Statistica - 13ª lezione

11 maggio 2021

Inferenza non parametrica

Supponiamo che

$$X_1, X_2, \ldots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim f, \quad f \text{ densità incognita}$$

Inferenza non parametrica

Supponiamo che

$$X_1, X_2, \ldots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim f, \quad f \text{ densità incognita}$$

Vogliamo fare inferenza su tutta f (e non solo su un suo parmetro)

Inferenza non parametrica

Supponiamo che

$$X_1, X_2, \ldots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim f, \quad f \text{ densità incognita}$$

Vogliamo fare inferenza su tutta f (e non solo su un suo parmetro)

Ci sono due modi:

- metodi grafici
- test non parametrici

Normal qq-plot

OSSERVAZIONE I: Se $X \sim N(\mu, \sigma^2)$, allora

$$X = \sigma Z + \mu \quad \text{con} \quad Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

Normal qq-plot

OSSERVAZIONE I: Se $X \sim N(\mu, \sigma^2)$, allora

$$X = \sigma Z + \mu \quad \text{con} \quad Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

$$\Rightarrow \quad q_\gamma^X = \sigma z_\gamma + \mu \quad \text{per ogni } \gamma$$

Normal qq-plot

OSSERVAZIONE I: Se $X \sim N(\mu, \sigma^2)$, allora

$$X = \sigma Z + \mu \quad \text{con} \quad Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$
$$\Rightarrow \quad q_\gamma^X = \sigma z_\gamma + \mu \quad \text{per ogni} \ \gamma$$

OSSERVAZIONE II (più difficile)**:** Se inoltre

$$X_1, X_2, \ldots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim N(\mu, \sigma^2)$$

Normal qq-plot

OSSERVAZIONE I: Se $X \sim N(\mu, \sigma^2)$, allora

$$X = \sigma Z + \mu \quad \text{con} \quad Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

$$\Rightarrow \quad q_\gamma^X = \sigma z_\gamma + \mu \quad \text{per ogni } \gamma$$

OSSERVAZIONE II (più difficile)**:** Se inoltre

$$X_1, X_2, \ldots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim N(\mu, \sigma^2)$$

allora i *quantili empirici*

$$\hat{Q}_\gamma^X := \begin{cases} X_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2}\Big(X_{(n\gamma)} + X_{(n\gamma+1)}\Big) & \text{se } n\gamma \in \mathbb{N} \end{cases}$$

Normal qq-plot

OSSERVAZIONE I: Se $X \sim N(\mu, \sigma^2)$, allora

$$X = \sigma Z + \mu \quad \text{con} \quad Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

$$\Rightarrow \quad q_\gamma^X = \sigma z_\gamma + \mu \quad \text{per ogni } \gamma$$

OSSERVAZIONE II (più difficile)**:** Se inoltre

$$X_1, X_2, \ldots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim N(\mu, \sigma^2)$$

allora i *quantili empirici*

$$\hat{Q}_\gamma^X := \begin{cases} X_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2}\left(X_{(n\gamma)} + X_{(n\gamma+1)}\right) & \text{se } n\gamma \in \mathbb{N} \end{cases}$$

convergono in probabilità a q_γ^X

Normal qq-plot

OSSERVAZIONE I: Se $X \sim N(\mu, \sigma^2)$, allora

$$X = \sigma Z + \mu \quad \text{con} \quad Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

$$\Rightarrow \quad q_\gamma^X = \sigma z_\gamma + \mu \quad \text{per ogni } \gamma$$

dalle tavole

OSSERVAZIONE II (più difficile)**:** Se inoltre

$$X_1, X_2, \dots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim N(\mu, \sigma^2)$$

allora i *quantili empirici*

$$\hat{Q}_\gamma^X := \begin{cases} X_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2}\Big(X_{(n\gamma)} + X_{(n\gamma+1)}\Big) & \text{se } n\gamma \in \mathbb{N} \end{cases}$$

convergono in probabilità a q_γ^X

Normal qq-plot

OSSERVAZIONE I: Se $X \sim N(\mu, \sigma^2)$, allora

$$X = \sigma Z + \mu \quad \text{con} \quad Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

$$\Rightarrow \quad q^X_\gamma = \sigma z_\gamma + \mu \quad \text{per ogni } \gamma$$

