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

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

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

Eigenvectors of a Symmetric Matrix

Theorem

Orthogonality of Eigenvectors of a Symmetric Matrix Corresponding to Distinct Eigenvalues. If A is symmetric, then any two eigenvectors from different eigenspace are orthogonal.

$$\left.\begin{array}{c} Av_1 = \lambda_1 v_1 \\ Av_2 = \lambda_2 v_2 \\ \lambda_1 \neq \lambda_2 \end{array}\right\} \Rightarrow \; {v_1}^T v_2 = 0$$

Proof?

Orthogonally Diagonalizable

Definition

A square matrix A is orthogonally diagonalizable if its eigenvectors are orthogonal

Orthogonally Diagonalizable

Theorem

An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is a symmetric matrix.

$(\Rightarrow)$:

$$A = A^T \;\Rightarrow A = Q\Lambda Q^T, \Lambda = diag\{\lambda_1, \cdots, \lambda_n\}$$

$(\Leftarrow)$:

$$A = A^T \Leftarrow A = Q\Lambda Q^T, \Lambda = diag\{\lambda_1, \cdots, \lambda_n\}, Q\ is\ orthogonal \Rightarrow Q^T = Q^{-1}$$

$$A^T = (Q\Lambda Q^{-1})^T = (Q\Lambda Q^T)^T = Q\Lambda^T Q^T = Q\Lambda Q^T = A$$

Theorem

All the **eigenvalues** of matrix A (a real symmetric matrix) are real.

Proof?

Relationship between eigenvalue and pivot signs

Theorem

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

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

- We know that determinant of matrix is product of pivots.
- We know that determinant of matrix is product of eigenvalues.

Conclusion: Spectral Theorem

The Spectral Theorem for Symmetric Matrices

An $n \times n$ symmetric matrix A has the following properties:

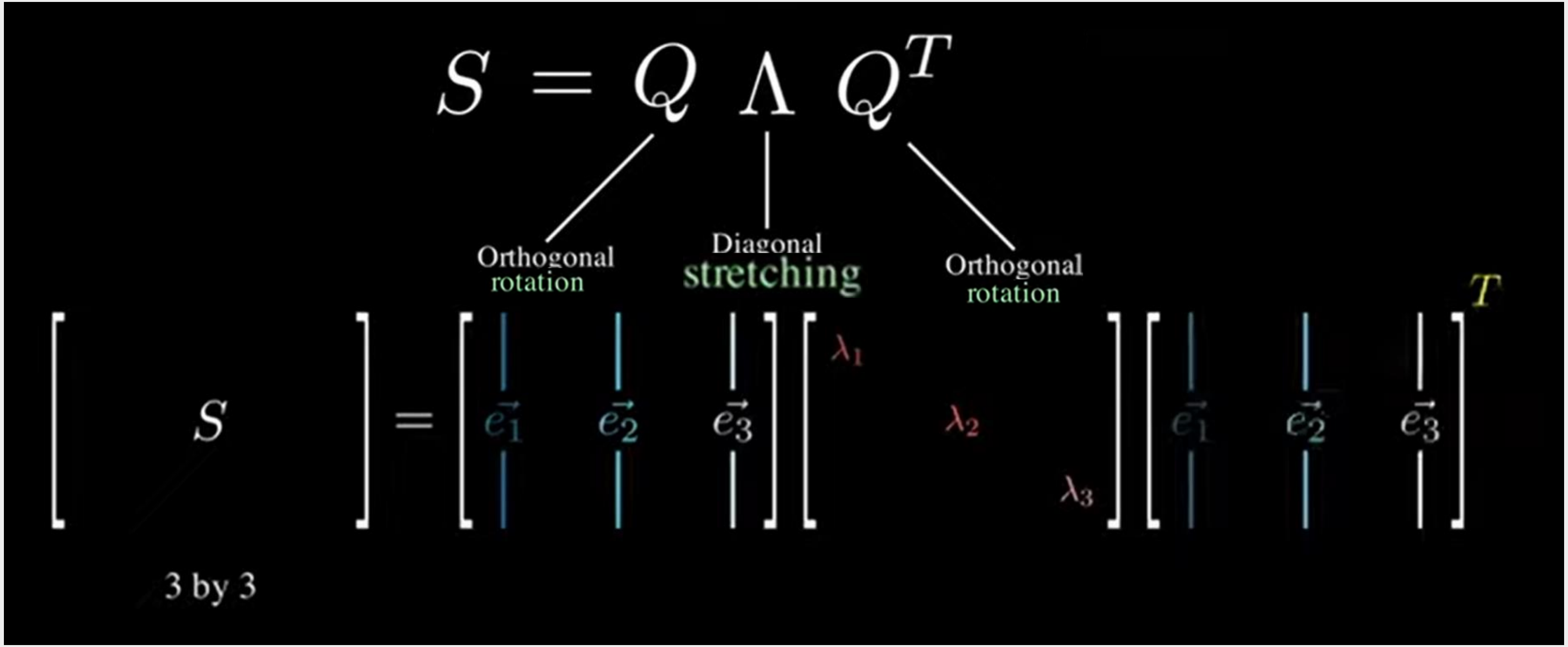
a. A has n real eigenvalues, counting multiplicities.

