Matrix Theory

Daniel Mao

Contents

Chapter 1

Fundamentals

1.1 Definitions

DEFINITION 1.1 (Column Space)**.** Let A be an $m \times n$ matrix. We define the **column space** of A, denoted by $\operatorname{col}(A)$, to be the set given by

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

DEFINITION 1.2 (Row Space)**.** Let A be an $m \times n$ matrix. We define the **row space** of A, denoted by $\operatorname{row}(A)$, to be the set given by

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

DEFINITION 1.3 (Nullspace)**.** Let A be an $m \times n$ matrix. We define the **nullspace** of A, denoted by $\operatorname{null}(A)$, to be the set given by

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

DEFINITION 1.4 (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 1.5 (The Fundamental Theorem of Linear Algebra)**.** Let A be an $m \times n$ matrix. Then $\mathrm{col}(A)^\perp = \mathrm{null}(A^\top)$ and $\mathrm{row}(A)^\perp = \mathrm{null}(A)$.

Chapter 2

Matrix Inverse

2.1 Definitions

DEFINITION 2.1 (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 2.2. 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 2.3 (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 2.4 (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 $AB = I_n$.

2.2 Characterization

PROPOSITION 2.5. 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 2.6. 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 2.7. Let A be an $n \times n$ matrix over field K. Then A is invertible if and only if $\det(A) \neq 0$.

PROPOSITION 2.8. 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.

2.3 Arithmetic Properties

PROPOSITION 2.9. Let A be an invertible matrix. Then

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

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

2.4 Pseudo-Inverse

DEFINITION 2.10 (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

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

Chapter 3

Rank

3.1 Definitions

DEFINITION 3.1 (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 3.2 (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 3.3 (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 3.4 (Full Rank)**.** Let A be an $m \times n$ matrix. We say that A has **full rank** if $\operatorname{rank}(A) = \min\{m, n\}$.

3.2 Properties

PROPOSITION 3.5. 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 3.6. Let A and B be matrices with appropriate dimensions. Then

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

PROPOSITION 3.7. 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 3.8 (Subadditivity)**.** Let A and B be matrices with appropriate dimensions. Then

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

PROPOSITION 3.9. 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

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

Chapter 4

Determinant

4.1 Definitions

DEFINITION 4.1 (Cofactor). Let M be an $n \times n$ matrix where $n \geq 2$. We define the $(i,j)^{\text{th}}$ **cofactor** of M, denoted by $C_{i,j}(M)$, to be a number given by

$$C_{i,j}(M) := (-1)^{i+j} \det(M(i,j))$$

where $M(i,j)$ denotes the submatrix obtained from M by removing the i^{th} row and the j^{th} column.

DEFINITION 4.2 (Determinant). Let M be an $n \times n$ matrix where $n \geq 2$. We define the **determinant** of M, denoted by $\det(M)$, to be a number given by

$$\det(M) := \sum_{i=1}^{n} [M]_{i,j} C_{i,j}(M)$$

where j can be anything in $\{1, ..., n\}$, $[M]_{i,j}$ denotes the $(i,j)^{\text{th}}$ entry of M, and $C_{i,j}(M)$ denotes the $(i,j)^{\text{th}}$ cofactor of M. Equivalently,

$$\det(M) := \sum_{j=1}^{n} [M]_{i,j} C_{i,j}(M)$$

where i can be anything in $\{1, ..., n\}$, $[M]_{i,j}$ denotes the $(i,j)^{\text{th}}$ entry of M, and $C_{i,j}(M)$ denotes the $(i,j)^{\text{th}}$ cofactor of M.

We define the determinant of an 1×1 matrix to be the number itself.

4.2 Properties

PROPOSITION 4.3. Let A be a matrix. Then

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

PROPOSITION 4.4. Let A and B be positive semi-definite matrices with appropriate dimensions. Then

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

PROPOSITION 4.5. Let A be an $n \times n$ matrix. Let c be some scalar. Then

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

PROPOSITION 4.6. Let A be an invertible matrix. Then

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

PROPOSITION 4.7. Let A and B be matrices with appropriate dimensions. Then

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

PROPOSITION 4.8. The determinant operator is a multi-linear operator on the rows/columns.

