Statistica - 16ª lezione

25 maggio 2021

Regressione multipla in 13 righe

Modello	
Era	Sarà
$Y = a + b\,x + E$	$Y = a + b_1 x_1 + \ldots + b_k x_k + E$

Regressione multipla in 13 righe

Modello	
Era	Sarà
$Y = a + b\,x + E$	$Y = a + b_1 x_1 + \ldots + b_k x_k + E$

Variabilità spiegata	
Era	Sarà
r^2	r^2_{adjusted}

Regressione multipla in 13 righe

Modello	
Era	Sarà
$Y = a + b\,x + E$	$Y = a + b_1 x_1 + \ldots + b_k x_k + E$

Variabilità spiegata	
Era	Sarà
r^2	r^2_{adjusted}

Test sui coefficienti	
Erano	Saranno
$H_0 : a = 0 \quad \text{vs.} \quad H_1 : a \neq 0$	$H_0 : a = 0 \quad \text{vs.} \quad H_1 : a \neq 0$
$H_0 : b = 0 \quad \text{vs.} \quad H_1 : b \neq 0$	$H_0 : b_i = 0 \quad \text{vs.} \quad H_1 : b_i \neq 0$
	H_0 : tutti i b_i sono 0 vs. H_1 : almeno uno dei b_i non è 0

Regressione lineare multipla

$$\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

Legge fisica

$$y = a + b_1 x_1 + \ldots + b_k x_k$$

input
(vettore in $\mathbb{R}^k$)

Regressione lineare multipla

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

$\longrightarrow$

input
(vettore in $\mathbb{R}^k$)

Legge fisica

$$y = a + b_1 x_1 + \ldots + b_k x_k$$
$$= a + \mathbf{x}^t \mathbf{b}$$
$$(a \in \mathbb{R},\ \mathbf{b} \in \mathbb{R}^k)$$

Regressione lineare multipla

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

input
(vettore in $\mathbb{R}^k$)

Legge fisica

$$\begin{aligned} y &= a + b_1 x_1 + \ldots + b_k x_k \\ &= a + \mathbf{x}^t \mathbf{b} \end{aligned}$$

$$(a \in \mathbb{R},\ \mathbf{b} \in \mathbb{R}^k)$$

$$Y = a + \mathbf{x}^t \mathbf{b} + E$$

output
(variabile aleatoria)

Regressione lineare multipla

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

input
(vettore in $\mathbb{R}^k$)

Legge fisica

$$y = a + b_1 x_1 + \ldots + b_k x_k$$
$$= a + \mathbf{x}^t \mathbf{b}$$
$$(a \in \mathbb{R},\ \mathbf{b} \in \mathbb{R}^k)$$

$$Y = a + \mathbf{x}^t \mathbf{b} + E$$

output
(variabile aleatoria)

predittori
(o *regressori*)

Regressione lineare multipla

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

input
(vettore in $\mathbb{R}^k$)

Legge fisica

$$y = a + b_1 x_1 + \ldots + b_k x_k$$
$$= a + \mathbf{x}^t \mathbf{b}$$
$$(a \in \mathbb{R},\ \mathbf{b} \in \mathbb{R}^k)$$

$$Y = a + \mathbf{x}^t \mathbf{b} + E$$

output
(variabile aleatoria)

Se facciamo n misure:

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \quad \longrightarrow \quad Y_1, Y_2, \ldots, Y_n \quad \text{con} \quad Y_i = a + \mathbf{x}_i^t \mathbf{b} + E_i$$

Regressione lineare multipla

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

input
(vettore in $\mathbb{R}^k$)

$\longrightarrow$

Legge fisica

$$y = a + b_1 x_1 + \ldots + b_k x_k$$
$$= a + \mathbf{x}^t \mathbf{b}$$
$$(a \in \mathbb{R},\ \mathbf{b} \in \mathbb{R}^k)$$

$\longrightarrow$

$$Y = a + \mathbf{x}^t \mathbf{b} + E$$

output
(variabile aleatoria)

Se facciamo n misure:

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \quad \longrightarrow \quad Y_1, Y_2, \ldots, Y_n \quad \text{con} \quad Y_i = a + \mathbf{x}_i^t \mathbf{b} + E_i$$

Ipotesi fondamentale del modello multilineare:

- $E_1, \ldots, E_n$ indipendenti $\quad \Leftrightarrow \quad$ $Y_1, \ldots, Y_n$ indipendenti
- $E_i \sim N(0, \sigma^2)\ \ \forall i$ $\quad \Leftrightarrow \quad$ $Y_i \sim N(a + \mathbf{x}_i^t \mathbf{b}\, ,\, \sigma^2)\ \ \forall i$

