

# Statistica - 16<sup>a</sup> lezione

25 maggio 2021

# Regressione multipla in 13 righe

Modello	
Era	Sarà
$Y = a + bx + E$	$Y = a + b_1x_1 + \dots + b_kx_k + E$

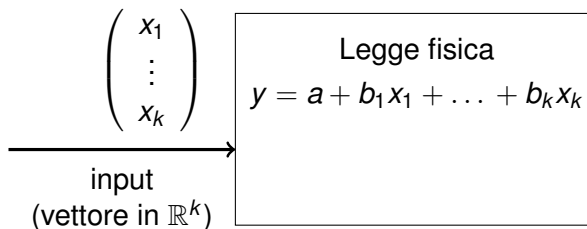
# Regressione multipla in 13 righe

Modello	
Era	Sarà
$Y = a + b x + E$	$Y = a + b_1 x_1 + \dots + b_k x_k + E$
Variabilità spiegata	
Era	Sarà
$r^2$	$r^2_{\text{adjusted}}$

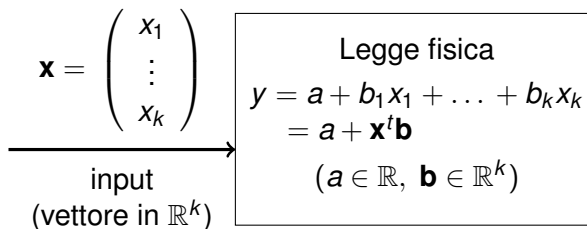
# Regressione multipla in 13 righe

Modello	
Era	Sarà
$Y = a + bx + E$	$Y = a + b_1x_1 + \dots + b_kx_k + E$
Variabilità spiegata	
Era	Sarà
$r^2$	$r^2_{\text{adjusted}}$
Test sui coefficienti	
Erano	Saranno
$H_0 : a = 0$ vs. $H_1 : a \neq 0$	$H_0 : a = 0$ vs. $H_1 : a \neq 0$
$H_0 : b = 0$ vs. $H_1 : b \neq 0$	$H_0 : b_i = 0$ vs. $H_1 : b_i \neq 0$
	$H_0 : \text{tutti i } b_i \text{ sono } 0$ vs. $H_1 : \text{almeno uno dei } b_i \text{ non è } 0$