OSSERVAZIONE II (più difficile)**:** Se inoltre

$$X_1, X_2, \dots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim N(\mu, \sigma^2)$$

allora i *quantili empirici*

$$\hat{Q}^X_\gamma := \begin{cases} X_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2}\left(X_{(n\gamma)} + X_{(n\gamma+1)}\right) & \text{se } n\gamma \in \mathbb{N} \end{cases}$$

convergono in probabilità a q^X_γ

dall'esperimento

Normal qq-plot

OSSERVAZIONE I: Se $X \sim N(\mu, \sigma^2)$, allora

$$X = \sigma Z + \mu \quad \text{con} \quad Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$
$$\Rightarrow \quad q_\gamma^X = \sigma z_\gamma + \mu \quad \text{per ogni } \gamma$$

OSSERVAZIONE II (più difficile)**:** Se inoltre

$$X_1, X_2, \ldots, X_n \quad \text{i.i.d.} \quad \text{con} \quad X_i \sim N(\mu, \sigma^2)$$

allora i *quantili empirici*

$$\hat{Q}_\gamma^X := \begin{cases} X_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2}\Big(X_{(n\gamma)} + X_{(n\gamma+1)}\Big) & \text{se } n\gamma \in \mathbb{N} \end{cases}$$

convergono in probabilità a q_γ^X

$$\Rightarrow \quad \hat{q}_\gamma^X \text{ e } z_\gamma \text{ si devono allineare !}$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}_{\gamma}^{X} = \hat{q}_{\frac{k}{n}}^{X}$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}_{\gamma}^{X} = \hat{q}_{\frac{k}{n}}^{X} = \frac{1}{2}\left(x_{(k)} + x_{(k+1)}\right)$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma = \hat{q}^X_{\frac{k}{n}} = \frac{1}{2}\big(x_{(k)} + x_{(k+1)}\big) \qquad \text{complicato!}$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma \;=\; \hat{q}^X_{\frac{k}{n}} \;=\; \tfrac{1}{2}\big(x_{(k)} + x_{(k+1)}\big) \qquad \text{complicato!}$$

Se invece $\gamma = \frac{k-0.5}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma \;=\; \hat{q}^X_{\frac{k-0.5}{n}}$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma = \hat{q}^X_{\frac{k}{n}} = \frac{1}{2}\left(x_{(k)} + x_{(k+1)}\right) \qquad \text{complicato!}$$

Se invece $\gamma = \frac{k-0.5}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma = \hat{q}^X_{\frac{k-0.5}{n}} = x_{(k)}$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}_{\gamma}^{X} \;=\; \hat{q}_{\frac{k}{n}}^{X} \;=\; \frac{1}{2}\big(x_{(k)} + x_{(k+1)}\big) \qquad \text{complicato!}$$

Se invece $\gamma = \frac{k-0.5}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}_{\gamma}^{X} \;=\; \hat{q}_{\frac{k-0.5}{n}}^{X} \;=\; x_{(k)} \qquad \text{più semplice\ldots}$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma = \hat{q}^X_{\frac{k}{n}} = \frac{1}{2}\left(x_{(k)} + x_{(k+1)}\right) \qquad \text{complicato!}$$

Se invece $\gamma = \frac{k-0.5}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma = \hat{q}^X_{\frac{k-0.5}{n}} = x_{(k)} \qquad \text{più semplice\ldots}$$

