Convexity

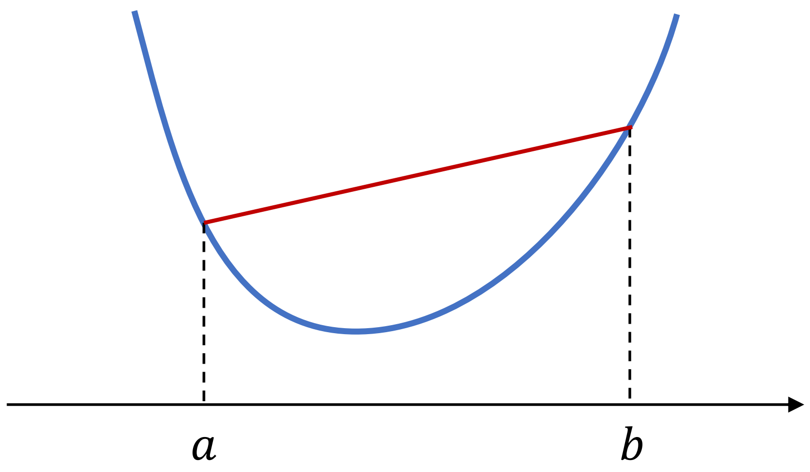
Hao-Wen Dong

Material based on Intro to Machine Learning (CSE 251A), Fall 2021

Definition – Convex function

- A function $f: \mathbb{R}^d \to \mathbb{R}$ is *convex* if for all $a, b \in \mathbb{R}^d$ and $0 < \theta < 1$,

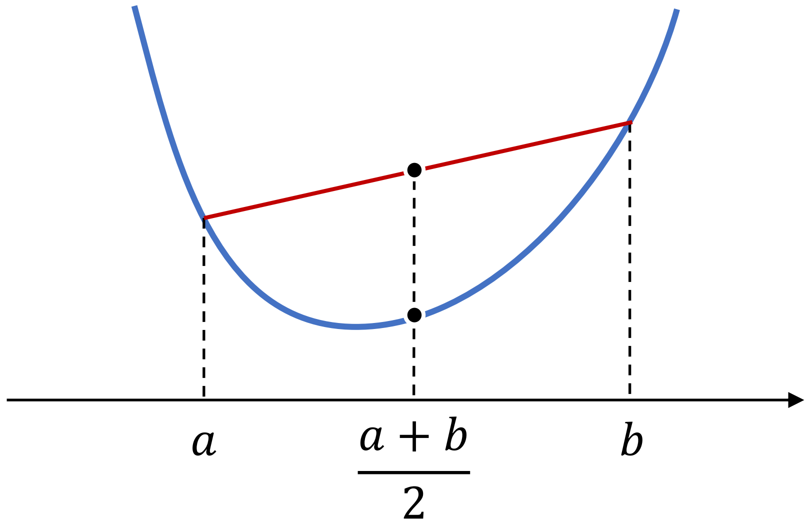
$$f(\theta a + (1-\theta)b) \leq \theta f(a) + (1-\theta)f(b)$$



Definition – Convex function

- A function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is *convex* if for all $a, b \in \mathbb{R}^d$ and $0 < \theta < 1$,

