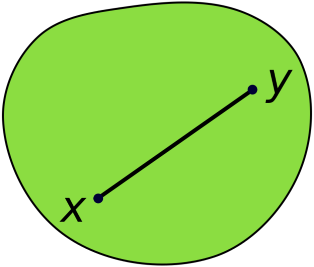
Optimization

Convexity



Learning goals

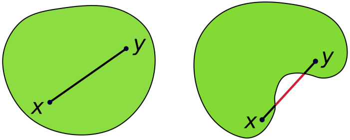
- Convex sets
- Convex functions

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 $\mathcal{S}$.



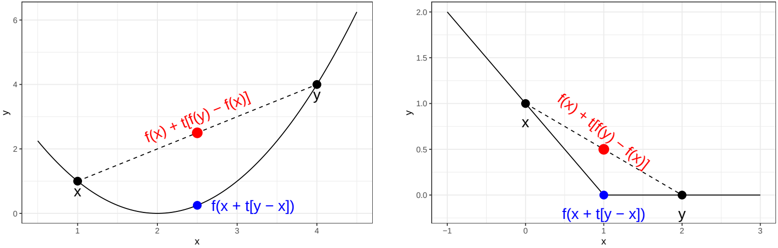
The set in the left image is convex, the set in the right image is not convex (concave). Source: Wikipedia.

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.

CONCAVE FUNCTIONS

Inversely, a function $f : \mathcal{S} \to \mathbb{R}$, $\mathcal{S}$ *concave*, is **concave** if for all $\mathbf{x}, \mathbf{y} \in \mathcal{S}$ and all $t \in [0, 1]$

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

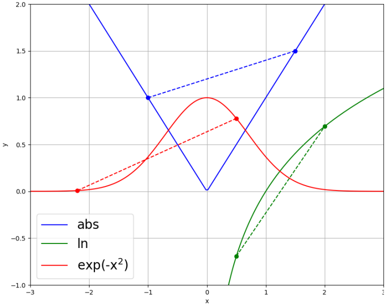
It is called **strictly convex**, if this is valid for "$>$" instead of "$\geq$". It holds that f (strictly) concave $\Leftrightarrow -f$ (strictly) convex.

EXAMPLES

Convex function:: $f(x) = |x|$. **Proof:**

Concave function: $f(x) = \log(x)$.

Non-convex, non-concave: $f(x) = \exp(-x^2)$ **Proof:** (plug in respective values to show that it is not convex and not concave)



CONVEXITY AND HESSIAN MATRIX

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:

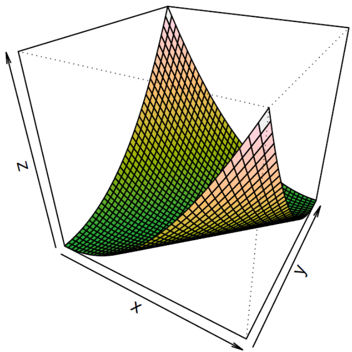
$$\mathbf{d}^\top \nabla^2 f(\mathbf{x}) \mathbf{d} \geq 0.$$

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.

CONVEXITY AND HESSIAN MATRIX

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



CONVEXITY AND HESSIAN MATRIX

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

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

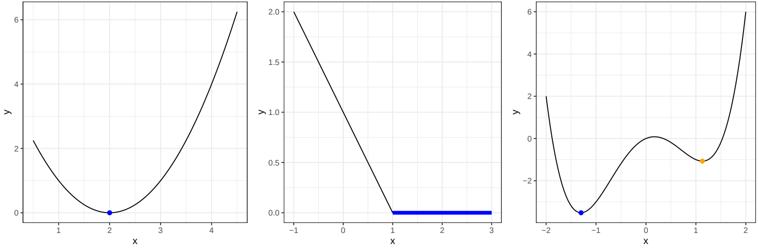
The matrix is positive semidefinite, since

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

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

CONVEX FUNCTIONS IN OPTIMIZATION

- For a convex function, every local optimum is a global one
- A strictly convex function at most one optimal point



Left: Example for a function that is strictly convex. It has one local minimum. Middle: A function that is convex, but not strictly convex. All local optima are global ones, but the optimum is not unique. Right: A function that is not convex.

CONVEX FUNCTIONS IN OPTIMIZATION

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

- R. Tyrrell Rockafellar, in SIAM Review, 1993