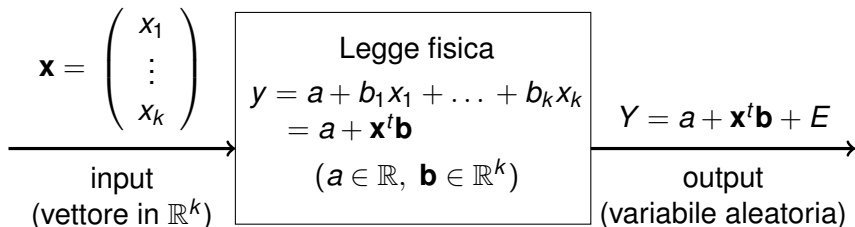
# Regressione lineare multipla



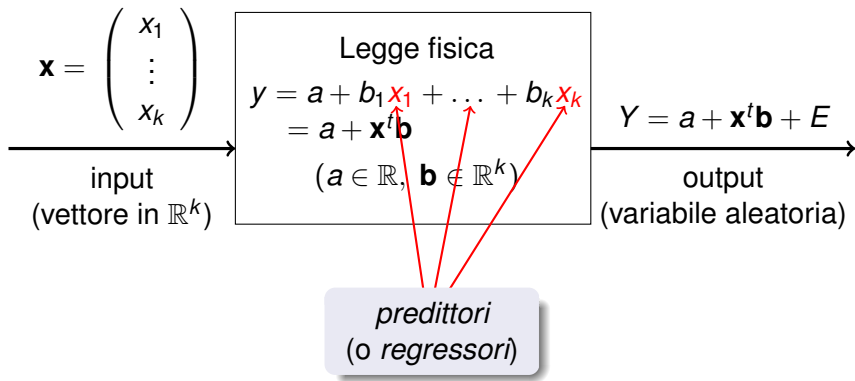
# Regressione lineare multipla



# Regressione lineare multipla

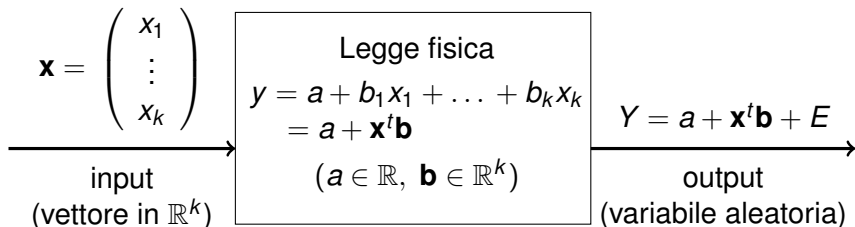


# Regressione lineare multipla





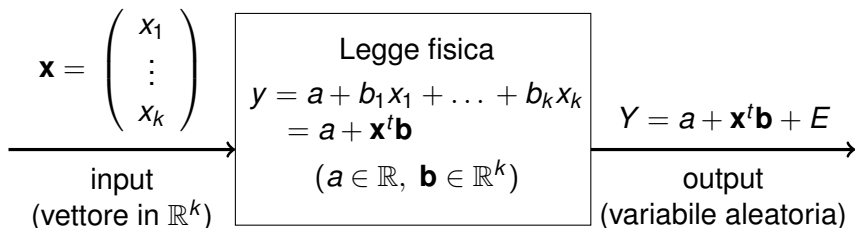
# Regressione lineare multipla



Se facciamo  $n$  misure:

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

# Regressione lineare multipla



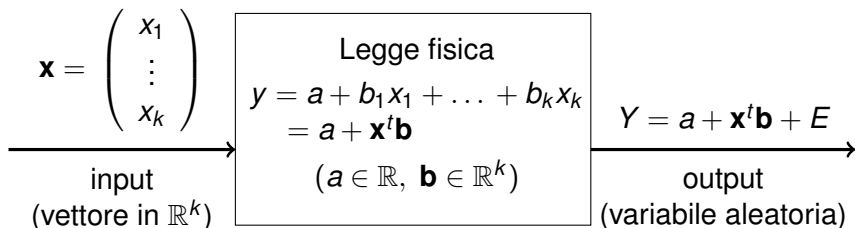
Se facciamo  $n$  misure:

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

Ipotesi fondamentale del modello multilineare:

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

# Regressione lineare multipla



Se facciamo  $n$  misure:

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

Ipotesi fondamentale del modello multilineare:

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

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

# Iperpiano dei minimi quadrati

**PROBLEMA:** Valutare quanto bene l'iperpiano

$$y = a + b_1x_1 + \dots + b_kx_k = a + \mathbf{x}^t\mathbf{b}$$

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

# Iperpiano dei minimi quadrati

**PROBLEMA:** Valutare quanto bene l'iperpiano

$$y = a + b_1x_1 + \dots + b_kx_k = a + \mathbf{x}^t\mathbf{b}$$

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

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

# Iperpiano dei minimi quadrati

**PROBLEMA:** Valutare quanto bene l'iperpiano

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

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

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

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

# Iperpiano dei minimi quadrati

**PROBLEMA:** Valutare quanto bene l'iperpiano

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

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

$$e_i := y_i - (a + \mathbf{x}_i^t \mathbf{b})$$

$i$ -esimo residuo

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

funzionale di errore

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

# Iperpiano dei minimi quadrati

**PROBLEMA:** Valutare quanto bene l'iperpiano

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

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

$$e_i := y_i - (a + \mathbf{x}_i^t \mathbf{b})$$

$i$ -esimo residuo

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

funzionale di errore

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



# Iperpiano dei minimi quadrati

**PROBLEMA:** Valutare quanto bene l'iperpiano

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

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

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

$$\begin{aligned} L(a, \mathbf{b}) &:= \sum_{i=1}^n e_i^2 = \sum_{i=1}^n [y_i - (a + \mathbf{x}_i^t \mathbf{b})]^2 && \text{funzionale di errore} \\ &= \sum_i (y_i - (\tilde{\mathbf{x}}\boldsymbol{\beta})_i)^2 \end{aligned}$$

