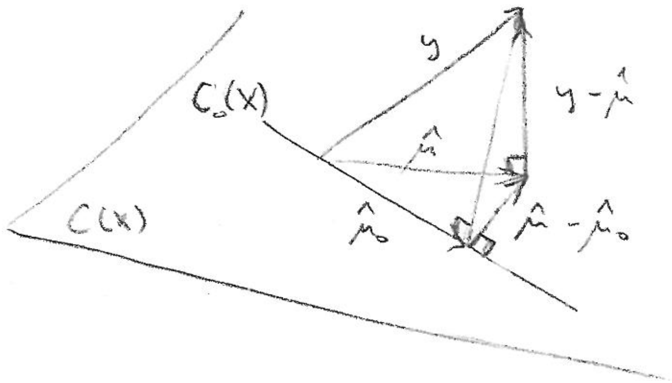
Estimation under bibetingelser

Lineære bånd på parametrene kan udtrykkes som

$$H\beta = 0, \quad H_{q\times p} \text{ med fuld rang } q \le p$$

Dette begrænser μ til at ligge i underrummet $C_0(X) = C(X) \cap N(H(X^TX)^{-1}X^T)$*, $\dim C_0(X) = p - q$.



MK-estimaterne under bibetingelserne kan findes ved hjælp af Lagranges multiplikatormetode:

$$f(\beta;\alpha) = (y - X\beta)^T(y - X\beta) - (H\beta)^T(2\alpha)$$

2α er vektoren bestående af multiplikatorerne (totallet er medtaget for at give simplere regninger).

* $X\beta = \mu \Rightarrow X^TX\beta = X^T\mu \Rightarrow \beta = (X^TX)^{-1}X^T\mu$

$H\beta = 0 \Rightarrow H(X^TX)^{-1}X^T\mu = 0 \Rightarrow \mu \in N(H(X^TX)^{-1}X^T)$

Differentiation mht. β og derefter sat lig 0 giver (efter division med 2)

$$X^T X \beta = X^T y + H^T \alpha \qquad (1)$$

Desuden gælder

$$H\beta = 0 \qquad (2)$$

Multiplikation af (1) med $(X^T X)^{-1}$ giver

$$\beta = (X^T X)^{-1} X^T y + (X^T X)^{-1} H^T \alpha$$

dvs.

$$\beta = \hat{\beta} + (X^T X)^{-1} H^T \alpha , \qquad (3)$$

som indsættes i (2):

$$H(\hat{\beta} + (X^T X)^{-1} H^T \alpha) = 0$$

$$\Rightarrow \quad H(X^T X)^{-1} H^T \alpha = - H\hat{\beta}$$

$$\Rightarrow \quad \alpha = -(H(X^T X)^{-1} H^T)^{-1} H\hat{\beta}$$

Når α indsættes i (3) får vi løsningen

$$\hat{\beta}_0 = \hat{\beta} - (X^T X)^{-1} H^T \underbrace{(H(X^T X)^{-1} H^T)^{-1}}_{K} H\hat{\beta} \qquad (4)$$

Heraf

$$\hat{\mu}_0 = X\hat{\beta}_0 = X\hat{\beta} - X(X^T X)^{-1} H^T K H \hat{\beta}$$

$$= \hat{\mu} - \underbrace{X(X^T X)^{-1} H^T K H (X^T X)^{-1} X^T}_{P_H} y$$

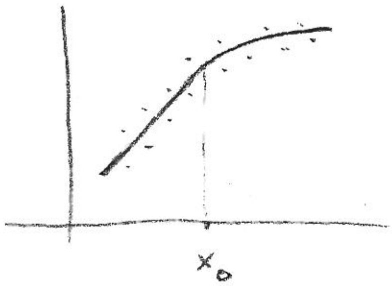
$$= P y - P_H y$$
$$= (P - P_H) y$$
$$= P_0 y \qquad\qquad P_H = P - P_0$$

P_H og P_0 er projektionsmatricer.

$P_H \sim$ projektion på $C_0(X)^\perp \cap C(X)$

$P_0 \sim$ projektion på $C_0(X)$

eks.



$$y = r(x) + u$$

$$r(x) = \begin{cases} \beta_1 + \beta_2 x & \text{for } x \le x_0 \\ \beta_3 + \beta_4 x + \beta_5 x^2 & \text{for } x > x_0 \end{cases}$$

bibetingelser (lineære i parametrene):

$$\beta_1 + \beta_2 x_0 = \beta_3 + \beta_4 x_0 + \beta_5 x_0^2 \quad (r \text{ kont. i } x_0)$$
$$\beta_2 = \beta_4 + 2\beta_5 x_0 \quad (r' \text{ kont. i } x_0)$$

designmatrix:

$$X = \begin{bmatrix} 1 & x_1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & 0 & 0 & 0 \\ 0 & 0 & 1 & x_{m+1} & x_{m+1}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & x_n & x_n^2 \end{bmatrix}$$

Bemærk, at $X = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}$, dvs.

