

Fundamental concepts

Tiago M. Magalhães

Departamento de Estatística - ICE-UFJF

Juiz de Fora, August 26, 2026



Outline

- 1 Operations
- 2 Matrix Results
- 3 Estimation
- 4 References



Outline

- 1 Operations
- 2 Matrix Results
- 3 Estimation
- 4 References



Summations

Let $Y_1, Y_2, \dots, Y_n$ be n arbitrary numbers. Then,

$$\sum_{i=1}^n Y_i = Y_1 + Y_2 + \dots + Y_n, \quad (1)$$



Summations

Let $Y_1, Y_2, \dots, Y_n$ be n arbitrary numbers. Then,

$$\sum_{i=1}^n Y_i = Y_1 + Y_2 + \dots + Y_n, \quad (1)$$

where $\sum$ is called the summation operator.



Summations

Let $Y_1, Y_2, \dots, Y_n$ be n arbitrary numbers. Then,

$$\sum_{i=1}^n Y_i = Y_1 + Y_2 + \dots + Y_n, \quad (1)$$

where $\sum$ is called the summation operator.



Summations

From (1), we have the following properties:

- $\sum_{i=1}^n a = na;$



Summations

From (1), we have the following properties:

- $\sum_{i=1}^n a = na;$
- $\sum_{i=1}^n (Y_i + Z_i) = \sum_{i=1}^n Y_i + \sum_{i=1}^n Z_i;$



Summations

From (1), we have the following properties:

- $\sum_{i=1}^n a = na;$
- $\sum_{i=1}^n (Y_i + Z_i) = \sum_{i=1}^n Y_i + \sum_{i=1}^n Z_i;$
- $\sum_{i=1}^n (b + cZ_i) = nb + c \sum_{i=1}^n Z_i;$



Summations

From (1), we have the following properties:

- $\sum_{i=1}^n a = na$;
- $\sum_{i=1}^n (Y_i + Z_i) = \sum_{i=1}^n Y_i + \sum_{i=1}^n Z_i$;
- $\sum_{i=1}^n (b + cZ_i) = nb + c \sum_{i=1}^n Z_i$;

where a, b, c are constants and $Z_1, Z_2, \dots, Z_n$ are arbitrary numbers.



Summations

From (1), we have the following properties:

- $\sum_{i=1}^n a = na$;
- $\sum_{i=1}^n (Y_i + Z_i) = \sum_{i=1}^n Y_i + \sum_{i=1}^n Z_i$;
- $\sum_{i=1}^n (b + cZ_i) = nb + c \sum_{i=1}^n Z_i$;

where a, b, c are constants and $Z_1, Z_2, \dots, Z_n$ are arbitrary numbers.



Products

Let $Y_1, Y_2, \dots, Y_n$ be n arbitrary numbers. Then,

$$\prod_{i=1}^n Y_i = Y_1 \times Y_2 \times \cdots \times Y_n, \quad (2)$$

Products

Let $Y_1, Y_2, \dots, Y_n$ be n arbitrary numbers. Then,

$$\prod_{i=1}^n Y_i = Y_1 \times Y_2 \times \cdots \times Y_n, \quad (2)$$

where $\prod$ is called the product operator.



Products

Let $Y_1, Y_2, \dots, Y_n$ be n arbitrary numbers. Then,

$$\prod_{i=1}^n Y_i = Y_1 \times Y_2 \times \cdots \times Y_n, \quad (2)$$

where $\prod$ is called the product operator.



Products

From (2), we have the following properties:

- $\prod_{i=1}^n a = a^n;$



Products

From (2), we have the following properties:

- $\prod_{i=1}^n a = a^n$;
- $\prod_{i=1}^n (Y_i \times Z_i) = \prod_{i=1}^n Y_i \times \prod_{i=1}^n Z_i$;



Products

From (2), we have the following properties:

- $\prod_{i=1}^n a = a^n$;
- $\prod_{i=1}^n (Y_i \times Z_i) = \prod_{i=1}^n Y_i \times \prod_{i=1}^n Z_i$;
- $\prod_{i=1}^n (bZ_i) = b^n \times \prod_{i=1}^n Z_i$;



Products

From (2), we have the following properties:

- $\prod_{i=1}^n a = a^n$;
- $\prod_{i=1}^n (Y_i \times Z_i) = \prod_{i=1}^n Y_i \times \prod_{i=1}^n Z_i$;
- $\prod_{i=1}^n (bZ_i) = b^n \times \prod_{i=1}^n Z_i$;

where a, b are constants and $Z_1, Z_2, \dots, Z_n$ are arbitrary numbers.



Products

From (2), we have the following properties:

- $\prod_{i=1}^n a = a^n$;
- $\prod_{i=1}^n (Y_i \times Z_i) = \prod_{i=1}^n Y_i \times \prod_{i=1}^n Z_i$;
- $\prod_{i=1}^n (bZ_i) = b^n \times \prod_{i=1}^n Z_i$;

where a, b are constants and $Z_1, Z_2, \dots, Z_n$ are arbitrary numbers.



