# Eigenvalues and Characteristic Polynomial

## Theory and Definitions

**Definition**

Let $A$ be an $n \times n$ matrix. A scalar $\lambda$ is called an **eigenvalue** of $A$ if there exists a nonzero vector $\mathbf{v}$ such that:

$$A\mathbf{v} = \lambda\mathbf{v}$$

Equivalently, $\lambda$ satisfies:

$$\det(A - \lambda I) = 0$$

This equation is called the **characteristic equation**, and its left-hand side is the **characteristic polynomial** of $A$.

**Properties:**

- The characteristic polynomial of an $n \times n$ matrix is a degree-$n$ polynomial.
- The **sum** of the eigenvalues equals $\mathrm{tr}(A)$ (the trace of $A$).
- The **product** of the eigenvalues equals $\det(A)$ (the determinant of $A$).
- Real symmetric matrices have only **real** eigenvalues.

**Skills:**

- Compute eigenvalues by solving $\det(A - \lambda I) = 0$.
- Interpret the characteristic polynomial.
- Factor and analyze for multiplicities.

## Examples

### Example 1: 2x2 Matrix

Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Find its eigenvalues.

$$\det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}$$
$$= (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3$$

Solve: $\lambda^2 - 4\lambda + 3 = 0 \Rightarrow \lambda = 1, 3$

### Example 2: 3x3 Matrix

Let $A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

This matrix is upper triangular, so its eigenvalues are the diagonal entries:

$$\lambda = 3, 2, 1$$

### Example 3: Symmetric Matrix

Let $A = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}$.

$$\det(A - \lambda I) = \det \begin{bmatrix} 4-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix}$$
$$= (4-\lambda)(3-\lambda) - 1 = \lambda^2 - 7\lambda + 11$$

Solve: $\lambda^2 - 7\lambda + 11 = 0 \Rightarrow \lambda = \frac{7\pm\sqrt{5}}{2}$ (real eigenvalues)

## Practice Problems

**P1** Find the eigenvalues of $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

**P2** Find the eigenvalues of $\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$.

**P3** A 3x3 matrix has characteristic polynomial $\lambda^3 - 6\lambda^2 + 11\lambda - 6$. What are its eigenvalues?

**P4** If the trace of $A$ is 10 and its determinant is 24, and one eigenvalue is 2, what are the other two eigenvalues (assume real roots)?

*Answers:*

P1: $\lambda = 3, -1$

P2: $\lambda = -1, -2$

P3: $\lambda = 1, 2, 3$

P4: Remaining sum of eigenvalues = 8. Product of remaining two = 24/2 = 12. So roots of $x^2 - 8x + 12 = 0 \Rightarrow x = 2, 6$

# Diagonalization

## Theory and Definitions

A square matrix $A \in \mathbb{R}^{n\times n}$ is said to be **diagonalizable** if there exists an invertible matrix $P$ and a diagonal matrix $D$ such that:

$$A = PDP^{-1}$$

This means $A$ is similar to a diagonal matrix.

### When is a matrix diagonalizable?

A matrix $A$ is diagonalizable if and only if it has $n$ linearly independent eigenvectors. This is always true if:

- $A$ has $n$ distinct eigenvalues.
- $A$ is a real symmetric matrix.

### Steps to Diagonalize a Matrix

1. Compute the eigenvalues $\lambda_1, \lambda_2, \ldots$ of $A$.
2. Find corresponding eigenvectors $v_1, v_2, \ldots$.
3. Form matrix $P = [v_1 \ v_2 \ \ldots \ v_n]$.
4. Form diagonal matrix $D$ with eigenvalues on the diagonal.
5. Then $A = PDP^{-1}$.

## Examples

### Example 1: Diagonalizable Matrix

Let

$$A = \begin{bmatrix} 4 & 1 \\ 0 & 2 \end{bmatrix}$$

Find if $A$ is diagonalizable.

*Solution:* Eigenvalues are $\lambda = 4, 2$. Two linearly independent eigenvectors exist, so matrix is diagonalizable.