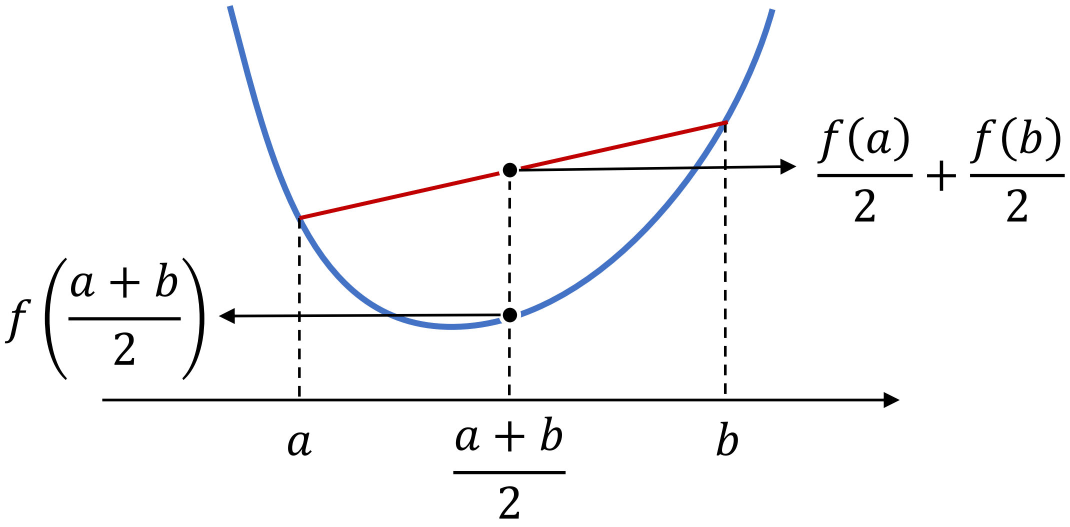
$$f(\theta a + (1-\theta)b) \;\leq\; \theta f(a) + (1-\theta)f(b)$$



Definition – Convex function

- A function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is *convex* if for all $a, b \in \mathbb{R}^d$ and $0 < \theta < 1$,

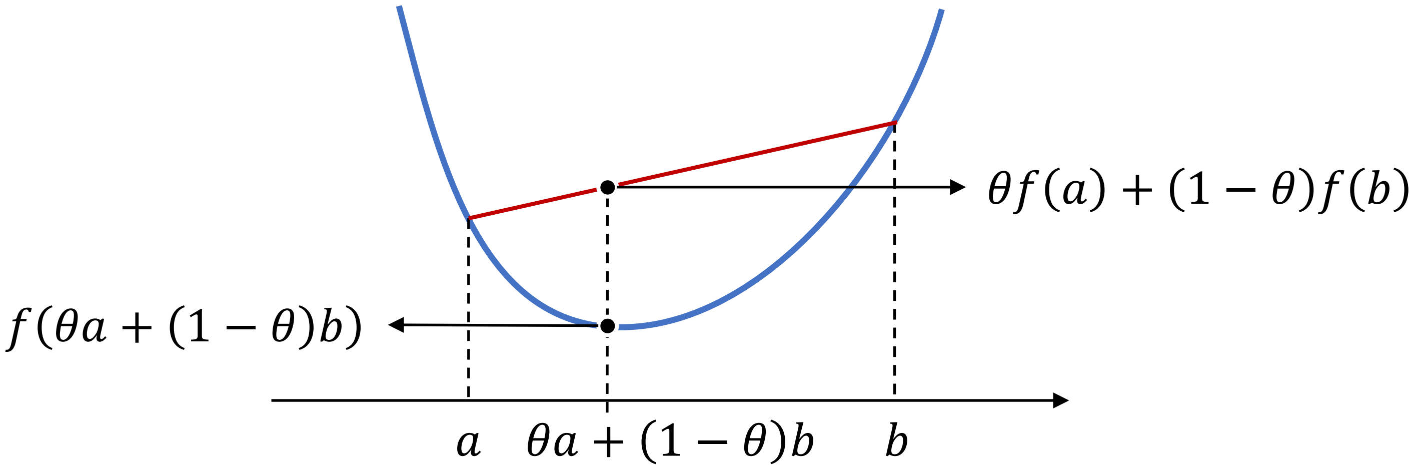
$$f(\theta a + (1-\theta)b) \;\leq\; \theta f(a) + (1-\theta)f(b)$$



Definition – Convex function

- A function $f: \mathbb{R}^d \to \mathbb{R}$ is *convex* if for all $a, b \in \mathbb{R}^d$ and $0 < \theta < 1$,