Outline

- 1 Operations
- 2 Matrix Results**
- 3 Estimation
- 4 References



Linear Equations

A system of n equations and p unknowns:

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p & = & y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p & = & y_2 \\ \vdots & = & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{np}x_p & = & y_n \end{array} \right. ,$$

can be written in matrix form,



Linear Equations

A system of n equations and p unknowns:

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p & = & y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p & = & y_2 \\ \vdots & = & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{np}x_p & = & y_n \end{array} \right. ,$$

can be written in matrix form, as

$$Ax = y,$$



Linear Equations

A system of n equations and p unknowns:

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p & = & y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p & = & y_2 \\ \vdots & = & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{np}x_p & = & y_n \end{array} \right. ,$$

can be written in matrix form, as

$$\mathbf{Ax} = \mathbf{y},$$



Linear Equations

where $\mathbf{A}$ is an $n \times p$ coefficient matrix, $\mathbf{x}$ is a $p \times 1$ vector



Linear Equations

where $\mathbf{A}$ is an $n \times p$ coefficient matrix, $\mathbf{x}$ is a $p \times 1$ vector and $\mathbf{y}$ is an $n \times 1$ vector,



Linear Equations

where $\mathbf{A}$ is an $n \times p$ coefficient matrix, $\mathbf{x}$ is a $p \times 1$ vector and $\mathbf{y}$ is an $n \times 1$ vector, that is,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Linear Equations

where $\mathbf{A}$ is an $n \times p$ coefficient matrix, $\mathbf{x}$ is a $p \times 1$ vector and $\mathbf{y}$ is an $n \times 1$ vector, that is,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Inverse of a 2×2 Matrix

Let

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

If $\det(\mathbf{A}) \neq 0$, then $\mathbf{A}$ is nonsingular and its inverse is given by

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$



Inverse of a 2×2 Matrix

Let

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

If $\det(\mathbf{A}) \neq 0$, then $\mathbf{A}$ is nonsingular and its inverse is given by

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

In fact, $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_2$, where $\mathbf{I}_2$ is the identity matrix of order



Inverse of a 2×2 Matrix

Let

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

If $\det(\mathbf{A}) \neq 0$, then $\mathbf{A}$ is nonsingular and its inverse is given by

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

In fact, $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_2$, where $\mathbf{I}_2$ is the identity matrix of order 2.



Idempotent matrix

Definition

A matrix **M** is said to be idempotent if and only if **MM = M**.



Characteristic Roots

The characteristic roots of a matrix $\mathbf{A}$, $n \times n$, are the solutions in λ of the equation

$$\det(\mathbf{A} - \lambda \mathbf{I}_n) = 0,$$



Characteristic Roots

The characteristic roots of a matrix $\mathbf{A}$, $n \times n$, are the solutions in λ of the equation

$$\det(\mathbf{A} - \lambda \mathbf{I}_n) = 0,$$

where $\mathbf{I}_n$ is the identity matrix of order n .



Characteristic Roots

The characteristic roots of a matrix $\mathbf{A}$, $n \times n$, are the solutions in λ of the equation

$$\det(\mathbf{A} - \lambda \mathbf{I}_n) = 0,$$

where $\mathbf{I}_n$ is the identity matrix of order n . This determinant is a polynomial of degree n in λ and the equation will have n solutions.



Characteristic Roots

The characteristic roots of a matrix $\mathbf{A}$, $n \times n$, are the solutions in λ of the equation

$$\det(\mathbf{A} - \lambda \mathbf{I}_n) = 0,$$

where $\mathbf{I}_n$ is the identity matrix of order n . This determinant is a polynomial of degree n in λ and the equation will have n solutions. If λ_i is a characteristic root of $\mathbf{A}$, the solution in $\mathbf{x}$ of the equation $(\mathbf{A} - \lambda_i \mathbf{I}_n)\mathbf{x} = 0$ is the corresponding characteristic vector, $\mathbf{x}_i$.



Characteristic Roots

The characteristic roots of a matrix $\mathbf{A}$, $n \times n$, are the solutions in λ of the equation

$$\det(\mathbf{A} - \lambda \mathbf{I}_n) = 0,$$

where $\mathbf{I}_n$ is the identity matrix of order n . This determinant is a polynomial of degree n in λ and the equation will have n solutions. If λ_i is a characteristic root of $\mathbf{A}$, the solution in $\mathbf{x}$ of the equation $(\mathbf{A} - \lambda_i \mathbf{I}_n)\mathbf{x} = 0$ is the corresponding characteristic vector, $\mathbf{x}_i$.



Characteristic Roots

Results:

- The product of the characteristic roots of $\mathbf{A}$ is $\det(\mathbf{A})$;



Characteristic Roots

Results:

- The product of the characteristic roots of $\mathbf{A}$ is $\det(\mathbf{A})$;
- The sum of the characteristic roots of $\mathbf{A}$ is the trace of $\mathbf{A}$;



Characteristic Roots

Results:

- The product of the characteristic roots of $\mathbf{A}$ is $\det(\mathbf{A})$;
- The sum of the characteristic roots of $\mathbf{A}$ is the trace of $\mathbf{A}$;
- If the rank of $\mathbf{A}$ is equal to r , $r(\mathbf{A}) = r$, zero is a root of the equation with multiplicity $n - r$.



Characteristic Roots

Results:

- The product of the characteristic roots of $\mathbf{A}$ is $\det(\mathbf{A})$;
- The sum of the characteristic roots of $\mathbf{A}$ is the trace of $\mathbf{A}$;
- If the rank of $\mathbf{A}$ is equal to r , $r(\mathbf{A}) = r$, zero is a root of the equation with multiplicity $n - r$.



Quadratic Forms

The function $f(x_1, x_2, \dots, x_n)$ of n real variables $x_1, x_2, \dots, x_n$ is a quadratic form if

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \mathbf{x}^\top \mathbf{A} \mathbf{x},$$



Quadratic Forms

The function $f(x_1, x_2, \dots, x_n)$ of n real variables $x_1, x_2, \dots, x_n$ is a quadratic form if

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \mathbf{x}^\top \mathbf{A} \mathbf{x},$$

where $\mathbf{x} = (x_1 x_2 \cdots x_n)^\top$



Quadratic Forms

The function $f(x_1, x_2, \dots, x_n)$ of n real variables $x_1, x_2, \dots, x_n$ is a quadratic form if

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \mathbf{x}^\top \mathbf{A} \mathbf{x},$$

where $\mathbf{x} = (x_1 x_2 \cdots x_n)^\top$ and $\mathbf{A} = (a_{ij})$, a symmetric $n \times n$ matrix, is called the matrix of the quadratic form.