### Example 2: Non-diagonalizable Matrix

Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

*Solution:* Has only one eigenvalue $\lambda = 1$, and only one linearly independent eigenvector. Not diagonalizable.

### Example 3: Symmetric Matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

*Solution:* Symmetric  diagonalizable. Compute eigenvalues and eigenvectors, form $P$ and $D$.

## Practice Problems

1. Diagonalize $A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$
2. Is $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ diagonalizable?
3. Find a matrix $A$ that is **not** diagonalizable and explain why.
4. Diagonalize the symmetric matrix $A = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$

## Answers

1. Eigenvalues $\lambda = 6, 1$, diagonalizable.
2. Not diagonalizable. One eigenvalue, one eigenvector.
3. Example: $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, only one eigenvector.
4. Diagonalizable, real symmetric  guaranteed.

# Famous Matrix Decompositions

## Theory and Definitions

Matrix decompositions are powerful tools in linear algebra, used to simplify solving systems, inverting matrices, and analyzing structure.

### LU Decomposition

LU decomposition expresses a matrix $A$ as the product:

$$A = LU$$

where:

- $L$ is a lower triangular matrix
- $U$ is an upper triangular matrix

**Conditions:**

- LU decomposition exists if no row swaps are needed during Gaussian elimination.
- Otherwise, a permutation matrix $P$ is included: $PA = LU$

### LDV Decomposition

An LU-type decomposition where:

$$A = LDV$$

- $L$: unit lower triangular
- $D$: diagonal
- $V$: unit upper triangular

### LDL$^\top$ Decomposition

For symmetric matrices:

$$A = LDL^\top$$

- $L$: lower unit triangular
- $D$: diagonal
- $L^\top$: transpose of $L$

### Cholesky Decomposition

For symmetric positive-definite matrices:

$$A = LL^\top$$

- $L$: lower triangular with positive diagonal entries

## Examples

### Example 1: LU Decomposition

Decompose

$$A = \begin{bmatrix} 4 & 3 \\ 6 & 3 \end{bmatrix}$$

into $LU$.

*Solution:*

Using Gaussian elimination:

$$L = \begin{bmatrix} 1 & 0 \\ \frac{3}{2} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 4 & 3 \\ 0 & -1.5 \end{bmatrix}$$

### Example 2: Cholesky Decomposition

Find the Cholesky decomposition of:

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$$

*Solution:*

$$L = \begin{bmatrix} 2 & 0 \\ 1 & \sqrt{2} \end{bmatrix}, \quad LL^\top = A$$

## Practice Problems

1. Find the LU decomposition of $\begin{bmatrix} 2 & 3 \\ 4 & 7 \end{bmatrix}$
2. Find the Cholesky decomposition of $A = \begin{bmatrix} 25 & 15 \\ 15 & 13 \end{bmatrix}$
3. Which decomposition would be most appropriate for a symmetric, positive-definite matrix?
4. True or False: Every matrix has a Cholesky decomposition.

## Answers

1. $L = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, U = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$

2. $L = \begin{bmatrix} 5 & 0 \\ 3 & 1 \end{bmatrix}$

3. Cholesky Decomposition

4. False

Famous Matrix Decompositions

# Examples

## Example 1: LU Decomposition

Decompose

$$A = \begin{bmatrix} 4 & 3 \\ 6 & 3 \end{bmatrix}$$

into $LU$.

*Solution:*

$$L = \begin{bmatrix} 1 & 0 \\ \frac{3}{2} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 4 & 3 \\ 0 & -1.5 \end{bmatrix}$$

## Example 2: LU Decomposition with 3x3 matrix

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 7 & 3 \\ 6 & 18 & 5 \end{bmatrix}$$

*Solution:* Use Gaussian elimination to find $L$ and $U$ matrices.

