Prüfung: Lineare Algebra und Statistik

Musterlösung

1. **a)** Es gilt

$$\left(\begin{array}{rrrr|r} 1 & 0 & 1 & -2 & 2 \\ 2 & 2 & 6 & 0 & 0 \\ -2 & 2 & 2 & 8 & \beta \\ 4 & 0 & 5 & -3 & 1 \end{array}\right) \to \left(\begin{array}{rrrr|r} 1 & 0 & 1 & -2 & 2 \\ 0 & 2 & 4 & 4 & -4 \\ 0 & 2 & 4 & 4 & \beta+4 \\ 0 & 0 & 1 & 5 & -7 \end{array}\right)$$

$$\to \left(\begin{array}{rrrr|r} 1 & 0 & 1 & -2 & 2 \\ 0 & 2 & 4 & 4 & -4 \\ 0 & 0 & 0 & 0 & \beta+8 \\ 0 & 0 & 1 & 5 & -7 \end{array}\right) \to \left(\begin{array}{rrrr|r} 1 & 0 & 1 & -2 & 2 \\ 0 & 2 & 4 & 4 & -4 \\ 0 & 0 & 1 & 5 & -7 \\ 0 & 0 & 0 & 0 & \beta+8 \end{array}\right).$$

b) Aus der Zeilenstufenform kann man ablesen, dass

1. die Anzahl Pivots des Gleichungssystems $Ax = b$ gleich 3 ist;
2. der Rang von (A, b) von β abhängt;
3. die Dimension des homogenen Lösungsraums L von $Ax = 0$ gleich 1 ist;
4. die Dimension des Vektorraums S, der von den Spalten von A aufgespannt wird, gleich 3 ist.
5. das Gleichungssystem $Ax = b$ genau dann lösbar ist, wenn $\beta = -8$ gilt.

c) Für $\beta = -8$ gilt, mit der Wahl des freien Parameters $x_4 = s$

$$\begin{aligned} x_3 &= -7 - 5s \\ x_2 &= -2 - 2x_3 - 2x_4 = -2 - 2(-7-5s) - 2s = 12 + 8s \\ x_1 &= 2 - x_3 + 2x_4 = 2 - (-7-5s) + 2s = 9 + 7s, \end{aligned}$$

anders gesagt ist die Lösungsmenge gleich

$$\left\{ \begin{pmatrix} 9+7s \\ 12+8s \\ -7-5s \\ s \end{pmatrix} : s \in \mathbb{R} \right\}.$$

2. **a)** Es gilt

$$CD = \begin{pmatrix} -1 & -2 & 7 \\ 0 & -2 & 4 \\ 4 & -9 & 6 \end{pmatrix}$$

und $\det(CD) = 0$. Dies kann man entweder durch direkte Rechnung verifizieren, oder aus der Ungleichung $\operatorname{rang}(CD) \leq \operatorname{rang}(C) = 2$ herleiten.

b) Es gilt
$$DC = \begin{pmatrix} 4 & 1 \\ 4 & -1 \end{pmatrix}.$$
Da $\det(DC) = -8$ gilt, ist DC invertierbar. Wir haben
$$\left(\begin{array}{cc|cc} 4 & 1 & 1 & 0 \\ 4 & -1 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 4 & 1 & 1 & 0 \\ 0 & -2 & -1 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 4 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} \end{array}\right)$$
$$\to \left(\begin{array}{cc|cc} 4 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{8} & \frac{1}{8} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} \end{array}\right).$$
Es folgt
$$(CD)^{-1} = \begin{pmatrix} \frac{1}{8} & \frac{1}{8} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

c) Wir haben
$$\det(M_n - \lambda I_3) = \det \begin{pmatrix} n-\lambda & 0 & n \\ 0 & n-\lambda & 0 \\ n & 0 & n-\lambda \end{pmatrix}$$
$$\overset{\text{Laplace 2.Z.}}{=} (n-\lambda) \det \begin{pmatrix} n-\lambda & n \\ n & n-\lambda \end{pmatrix} = (n-\lambda)\left((n-\lambda)^2 - n^2\right) = \lambda(n-\lambda)(\lambda-2n).$$
Die Eigenwerte von M_n sind die Nullstellen des obigen Ausdrucks, nämlich $0, n$ und $2n$.

d) Da $n \neq 0$ ist, hat die 3×3-Matrix M_n drei verschiedene Eigenwerte. Folglich ist sie für jeden betrachteten Wert von n diagonalisierbar.

3. a) Für $\sigma = 5000$ und $n = 100$ gilt
$$\overline{x} \pm z_{1-\frac{0.05}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 49200 \pm 1.96 \cdot \frac{5000}{10} = 49200 \pm 980.$$
Folglich ist das gesuchte Konfidenzintervall
$$[49200 - 980, 49200 + 980] = [48220, 50180].$$

b) $H_0 : \mu = 50000$, $H_A : \mu \neq 50000$. Da μ im oben gefundenen Konfidenzintervall liegt, wird die Nullhypothese beibehalten.

c) Der Wert der Test-statistik $T = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ ist $t = \frac{(49200-50000)}{500} = -\frac{8}{5} = -1.6$. Der P-Wert ist gleich
$$P(T \leq -1.6) + P(T \geq 1.6) = 2(1 - P(T \leq 1.6)) = 2(1-0.9452) = 0.1096 \approx 11\%.$$

Siehe nächstes Blatt!

d) In den vorherigen Gleichungen müsste $z_{0.975}$ durch $t_{99;0.975}$ ersetzt werden. Dieser Wert erscheint in der Tabelle nicht, zu diesen Aufgabenteil genügt jedoch die Abschätzung $t_{99;0.975} \geq t_{120;0.975} > z_{0.975} = 1.96$. Folglich wird das Konfidenzintervall grösser; die Testentscheidung (H_0 beibehalten) ändert sich also nicht.

4. a) Im Output der Residuen stehen 20 Komponenten, also ist Antwort (i) richtig.

b) Die Regressionsgleichung lautet

$$Y = -0.2 + 0.2T_1 + 0.3T_2 + 0.6T_3.$$

c) Der grösste Koeffizient gehört zu T_3; folglich ist die richtige Antwort (iii).

d) Einsetzen in die Regressionsgleichung ergibt

$$Y = -0.2 + 0.2 \cdot 4.3 + 0.3 \cdot 4.2 + 0.6 \cdot 4.6 = 4.68 \approx 4.7.$$

e) Zum Beispiel ist $(T_1, T_2, T_3) = (3.6, 3.6, 4)$ ein Tripel, für welches $Y = 4$ gilt.

f) Wir arbeiten mit der $t_{n-k-1;1-\frac{\alpha}{2}}$-Verteilung, wobei n die Anzahl Studenten, k die Anzahl erklärende Variablen, und $\alpha = 10\%$ ist. Es gilt also $n - k - 1 = 20 - 3 - 1 = 16$ und $1 - \frac{\alpha}{2} = 0.95$. Das Konfidenzintervall wird folglich durch

$$\beta_2 \pm t_{16;0.95} \cdot \text{Std.Err}(\beta_2) = 0.3 \pm 1.746 \cdot 0.08 = 0.3 \pm 0.14$$

berechnet und ist damit gleich $[0.16, 0.44]$.

Da der P-Wert 0.0013, der β_2 entspricht, kleiner als 10% ist, ist der Zusammenhang zwischen T_2 und Y signifikant auf dem 10%-Niveau. Dies bedeutet, das 0 im 90%-Konfidenzintervall nicht enthalten ist.