Symmetric Matrices and Quadratic Forms

Linear Algebra

Department of Computer Engineering

Sharif University of Technology

Hamid R. Rabiee rabiee@sharif.edu

Maryam Ramezani maryam.ramezani@sharif.edu

Symmetric Matrix

- A symmetric matrix is a matrix A such that $A^T = A$. Such a matrix is necessarily square. Its main diagonal entries are arbitrary, but its other entries occur in pairs – on opposite sides of the main diagonal.

Symmetric: $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{bmatrix}, \quad \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$

Nonsymmetric: $\begin{bmatrix} 1 & -3 \\ 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -4 & 0 \\ -6 & 1 & -4 \\ 0 & -6 & 1 \end{bmatrix}, \quad \begin{bmatrix} 5 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$

- A quadratic form is any homogeneous polynomial of degree two in any number of variables. In this situation, **homogeneous** means that all the terms are of degree two. For example, the expression $7x_1x_2 + 3x_2x_4$ is homogeneous, but the expression $x_1 - 3x_1x_2$ is not. The square of the distance between two points in an inner–product space is a quadratic form. Quadratic forms were introduced by Hermite, and 70 years later they turned out to be essential in the theory of quantum mechanics! The formal definition follows.

Quadratic Form

- Given a square matrix $A \in \mathbb{R}^{n \times n}$ and a vector $x \in \mathbb{R}^n$, the scalar value $x^T A x$ is called a **quadratic form**.

$$x^T A x = \sum_{i=1}^{n} x_i (Ax)_i = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} A_{ij} x_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_i x_j$$

- ❑ A quadratic form on $\mathbb{R}^n$ is a function Q defined on $\mathbb{R}^n$ whose value at a vector x in $\mathbb{R}^n$ can be computed by an expression of the form $Q(x) = x^T A x$, where A is an $n \times n$ symmetric matrix. The matrix A is called the **matrix of the quadratic form**.

Quadratic Form

Definition

- Suppose $\mathcal{X}$ is a vector space over $\mathbb{R}$. Then a function $Q: \mathcal{X} \to \mathbb{R}$ is called a quadratic form if there exists a bilinear form $f: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that:

$$Q(x) = f(x, x) \text{ for all } x \in \mathcal{X}$$

Example

Simplest example of a nonzero quadratic form is …

Quadratic Form

Example

Without cross-product term:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$$

With cross-product term:

$$A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$$

Tip

- Quadratic forms are easier to use when they have no cross-product terms; that is, when the matrix of the quadratic form is a diagonal matrix.

Quadratic Form

Example

For x in $\mathbb{R}^3$, let $Q(x) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$. Write this quadratic form as x^TAx.

Change of Variable in QF

- If x represents a variable in $\mathbb{R}^n$, then a **change of variable** is an equation of the form:

$$\boldsymbol{x} = P\boldsymbol{y} \qquad \text{or equivalently,} \qquad \boldsymbol{y} = P^{-1}\boldsymbol{x}$$

where P is an **invertible matrix** and $\boldsymbol{y}$ is a new variable vector in $\mathbb{R}^n$.

Note

$\boldsymbol{y}$ can be regarded as the **coordinate vector** of $\boldsymbol{x}$ relative to the basis of $\mathbb{R}^n$ determined by the columns of P.

Change of Variable in QF

- If the change of variable is made in a quadratic form $x^T Ax$, then

$$x^T A\boldsymbol{x} = (P\boldsymbol{y})^T A(P\boldsymbol{y}) = \boldsymbol{y}^T P^T AP\boldsymbol{y} = \boldsymbol{y}^T (P^T AP)\boldsymbol{y}$$

- The new matrix of the quadratic form is $P^T AP$.
- A is symmetric, so there is an orthogonal matrix P such that $P^T AP$ is a diagonal matrix D.
- Then the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$ becomes $\boldsymbol{y}^T D\boldsymbol{y}$. There is no cross-product.