Regressione lineare multipla

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

input
(vettore in $\mathbb{R}^k$)

$\longrightarrow$

Legge fisica

$$y = a + b_1 x_1 + \ldots + b_k x_k$$
$$= a + \mathbf{x}^t \mathbf{b}$$
$$(a \in \mathbb{R},\ \mathbf{b} \in \mathbb{R}^k)$$

$\longrightarrow$

$$Y = a + \mathbf{x}^t \mathbf{b} + E$$

output
(variabile aleatoria)

Se facciamo n misure:

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \quad \longrightarrow \quad Y_1, Y_2, \ldots, Y_n \quad \text{con} \quad Y_i = a + \mathbf{x}_i^t \mathbf{b} + E_i$$

Ipotesi fondamentale del modello multilineare:

- $E_1, \ldots, E_n$ indipendenti $\quad \Leftrightarrow \quad$ $Y_1, \ldots, Y_n$ indipendenti
- $E_i \sim N(0, \sigma^2)\ \ \forall i$ $\quad \Leftrightarrow \quad$ $Y_i \sim N(a + \mathbf{x}_i^t \mathbf{b}\,,\, \sigma^2)\ \ \forall i$

Parametri incogniti: $\quad a, \quad \mathbf{b}, \quad \sigma^2$

Iperpiano dei minimi quadrati

PROBLEMA: Valutare quanto bene l'iperpiano

$$y = a + b_1 x_1 + \ldots + b_k x_k = a + \mathbf{x}^t \mathbf{b}$$

interpola n punti $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ assegnati in $\mathbb{R}^{k+1}$

Iperpiano dei minimi quadrati

PROBLEMA: Valutare quanto bene l'iperpiano

$$y = a + b_1 x_1 + \ldots + b_k x_k = a + \mathbf{x}^t \mathbf{b}$$

interpola n punti $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ assegnati in $\mathbb{R}^{k+1}$

$e_i := y_i - (a + \mathbf{x}_i^t \mathbf{b})$ $\qquad$ *i*-esimo residuo

Iperpiano dei minimi quadrati

PROBLEMA: Valutare quanto bene l'iperpiano

$$y = a + b_1 x_1 + \ldots + b_k x_k = a + \mathbf{x}^t \mathbf{b}$$

interpola n punti $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ assegnati in $\mathbb{R}^{k+1}$

$$e_i := y_i - (a + \mathbf{x}_i^t \mathbf{b}) \qquad i\text{-esimo residuo}$$

$$L(a, \mathbf{b}) := \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[y_i - (a + \mathbf{x}_i^t \mathbf{b}) \right]^2 \qquad \text{funzionale di errore}$$

Iperpiano dei minimi quadrati

PROBLEMA: Valutare quanto bene l'iperpiano

$$y = a + b_1 x_1 + \ldots + b_k x_k = a + \mathbf{x}^t \mathbf{b}$$

interpola n punti $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ assegnati in $\mathbb{R}^{k+1}$

$$e_i := y_i - (a + \mathbf{x}_i^t \mathbf{b}) \qquad i\text{-esimo residuo}$$

$$L(a, \mathbf{b}) := \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[y_i - (a + \mathbf{x}_i^t \mathbf{b}) \right]^2 \qquad \text{funzionale di errore}$$

$$\tilde{\mathbf{x}} = \underbrace{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix}}_{n \times (k+1)} \qquad \boldsymbol{\beta} = \underbrace{\begin{pmatrix} a \\ b_1 \\ \vdots \\ b_k \end{pmatrix}}_{(k+1) \times 1}$$

Iperpiano dei minimi quadrati

PROBLEMA: Valutare quanto bene l'iperpiano

$$y = a + b_1 x_1 + \ldots + b_k x_k = a + \mathbf{x}^t \mathbf{b}$$

interpola n punti $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ assegnati in $\mathbb{R}^{k+1}$

$$e_i := y_i - (a + \mathbf{x}_i^t \mathbf{b}) \qquad i\text{-esimo residuo}$$

$$L(a, \mathbf{b}) := \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[y_i - (a + \mathbf{x}_i^t \mathbf{b}) \right]^2 \qquad \text{funzionale di errore}$$

$$\tilde{\mathbf{x}} = \underbrace{\begin{pmatrix} 1 & \mathbf{x}_1^t \\ 1 & \mathbf{x}_2^t \\ \cdots & \cdots \\ 1 & \mathbf{x}_n^t \end{pmatrix}}_{n \times (k+1)} \qquad \beta = \underbrace{\begin{pmatrix} a \\ \\ \mathbf{b} \\ \\ \end{pmatrix}}_{(k+1) \times 1}$$