Quadratic Forms

The function $f(x_1, x_2, \dots, x_n)$ of n real variables $x_1, x_2, \dots, x_n$ is a quadratic form if

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \mathbf{x}^\top \mathbf{A} \mathbf{x},$$

where $\mathbf{x} = (x_1 x_2 \cdots x_n)^\top$ and $\mathbf{A} = (a_{ij})$, a symmetric $n \times n$ matrix, is called the matrix of the quadratic form.



Quadratic Forms

We can classify quadratic forms as:

- ① $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$, for every $\mathbf{x} \neq \mathbf{0}$;



Quadratic Forms

We can classify quadratic forms as:

- ① $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ② $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x} \neq \mathbf{0}$;



Quadratic Forms

We can classify quadratic forms as:

- ① $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ② $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ③ $\mathbf{A}$ is said to be positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x}$;



Quadratic Forms

We can classify quadratic forms as:

- ① $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ② $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ③ $\mathbf{A}$ is said to be positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x}$;
- ④ $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is negative definite if $-\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite



Quadratic Forms

We can classify quadratic forms as:

- ① $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ② $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ③ $\mathbf{A}$ is said to be positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x}$;
- ④ $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is negative definite if $-\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite and is negative semidefinite if $-\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive semidefinite.



Quadratic Forms

We can classify quadratic forms as:

- ① $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ② $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x} \neq \mathbf{0}$;
- ③ $\mathbf{A}$ is said to be positive semidefinite if $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, for every $\mathbf{x}$;
- ④ $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is negative definite if $-\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite and is negative semidefinite if $-\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive semidefinite.



Quadratic Forms

Example. Let $\mathbf{X} = (X_1, \dots, X_n)^\top$ be an n -dimensional random vector with covariance matrix Σ . The matrix Σ is positive semidefinite,



Quadratic Forms

Example. Let $\mathbf{X} = (X_1, \dots, X_n)^\top$ be an n -dimensional random vector with covariance matrix Σ . The matrix Σ is positive semidefinite, since, if $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a non-random vector

$$\text{Var}(a_1 X_1 + \dots + a_n X_n) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{Cov}(X_i, X_j) = \mathbf{a}^\top \Sigma \mathbf{a} \geq 0, \quad \forall \mathbf{a} \neq \mathbf{0}.$$

Quadratic Forms

Example. Let $\mathbf{X} = (X_1, \dots, X_n)^\top$ be an n -dimensional random vector with covariance matrix Σ . The matrix Σ is positive semidefinite, since, if $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a non-random vector

$$\text{Var}(a_1 X_1 + \dots + a_n X_n) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{Cov}(X_i, X_j) = \mathbf{a}^\top \Sigma \mathbf{a} \geq 0, \quad \forall \mathbf{a} \neq \mathbf{0}.$$

Remark: The matrix Σ will be positive definite if and only if there is no nontrivial linear combination of the random variables $X_1, \dots, X_n$ that is constant almost surely.



Quadratic Forms

Example. Let $\mathbf{X} = (X_1, \dots, X_n)^\top$ be an n -dimensional random vector with covariance matrix Σ . The matrix Σ is positive semidefinite, since, if $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a non-random vector

$$\text{Var}(a_1 X_1 + \dots + a_n X_n) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{Cov}(X_i, X_j) = \mathbf{a}^\top \Sigma \mathbf{a} \geq 0, \quad \forall \mathbf{a} \neq \mathbf{0}.$$

Remark: The matrix Σ will be positive definite if and only if there is no nontrivial linear combination of the random variables $X_1, \dots, X_n$ that is constant almost surely.



Quadratic Forms

Transformations. Let $\mathbf{y} = \mathbf{B}^{-1}\mathbf{x}$ be a transformation of $\mathbf{x}$,



Quadratic Forms

Transformations. Let $\mathbf{y} = \mathbf{B}^{-1}\mathbf{x}$ be a transformation of $\mathbf{x}$, where $\mathbf{B}$ is a nonsingular matrix, that is $\mathbf{x} = \mathbf{B}\mathbf{y}$.



Quadratic Forms

Transformations. Let $\mathbf{y} = \mathbf{B}^{-1}\mathbf{x}$ be a transformation of $\mathbf{x}$, where $\mathbf{B}$ is a nonsingular matrix, that is $\mathbf{x} = \mathbf{B}\mathbf{y}$. Then

$$Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{B}^\top \mathbf{A} \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{C} \mathbf{y}, \quad (3)$$



Quadratic Forms

Transformations. Let $\mathbf{y} = \mathbf{B}^{-1}\mathbf{x}$ be a transformation of $\mathbf{x}$, where $\mathbf{B}$ is a nonsingular matrix, that is $\mathbf{x} = \mathbf{B}\mathbf{y}$. Then

$$Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{B}^\top \mathbf{A} \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{C} \mathbf{y}, \quad (3)$$

where $\mathbf{C} = \mathbf{B}^\top \mathbf{A} \mathbf{B}$ is the quadratic form in the new variables.



Quadratic Forms

Transformations. Let $\mathbf{y} = \mathbf{B}^{-1}\mathbf{x}$ be a transformation of $\mathbf{x}$, where $\mathbf{B}$ is a nonsingular matrix, that is $\mathbf{x} = \mathbf{B}\mathbf{y}$. Then

$$Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{B}^\top \mathbf{A} \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{C} \mathbf{y}, \quad (3)$$

where $\mathbf{C} = \mathbf{B}^\top \mathbf{A} \mathbf{B}$ is the quadratic form in the new variables.



Quadratic Forms

Results. From (3), we have the following results



Quadratic Forms

Results. From (3), we have the following results

① There exists a diagonal matrix $\mathbf{D}$ $n \times n$ such that

$$Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{B}^\top \mathbf{A} \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{D} \mathbf{y} = d_1 y_1^2 + \cdots + d_n y_n^2,$$



Quadratic Forms

Results. From (3), we have the following results

- ① There exists a diagonal matrix $\mathbf{D}$ $n \times n$ such that

$$Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{B}^\top \mathbf{A} \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{D} \mathbf{y} = d_1 y_1^2 + \cdots + d_n y_n^2,$$

where $d_1, \dots, d_n$ are the diagonal elements of $\mathbf{D}$.