Quadratic Form

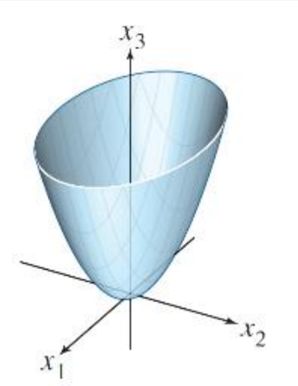
- If A and B are $n \times n$ real matrices connected by the relation

$$B = \frac{1}{2}\left(A + A^T\right)$$

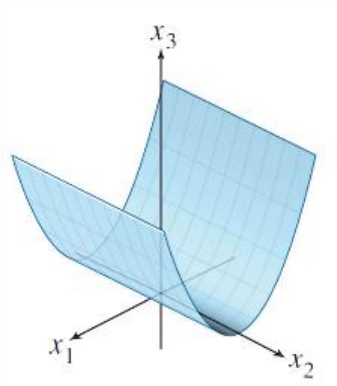
then the corresponding quadratic forms of A and B are identical, and B is symmetric

- When A is an $n \times n$ matrix, the quadratic form $Q(x) = x^T A x$ is a real-valued function with domain $\mathbb{R}^n$.

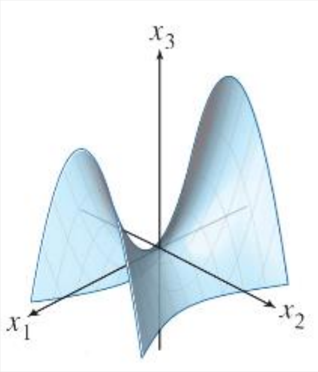
point (x_1, x_2, z) where $z = Q(x)$



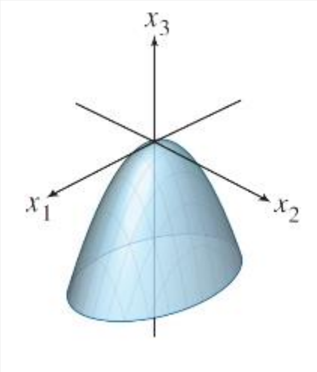
(a) $z = 3x_1^2 + 7x_2^2$



(b) $z = 3x_1^2$



(c) $z = 3x_1^2 - 7x_2^2$



(d) $z = -3x_1^2 - 7x_2^2$

Classifying Quadratic Forms

- A symmetric matrix $A \in \mathbb{S}^n$ is **positive definite (PD)** if for all non zero vectors $A \in \mathbb{R}^n$, $x^T A x > 0$. This is usually denoted $A \succ 0$, and often times the set of all positive definite matrices is denoted $\mathbb{S}^n_{++}$.
- A symmetric matrix $A \in \mathbb{S}^n$ is **positive semidefinite (PSD)** if for all vectors $x^T A x \geq 0$. This is written $A \succcurlyeq 0$, and the set of all positive semidefinite matrices is often denoted $\mathbb{S}^n_+$.
- Likewise, a symmetric matrix $A \in \mathbb{S}^n$ is **negative definite (ND)**, denoted $A \prec 0$ if for all non-zero $x \in \mathbb{R}^n$, $x^T A x < 0$.
- Similarly, a symmetric matrix $A \in \mathbb{S}^n$ is **negative semidefinite (NSD)**, denoted $A \preccurlyeq 0$ if for all $x \in \mathbb{R}^n$, $x^T A x \leq 0$.
- Finally, a symmetric matrix $A \in \mathbb{S}^n$ is **indefinite**, if it is neither positive semidefinite nor negative semidefinite; i.e., if there exists $x_1, x_2 \in \mathbb{R}^n$ such that $x_1^T A x_1 > 0$ and $x_2^T A x_2 < 0$.

