7. Unconstrained optimization and quadratic functions

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

Example

$$\underset{x,y \in \mathbf{R}}{\text{minimize}} \quad f(x,y) := \frac{x+y}{x^2+y^2+1}$$

How many saddles point does $f(x)$ have?

A. 0

B. 1

C. 2

Example

$$\underset{x,y}{\text{minimize}} \quad f(x,y) := \frac{x+y}{x^2+y^2+1}$$

Gradient of f:

$$\nabla f(x,y) = \frac{1}{(x^2+y^2+1)^2} \begin{bmatrix} y^2 - 2xy - x^2 + 1 \\ x^2 - 2xy - y^2 + 1 \end{bmatrix}$$

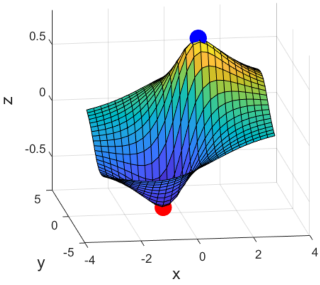
Where is $\nabla f(x,y) = 0$?

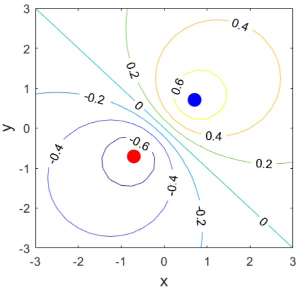
$$(x^*, y^*) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \quad \text{and} \quad (x^*, y^*) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

Hessian of f at these points:

$$\nabla^2 f(x,y) = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \text{and} \quad \nabla^2 f(x,y) = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Example





$$\underbrace{(x^*, y^*) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)}_{\text{minimum}} \quad \text{and} \quad \underbrace{(x^*, y^*) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)}_{\text{maximum}}$$

$$\underbrace{\nabla^2 f(x,y) = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\text{positive definite}} \quad \text{and} \quad \underbrace{\nabla^2 f(x,y) = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{\text{negative definite}}$$

Sufficient conditions for optimality

$$\underset{x \in \mathcal{S}}{\text{minimize}} \quad f(x), \qquad f : \mathbf{R}^n \to \mathbf{R}$$

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

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

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

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

- If $\nabla f(x^*) = 0$, $\nabla^2 f(x^*)$ is neither positive nor negative definite, it is **indefinite**
  - x^* is a saddle point (not a minimizer or a maximizer)

Motivating proof

- If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite, then for any x and $z = x - x^*$:

$$f(x) = f(x^*) + \underbrace{\nabla f(x^*)^T z}_{=0} + \underbrace{\frac{1}{2} z^T \nabla^2 f(x^*) z}_{\text{strictly positive}} + \underbrace{O(\|z\|^3)}_{\text{really small}} > f(x^*)$$

  when x is close enough to x^* (local minimum)

- If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is negative definite, then for any x and $z = x - x^*$:

$$f(x) = f(x^*) + \underbrace{\nabla f(x^*)^T z}_{=0} + \underbrace{\frac{1}{2} z^T \nabla^2 f(x^*) z}_{\text{strictly negative}} + \underbrace{O(\|z\|^3)}_{\text{really small}} < f(x^*)$$

  when x is close enough to x^* (local maximum)

Convexity and optimality: 1-D

$$\underset{x \in \mathcal{S}}{\text{minimize}} \quad f(x), \qquad f(x) \quad \text{is differentiable everywhere.}$$

- Suppose for some point x^* in the interior of $\mathcal{S}$, $f'(x^*) = 0$.
- Then,
  - if $f''(x^*) > 0$, x^* is a local minimum
  - if $f''(x^*) < 0$, x^* is a local maximum
  - if $f''(x^*) = 0$, x^* could be a local minimum, maximum, or saddle point.
- Example of third case:
  - $f(x) = x^4$, $x^* = 0$ is a local minimum
  - $f(x) = -x^4$, $x^* = 0$ is a local maximum
  - $f(x) = x^3$, $x^* = 0$ is a saddle point

Convexity and optimality: 1-D

$$\underset{x \in \mathcal{S}}{\text{minimize}} \quad f(x), \qquad f(x) \quad \text{is differentiable everywhere.}$$

Claim: If $f''(x) \geq 0$ for all $x \in \mathcal{S}$, then

$$f'(x^*) = 0 \iff x^* \quad \text{is a global minimum (maybe not unique)}$$

Convexity and optimality: 1-D

$$\underset{x \in \mathcal{S}}{\text{minimize}} \quad f(x), \qquad f(x) \quad \text{is differentiable everywhere.}$$

Claim: If $f''(x) \geq 0$ for all $x \in \mathcal{S}$, then $\leftarrow$ We don't need strict inequality!

$$f'(x^*) = 0 \iff x^* \quad \text{is a global minimum (maybe not unique)}$$

Convexity and optimality: 1-D

$$\underset{x \in \mathcal{S}}{\text{minimize}} \quad f(x), \qquad f(x) \quad \text{is differentiable everywhere.}$$

Claim: If $f''(x) \geq 0$ for all $x \in \mathcal{S}$, then $\leftarrow$ We don't need strict inequality!

$$f'(x^*) = 0 \iff x^* \quad \text{is a global minimum (maybe not unique)}$$

Proof: use mean value theorem from calculus.

- Suppose there exists some $\bar{x}$ where $f(\bar{x}) < f(x^*)$
- Without loss of generality, assume $\bar{x} < x^*$. (Just reverse the proof otherwise.)
- By MVT, there exists $\tilde{x} \in (\bar{x}, x^*)$ where $f'(\tilde{x}) > 0$
- By MVT, there exists $\hat{x} \in (\tilde{x}, x^*)$ where $f''(\hat{x}) < 0$
- Contradiction!