## Example 3: Cholesky Decomposition

Find the Cholesky decomposition of:

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$$

*Solution:*

$$L = \begin{bmatrix} 2 & 0 \\ 1 & \sqrt{2} \end{bmatrix}$$

### Example 4: Cholesky Decomposition (3x3)

$$A = \begin{bmatrix} 6 & 3 & 4 \\ 3 & 6 & 5 \\ 4 & 5 & 10 \end{bmatrix}$$

*Solution:* Matrix is symmetric positive-definite. Apply Cholesky algorithm to compute $L$.

### Example 5: LDL$^\top$ Decomposition

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$$

*Solution:*

$$L = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$

### Example 6: LDV Decomposition (outline)

$$A = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}$$

*Solution:* Decompose into $A = LDV$, where $L$ and $V$ are unit triangular, and $D$ is diagonal.

## Practice Problems

1. Find the LU decomposition of $\begin{bmatrix} 2 & 3 \\ 4 & 7 \end{bmatrix}$
2. Find the Cholesky decomposition of $A = \begin{bmatrix} 25 & 15 \\ 15 & 13 \end{bmatrix}$
3. Which decomposition would be most appropriate for a symmetric, positive-definite matrix?
4. True or False: Every matrix has a Cholesky decomposition.
5. Perform LU decomposition of:
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix}$$
6. Determine the LDL$^\top$ decomposition of:
$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

7. Explain why LU decomposition might require row swapping.
8. Find the LDV decomposition of:
$$\begin{bmatrix} 3 & 1 \\ 6 & 4 \end{bmatrix}$$

# Answers

1. $L = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, U = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$
2. $L = \begin{bmatrix} 5 & 0 \\ 3 & 1 \end{bmatrix}$
3. Cholesky Decomposition
4. False
5. Use Gaussian elimination
6. Tridiagonal symmetric $\rightarrow$ LDL$^\top$ applicable
7. If a pivot is zero, row swaps are required to avoid division by zero.
8. Solve using elementary transformations to extract L, D, V

# Positivity and Positive-Definite Matrices

Based on Oliver Shakiban's Text

## Theory and Definitions

A square matrix $A \in \mathbb{R}^{n \times n}$ is said to be **positive-definite** if it satisfies any of the following equivalent conditions:

- For all nonzero vectors $x \in \mathbb{R}^n$, the quadratic form $x^\top Ax > 0$.
- All eigenvalues of $A$ are positive.
- All leading principal minors of $A$ are positive.

If a matrix satisfies only that $x^\top Ax \geq 0$ for all $x \in \mathbb{R}^n$, then it is called **positive semi-definite**.

### Symmetric Matrices and Cholesky Decomposition

A matrix $A$ must be symmetric to be tested for positive-definiteness using the **Cholesky decomposition**. A symmetric matrix $A$ is positive-definite if and only if there exists a lower triangular matrix $L$ such that

$$A = LL^\top.$$

## Examples

### Example 1: A Positive-Definite Matrix

Let

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

We check if $x^\top Ax > 0$ for all $x \neq 0$. Also compute eigenvalues:

$$\det(A - \lambda I) = (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0.$$

Eigenvalues: $\lambda = 1, 3$ (positive). So $A$ is positive-definite.

### Example 2: Not Positive-Definite

Let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

Compute eigenvalues:

$$\det(A - \lambda I) = (1 - \lambda)^2 - 4 = \lambda^2 - 2\lambda - 3.$$

Eigenvalues: $\lambda = 3, -1$ (one negative). So $A$ is not positive-definite.

# Practice Problems

1. Determine if the matrix is positive-definite:

$$A = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}$$

2. Determine if the matrix is positive-definite using leading principal minors:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

3. Use Cholesky decomposition to verify if the following matrix is positive-definite:

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

# Answers

1. Eigenvalues are positive: $5, 2 \Rightarrow$ Positive-definite.
2. All leading principal minors are positive: $\Rightarrow$ Positive-definite.
3. Cholesky exists: $A = LL^\top$ with

$$L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix} \Rightarrow \text{Positive-definite.}$$

# Gram-Schmidt Process and QR Decomposition

Based on Oliver Shakiban's Text

## Gram-Schmidt Process

The **Gram-Schmidt Process** takes a linearly independent set of vectors and converts it into an **orthonormal basis** for the same subspace. It is foundational for QR decomposition and orthogonal projections.