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

# Iperpiano dei minimi quadrati

**PROBLEMA:** Valutare quanto bene l'iperpiano

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

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

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

$$\begin{aligned} L(a, \mathbf{b}) &:= \sum_{i=1}^n e_i^2 = \sum_{i=1}^n [y_i - (a + \mathbf{x}_i^t \mathbf{b})]^2 && \text{funzionale di errore} \\ &= \sum_i (y_i - (\tilde{\mathbf{x}}\boldsymbol{\beta})_i)^2 = \|\mathbf{y} - \tilde{\mathbf{x}}\boldsymbol{\beta}\|^2 \end{aligned}$$

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

# Iperpiano dei minimi quadrati

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

$$\nabla_{\beta} L(\beta) = \text{?????}$$

# Iperpiano dei minimi quadrati

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

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

# Iperpiano dei minimi quadrati

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

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

# Iperpiano dei minimi quadrati

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

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

# Iperpiano dei minimi quadrati

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

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

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

# Iperpiano dei minimi quadrati

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

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

se  $\tilde{\mathbf{x}}^t\tilde{\mathbf{x}}$  è invertibile  $\implies \beta = (\tilde{\mathbf{x}}^t\tilde{\mathbf{x}})^{-1}\tilde{\mathbf{x}}^t\mathbf{y}$

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

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

Altrimenti: *collinearità*



# Iperpiano dei minimi quadrati

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

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

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

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k = (1 \ \mathbf{x}^t) \hat{\beta} \quad \text{iperpiano dei minimi quadrati (LSH)}$$

# Iperpiano dei minimi quadrati

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

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

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

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k = (1 \ \mathbf{x}^t) \hat{\beta} \quad \text{iperpiano dei minimi quadrati (LSH)}$$

Come nel caso semplice:

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

# Iperpiano dei minimi quadrati

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

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

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

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k = (1 \ \mathbf{x}^t) \hat{\beta} \quad \text{iperpiano dei minimi quadrati (LSH)}$$

Come nel caso semplice:

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

output dell'LSH

$$\hat{e}_i := y_i - \hat{y}_i$$

residui dell'LSH

# Iperpiano dei minimi quadrati

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

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

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

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k = (1 \ \mathbf{x}^t) \hat{\beta} \quad \text{iperpiano dei minimi quadrati (LSH)}$$

Come nel caso semplice:

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

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

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

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

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k = (1 \ \mathbf{x}^t) \hat{\beta} \quad \text{iperpiano dei minimi quadrati (LSH)}$$

Come nel caso semplice:

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

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

# Iperpiano dei minimi quadrati

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

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

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

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k = (1 \ \mathbf{x}^t) \hat{\beta} \quad \text{iperpiano dei minimi quadrati (LSH)}$$

Come nel caso semplice:

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

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

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

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

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k = (1 \ \mathbf{x}^t) \hat{\beta} \quad \text{iperpiano dei minimi quadrati (LSH)}$$

Come nel caso semplice:

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

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

**OVERFITTING** = ridurre  $ss_e$  aumentando i regressori



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:

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

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

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

# Iperpiano dei minimi quadrati

OVERFITTING = ridurre  $ss_e$  aumentando i regressori:

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

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

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

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

# Iperpiano dei minimi quadrati

OVERFITTING = ridurre  $ss_e$  aumentando i regressori:

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

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

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

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

# Iperpiano dei minimi quadrati

OVERFITTING = ridurre  $ss_e$  aumentando i regressori:

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

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

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

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

Perciò si preferisce usare l'*r<sup>2</sup>-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:

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

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

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

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

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

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