Quadratic Forms

Results. From (3), we have the following results

- ① There exists a diagonal matrix $\mathbf{D}$ $n \times n$ such that

$$Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{B}^\top \mathbf{A} \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{D} \mathbf{y} = d_1 y_1^2 + \cdots + d_n y_n^2,$$

where $d_1, \dots, d_n$ are the diagonal elements of $\mathbf{D}$.

- ② If $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite, there exists a transformation $\mathbf{y} = \mathbf{B}^{-1} \mathbf{x}$ such that



Quadratic Forms

Results. From (3), we have the following results

- ① There exists a diagonal matrix $\mathbf{D}$ $n \times n$ such that

$$Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{B}^\top \mathbf{A} \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{D} \mathbf{y} = d_1 y_1^2 + \cdots + d_n y_n^2,$$

where $d_1, \dots, d_n$ are the diagonal elements of $\mathbf{D}$.

- ② If $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite, there exists a transformation $\mathbf{y} = \mathbf{B}^{-1} \mathbf{x}$ such that

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{y} = y_1^2 + \cdots + y_n^2.$$



Quadratic Forms

Results. From (3), we have the following results

- ① There exists a diagonal matrix $\mathbf{D}$ $n \times n$ such that

$$Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{B}^\top \mathbf{A} \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{D} \mathbf{y} = d_1 y_1^2 + \cdots + d_n y_n^2,$$

where $d_1, \dots, d_n$ are the diagonal elements of $\mathbf{D}$.

- ② If $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is positive definite, there exists a transformation $\mathbf{y} = \mathbf{B}^{-1} \mathbf{x}$ such that

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{y} = y_1^2 + \cdots + y_n^2.$$



Quadratic Forms

Theorem 1. Let $\mathbf{A}$, $n \times n$ be a positive semidefinite matrix. Then



Quadratic Forms

Theorem 1. Let $\mathbf{A}$, $n \times n$ be a positive semidefinite matrix. Then

$$\textcircled{1} \quad r(\mathbf{A}) \leq n;$$



Quadratic Forms

Theorem 1. Let $\mathbf{A}$, $n \times n$ be a positive semidefinite matrix. Then

- ① $r(\mathbf{A}) \leq n$;
- ② $a_{ii} \geq 0$, $\forall i = 1, \dots, n$ and if $a_{jj} = 0$, every element of the j th row and j th column is zero;



Quadratic Forms

Theorem 1. Let $\mathbf{A}$, $n \times n$ be a positive semidefinite matrix. Then

- ① $r(\mathbf{A}) \leq n$;
- ② $a_{ii} \geq 0$, $\forall i = 1, \dots, n$ and if $a_{jj} = 0$, every element of the j th row and j th column is zero;
- ③ $\mathbf{P}^T \mathbf{A} \mathbf{P}$ is positive semidefinite, for every $n \times n$ matrix $\mathbf{P}$.



Quadratic Forms

Theorem 1. Let $\mathbf{A}$, $n \times n$ be a positive semidefinite matrix. Then

- ① $r(\mathbf{A}) \leq n$;
- ② $a_{ii} \geq 0$, $\forall i = 1, \dots, n$ and if $a_{jj} = 0$, every element of the j th row and j th column is zero;
- ③ $\mathbf{P}^\top \mathbf{A} \mathbf{P}$ is positive semidefinite, for every $n \times n$ matrix $\mathbf{P}$.



Quadratic Forms

Theorem 2. Let $\mathbf{A}$, $n \times n$ be a positive definite matrix. Then



Quadratic Forms

Theorem 2. Let $\mathbf{A}$, $n \times n$ be a positive definite matrix. Then

① $r(\mathbf{A}) = n;$



Quadratic Forms

Theorem 2. Let $\mathbf{A}$, $n \times n$ be a positive definite matrix. Then

① $r(\mathbf{A}) = n$;

② $a_{ii} > 0, \forall i = 1, \dots, n$;



Quadratic Forms

Theorem 2. Let $\mathbf{A}$, $n \times n$ be a positive definite matrix. Then

- ① $r(\mathbf{A}) = n$;
- ② $a_{ii} > 0$, $\forall i = 1, \dots, n$;
- ③ $\mathbf{P}^\top \mathbf{A} \mathbf{P}$ is positive definite, for any nonsingular $n \times n$ matrix $\mathbf{P}$.



Quadratic Forms

Theorem 2. Let $\mathbf{A}$, $n \times n$ be a positive definite matrix. Then

- ① $r(\mathbf{A}) = n$;
- ② $a_{ii} > 0$, $\forall i = 1, \dots, n$;
- ③ $\mathbf{P}^\top \mathbf{A} \mathbf{P}$ is positive definite, for any nonsingular $n \times n$ matrix $\mathbf{P}$.



Quadratic Forms

Theorem 3. An $n \times n$ matrix $\mathbf{A}$ is positive semidefinite if and only if



Quadratic Forms

Theorem 3. An $n \times n$ matrix $\mathbf{A}$ is positive semidefinite if and only if

- ① There exists $\mathbf{B}$, $n \times n$, such that $\mathbf{B}^\top \mathbf{B} = \mathbf{A}$;



Quadratic Forms

Theorem 3. An $n \times n$ matrix $\mathbf{A}$ is positive semidefinite if and only if

- ① There exists $\mathbf{B}$, $n \times n$, such that $\mathbf{B}^\top \mathbf{B} = \mathbf{A}$;
- ② The characteristic roots of $\mathbf{A}$ are nonnegative.



Quadratic Forms

Theorem 3. An $n \times n$ matrix $\mathbf{A}$ is positive semidefinite if and only if

- ① There exists $\mathbf{B}$, $n \times n$, such that $\mathbf{B}^\top \mathbf{B} = \mathbf{A}$;
- ② The characteristic roots of $\mathbf{A}$ are nonnegative.