Classifying Quadratic Forms

Definition $Q(x) = x^T A x$

A quadratic form Q is:

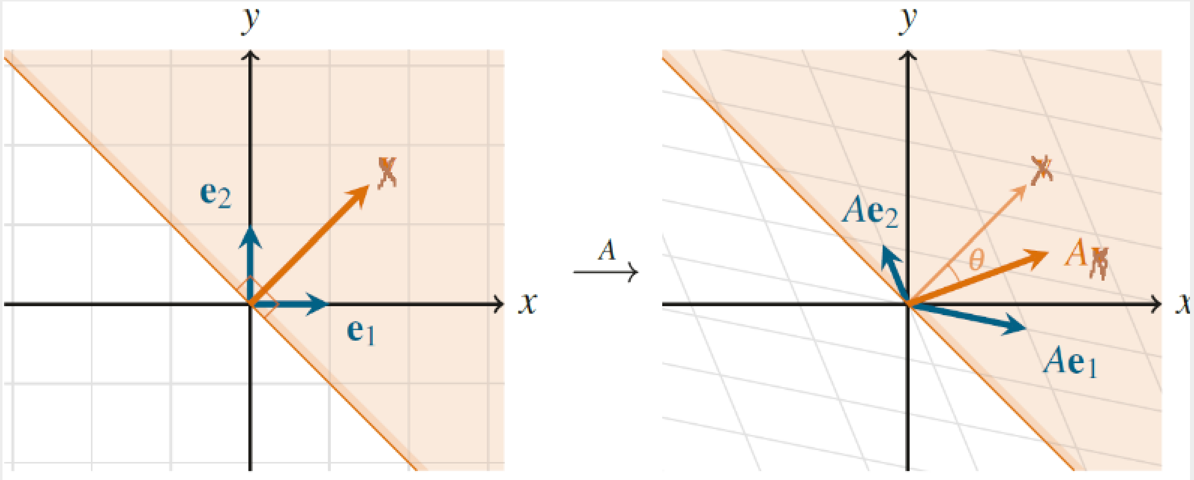
- **positive definite** if $Q(x) > 0$ for all $x \neq 0$;
- **negative definite** if $Q(x) < 0$ for all $x \neq 0$;
- **indefinite** if $Q(x)$ assumes both positive and negative values;
- **positive semidefinite** if $Q(x) \geq 0$ for all x;
- **negative semidefinite** if $Q(x) \leq 0$ for all x;

❑ For diagonal matrix $A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix} \Rightarrow x^T A x = a_1 x_1^2 + a_2 x_2^t + \cdots + a_n x_n^2$.

Geometric Interpretation

- $Q(x) = x^T A x$

- $\theta = \arccos(\frac{(Ax).x}{\|x\|\|Ax\|})$



Characterization of Positive Semidefinite Matrices

Suppose $A \in \mathcal{M}_n(\mathbb{F})$ is self-adjoint. The following are equivalent:

a) A is positive semidefinite.

b) All of the eigenvalues of A are non-negative.

c) There is a matrix $\mathrm{B} \in \mathcal{M}_n(\mathbb{F})$ such that $A = B^*B$, and

d) There is a diagonal matrix $D \in \mathcal{M}_n(\mathbb{R})$ with non-negative diagonal entries and a unitary matrix $U \in \mathcal{M}_n(\mathbb{F})$ such that $A = UDU^*$.

Characterization of Positive Definite Matrices

Suppose $A \in \mathcal{M}_n(\mathbb{F})$ is self-adjoint. The following are equivalent:

a) A is positive *definite*.

b) All of the eigenvalues of A are *strictly positive*.

c) There is an *invertible* matrix $\mathrm{B} \in \mathcal{M}_n(\mathbb{F})$ such that $A = B^*B$, and

d) There is a diagonal matrix $D \in \mathcal{M}_n(\mathbb{R})$ with *strictly positive* diagonal entries and a unitary matrix $U \in \mathcal{M}_n(\mathbb{F})$ such that $A = UDU^*$.

Quadratic Form

Theorem

Let A be an $n \times n$ symmetric matrix. Then a quadratic form $x^T Ax$ is:

- **positive definite** if and only if the eigenvalues of A are **all positive**;
- **negative definite** if and only if the eigenvalues of A are **all negative**;
- **indefinite** if and only if A has **both positive and negative** eigenvalues;

- How about semidefinite?

