Video 11.1: Critical points, global and local extrema

Optimization and Computational Linear Algebra for Data Science

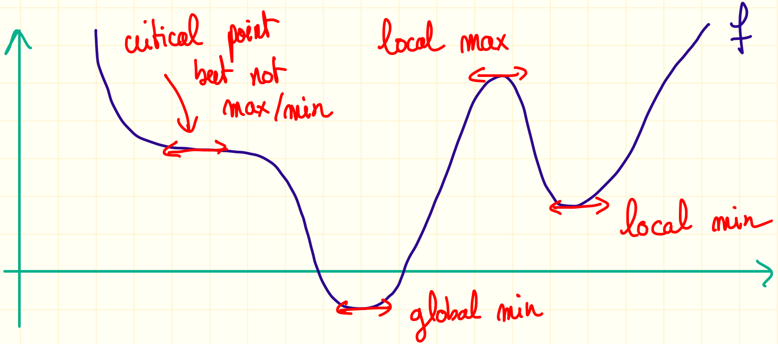
Léo Miolane

Definitions

Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. We say that $x \in \mathbb{R}^n$ is

- a critical point of f if $\nabla f(x) = 0$,
- a *global* minimizer of f if for all $x' \in \mathbb{R}^n$, $f(x) \leq f(x')$,
- a *local* minimizer of f if there exists $\delta > 0$ such that we have $f(x) \leq f(x')$ for all x' verifying $\|x - x'\| \leq \delta$.



Local extrema are critical points

Proposition

$$x \text{ is a local minimizer of } f \implies \nabla f(x) = 0.$$

Proposition

Assume that f is convex. Then

$$\nabla f(x) = 0 \iff x \text{ is a global minimizer of } f.$$

Looking at the Hessian

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. Let $x \in \mathbb{R}^n$ be a critical point of f, i.e. $\nabla f(x) = 0$.
Then, if $H_f(x)$ is positive definite (that is, if all the eigenvalues of $H_f(x)$ are strictly positive), then x is a local minimizer of f.

Idea: Use Taylor's formula "for small h"

$$f(x+h) \simeq f(x) + \cancel{\langle h, \nabla f(x)\rangle} + \frac{1}{2} \underbrace{h^T H_f(x) h}_{>0 \text{ for all } h \neq 0}.$$

$\to$ for all "h small"

$$f(x+h) \gtrsim f(x)$$

Looking at the Hessian

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. Let $x \in \mathbb{R}^n$ be a critical point of f, i.e. $\nabla f(x) = 0$.
Then, if $H_f(x)$ is negative definite (that is, if all the eigenvalues of $H_f(x)$ are strictly negative), then x is a local maximizer of f.

Saddle points

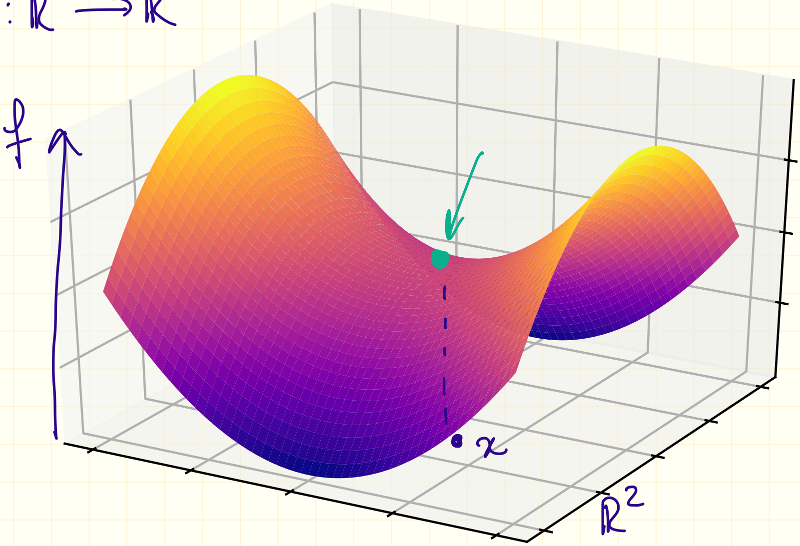
Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. Let $x \in \mathbb{R}^n$ be a critical point of f, i.e. $\nabla f(x) = 0$.
Then, if $H_f(x)$ admits strictly positive eigenvalues and strictly negative eigenvalues, then x is neither a local maximum nor a local minimum. We call x a saddle point.

$$f(x+h) \simeq f(x) + \frac{1}{2} \underbrace{h^T H_f(x) h}_{\text{can be either } > 0 \text{ or } < 0}.$$

Saddle points

$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$



Example

Study the critical points of $f(x,y) = x^2 + xy^2 - x + 1$.

$$\nabla f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y^2 - 1 \\ 2xy \end{pmatrix}$$

Let's find the critical points of f : we solve:

$$\begin{cases} 2x + y^2 - 1 = 0 \\ 2xy = 0 \end{cases} \iff \begin{pmatrix} y = 0 \\ 2x - 1 = 0 \end{pmatrix} \text{ or } \begin{pmatrix} x = 0 \\ y^2 - 1 = 0 \end{pmatrix}$$

3 critical points: $\begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$.

$$Hf\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 2y \\ 2y & 2x \end{pmatrix} = 2\begin{pmatrix} 1 & y \\ y & x \end{pmatrix}$$

Example

Study the critical points of $f(x, y) = x^2 + xy^2 - x + 1$.

- $Hf\begin{pmatrix} 1/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ is positive definite: $\begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$ loc min
- $Hf\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2\overbrace{\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}}^{M}$
- $Hf\begin{pmatrix} 0 \\ -1 \end{pmatrix} = 2\begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$

these two matrices admits one >0 eigenvalue and one <0 neigenvalue $\longrightarrow$ saddle points

M is symmetric, it has 2 eigenvalues $\lambda_1 \geq \lambda_2$

- $Tr(M) = 1 = \lambda_1 + \lambda_2 \longrightarrow \lambda_1 > 0$
- $\lambda_2 < 0$ because for $v = (1, -1)$ we have $v^T M v = -1$ therefore $\lambda_2 < 0$