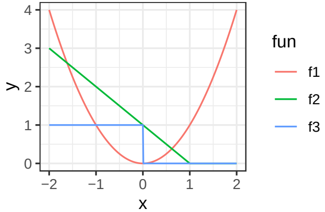
Optimization

Convexity



Learning goals

- TODO
- TODO

CONVEX VS. NON-CONVEX

"...in fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."

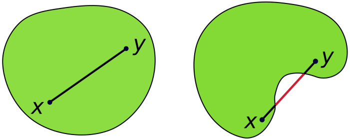
- R. Tyrrell Rockafellar, in SIAM Review, 1993

CONVEX SETS

A set of $\mathcal{S}$ is **convex**, if for all $\mathbf{x}, \mathbf{y} \in \mathcal{S}$ and all $t \in [0, 1]$ the following applies:

$$\mathbf{x} + t(\mathbf{y} - \mathbf{x}) \in \mathcal{S}$$

Intuitively: If $\mathbf{x}, \mathbf{y}$ are in $\mathcal{S}$, then the connecting line is also in $\mathcal{S}$.



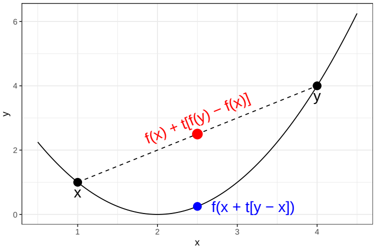
The set in the left image is convex, the set in the right image is not convex (concave).
Source: Wikipedia.

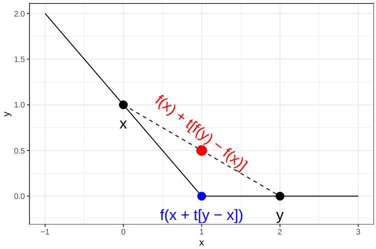
CONVEX FUNCTIONS

Consider $f : \mathcal{S} \to \mathbb{R}$, where $\mathcal{S}$ convex. The function is **convex** if for all $\mathbf{x}, \mathbf{y} \in \mathcal{S}$ and all $t \in [0, 1]$

$$f(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) \leq f(\mathbf{x}) + t(f(\mathbf{y}) - f(\mathbf{x})).$$

It is called **strictly convex**, if this is valid for "$<$" instead of "$\leq$".





Left: A differentiable and strictly convex function. Right: A convex function that is non-differentiable in $x = 0$, which is not strictly convex.

CONVEX FUNCTIONS

For a twice differentiable function f, convexity can determined from the **Hessian matrix**.

The function $f : \mathcal{S} \to \mathbb{R}$ is **convex iff** the Hessian matrix $\nabla^2 f(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in \mathcal{S}$, i.e. if for all points $\mathbf{x}$ and all vectors $\mathbf{d} \neq 0$ it applies:

$$\mathbf{d}^\top \nabla^2 f(\mathbf{x}) \mathbf{d} \geq 0.$$

If the Hessian matrix is positive definite (strict "$>$"), the function f is strictly convex.

Equivalent definition: A matrix is positive semidefinite if all eigenvalues are non-negative.

CONVEX FUNCTIONS

Convex functions play a special role in optimization.

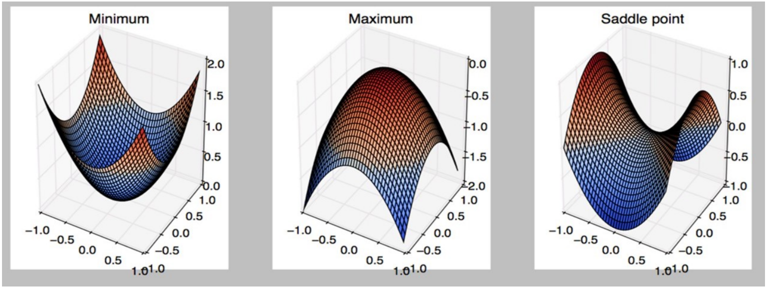
Let $f : \mathcal{S} \to \mathbb{R}$ be convex on the convex set $\mathcal{S}$. Then the following applies:

- Any local minimum of f is also a global minimum (see chapter *Conditions for Optimality*).
- If f is strictly convex, f has exactly one local minimum on $\mathcal{S}$ and it is also the unique global minimum of f on $\mathcal{S}$ (see chapter *Conditions for Optimality*).
- Sublevel sets $S_1 = \{\mathbf{x} \mid f(\mathbf{x}) < a\}$ and $S_2 = \{\mathbf{x} \mid f(\mathbf{x}) \leq a\}$, $a \in \mathbb{R}$, form convex sets.

CONVEX FUNCTIONS

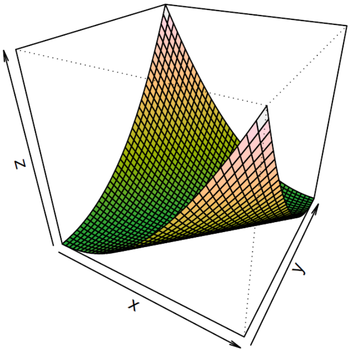
Example:

1. Local min: If [f convex $\Leftrightarrow$ All eigenvalues positive], then global min
2. Local max: If [f concave $\Leftrightarrow$ All eigenvalues negative], then global max
3. Some eigenvalues positive and some negative $\Leftrightarrow$ saddle point



CONVEX FUNCTIONS

Example: Consider the function $f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1x_2$.



CONVEX FUNCTIONS

The gradient of the function is $\nabla f(x) = (2x_1 - 2x_2, 2x_2 - 2x_1)$ and the Hessian is

$$\nabla^2 f(x) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

The matrix is positive semidefinite, since

$$\begin{aligned} \mathbf{d}^\top \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \mathbf{d} &= \mathbf{d}^\top \begin{pmatrix} 2d_1 - 2d_2 \\ -2d_1 + 2d_2 \end{pmatrix} \\ &= 2d_1^2 - 2d_1 d_2 - 2d_1 d_2 + 2d_2^2 \\ &= 2d_1^2 - 4d_1 d_2 + 2d_2^2 = 2(d_1 - d_2)^2 \geq 0. \end{aligned}$$

So the function f is convex and every local minimum is also a global minimum.