Iperpiano dei minimi quadrati

PROBLEMA: Valutare quanto bene l'iperpiano

$$y = a + b_1 x_1 + \ldots + b_k x_k = a + \mathbf{x}^t \mathbf{b}$$

interpola n punti $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ assegnati in $\mathbb{R}^{k+1}$

$$e_i := y_i - (a + \mathbf{x}_i^t \mathbf{b}) = y_i - (\tilde{\mathbf{x}}\boldsymbol{\beta})_i \qquad i\text{-esimo residuo}$$

$$L(a, \mathbf{b}) := \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[y_i - (a + \mathbf{x}_i^t \mathbf{b}) \right]^2 \qquad \text{funzionale di errore}$$

$$= \sum_i (y_i - (\tilde{\mathbf{x}}\boldsymbol{\beta})_i)^2$$

$$\tilde{\mathbf{x}} = \underbrace{\begin{pmatrix} 1 & \mathbf{x}_1^t \\ 1 & \mathbf{x}_2^t \\ \cdots & \cdots \\ 1 & \mathbf{x}_n^t \end{pmatrix}}_{n \times (k+1)} \qquad \boldsymbol{\beta} = \underbrace{\begin{pmatrix} a \\ \\ \mathbf{b} \\ \\ \end{pmatrix}}_{(k+1) \times 1} =: \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

Iperpiano dei minimi quadrati

PROBLEMA: Valutare quanto bene l'iperpiano

$$y = a + b_1 x_1 + \ldots + b_k x_k = a + \mathbf{x}^t \mathbf{b}$$

interpola n punti $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ assegnati in $\mathbb{R}^{k+1}$

$$e_i := y_i - (a + \mathbf{x}_i^t \mathbf{b}) = y_i - (\tilde{\mathbf{x}}\boldsymbol{\beta})_i \qquad i\text{-esimo residuo}$$

$$L(a, \mathbf{b}) := \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[y_i - (a + \mathbf{x}_i^t \mathbf{b})\right]^2 \qquad \text{funzionale di errore}$$

$$= \sum_i (y_i - (\tilde{\mathbf{x}}\boldsymbol{\beta})_i)^2 = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$$

$$\tilde{\mathbf{x}} = \underbrace{\begin{pmatrix} 1 & \mathbf{x}_1^t \\ 1 & \mathbf{x}_2^t \\ \cdots & \cdots \\ 1 & \mathbf{x}_n^t \end{pmatrix}}_{n \times (k+1)} \qquad \boldsymbol{\beta} = \underbrace{\begin{pmatrix} a \\ \\ \mathbf{b} \\ \\ \end{pmatrix}}_{(k+1)\times 1} =: \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \qquad \mathbf{y} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{n \times 1}$$

Iperpiano dei minimi quadrati

Minimizziamo $L(\beta) = \|\mathbf{y} - \tilde{\mathbf{x}}\beta\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \; = \; ?????$$

Iperpiano dei minimi quadrati

Minimizziamo $L(\beta) = \|\mathbf{y} - \tilde{\mathbf{x}}\beta\|^2$ rispetto a β:

$$\nabla_\beta L(\beta) = -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\beta)$$

Iperpiano dei minimi quadrati

Minimizziamo $L(\beta) = \|\mathbf{y} - \tilde{\mathbf{x}}\beta\|^2$ rispetto a β:

$$\nabla_{\beta} L(\beta) \ = \ -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\beta) \ = \ -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\beta$$

Iperpiano dei minimi quadrati

Minimizziamo $L(\beta) = \|\mathbf{y} - \tilde{\mathbf{x}}\beta\|^2$ rispetto a β:

$$\nabla_{\beta} L(\beta) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\beta) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\beta \;\equiv\; 0$$

Iperpiano dei minimi quadrati

Minimizziamo $L(\boldsymbol{\beta}) = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\boldsymbol{\beta} \;\equiv\; 0$$

$$\overset{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}}{\Longrightarrow} \qquad \boldsymbol{\beta} = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y}$$