Quadratic functions

Quadratic functions

Quadratic functions over $\mathbf{R}^n$ have the form

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

where A is an $n \times n$ symmetric matrix, $b \in \mathbf{R}^n$, $c \in \mathbf{R}$

$n = 1$,

$$f(x) = \frac{1}{2}ax^2 + bx + c, \quad A = [a]$$

$n = 2$,

$$\begin{aligned} f(x) &= \frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix} \overbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}}^{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \overbrace{\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}^T}^{b^T} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + c \\ &= \frac{1}{2}a_{11}x_1^2 + \frac{1}{2}a_{22}x_2^2 + a_{12}x_1x_2 + b_1x_1 + b_2x_2 + c \end{aligned}$$

Question: how to minimize $f(x)$? Local / global minimizer?

Quadratic functions and symmetry

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

We can always assume without loss of generality that

$$A = A^T \qquad \text{(symmetric)}$$

Suppose that $A \neq A^T$. Then

$$x^T A x = \frac{1}{2}x^T A x + \frac{1}{2}x^T A^T x = \frac{1}{2}x^T \underbrace{(A + A^T)}_{\substack{\text{always} \\ \text{symm.}}} x$$

e.g. we could replace A with $\frac{1}{2}(A + A^T)$ and not change the function value.

Gradients and hessians of quadratic function

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c, \qquad A \quad \text{is symmetric}$$

Recall $x = x^*$ is a

- local minimizer if $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite ($\nabla^2 f(x^*) \succ 0$

$$\forall z \in \mathbf{R}^n, \quad z^T \nabla^2 f(x^*) z > 0.$$

- global minimizer if there is only one such point satisfying this

Note that this condition is **sufficient** but **not necessary**

Question: What is the gradient and Hessian of $f(x)$?

Gradients and hessians of quadratic function

$$f(x) = \frac{1}{2}\underbrace{x^T A x}_{h(x)} + \underbrace{b^T x}_{g(x)} + c, \qquad A \quad \text{is symmetric}$$

Gradient and Hessian?

Gradients and hessians of quadratic function

$$f(x) = \frac{1}{2}\underbrace{x^T A x}_{h(x)} + \underbrace{b^T x}_{g(x)} + c, \qquad A \quad \text{is symmetric}$$

Gradient and Hessian?

$$g(x) = b^T x = \sum_{i=1}^{n} b_i x_i, \qquad \frac{\partial g}{\partial x_i} = b_i, \qquad \frac{\partial^2 g}{\partial x_i \partial x_j} = 0$$

$$\nabla g(x) = b, \qquad \nabla^2 g(x) = 0$$

Gradients and hessians of quadratic function

$$f(x) = \frac{1}{2} \underbrace{x^T A x}_{h(x)} + \underbrace{b^T x}_{g(x)} + c, \qquad A \quad \text{is symmetric}$$

Gradient and Hessian?

$$h(x) = x^T A x = \sum_{i=1}^{n} \sum_{j \neq i} A_{ij} x_i x_j + \sum_{i=1}^{n} A_{ii} x_i^2$$

$$\frac{\partial h}{\partial x_i} = 2 \sum_{j \neq i} A_{ij} x_j + 2 A_{ii} x_i, \qquad \frac{\partial^2 h}{\partial x_i \partial x_i} = 2 A_{ij} x_j + 2 A_{ii}$$

$$\nabla h(x) = 2Ax, \qquad \nabla^2 h(x) = 2A$$

Gradients and hessians of quadratic function

$$f(x) = \frac{1}{2}\underbrace{x^T A x}_{h(x)} + \underbrace{b^T x}_{g(x)} + c, \qquad A \quad \text{is symmetric}$$

Gradient and Hessian?

$$\nabla g(x) = b, \qquad \nabla^2 g(x) = 0$$

$$\nabla h(x) = 2Ax, \qquad \nabla^2 h(x) = 2A$$

By linearity of derivatives,

$$\boxed{\nabla f(x) = Ax + b, \qquad \nabla^2 f(x) = A}$$

Gradients and hessians of quadratic function

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c, \qquad A \quad \text{is symmetric}$$

Which of the following statement is true?

A. The solution to $Ax = b$ are the minimizer of $f(x)$ and is unique minimizer if A is invertible.

B. $f(x)$ has a unique minimizer if A is positive definite.

C. Assume null$(A) \neq \{0\}$. If x^* is a minimizer, then there exists an $\alpha \in \mathbf{R}$ and $d \in \mathcal{N}(A)$ such that $x^* + \alpha d$ is a maximizer of $f(x)$

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. From before: If $A \succ 0$, then $x = x^*$ is a **local minimum**

   for all points x close enough to x^*, $\quad f(x) > f(x^*)$.

3. Can we generalize to global optimality?

   Let's take a closer look at A.

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:

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

Example 1

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

Claim: $A \succ 0$ Proof:

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

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

Proof:

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

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 $x = e_i$.

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

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 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:

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

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

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

Eigenvectors are orthogonal

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

Matrix is diagonalized by eigenvectors

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

Eigenvalues and definiteness

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

Proof: $X^T A X = \Lambda = \mathbf{diag}(\lambda_1)$ 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

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

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

$$\underset{x \in \mathcal{S}}{\text{minimize}} \quad f(x)$$

- $x = x^* \in \mathcal{S}$ 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

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

- $x = x^* \in \mathcal{S}$ 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**.