(β_1, β_2) og $(\beta_3, \beta_4, \beta_5)$ kan estimeres hver for sig; kombineret får vi $\hat{\beta}$.

De lineære bånd kan beskrives ved hjælp af matricen

$$H = \begin{bmatrix} 1 & x_0 & -1 & -x_0 & -x_0^2 \\ 0 & 1 & 0 & -1 & -2x_0 \end{bmatrix},$$

som kan indsættes i (4) til bestemmelse af $\hat{\beta}_0$. $\square$

Projektion på $C(X)$ og $C_0(X)$, jf. figur side 1.

Bemærk, at

$$y - \hat{\mu} \perp \hat{\mu}$$
$$y - \hat{\mu}_0 \perp \hat{\mu}_0$$
$$\hat{\mu} - \hat{\mu}_0 \perp \hat{\mu}_0$$

Desuden:

$$y = \hat{\mu}_0 + (\hat{\mu} - \hat{\mu}_0) + (y - \hat{\mu})$$

er en ortogonal dekomposition af y.

Heraf

$$\|y\|^2 = \|\hat{\mu}_0\|^2 + \|\hat{\mu} - \hat{\mu}_0\|^2 + \|y - \hat{\mu}\|^2$$

(Pythagoras)

Ækvivalent skrivemåde:

$$y^T y = y^T P_0 y + y^T (P - P_0) y + y^T (I - P) y$$

Bemærk, at

$$P_0 (P - P_0) = P_0 P - P_0^2 = P_0 - P_0 = 0$$

$$P_0 (I - P) = P_0 - P_0 P = P_0 - P_0 = 0$$

$$(P - P_0)(I - P) = P - P^2 - P_0 - P_0 P$$

$$= P - P - P_0 + P_0 = 0$$

og at

$$\text{rang}\,(P - P_0) = \text{rang}\,P_H = q$$

$$\text{rang}\,P_0 = \text{rang}\,(P - (P - P_0)) = r - q$$

Notation

De enkelte led i opspaltningen af $\|y\|^2$ betegnes ofte som kvadratsummer (eng. sum of squares), forkortet SS med et passende indeks:

$$\|y\|^2 = SS_{tot} \quad (\text{tot for total})$$

$$\|\hat{\mu}\|^2 = SS_{reg} \quad (\text{reg for regression})$$

$$\|\hat{\mu}_0\|^2 = SS_{reg(C_0)}$$

$$\|\hat{\mu} - \hat{\mu}_0\|^2 = SS_{reg(\perp C_0)}$$

$$\|y - \hat{\mu}\|^2 = SS_{res} \quad (\text{res for residual})$$

Altså

$$SS_{tot} = SS_{reg} + SS_{res} = SS_{reg(C_0)} + SS_{reg(\perp C_0)} + SS_{res}$$

Lineære normale modeller

Antag, at $U \sim N_n(0, \sigma^2 I_n)$

dvs. $Y \sim N_n(\underset{X\beta}{\underset{\uparrow}{\mu}}, \sigma^2 I_n)$

parametre : $\theta = (\beta_1, \ldots, \beta_p, \sigma^2)$

parameterrum $\Theta = R^p \times R_+$, $\dim \Theta = p+1$

Loglikelihoodfunktionen

$$l(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\|y - X\beta\|^2$$

$$= \text{——} \;\text{"}\; \text{——} - \frac{1}{2\sigma^2}(y^Ty - 2y^TX\beta + \beta^TX^TX\beta)$$

$$= \text{——} \;\text{"}\; \text{——} - \frac{1}{2\sigma^2}Q(\beta)$$

Bemærk, at (y^TX, y^Ty) er minimal sufficient for θ, jf. sætn. s. 40.

Antal komponenter : $p+1$, dvs. der foreligger en regulær eksponentiel familie af fordelinger.

Maksimering af $l(\theta)$ mht. β er ækvivalent med minimering af $Q(\beta)$, dvs. ML-estimatet for β bliver identisk med MK-estimatet $\hat{\beta} = (X^TX)^{-1}X^Ty$.

