Aufgabe 22 (Bayesianisches lineares Modell; Gibbs-Sampler)

Betrachtet wird das lineare Modell in Matrixform

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad \text{mit} \quad \boldsymbol{\epsilon} \sim \mathrm{N}_n\left(\boldsymbol{0}, \sigma^2 \boldsymbol{I}\right),$$

wobei $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$.

Für $\boldsymbol{\beta}$ wird als Priori-Verteilung eine multivariate Normalverteilung und für σ^2 wird als Priori-Verteilung eine Inverse Gammaverteilung angenommen, wobei $\boldsymbol{\beta}$ und σ^2 a priori unabhängig sind

$$\begin{aligned} \boldsymbol{\beta} &\sim \mathrm{N}_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \\ \sigma^2 &\sim \mathrm{IG}(a, b). \end{aligned}$$

Als vollständig bedingte Dichten ergeben sich für $\boldsymbol{\beta}$ eine multivariate Normalverteilung

$$\boldsymbol{\beta} | \boldsymbol{y}, \boldsymbol{X}, \sigma^2 \sim \mathrm{N}_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$$

mit

$$\begin{aligned} \boldsymbol{\mu}_1 &= \left[\frac{1}{\sigma^2}\boldsymbol{X}^\top\boldsymbol{X} + \boldsymbol{\Sigma}_0^{-1}\right]^{-1} \left[\frac{1}{\sigma^2}\boldsymbol{X}^\top\boldsymbol{y} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0\right] \\ \boldsymbol{\Sigma}_1 &= \left[\frac{1}{\sigma^2}\boldsymbol{X}^\top\boldsymbol{X} + \boldsymbol{\Sigma}_0^{-1}\right]^{-1} \end{aligned}$$

und für σ^2 eine Inverse Gammaverteilung (siehe dazu auch die Tutorien 9 und 10)

$$\sigma^2 | \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{\beta} \sim \mathrm{IG}\left(a + \frac{n}{2}, b + \frac{1}{2}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})\right).$$

Die Bayes-Schätzer können mittels Gibbs-Sampling berechnet werden.

(a) Beschreiben Sie den Gibbs-Sampling-Algorithmus zur Simulation aus der gemeinsamen Posteriori-Verteilung von $\boldsymbol{\beta}, \sigma^2 | \boldsymbol{y}, \boldsymbol{X}$ in eigenen Worten oder Pseudocode.

(b) Wie können Sie mit diesen Zufallsziehungen Punktschätzer für $\boldsymbol{\beta}$ und σ^2 bestimmen?

(c) Wie können Sie mit diesen Zufallsziehungen Kredibilitätsintervalle für $\boldsymbol{\beta}$ und σ^2 bestimmen?

(d) Angenommen, ein Anwender interessiert sich für die Frage, ob die Kovariablen x_2 und x_3 den gleichen Effekt auf die abhängige Variable y haben oder nicht: Wie können Sie diese Frage anhand der Simulationen aus der gemeinsamen Posteriori-Verteilung von $\boldsymbol{\beta}, \sigma^2 | \boldsymbol{y}, \boldsymbol{X}$ beantworten?