3. Regularized linear regression

lunedì 28 ottobre 2024 16:50

$$J(\theta) = \frac{1}{2m}\left(\sum_{i=1}^{m}\left(h\left(x^{(i)}\right) - y^{(i)}\right)^2 + \lambda \sum_{j=1}^{M} \theta_j^2\right)$$

Senza regolarizzazione

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2 \cdot x_0^{(i)}$$

CASO BASE

Nel caso della regolarizzazione abbiamo:

$$\theta_j = \theta_0 - \alpha \left[\frac{1}{m} \sum \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right) x_j^{(i)} + \frac{\lambda}{m}\theta_j\right]$$

$\alpha =$ Tasso di apprendimento

$\frac{\lambda}{m}\theta_j =$ Componente aggiunta per evitare problemi di overfitting (eccessivo adattamento dei dati)

Ora possiamo riscrivere la nostra funzione in forma aggiornata

$$\theta_j = \theta_j \underbrace{\left(1 - \alpha \frac{\lambda}{m}\right)} - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right) \cdot x_j^{(i)}$$

Va a ridurre proprio il termine θ_j Ovviamente ad ogni iterazione

$1 - \alpha \frac{\lambda}{m}$ = Fa tendere i valori verso 0, tale da ridurre il rischio di overfitting.

$\lambda \to +\infty$ = Il termine $\left(1 - \alpha \frac{\lambda}{m}\right)$ Diventa piccolo costringendo

; θ_j Vicino allo zero, tale azione porta a una riduzione della complessità del modello.

Per quanto riguarda il contrario invece permette a. A crescere di piu.