NORMAL QQ-PLOT = grafico dei punti

$$\left(\hat{q}^X_{\frac{k-0.5}{n}}\,,\; z_{\frac{k-0.5}{n}}\right) \qquad \text{con } k = 1, 2, \ldots, n$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma = \hat{q}^X_{\frac{k}{n}} = \frac{1}{2}\big(x_{(k)} + x_{(k+1)}\big) \qquad \text{complicato!}$$

Se invece $\gamma = \frac{k-0.5}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma = \hat{q}^X_{\frac{k-0.5}{n}} = x_{(k)} \qquad \text{più semplice\ldots}$$

NORMAL QQ-PLOT $=$ grafico dei punti

$$\left(\hat{q}^X_{\frac{k-0.5}{n}}\,,\; z_{\frac{k-0.5}{n}}\right) = \left(x_{(k)}\,,\; z_{\frac{k-0.5}{n}}\right) \qquad \text{con } k = 1, 2, \ldots, n$$

Normal qq-plot

Se $\gamma = \frac{k}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma \;=\; \hat{q}^X_{\frac{k}{n}} \;=\; \tfrac{1}{2}\big(x_{(k)} + x_{(k+1)}\big) \qquad \text{complicato!}$$

Se invece $\gamma = \frac{k-0.5}{n}$ con $k = 1, 2, \ldots, n$:

$$\hat{q}^X_\gamma \;=\; \hat{q}^X_{\frac{k-0.5}{n}} \;=\; x_{(k)} \qquad \text{più semplice\ldots}$$

NORMAL QQ-PLOT $=$ grafico dei punti

$$\left(\hat{q}^X_{\frac{k-0.5}{n}}\,,\; z_{\frac{k-0.5}{n}}\right) \;=\; \left(x_{(k)}\,,\; z_{\frac{k-0.5}{n}}\right) \qquad \text{con } k = 1, 2, \ldots, n$$

punti quasi allineati $\quad\Rightarrow\quad$ è verosimile che $X_i \sim N(\mu, \sigma^2)$

Test di Shapiro-Wilks

Si può fare (con R) un test per le ipotesi

$$H_0 : \text{le } X_i \text{ hanno densità normale} \qquad \text{vs.} \qquad H_1 : H_0 \text{ è falsa}$$

Test di Shapiro-Wilks

Si può fare (con R) un test per le ipotesi

$$H_0 : \text{le } X_i \text{ hanno densità normale} \qquad \text{vs.} \qquad H_1 : H_0 \text{ è falsa}$$

$$p\text{-value alto} \quad \Rightarrow \quad \text{non possiamo escludere } X_i \text{ normali}$$

Test di Shapiro-Wilks

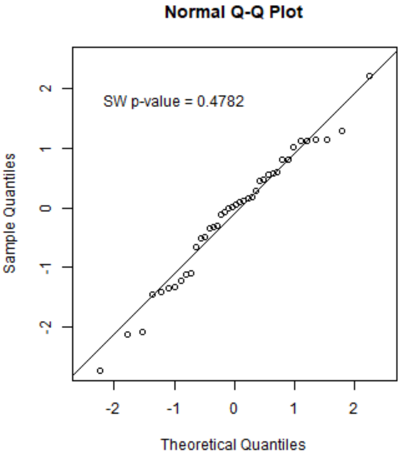
Si può fare (con R) un test per le ipotesi

$$H_0 : \text{le } X_i \text{ hanno densità normale} \qquad \text{vs.} \qquad H_1 : H_0 \text{ è falsa}$$

p-value alto $\Rightarrow$ non possiamo escludere X_i normali

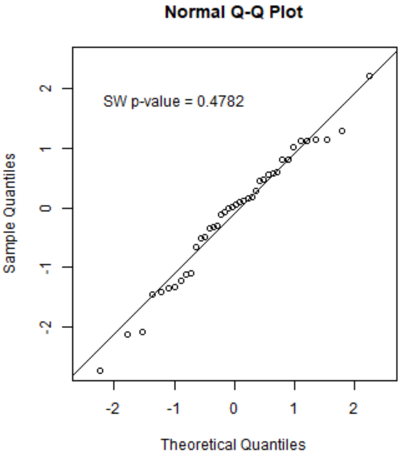
p-value basso $\Rightarrow$ normalità delle X_i poco verosimile