4.3 Adjoint of a Matrix

DEFINITION 4.9 (Adjoint)**.** Let M be an $n \times n$ matrix. We define the **adjoint** of

M, denoted by $\text{adj}(M)$, to be an $n \times n$ matrix given by

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

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

PROPOSITION 4.10. Let M be an $n \times n$ matrix. Then

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

Chapter 5

Trace

DEFINITION 5.1. 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 Basic Properties

PROPOSITION 5.2. Trace is a linear operator.

PROPOSITION 5.3. The trace of an idempotent matrix is equal to its rank.

PROPOSITION 5.4. The trace of a matrix equals the sum of its eigenvalues.

5.2 Invariant Properties

PROPOSITION 5.5 (Transpose Invariant)**.** Let $M \in \mathbb{C}^{n\times n}$. Then we have

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

PROPOSITION 5.6 (Cyclical Permutation Invariant)**.** Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$. Then we have

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

PROPOSITION 5.7 (Similarity Invariant)**.** If A is similar to B, then $\operatorname{tr}(A) = \operatorname{tr}(B)$.

Chapter 6

Eigenvalues and Eigenvectors

6.1 Definitions

DEFINITION 6.1 (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

$$Ax = \lambda x.$$

6.2 Properties

PROPOSITION 6.2. 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 6.3. 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 6.4. 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 $A^{n+1}x = \lambda^{n+1}x$.

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

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 6.5. 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 $A^2x = \lambda^2 x$, 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 $\lambda \in \{0, 1\}$. □

6.3 Eigenspace

DEFINITION 6.6 (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 6.7. Eigenspaces are linear subspaces.

Chapter 7

Singular Values and Singular Vectors

7.1 Definitions

DEFINITION 7.1 (Singular Value, Singular Vector)**.** Let $A \in \mathbb{F}^{m \times n}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. We define a **singular value** for A 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 $Av = \sigma u$ and $A^* u = \sigma v$. We call u the **left-singular vector** for σ and v the **right-singular vector** for σ.

7.2 Properties

PROPOSITION 7.2. Let $A \in \mathbb{F}^{m \times n}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Then $\forall i \in [\min\{m, n\}]$, we have

$$\sigma_i(A) = \sigma_i(A^\top) = \sigma_i(A^-) = \sigma_i(A^*)$$

where $A^\top$ denotes the transpose of A, A^- denotes the complex conjugate of A, and A^* denote the conjugate transpose of A.

PROPOSITION 7.3. Let $A \in \mathbb{F}^{m \times n}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Let $U \in \mathbb{F}^{m \times m}$ and

$V \in \mathbb{F}^{n\times n}$ be unitary. Then $\forall i \in [\min\{m,n\}]$, we have

$$\sigma_i(A) = \sigma_i(UAV).$$

PROPOSITION 7.4. Let $A \in \mathbb{F}^{m\times n}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Then $\forall i \in [\min\{m,n\}]$, we have

$$\sigma_i^2(A) = \lambda_i(AA^*) = \lambda_i(A^*A).$$

PROPOSITION 7.5 (Singular Value of Sum of Matrices)**.** Let $A, B \in \mathbb{F}^{m\times n}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Then $\forall i, j \in \mathbb{Z}_{++}$ and $i + j - 1 \leq \min\{m,n\}$, we have

$$\sigma_{i+j-1}(A + B) \leq \sigma_i(A) + \sigma_j(B).$$

PROPOSITION 7.6 (Singular Value of Sum of Matrices)**.** Let $A, B \in \mathbb{F}^{m\times n}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Then we have

$$\sum_{i=1}^{k} \sigma_i(A + B) \leq \sum_{i=1}^{k} (\sigma_i(A) + \sigma_i(B))$$

where $k := \min\{m,n\}$.

For more see `https://en.wikipedia.org/wiki/Singular_value`.

