Übung 6

1. Wir betrachten das lineare Regressionsmodell

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

mit

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 5 \\ 1 & 7 \end{pmatrix} \;,\; \mathbf{Y} = \begin{pmatrix} 34 \\ 47 \\ 55 \\ 64 \end{pmatrix} \;,\; \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \;,\; \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix}$$

a) Berechnen Sie $\mathbf{X}^t\mathbf{X}$, $\mathbf{X}^t\mathbf{Y}$, $\mathbf{Y}^t\mathbf{Y}$ und $(\mathbf{X}^t\mathbf{X})^{-1}$

b) Der least–squares Schätzer für $\boldsymbol{\beta}$ ist gegeben durch

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{Y}$$

Berechnen Sie $\hat{\boldsymbol{\beta}}$.

c) Führen Sie die Regressionsanalyse mit S-PLUS durch.

2. Schreiben Sie für die folgenden Regressionsmodelle die Designmatrix und den Parametervektor auf. Um welchen speziellen Typ von Regression handelt es sich jeweils?

a) $Y_i = \beta_1 x_i + \epsilon_i \qquad i = 1, \ldots, 4$

b) $Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i \qquad i = 1, \ldots, n$

c) $Y_i = \begin{cases} \mu_1 + \epsilon_i & \text{für } i = 1, \ldots, 8 \\ \mu_2 + \epsilon_i & \text{für } i = 9, \ldots, 20 \end{cases}$

3. Die Datei `/afs/ethz.ch/users/s/sfs/datasets/farm.dat` enthält von 20 Bauernhöfen die Hofgrösse A (in acres), die Anzahl Kühe C und das Einkommen I (in $).

a) Rechnen Sie eine einfache Regression von I auf C. Hängt das Einkommen von der Anzahl Kühe ab?

b) Berechnen Sie ein Vertrauensintervall für das erwartete Einkommen ohne Kühe, mit 20 Kühen, mit $\bar{C} = 8.85$ Kühen. Berechnen Sie ein Prognoseintervall für das Einkommen einer Farm ohne Kühe, mit $\bar{C} = 8.85$ Kühen.

c) Rechnen Sie einfache lineare Regressionen von I auf A und von A auf C, sowie eine multiple Regression von I auf A und C. Wie erklären Sie sich die verschiedenen Ergebnisse?

Abgabe: Montag, 13.12.99