8. Unconstrained optimization and quadratic functions

- positive definite and positive semidefinite matrices
- eigenvalues and eigenvectors
- sufficient conditions for quadratic functions

8. Unconstrained optimization and quadratic functions

- positive definite and positive semidefinite matrices
- eigenvalues and eigenvectors
- sufficient conditions for quadratic functions

Positive definite and positive semidefinite matrices

Types of symmetric matrices

Consider a square symmetric matrix $A = A^T \in \mathbf{R}^{n \times n}$

- A is positive definite ($A \succ 0$) if

$$x^T A x > 0, \quad \forall x \neq 0 \in \mathbf{R}^n$$

- A is positive semidefinite ($A \succeq 0$) if

$$x^T A x \geq 0, \quad \forall x \in \mathbf{R}^n$$

- The matrix A is **negative** definite iff $-A$ is **positive** definite

$$A \prec 0 \iff -A \succ 0$$

- The matrix A is **negative** semidefinite iff $-A$ is **positive** semidefinite

$$A \preceq 0 \iff -A \succeq 0$$

- The matrix A is **indefinite** if $x^T A x > 0$ and $y^T A y < 0$ for some $x \neq y \in \mathbf{R}^n$.

Example 1

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

Claim: $A \succ 0$

Example 1

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

Claim: $A \succ 0$ Proof: $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\begin{aligned} x^T A x &= 2x_1^2 + x_2^2 - 2x_1x_2 \\ &= x_1^2 + (x_1^2 - 2x_1x_2 + x_2^2) \\ &= x_1^2 + (x_1 - x_2)^2 \geq 0 \quad \text{(sum of squares)} \end{aligned}$$

Example 1

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

Claim: $A \succ 0$ Proof:

$x > 0 \iff x_i > 0$ for all i

$$\begin{aligned} x^T A x &= 2x_1^2 + x_2^2 - 2x_1 x_2 \\ &= x_1^2 + (x_1^2 - 2x_1 x_2 + x_2^2) \\ &= x_1^2 + (x_1 - x_2)^2 \geq 0 \quad \text{(sum of squares)} \end{aligned}$$

Can $x^T A x = 0$ for $x \neq 0$? (Why?)

Example 2

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

This matrix is **indefinite**. (Why?)

Example 2

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

This matrix is **indefinite**. (Why?)

$$x^T A x = x_1^2 + x_2^2 + 4x_1 x_2$$

Pick $x = (1, 1)$,

$$x^T A x = 6$$

Pick $x = (1, -1)$,

$$x^T A x = -2$$

Example 3: Diagonal matrix

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \in \mathbf{R}^{n \times n}$$

Then

- $A \succ 0 \iff a_{ii} > 0$ for all i
- $A \succeq 0 \iff a_{ii} \geq 0$ for all i

Proof:

Example 3: Diagonal matrix

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \in \mathbf{R}^{n \times n}$$

Then

- $A \succ 0 \iff a_{ii} > 0$ for all i here $x \neq 0$
- $A \succeq 0 \iff a_{ii} \geq 0$ for all i here $x \in \mathbb{R}^n$

Proof:

$$x^T A x = \sum_{i=1}^{n} a_{ii} x_i^2 \quad \begin{cases} > 0 & \text{if} \quad a_{ii} > 0, x \neq 0 \\ \geq 0 & \text{if} \quad a_{ii} \geq 0, x \neq 0 \end{cases}$$

Example 3: Diagonal matrix

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \in \mathbf{R}^{n \times n}$$

Then

- $A \succ 0 \iff a_{ii} > 0$ for all i
- $A \succeq 0 \iff a_{ii} \geq 0$ for all i

Proof:

$$x^T A x = \sum_{i=1}^{n} a_{ii} x_i^2 \quad \begin{cases} > 0 & \text{if} \quad a_{ii} > 0, x \neq 0 \\ \geq 0 & \text{if} \quad a_{ii} \geq 0, x \neq 0 \end{cases}$$

Now suppose that $a_{ii} < 0$. Then pick $x = e_i$.

$$x^T A x = a_{ii} < 0.$$

Diagonal dominant matrices:

A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is

1. diagonally dominant if

$$|A_{ii}| \geq \sum_{j \neq i} |A_{ij}|$$

2. strictly diagonally dominant if

$$|A_{ii}| > \sum_{j \neq i} |A_{ij}|$$

Eg:

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

Diagonally dominant matrix.

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.

1. A is positive semidefinite if $A_{ii} \geq 0$ for all $i \in \{1, \ldots, n\}$ and A is diagonally dominant.
2. A is positive definite if diagonal elements are positive and A is strictly diagonal dominant.

eg: $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ $\quad A = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$

Eigenvalues and eigenvectors

Eigenvalues and eigenvectors

Let A be a square $n \times n$ and $x \neq 0 \in \mathbf{R}^n$. Then

$$Ax = \lambda x, \quad x \in \mathbf{R}^n, \quad \lambda \in \mathbf{R}$$

where $x \in N(A - \lambda I)$

- x is an eigenvector
- λ is an eigenvalue

Examples:

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

Eigenvalues and eigenvectors

Let A be a square $n \times n$ and $x \neq 0 \in \mathbf{R}^n$. Then

$$Ax = \lambda x, \quad x \in \mathbf{R}^n, \quad \lambda \in \mathbf{R}$$

where

- x is an eigenvector
- λ is an eigenvalue

Examples:

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

$$\lambda = 1, \quad x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{or} \qquad \lambda = 2, \quad x = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