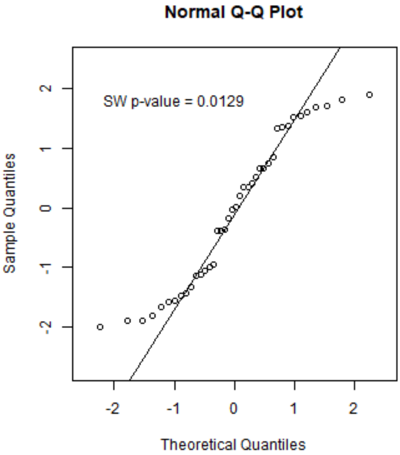
Test di Shapiro-Wilks



Non posso rifiutare la gaussianità

Test di Shapiro-Wilks





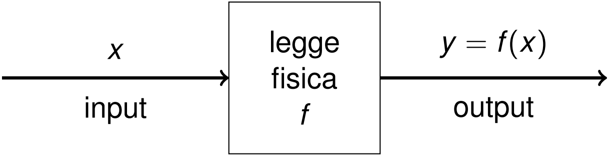
Non posso rifiutare la gaussianità

Devo rifiutare la gaussianità

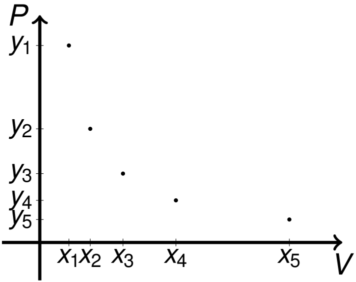
Programma

- Statistica descrittiva
  (riassumere molti dati attraverso poche caratteristiche essenziali)
- Probabilità
  (costruire un modello che preveda il risultato di un esperimento)
- Inferenza statistica
  (tarare i parametri del modello in base ai risultati dell'esperimento)
- Regressione lineare
  (riconoscere relazioni tra dati di tipo diverso)

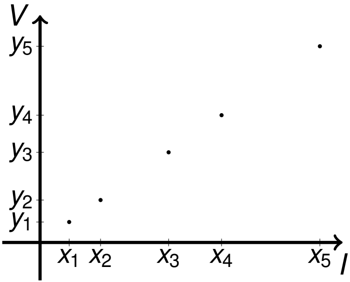
Modelli empirici



Idealmente:

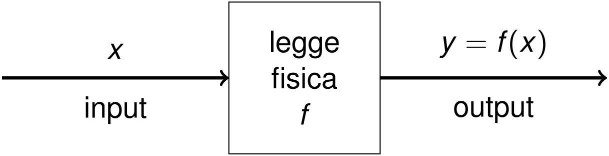


Gas perfetto

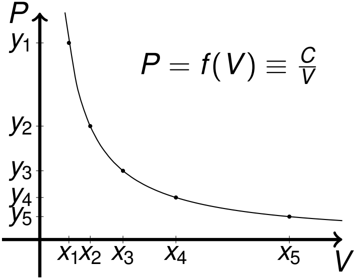


Corrente-tensione

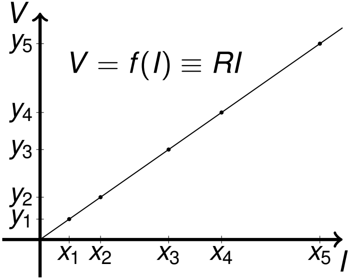
Modelli empirici



Idealmente:



