

Reti Neurali (Derivazione)

Corso di AA, anno 2018/19, Padova



Fabio Aioli

12 Novembre 2018



- $y = c, c \in \mathbb{R} \Rightarrow y' = 0$



- $y = c, c \in \mathbb{R} \Rightarrow y' = 0$
- $y = c \cdot x, c \in \mathbb{R} \Rightarrow y' = c$



- $y = c, c \in \mathbb{R} \Rightarrow y' = 0$
- $y = c \cdot x, c \in \mathbb{R} \Rightarrow y' = c$
- $y = x^n, n \in \mathbb{N} \Rightarrow y' = nx^{n-1}$



- $y = c, c \in \mathbb{R} \Rightarrow y' = 0$
- $y = c \cdot x, c \in \mathbb{R} \Rightarrow y' = c$
- $y = x^n, n \in \mathbb{N} \Rightarrow y' = nx^{n-1}$
- $y = e^x \Rightarrow y' = e^x$



- $D(k \cdot f(x)) = k \cdot f'(x)$



- $D(k \cdot f(x)) = k \cdot f'(x)$
- $D(\sum_i f_i(x)) = \sum_i f'_i(x)$



- $D(k \cdot f(x)) = k \cdot f'(x)$
- $D(\sum_i f_i(x)) = \sum_i f'_i(x)$
- $D(f(x) \cdot g(x)) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$



- $D(k \cdot f(x)) = k \cdot f'(x)$
- $D(\sum_i f_i(x)) = \sum_i f'_i(x)$
- $D(f(x) \cdot g(x)) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$
- $D\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x)^2}$



- $D(k \cdot f(x)) = k \cdot f'(x)$
- $D(\sum_i f_i(x)) = \sum_i f'_i(x)$
- $D(f(x) \cdot g(x)) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$
- $D\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x)^2}$
- $D\left(\frac{1}{g(x)}\right) = -\frac{g'(x)}{g(x)^2}$



- $D(k \cdot f(x)) = k \cdot f'(x)$
- $D(\sum_i f_i(x)) = \sum_i f'_i(x)$
- $D(f(x) \cdot g(x)) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$
- $D\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x)^2}$
- $D\left(\frac{1}{g(x)}\right) = -\frac{g'(x)}{g(x)^2}$
- $D(g(f(x))) = g'(f(x)) \cdot f'(x)$



Definizione di Gradiente

Nel calcolo differenziale vettoriale, il gradiente di una funzione a valori reali $f : \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione vettoriale $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (a valori in \mathbb{R}^n). Il gradiente di una funzione è spesso definito come il vettore che ha come componenti le derivate parziali della funzione, anche se questo vale solo se si utilizzano coordinate cartesiane ortonormali. Il simbolo ∇ si legge *nabla*.



Minimizzazione con discesa di gradiente

Problema di minimizzazione

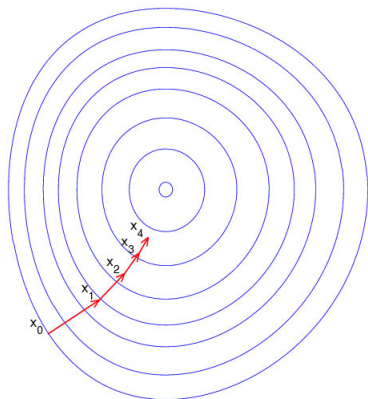
Sia $f(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione a valori scalari, trovare il vettore $\mathbf{x}^* \in \mathbb{R}^n$ che minimizza la funzione f , ovvero:

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x})$$

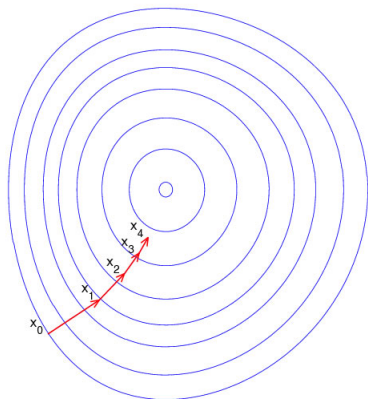
Algoritmo di discesa di gradiente

- $k \leftarrow 0, \mathbf{x}_0 \in \mathbb{R}^n$
- **while** $\nabla f(\mathbf{x}_k) \neq 0$
 - Calcolare la direzione di discesa $\mathbf{p}_k := -\nabla f(\mathbf{x}_k)$
 - Calcolare il passo di discesa η_k
 - $\mathbf{x}_{k+1} \leftarrow \mathbf{x}_k + \eta_k \mathbf{p}_k$
 - $k \leftarrow k + 1$

Minimizzazione con discesa di gradiente



Minimizzazione con discesa di gradiente



N.B. Non è detto che converga ad un minimo globale!