

Reparametrisierung des Modells

Gegeben sei das Modell (2.1) mit (2.5) unter der linearen Restriktion $A\beta = c$ mit $A \in R^{a \times p'}$, $\text{rg}(A) = a$, $c \in R^a$.

Wir betrachten die lineare Restriktion als Gleichungssystem. Die Lösungsmenge des Gleichungssystems ist wie folgt darstellbar:

$$\beta = B\gamma + d \quad (3.44)$$

mit Matrizen $B \in R^{p' \times (p' - a)}$, $d \in R^{p'}$ und $\gamma \in R^{p' - a}$. Es folgt:

$$Y = X\beta + \varepsilon = XB\gamma + Xd + \varepsilon$$

$$V = Z\gamma + \varepsilon$$

$$V := Y - Xd$$

$$Z := XB$$

Das Modell ist das **reparametrisierte** Modell mit:

Zielgröße: $V = Y - Xd \in R^n$

Design-Matrix: Z , $Z \in R^{n \times (p' - a)}$; $\text{rg}Z = p' - a$

Parameter: $\gamma \in R^{p' - a}$

Störterm: ε stimmt mit dem aus dem Grundmodell überein!

Zusammenhang Reparametrisierung und Modell unter linearer Restriktion

Es gilt:

$$\hat{\hat{\beta}} = B\hat{\gamma} + d \quad (3.45)$$

$$SSH = \hat{\hat{\varepsilon}}'\hat{\hat{\varepsilon}} - \hat{\varepsilon}'\hat{\varepsilon} \quad (3.46)$$

mit $\hat{\hat{\beta}}$: KQ-Schätzer unter Restriktion $A\beta = c$
 $\hat{\gamma}$: KQ-Schätzer aus Modell (3.45)
 $\hat{\hat{\varepsilon}}$: Residuenvektor aus Modell (3.45) =
Residuenvektor aus KQ-Schätzung unter Restriktion.

Die maximierte Likelihood des Modells ist:

$$(2\pi)^{-n/2} \cdot \hat{\sigma}^{-n} \cdot \exp \left[-\sum_{i=1}^n \frac{\hat{\varepsilon}_i^2}{2\hat{\sigma}^2} \right]$$

Es folgt (siehe Nachweis ML = KQ aus Kapitel 1)

$$\hat{\sigma}^2 = \frac{\sum \hat{\varepsilon}_i^2}{n} \quad (3.47)$$

$$MaxL = C * \hat{\sigma}^n = C * (SSE/n)^{(n/2)} \quad (3.48)$$

Likelihood Quotienten-Test

Grundidee des Likelihood-Quotienten-Tests:

Vergleiche (Bilde den Quotienten) maximierte Likelihood des Modells unter H_0 mit maximierter Likelihood ohne H_0

Wir betrachten also den ML - Schätzer mit und ohne die Restriktion $A\beta = c$:

$\hat{\hat{\varepsilon}}$: Residuen unter dem Modell mit H_0

$\hat{\varepsilon}$: Residuen unter dem Modell ohne Einschränkung

Die LQ- Teststatistik lautet dann:

$$\tau_{LQ} = \left(\frac{\hat{\hat{\sigma}}}{\hat{\sigma}} \right)^{-n} = \left(\frac{\hat{\hat{\varepsilon}}' \hat{\hat{\varepsilon}}}{\hat{\varepsilon}' \hat{\varepsilon}} \right)^{-n/2} \quad (3.49)$$

Wald-Test ist Likelihood-Quotienten-Test

Sei das Modell (2.1) mit (2.5) und $A \in R^{a \times p'}$, $rg(A) = a$, $c \in R^a$ gegeben.

$$H_0 : A\beta = c \text{ gegen } H_1 : A\beta \neq c$$

Dann ist der Wald-Test zu dem Likelihood-Quotienten-Test äquivalent, d. h. :