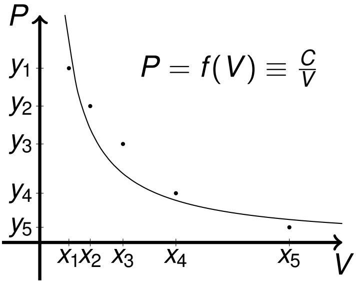
Gas perfetto



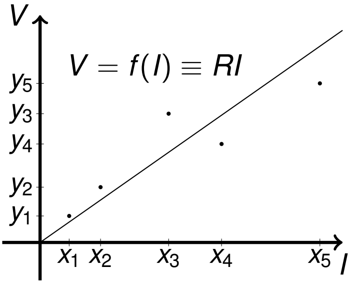
Corrente-tensione

Modelli empirici

Realisticamente:



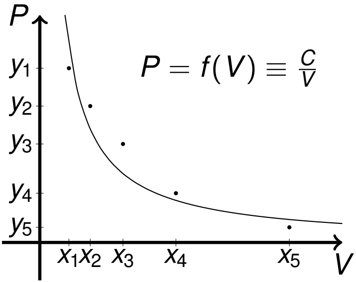
Gas perfetto



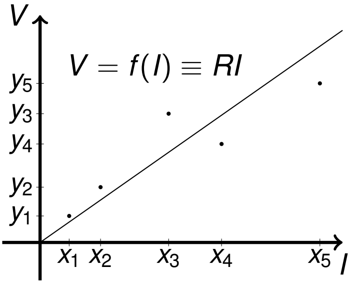
Corrente-tensione

Modelli empirici

Vogliamo trovare f!



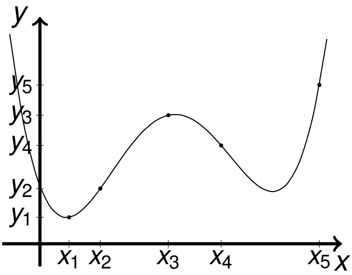
Gas perfetto



Corrente-tensione

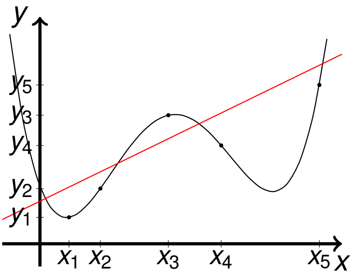
Modello lineare

Per n punti passa sempre un polinomio di grado $n-1$:



Modello lineare

Per n punti passa sempre un polinomio di grado $n-1$:



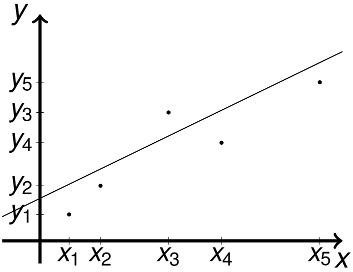
Ma noi vogliamo la funzione più semplice possibile: la retta

$$y = a + bx$$

parametrizzata solo da (a, b).

Modello lineare

Per n punti passa sempre un polinomio di grado $n-1$:



Ma noi vogliamo la funzione più semplice possibile: la retta

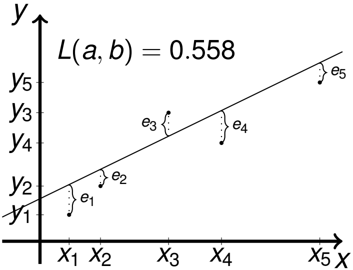
$$y = a + bx$$

parametrizzata solo da (a, b).

Come si capisce se la retta di parametri (a, b) approssima bene gli n punti?

Modello lineare

Per n punti passa sempre un polinomio di grado $n-1$:



Ma noi vogliamo la funzione più semplice possibile: la retta

$$y = a + bx$$

parametrizzata solo da (a, b).

Come si capisce se la retta di parametri (a, b) approssima bene gli n punti?

