Análisis Matemático para Inteligencia Artificial

Martín Errázquin (merrazquin@fi.uba.ar)

Especialización en Inteligencia Artificial

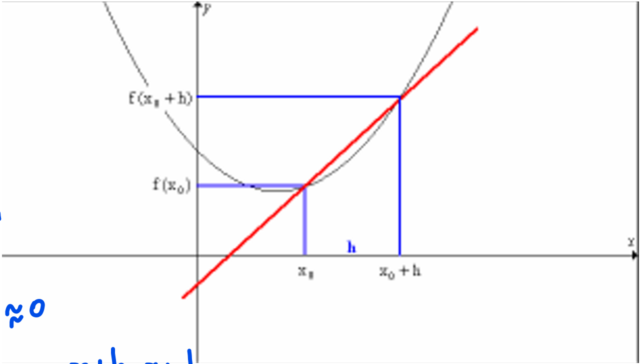
Derivadas de funciones multivariadas

Repaso: derivando escalares

(x_1, y_1)

(x_2, y_2)

$\Rightarrow m = \dfrac{\Delta y}{\Delta x} = \dfrac{y_2 - y_1}{x_2 - x_1}$



$x_2 = x_1 + h$ con $h \approx 0$

$\Rightarrow \Delta x = x + h - x = h$

Recordemos siempre que la derivada busca calcular $\Delta y / \Delta x$ para una función dada $y = f(x)$ en un punto x.

$$f'(x) \doteq \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Derivadas de orden 1: campos escalares

Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$, $(x_1, ..., x_n)^T \mapsto f((x_1, ..., x_n)^T)$, se definen las derivadas parciales como:

n deriv. parciales

$$\frac{\partial f}{\partial x_1} = \lim_{h \to 0} \frac{f(x_1 + h, x_2, ..., x_n) - f(x_1, x_2, ..., x_n)}{h} \quad \in \mathbb{R}$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n} = \lim_{h \to 0} \frac{f(x_1, x_2, ..., x_n + h) - f(x_1, x_2, ..., x_n)}{h} \quad \in \mathbb{R}$$

Se define el gradiente como: $\nabla f = \left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n} \right)$. $\in \mathbb{R}^n$

Importante: El gradiente apunta en la dirección de máximo crecimiento.

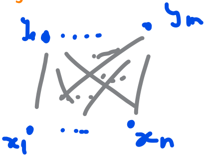
$f(x,y) = x^2 \operatorname{sen}(y)$

$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{(x+h)^2 \operatorname{sen}(y) - x^2 \operatorname{sen}(y)}{h} = \dots = 2x \operatorname{sen}(y)$

$\frac{\partial f}{\partial y} = x^2 \cos(y)$

$\nabla f(x,y) = (2x \operatorname{sen}(y), x^2 \cos(y))$

Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$,
$(x_1, \ldots, x_n)^T \mapsto (f_1((x_1, \ldots, x_n)^T), \ldots, f_m((x_1, \ldots, x_n)^T)$, se define el jacobiano como:



$$J_f = \begin{array}{cc} & \xrightarrow{n} \\ m \downarrow & \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ & \vdots & \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \end{array} = \begin{pmatrix} -\nabla f_1- \\ \vdots \\ -\nabla f_m- \end{pmatrix} \in \mathbb{R}^{m \times n}$$

$$J_{f(i,j)} = \frac{\partial f_i}{\partial x_j}$$

$n \cdot m$ deriv. parciales

Ej. $\vec{f}(x,y) = (\underbrace{xy}_{f_1}, \underbrace{x^2 y}_{f_2}, \underbrace{\cos(x)}_{f_3})$

$$\Rightarrow J_f(x,y) = \begin{pmatrix} y & x \\ 2xy & x^2 \\ -\mathrm{sen}(x) & 0 \end{pmatrix}$$

$\nabla f_1(x,y) = (y, x)$ $\quad \nabla f_2(x,y) = (2xy, x^2)$ $\quad \nabla f_3(x,y) = (-\mathrm{sen}(x), 0)$

Derivadas de orden 2: Matriz Hessiana

La matriz Hessiana es aquella cuyas derivadas de orden 2 de f respecto a $x \in \mathbb{R}^n$ se ubican:

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} = H_f \in \mathbb{R}^{n \times n}$$

(reg) $\Rightarrow \dfrac{\partial f}{\partial x_i \partial x_j} = \dfrac{\partial f}{\partial x_j \partial x_i}$

$\Rightarrow H_f$ simétrica

• es el $J_{\nabla f}$

Ej.: si $f(x,y) = x^2 \operatorname{sen}(y)$

$\Rightarrow \nabla f(x,y) = (\underbrace{2x \operatorname{sen}(y)}_{f_1}, \underbrace{x^2 \cos(y)}_{f_2})$ $\Rightarrow H_f(x,y) = \begin{pmatrix} 2 \operatorname{sen}(y) & 2x \cos(y) \\ 2x \cos(y) & -x^2 \operatorname{sen}(y) \end{pmatrix}$

Una aplicación común del Hessiano es el polinomio de Taylor de orden 2 para campos escalares. Sea f un campo escalar $f : \mathbb{R}^n \to \mathbb{R}$, asumiendo que posee derivadas parciales de todo orden en un entorno de un punto $a \in \mathbb{R}^n$, se define el polinomio de Taylor de orden 2:

$$P_2(x) = \overbrace{f(a)} + \overbrace{\nabla_f(a)^T(x-a)}^{y^Tx} + \frac{1}{2}\overbrace{(x-a)^T H_f(a)(x-a)}^{z^TAx}$$

escalares:

$$P_2(x) = \underbrace{f(a)}_{\substack{\text{comp.} \\ \text{const.} \\ (0)}} + \underbrace{f'(a)\cdot(x-a)}_{\substack{\text{comp.} \\ \text{lineal} \\ (1)}} + \underbrace{\frac{1}{2}f''(a)(x-a)^2}_{\substack{\text{comp.} \\ \text{cuadrática} \\ (2)}}$$

Ejemplo

Sea $f(x_1, x_2) = cos(x_1) + x_1x_2 + 3x_1^2$

$\mathbb{R}^2 \to \mathbb{R}$

$a = \left(\frac{\pi}{2}, 0\right)$

$f(a) = \cos\left(\frac{\pi}{2}\right) + \frac{\pi}{2} \cdot 0 + \frac{3 \cdot \pi^2}{4}$

$= \frac{3}{4}\pi^2$

$\nabla_f(x_1, x_2) = \left(-\sin(x_1) + x_2 + 3x_1,\ x_1\right)$

$\nabla_f(a) = \left(-\sin\left(\frac{\pi}{2}\right) + 0 + 3 \cdot \frac{\pi}{2},\ \frac{\pi}{2}\right)$

$= \left(\frac{3}{2}\pi - 1,\ \frac{\pi}{2}\right)$

$H_f(x_1, x_2) = \begin{pmatrix} -\cos(x_1) + 3 & 1 \\ 1 & 0 \end{pmatrix}$

$H_f(a) = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$

$\nabla_0 (x - a) = \left(x_1 - \frac{\pi}{2},\ x_2\right)$

$P_2(x) = \frac{3}{4}\pi^2 + \left(\frac{3}{2}\pi - 1, \frac{\pi}{2}\right) \cdot \left(x_1 - \frac{\pi}{2}, x_2\right) + \left(x_1 - \frac{\pi}{2}, x_2\right) \cdot \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 - \frac{\pi}{2} \\ x_2 \end{pmatrix}$