Die Testgröße des LQ-Tests ist eine streng monotone Funktion der Testgröße des Wald-Tests:

$$\tau_{LQ} = g \left[\frac{MSH}{MSE} \right] \quad (3.50)$$

g streng monoton

Damit ist τ_{LQ} eine monotone Funktion der Testgröße des Wald-Tests:

$$\frac{MSH}{MSE} = \frac{(\hat{\hat{\varepsilon}}' \hat{\hat{\varepsilon}} - \hat{\varepsilon}' \hat{\varepsilon})/a}{\hat{\varepsilon}' \hat{\varepsilon}/(n - p')} =: \tau_W$$
$$\Rightarrow \left[\tau_W \cdot \frac{a}{(n - p)'} + 1 \right]^{-n/2} = \tau_{LQ}$$

Sequentielle und Partielle Quadratsummen

Gegeben sei das Modell (2.1). Wir betrachten nun Teilmodelle, die durch Nullrestriktionen von Komponenten des Vektors β entstehen und deren Residuenquadratsummen.

$$R(\beta_{i1}, \dots, \beta_{ik} | \beta_{j1}, \dots, \beta_{jl}) = SSH = SSE(M1) - SSE(M2) \quad (3.51)$$

M2: Modell, das die Parameter $\beta_{i1}, \dots, \beta_{ik}, \beta_{j1}, \dots, \beta_{jl}$ enthält.

M1: Modell, das die Parameter $\beta_{j1}, \dots, \beta_{jl}$ enthält.

SSH: Hypothesenquadratsumme zur Hypothese $(\beta_{i1} = \dots = \beta_{ik} = 0)$ im Modell M2.

Partielle Quadratsummen

Die zu der Hypothese $\beta_i = 0$ gehörigen Quadratsummen bzgl. des Gesamtmodells heißen **partielle Quadratsummen**:

$$R(\beta_i | \beta_0, \dots, \beta_{i-1}, \beta_{i+1}, \dots, \beta_p) = SSE(M_{-i}) - SSE \quad (3.52)$$

M_{-k} : Modell mit $\beta_k = 0$.

Beachte: Die partiellen Quadratsummen führen im Allgemeinen zu keiner Zerlegung von SST, d.h.

$$SST \neq \sum_{k=1}^p R(\beta_k | \beta_0, \dots, \beta_{k-1}, \beta_{k+1}, \dots, \beta_p) + SSE \quad (3.53)$$

Sequentielle Quadratsummen

Wir betrachten die Folge von Modellen:

$$M_0 : \quad Y = \beta_0 + \varepsilon \quad (3.54)$$

$$M_1 : \quad Y = \beta_0 + \beta_1 x_1 + \varepsilon \quad (3.55)$$

...

$$M_p : \quad Y = X\beta + \varepsilon \quad (3.56)$$

$$R(\beta_k | \beta_0, \dots, \beta_{k-1}) = SSE(M_{k-1}) - SSE(M_k) \quad (3.57)$$

heißen **sequentielle Quadratsummen** und es gilt:

$$SST = \sum_{k=1}^p R(\beta_k | \beta_0, \dots, \beta_{k-1}) + SSE \quad (3.58)$$

Bereinigung

Wir betrachten folgende Zerlegung des Modells (2.1):

$$Y = X\beta + \varepsilon \quad (3.59)$$

$$Y = (X_1 X_2)(\beta_1' \beta_2')' + \varepsilon = X_1 \beta_1 + X_2 \beta_2 + \varepsilon \quad (3.60)$$

$$\hat{\beta} = (X'X)^{-1}X'Y = (\hat{\beta}_1', \hat{\beta}_2')' \quad (3.61)$$

$$Q_2 := I - X_2(X_2'X_2)^{-1}X_2' \quad (3.62)$$

$$Y^* := Q_2 Y \quad (3.63)$$

$$X_1^* := Q_2 X_1 \quad (3.64)$$

$$\Rightarrow \hat{\beta}_1 = (X_1^{*'}X_1^*)^{-1}X_1^{*'}Y^* \quad (3.65)$$

Y^*, X_1^* : von X_2 bereinigte Variablen

Beweis von (3.67) I

$$Y = X_1\beta_1 + X_2\beta_2 + \varepsilon \quad X = (X_1 \ X_2) \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

KQ:

$$Y = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + \hat{\varepsilon}$$

$$Q_2 = I - P_2 = I - X_2(X_2'X)^{-1}X_2'$$

$$\text{Es gilt: } Q_2Q = (I - P_2)(I - P) = I - P_2 - P + \underbrace{P_2P}_{P_2} = I - P = Q$$