Chapter 8

Orthogonal and Unitary Matrices

8.1 Definitions

DEFINITION 8.1 (Orthogonal)**.** Let $U \in \mathbb{R}^{n\times n}$. We say that U is **orthogonal** if and only if

$$UU^\top = U^\top U = I$$

where $U^\top$ denotes the transpose of U and I denotes the $n \times n$ identity matrix. i.e., the transpose equals the inverse.

DEFINITION 8.2 (Unitary - 1)**.** Let $U \in \mathbb{C}^{n\times n}$. We say that U is **unitary** if and only if

$$UU^* = U^* U = I$$

where U^* denotes the conjugate transpose of U and I denotes the $n \times n$ identity matrix. i.e., the conjugate transpose equals the inverse.

DEFINITION 8.3 (Unitary - 2)**.** Let $U \in \mathbb{C}^{n\times n}$. We say that U is **unitary** if and only if the <u>columns</u> of U form an orthonormal basis for $\mathbb{C}^n$, or equivalently, the <u>rows</u> of U form an orthonormal basis for $\mathbb{C}^n$.

8.2 Stability of Unitary Matrices

PROPOSITION 8.4. The product of two unitary matrices is still unitary.

8.3 Properties of Unitary Matrices

PROPOSITION 8.5 (Unitary Matrices Preserve Inner Products)**.** Let $U \in \mathbb{C}^{n\times n}$. Then U is unitary if and only if

$$\forall x, y \in \mathbb{C}^n, \quad \langle Ux, Uy \rangle = \langle x, y \rangle .$$

PROPOSITION 8.6 (Eigenvalues)**.** The eigenvalues of a unitary matrix are all unimodular.

Proof. Let U be a unitary matrix. Let (λ, v) be an arbitrary eigenpair of U. Since U is a unitary matrix, we get

$$\langle Uv, Uv \rangle = \langle v, v \rangle .$$

Since (λ, v) is an eigenpair of U, we get

$$\langle Uv, Uv \rangle = \langle \lambda v, \lambda v \rangle = \lambda^2 \langle v, v \rangle .$$

So $\langle v, v \rangle = \lambda^2 \langle v, v \rangle$. Since v is an eigenvector, $v \neq 0$ and hence $\langle v, v \rangle \neq 0$. So $\lambda^2 = 1$.

$\square$

Chapter 9

Definite Matrices

9.1 Definitions

DEFINITION 9.1 (Definite Matrices)**.** Let $M \in \mathbb{C}^{n\times n}$ be Hermitian.

- We say that M is **positive semidefinite**, denoted by $M \succeq 0$, if

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

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

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

- We say that M is **negative semidefinite**, denoted by $M \preceq 0$, if

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

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

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

where x^* denotes the conjugate transpose of x.

PROPOSITION 9.2 (Characterization by Eigenvalues)**.** Let $M \in \mathbb{C}^{n\times n}$ be Hermitian. Then

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

Proof of (2). **Forward Direction**: Assume that M is positive definite. I will show that the eigenvalues of M are all positive. Let (λ, x) be an arbitrary eigenpair of M. Then we have $Mx = \lambda x$. Since M is positive definite, we have $x^* M x > 0$. So $x^* \lambda x = \lambda x^* x > 0$. Note that $x^* x \geq 0$. So $\lambda > 0$.

Backward Direction: □

PROPOSITION 9.3 (Equivalent Formulations of PSD Matrices)**.** Let $X \in \mathbb{S}^n$. Then the following statements are equivalent.

1. $X \in \mathbb{S}^n_+$.
2. $\forall j \in \{1, ..., n\}$, $\lambda_j(X) \geq 0$ where $\lambda_j(X)$ denotes the j^{th} eigenvalue of X.
3. $\exists \mu \in \mathbb{R}^n_+$ and $h^{(1)}, h^{(2)}, ..., h^{(n)} \in \mathbb{R}^n$ such that
$$X = \sum_{i=1}^{n} \mu_i h^{(i)} h^{(i)\top}.$$
4. $\exists B \in \mathbb{R}^{n \times n}$ such that $X = BB^\top$.
5. $\forall J \subseteq \{1, 2, ..., n\} : J \neq \emptyset$, $\det(X_J) \geq 0$ where X_J denotes the symmetric minor of X defined by J.
6. $\forall Y \in \mathbb{S}^n_+$, $\operatorname{tr}(XY) \geq 0$.