Quadratic Forms

Theorem 4. An $n \times n$ matrix $\mathbf{A}$ is positive definite if and only if



Quadratic Forms

Theorem 4. An $n \times n$ matrix $\mathbf{A}$ is positive definite if and only if

- ① There exists $\mathbf{B}$, $n \times n$, $r(\mathbf{B}) = n$ such that $\mathbf{B}^\top \mathbf{B} = \mathbf{A}$;



Quadratic Forms

Theorem 4. An $n \times n$ matrix $\mathbf{A}$ is positive definite if and only if

- ① There exists $\mathbf{B}$, $n \times n$, $r(\mathbf{B}) = n$ such that $\mathbf{B}^\top \mathbf{B} = \mathbf{A}$;
- ② The characteristic roots of $\mathbf{A}$ are all positive;



Quadratic Forms

Theorem 4. An $n \times n$ matrix $\mathbf{A}$ is positive definite if and only if

- ① There exists $\mathbf{B}$, $n \times n$, $r(\mathbf{B}) = n$ such that $\mathbf{B}^\top \mathbf{B} = \mathbf{A}$;
- ② The characteristic roots of $\mathbf{A}$ are all positive;
- ③ And

$$a_{11} > 0, \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0, \dots, \det(\mathbf{A}) > 0.$$

Quadratic Forms

Theorem 4. An $n \times n$ matrix $\mathbf{A}$ is positive definite if and only if

- ① There exists $\mathbf{B}$, $n \times n$, $r(\mathbf{B}) = n$ such that $\mathbf{B}^\top \mathbf{B} = \mathbf{A}$;
- ② The characteristic roots of $\mathbf{A}$ are all positive;
- ③ And

$$a_{11} > 0, \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0, \dots, \det(\mathbf{A}) > 0.$$



Matrix Derivatives

Let $\mathbf{x} = (x_1, \dots, x_n)^\top$ and let $f(\mathbf{x})$ be a real-valued function of $x_1, \dots, x_n$.

The derivative of $f(\mathbf{x})$ with respect to $\mathbf{x}$, which represents the gradient of f , is



Matrix Derivatives

Let $\mathbf{x} = (x_1, \dots, x_n)^\top$ and let $f(\mathbf{x})$ be a real-valued function of $x_1, \dots, x_n$. The derivative of $f(\mathbf{x})$ with respect to $\mathbf{x}$, which represents the gradient of f , is

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \dots \quad \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^\top.$$



Matrix Derivatives

Let $\mathbf{x} = (x_1, \dots, x_n)^\top$ and let $f(\mathbf{x})$ be a real-valued function of $x_1, \dots, x_n$. The derivative of $f(\mathbf{x})$ with respect to $\mathbf{x}$, which represents the gradient of f , is

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \dots \quad \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^\top.$$



Matrix Derivatives

Theorem 1. If $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a vector of constants



Matrix Derivatives

Theorem 1. If $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a vector of constants and $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{a}$ (linear),



Matrix Derivatives

Theorem 1. If $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a vector of constants and $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{a}$ (linear), then

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}.$$



Matrix Derivatives

Theorem 1. If $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a vector of constants and $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{a}$ (linear), then

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}.$$

Theorem 2.



Matrix Derivatives

Theorem 1. If $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a vector of constants and $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{a}$ (linear), then

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}.$$

Theorem 2. If $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$ (quadratic form),



Matrix Derivatives

Theorem 1. If $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a vector of constants and $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{a}$ (linear), then

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}.$$

Theorem 2. If $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$ (quadratic form), $\mathbf{A}$ is a matrix of constants,



Matrix Derivatives

Theorem 1. If $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a vector of constants and $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{a}$ (linear), then

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}.$$

Theorem 2. If $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$ (quadratic form), $\mathbf{A}$ is a matrix of constants, then

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}.$$



Matrix Derivatives

Theorem 1. If $\mathbf{a} = (a_1, \dots, a_n)^\top$ is a vector of constants and $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{a}$ (linear), then

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}.$$

Theorem 2. If $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$ (quadratic form), $\mathbf{A}$ is a matrix of constants, then

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}.$$



Decompositions

A symmetric matrix **A** can be factored as

$$\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^T,$$



Decompositions

A symmetric matrix **A** can be factored as

$$\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^T,$$

where **L** is a lower triangular matrix, $\mathbf{L}^T$ is an upper triangular matrix and **D** is a diagonal matrix.



Decompositions

A symmetric matrix **A** can be factored as

$$\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^{\top},$$

where **L** is a lower triangular matrix, $\mathbf{L}^{\top}$ is an upper triangular matrix and **D** is a diagonal matrix.



Decompositions

Let a matrix $\mathbf{A}$, $p \times p$, be positive definite. Then there exists a unique upper triangular matrix $\mathbf{T}$, of rank p ,



Decompositions

Let a matrix $\mathbf{A}$, $p \times p$, be positive definite. Then there exists a unique upper triangular matrix $\mathbf{T}$, of rank p , such that

$$\mathbf{A} = \mathbf{T}^\top \mathbf{T}$$



Decompositions

Let a matrix $\mathbf{A}$, $p \times p$, be positive definite. Then there exists a unique upper triangular matrix $\mathbf{T}$, of rank p , such that

$$\mathbf{A} = \mathbf{T}^T \mathbf{T}$$

and $t_{ii} > 0$, $i = 1, \dots, p$.



Decompositions

Let a matrix $\mathbf{A}$, $p \times p$, be positive definite. Then there exists a unique upper triangular matrix $\mathbf{T}$, of rank p , such that

$$\mathbf{A} = \mathbf{T}^T \mathbf{T}$$

and $t_{ii} > 0$, $i = 1, \dots, p$. For the determination of $\mathbf{T}$, see Graybill (1983), for example.