Positive Definite Matrices

- For a symmetric matrix the signs of the pivots are the signs of the eigenvalues.

$$number\ of\ positive\ pivots = number\ of\ positive\ eigenvalues$$

Important

A symmetric matrix $\boldsymbol{A}$ is to be **positive definite** if:

- all the eigenvalues are positive
- all the pivots are positive
- all the determinants are positive
- $x^T Ax > 0\ \forall x$ except $x = 0$

If any of the eigenvalues or pivots or determinants is zero, that matrix is called a **positive semidefinite** matrix.

Positive Definite Matrices

Five tests to see whether a matrix is positive definite or not:

1. $x^T A x > 0$ for all x (other than zero-vector)
2. If A is positive definite, $A = S^T S$ (S must have independent columns.)
3. All eigen values are greater than 0
4. Sylvester's Criterion: All upper left determinants must be > 0.
5. Every pivot must be > 0

Note

A positive definite matrix A has positive eigenvalues, positive pivots, positive determinants, and positive energy $v^T A v$ for every vector v. $A = S^T S$ is always positive definite if S has independent columns.

Positive Definite Matrices

For positive definite matrices we had:

- *If A is positive definite, $A = S^T S$ (S must have independent columns.)*

Theorem

If S is positive definite $S = A^T A$ (A must have independent columns): $A^T A$ is positive definite iff the columns of A are linearly independent.

- Proof?

Positive Definite Matrices

For positive definite matrices we had:

- *All eigen values are greater than 0*

Theorem

If a matrix is positive definite, then its eigenvalues are positive.

- ❑ Proof?

Theorem

If a matrix has positive eigenvalues, then it is positive definite.

- Proof?

Positive Definite Matrices

For positive definite matrices we had:

- *Sylvester's Criterion: All upper left determinants must be* > 0.

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

Theorem

If a matrix is positive definite, then it has positive determinant.

- Proof?

Theorem

Suppose $A \in \mathcal{M}_n$ is self-adjoint. Then A is positive definite if and only if, for all $1 \leq k \leq n$, the determinant of the top-left $k \times k$ block of A is strictly positive.

Sylvester's Criterion for Positive Semidefinite Matrices

- A principal minor of a square matrix is the determinant of a submatrix of A that is obtained by deleting some (or none) of its rows as well as the corresponding columns.
- A matrix is positive semidefinite if and only if all of its principal minors are non-negative.

$$B = \begin{bmatrix} a & b & c \\ \bar{b} & d & e \\ \bar{c} & \bar{e} & f \end{bmatrix}$$

are $a, d, f, \det(B)$ itself, as well as

$$det\left(\begin{bmatrix} a & b \\ \bar{b} & d \end{bmatrix}\right) = ad - |b|^2 \qquad det\left(\begin{bmatrix} a & c \\ \bar{c} & f \end{bmatrix}\right) = af - |c|^2 \qquad det\left(\begin{bmatrix} d & e \\ \bar{e} & f \end{bmatrix}\right) = df - |e|^2$$

Positive Definite Matrices

For positive definite matrices we had:

- *Every pivot must be* > 0.

  - Pivots are, in general, way easier to calculate than eigenvalues.
  - Just perform elimination and examine the diagonal terms.

Example

Is the following matrix positive definite matrix? $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

Note

Number of positive (negative) pivots = number of positive (negative) eigenvalues.

Pivots & Positive Definite Matrices

Theorem

If a matrix has positive pivots, then it is positive definite.

- Proof?

Properties

Important

- If A is positive definite, A^{-1} will also be positive definite.
- If A and B are positive definite matrices, $A + B$ will also be a positive definite matrix.
- Positive definite and negative definite matrices are always full rank, and hence, invertible.
- For $A \in \mathbb{R}^{m \times n}$ gram matrix is always positive semidefinite. Further, if $m \geq n$ (and we assume for convenience that A is full rank), then gram matrix is positive definite.

Properties

Important

Suppose $A, B \in \mathcal{M}_n$ are positive (semi)definite, $P \in \mathcal{M}_{n,m}$ is any matrix, and $c > 0$ is real scalar. Then

a) $A + B$ is positive (semi)definite.

b) cA is positive (semi)definite.

c) A^T is positive (semi)definite, and

d) P^*AP is positive semidefinite. Furthermore, if A is positive definite then P^*AP is positive definite if and only if $rank(P) = m$.