Given a set of linearly independent vectors $u_1, u_2, \ldots, u_n$, the process generates an orthonormal set $q_1, q_2, \ldots, q_n$ as follows:

$$v_k = u_k - \sum_{j=1}^{k-1} \langle u_k, q_j \rangle q_j, \quad q_k = \frac{v_k}{\|v_k\|}$$

### Example

Given $u_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, perform Gram-Schmidt:

$$\begin{aligned} q_1 &= \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ v_2 &= u_2 - \langle u_2, q_1 \rangle q_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix} \\ q_2 &= \frac{v_2}{\|v_2\|} = \frac{1}{\sqrt{0.5}} \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

## QR Decomposition

**QR Decomposition** expresses a full-rank matrix $A$ as a product $A = QR$ where:

- $Q$ has orthonormal columns (i.e., $Q^\top Q = I$),
- $R$ is an upper triangular matrix.

This decomposition is especially useful for:

- Solving **least squares problems**
- Connecting to **Gram-Schmidt**: the $Q$ matrix contains the orthonormal vectors produced by Gram-Schmidt, and $R$ gives the coefficients used in forming the original vectors.

### Example

Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Applying Gram-Schmidt on the columns gives $Q$, and $R = Q^\top A$.

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad R = Q^\top A = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

# Practice Problems

1. Use Gram-Schmidt to orthonormalize the set:

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$$

2. Find the QR decomposition of:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

3. Explain why $Q$ in QR decomposition is always orthogonal when $A$ is full rank.

# Answers

1. Apply the Gram-Schmidt formula step-by-step to obtain orthonormal vectors.
2. Use $Q$ from Gram-Schmidt; compute $R = Q^\top A$.
3. Because Gram-Schmidt constructs an orthonormal set, $Q^\top Q = I$.

# Inner Products and Norms

Based on Oliver Shakiban's Text

## Inner Products

An **inner product** is a generalization of the dot product. It is a function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ that satisfies:

- **Linearity**: $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$
- **Symmetry**: $\langle x, y \rangle = \langle y, x \rangle$
- **Positivity**: $\langle x, x \rangle \geq 0$ and equals 0 only if $x = 0$

  The standard inner product in $\mathbb{R}^n$ is:

$$\langle x, y \rangle = x^\top y = \sum_{i=1}^{n} x_i y_i$$

**Weighted inner products** use a symmetric positive-definite matrix $W$:

$$\langle x, y \rangle_W = x^\top W y$$

## Norms

A **norm** is a function $\| \cdot \| : \mathbb{R}^n \to \mathbb{R}$ that satisfies:

- $\|x\| \geq 0$ and $\|x\| = 0 \iff x = 0$
- $\|\alpha x\| = |\alpha| \|x\|$ (homogeneity)
- $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)

### Vector Norms

- $\|x\|_1 = \sum_i |x_i|$ (Manhattan norm)
- $\|x\|_2 = \sqrt{\sum_i x_i^2}$ (Euclidean norm)
- $\|x\|_\infty = \max_i |x_i|$ (Maximum norm)

### Matrix Norms

A matrix norm assigns a size to matrices. Common types include:

- **Frobenius norm**: $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$
- **Induced 2-norm**: $\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2$ = largest singular value of $A$

# Examples

## Example 1: Standard Inner Product

For $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $y = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$:

$$\langle x, y \rangle = 1 \cdot 3 + 2 \cdot 4 = 11$$

## Example 2: Weighted Inner Product

With $W = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$,

$$\langle x, y \rangle_W = x^\top W y = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 1 \cdot 2 \cdot 3 + 2 \cdot 4 = 6 + 8 = 14$$

## Example 3: Induced 2-Norm

For $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, the induced 2-norm is the maximum singular value = 2.

