14 Lineare statistische Modelle

14.1 Definition

Es seien $X = (X_1, \dots, X_n)^T$ ein (beobachtbarer) Zufallsvektor,
$C = (c_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,s}}$ eine bekannte $n \times s$-Matrix mit $\mathrm{Rang}(C) = s$
(insbesondere $n \geq s$), $\vartheta = (\vartheta_1, \dots, \vartheta_s)^T$ unbekannter Parametervektor,
$\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$ ein (nicht beobachtbarer) Zufallsvektor mit

$$E(\varepsilon) = 0,\ \mathrm{Var}(\varepsilon) = E(\varepsilon \cdot \varepsilon^T) = \sigma^2 \cdot I_n$$

σ^2 unbekannt.
Ein **lineares Modell (LM)** wird beschrieben durch die Gleichung

$$X = C\vartheta + \varepsilon \qquad (1)$$

also

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1s} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{ns} \end{pmatrix} \begin{pmatrix} \vartheta_1 \\ \vdots \\ \vartheta_s \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

C heißt „Designmatrix“.
(1) heißt klassisch, falls $\varepsilon \sim \mathcal{N}_n(0, \sigma^2 I_n)$.

Bemerkungen:

a) Im klassischen LM gilt: $X \sim \mathcal{N}_n(C\vartheta, \sigma^2 I_n)$.
Die Beobachtungen $X_1, \dots, X_n$ sind also unabhängig, aber nicht identisch verteilt.

b) $\mathrm{Rang}(C) = s \;\Leftrightarrow\; C^T C$ nicht singulär
Denn[39]:

$$\begin{aligned} C^T C \text{ singulär} &\Leftrightarrow \exists u \in \mathbb{R}^s, u \neq 0 :\ C^T C u = 0 \\ &\Leftrightarrow \exists u \in \mathbb{R}^s, u \neq 0 :\ u^T C^T C u = (Cu)^T C u = 0 \\ &\Leftrightarrow \exists u \in \mathbb{R}^s, u \neq 0 :\ Cu = 0 \\ &\Leftrightarrow \mathrm{Rang}(C) < s \end{aligned}$$

[39] In der Hinrichtung multipliziere $C^T C u = 0$ mit u^T, in der Rückrichtung multipliziere $Cu = 0$ mit C^T.

14.2 Beispiele

a) $X_i = \vartheta + \varepsilon_i,\ i = 1, \dots, n$

$$(s = 1,\ C = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix})$$

(wiederholte Messung)

b) $X_i = a + bt_i + \varepsilon_i,\ i = 1, \dots, n$

$$(s = 2,\ a = \vartheta_1,\ b = \vartheta_2,\ C = \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix})$$

(einfache lineare Regression)

c) $X_i = a + bt_i + ct_i^2 + \varepsilon_i,\ i = 1, \dots, n$

$$(s = 3,\ \vartheta = (a, b, c)^T,\ C = \begin{pmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 \end{pmatrix})$$

(einfache quadratische Regression)

d) $X_i = \sum_{j=1}^{s} \vartheta_j \cdot f_j(t_i) + \varepsilon_i,\ i = 1, \dots, n$
$f_1, \dots, f_s$ beliebige gegebene Funktionen! (allgemeine (lineare) Regression)
z.B. $f_j(t) = \sin(\omega_j \cdot t)$ (trigonometrische Regression)

e) $X_i = a + bu_i + cv_i + \ldots + gz_i + \varepsilon_i,\ i = 1, \dots, n$

$$\vartheta = \begin{pmatrix} a \\ \vdots \\ g \end{pmatrix},\ C = \begin{pmatrix} 1 & u_1 & v_1 & \cdots & z_1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & u_n & v_n & \cdots & z_n \end{pmatrix})$$

(multiple lineare Regression)

f) $X_{1,i} = \vartheta_1 + \varepsilon_{1,i},\ i = 1, \dots, n_1$
$X_{2,i} = \vartheta_2 + \varepsilon_{2,i},\ i = 1, \dots, n_2$

$$\begin{pmatrix} X_{1,1} \\ \vdots \\ X_{1,n_1} \\ X_{2,1} \\ \vdots \\ X_{2,n_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \vartheta_1 \\ \vartheta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,1} \\ \vdots \\ \varepsilon_{1,n_1} \\ \varepsilon_{2,1} \\ \vdots \\ \varepsilon_{2,n_2} \end{pmatrix}$$