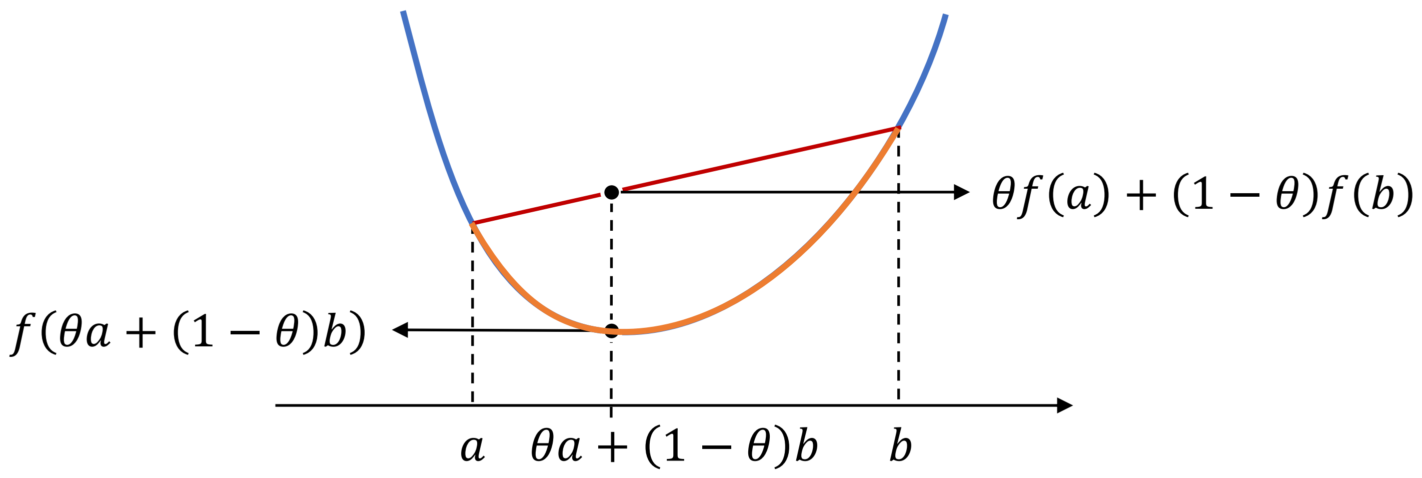
$$f(\theta a + (1-\theta)b) \;\leq\; \theta f(a) + (1-\theta) f(b)$$



Definition – Convex function

- A function $f: \mathbb{R}^d \to \mathbb{R}$ is *convex* if for all $a, b \in \mathbb{R}^d$ and $0 < \theta < 1$,

$$f(\theta a + (1-\theta)b) \;\leq\; \theta f(a) + (1-\theta) f(b)$$



Definition – Strictly convex function

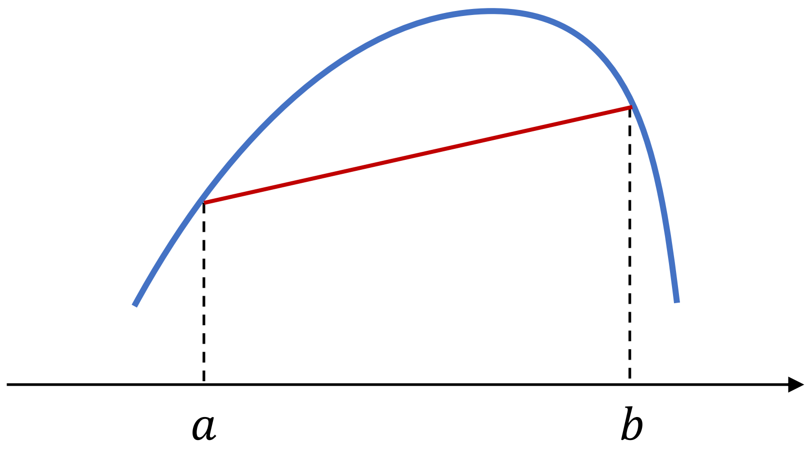
- A function $f: \mathbb{R}^d \to \mathbb{R}$ is *strictly convex* if for all $a \neq b \in \mathbb{R}^d$ and $0 < \theta < 1$,