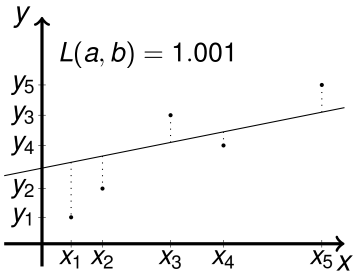
Dobbiamo introdurre gli errori (o *residui*)

$$e_i := y_i - (a + bx_i)$$

e un funzionale di errore

$$L(a, b) := \sum_{i=1}^{n} (y_i - a - bx_i)^2 = \sum_{i=1}^{n} e_i^2$$

Per n punti passa sempre un polinomio di grado $n-1$:



Ma noi vogliamo la funzione più semplice possibile: la retta

$$y = a + bx$$

parametrizzata solo da (a, b).

Come si capisce se la retta di parametri (a, b) approssima bene gli n punti?

Dobbiamo introdurre gli errori (o *residui*)

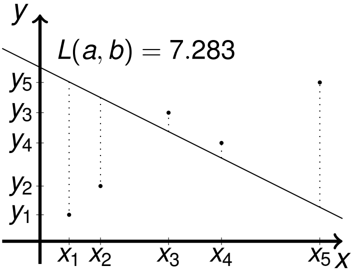
$$e_i := y_i - (a + bx_i)$$

e un funzionale di errore

$$L(a,b) := \sum_{i=1}^{n} (y_i - a - bx_i)^2 = \sum_{i=1}^{n} e_i^2$$

Modello lineare

Per n punti passa sempre un polinomio di grado $n-1$:



Ma noi vogliamo la funzione più semplice possibile: la retta

$$y = a + bx$$

parametrizzata solo da (a,b).

Come si capisce se la retta di parametri (a,b) approssima bene gli n punti?

Dobbiamo introdurre gli errori (o *residui*)

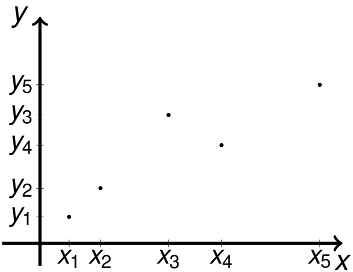
$$e_i := y_i - (a + bx_i)$$

e un funzionale di errore

$$L(a,b) := \sum_{i=1}^{n} (y_i - a - bx_i)^2 = \sum_{i=1}^{n} e_i^2$$

Modello lineare

Per n punti passa sempre un polinomio di grado $n-1$:



Ma noi vogliamo la funzione più semplice possibile: la retta

$$y = a + bx$$

parametrizzata solo da (a, b).

Come si capisce se la retta di parametri (a, b) approssima bene gli n punti?

Dobbiamo introdurre gli errori (o *residui*)

$$e_i := y_i - (a + bx_i)$$

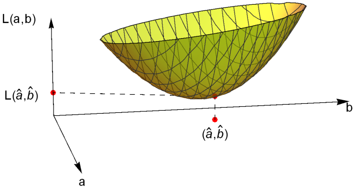
e un funzionale di errore

Perché si mette il quadrato?

$$L(a, b) := \sum_{i=1}^{n} (y_i - a - bx_i)^2 = \sum_{i=1}^{n} e_i^2$$

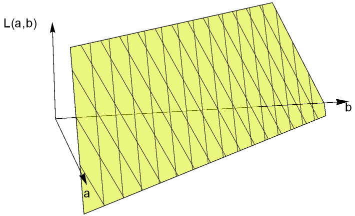
Modello lineare

Col quadrato:



- $L(a, b) \geq 0 \ \forall (a, b)$
- $L(\infty) = +\infty$
- $\exists!$ punto di minimo $(\hat{a}, \hat{b})$

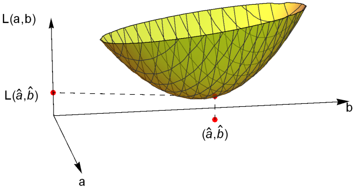
Senza quadrato:



