Prof. Dr. E. Spodarev
W. Karcher

Wintersemester 2011 / 2012
16. November 2011

Stochastik III – Übungsblatt 3

Abgabe: 30. November 2011, vor den Übungen

Aufgabe 1 (4)

Betrachte den Spezialfall $m = 2$ des multiplen linearen Regressionsmodells, d.h. $Y = X\beta + \varepsilon$ mit

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Zeige, dass dann der Kleinste-Quadrate-Schätzer $\left(\widehat{\beta}_1, \widehat{\beta}_2\right)^\top = \left(X^\top X\right)^{-1} X^\top Y$ dem aus Statistik I / Stochastik I bekannten Schätzer im einfachen linearen Regressionsmodell entspricht, d.h.

$$\widehat{\beta}_2 = \frac{S_{XY}^2}{s_{XX}^2}, \quad \widehat{\beta}_1 = \overline{Y}_n - \widehat{\beta}_2 \overline{x}_n,$$

wobei $\overline{x}_n$ bzw. $\overline{Y}_n$ die Stichprobenmittel bezeichnen, d.h.

$$\overline{x}_n = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{bzw.} \quad \overline{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i,$$

und die Stichprobenvarianz S_{XX}^2 bzw. die Stichprobenkovarianz S_{XY}^2 gegeben sind durch

$$S_{XX}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \overline{x}_n)^2 \quad \text{bzw.} \quad s_{XY}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \overline{x}_n)\left(Y_i - \overline{Y}_n\right).$$

Aufgabe 2

Das Produktionsvolumen der USA zwischen 1932 und 1953 lässt sich mit Hilfe der Cobb-Douglas Produktionsfunktion

$$Y_t = c \cdot K_t^{\beta_1} \cdot A_t^{\beta_2} \cdot \varepsilon_t, \quad t \in \{1, \ldots, 22\}$$

mit unbekannten Konstanten c, β_1, β_2 beschreiben. Dabei bezeichnet Y die Produktion (in Mrd. Dollar), K den Kapitaleinsatz (in Mrd. Dollar) und A den Arbeitseinsatz (in Mill. Arbeitskräften). Die Datei *production.txt* auf der Homepage der Vorlesung enthält folgende Daten:

Jahr	t	Produktion	Kapital	Arbeit
1932	1	60.3	297.1	39.3
1933	2	58.2	290.1	39.6
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

a) Führe den Modellansatz in ein geeignetes lineares Modell über und bestimme mit `R` den Kleinste-Quadrate-Schätzer für $\beta = (\log c, \beta_1, \beta_2)$. (3)

b) Gib die geschätzten Werte $\widehat{Y}_t$ sowie die Residuen $\widehat{\varepsilon}_t$ an. (2)

c) Erstelle *ein* Schaubild, das die tatsächliche Entwicklung des Produktionsvolumens Y über die Jahre 1932 – 1953 darstellt, sowie die „Verlaufskurve“ der geschätzten Daten, d.h. eine stückweise lineare Funktion durch die Punkte $(t, \widehat{Y}_t)$. (2)

Hinweis: Auf der Homepage ist eine Einführung zu `R` verfügbar. Mit dem Befehl `lines()` kann ein weiterer Polygonzug zu einem Schaubild hinzugefügt werden.

Aufgabe 3

Gegeben sei das lineare Modell $Y = X\beta + \varepsilon$ mit $\mathbb{E}\varepsilon = 0$ und $\operatorname{Cov}(\varepsilon) = \sigma^2 V$, wobei V eine symmetrische, positiv definite $n \times n$ - Matrix ist und X eine $n \times m$ - Matrix mit $\operatorname{rg}(X) = m$.

a) Sei A eine symmetrische, positiv definite $m \times m$ - Matrix und sei B eine $r \times m$ - Matrix mit $\operatorname{rg}(B) = r \leq m$. Zeige, dass dann auch die Matrizen $BAB^\top$ und A^{-1} positiv definit sind. (3)

b) Zeige, dass $\widehat{\beta} = \left(X^\top V^{-1} X\right)^{-1} X^\top V^{-1} Y$ der Kleinste-Quadrate-Schätzer für β bezüglich der Norm $\|x\|_V = \sqrt{x^\top V^{-1} x}$ ist, d.h. zeige dass gilt: (5)

$$\left\|Y - X\widehat{\beta}\right\|_V^2 = \min_{\beta \in \mathbb{R}^m} \|Y - X\beta\|_V^2.$$

c) Zeige, dass $\widehat{\beta}$ erwartungstreu für β ist. (2)

d) Zeige, dass gilt: $\operatorname{Cov}\left(\widehat{\beta}\right) = \sigma^2 \left(X^\top V^{-1} X\right)^{-1}$. (2)