Conceptos de Cálculo Multivariable para ML

I302 - Aprendizaje Automático y Aprendizaje Profundo

Roberto Bunge

Universidad de San Andrés

Gradiente (de una función escalar)

Sea $f: \mathbb{R}^n \to \mathbb{R}$ diferenciable

$$\nabla_x f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Jacobiano (de una función vectorial)

Sea $f : \mathbb{R}^n \to \mathbb{R}^m$

$$J_x f = \nabla_x f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Hessiano (de una función escalar)

Sea $f : \mathbb{R}^n \to \mathbb{R}$ doblemente diferenciable

$$H_x f = \nabla_x^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

$\nabla_x^2 f(x)$ es una matriz simétrica.

Algunas gradientes comunes

$$\nabla_x (a^T x) = a$$

$$\nabla_x (x^T a) = a$$

$$\nabla_x (x^T A x) = (A + A^T) \cdot x \qquad \text{Si } A = A^T \Rightarrow \nabla_x (x^T A x) = 2Ax$$

$$\nabla_x^2 (x^T A x) = (A + A^T)$$