(2-Stichproben-Modell)

g) $X_{i,j} = \vartheta_i + \varepsilon_{i,j}$, $i = 1, \dots, k$, $j = 1, \dots, n_i$
(Modell der einfachen Varianzanalyse, 1-faktorielle ANOVA)
z.B. Effekt $X_{i,j}$ bei k unterschiedlichen Behandlungen

14.3 Schätzung von ϑ

Sei $R(C) := \{C\vartheta : \vartheta \in \mathbb{R}^s\}$ s-dimensionaler Unterraum des $\mathbb{R}^n$.
14.1(1) besagt $EX \in R(C)$.
Forderung: $\|X - C\vartheta\|^2 = \min_{\vartheta}!$ (kleinste-Quadrate-Methode; vgl. 4.6)

Lösung:

$$\hat{\vartheta} = \hat{\vartheta}(X) = (C^T C)^{-1} \cdot C^T X$$

Beweis:
Wegen $\mu(\vartheta) = C\vartheta$ folgt $M(\vartheta) = \left(\frac{\partial \mu_i}{\partial \vartheta_j}\right)_{i,j} = C$ in 4.6 und somit die Normalengleichung $C^T C\vartheta = C^T X$.
Da $C^T C$ nach Bemerkung 14.1(b) invertierbar ist, ist

$$\hat{\vartheta} = \hat{\vartheta}(X) = (C^T C)^{-1} \cdot C^T X$$

die (einzige) Lösung.

Bemerkung:
Es gilt[40]:

$$E_{\vartheta,\sigma^2}(\hat{\vartheta}) = (C^T C)^{-1} C^T \underbrace{E_{\vartheta,\sigma^2}(X)}_{=C\vartheta} = \vartheta$$

d.h. $\hat{\vartheta}$ ist erwartungstreu für ϑ.

$$\mathrm{Var}_{\vartheta,\sigma^2}(\hat{\vartheta}) = (C^T C)^{-1} C^T \underbrace{\mathrm{Var}_{\vartheta,\sigma^2}(X)}_{=\mathrm{Var}_{\vartheta,\sigma^2}(\varepsilon)=\sigma^2 \cdot I_n} \cdot C(C^T C)^{-1} = \sigma^2 (C^T C)^{-1}$$

Beispiele:

a) In 14.2(b) (einfache lineare Regression) ist (vgl. 4.7)

$$\hat{\vartheta}_1 = \bar{X} - \hat{\vartheta}_2 \bar{t},\ \hat{\vartheta}_2 = \frac{\sum_{i=1}^n t_i X_i - n \cdot \bar{t} \cdot \bar{X}}{\sum_{i=1}^n (t_i - \bar{t})^2}$$

[40] Beachte: $(A^T)^{-1} = (A^{-1})^T$

b) In 14.2(g) (ANOVA) ist

$$C = \begin{pmatrix} 1 & & & \\ \vdots & & & \\ 1 & & & \\ & 1 & & \\ & \vdots & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ & & & \vdots \\ & & & 1 \end{pmatrix}, \; C^T C = \begin{pmatrix} n_1 & & & 0 \\ & n_2 & & \\ & & \ddots & \\ 0 & & & n_k \end{pmatrix}$$

und somit

$$\hat{\vartheta} = \begin{pmatrix} \hat{\vartheta}_1 \\ \vdots \\ \hat{\vartheta}_k \end{pmatrix} = \begin{pmatrix} \frac{1}{n_2} \sum_{j=1}^{n_1} X_{1,j} \\ \vdots \\ \frac{1}{n_k} \sum_{j=1}^{n_k} X_{k,j} \end{pmatrix} =: \begin{pmatrix} \bar{X}_{1+} \\ \vdots \\ \bar{X}_{k+} \end{pmatrix}$$

(+ bedeutet, dass hier summiert wird)