# Practice Problems

1. Compute $\langle x, y \rangle$ for $x = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $y = \begin{bmatrix} 4 \\ -2 \end{bmatrix}$.
2. Compute $\|x\|_2$ for $x = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.
3. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Find $\|A\|_F$.

# Answers

1. $\langle x, y \rangle = 1 \cdot 4 + 3 \cdot (-2) = 4 - 6 = -2$
2. $\|x\|_2 = \sqrt{3^2 + 4^2} = 5$
3. $\|A\|_F = \sqrt{0^2 + 1^2 + (-1)^2 + 0^2} = \sqrt{2}$

# Cholesky Factorization, Equivalence of Norms, Orthogonal Matrices

Based on Oliver Shakiban's Text

## Cholesky Factorization

The **Cholesky Factorization** is a matrix decomposition applicable only to **symmetric positive-definite matrices**. For such a matrix $A$, there exists a unique lower triangular matrix $L$ with positive diagonal entries such that:

$$A = LL^\top$$

**Key Features**:

- Works only for symmetric positive-definite matrices.
- More efficient and numerically stable than LU decomposition.
- Useful in solving systems $Ax = b$ by forward and backward substitution.

### Example

Let

$$A = \begin{bmatrix} 4 & 12 & -16 \\ 12 & 37 & -43 \\ -16 & -43 & 98 \end{bmatrix}$$

Cholesky factorization gives:

$$L = \begin{bmatrix} 2 & 0 & 0 \\ 6 & 1 & 0 \\ -8 & 5 & 3 \end{bmatrix}, \quad \text{since } A = LL^\top$$

## Equivalence of Norms on $\mathbb{R}^n$

**Theorem (3.17):** All norms on $\mathbb{R}^n$ are equivalent.

This means that for any two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ on $\mathbb{R}^n$, there exist constants $c, C > 0$ such that:

$$c\|x\|_a \leq \|x\|_b \leq C\|x\|_a \quad \text{for all } x \in \mathbb{R}^n$$

**Implications:**

- Convergence in one norm implies convergence in any other.
- Boundedness is preserved under any norm.
- Topological properties of $\mathbb{R}^n$ do not depend on the choice of norm.

### Example

Although $\|x\|_1$, $\|x\|_2$, and $\|x\|_\infty$ differ in value, they all define the same notion of convergence and bounded sets in $\mathbb{R}^n$.

# Orthogonal Matrices

A matrix $Q \in \mathbb{R}^{n\times n}$ is **orthogonal** if:

$$Q^\top Q = I \quad \text{(equivalently, } Q^{-1} = Q^\top\text{)}$$

**Properties**:

- Columns (and rows) of $Q$ form an orthonormal set.
- $\|Qx\| = \|x\|$ for all $x$ (length-preserving).
- Preserves angles between vectors.
- $\det(Q) = \pm 1$.

## Example

Let

$$Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Then $Q^\top Q = I$, so $Q$ is orthogonal.

# Practice Problems

1. Perform Cholesky decomposition of:

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

2. Prove that $\|x\|_1 \leq \sqrt{n}\|x\|_2$ for $x \in \mathbb{R}^n$.
3. Show that if $Q$ is orthogonal, then $\|Qx\|_2 = \|x\|_2$.

# Answers

1. Cholesky factorization exists. $L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix}$
2. Follows from Cauchy-Schwarz: $\|x\|_1 \leq \sqrt{n}\|x\|_2$
3. $\|Qx\|_2^2 = x^\top Q^\top Q x = x^\top x = \|x\|_2^2$

Apostol Volume II - Chapter 8
Differentiation of Functions of Several Variables

## 1. Directional Derivatives and Partial Derivatives

Let $f : \mathbb{R}^n \to \mathbb{R}$. The **directional derivative** of $f$ at a point $\mathbf{a}$ in the direction of a unit vector $\mathbf{u}$ is:

$$D_{\mathbf{u}}f(\mathbf{a}) = \lim_{h\to 0} \frac{f(\mathbf{a}+h\mathbf{u}) - f(\mathbf{a})}{h}$$

If $f$ is differentiable at $\mathbf{a}$, then:

$$D_{\mathbf{u}}f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{u}$$

Partial derivatives are special cases where $\mathbf{u}$ is a standard basis vector.