PROPOSITION 9.4 (Equivalent Formulations of PD Matrices)**.** Let $X \in \mathbb{S}^n$. Then the following statements are equivalent.

1. $X \in \mathbb{S}^n_{++}$.
2. $\forall j \in \{1, ..., n\}$, $\lambda_j(X) > 0$ where $\lambda_j(X)$ denotes the j^{th} eigenvalue of X.

3. $\exists \mu \in \mathbb{R}^n_{++}$ and $h^{(1)}, ..., h^{(n)} \in \mathbb{R}^n$ linearly independent such that

$$X = \sum_{i=1}^{n} \mu_i h^{(i)} h^{(i)\top}.$$

4. $\exists B \in \mathbb{R}^{n \times n}$ non-singular such that $X = BB^\top$.
5. $\forall k \in \{1, ..., n\}$, $\det(X_{J_k}) > 0$ where $J_k := \{1, ..., k\}$ and X_{J_k} denotes the leading principle minor of X defined by J_k.
6. $\forall Y \in \mathbb{S}^n_+ \setminus \{0\}$, $\operatorname{tr}(XY) > 0$.
7. $X \in \mathbb{S}^n_+$ and $\operatorname{rank}(X) = n$.

PROPOSITION 9.5. Let $M \in \mathbb{S}^n$. Then the following statements are equivalent:

1. $\forall x \in \mathbb{R}^n \setminus \{0\}$, $x^\top M x > 0$;
2. $\exists \alpha > 0$ such that $\forall x \in \mathbb{R}^n$, $x^\top M x \geq \alpha x^\top x$.

Proof. **Forward Inclusion**: Assume that $\forall x \in \mathbb{R}^n \setminus \{0\}$, $x^\top M x > 0$. Let $S := \{x \in \mathbb{R}^n : \|x\| = 1\}$. Let $f(x) := x^\top M x$. Notice $S \subseteq \mathbb{R}^n$ is nonempty and compact and $f : \mathbb{R}^n \to \mathbb{R}$ is continuous. So $f(S) \subseteq \mathbb{R}$ is compact and hence $\min(f(S))$ exists. Let $\alpha := \min(f(S))$ and $x_0 \in \mathbb{R}^n$ be such that $\|x_0\| = 1$ and $f(x_0) = \alpha$. Then $\forall x \in \mathbb{R}^n \setminus \{0\}$, we have

$$\left(\frac{x}{\|x\|}\right)^\top M \left(\frac{x}{\|x\|}\right) \geq \alpha.$$

It follows that $x^\top M x \geq \alpha \|x\|^2 = \alpha x^\top x$. For $x = 0$, it is clear that $0^\top M 0 \geq \alpha 0^\top 0$.

Backward Inclusion: Assume that $\exists \alpha > 0$ such that $\forall x \in \mathbb{R}^n$, $x^\top M x \geq \alpha x^\top x$. Now consider an arbitrary $x \in \mathbb{R}^n \setminus \{0\}$. Since $x \neq 0$, $x^\top x \neq 0$. So $x^\top M x \geq \alpha x^\top x > 0$. $\square$

9.2 Properties

PROPOSITION 9.6. 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

$$y^T A^{-1} y$$

$$
\begin{aligned}
&= x^T A^\top A^{-1} A x \\
&= x^T A^\top x \\
&= x^T A x > 0.
\end{aligned}
$$

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. ☐

PROPOSITION 9.7. Let $A, B \in \mathbb{S}^n_+$. Then $\operatorname{tr}(AB) \leq \operatorname{tr}(A)\operatorname{tr}(B)$.