14.4 Satz (Gauß-Markov-Theorem)

Es sei $a \in \mathbb{R}^s$. Dann ist $T := a^T \hat{\vartheta}$ **bester linearer erwartungstreuer** Schätzer für $a^T \vartheta$. (BLUE)

<u>Beweis:</u>
Sei $S = S(X)$ linearer Schätzer für $a^T \vartheta$.

$$\Rightarrow \exists b \in \mathbb{R}^n : S = b^T X$$

S erwartungstreu für $a^T \vartheta \Rightarrow$

$$E_{\vartheta,\sigma^2} S = b^T E_{\vartheta,\sigma^2} X = b^T C \vartheta \stackrel{!}{=} a^T \vartheta \; \forall \vartheta$$

$\Rightarrow b^T C = a^T \;\; (*)$

$$\mathrm{Var}_{\vartheta,\sigma^2}(S)) b^T \underbrace{\mathrm{Var}_{\vartheta,\sigma^2} X}_{\sigma^2 I_n} \cdot b = \sigma^2 b^T b$$

$$\mathrm{Var}_{\vartheta,\sigma^2}(T) = a^T \, \mathrm{Var}_{\vartheta,\sigma^2}(\hat{\vartheta}) a = \sigma^2 a^T (C^T C)^{-1} a \stackrel{(*)}{=} \sigma^2 b^T C (C^T C)^{-1} C^T b$$

$$\Rightarrow \mathrm{Var}_{\vartheta,\sigma^2}(S) - \mathrm{Var}_{\vartheta,\sigma^2}(T) = \sigma^2 b^T \underbrace{(I_n - \underbrace{C(C^TC)^{-1}C^T}_{=:P})}_{=:Q} b$$

Wegen $P = P^T = P^2$ folgt $Q = Q^T = Q^2$ (vgl. Aufgabe 44) folgt

$$b^TQb = b^TQ^2b = b^TQ^TQb = \|Qb\|^2 \geq 0$$

$\Rightarrow$ Behauptung

Beispiele:

a) 1-faktorielle ANOVA (14.2(g), Beispiel 14.3(b))
$a^T = (0, \ldots, 0, \underbrace{1}_{a_i}, 0, \ldots, 0, \underbrace{-1}_{a_j}, 0, \ldots, 0),\ a^T\vartheta = \vartheta_i - \vartheta_j$
Differenz der Erwartungswerte der i-ten und j-ten Gruppe.
$T = a^T\hat{\vartheta} = \bar{X}_{i+} - \bar{X}_{j+}$ ist BLUE für $a^T\vartheta$.

b) einfache lineare Regression
$a = \begin{pmatrix} 1 \\ t^* \end{pmatrix},\ a^T\vartheta = \vartheta_1 + \vartheta_2 - t^*$
$T = a^T\hat{\vartheta} = \hat{\vartheta}_1 + \hat{\vartheta}_2 t^*$ ist BLUE.
Hier:

$$C = \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix}$$

$$C^TC = \begin{pmatrix} n & n\bar{t} \\ n\bar{t} & \sum t_i^2 \end{pmatrix},\ (C^TC)^{-1} = \frac{1}{\sum(t_i - \bar{t})^2} \begin{pmatrix} \frac{1}{n}\sum t_i^2 & -\bar{t} \\ -\bar{t} & 1 \end{pmatrix}$$

$$\Rightarrow \mathrm{Var}_{\vartheta,\sigma^2}(\hat{\vartheta}_1) = \sigma^2 \frac{\frac{1}{n}\sum t_i^2}{\sum(t_i - \bar{t})^2}$$

$$\mathrm{Var}_{\vartheta,\sigma^2}(\hat{\vartheta}_2) = \sigma^2 \frac{1}{\sum(t_i - \bar{t})^2} \text{ (vgl. 4.7)}$$

$$\mathrm{Cov}_{\vartheta,\sigma^2}(\hat{\vartheta}_1, \hat{\vartheta}_2) = \frac{-\sigma^2\bar{t}}{\sum(t_i - \bar{t})^2} \text{ (}= 0,\text{ falls } \bar{t} = 0\text{)}$$

$$\begin{aligned} \mathrm{Var}_{\vartheta,\sigma^2}(T) &= \sigma^2 a^T(C^TC)^{-1}a \\ &= \frac{\sigma^2}{\sum(t_i - \bar{t})^2}(\frac{1}{n}\sum t_i^2 - 2t^*\bar{t} + (t^*)^2) \\ &= \sigma^2(\frac{1}{n} + \frac{(t^* - \bar{t})^2}{\sum(t_i - \bar{t})^2}) \end{aligned}$$