(Residuen von Regression auf $(Y_1 \ Y_2)$ und Residuen bzgl. Y_2)

$$\text{Analog: } QQ_2 = Q$$

Beweis von (3.67) II

$$Q_2 Y = Q_2 X_1 \hat{\beta}_1 + \underbrace{Q_2 X_2}_{0} \hat{\beta}_2 + \underbrace{Q_2 \hat{\varepsilon}}_{\underbrace{Q_2 Q Y}_Q}$$

$$\Rightarrow Q_2 Y = Q_2 X_1 \hat{\beta}_1 + \hat{\varepsilon}$$

$$Y^* = X_1^* \hat{\beta}_1 + \hat{\varepsilon} \quad \text{Bereinigte Variablen}$$

$$X_1^{*'} \hat{\varepsilon} = (Q_2 X_1)' \hat{\varepsilon} = X_1' Q_2 Q \hat{\varepsilon} = X_1' \hat{\varepsilon} = 0$$

$\Rightarrow \hat{\beta}_1$ ist KQ-Lösung der Regression von Y^* auf X^* .

Beispiele zur Bereinigung

Mittelwerts-Bereinigung beim einfachen linearen Regressionsmodell

$$\begin{aligned}X_2 &= (1 \dots 1)' \\X_1 &= (x_1 \dots x_n)' \\Y^* &= Y - \bar{y} \\X_1^* &= x - \bar{x} \\(X'^* X^*)^{-1} X'^* Y^* &= S_x^{-2} S_{xy}\end{aligned}$$

Geschlechtseffekt, Trendbereinigung

Beispiel: Zusammenhang zwischen Fehleranzahl bei Test zu starken Verben

Zielgröße : Anzahl der Fehler bei Test

Einflussgrößen : Geschlecht Alter Leseverhalten, Fernsehverhalten, etc.

Regression auf binäre Variable Geschlecht entspricht
Mittelwertsschätzung

Bereinigung nach Geschlecht: Abziehen des jeweiligen
Gruppenmittelwertes

Prognose

In Modell (2.1) betrachten wir eine weitere Beobachtung mit bekanntem Vektor der x-Werte x_{n+1} und zugehörigem unbekanntem Y_{n+1} . Der Prognosewert von Y_{n+1} ist gegeben durch

$$\hat{Y}_{n+1} = x'_{n+1} \hat{\beta} \quad (3.66)$$

Für den Erwartungswert und die Varianz des Prognosefehlers gilt:

$$E \left[\hat{Y}_{n+1} - Y_{n+1} \right] = 0 \quad (3.67)$$

$$V \left[\hat{Y}_{n+1} - Y_{n+1} \right] = \sigma^2 (1 + x'_{n+1} (X'X)^{-1} x_{n+1}) \quad (3.68)$$

Der Nachweis verläuft analog zum einfachen linearen Modell unter Berücksichtigung von

$$V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

Prognoseintervalle und Konfidenzintervalle

Das Prognoseintervall für eine zukünftige Beobachtung y_{n+1} bei gegebenem x_{n+1} erhält man mit Hilfe der Normalverteilung von $\hat{\beta}$ und standardisieren:

$$[\hat{Y}_{n+1} - \hat{\sigma}_{\hat{Y}_{n+1}} t_{1-\alpha/2}(n-p'); \hat{Y}_{n+1} + \hat{\sigma}_{\hat{Y}_{n+1}} t_{1-\alpha/2}(n-p')] \quad (3.69)$$

$$\text{mit } \hat{\sigma}_{\hat{Y}_{n+1}}^2 = \hat{\sigma}^2 (1 + x'_{n+1}(X'X)^{-1}x_{n+1})$$