Iperpiano dei minimi quadrati

Minimizziamo $L(\beta) = \|\mathbf{y} - \tilde{\mathbf{x}}\beta\|^2$ rispetto a β:

$$\nabla_\beta L(\beta) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\beta) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\beta \;\equiv\; 0$$

$$\overset{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}}{\Longrightarrow} \qquad \beta = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y}$$

P. es., con $k = 1$ predittore:

$$\tilde{\mathbf{x}}^t\tilde{\mathbf{x}} = \begin{pmatrix} n & n\bar{x} \\ n\bar{x} & s_{xx} + n\bar{x}^2 \end{pmatrix} \quad \text{è invertibile} \quad \Leftrightarrow \quad s_{xx} \neq 0$$

Altrimenti: *collinearità*

Iperpiano dei minimi quadrati

Minimizziamo $L(\boldsymbol{\beta}) = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\boldsymbol{\beta} \;\equiv\; 0$$

$$\xrightarrow{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}} \qquad \boldsymbol{\beta} = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y} =: \hat{\boldsymbol{\beta}}$$

$$y \;=\; \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k \;=\; (1\ \mathbf{x}^t)\,\hat{\boldsymbol{\beta}}$$

iperpiano dei minimi quadrati (LSH)

Iperpiano dei minimi quadrati

Minimizziamo $L(\boldsymbol{\beta}) = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\boldsymbol{\beta} \;\equiv\; 0$$

$$\xrightarrow{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}} \qquad \boldsymbol{\beta} = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y} =: \hat{\boldsymbol{\beta}}$$

$$y \;=\; \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k \;=\; (1\ \mathbf{x}^t)\,\hat{\boldsymbol{\beta}}$$

iperpiano dei minimi quadrati (LSH)

Come nel caso semplice:

$$\hat{y}_i \;:=\; (1\ \mathbf{x}_i^t)\hat{\boldsymbol{\beta}}$$

output dell'LSH

Iperpiano dei minimi quadrati

Minimizziamo $L(\boldsymbol{\beta}) = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\boldsymbol{\beta} \;\equiv\; 0$$

$$\overset{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}}{\Longrightarrow} \qquad \boldsymbol{\beta} = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y} =: \hat{\boldsymbol{\beta}}$$

$$y \;=\; \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k \;=\; (1\ \mathbf{x}^t)\,\hat{\boldsymbol{\beta}}$$

iperpiano dei minimi quadrati (LSH)

Come nel caso semplice:

$$\hat{y}_i \;:=\; (1\ \mathbf{x}_i^t)\hat{\boldsymbol{\beta}} \qquad \text{output dell'LSH}$$

$$\hat{e}_i \;:=\; y_i - \hat{y}_i \qquad \text{residui dell'LSH}$$

Iperpiano dei minimi quadrati

Minimizziamo $L(\boldsymbol{\beta}) = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\boldsymbol{\beta} \;\equiv\; 0$$

$$\overset{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}}{\Longrightarrow} \qquad \boldsymbol{\beta} = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y} =: \hat{\boldsymbol{\beta}}$$

$$y \;=\; \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k \;=\; (1\ \mathbf{x}^t)\,\hat{\boldsymbol{\beta}}$$

iperpiano dei minimi quadrati (LSH)

Come nel caso semplice:

$$\hat{y}_i \;:=\; (1\ \mathbf{x}_i^t)\hat{\boldsymbol{\beta}}$$ output dell'LSH

$$\hat{e}_i \;:=\; y_i - \hat{y}_i$$ residui dell'LSH

$$ss_r \;:=\; \sum(\bar{y} - \hat{y}_i)^2$$ varianza spiegata

Iperpiano dei minimi quadrati

Minimizziamo $L(\boldsymbol{\beta}) = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\boldsymbol{\beta} \;\equiv\; 0$$

$$\xRightarrow{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}} \qquad \boldsymbol{\beta} = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y} =: \hat{\boldsymbol{\beta}}$$

$$y \;=\; \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k \;=\; (1\ \mathbf{x}^t)\,\hat{\boldsymbol{\beta}}$$

iperpiano dei minimi quadrati (LSH)

Come nel caso semplice:

$$\begin{aligned} \hat{y}_i &:= (1\ \mathbf{x}_i^t)\hat{\boldsymbol{\beta}} && \text{output dell'LSH} \\ \hat{e}_i &:= y_i - \hat{y}_i && \text{residui dell'LSH} \\ ss_r &:= \sum(\bar{y} - \hat{y}_i)^2 && \text{varianza spiegata} \\ ss_e &:= \sum(y_i - \hat{y}_i)^2 && \text{varianza residua} \end{aligned}$$