14.5 Schätzung von σ^2

$\hat{\sigma}^2 = \hat{\sigma}^2(X) = \frac{1}{n} \| \underbrace{X - C\hat{\vartheta}}_{=:\hat{\varepsilon}} \|^2 = \frac{1}{n} \|\hat{\varepsilon}\|^2 = \frac{\hat{\varepsilon}^T\hat{\varepsilon}}{n}$
($\hat{\varepsilon}$ Residuenvektor)

Bemerkung:
$\hat{\sigma}^2$ ist asymptotisch erwartungstreu, aber nicht erwartungstreu für σ^2, da nach Aufgabe 44

$$\hat{S}^2 = \frac{n}{n-s}\hat{\sigma}^2 = \frac{1}{n-s} \|\hat{\varepsilon}\|^2$$

erwartungstreu für σ^2 ist.

Ab jetzt stets klassisches lineares Modell ($\varepsilon \sim \mathcal{N}_n(0, \sigma^2 I_n)$)!

14.6 Satz

Im (klassischen) linearen Modell gilt:

a) $(\hat{\vartheta}, \hat{\sigma})$ ist ML-Schätzer für (ϑ, σ^2)

b) $\hat{\vartheta} \sim \mathcal{N}_s(\vartheta, \sigma^2(C^TC)^{-1})$

c) $\frac{n}{\sigma^2}\hat{\sigma}^2 \sim \chi^2_{n-s}$

d) $\hat{\vartheta}$ und $\hat{\sigma}^2$ sind stochastisch unabhängig

Beweis:

a) $X \sim \mathcal{N}_n(C\vartheta, \sigma^2 I_n)$

$$\begin{aligned} \Rightarrow f(x, \vartheta, \sigma^2) &= \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - (C\vartheta)_i)^2\} \\ &= \frac{1}{(\sigma^2 2\pi)^{\frac{n}{2}}} \exp\{-\frac{\|x - C\vartheta\|^2}{2\sigma^2}\} \\ &=: L_x(\vartheta, \sigma^2) \end{aligned}$$

Maximieren von L_x bezüglich ϑ bei festem σ^2 führt auf Minimierung von $\|x - C\vartheta\|^2$, Lösung ist $\hat{\vartheta}$.

$$\frac{\partial \log L_x(\hat{\vartheta}, \sigma^2)}{\partial \sigma^2} \stackrel{!}{=} 0$$

$$\hat{\sigma}^2 = \frac{1}{n} \left\| x - C\hat{\vartheta} \right\|^2$$

b) folgt aus Bemerkung 14.3 und Normalverteilungs-Annahme

c)

$$\begin{aligned}
\varepsilon^T \varepsilon &= (X - C\vartheta)^T (X - C\vartheta) \\
&= (X - C\hat{\vartheta} + C(\hat{\vartheta} - \vartheta))^T (X - C\hat{\vartheta} + C(\hat{\vartheta} - \vartheta)) \\
&= (\hat{\varepsilon} + C(\hat{\vartheta} - \vartheta))^T (\hat{\varepsilon} + C(\hat{\vartheta} - \vartheta))
\end{aligned}$$

$$\Rightarrow \underbrace{\frac{\varepsilon^T \varepsilon}{\sigma^2}}_{\sim \chi^2_n} = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{\sigma^2} + \underbrace{(\hat{\vartheta} - \vartheta)^T \frac{C^T C}{\sigma^2} (\hat{\vartheta} - \vartheta)}_{\sim \chi^2_s \ (1)} + 2 \underbrace{\hat{\varepsilon}^T C}_{(2)} \frac{(\hat{\vartheta} - \vartheta)}{\sigma^2}$$

(1) nach Hilfssatz 13.6 und (b)
(2) $= \varepsilon^T (I_n - P)^T C = \varepsilon^T (I_n - P) C = 0$

