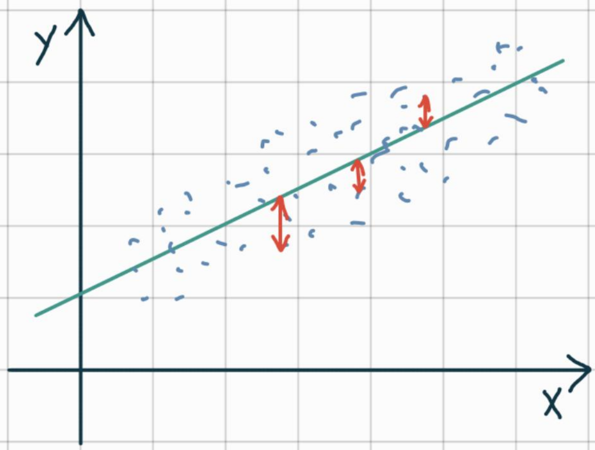
UVOD

(Francis Galton, 1822-1911)



x - pojasnjevalna / explanatory
y - odvisna/dependent

To je enostavna linearna regresija, ki se računa po predlogi $y = a + bx$, kjer sta a, b koeficienta/parametra.

Splošna linearna regresija:

$$y = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p$$

$x_1, x_2, \ldots, x_p$ ... pojasnjevalne spremenljivke
y ... odvisna spremenljivka
$\beta_1, \beta_2, \ldots, \beta_p$ ... koeficienti/parametri

Podatki:

$x_{11}, x_{12}, \ldots, x_{1p} \; ; \; y_1$

$x_{21}, x_{22}, \ldots, x_{2p} \; ; \; y_2$

$\vdots$

$x_{n1}, x_{n2}, \ldots, x_{np} \; ; \; y_n$

$$\Rightarrow y_1 = \beta_1 x_{11} + \ldots + \beta_p x_{1p}$$
$$\vdots$$
$$y_n = \beta_n x_{n1} + \ldots + \beta_p x_{np}$$

$n > p \Rightarrow$ Predoločen sistem

$$\underline{\underline{X}} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}, \quad \underline{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

$$\Rightarrow \underline{y} = \underline{\underline{X}}\,\underline{\beta}$$

TOČKOVNO OCENJEVANJE KOEFICIENTOV

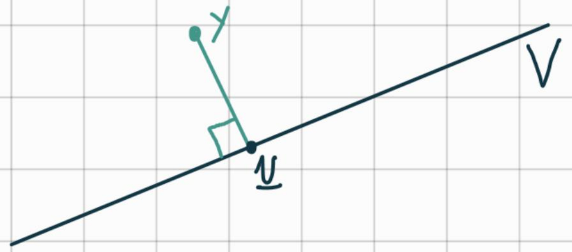
Uporabimo metodo najmanjših kvadratov (ordinary least squares, OLE):

$$\underline{y} = \underline{\underline{X}}\,\underline{\beta} + \underline{\varepsilon}$$

$\underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$ je vektor šumov.

$$V := \operatorname{Im} \underline{\underline{X}} \subseteq \mathbb{R}^n$$

Poiščemo $\underline{v} \in V$, za katerega je $\|\underline{y} - \underline{v}\|^2$ minimalna.



Nato pa rešimo $\underline{v} = \underline{\underline{X}}\,\underline{\beta}$.

Privzememo, da ima $\underline{\underline{X}}$ ničelno jedro, tako da je rešitev enolična.

Računsko pride:

$$\underline{\beta} = (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T \underline{y}$$

Statistični pogled:

$$\underbrace{\underline{Y}}_{\substack{\text{Slučajen}\\ \text{opazljiv}}} = \underbrace{\underline{\underline{X}}}_{\substack{\text{konstantna}\\ \text{matrika}\\ \text{(design}\\ \text{matrix)}}}\ \underbrace{\underline{\beta}}_{\substack{\text{parameter}\\ \text{(determinističen}\\ \text{poznan)}}} + \underbrace{\underline{\varepsilon}}_{\substack{\text{šum}\\ \text{(slučajen}\\ \text{neznan)}}}$$

Minimalne predpostavke:

$\underline{\varepsilon}$ je beli šum:

$$E(\underline{\varepsilon}) = 0$$

$$Var(\underline{\varepsilon}) = \sigma^2 \underline{\underline{I}} \quad \text{(homoskledastičnost, nekoreliranost)}$$

V luči statistike iščemo cenilko $\hat{\underline{\beta}}$ parametra $\underline{\beta}$.

