

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light green. They are positioned diagonally, with the blue one partially covering the green one.

Python per il Calcolo Scientifico

Angelo Cardellicchio

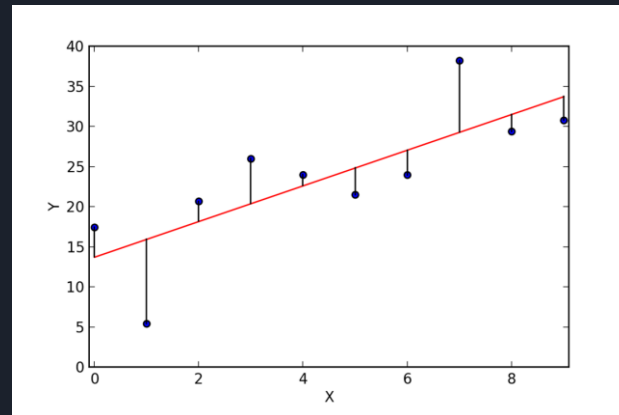


Regressione

- Cosa è la regressione?
- Multi-output regression
- Metriche (alcuni esempi)
- Algoritmi (giusto un paio...)

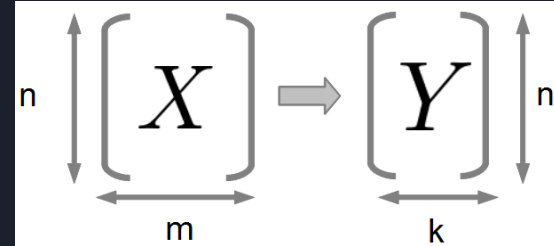
Cosa è la regressione?

- Differisce dalla classificazione in quanto non abbiamo una classe, ma un **valore numerico**
- Ad esempio, è *possibile individuare la relazione tra genere ed età di un passeggero del Titanic e costo del biglietto?*
- Abbiamo, in questo caso, due *variabili indipendenti* (genere ed età) ed una *variabile dipendente* (costo del biglietto)
- In altri termini, stiamo anche in questo caso cercando una relazione tra ingresso ed uscita!



Multi-output regression

- Si differenzia dalla regressione classica in quanto i valori predetti sono più di uno.
- Tipicamente, questi valori sono interdipendenti.
- Ad esempio, qual è il rapporto tra **umidità**, **pressione** e **temperatura** (supposte variabili indipendenti) e **direzione** e **velocità** del vento (variabili dipendenti)?





Metriche (alcuni esempi)

- **Mean Absolute Error (MAE):** è lo scarto tra i valori veri e quelli predetti dal regressore.

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_i^{n-1} |y_i - \hat{y}_i|$$

- **Mean Squared Error (MSE):** molto simile al precedente, anche se stavolta consideriamo lo scarto quadratico medio.

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_i^{n-1} (y_i - \hat{y}_i)^2$$

- **Mean Absolute Percentage Error (MAPE):** tiene conto del valore di y_i .



Algoritmi (giusto un paio)

- *Regressore lineare*
 - Sfrutta il metodo dei minimi quadrati per trovare la retta interpolante.
- *Alberi decisionali*
 - Come nel caso della classificazione, possono essere usati anche a scopi di regressione.