### Example

Let $f(x, y) = x^2y + y^3$. Find the directional derivative at $(1, 2)$ in the direction of $\mathbf{u} = \frac{1}{\sqrt{2}}(1, 1)$.

$$\nabla f = (2xy, x^2 + 3y^2), \quad \nabla f(1, 2) = (4, 13)$$
$$D_{\mathbf{u}}f(1, 2) = \frac{1}{\sqrt{2}}(4 + 13) = \frac{17}{\sqrt{2}}$$

## 2. Partial Derivatives of Higher Order

Higher-order partial derivatives are derivatives taken multiple times with respect to one or more variables:

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \text{etc.}$$

### Example

$f(x, y) = x^2y^3$

$$\frac{\partial f}{\partial x} = 2xy^3, \quad \frac{\partial^2 f}{\partial x \partial y} = 6xy^2$$

## 3. Directional Derivatives and Continuity

A function may have directional derivatives at a point without being continuous there.

### Example

$f(x, y) = \frac{xy}{x^2+y^2}$ for $(x, y) \neq (0, 0)$, and $f(0, 0) = 0$.

Directional derivatives at the origin exist but $f$ is not continuous at $(0, 0)$.

## 4. Total Derivative

The total derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the linear map $Df(\mathbf{a})$ such that:

$$f(\mathbf{a} + \mathbf{h}) \approx f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{h})$$

If $f$ has continuous partial derivatives, then the total derivative exists and is given by the gradient:

$$Df(\mathbf{a}) = \nabla f(\mathbf{a})$$

## 5. A Sufficient Condition for Differentiability

If all partial derivatives of $f$ exist and are continuous near $\mathbf{a}$, then $f$ is differentiable at $\mathbf{a}$.

## 6. A Chain Rule for Derivatives of Scalar Fields

Let $f : \mathbb{R}^n \to \mathbb{R}$, and $\mathbf{x}(t) : \mathbb{R} \to \mathbb{R}^n$, then:

$$\frac{d}{dt} f(\mathbf{x}(t)) = \nabla f(\mathbf{x}(t)) \cdot \frac{d\mathbf{x}}{dt}$$

### Example

If $f(x, y) = x^2 + y^2$, $x = t^2, y = \sin t$, then:

$$\frac{d}{dt} f = 2x \cdot \frac{dx}{dt} + 2y \cdot \frac{dy}{dt} = 2t^2 \cdot 2t + 2 \sin t \cdot \cos t$$

## 7. Derivatives of Vector Fields

Let $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$, then the derivative at a point is the Jacobian matrix:

$$J_{\mathbf{F}} = \left[ \frac{\partial F_i}{\partial x_j} \right]$$

### Example

$\mathbf{F}(x, y) = (x^2 + y, xy^2) \Rightarrow J = \begin{bmatrix} 2x & 1 \\ y^2 & 2xy \end{bmatrix}$

## 8. Differentiability Implies Continuity

If a function is differentiable at a point, it is also continuous at that point.

## 9. Chain Rule for Derivatives of Vector Fields

Let $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{G} : \mathbb{R}^m \to \mathbb{R}^p$, then:

$$D(\mathbf{G} \circ \mathbf{F})(x) = D\mathbf{G}(\mathbf{F}(x)) \cdot D\mathbf{F}(x)$$

## 10. Matrix Form of the Chain Rule

The chain rule in matrix form for compositions of vector functions:

$$J_{\mathbf{G}\circ\mathbf{F}} = J_{\mathbf{G}} \cdot J_{\mathbf{F}}$$

## 11. Sufficient Conditions for Equality of Mixed Partial Derivatives

If the mixed partial derivatives are continuous at a point, then they are equal:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

### Example

$f(x, y) = x^2 y^3 \Rightarrow \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 6xy^2$