Zu zeigen: $\frac{\hat{\varepsilon}^T \hat{\varepsilon}}{\sigma^2} \sim \chi^2_{n-s}$
Die charakteristische Funktion von χ^2_k ist

$$\varphi_{\chi^2_k}(t) = \int_{-\infty}^{\infty} e^{itx} f_k(x) dx = (1 - 2it)^{-\frac{k}{2}}$$

Unabhängigkeit von $\hat{\vartheta}$ und $\hat{\varepsilon}$ nach (d)

$$\Rightarrow (1 - 2it)^{-\frac{n}{2}} = \varphi_{\frac{\hat{\varepsilon}^T \hat{\varepsilon}}{\sigma^2}}(t) \cdot (1 - 2it)^{-\frac{s}{2}}$$

$$\Rightarrow \varphi_{\frac{\hat{\varepsilon}^T \hat{\varepsilon}}{\sigma^2}}(t) = (1 - 2it)^{-\frac{n-s}{2}}$$

Eindeutigkeitssatz für charakteristische Funktionen

$$\Rightarrow \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{\sigma^2} = \frac{n}{\sigma^2} \hat{\sigma}^2 \sim \chi^2_{n-s}$$

d) $\hat{\vartheta} = (C^T C)^{-1} C^T X = (C^T C)^{-1} C^T (C\vartheta + \varepsilon) = \vartheta + (C^T C)^{-1} C^T \varepsilon$

$$\begin{aligned}
\hat{\varepsilon} &= X - C\hat{\vartheta} \\
&= (I_n - C(C^T C)^{-1} C^T) X \\
&= (I_n - P)(C\vartheta + \varepsilon) \\
&= \underbrace{(I_n - P)C}_{C - C = 0} \vartheta + (I_n - P)\varepsilon \\
&= (I_n - P)\varepsilon \\
(&= Q\varepsilon)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \underbrace{\mathrm{Cov}(\hat\vartheta,\hat\varepsilon)}_{s\times n \text{ Matrix}} &= \mathrm{Cov}(\vartheta+(C^TC)^{-1}C^T\varepsilon,(I_n-P)\varepsilon)\\
&= \mathrm{Cov}((C^TC)^{-1}C^T\varepsilon,(I_n-P)\varepsilon)\\
&= \underbrace{(C^TC)^{-1}C^T}_{s\times n}\cdot \underbrace{\mathrm{Cov}(\varepsilon,\varepsilon)}_{=\mathrm{Var}(\varepsilon)=\sigma^2 I_n}\cdot \underbrace{(I_n-P)^T}_{n\times n}\\
&= \sigma^2(C^TC)^{-1}(\underbrace{(I_n-P)C}_{=0})^T\\
&= 0
\end{aligned}$$

$\hat\varepsilon=(I_n-P)\varepsilon\sim\mathcal{N}_n(0,(I_n-P)\sigma^2(I_n-P)^T)=\mathcal{N}_n(0,\sigma^2(I_n-P))$
$\hat\varepsilon,\hat\vartheta$ normalverteilt und unkorreliert $\Rightarrow \hat\vartheta,\hat\varepsilon$ unabhängig
$\Rightarrow \hat\vartheta,\hat\sigma^2$ stochastisch unabhängig.

Bemerkung:
$\hat\varepsilon\sim\mathcal{N}_n(0,(I_n-P)\sigma^2)$, d.h. die $\hat\varepsilon_i$ haben nicht die gleiche Varianz.

14.7 Konfidenzbereiche für ϑ

a) elliptischer Konfidenzbereich für ϑ:

$$\hat\vartheta-\vartheta\sim\mathcal{N}(0,\sigma^2(C^TC)^{-1})$$

$$\Rightarrow (\hat\vartheta-\vartheta)^T\frac{C^TC}{\sigma^2}(\hat\vartheta-\vartheta)\sim\chi^2_s;\ \frac{n\hat\sigma^2}{\sigma^2}\sim\chi^2_{n-s}$$

Beide Größen sind stochastisch unabhängig.