- i residui si compensano
- $\nexists$ minimo

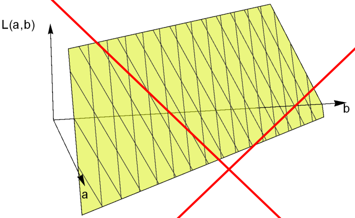
Modello lineare

Col quadrato:



- $L(a, b) \geq 0 \; \forall (a, b)$
- $L(\infty) = +\infty$
- $\exists!$ punto di minimo $(\hat{a}, \hat{b})$

Senza quadrato:



- i residui si compensano
- $\nexists$ minimo

$y = \hat{a} + \hat{b}x$ è la *retta dei minimi quadrati* (LSL = *least square line*)

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$\frac{\partial L}{\partial a} = \frac{\partial}{\partial a} \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$\begin{aligned} \frac{\partial L}{\partial a} &= \frac{\partial}{\partial a} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\ &= -2 \sum_{i=1}^{n} (y_i - a - bx_i) \end{aligned}$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$\begin{aligned}\frac{\partial L}{\partial a} &= \frac{\partial}{\partial a}\sum_{i=1}^{n}(y_i - a - bx_i)^2 \\ &= -2\sum_{i=1}^{n}(y_i - a - bx_i) \\ &= -2(n\bar{y} - na - nb\bar{x})\end{aligned}$$

dove

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i .$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$\begin{aligned}
\frac{\partial L}{\partial a} &= \frac{\partial}{\partial a} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= -2 \sum_{i=1}^{n} (y_i - a - bx_i) \\
&= -2(n\bar{y} - na - nb\bar{x}) \\
&\equiv 0
\end{aligned}$$

dove

$$\bar{x} := \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \bar{y} := \frac{1}{n} \sum_{i=1}^{n} y_i .$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$\begin{aligned}
\frac{\partial L}{\partial a} &= \frac{\partial}{\partial a} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= -2 \sum_{i=1}^{n} (y_i - a - bx_i) \\
&= -2(n\bar{y} - na - nb\bar{x}) \\
&\equiv 0 \\
&\qquad \Rightarrow \quad a = \bar{y} - b\bar{x}
\end{aligned}$$

dove

$$\bar{x} := \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \bar{y} := \frac{1}{n} \sum_{i=1}^{n} y_i .$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$\frac{\partial L}{\partial b} = \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$\begin{aligned} \frac{\partial L}{\partial b} &= \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\ &= -2 \sum_{i=1}^{n} x_i (y_i - a - bx_i) \end{aligned}$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$
\begin{aligned}
\frac{\partial L}{\partial b} &= \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= -2 \sum_{i=1}^{n} x_i (y_i - a - bx_i) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - na\bar{x} - b \sum_{i=1}^{n} x_i^2 \right)
\end{aligned}
$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$\begin{aligned}
\frac{\partial L}{\partial b} &= \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= -2 \sum_{i=1}^{n} x_i (y_i - a - bx_i) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - na\bar{x} - b \sum_{i=1}^{n} x_i^2 \right) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - n\bar{x}(\bar{y} - b\bar{x}) - b \sum_{i=1}^{n} x_i^2 \right) \qquad \text{con } a = \bar{y} - b\bar{x}
\end{aligned}$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$
\begin{aligned}
\frac{\partial L}{\partial b} &= \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= -2 \sum_{i=1}^{n} x_i (y_i - a - bx_i) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - na\bar{x} - b \sum_{i=1}^{n} x_i^2 \right) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - n\bar{x}(\bar{y} - b\bar{x}) - b \sum_{i=1}^{n} x_i^2 \right) \qquad \text{con } a = \bar{y} - b\bar{x} \\
&\equiv 0
\end{aligned}
$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$
\begin{aligned}
\frac{\partial L}{\partial b} &= \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= -2 \sum_{i=1}^{n} x_i (y_i - a - bx_i) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - na\bar{x} - b \sum_{i=1}^{n} x_i^2 \right) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - n\bar{x}(\bar{y} - b\bar{x}) - b \sum_{i=1}^{n} x_i^2 \right) \qquad \text{con } a = \bar{y} - b\bar{x} \\
&\equiv 0 \\
&\qquad \Rightarrow \quad b \; = \; \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_j^2 - n\bar{x}^2}
\end{aligned}
$$