# Apostol Volume II - Chapter 9
# Extrema of Functions of Several Variables

## 1. Maxima, Minima and Saddle Points

A point $\mathbf{a} \in \mathbb{R}^n$ is a:

- **Local maximum** if $f(\mathbf{a}) \geq f(\mathbf{x})$ for all $\mathbf{x}$ near $\mathbf{a}$,
- **Local minimum** if $f(\mathbf{a}) \leq f(\mathbf{x})$ for all $\mathbf{x}$ near $\mathbf{a}$,
- **Saddle point** if it is not a max/min but is a stationary point (i.e., $\nabla f(\mathbf{a}) = 0$).

### Example

$f(x, y) = x^2 - y^2$: has a saddle point at (0,0) since

$$\nabla f = (2x, -2y), \quad \nabla f(0, 0) = (0, 0)$$

## 2. Second Order Taylor Formula for Scalar Fields

The second-order Taylor expansion of $f$ near $\mathbf{a}$ is:

$$f(\mathbf{a} + \mathbf{h}) \approx f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h} + \frac{1}{2}\mathbf{h}^T H_f(\mathbf{a})\mathbf{h}$$

Where $H_f(\mathbf{a})$ is the **Hessian matrix** of second partial derivatives.

### Example

Let $f(x, y) = x^2 + y^2$, compute Taylor expansion at $(0, 0)$:

$$\nabla f = (2x, 2y), \quad H_f = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$f(h, k) \approx 0 + 0 + \frac{1}{2}(h^2 + k^2) = \frac{1}{2}(h^2 + k^2)$$

## 3. Nature of Stationary Point Determined by Eigenvalues of Hessian Matrix

If $\nabla f(\mathbf{a}) = 0$, then the behavior of $f$ near $\mathbf{a}$ is determined by the eigenvalues of the Hessian matrix:

- All eigenvalues positive: local minimum
- All eigenvalues negative: local maximum
- Mixed signs: saddle point
- Zero eigenvalue(s): test is inconclusive

# 4. Second Derivative Test for Extrema of Functions of Two Variables

Let $f(x, y)$ be $C^2$ and let $(a, b)$ be a critical point. Define:

$$D = f_{xx}(a, b)f_{yy}(a, b) - (f_{xy}(a, b))^2$$

- If $D > 0$ and $f_{xx}(a, b) > 0$: local minimum
- If $D > 0$ and $f_{xx}(a, b) < 0$: local maximum
- If $D < 0$: saddle point
- If $D = 0$: test is inconclusive

## Example

Let $f(x, y) = x^3 - 3xy^2$

$$f_x = 3x^2 - 3y^2, \quad f_y = -6xy$$

Critical point at $(0, 0)$

$$f_{xx} = 6x, \quad f_{yy} = -6x, \quad f_{xy} = -6y \Rightarrow D = f_{xx}f_{yy} - (f_{xy})^2 = -36x^2 - 36y^2 \Rightarrow D(0, 0) = 0$$

So test is inconclusive.

Practice Problems
Apostol Volume II Chapters 8 and 9

# Practice Problems

## Chapter 8: Differentiation of Functions of Several Variables

**8.1** Compute the directional derivative of $f(x, y) = x^2y + y^3$ at $(1, 2)$ in the direction of $\mathbf{u} = \frac{1}{\sqrt{2}}(1, 1)$.

**8.2** Compute the gradient of $f(x, y) = \ln(x^2 + y^2)$ and evaluate it at $(1, 1)$.

**8.3** Find all second-order partial derivatives of $f(x, y) = e^{xy}$.

**8.4** Check whether the function $f(x, y) = \frac{2xy}{x^2+y^2}$ (with $f(0, 0) = 0$) is continuous at the origin.

**8.5** Find the total derivative of $f(x, y) = x^2 + xy$ at the point $(1, 2)$.

**8.6** State a sufficient condition for differentiability of a function $f(x, y)$ at a point.

**8.7** Let $f(x, y) = \sin(x^2 + y^2)$, and $x = t^2, y = t$. Use the chain rule to compute $\frac{df}{dt}$.

**8.8** Find the Jacobian matrix of $\mathbf{F}(x, y) = (xy, x^2 + y^2)$.

**8.9** If $f(x, y)$ is differentiable at $(1, 2)$, is it necessarily continuous at that point? Justify.

**8.10** Given $\mathbf{F}(x, y) = (x + y, xy)$ and $\mathbf{G}(u, v) = (u^2, v^2)$, compute $D(\mathbf{G} \circ \mathbf{F})$.