Proof Approach 1.

$$\operatorname{tr}(AB) = \operatorname{tr}(B^{1/2} A B^{1/2}) \leq \operatorname{tr}(B^{1/2}(\operatorname{tr}(A) I) B^{1/2}) = \operatorname{tr}(A)\operatorname{tr}(B^{1/2} B^{1/2}) = \operatorname{tr}(A)\operatorname{tr}(B).$$

☐

Proof Approach 2.

$$\operatorname{tr}(AB) \leq \sum_{i \in [n]} \lambda_i(A)\lambda_i(B) \leq \left(\sum_{i \in [n]} \lambda_i(A)\right)\left(\sum_{i \in [n]} \lambda_i(B)\right) = \operatorname{tr}(A)\operatorname{tr}(B).$$

☐

Proof Approach 3. Let (e_i) be an orthonormal basis of eigenvectors of B and let (λ_i) be the corresponding eigenvalues. Then

$$\operatorname{tr}(AB) = \sum_{i \in [n]} e_i^\top A B e_i = \sum_{i \in [n]} \lambda_i e_i^\top A e_i \leq \lambda_1 \sum_{i \in [n]} e_i^\top A e_i = \lambda_1 \operatorname{tr}(A) \leq \operatorname{tr}(A)\operatorname{tr}(B).$$

☐

9.3 Ordering of Symmetric Matrices

PROPOSITION 9.8. Let $A, B \in \mathbb{S}^n_+$. Let $U \in \mathbb{R}^{n \times n}$ be orthonormal. Then $A \succeq B$ if and only if $UAU^\top \succeq UBU^\top$.

Proof. **Forward Direction**: Assume that $A \succeq B$. Then $(A - B) \succeq 0$. Let $h \in \mathbb{R}^n$ be arbitrary. Then

$$h^\top (UAU^\top - UBU^\top)h = (h^\top U)(A - B)(U^\top h) \geq 0.$$

So $UAU^\top \succeq UBU^\top$.

Backward Direction: Assume that $UAU^\top \succeq UBU^\top$. Then using the forward direction, we get $U^\top(UAU^\top)U \succeq U^\top(UBU^\top)U$. Since $UU^\top = U^\top U = I$, the above is equivalent to, $A \succeq B$. $\square$

PROPOSITION 9.9. Let $A, B \in \mathbb{S}^n_{++}$. Then if $A \succeq B$, we have $A^{-1} \preceq B^{-1}$.

PROPOSITION 9.10. Let $A, B \in \mathbb{S}^n_{+}$. If $A \succeq B$, then

- $\lambda(A) \geq \lambda(B)$;
- $\mathrm{tr}(A) \geq \mathrm{tr}(B)$;
- $\det(A) \geq \det(B)$.

Chapter 10

Special Types of Matrices

10.1 Elementary Matrices

PROPOSITION 10.1. The inverse of an elementary matrix can be obtained by multiplying its off-diagonal entries by -1.

Unconfirmed...

10.2 Triangular Matrix

PROPOSITION 10.2. 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 $U := U_1 U_2$ is also upper triangular.

PROPOSITION 10.3. 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.

10.3 Symmetric and Hermitian Matrices

10.3.1 Definitions

DEFINITION 10.4 (Symmetric Matrix)**.** Let $M \in \mathcal{M}_{n\times n}(\mathbb{R})$ (a real square matrix). We say that M is **symmetric**, denoted by $M \in \mathbb{S}^n$, if and only if $M = M^\top$, where $M^\top$ denotes the transpose of M.

DEFINITION 10.5 (Hermitian Matrix)**.** Let $M \in \mathbb{C}^{n\times n}$. We say that M is **Hermitian**, or **self-adjoint**, denoted by $M \in \mathbb{H}^n$ if and only if $M = M^*$, where M^* denotes the conjugate transpose of M.

PROPOSITION 10.6 (Equivalent Conditions of Hermitian)**.** Let $M \in \mathbb{C}^{n\times n}$. Then the following statements are equivalent:

1. $M = M^*$.
2. $\forall x, y \in \mathbb{C}^n$, $\langle x, My\rangle = \langle Mx, y\rangle$.
3. $\forall x \in \mathbb{C}^n$, $\langle x, Mx\rangle \in \mathbb{R}$.