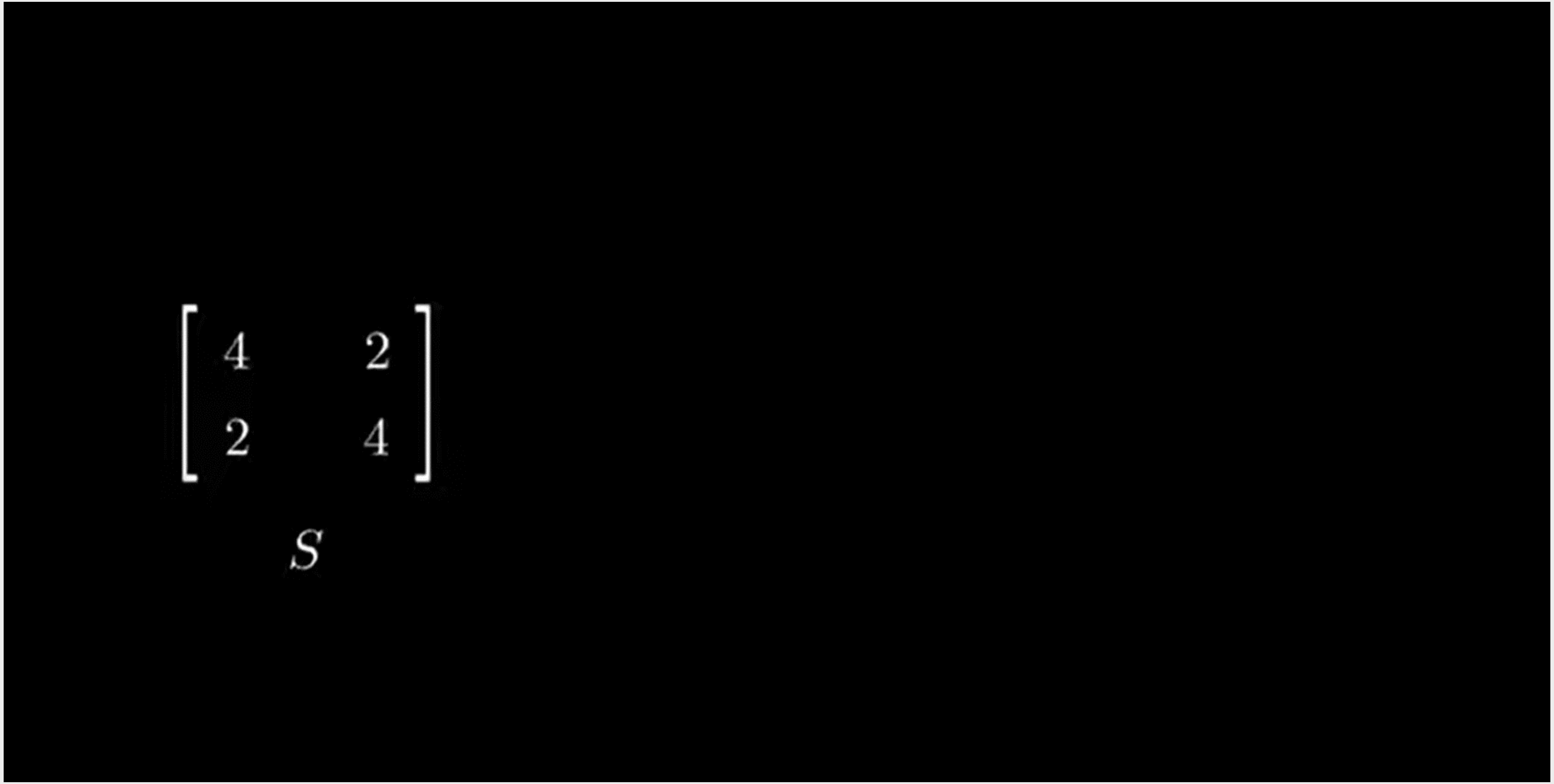
b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.

c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.

d. A is orthogonally diagonalizable.

Spectral Decomposition (Real)





$$\begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$
S

- Spectral Decomposition is nice and pretty, but with loss of generality:

  Real Field: For square and symmetric matrices!
  Complex Field: For square and normal matrices!

  For General?? SVD!!!

Quadratic Form

Quadratic Form

- 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 Form

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

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

- 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 $Q(x) = x^T Ax$, 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 (A) is a diagonal matrix.

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.

- 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 $x^T A x$, 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.

