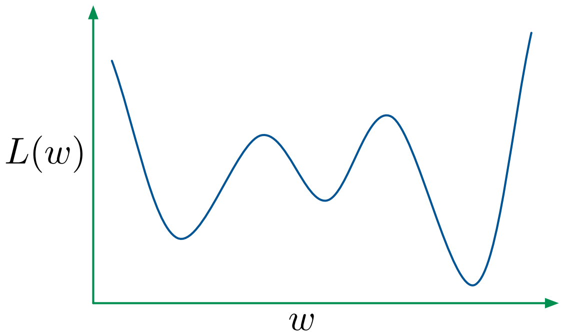
Convexity I

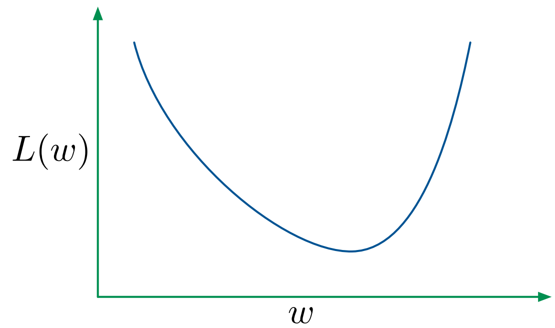
Topics we'll cover

1. Definition of convexity
2. The second-derivative test for convexity

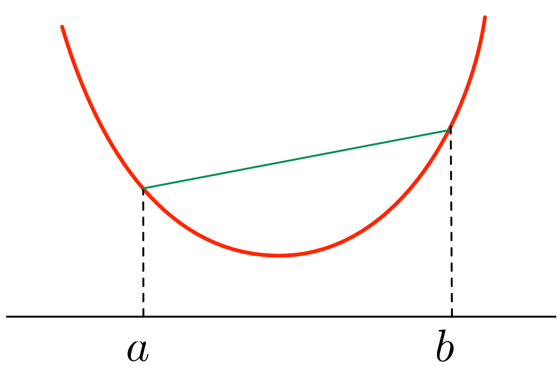
Is our loss function convex?



versus



Convexity



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

$$f(\theta a + (1-\theta)b) \ \le \ \theta f(a) + (1-\theta) f(b).$$

It is **strictly convex** if strict inequality holds for all $a \neq b$.

f is **concave** $\Leftrightarrow$ $-f$ is convex

Checking convexity for functions of one variable

A function $f : \mathbb{R} \to \mathbb{R}$ is convex if its second derivative is ≥ 0 everywhere.

Example: $f(z) = z^2$

Checking convexity

Function of one variable

$$F : \mathbb{R} \to \mathbb{R}$$

- Value: number
- Derivative: number
- Second derivative: number

Convex if second derivative is always ≥ 0

Function of d variables

$$F : \mathbb{R}^d \to \mathbb{R}$$

- Value: number
- Derivative: d-dimensional vector
- Second derivative: $d \times d$ matrix

Convex if second derivative matrix is always positive semidefinite

First and second derivatives of multivariate functions

For a function $f : \mathbb{R}^d \to \mathbb{R}$,

- the first derivative is a vector with d entries:

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

- the second derivative is a $d \times d$ matrix, the **Hessian** $H(z)$:

$$H_{jk} \ = \ \frac{\partial^2 f}{\partial z_j \partial z_k}$$

Example

Find the second derivative matrix of $f(z) = \|z\|^2$.

Checking convexity

Function of one variable

$$F : \mathbb{R} \to \mathbb{R}$$

- Value: number
- Derivative: number
- Second derivative: number

Convex if second derivative is always ≥ 0

Function of d variables

$$F : \mathbb{R}^d \to \mathbb{R}$$

- Value: number
- Derivative: d-dimensional vector
- Second derivative: $d \times d$ matrix

Convex if second derivative matrix is always positive semidefinite