Decompositions

Let a matrix $\mathbf{A}$, $p \times p$, be positive definite. Then there exists a unique upper triangular matrix $\mathbf{T}$, of rank p , such that

$$\mathbf{A} = \mathbf{T}^T \mathbf{T}$$

and $t_{ii} > 0$, $i = 1, \dots, p$. For the determination of $\mathbf{T}$, see Graybill (1983), for example.



Decompositions

Cholesky Decomposition. A matrix $\mathbf{A}$ is positive definite if and only if there exists a nonsingular $\mathbf{W}$



Decompositions

Cholesky Decomposition. A matrix **A** is positive definite if and only if there exists a nonsingular **W** such that

$$\mathbf{A} = \mathbf{W}\mathbf{W}^{\top}.$$



Decompositions

Cholesky Decomposition. A matrix **A** is positive definite if and only if there exists a nonsingular **W** such that

$$\mathbf{A} = \mathbf{W}\mathbf{W}^{\top}.$$



Decompositions

Spectral Decomposition. Any symmetric matrix $\mathbf{A}$, $p \times p$,



Decompositions

Spectral Decomposition. Any symmetric matrix $\mathbf{A}$, $p \times p$, can be written as

$$\mathbf{A} = \mathbf{\Gamma} \mathbf{D} \mathbf{\Gamma}^T,$$



Decompositions

Spectral Decomposition. Any symmetric matrix $\mathbf{A}$, $p \times p$, can be written as

$$\mathbf{A} = \mathbf{\Gamma} \mathbf{D} \mathbf{\Gamma}^T,$$

where $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_p)$, $\mathbf{\Gamma} = (\mathbf{x}_1 \cdots \mathbf{x}_p)$, $\lambda_1, \dots, \lambda_p$ are the characteristic roots of $\mathbf{A}$ and $\mathbf{x}_1, \dots, \mathbf{x}_p$ are the corresponding normalized characteristic vectors, that is, $\mathbf{x}_i^T \mathbf{x}_i = 1$, $i = 1, \dots, p$.



Decompositions

Spectral Decomposition. Any symmetric matrix $\mathbf{A}$, $p \times p$, can be written as

$$\mathbf{A} = \mathbf{\Gamma} \mathbf{D} \mathbf{\Gamma}^\top,$$

where $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_p)$, $\mathbf{\Gamma} = (\mathbf{x}_1 \cdots \mathbf{x}_p)$, $\lambda_1, \dots, \lambda_p$ are the characteristic roots of $\mathbf{A}$ and $\mathbf{x}_1, \dots, \mathbf{x}_p$ are the corresponding normalized characteristic vectors, that is, $\mathbf{x}_i^\top \mathbf{x}_i = 1$, $i = 1, \dots, p$.



Decompositions

It can be verified that

$$\mathbf{A} = \mathbf{\Gamma} \mathbf{D} \mathbf{\Gamma}^T = \lambda_1 \mathbf{x}_1 \mathbf{x}_1^T + \cdots + \lambda_p \mathbf{x}_p \mathbf{x}_p^T.$$



Decompositions

It can be verified that

$$\mathbf{A} = \mathbf{\Gamma} \mathbf{D} \mathbf{\Gamma}^T = \lambda_1 \mathbf{x}_1 \mathbf{x}_1^T + \cdots + \lambda_p \mathbf{x}_p \mathbf{x}_p^T.$$

Remark: If $\mathbf{x}_i$ and $\mathbf{x}_j$ are two characteristic vectors associated with different characteristic roots, then $\mathbf{x}_i^T \mathbf{x}_j = 0$.



Decompositions

It can be verified that

$$\mathbf{A} = \mathbf{\Gamma} \mathbf{D} \mathbf{\Gamma}^T = \lambda_1 \mathbf{x}_1 \mathbf{x}_1^T + \cdots + \lambda_p \mathbf{x}_p \mathbf{x}_p^T.$$

Remark: If $\mathbf{x}_i$ and $\mathbf{x}_j$ are two characteristic vectors associated with different characteristic roots, then $\mathbf{x}_i^T \mathbf{x}_j = 0$. We say that $\mathbf{x}_i$ and $\mathbf{x}_j$ are orthogonal.



Decompositions

It can be verified that

$$\mathbf{A} = \mathbf{\Gamma} \mathbf{D} \mathbf{\Gamma}^\top = \lambda_1 \mathbf{x}_1 \mathbf{x}_1^\top + \cdots + \lambda_p \mathbf{x}_p \mathbf{x}_p^\top.$$

Remark: If $\mathbf{x}_i$ and $\mathbf{x}_j$ are two characteristic vectors associated with different characteristic roots, then $\mathbf{x}_i^\top \mathbf{x}_j = 0$. We say that $\mathbf{x}_i$ and $\mathbf{x}_j$ are orthogonal.



Computational Cost

Consider the matrix multiplication

$$\mathbf{A}_{m \times p} \mathbf{B}_{p \times n} = \mathbf{C}_{m \times n}.$$

To compute each element of $\mathbf{C}$, we need p multiplications and $p - 1$ additions.



Computational Cost

Consider the matrix multiplication

$$\mathbf{A}_{m \times p} \mathbf{B}_{p \times n} = \mathbf{C}_{m \times n}.$$

To compute each element of $\mathbf{C}$, we need p multiplications and $p - 1$ additions. Since $\mathbf{C}$ has mn elements, the computational cost is of order

$$\mathcal{O}(mpn).$$



Computational Cost

Consider the matrix multiplication

$$\mathbf{A}_{m \times p} \mathbf{B}_{p \times n} = \mathbf{C}_{m \times n}.$$

To compute each element of $\mathbf{C}$, we need p multiplications and $p - 1$ additions. Since $\mathbf{C}$ has mn elements, the computational cost is of order

$$\mathcal{O}(mpn).$$

In particular, for two $n \times n$ matrices, the computational cost is of order $\mathcal{O}(n^3)$.