$$\Rightarrow \frac{\frac{1}{s}(\hat\vartheta-\vartheta)^TC^TC(\hat\vartheta-\vartheta)}{\frac{n}{n-s}\hat\sigma^2}\sim F_{s,n-s}$$

$$\Rightarrow C_E:=\{y\in\mathbb{R}^s:\ \frac{\frac{1}{s}(\hat\vartheta-y)^TC^TC(\hat\vartheta-y)}{\hat s^2}\le F_{s,n-s,1-\alpha}\}$$

erfüllt $P_{\vartheta,\sigma^2}(C_E(X)\ni\vartheta)=1-\alpha\ \forall\vartheta,\sigma^2$, d.h. C_E ist ein (exakter) $(1-\alpha)$-Konfidenzbereich für ϑ.

b) Konfidenzintervall für ϑ_j:
Sei $(C^TC)^{-1}=:(b_{ij})_{s\times s}$. $\hat\vartheta_j\sim\mathcal{N}(\vartheta_j,b_{jj}\sigma^2)$

$$\overset{14.6(c),(d),\ 2.1}{\Rightarrow} \frac{\frac{\hat\vartheta_j-\vartheta_j}{\sigma\sqrt{b_{jj}}}}{\sqrt{\frac{n}{n-s}\frac{\hat\sigma^2}{\sigma^2}}}=\frac{\hat\vartheta_j-\vartheta_j}{\hat s\cdot\sqrt{b_{jj}}}\sim t_{n-s}(\sim\sqrt{F_{1,n-s}})$$

$$\Rightarrow P_{\vartheta,\sigma^2}(|\hat\vartheta_j-\vartheta_j|\le t_{n-s,1-\frac{\alpha}{2}}\cdot\hat s\sqrt{b_{jj}})=1-\alpha$$

d.h. $\hat\vartheta_j\pm t_{n-s,1-\frac{\alpha}{2}}\cdot\hat s\sqrt{b_{jj}}$ ist zweiseitiges $(1-\alpha)$-Konfidenzintervall für ϑ_j.

c) quaderförmiger Konfidenzbereich für ϑ („Bonferroni-Methode"):
Regel von den kleinen Ausnahmewahrscheinlichkeiten:

$$P(A_j) \geq 1 - \frac{a}{s},\ j = 1, \dots, s \ \Rightarrow\ P(\bigcap_{j=1}^{s} A_j) \geq 1 - \alpha$$

Denn:

$$P(\bigcap_{j=1}^{s} A_j) = 1 - P((\bigcap_{j=1}^{s} A_j)^C) = 1 - P(\bigcup_{j=1}^{s} A_j^C) \geq 1 - \sum_{j=1}^{s} \underbrace{P(A_j^C)}_{\leq \frac{a}{s}} \geq 1 - \alpha$$

Somit gilt für

$$C_Q(x) := \times_{j=1}^{s} [\hat{\vartheta}_j(x) - r(x), \hat{\vartheta}_j(x) + r(x)]$$

mit $r(x) := t_{n-s,1-\frac{\alpha}{2s}} \cdot \hat{s}\sqrt{b_{jj}}$:

$$P_{\vartheta,\sigma^2}(C_Q(X) \ni \vartheta) \geq 1 - \alpha \ \forall \vartheta, \sigma^2$$

d.h. C_Q ist quaderförmiger $(1-\alpha)$-Konfidenzbereich für ϑ.

<u>Bemerkung:</u>
C_E hat kleineres Volumen wie C_Q, aber C_Q ist leichter zu interpretieren.

d) Konfidenzintervall für $a^T\vartheta$:

$$a^T\hat{\vartheta} \sim \mathcal{N}(a^T\vartheta, \sigma^2 \cdot a^T(C^TC)^{-1}a)$$

$$\Rightarrow \frac{a^T(\hat{\vartheta} - \vartheta)}{\hat{s}\sqrt{a^T(C^TC)^{-1}a}} = \frac{\frac{a^T(\hat{\vartheta}-\vartheta)}{\sigma\sqrt{a^T(C^TC)^{-1}a}}}{\sqrt{\frac{\hat{s}^2}{\sigma^2}}} \sim t_{n-s}$$

$\Rightarrow$ Mit $r := t_{n-s,1-\frac{\alpha}{2}} \cdot \hat{s}\sqrt{a^T(C^TC)^{-1}a}$ ist $[a^T\hat{\vartheta} - r, a^T\hat{\vartheta} + r]$ $(1-\alpha)$-Konfidenzintervall für $a^T\vartheta$.