Iperpiano dei minimi quadrati

Minimizziamo $L(\boldsymbol{\beta}) = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\boldsymbol{\beta} \;\equiv\; 0$$

$$\xRightarrow{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}} \qquad \boldsymbol{\beta} = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y} =: \hat{\boldsymbol{\beta}}$$

$$y \;=\; \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k \;=\; (1\ \mathbf{x}^t)\,\hat{\boldsymbol{\beta}}$$

iperpiano dei minimi quadrati (LSH)

Come nel caso semplice:

$\hat{y}_i := (1\ \mathbf{x}_i^t)\hat{\boldsymbol{\beta}}$		output dell'LSH
$\hat{e}_i := y_i - \hat{y}_i$		residui dell'LSH
$ss_r := \sum(\bar{y} - \hat{y}_i)^2$		varianza spiegata
$ss_e := \sum(y_i - \hat{y}_i)^2$		varianza residua
$ss_t := \sum(y_i - \bar{y})^2$		varianza totale

Iperpiano dei minimi quadrati

Minimizziamo $L(\boldsymbol{\beta}) = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2$ rispetto a $\boldsymbol{\beta}$:

$$\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t(\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}) \;=\; -2\tilde{\mathbf{x}}^t\mathbf{y} + 2\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}\boldsymbol{\beta} \;\equiv\; 0$$

$$\stackrel{\text{se } \tilde{\mathbf{x}}^t\tilde{\mathbf{x}} \text{ è invertibile}}{\Longrightarrow} \qquad \boldsymbol{\beta} = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y} =: \hat{\boldsymbol{\beta}}$$

$$y \;=\; \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k \;=\; (1\ \mathbf{x}^t)\,\hat{\boldsymbol{\beta}}$$

iperpiano dei minimi quadrati (LSH)

Come nel caso semplice:

$\hat{y}_i \;:=\; (1\ \mathbf{x}_i^t)\hat{\boldsymbol{\beta}}$	output dell'LSH
$\hat{e}_i \;:=\; y_i - \hat{y}_i$	residui dell'LSH
$ss_r \;:=\; \sum(\bar{y} - \hat{y}_i)^2$	varianza spiegata
$ss_e \;:=\; \sum(y_i - \hat{y}_i)^2$	varianza residua
$ss_t \;:=\; \sum(y_i - \bar{y})^2 \;\equiv\; ss_r + ss_e$	varianza totale

Iperpiano dei minimi quadrati

OVERFITTING = ridurre ss_e aumentando i regressori

Iperpiano dei minimi quadrati

OVERFITTING = ridurre ss_e aumentando i regressori:

$$ss_e = L(\hat{\beta}) = \min_{\beta} L(\beta)$$

Iperpiano dei minimi quadrati

OVERFITTING = ridurre ss_e aumentando i regressori:

$$ss_e \;=\; L(\hat{\boldsymbol{\beta}}) \;=\; \min_{\boldsymbol{\beta}} L(\boldsymbol{\beta})$$

$$L_k(\boldsymbol{\beta}) \;=\; \sum_i (y_i - \beta_0 - \beta_1 x_{i,1} - \ldots - \beta_k x_{i,k})^2 \qquad \boldsymbol{\beta} \in \mathbb{R}^{k+1}$$

$$L_{k+1}(\boldsymbol{\beta}') \;=\; \sum_i (y_i - \beta'_0 - \beta'_1 x_{i,1} - \ldots - \beta'_{k+1} x_{i,k+1})^2 \qquad \boldsymbol{\beta}' \in \mathbb{R}^{k+2}$$

Iperpiano dei minimi quadrati

OVERFITTING = ridurre ss_e aumentando i regressori:

$$ss_e \;=\; L(\hat{\boldsymbol{\beta}}) \;=\; \min_{\boldsymbol{\beta}} L(\boldsymbol{\beta})$$

$$L_k(\boldsymbol{\beta}) \;=\; \sum_i (y_i - \beta_0 - \beta_1 x_{i,1} - \ldots - \beta_k x_{i,k})^2 \qquad \boldsymbol{\beta} \in \mathbb{R}^{k+1}$$

$$L_{k+1}(\boldsymbol{\beta}') \;=\; \sum_i (y_i - \beta_0' - \beta_1' x_{i,1} - \ldots - \beta_{k+1}' x_{i,k+1})^2 \qquad \boldsymbol{\beta}' \in \mathbb{R}^{k+2}$$

$$L_k(\boldsymbol{\beta}) \;\equiv\; L_{k+1}\begin{pmatrix} \boldsymbol{\beta} \\ 0 \end{pmatrix}$$