Cholesky Factorization

Important

Every positive definite matrix $A \epsilon \mathbb{R}^{n \times n}$ can be factored as

$$A = \mathbb{R}^T \mathbb{R}$$

where $\mathbb{R}$ is upper triangular with positive diagonal elements

- ❑ complexity of computing $\mathbb{R}$ is $(1/3)n^3$ flops
- ❑ $\mathbb{R}$ is called the $Cholesky\ factor$ of A
- ❑ can be interpreted as "square root" of a positive definite matrix
- ❑ gives a practical method for testing positive definiteness

Example

$$\begin{bmatrix} A_{11} & A_{1,2:n} \\ A_{2:n,1} & A_{2:n,2:n} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ R_{1,2:n}^T & R_{2:n,2:n}^T \end{bmatrix} \begin{bmatrix} R_{11} & R_{1,2:n} \\ 0 & R_{2:n,2:n} \end{bmatrix}$$

$$= \begin{bmatrix} R_{11}^2 & R_{11}R_{1,2:n} \\ R_{11}R_{1,2:n}^T & R_{1,2:n}^T R_{1,2:n} + R_{2:n,2:n}^T R_{2:n,2:n} \end{bmatrix}$$

1. compute first row of R:

$$R_{11} = \sqrt{A_{11}}, \qquad R_{1,2:n} = \frac{1}{R_{11}} A_{1,2:n} \qquad \boxed{A_{11} > 0}$$

if A is positive definite

2. compute 2, 2 block $R_{2:n,2:n}$ from

$$A_{2:n,2:n} - R_{1,2:n}^T R_{1,2:n} = R_{2:n,2:n}^T R_{2:n,2:n} = \boxed{A_{2:n,2:n} - \frac{1}{A_{11}} A_{2:n,1} A_{2:n,1}^T}$$

this is a Cholesky factorization of order $n-1$

Example

$$\begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & 0 \\ R_{12} & R_{22} & 0 \\ R_{13} & R_{23} & R_{33} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \end{bmatrix}$$

- first row of R

$$\begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 3 & R_{22} & 0 \\ -1 & R_{23} & R_{33} \end{bmatrix} \begin{bmatrix} 5 & 3 & -1 \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \end{bmatrix}$$

- second row of R

$$\begin{bmatrix} 18 & 0 \\ 0 & 11 \end{bmatrix} - \begin{bmatrix} 3 \\ -1 \end{bmatrix} [3 \quad -1] = \begin{bmatrix} R_{22} & 0 \\ R_{23} & R_{33} \end{bmatrix} \begin{bmatrix} R_{22} & R_{23} \\ 0 & R_{33} \end{bmatrix}$$

$$\begin{bmatrix} 9 & 3 \\ 3 & 10 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 1 & R_{33} \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & R_{33} \end{bmatrix}$$

- third column of R: $10 - 1 = R_{33}^2, i.e., R_{33} = 3$

Example

- Let $B = \{b_1, \dots, b_r\} \subset \mathbb{R}^m$ with $r = \text{rank}(A)$ be basis of range(A). Then each of the columns of $A = [a_1, a_2, \dots, a_n]$ can be expressed as linear combination of B:

$$a_i = b_1 c_{i1} + b_2 c_{i2} + \cdots + b_r c_{ir} = [b_1, \dots, b_r] \begin{bmatrix} c_{i1} \\ \vdots \\ c_{ir} \end{bmatrix},$$

for some coefficients $c_{ij} \epsilon \ \mathbb{R}$ with $i = 1, \dots, n, j = 1, \dots, r$.

Stacking these relations column by column →

$$[a_1, \dots, a_n] = [b_1, \dots, b_r] \begin{bmatrix} c_{11} & \cdots & c_{n1} \\ \vdots & & \vdots \\ c_{1r} & \cdots & c_{nr} \end{bmatrix}$$