Smiselna možnost:

$$\hat{\underline{\beta}} = (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T \underline{Y} \quad \text{iz metode najmanjših kvadratov}$$

Kako dobra je ta cenilka?

Če je $\underline{c} \in \mathbb{R}^p$ konstanten vektor, kako dobra je cenilka $\underline{c}^T \hat{\underline{\beta}}$ karakteristike $\underline{c}^T \underline{\beta}$?

Definicija: Linearna ali afina cenilka za količino g z vrednostmi v $\mathbb{R}$ na podlagi opažanja $\underline{Y}$ z vrednostmi v $\mathbb{R}^n$ je cenilka oblike $\hat{g} = \underline{\Psi}^T \underline{Y}^T + \rho$, kjer je $\underline{\Psi} \in \mathbb{R}^n$ konstanten vektor in $\rho \in \mathbb{R}$ konstanta.

Definicija: Najboljša nepristranska linearna cenilka (NNLC, best linear unbiased estimator) za g je taka nepristranska linearna cenilka $\hat{g}$, da za poljubno alternativno nepristransko linearno ceniko Q velja $\mathrm{Var}_\theta(Q) > \mathrm{Var}_\theta(g)$ za vse θ.

Izrek 9.1 (Gauss, Markov):

Če je $\underline{c} \in \mathbb{R}^p$ konstanten vektor, je $\hat{g} := \underline{c}^T \hat{\underline{\beta}}$ edina NNLC za $g := \underline{c}^T \beta$.

Dokaz opustimo.

Minimiziramo $\|\underline{Y} - \underline{\underline{X}}\underline{\beta}\|^2$ in pride $\|(\underline{\underline{I}}_n - \underline{\underline{H}})\|^2 = \sum_{i=1}^{n} (Y_i - \sum_{j=1}^{p} x_{ij} \hat{\beta}_j)^2$.

$\underline{\underline{H}} = \underline{\underline{X}}(\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T$ je ortogonalni projektor na $\mathrm{Im}\, \underline{\underline{X}}$.

$\underline{Y} = \underline{\underline{X}}\underline{\beta} + \underline{\varepsilon}$

$\underline{Y} = \underline{\underline{X}}\hat{\underline{\beta}} + \hat{\underline{\varepsilon}}$

$\hat{\underline{\varepsilon}}$ je vektor rezidualov ali ostankov.

$$\|\underbrace{(\underline{\underline{I}}_n - \underline{\underline{H}})\underline{Y}}_{\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}}}\|^2 = \|\hat{\underline{\varepsilon}}\|^2 = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 =: RSS$$

Trditev 9.2: $\hat{\sigma}_+^2 = \frac{RSS}{n-p}$ je nepristranska cenilka za σ^2.

Dokaz opustimo.

Močnejše predpostavke:

Dodamo $\underline{\varepsilon} \sim N(0, \sigma^2 \underline{\underline{I}}_n)$.

Trditev 9.3: Tedaj je $\hat{\beta}$ tudi cenilka po metodi največjega verjetja.

$$\mathrm{var}(\underline{c}^T \hat{\underline{\beta}}) = \sigma \, \underline{c}^T (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{c}$$

$$SE = \sqrt{\mathrm{var}(\underline{c}^T \hat{\underline{\beta}})}$$

Trditev 9.4: $\frac{\underline{c}^T \hat{\underline{\beta}} - \underline{c}^T \underline{\beta}}{SE} \sim N(0,1)$

Trditev 9.5: $\hat{\beta}$ in $\hat{\sigma}_+$ sta neodvisni.

Posledica 9.6: $\frac{\underline{c}^T \hat{\underline{\beta}} - \underline{c}^T \underline{\beta}}{\hat{SE}} \sim \mathrm{Student}(n-p)$

$$\hat{SE} = \hat{\sigma}_+ \sqrt{\underline{c}^T (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{c}}$$