10.3.2 Stability of Hermitian Matrices

PROPOSITION 10.7 (Sum of Two Hermitian Matrices)**.** Let A and B be Hermitian matrices. Then $A + B$ is also Hermitian.

PROPOSITION 10.8 (Associative Product)**.** Let A and B be Hermitian matrices. Suppose that $AB = BA$. Then AB is also Hermitian.

PROPOSITION 10.9 (Inverse of a Hermitian Matrix)**.** Let M be a Hermitian matrix. Suppose that M is invertible. Then M^{-1} is also Hermitian.

10.3.3 Properties of Hermitian Matrices

PROPOSITION 10.10. Hermitian matrices are normal.

PROPOSITION 10.11. The determinant of a Hermitian matrix is real.

Proof. Let M be a Hermitian matrix. Then

$$\det(M) = \det(M^*) = \det(\overline{M}^\top) = \det(\overline{M}) = \overline{\det(M)}.$$

That is, $\det(M) = \overline{\det(M)}$. So $\det(M) \in \mathbb{R}$. $\square$

PROPOSITION 10.12 (Eigenvalues)**.** 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. Then we have $Av = \lambda v$ and hence

$$v^*Av = v^*\lambda v = \lambda v^*v. \tag{1}$$

Note that v^*Av has size 1×1. So $v^*Av = [a]$ for some $a \in \mathbb{C}$.

$$\begin{aligned} &(v^*Av)^* = v^*A^*v^{**} = v^*Av \\ &\implies v^*Av \text{ is Hermitian} \iff [a] \text{ is Hermitian} \\ &\implies a = \bar{a} \implies a \in \mathbb{R}. \end{aligned}$$

That is,

$$v^*Av = a \in \mathbb{R}. \tag{2}$$

Note that v^*v has size 1×1. So $v^*v = [b]$ for some $b \in \mathbb{C}$.

$$\begin{aligned} &(v^*v)^* = v^*v^{**} = v^*v \\ &\implies v^*v \text{ is Hermitian} \iff [b] \text{ is Hermitian} \\ &\implies b = \bar{b} \implies b \in \mathbb{R}. \end{aligned}$$

That is,

$$v^*v = b \in \mathbb{R}. \tag{3}$$

From (1), (2), and (3), we get $a = \lambda b$. It follows that $\lambda \in \mathbb{R}$.

$\square$

Proof Approach 2. Let A be a Hermitian matrix. Let (λ, v) be an arbitrary eigenpair of A.

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

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.

□

LEMMA 10.13. Let $M \in \mathbb{C}^{n\times n}$ be Hermitian. Then

$$\forall x \in \mathbb{C}^n, \quad x^* M x \in \mathbb{R}.$$

PROPOSITION 10.14. The eigenvectors of a Hermitian matrix are orthogonal.

10.4 Normal Matrices

10.4.1 Definitions

DEFINITION 10.15 (Normal Matrix - 1)**.** Let $M \in \mathbb{C}^{n\times n}$. We say that M is **normal** if

$$MM^* = M^* M,$$

where M^* denotes the conjugate transpose of M.

DEFINITION 10.16 (Normal Matrix - 2)**.** Let $M \in \mathbb{C}^{n\times n}$. We say that M is **normal** if $\exists \mathcal{B} \subseteq \mathcal{E}(M)$ such that $\mathcal{B}$ is a orthonormal basis for $\mathbb{C}^n$ where $\mathcal{E}(M)$ denotes the set of eigenvectors of M.

PROPOSITION 10.17. Definitions (1) and (2) of normal matrices are equivalent.

Proof. Let $M \in \mathbb{C}^{n \times n}$.

Forward Direction Assume that $MM^* = M^*M$. I will show that M has an orthonormal basis of eigenvectors.

□

DEFINITION 10.18 (Normal Matrix - 3)**.** Let $M \in \mathbb{C}^{n \times n}$. We say that M is **normal** if M is diagonalizable by a unitary matrix.

10.4.2 Stability of Normal Matrices

PROPOSITION 10.19. Let A and B be normal matrices. Suppose that $AB = BA$. Then

1. $A + B$ is also normal.
2. AB is also normal.