**8.11** Verify whether $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ for $f(x, y) = \ln(x^2 + y^2)$.

## Chapter 9: Extrema of Functions of Several Variables

**9.1** Identify and classify the critical points of $f(x, y) = x^2 - y^2$.

**9.2** Find and classify the critical points of $f(x, y) = x^2 + y^2 + xy$.

**9.3** Use the second-order Taylor formula to approximate $f(x, y) = e^x \cos y$ at $(0, 0)$ for small $x, y$.

**9.4** Compute the Hessian matrix of $f(x, y) = x^2 + 4xy + y^2$, and determine the nature of the stationary point at the origin.

**9.5** Use the second derivative test to classify the critical point of $f(x, y) = x^3 - 3xy^2$ at $(0, 0)$.

**9.6** Find all critical points and use the second derivative test for $f(x, y) = x^4 + y^4 - 4xy$.

**9.7** Let $f(x, y) = x^2 + 3xy + y^2$. Compute the second-order Taylor polynomial at $(1, 1)$.

**9.8** Explain how eigenvalues of the Hessian matrix help classify stationary points.

**9.9** Determine whether the function $f(x, y) = x^2 + y^2$ has a saddle point.

# Solutions

**8.1** $\nabla f = (2xy, x^2 + 3y^2) = (4, 13)$, Answer: $\frac{17}{\sqrt{2}}$

**8.2** $\nabla f = \left(\frac{2x}{x^2+y^2}, \frac{2y}{x^2+y^2}\right) = (1, 1)$

**8.3** $f_x = ye^{xy}, f_y = xe^{xy}, f_{xx} = y^2e^{xy}, f_{yy} = x^2e^{xy}, f_{xy} = (1 + xy)e^{xy}$

**8.4** Not continuous at origin (limit depends on path)

**8.5** $Df = (2x + y, x)$, at $(1, 2) : (4, 1)$

**8.6** If all partials exist and are continuous → function is differentiable.

**8.7** Chain rule: $df/dt = \nabla f \cdot dx/dt = 2t\cos(t^4 + t^2) + 2t^2\cos(t^4 + t^2)$

**8.8** $J = \begin{bmatrix} y & x \\ 2x & 2y \end{bmatrix}$

**8.9** Yes, differentiability implies continuity.

**8.10** $D\mathbf{F} = \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix}, D\mathbf{G} = \begin{bmatrix} 2u & 0 \\ 0 & 2v \end{bmatrix} \Rightarrow D(\mathbf{G} \circ \mathbf{F})$ via chain rule

**8.11** Mixed partials are equal if function is smooth — here they are.

**9.1** Saddle point at (0,0)

**9.2** $\nabla f = (2x + y, 2y + x) = 0 \Rightarrow (0, 0) \to \min$

**9.3** $f(x, y) \approx 1 + x - \frac{y^2}{2}$

**9.4** $H = \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}$, eigenvalues: $6, -2 \Rightarrow$ saddle

**9.5** $D = -36x^2 - 36y^2 = 0 \Rightarrow$ inconclusive

**9.6** $\nabla f = (4x^3 - 4y, 4y^3 - 4x)$, critical points at $(0, 0)$ and more → test each using Hessian.

**9.7** Gradient = (2x + 3y, 3x + 2y), Hessian = $[[2, 3], [3, 2]]$ → Apply Taylor formula

**9.8** Positive definite → local min, negative definite → max, mixed → saddle

**9.9** No, has a local minimum