Iperpiano dei minimi quadrati

OVERFITTING = ridurre ss_e aumentando i regressori:

$$ss_e = L(\hat{\boldsymbol{\beta}}) = \min_{\boldsymbol{\beta}} L(\boldsymbol{\beta})$$

$$L_k(\boldsymbol{\beta}) = \sum_i (y_i - \beta_0 - \beta_1 x_{i,1} - \ldots - \beta_k x_{i,k})^2 \qquad \boldsymbol{\beta} \in \mathbb{R}^{k+1}$$

$$L_{k+1}(\boldsymbol{\beta}') = \sum_i (y_i - \beta_0' - \beta_1' x_{i,1} - \ldots - \beta_{k+1}' x_{i,k+1})^2 \qquad \boldsymbol{\beta}' \in \mathbb{R}^{k+2}$$

$$L_k(\boldsymbol{\beta}) \equiv L_{k+1}\begin{pmatrix} \boldsymbol{\beta} \\ 0 \end{pmatrix} \quad \Rightarrow \quad \min_{\boldsymbol{\beta}} L_k(\boldsymbol{\beta}) \geq \min_{\boldsymbol{\beta}'} L_{k+1}(\boldsymbol{\beta}')$$

Iperpiano dei minimi quadrati

OVERFITTING = ridurre ss_e aumentando i regressori:

$$ss_e \;=\; L(\hat{\boldsymbol{\beta}}) \;=\; \min_{\boldsymbol{\beta}} L(\boldsymbol{\beta})$$

$$L_k(\boldsymbol{\beta}) \;=\; \sum_i (y_i - \beta_0 - \beta_1 x_{i,1} - \ldots - \beta_k x_{i,k})^2 \qquad \boldsymbol{\beta} \in \mathbb{R}^{k+1}$$

$$L_{k+1}(\boldsymbol{\beta}') \;=\; \sum_i (y_i - \beta'_0 - \beta'_1 x_{i,1} - \ldots - \beta'_{k+1} x_{i,k+1})^2 \qquad \boldsymbol{\beta}' \in \mathbb{R}^{k+2}$$

$$L_k(\boldsymbol{\beta}) \;\equiv\; L_{k+1}\begin{pmatrix} \boldsymbol{\beta} \\ 0 \end{pmatrix} \quad \Rightarrow \quad \min_{\boldsymbol{\beta}} L_k(\boldsymbol{\beta}) \geq \min_{\boldsymbol{\beta}'} L_{k+1}(\boldsymbol{\beta}')$$

Perciò si preferisce usare l'*r^2-adjusted*

$$r_A^2 \;:=\; 1 - \frac{ss_e}{ss_t}\,\frac{n-1}{n-1-k}$$

Iperpiano dei minimi quadrati

OVERFITTING = ridurre ss_e aumentando i regressori:

$$ss_e \; = \; L(\hat{\boldsymbol{\beta}}) \; = \; \min_{\boldsymbol{\beta}} L(\boldsymbol{\beta})$$

$$L_k(\boldsymbol{\beta}) \; = \; \sum_i (y_i - \beta_0 - \beta_1 x_{i,1} - \ldots - \beta_k x_{i,k})^2 \qquad \boldsymbol{\beta} \in \mathbb{R}^{k+1}$$

$$L_{k+1}(\boldsymbol{\beta}') \; = \; \sum_i (y_i - \beta_0' - \beta_1' x_{i,1} - \ldots - \beta_{k+1}' x_{i,k+1})^2 \qquad \boldsymbol{\beta}' \in \mathbb{R}^{k+2}$$

$$L_k(\boldsymbol{\beta}) \; \equiv \; L_{k+1}\begin{pmatrix} \boldsymbol{\beta} \\ 0 \end{pmatrix} \quad \Rightarrow \quad \min_{\boldsymbol{\beta}} L_k(\boldsymbol{\beta}) \geq \min_{\boldsymbol{\beta}'} L_{k+1}(\boldsymbol{\beta}')$$

Perciò si preferisce usare l'r^2-*adjusted*

$$r_A^2 \; := \; 1 - \frac{ss_e}{ss_t} \, \frac{n-1}{n-1-k} \begin{cases} \leq r^2 \\ \text{decrescente in } k \end{cases}$$