Computational Cost

Consider the matrix multiplication

$$\mathbf{A}_{m \times p} \mathbf{B}_{p \times n} = \mathbf{C}_{m \times n}.$$

To compute each element of $\mathbf{C}$, we need p multiplications and $p - 1$ additions. Since $\mathbf{C}$ has mn elements, the computational cost is of order

$$\mathcal{O}(mpn).$$

In particular, for two $n \times n$ matrices, the computational cost is of order $\mathcal{O}(n^3)$.



Computational Cost

Consider the inverse of a nonsingular matrix $\mathbf{A}$, $n \times n$. Using conventional matrix factorization methods, the computational complexity of computing $\mathbf{A}^{-1}$ is $\mathcal{O}(n^3)$.



Computational Cost

Consider the inverse of a nonsingular matrix $\mathbf{A}$, $n \times n$. Using conventional matrix factorization methods, the computational complexity of computing $\mathbf{A}^{-1}$ is $\mathcal{O}(n^3)$.

Thus, for an $n \times n$ matrix, both matrix multiplication and matrix inversion have computational complexity of order $\mathcal{O}(n^3)$.



Computational Cost

Consider the inverse of a nonsingular matrix $\mathbf{A}$, $n \times n$. Using conventional matrix factorization methods, the computational complexity of computing $\mathbf{A}^{-1}$ is $\mathcal{O}(n^3)$.

Thus, for an $n \times n$ matrix, both matrix multiplication and matrix inversion have computational complexity of order $\mathcal{O}(n^3)$.



Computational Cost

Example. Let $\mathbf{A}$ be an $n \times n$ matrix. Consider $(\mathbf{A}^\top \mathbf{A})^{-1}$. The multiplication $\mathbf{A}^\top \mathbf{A}$ has computational complexity $\mathcal{O}(n^3)$.



Computational Cost

Example. Let $\mathbf{A}$ be an $n \times n$ matrix. Consider $(\mathbf{A}^\top \mathbf{A})^{-1}$. The multiplication $\mathbf{A}^\top \mathbf{A}$ has computational complexity $\mathcal{O}(n^3)$. The inversion of the resulting $n \times n$ matrix has computational complexity $\mathcal{O}(n^3)$.



Computational Cost

Example. Let $\mathbf{A}$ be an $n \times n$ matrix. Consider $(\mathbf{A}^\top \mathbf{A})^{-1}$. The multiplication $\mathbf{A}^\top \mathbf{A}$ has computational complexity $\mathcal{O}(n^3)$. The inversion of the resulting $n \times n$ matrix has computational complexity $\mathcal{O}(n^3)$.

Observation: The computational complexity of the complete operation is not expressed as $\mathcal{O}(n^3) + \mathcal{O}(n^3)$,



Computational Cost

Example. Let $\mathbf{A}$ be an $n \times n$ matrix. Consider $(\mathbf{A}^\top \mathbf{A})^{-1}$. The multiplication $\mathbf{A}^\top \mathbf{A}$ has computational complexity $\mathcal{O}(n^3)$. The inversion of the resulting $n \times n$ matrix has computational complexity $\mathcal{O}(n^3)$.

Observation: The computational complexity of the complete operation is not expressed as $\mathcal{O}(n^3) + \mathcal{O}(n^3)$, but simply as $\mathcal{O}(n^3)$.



Computational Cost

Example. Let $\mathbf{A}$ be an $n \times n$ matrix. Consider $(\mathbf{A}^\top \mathbf{A})^{-1}$. The multiplication $\mathbf{A}^\top \mathbf{A}$ has computational complexity $\mathcal{O}(n^3)$. The inversion of the resulting $n \times n$ matrix has computational complexity $\mathcal{O}(n^3)$.

Observation: The computational complexity of the complete operation is not expressed as $\mathcal{O}(n^3) + \mathcal{O}(n^3)$, but simply as $\mathcal{O}(n^3)$.



Expectation of a quadratic form

Let $\mathbf{v}$ be a random vector of length v , with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{C}$, and let $\mathbf{W}$ be a matrix of dimension $v \times v$, then

$$\mathbb{E} \left(\mathbf{v}^\top \mathbf{W} \mathbf{v} \right) = \text{tr}(\mathbf{W} \mathbf{C}) + \mathbf{m}^\top \mathbf{W} \mathbf{m},$$

where $\text{tr}(\cdot)$ is the trace operator.



Expectation of a quadratic form

Let $\mathbf{v}$ be a random vector of length v , with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{C}$, and let $\mathbf{W}$ be a matrix of dimension $v \times v$, then

$$\mathbb{E} \left(\mathbf{v}^\top \mathbf{W} \mathbf{v} \right) = \text{tr}(\mathbf{W} \mathbf{C}) + \mathbf{m}^\top \mathbf{W} \mathbf{m},$$

where $\text{tr}(\cdot)$ is the trace operator.



Outline

- 1 Operations
- 2 Matrix Results
- 3 Estimation**
- 4 References



Properties of Estimators

An estimator $\hat{\theta}$ of an unknown parameter θ is unbiased if

$$\mathbb{E}(\hat{\theta}) = \theta.$$



Properties of Estimators

An estimator $\hat{\theta}$ of an unknown parameter θ is unbiased if

$$\mathbb{E}(\hat{\theta}) = \theta.$$

An estimator $\hat{\theta}$ is a consistent estimator of θ



Properties of Estimators

An estimator $\hat{\theta}$ of an unknown parameter θ is unbiased if

$$\mathbb{E}(\hat{\theta}) = \theta.$$

An estimator $\hat{\theta}$ is a consistent estimator of θ if

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\hat{\theta} - \theta| \geq \varepsilon) = 0, \text{ for any } \varepsilon > 0.$$



Properties of Estimators

An estimator $\hat{\theta}$ of an unknown parameter θ is unbiased if

$$\mathbb{E}(\hat{\theta}) = \theta.$$

An estimator $\hat{\theta}$ is a consistent estimator of θ if

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\hat{\theta} - \theta| \geq \varepsilon) = 0, \text{ for any } \varepsilon > 0.$$



Properties of Estimators

An estimator $\hat{\theta}$ is a sufficient estimator of θ if the conditional distribution of the observations given $\hat{\theta}$ does not depend on the parameter θ .