10.4.3 Properties of Normal Matrices

PROPOSITION 10.20. Let M be a normal matrix. Then if M is triangular, M is diagonal.

PROPOSITION 10.21. Let M be a normal matrix. Then M is Hermitian if and only if $\sigma(M) \subseteq \mathbb{R}$ where $\sigma(M)$ denotes the set of eigenvalues of M.

Proof. **Forward Direction** Assume that M is Hermitian. I will show hat $\sigma(M) \subseteq \mathbb{R}$. Since M is Hermitian, we get $\sigma(M) \subseteq \mathbb{R}$.

Backward Direction Assume that $\sigma(M) \subseteq \mathbb{R}$. I will show that M is Hermitian. Since M is normal, it is diagonalizable by a unitary matrix. Say $M = U^*DU$ where U is unitary and D is diagonal. Then the diagonal entries of D are the eigenvalues of M and hence are real. So $D^* = D$. Then

$$M^* = (U^*DU)^* = U^*D^*U^{**} = U^*D^*U = U^*DU = M.$$

So M is Hermitian.

$\square$

PROPOSITION 10.22. Let M be a normal matrix. Then M is unitary if and only if $\sigma(M) \subseteq \mathbb{T}$ where $\sigma(M)$ denotes the set of eigenvalues of M and $\mathbb{T}$ denotes the unit circle of the complex plane.

Chapter 11

Matrix Norm

11.1 Operator Norm

DEFINITION 11.1 (Operator Norm)**.** Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be a field. Let $A \in \mathbb{K}^{m\times n}$. Let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ denote the vector norms on $\mathbb{K}^n$ and $\mathbb{K}^m$, respectively. We define the **operator norm** of A, denoted by $\|A\|_{\alpha,\beta}$, to be the number in $\mathbb{R}_+$ given by

$$\|A\|_{\alpha,\beta} := \sup\left\{\frac{\|Ax\|_\beta}{\|x\|_\alpha} : x \in \mathbb{K}^n \setminus \{0\}\right\}.$$

In the case $\alpha = \beta$, we simply denote the operator norm of A by $\|A\|_\alpha$.

Operator norms defined by the 1-norms, 2-norms, and ∞-norms on the spaces are of particular importance.

DEFINITION 11.2 (Spectral Norm)**.** Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be a field. Let $A \in \mathbb{K}^{m\times n}$. We define the **spectral norm** of A to be $\|A\|_2$.

PROPOSITION 11.3. Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be a field. Let $A \in \mathbb{K}^{m\times n}$. Then the following statements hold:

1. $\|A\|_1 = \max\limits_{j\in\{1,\ldots,n\}} \sum\limits_{i=1}^{m} |A_{ij}|$.

2. $\|A\|_2 = \sigma_{\max}(A) = \lambda_{\max}((A^*A)^{1/2})$;

3. $\|A\|_\infty = \max_{i \in \{1,\ldots,m\}} \sum_{j=1}^{n} |A_{ij}|$.

PROPOSITION 11.4. Let $A \in \mathbb{S}^n$. Then

$$\|A\|_2 = \max\{|\lambda_1(A)|, \ldots, |\lambda_n(A)|\}$$

where $\lambda(A)$ denotes the ordered vector of eigenvalues of A.

11.2 Frobenius Norm

DEFINITION 11.5 (Frobenius Norm)**.** Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be a field. Let $A \in \mathbb{K}^{m \times n}$. We define the **Frobenius norm** of A, denoted by $\|A\|_F$, to be the number in $\mathbb{R}_+$ given by

$$\|A\|_F := \sqrt{\operatorname{tr}(AA^*)}$$

where A^* denotes the complex conjugate of A.

PROPOSITION 11.6. Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be a field. Let $A \in \mathbb{K}^{m \times n}$. Then

$$\|A\|_F = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2(A)}$$

where $\sigma(A)$ denotes the vector of singular values of A.

PROPOSITION 11.7. Let $A \in \mathbb{S}^n$. Then

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \lambda_i^2(A)}$$

where $\lambda(A)$ denotes the eigenvalues of A.