$$f(\theta a + (1-\theta)b) < \theta f(a) + (1-\theta)f(b)$$

Definition – Concave function

- A function $f: \mathbb{R}^d \to \mathbb{R}$ is *concave* if for all $a, b \in \mathbb{R}^d$ and $0 < \theta < 1$,

$$f(\theta a + (1-\theta)b) \;\geq\; \theta f(a) + (1-\theta) f(b)$$



Properties of a convex function

- f is convex $\Leftrightarrow -f$ is concave
- f, g are both convex $\Rightarrow f + g$ is convex
- f, g are both convex $\Rightarrow \max(f, g)$ is convex

Second-derivative test for convexity

- A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is convex if its second derivative is nonnegative everywhere
  - Example: $f(x) = x^2$

Second-derivative test for convexity

- How about a *multivariate* function $f: \mathbb{R}^d \to \mathbb{R}$?

Second-derivative test for convexity

- How about a *multivariate* function $f: \mathbb{R}^d \to \mathbb{R}$?
- A function $f: \mathbb{R}^d \to \mathbb{R}$ is convex if its matrix of second derivatives is positive semidefinite everywhere

First derivative of multivariate functions

- For a function $f: \mathbb{R}^d \to \mathbb{R}$, its first derivative is a vector with d entries, called the *gradient*

$$\nabla f(z) = \begin{bmatrix} \frac{\partial f}{\partial z_1} \\ \vdots \\ \frac{\partial f}{\partial z_d} \end{bmatrix}$$

First derivative of multivariate functions

- For a function $f: \mathbb{R}^d \rightarrow \mathbb{R}$, its second derivative is a $d \times d$ matrix, called the *Hessian* matrix

$$H_f = \begin{bmatrix} \frac{\partial f}{\partial z_1 \partial z_1} & \cdots & \frac{\partial f}{\partial z_1 \partial z_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial z_d \partial z_1} & \cdots & \frac{\partial f}{\partial z_d \partial z_d} \end{bmatrix}$$

  - It's the Jacobian matrix of $\nabla f(z)$

Second-derivative test for convexity

- A function $f: \mathbb{R}^d \to \mathbb{R}$ is convex if its Hessian is positive semidefinite everywhere

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

  - Note that

$$z^T M z = \sum_{i,j=1}^{d} M_{ij} z_i z_j$$

A hierarchy of square matrices

Square	$M \in \mathbb{R}^{d \times d}$
$\cup$	
Symmetric	$M = M^T$
$\cup$	
Positive semidefinite	$z^T M z \geq 0$ for all $z \in \mathbb{R}^d$
$\cup$	
Positive definite	$z^T M z > 0$ for all $z \neq 0$

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- Is $M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ positive semidefinite?

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- Is $M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ positive semidefinite?

For any $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$, we have

$$z^T M z = \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 & z_1 + z_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = (z_1 + z_2)^2 \geq 0$$

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- Is $M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ positive semidefinite?

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- Is $M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ positive semidefinite?

For any $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$, we have

$$z^T M z = \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} z_1 + 2z_2 & 2z_1 + z_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = z_1^2 + 5z_1 z_2 + {z_2}^2$$

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- Is $M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ positive semidefinite? No! $z = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is a counterexample

For any $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$, we have

$$z^T M z = \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} z_1 + 2z_2 & 2z_1 + z_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = z_1^2 + 5z_1 z_2 + {z_2}^2$$

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- Is $M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ positive semidefinite? No! $z = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is a counterexample

  Recall: a matrix corresponds to a *linear transformation*

  A geometric view of PSD: Any transformed vector Mz must have a nonnegative scalar projection on the original vector z.