Profilloglikelihoodfunktionen

$$l^*(\sigma^2; \hat\beta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\|y - X\hat\beta\|^2$$

differentieres mht. σ^2 og sættes lig 0. Vi får $\hat{\sigma^2} = \frac{1}{n}\|y - X\hat\beta\|^2$. Ved yderligere differentiation kontrollerer vi, at $\hat{\sigma^2}$ svarer til maksimum.

Normalt benytter vi s^2 som estimat for σ^2 i stedet for $\hat{\sigma^2}$:

$$s^2 = \frac{n}{n-p}\hat{\sigma^2} = \frac{1}{n-p}\|y - X\hat\beta\|^2$$

$E[S^2] = \sigma^2$, dvs. s^2 er et centralt estimat for σ^2.

Kvadratsummer

$$Y^T Y \sim \sigma^2\chi^2(n; \delta), \quad \delta = \frac{1}{\sigma^2}\beta^T X^T X\beta$$

$$Y^T P Y \sim \sigma^2\chi^2(p; \delta), \quad \delta = \frac{1}{\sigma^2}\beta^T X^T X\beta$$

$$Y^T P_0 Y \sim \sigma^2\chi^2(p-q; \delta), \quad \delta = \frac{1}{\sigma^2}\beta^T X^T P_0 X\beta$$

$$Y^T P_H Y \sim \sigma^2\chi^2(q; \delta), \quad \delta = \frac{1}{\sigma^2}\beta^T X^T P_H X\beta$$

$$Y^T(I-P)Y \sim \sigma^2\chi^2(n-p)$$

Bemærk, at $Y^T P_0 Y$, $Y^T P_H Y$ og $Y^T(I-P)Y$ er indbyrdes uafhængige, jf. s. 5

Kvotienttest

Test af $\mu = 0$, ækvivalent med $\beta = 0$

$H_0 : \beta = 0$

$H_1 : \beta \neq 0$

Under H_0 skal kun σ^2 estimeres:

$$\hat{\sigma}_0^2 = \frac{1}{n} \| y \|^2$$

Kvotientteststørrelsen

$$\lambda(y) = \frac{(2\pi e \hat{\sigma}_0^2)^{-\frac{n}{2}}}{(2\pi e \hat{\sigma}^2)^{-\frac{n}{2}}} = \left(\frac{\frac{1}{n} \| y - \hat{\mu} \|^2}{\frac{1}{n} \| y \|^2} \right)^{\frac{n}{2}}$$

$$= \left(\frac{\| y - \hat{\mu} \|^2}{\| y - \hat{\mu} + \hat{\mu} \|^2} \right)^{\frac{n}{2}} = \left(\frac{\| y - \hat{\mu} \|^2}{\| y - \hat{\mu} \|^2 + \| \hat{\mu} \|^2} \right)^{\frac{n}{2}}$$

$$= \left(\frac{1}{1 + \frac{\| \hat{\mu} \|^2}{\| y - \hat{\mu} \|^2}} \right)^{\frac{n}{2}}$$

$$F = \frac{\frac{1}{r} \| \hat{\mu} \|^2}{\frac{1}{n-r} \| Y - \hat{\mu} \|^2} \sim F(r, n-r)$$

under H_0

*

* Bemærk, at $\| \hat{\mu} \|^2 = Y^T P Y \sim \sigma^2 \chi^2(r)$ under H_0, og at $\| Y - \hat{\mu} \|^2 = Y^T (I - P) Y \sim \sigma^2 \chi^2(n-r)$

$F(Y)$ er en monoton transformation af $\lambda(Y)$

store værdier af F_{obs} er kritiske for H_0.

□

Test af en lineær hypotese

$$H_0: \quad H\beta = 0$$

$$H_1: \quad H\beta \neq 0$$

Kvotientteststørrelsen:

$$\lambda(y) = \frac{L(\hat{\theta}_0)}{L(\hat{\theta})} = \frac{(2\pi e \hat{\sigma}_0^2)^{-\frac{n}{2}}}{(2\pi e \hat{\sigma}^2)^{-\frac{n}{2}}}$$

$$= \left(\frac{\frac{1}{n}\|y-\hat{\mu}\|^2}{\frac{1}{n}\|y-\hat{\mu}_0\|^2} \right)^{\frac{n}{2}} = \left(\frac{\|y-\hat{\mu}\|^2}{\|y-\hat{\mu}+\hat{\mu}-\hat{\mu}_0\|^2} \right)^{\frac{n}{2}}$$