Eigenvalues

Let A be a square $n \times n$ and $x \neq 0 \in \mathbf{R}^n$. Then

$$Ax = \lambda x, \quad x \in \mathbf{R}^n, \quad \lambda \in \mathbf{R}$$

where

- x is an eigenvector
- λ is an eigenvalue

Examples:

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

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

Eigenvalues

Let A be a square $n \times n$ and $x \neq 0 \in \mathbf{R}^n$. Then

$$Ax = \lambda x, \quad x \in \mathbf{R}^n, \quad \lambda \in \mathbf{R}$$

where

- x is an eigenvector
- λ is an eigenvalue

Examples:

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

$$\lambda = 3, \quad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{or} \qquad \lambda = 1, \quad x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Eigenvalues

Let A be a square $n \times n$ and $x \neq 0 \in \mathbf{R}^n$. Then

$$Ax = \lambda x, \quad x \in \mathbf{R}^n, \quad \lambda \in \mathbf{R}$$

where

- x is an eigenvector
- λ is an eigenvalue

Examples:

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

Eigenvalues

Let A be a square $n \times n$ and $x \neq 0 \in \mathbf{R}^n$. Then

$$Ax = \lambda x, \quad x \in \mathbf{R}^n, \quad \lambda \in \mathbf{R}$$

where

- x is an eigenvector
- λ is an eigenvalue

Examples:

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

$$\lambda = 0,\ x = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad \text{or} \qquad \lambda = 2,\ x = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \qquad \text{or} \qquad \lambda = 8,\ x = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Eigenvalues of symmetric matrices

If A is **symmetric**, it has n eigenvectors :

$$Ax_1 = \lambda_1 x_1, \quad Ax_2 = \lambda_2 x_2, \quad \dots, \quad Ax_n = \lambda_n x_n$$

Matrix form

for orthonormal X, $X^T X = I$
$X X^T = I$

$$A \underbrace{[x_1, ..., x_n]}_{X} = [x_1, ..., x_n] \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}_{\Lambda} \qquad \text{or} \qquad AX = X\Lambda$$

Thm: Assume (λ, x_λ) and (γ, x_γ) as eigenpairs of A.

Eigenvectors are orthogonal If $\lambda \neq \gamma$ then $x_\lambda^T x_\gamma = 0$

$$x_i^T x_j = 0, \quad \forall i \neq j, \quad X^T X = I \iff X^{-1} = X^T \quad \text{if normalized.}$$

Matrix is diagonalized by eigenvectors

$$\Lambda = X^T A X = \quad \text{diagonal}$$

A matrix A is diagonalizable if $\exists$ an invertible matrix P s.t. $D = P^{-1} A P$.

Eigenvalues and definiteness

$n \times n$ symmetric matrix A is PD (positive definite) iff all eigenvalues are positive

Proof: $X^T A X = \Lambda = \mathbf{diag}(\lambda)$ eigenvalues

- For any vector $z \in \mathbf{R}^n$, take $y = X^T z \iff Xy = z$. Then

$$z^T A z = y^T X^T A X y = y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2$$

- Thus, $x \neq 0$, $x^T A x > 0 \iff \lambda_i > 0$ for all i
- Examples

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

Sufficient conditions for quadratic functions

Minimizing quadratic functions

$$\underset{x \in \mathcal{S}}{\text{minimize}} \; f(x) = \frac{1}{2} x^T A x + b^T x + c,$$

Gradient and Hessian

$$\nabla f(x) = Ax + b, \quad \nabla^2 f(x) = A$$

Finding optimal points

1. Find $x = x^*$ where $Ax^* + b = 0$ (stationary points).
2. If $A \succeq 0$, then $x = x^*$ is a global minimum
3. If $A \succ 0$, then $x = x^*$ is a **unique** global minimum

Proof:

Minimizing quadratic functions

$$\underset{x \in \mathcal{S}}{\text{minimize}} \;\; f(x) = \frac{1}{2}x^T A x + b^T x + c,$$

Gradient and Hessian

$$\nabla f(x) = Ax + b, \quad \nabla^2 f(x) = A$$

Finding optimal points

1. Find $x = x^*$ where $Ax^* + b = 0$ (stationary points).
2. If $A \succeq 0$, then $x = x^*$ is a global minimum
3. If $A \succ 0$, then $x = x^*$ is a **unique** global minimum

Proof: for all $x \neq x^*$,

$$f(x) = f(x^*) + (x - x^*)^T \underbrace{\nabla f(x^*)}_{=0} + \underbrace{\frac{1}{2}(x - x^*)^T \overbrace{\nabla^2 f(x^*)}^{=A \succeq 0}(x - x^*)}_{\geq 0} \geq f(x^*)$$

Sufficient optimality conditions

$$\underset{x \in \mathcal{S}}{\text{minimize}} \quad f(x) \quad f(x) = \frac{1}{2} a x^2 + bx + c$$

- $x = x^* \in \mathcal{S}$ is a **local minimum** of $f(x)$ if

$$\nabla f(x^*) = 0, \qquad \nabla^2 f(x^*) \succ 0$$

- $x = x^* \in \mathcal{S}$ is a **global minimum** of the **quadratic function**

$$f(x) = \frac{1}{2} x^T A x + b^T x + c$$

if

$$\nabla f(x^*) = 0, \qquad \nabla^2 f(x^*) = A \succeq 0$$

2nd order sufficiency cond.

- $x = x^* \in \mathcal{S}$ is a **global minimum** of the **general function** $f(x)$ if

$$\nabla f(x^*) = 0, \qquad \nabla^2 f(x) \succeq 0 \; \forall x \in \mathcal{S}$$

e.g. $f(x)$ is **convex**.

$$\nabla^2 f(x^*) \succ 0$$

$$f(x) = x^4$$

$$f''(x) = 12x^2$$