11.3 Nuclear Norm

DEFINITION 11.8 (Nuclear Norm)**.** Let $A \in \mathbb{F}^{m \times n}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. We define the **nuclear norm** of A, denoted by $\|A\|_*$, to be the number given by $\|A\|_* := \displaystyle\sum_{i=1}^{k} \sigma_i(A)$ where $k := \min\{m, n\}$.

Chapter 12

Matrix Diagonalization

12.1 Diagonalization in General

DEFINITION 12.1 (Diagonalizable Matrix)**.** Let $M \in \mathcal{M}_{n\times n}(\mathbb{C})$. We say that M is **diagonalizable** if and only if $P^{-1}MP = D$ for some invertible matrix $P \in \mathcal{M}_{n\times n}(\mathcal{C})$ and some diagonal matrix $D \in \mathcal{M}_{n\times n}(\mathcal{C})$.

PROPOSITION 12.2. Let $M \in \mathcal{M}_{n\times n}(\mathbb{C})$. Then M is diagonalizable if and only if $\exists$ eigenpairs $((\lambda_i, v_i))_{i=1}^{n}$ of M such that the matrix $P = [v_1, ..., v_n]$ is invertible. In this case, we have

$$P^{-1}MP = \operatorname{diag}(\lambda_1, ..., \lambda_n).$$

12.2 Unitary Diagonalization

12.2.1 Definitions

DEFINITION 12.3 (Unitarily Similar)**.** Let $A, B \in \mathcal{M}_{n\times n}(\mathbb{C})$. We say that A and B are **unitarily similar** if there exists a unitary matrix U such that

$$U^*AU = B.$$

THEOREM 12.4 (Schur)**.** Any matrix is unitarily similar to an upper triangular matrix.

DEFINITION 12.5 (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.

12.2.2 Properties

PROPOSITION 12.6. Unitarily diagonalizable matrices are normal.

12.3 Sufficient Conditions

PROPOSITION 12.7. Hermitian matrices are unitarily diagonalizable.

PROPOSITION 12.8. Normal matrices are unitarily diagonalizable.

Chapter 13

Matrix Decomposition

13.1 Lower-Upper Decomposition

DEFINITION 13.1 (Lower-Upper (LU) Decomposition)**.** Let A be some square matrix. In the following let L denote lower triangular matrices, U denote upper triangular matrices, P denote permutation matrices, and D denote diagonal matrices. We define the followings:

- **LU decomposition**:
$$A = LU.$$
- **LUP decomposition**:
$$A = LUP.$$
- **PLU decomposition**:
$$A = PLU.$$
- **LDU decomposition**:
$$A = LDU$$
where L and U are required to be unitriangular.

THEOREM 13.2 (Lower-Upper (LU) Decomposition)**.**

- All square matrices admit LUP and PLU decompositions.

LU decomposition can be viewed as the matrix form of Gaussian elimination.

13.2 Cholesky Decomposition

DEFINITION 13.3 (Cholesky Decomposition)**.** Let A be some square matrix. In the following let L denote real lower triangular matrices and D denote diagonal matrices. We define the followings:

- **Cholesky decomposition**:
$$A = LL^*.$$
- **Square-Root-Free Cholesky (LDL) decomposition**:
$$A = LDL$$
where L is required to be unitriangular.

The diagonal elements of L are required to be 1 at the cost of introducing an additional diagonal matrix D in the decomposition.

THEOREM 13.4 (Existence and Uniqueness)**.** Let $X \in \mathbb{S}^n$.

- $X \in \mathbb{S}^n_+$ if and only if S admits a Cholesky decomposition matrix L with non-negative real diagonal entries.
- $X \in \mathbb{S}^n_{++}$ if and only if S admits a unique Cholesky decomposition matrix L with strictly positive real diagonal entries.

13.3 Eigenvalue Decomposition

DEFINITION 13.5 (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 13.6. Let A be an $n \times n$ matrix. Then A can be eigendecomposed if and only if A has n linearly independent eigenvectors.

13.4 Singular Value Decomposition

DEFINITION 13.7 (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.