An unbiased estimator $\hat{\theta}$ is a minimum-variance unbiased estimator of θ



Properties of Estimators

An estimator $\hat{\theta}$ is a sufficient estimator of θ if the conditional distribution of the observations given $\hat{\theta}$ does not depend on the parameter θ .

An unbiased estimator $\hat{\theta}$ is a minimum-variance unbiased estimator of θ if for any other unbiased estimator θ^* :

$$\text{Var}(\hat{\theta}) \leq \text{Var}(\theta^*).$$



Properties of Estimators

An estimator $\hat{\theta}$ is a sufficient estimator of θ if the conditional distribution of the observations given $\hat{\theta}$ does not depend on the parameter θ .

An unbiased estimator $\hat{\theta}$ is a minimum-variance unbiased estimator of θ if for any other unbiased estimator θ^* :

$$\text{Var}(\hat{\theta}) \leq \text{Var}(\theta^*).$$



Maximum Likelihood Estimator

Let Y be a random variable with probability density function (PDF) given by $f(y; \theta)$. Given independent observations $Y_1, \dots, Y_n$,



Maximum Likelihood Estimator

Let Y be a random variable with probability density function (PDF) given by $f(y; \theta)$. Given independent observations $Y_1, \dots, Y_n$, the joint PDF of these observations is

$$f(y_1, \dots, y_n) = \prod_{i=1}^n f(y_i; \theta). \quad (4)$$



Maximum Likelihood Estimator

Let Y be a random variable with probability density function (PDF) given by $f(y; \theta)$. Given independent observations $Y_1, \dots, Y_n$, the joint PDF of these observations is

$$f(y_1, \dots, y_n) = \prod_{i=1}^n f(y_i; \theta). \quad (4)$$



Maximum Likelihood Estimator

When the joint PDF (4) is viewed as a function of the parameter θ , it is called the likelihood function and written as

$$L(\theta) = \prod_{i=1}^n f(y_i; \theta).$$



Maximum Likelihood Estimator

When the joint PDF (4) is viewed as a function of the parameter θ , it is called the likelihood function and written as

$$L(\theta) = \prod_{i=1}^n f(y_i; \theta).$$

The maximum likelihood estimator (MLE) of θ is the value of the unknown parameter that maximizes $L(\theta)$.



Maximum Likelihood Estimator

When the joint PDF (4) is viewed as a function of the parameter θ , it is called the likelihood function and written as

$$L(\theta) = \prod_{i=1}^n f(y_i; \theta).$$

The maximum likelihood estimator (MLE) of θ is the value of the unknown parameter that maximizes $L(\theta)$. Under suitable regularity conditions, the MLE is consistent.



Maximum Likelihood Estimator

When the joint PDF (4) is viewed as a function of the parameter θ , it is called the likelihood function and written as

$$L(\theta) = \prod_{i=1}^n f(y_i; \theta).$$

The maximum likelihood estimator (MLE) of θ is the value of the unknown parameter that maximizes $L(\theta)$. Under suitable regularity conditions, the MLE is consistent.



Least Squares Estimator

Suppose that a sample observation has the following form:

$$Y_i = f_i(\theta) + \varepsilon_i, \quad i = 1, \dots, n,$$



Least Squares Estimator

Suppose that a sample observation has the following form:

$$Y_i = f_i(\theta) + \varepsilon_i, \quad i = 1, \dots, n,$$

where $f_i(\theta)$ is a known function of θ and ε_i is a random variable, usually assumed to have zero mean.



Least Squares Estimator

Suppose that a sample observation has the following form:

$$Y_i = f_i(\theta) + \varepsilon_i, \quad i = 1, \dots, n,$$

where $f_i(\theta)$ is a known function of θ and ε_i is a random variable, usually assumed to have zero mean.



Least Squares Estimator

In the least squares method, for a given random sample, the sum of squares

$$Q = \sum_{i=1}^n [Y_i - f_i(\theta)]^2$$



Least Squares Estimator

In the least squares method, for a given random sample, the sum of squares

$$Q = \sum_{i=1}^n [Y_i - f_i(\theta)]^2$$

is considered as a function of the parameter θ .

Least Squares Estimator

In the least squares method, for a given random sample, the sum of squares

$$Q = \sum_{i=1}^n [Y_i - f_i(\theta)]^2$$

is considered as a function of the parameter θ . The least squares estimator (LSE) of θ is obtained by minimizing Q with respect to θ .



Least Squares Estimator

In the least squares method, for a given random sample, the sum of squares

$$Q = \sum_{i=1}^n [Y_i - f_i(\theta)]^2$$

is considered as a function of the parameter θ . The least squares estimator (LSE) of θ is obtained by minimizing Q with respect to θ . In many cases, the LSE is unbiased and consistent.



Least Squares Estimator

In the least squares method, for a given random sample, the sum of squares

$$Q = \sum_{i=1}^n [Y_i - f_i(\theta)]^2$$

is considered as a function of the parameter θ . The least squares estimator (LSE) of θ is obtained by minimizing Q with respect to θ . In many cases, the LSE is unbiased and consistent.



Outline

- 1 Operations
- 2 Matrix Results
- 3 Estimation
- 4 References



References

Graybill, F. A. (1983). *Matrices with applications in Statistics*. Wadsworth International Group, Belmont, 2 edition.



Thank you!

✉ tiago.magalhaes@ufjf.br

📄 ufjf.br/tiago_magalhaes

🌐 Department of Statistics, Room 319

H Á 112 ANOS... NOSSA HISTÓRIA
TREMULA AO VENTO

