Lecture 8.1: Functions of n variables

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

Derivative / Gradient

$$
\begin{aligned}
f: \quad \mathbb{R} &\to \mathbb{R} \\
x &\mapsto f(x)
\end{aligned}
$$

Derivative at $x \in \mathbb{R}$:

$$f'(x) \in \mathbb{R}$$

$$
\begin{aligned}
f: \quad \mathbb{R}^n &\to \mathbb{R} \\
x &\mapsto f(x) = f(x_1, \ldots, x_n)
\end{aligned}
$$

Gradient at $x \in \mathbb{R}^n$:

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^n$$

Gradient and contour lines

Hessian matrix

What is the equivalent of the second derivative for function of n variables ?

$$\begin{aligned} f: \quad \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto f(x) = f(x_1, \ldots, x_n) \end{aligned}$$

Hessian at $x \in \mathbb{R}^n$:

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1\,\partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1\,\partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2\,\partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2\,\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\,\partial x_1}(x) & \frac{\partial^2 f}{\partial x_n\,\partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix} \in \mathbb{R}^{n \times n}$$

Example

Schwarz's Theorem

Theorem

If $f : \mathbb{R}^n \to \mathbb{R}$ is «twice differentiable», then for all $x \in \mathbb{R}^n$ and all $i, j \in \{1, \ldots, n\}$ we have:

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(x) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x).$$