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

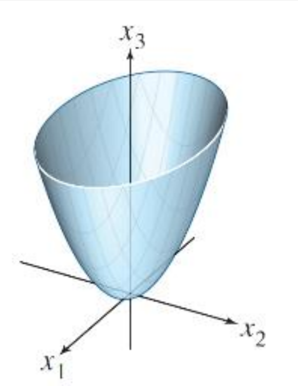
$$B = \frac{1}{2}\left(A + A^T\right)$$

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

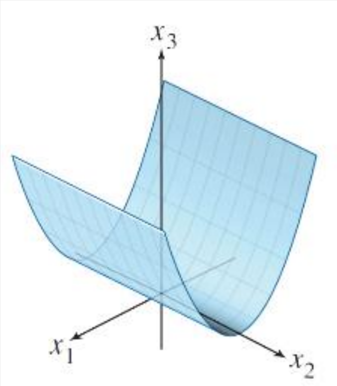
Classifying Quadratic Forms

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

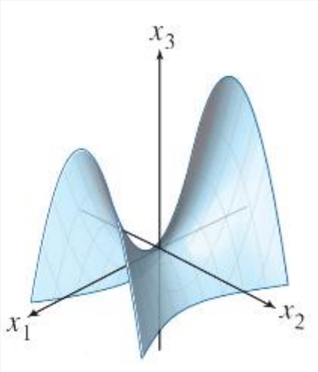
point (x_1, x_2, z) where $z = Q(x)$



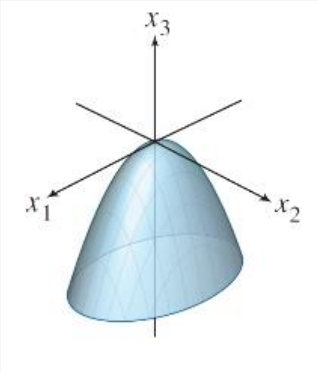
(a) $z = 3x_1^2 + 7x_2^2$



(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:

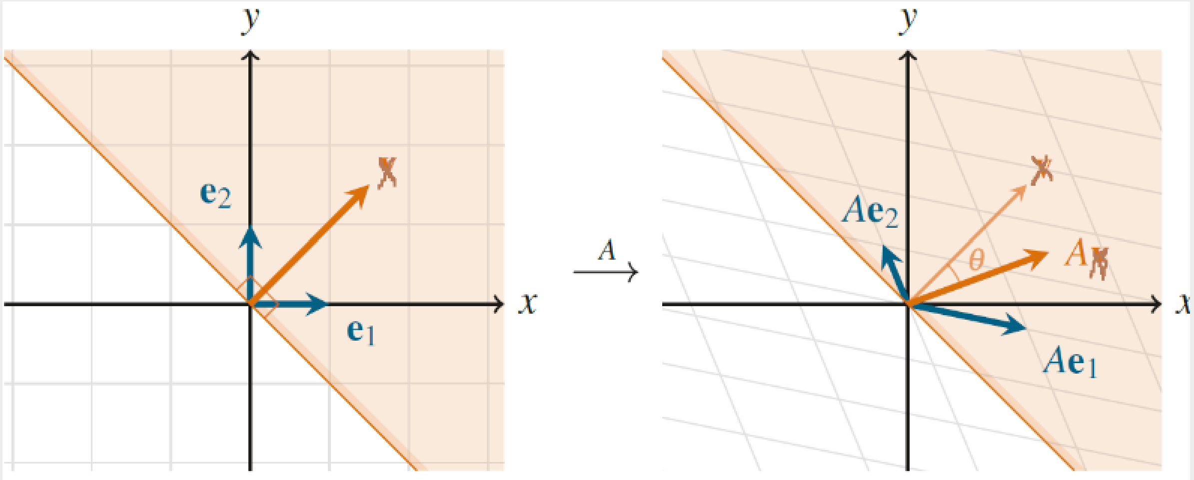
- **positive definite** if $Q(x) > 0$ for all $x \neq 0$;
- **negative definite** if $Q(x) < 0$ for all $x \neq 0$;
- **indefinite** if $Q(x)$ assumes both positive and negative values;
- **positive semidefinite** if $Q(x) \geq 0$ for all x;
- **negative semidefinite** if $Q(x) \leq 0$ for all x;

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

Geometric Interpretation

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

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



Characterization of Positive Definite Matrices

Suppose $A \in \mathcal{M}_n(\mathbb{F})$ is self-adjoint $(A^* = A)$.. 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 $A = B^*B$

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 $A = UDU^*$.

You can extend these facts to other categories!

Characterization of Positive Semidefinite Matrices

Suppose $A \in \mathcal{M}_n(\mathbb{F})$ is self-adjoint $(A^* = A)$. 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 $A = UDU^*$.

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 Tests

Positive Definite Matrices

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

1. $x^T Ax > 0$ 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 $v^T Av$ 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?

Sylvester’s Criterion

Theorem

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.

Proof?

- 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

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

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 $rank(P) = m$.

Gram Matrix

Gram Matrix

Gram(A) : $A^T A$

- symmetric
- non-negative eigenvalues
- real eigenvalues
- orthonormal eigenvectors
- positive semi-definite

Proof?