathbb{N}$. Let $\{(x_i, \lambda_i)\}_{i=1}^n$ be the eigenpairs of A. We define the* ***eigenvalue decomposition*** *of A to be a factorization of A given by*

$$A = Q\Lambda Q^{-1}$$

where $Q = \begin{bmatrix} q_1 & \dots & q_n \end{bmatrix}$ and $\Lambda = \operatorname{diag}(\{\lambda_i\}_{i=1}^n)$.

Proposition 11.2.1. *Let A be an $n \times n$ matrix. Then A can be eigendecomposed if and only if A has n linearly independent eigenvectors.*