  In this case, $Mz = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is in the opposite direction to $z = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- When is a diagonal matrix PSD?

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- When is a diagonal matrix PSD?

Suppose we have a diagonal matrix $M = \mathrm{diag}(a_1, a_2, \dots, a_d)$, then we have

$$z^T M z = [z_1 \quad \cdots \quad z_d] \begin{bmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_d \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_d \end{bmatrix} = [a_1 z_1 \quad \cdots \quad a_d z_d] \begin{bmatrix} z_1 \\ \vdots \\ z_d \end{bmatrix} = a_1 z_1^2 + \cdots + a_d z_d^2$$

Positive semidefinite (PSD)

- A symmetric matrix M is *positive semidefinite* if for all $z \in \mathbb{R}^d$

$$z^T M z \geq 0$$

- When is a diagonal matrix PSD? When all its elements are nonnegative

Suppose we have a diagonal matrix $M = \mathrm{diag}(a_1, a_2, \dots, a_d)$, then we have

$$z^T M z = \begin{bmatrix} z_1 & \dots & z_d \end{bmatrix} \begin{bmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_d \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_d \end{bmatrix} = \begin{bmatrix} a_1 z_1 & \dots & a_d z_d \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_d \end{bmatrix} = a_1 z_1^2 + \cdots + a_d z_d^2$$

Second-derivative test for convexity

- A function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is convex if its Hessian is positive semidefinite everywhere

- Is $f(z) = ||z||^2$ convex?

Second-derivative test for convexity

- A function $f: \mathbb{R}^d \to \mathbb{R}$ is convex if its Hessian is positive semidefinite everywhere

- Is $f(z) = \|z\|^2$ convex?

$$f(z) = \|z\|^2 = \sum_{i=1}^{d} z_i^2 \,,$$

Second-derivative test for convexity

- A function $f: \mathbb{R}^d \to \mathbb{R}$ is convex if its Hessian is positive semidefinite everywhere

- Is $f(z) = \|z\|^2$ convex?

$$f(z) = \|z\|^2 = \sum_{i=1}^{d} z_i^2 \, ,$$

$$\nabla f(z) = \begin{bmatrix} \frac{\partial f}{\partial z_1} \\ \vdots \\ \frac{\partial f}{\partial z_d} \end{bmatrix} = \begin{bmatrix} 2z_1 \\ \vdots \\ 2z_d \end{bmatrix} = 2z, \qquad H_f = \begin{bmatrix} \frac{\partial f}{\partial z_1 \partial z_1} & \cdots & \frac{\partial f}{\partial z_1 \partial z_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial z_d \partial z_1} & \cdots & \frac{\partial f}{\partial z_d \partial z_d} \end{bmatrix} = \begin{bmatrix} 2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2 \end{bmatrix} = 2I_d$$

Second-derivative test for convexity

- A function $f : \mathbb{R}^d \to \mathbb{R}$ is convex if its Hessian is positive semidefinite everywhere

- Is $f(z) = \|z\|^2$ convex? Yes!

$$f(z) = \|z\|^2 = \sum_{i=1}^{d} z_i^2 \,,$$

$$\nabla f(z) = \begin{bmatrix} \frac{\partial f}{\partial z_1} \\ \vdots \\ \frac{\partial f}{\partial z_d} \end{bmatrix} = \begin{bmatrix} 2z_1 \\ \vdots \\ 2z_d \end{bmatrix} = 2z, \qquad H_f = \begin{bmatrix} \frac{\partial f}{\partial z_1 \partial z_1} & \cdots & \frac{\partial f}{\partial z_1 \partial z_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial z_d \partial z_1} & \cdots & \frac{\partial f}{\partial z_d \partial z_d} \end{bmatrix} = \begin{bmatrix} 2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2 \end{bmatrix} = 2I_d$$