

Esercizi per il corso di Data Science - Laurea in Scienza dei Materiali

PROF. D. DI SANTE, DR. A. CONSIGLIO
SEMESTRE INVERNALE 2025/2026

2° Foglio, Ottimizzazione
08/10/2025

1 Discesa del gradiente e Regressione Polinomiale

Si desidera studiare il comportamento della **discesa del gradiente** nell'ambito di un problema di **regressione polinomiale**.

1. Generazione dei dati: Si creino 200 punti-dati (x_i, y_i) , dove ogni x_i è generato in modo casuale dall'intervallo $[0, 1]$ e

$$y_i = \sin(2\pi x_i) + \epsilon$$

con ϵ rumore casuale i cui valori appartengono all'intervallo $[-0.2, 0.2]$.

Suggerimento: visualizzare i dati generati insieme alla curva reale $\sin(2\pi x)$ per comprendere l'effetto del rumore e la distribuzione dei punti.

2. Modello di regressione polinomiale: Si consideri un modello di regressione polinomiale di ordine D :

$$\hat{y}(x; \mathbf{w}) = w_0 + w_1x + w_2x^2 + \dots + w_Dx^D$$

I coefficienti del polinomio $\mathbf{w} = (w_0, \dots, w_D)$ possono essere inizializzati casualmente all'interno dell'intervallo $[-0.5, 0.5]$.

Esperimento: provare diversi gradi del polinomio e discutere le differenze qualitative dei risultati.

3. Discesa del gradiente: Si implementi in Python un algoritmo di **discesa del gradiente** per minimizzare la seguente funzione costo:

$$J(\mathbf{w}) = \frac{1}{2N} \sum_{i=1}^N (y_i - \hat{y}(x_i; \mathbf{w}))^2$$

Partendo dai coefficienti iniziali, aggiornare iterativamente:

$$w_j \leftarrow w_j - \eta \frac{\partial J}{\partial w_j}$$

dove $\eta > 0$ è il **tasso di apprendimento**.

Esperimento: provare diversi valori di η (ad esempio 0.001, 0.01, 0.1, etc) per osservare come influisce sulla *convergenza* e sulla *stabilità* del processo di ottimizzazione.

4. Monitoraggio della convergenza: Durante l'allenamento:

- Registrare e rappresentare graficamente il valore della funzione costo $J(\mathbf{w})$ a ogni iterazione;
- Mostrare, a intervalli regolari, l'evoluzione della curva di regressione $\hat{y}(x)$ sovrapposta ai dati osservati.

Suggerimento: la visualizzazione del processo iterativo aiuta a comprendere come la discesa del gradiente “apprende” progressivamente i parametri del modello.

5. Variante: Come estensione, si implementi la variante dell'algoritmo denominato Mini-batch Stochastic Gradient Descent, nel quale l'aggiornamento del gradiente avviene su piccoli sottoinsiemi casuali dei dati (*batch*).

Confrontare, in funzione del numero di dati all'interno del batch:

- la *velocità di convergenza*,
- la *stabilità* della funzione costo,
- e il *rumore* nell'andamento della discesa.

Discussione: riflettere su vantaggi e svantaggi delle diverse strategie di ottimizzazione, in termini di efficienza e qualità della soluzione finale.