Retta dei minimi quadrati

Cerchiamo $(\hat{a}, \hat{b})$ ponendo $\nabla L(\hat{a}, \hat{b}) \equiv 0$:

$$
\begin{aligned}
\frac{\partial L}{\partial b} &= \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= -2 \sum_{i=1}^{n} x_i (y_i - a - bx_i) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - na\bar{x} - b \sum_{i=1}^{n} x_i^2 \right) \\
&= -2 \left(\sum_{i=1}^{n} x_i y_i - n\bar{x}(\bar{y} - b\bar{x}) - b \sum_{i=1}^{n} x_i^2 \right) \qquad \text{con } a = \bar{y} - b\bar{x} \\
&\equiv 0
\end{aligned}
$$

$$
\Rightarrow \quad b = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_j^2 - n\bar{x}^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_j - \bar{x})^2}
$$

Sintesi dei dati

Dai punti $(x_1, y_1), \ldots, (x_n, y_n)$ distilliamo le 5 quantità

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$s_{xx} := \sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad s_{xy} := \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \qquad s_{yy} := \sum_{i=1}^{n}(y_i - \bar{y})^2$$

Sintesi dei dati

Dai punti $(x_1, y_1), \ldots, (x_n, y_n)$ distilliamo le 5 quantità

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$s_{xx} := \sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad s_{xy} := \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \qquad s_{yy} := \sum_{i=1}^{n}(y_i - \bar{y})^2$$

Tutte le altre quantità della LSL sono funzioni di queste. P. es.:

$$\hat{a} = \bar{y} - \frac{s_{xy}}{s_{xx}}\,\bar{x} \qquad \hat{b} = \frac{s_{xy}}{s_{xx}} \qquad \text{parametri della LSL}$$

Sintesi dei dati

Dai punti $(x_1, y_1), \ldots, (x_n, y_n)$ distilliamo le 5 quantità

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$s_{xx} := \sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad s_{xy} := \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \qquad s_{yy} := \sum_{i=1}^{n}(y_i - \bar{y})^2$$

Tutte le altre quantità della LSL sono funzioni di queste. P. es.:

$$\hat{a} = \bar{y} - \frac{s_{xy}}{s_{xx}}\,\bar{x} \qquad \hat{b} = \frac{s_{xy}}{s_{xx}} \qquad \text{parametri della LSL}$$

$$\hat{y}_i := \hat{a} + \hat{b}x_i = \bar{y} + \frac{s_{xy}}{s_{xx}}(x_i - \bar{x}) \qquad \text{output della LSL}$$

Sintesi dei dati

Dai punti $(x_1, y_1), \ldots, (x_n, y_n)$ distilliamo le 5 quantità

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$s_{xx} := \sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad s_{xy} := \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \qquad s_{yy} := \sum_{i=1}^{n}(y_i - \bar{y})^2$$

Tutte le altre quantità della LSL sono funzioni di queste. P. es.:

$$\hat{a} = \bar{y} - \frac{s_{xy}}{s_{xx}}\,\bar{x} \qquad \hat{b} = \frac{s_{xy}}{s_{xx}} \qquad \text{parametri della LSL}$$

$$\hat{y}_i := \hat{a} + \hat{b}x_i = \bar{y} + \frac{s_{xy}}{s_{xx}}(x_i - \bar{x}) \qquad \text{output della LSL}$$

$$\hat{e}_i := y_i - \hat{y}_i = (y_i - \bar{y}) - \frac{s_{xy}}{s_{xx}}(x_i - \bar{x}) \qquad \text{residui della LSL}$$