<u>Beispiel:</u>
einfache lineare Regression (vgl. Beispiel 14.4(b))

$$a = \begin{pmatrix} 1 \\ t^* \end{pmatrix},\ r = t_{n-2,1-\frac{\alpha}{2}} \cdot \hat{s}\sqrt{\frac{1}{n} + \frac{(t^* - \bar{t})^2}{\sum_i (t_i - \bar{t})^2}}$$

$[\hat{\vartheta}_1 + \hat{\vartheta}_2 \cdot t^* - r, \hat{\vartheta}_1 + \hat{\vartheta}_2 \cdot t^* + r]$ ist $(1-\alpha)$-Konfidenzintervall für $a^T\vartheta = \vartheta_1 + \vartheta_2 \cdot t^*$.

14.8 Tests von linearen Hypothesen im linearen Modell

$X = C\vartheta + \varepsilon$, $\varepsilon \sim \mathcal{N}_n(0, \sigma^2 \cdot I_n)$
Zu testen sei „lineare Hypothese“

$$H_0: \ H\vartheta = h \ \text{ gegen } \ H_1: \ H\vartheta \neq h$$

Dabei: H $r \times s$-Matrix, Rang$(H) = r$ *(insbesondere $r \leq s$)*, $h \in \mathbb{R}^r$ gegeben

$$H_0 \hat{=} \Theta_0 := \{(\vartheta, \sigma^2) \in \underbrace{\mathbb{R}^s \times \mathbb{R}_{>0}}_{=\Theta}: \ H\vartheta = h\}, \ H_1 \hat{=} \Theta \backslash \Theta_0$$

14.9 Beispiele

a) $X_j = \vartheta_1 + \vartheta_2 \cdot t_j + \varepsilon_j$, $j = 1, \ldots, n$ (einfache lineare Regression)

$$H_0: \vartheta_2 = 0 \ \text{ gegen } \ H_1: \vartheta_2 \neq 0$$

„Lineare Hypothese“: $H = (0, 1)$, $h = 0$ $(s = 2, r = 1)$

$$H_0: \ H \cdot \begin{pmatrix} \vartheta_1 \\ \vartheta_2 \end{pmatrix} = 0$$

<u>Möglicher Test:</u> Verallgemeinerter Likelihood-Quotienten-Test
<u>Testgröße</u> Λ_n bzw. $\log \Lambda_n$.

$$\Lambda_n := \frac{\sup_{(\vartheta,\sigma^2) \in \Theta_0} f(x, \vartheta, \sigma^2)}{\sup_{\Theta} f(x, \vartheta, \sigma^2)}$$

<u>Unter H_0:</u> $X_j = \vartheta_1 + \varepsilon_j$, $X_j \sim \mathcal{N}(\vartheta_1, \sigma^2)$, ML-Schätzer für ϑ_1: $\bar{X}_n$
<u>Ohne Restriktion:</u> ML-Schätzer = KQ-Schätzer[41] $= \hat{\vartheta}$ (Satz 14.6(a))
Als Schätzer für σ^2 wird aber üblicherweise in beiden Fällen der Schätzer $\hat{\sigma}^2$ aus Obermodell verwendet!
Dann[42]:

$$\log \Lambda_n = -\frac{1}{2\hat{\sigma}^2}[\underbrace{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}_{=:SS_0} - \underbrace{\sum_{i=1}^{n}(X_i - (\hat{\vartheta}_1 + \hat{\vartheta}_2 t_i))^2}_{=SS_1 (=n\hat{\sigma}^2)}]$$

Als Testgröße wird

$$T := \frac{SS_0 - SS_1}{\frac{SS_1}{n-2}}$$

verwendet. Es gilt:

[41]Kleinste-Quadrate-Schätzer
[42]SS: sum of squares

(i) $\frac{SS_1}{\sigma^2} \sim \chi^2_{n-2}$ (nach 14.6(c))
(ii) $\frac{SS_0}{\sigma^2} \sim \chi^2_{n-1}$ unter H_0 (nach 2.2)
(iii) $SS_0 - SS_1$ und SS_1 stochastisch unabhängig (ohne Beweis)

$$\underbrace{\frac{SS_0}{\sigma^2}}_{\overset{H_0}{\sim}\chi^2_{n-1}} = \frac{SS_0 - SS_1}{\sigma^2} + \underbrace{\frac{SS_1}{\sigma^2}}_{\sim\chi^2_{n-2}}$$

$\Rightarrow \frac{SS_0-SS_1}{\sigma^2} \sim \chi^2_{n-1-(n-2)} = \chi^2_1$ unter H_0 (vgl. Beweis von 14.6(c))

Damit $T \sim F_{1,n-2}$ unter H_0.

b) $X_{i,j} = \vartheta_j + \varepsilon_{i,j}$ $(i = 1, \ldots, k,\ j = 1, \ldots, n_i)$ (einfache Varianzanalyse[43])

$$H_0 : \vartheta_1 = \ldots = \vartheta_k$$

(„kein Effekt des zu untersuchenden Faktors“)

$$\underbrace{\begin{pmatrix} 1 & & 0 & -1 \\ & \ddots & & \vdots \\ 0 & & 1 & -1 \end{pmatrix}}_{=:H\in\mathbb{R}^{k-1\times k}} \cdot \begin{pmatrix} \vartheta_1 \\ \vdots \\ \vartheta_k \end{pmatrix} = \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}}_{=:h\in\mathbb{R}^{k-1}}$$

$\mathrm{Rang}(H) = k - 1 (= r)$
Testgröße: (vgl. Aufgabe 45)
Sei $\bar{X}_{i+} = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{i,j}$, $\bar{X}_{++} = \frac{1}{n}\sum_{i,j} X_{i,j}$, $n = \sum_{i=1}^{k} n_i$,
$\mathrm{SQZ} = \sum_{i=1}^{k} n_i(\bar{X}_{i+} - \bar{X}_{++})^2$, $\mathrm{SQI} = \sum_{i,j}(X_{i,j} - \bar{X}_{i+})^2$

$$\sum_{i,j}(X_{i,j} - \bar{X}_{++})^2 = \mathrm{SQI} + \mathrm{SQZ}$$

$$T := \frac{\frac{\mathrm{SQZ}}{k-1}}{\frac{\mathrm{SQI}}{n-k}} \sim F_{k-1,n-k} \quad \text{unter } H_0$$

[43] $k \hat{=} s$

14.10 Die Testgröße bei allgemeinen linearen Hypothesen

$\hat{\vartheta} \sim \mathcal{N}_j(\vartheta, \sigma^2 (C^T C)^{-1})$

$$\Rightarrow H\hat{\vartheta} \sim \mathcal{N}_r(H\vartheta, \sigma^2 \underbrace{H(C^T C)^{-1} H^T}_{=:B})$$

$$\frac{\frac{1}{r} \cdot \frac{1}{\sigma^2}(H\hat{\vartheta} - H\vartheta)^T B^{-1} (H\hat{\vartheta} - H\vartheta)}{\frac{\hat{s}^2}{\sigma^2}} \sim \frac{\frac{\chi_r^2}{r}}{\frac{\chi_{n-s}^2}{n-s}} \sim F_{r,n-s}$$

(Zähler und Nenner sind stochastisch unabhängig.)
Sei

$$T := \frac{\frac{1}{r}(H\hat{\vartheta} - h)^T (H(C^T C)^{-1} H^T)^{-1} (H\hat{\vartheta} - h)}{\hat{s}^2} \sim F_{r,n-s} \quad \text{unter } H_0$$

Der sogenannte **F-Test** im linearen Modell besitzt die Gestalt:
H_0 ablehnen, falls $T \geq F_{r,n-s,1-\alpha}$.
Kein Widerspruch zu H_0, falls $T < F_{r,n-s,1-\alpha}$.

Bemerkung:
Für die Beispiele aus 14.9 stimmt die obige Testgröße mit den Testgrößen aus 14.9(a) bzw. (b) überein.