$$= \left(\frac{\|y-\hat{\mu}\|^2}{\|y-\hat{\mu}\|^2 + \|\hat{\mu}-\hat{\mu}_0\|^2} \right)^{\frac{n}{2}}$$

$$= \left(\frac{1}{1 + \frac{\|\hat{\mu}-\hat{\mu}_0\|^2}{\|y-\hat{\mu}\|^2}} \right)^{\frac{n}{2}}$$

$$F = \frac{\frac{1}{q}\|\hat{\mu}-\hat{\mu}_0\|^2}{\frac{1}{n-p}\|Y-\hat{\mu}\|^2} \sim F(q, n-p)$$

under H_0

$$f_{obs} = \frac{\frac{1}{q} y^T (P-P_0) y}{\frac{1}{n-p}\|y-\hat{\mu}\|^2} = \frac{(H\hat{\beta})^T K H\hat{\beta}}{q s^2}$$

Bemærk, at

$$Y^T(P-P_0)Y = Y^T P_H Y \sim \sigma^2 \chi^2(q)$$

under H_0

En variant af den lineære hypotese:

$H_0: \ H\beta = h$

$H_1: \ H\beta \neq h$

Analoge regninger giver

$$\hat{\beta}_0 = \hat{\beta} - (X^T X)^{-1} H^T K (H\hat{\beta} - h)$$

$$f_{obs} = \frac{(H\hat{\beta} - h)^T K (H\hat{\beta} - h)}{q s^2} \qquad *$$

eks $H_0: \ \beta_r = h \in \mathbb{R}$ (ofte $h = 0$)

$H_1: \ \beta_r \neq h$

$H = [0 \ldots 0 \; 1 \; 0 \ldots 0]$ dvs. $q = 1$

↑ r'te element

$$\|\hat{\mu} - \hat{\mu}_0\|^2 = (\beta_r - h) \frac{1}{v_{rr}} (\beta_r - h) = \frac{(\beta_r - h)^2}{v_{rr}},$$

$$v_{rr} = \left((X^T X)^{-1}\right)_{rr}$$

$$f_{obs} = \frac{(\beta_r - h)^2}{s^2 v_{rr}}$$

ækvivalent hermed er

$$t_{obs} = \frac{\beta_r - h}{s\sqrt{v_{rr}}}, \quad T \sim t(n-p) \text{ under } H_0.$$

Konfidensinterval for β_r:

$$\beta_r = \hat{\beta}_r \pm t_{1-\frac{\alpha}{2}}(n-p)\, s\sqrt{v_{rr}}$$

$\square$

Konfidensområde for β

Når vi har konstrueret et test på signifikansniveau α, kan vi ud fra dette få et konfidensområde for parameteren som mængden af alle de parameterværdier, for hvilke nulhypotesen vil blive accepteret.

Her indsætter vi $H\beta$ for h i $*$ s. 10 og får umiddelbart konfidensområdet

$$\left\{ \beta \;\middle|\; \frac{(H\hat{\beta} - H\beta)^T \cdot K \cdot (H\hat{\beta} - H\beta)}{q s^2} < f_{1-\alpha} \right\}$$

$$= \left\{ \beta \;\middle|\; \frac{(\hat{\beta} - \beta)^T H^T K H (\hat{\beta} - \beta)}{q s^2} < f_{1-\alpha} \right\}$$

Konfidensgrad $1-\alpha$

$f_{1-\alpha}$ er $(1-\alpha)$-fraktilen i $F(q, n-p)$-fordelingen

Residualkvadratsummer

$$Q(\hat{\beta}) = \|y - X\hat{\beta}\|^2 = \|y - \hat{\mu}\|^2 = y^T (I - P) y$$

$$Q(\hat{\beta}_0) = \|y - X\hat{\beta}_0\|^2 = \|y - \hat{\mu}_0\|^2 = y^T (I - P_0) y$$

Devians er et andet ord for residualkvadratsum.

Bemærk, at

$$(\lambda(y))^{-\frac{2}{n}} - 1 = \frac{\|\hat{\mu} - \hat{\mu}_0\|^2}{\|y - \hat{\mu}\|^2} = \frac{\|y - \hat{\mu}_0 - (y - \hat{\mu})\|^2}{\|y - \hat{\mu}\|^2}$$

$$= \frac{\|y - \hat{\mu}_0\|^2 - \|y - \hat{\mu}\|^2}{\|y - \hat{\mu}\|^2}$$

$$= \frac{Q(\hat{\beta}_0) - Q(\hat{\beta})}{Q(\hat{\beta})},$$

altså den relative forskel i devians.