Das Konfidenzintervall für den Erwartungswert μ_{n+1} einer zukünftige Beobachtung y_{n+1} bei gegebenem x_{n+1} ist:

$$[\hat{Y}_{n+1} - \hat{\sigma}_{\hat{\mu}_{n+1}} t_{1-\alpha/2}(n-p'); \hat{Y}_{n+1} + \hat{\sigma}_{\hat{\mu}_{n+1}} t_{1-\alpha/2}(n-p')] \quad (3.70)$$

$$\text{mit } \hat{\sigma}_{\hat{\mu}_{n+1}}^2 = \hat{\sigma}^2 (x'_{n+1}(X'X)^{-1}x_{n+1})$$

Bemerkungen

- Prognoseintervalle enthalten die Zufallsschwankung durch den Störterm ε_{n+1} **und** die Unsicherheit bezüglich der Modellparameter
- Die entsprechenden Konfidenzintervalle enthalten nur die Unsicherheit bezüglich der Modellparameter. Sie können gut zur Darstellung der Modellschätzung verwendet werden.
- Um den geschätzten Effekt einer Einflussgröße x_1 darzustellen, werden die anderen Einflussgrößen auf einen festen Wert (z.B. Modus, Median oder Mittelwert) gesetzt. Wann wird der Zusammenhang von $x_{1,n+1}$ und $\hat{Y}_{n+1}$ und die entsprechenden Konfidenzintervalle geplottet. (R- Paket effects)

Das Gauss-Markov-Theorem

Sei das Modell

$$\begin{aligned} Y &= X\beta + \varepsilon, & \text{rg } X &= p' \\ E(\varepsilon) &= 0 \\ V(\varepsilon) &= \sigma^2 I \end{aligned}$$

gegeben.

Dann ist der KQ-Schätzer $\hat{\beta}$ unter den erwartungsternen linearen Schätzern derjenige mit der kleinsten Varianz: $\hat{\beta}$ ist BLUE-Schätzer (**b**est **l**inear **u**nbiased **e**stimator).

Ist $\tilde{\beta}$ ein weiterer Schätzer von β mit $E(\tilde{\beta}) = \beta$ und $\tilde{\beta} = CY$, so gilt:

$$V(\tilde{\beta}) \geq V(\hat{\beta})$$

$$V(\tilde{\beta}) \geq V(\hat{\beta}) \Leftrightarrow V(\tilde{\beta}) = V(\hat{\beta}) + M \text{ mit } M \text{ positiv semidefinit.}$$

Konsistenz des KQ-Schätzers I

Sei das Modell (2.1) - (2.4) gegeben. Wir betrachten nun das Modell mit steigendem Stichprobenumfang n . Da die Einflussgrößen fest sind, gehen wir von einer gegebenen Folge x_n der Einflussgrößen aus.

Sei zu jedem $n > p'$

X_n : Designmatrix, die aus den ersten n Beobachtungen besteht
 $\hat{\beta}^{(n)}$: KQ- Schätzer aus den ersten n Beobachtungen

Vor.: X_n hat vollen Rang für alle $n \geq p$

$$\lim_{n \rightarrow \infty} (X_n' X_n)^{-1} = 0.$$

Konsistenz des KQ-Schätzers II

Dann folgt die schwache Konsistenz des KQ-Schätzers (Konvergenz in Wahrscheinlichkeit):

$$\hat{\beta}^{(n)} \xrightarrow{P} \beta. \quad (3.71)$$

Sind die Störgrößen zusätzlich identisch verteilt, so folgt die starke Konsistenz (fast sichere Konvergenz)

$$\hat{\beta}^{(n)} \xrightarrow{\text{f.s.}} \beta. \quad (3.72)$$

Asymptotische Normalität des KQ-Schätzers

Sei das Modell (2.1) - (2.4) gegeben.

Sei zu jedem $n > p'$

X_n : Designmatrix, die aus den ersten n Beobachtungen besteht.
 $\hat{\beta}^{(n)}$: KQ- Schätzer aus den ersten n Beobachtungen

Vor.: X_n hat vollen Rang für alle $n \geq p$

$$\lim_{n \rightarrow \infty} \max_{1 \leq i \leq n} x_i' (X_n' X_n)^{-1} x_i = 0.$$

Dann folgt die asymptotische Normalität von β

$$\sigma^{-1} (X_n' X_n)^{1/2} (\hat{\beta}^{(n)} - \beta) \xrightarrow{d} N(0, I). \quad (3.73)$$