University of Iceland

School of Engineering and Sciences

Faculty of Physical Sciences

Department of Mathematics

STÆ004F Random Effects Models

Spring 2024

Homework 1

Assigned: Friday January 12th 2024.

Due: Friday February 2nd 2024.

Exercises:

1. Here a few exercise on the matrix algebra involved in the statistical inference.

   (a) Show that
   $$\begin{pmatrix} Q_\epsilon & -Q_\epsilon Z \\ -Z^\mathsf{T} Q_\epsilon & Q_x + Z^\mathsf{T} Q_\epsilon Z \end{pmatrix}^{-1} = \begin{pmatrix} Z Q_x^{-1} Z^\mathsf{T} + Q_\epsilon^{-1} & Z Q_x^{-1} \\ Q_x^{-1} Z^\mathsf{T} & Q_x^{-1} \end{pmatrix}.$$

   (b) Show that if
   $$\pi(\boldsymbol{y}, \boldsymbol{x} \mid \boldsymbol{\theta}) = \mathcal{N}\left(\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{pmatrix} \middle| \begin{pmatrix} Z\boldsymbol{\mu}_x \\ \boldsymbol{\mu}_x \end{pmatrix}, \begin{pmatrix} Q_\epsilon & -Q_\epsilon Z \\ -Z^\mathsf{T} Q_\epsilon & Q_x + Z^\mathsf{T} Q_\epsilon Z \end{pmatrix}^{-1} \right)$$
   then the posterior distribution of $\boldsymbol{x}$ conditional on $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ is Gaussian with mean and precision matrix
   $$\begin{aligned} \boldsymbol{\mu}_{x|y} &= Q_{x|y}^{-1}(Q_x \boldsymbol{\mu}_x + Z^\mathsf{T} Q_\epsilon \boldsymbol{y}), \\ Q_{x|y} &= Q_x + Z^\mathsf{T} Q_\epsilon Z. \end{aligned}$$
   Hint: Theorem 2.5, p.37, in Rue and Held (2005).

   (c) Show that if
   $$\pi(\boldsymbol{y}, \boldsymbol{x} \mid \boldsymbol{\theta}) = \mathcal{N}\left(\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{pmatrix} \middle| \begin{pmatrix} Z\boldsymbol{\mu}_x \\ \boldsymbol{\mu}_x \end{pmatrix}, \begin{pmatrix} Z\Sigma_x Z^\mathsf{T} + \Sigma_\epsilon & Z\Sigma_x \\ \Sigma_x Z^\mathsf{T} & \Sigma_x \end{pmatrix} \right)$$
   then the posterior distribution of $\boldsymbol{x}$ conditional on $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ is Gaussian with mean and covariance matrix
   $$\begin{aligned} \boldsymbol{\mu}_{x|y} &= \boldsymbol{\mu}_x + \Sigma_x Z^\mathsf{T}(Z\Sigma_x Z^\mathsf{T} + \Sigma_\epsilon)^{-1}(\boldsymbol{y} - Z\boldsymbol{\mu}_x), \\ \Sigma_{x|y} &= \Sigma_x - \Sigma_x Z^\mathsf{T}(Z\Sigma_x Z^\mathsf{T} + \Sigma_\epsilon)^{-1} Z\Sigma_x. \end{aligned}$$
   Hint: Wikipedia: Multivariate normal distribution.

2. Let $y_{i,j}$ be an observation of individual j in category i. The following model is assumed for $y_{i,j}$
   $$y_{i,j}|\eta, u_i, \sigma^2 \sim \mathcal{N}(\eta + u_i, \sigma^2), \quad j \in \{1, ..., n\}, \quad i \in \{1, ..., Q\},$$
   where n is the number of observations from each category and Q is the number of categories. The $y_{i,j}$s are independent within each category i and they are also independent between categories. The u_is are modeled at the latent level as
   $$u_i|\tau^2 \sim \mathcal{N}(0, \tau^2), \quad i \in \{1, ..., Q\},$$

where the u_is are independent of each other (conditional on τ).

The parameter η is modeled at the latent level with a Gaussian density with mean zero and variance 100^2. Assume that σ is known. The parameter τ is the hyperparameter of the model, and it is assigned an exponential prior density with mean $\kappa_\tau = 2$.

The data are in the file `y_hw1.dat`. Each column in the file contains the observations from one of the categories. The number of columns is $Q = 53$ and the number of rows is $n = 41$. Here $\sigma = 2.51$.

(a) Specify the hierarchical model using a vector and matrix notation, that is, specify the densities at the response level, the latent level and the hyperparameter level. Let $\boldsymbol{y}_i = (y_{i1}, ..., y_{in})^\top$ denote the observations from category i, $i \in \{1, ..., Q\}$. At the latent level, specify the density of $\boldsymbol{x} = (\eta, u_1, ..., u_Q)^\top$ conditional on τ in term of a precision matrix. At the response level, specify the density of $\boldsymbol{y} = (\boldsymbol{y}_1{}^\top,\ ..., \boldsymbol{y}_Q{}^\top)^\top$ conditional on $\boldsymbol{x}$, τ and σ in term of a precision matrix.

(b) Find the posterior distribution of the unknown parameters in the hierarchical model in terms of vectors and matrices using the precision matrix representation.

(c) Find the mathematical form of the marginal posterior density of τ using the precision matrix representation. Draw it on the interval $[0, 3]$. Use a fine grid and scale the posterior density of τ such that its integral is approximately one.
Hint: Find first the density of $\boldsymbol{y}$ conditional on τ only.

(d) Find the conditional posterior density of $\boldsymbol{x}$ given τ using the precision matrix representation.

(e) Reparameterize with $\theta = \log(\tau)$. The prior density for θ is based on the prior density for τ and given by

$$p(\theta) = \kappa_\tau^{-1} \exp(-\kappa_\tau^{-1} \exp(\theta) + \theta), \quad \theta \in \mathbb{R}, \quad \theta \in \mathbb{R}.$$

Write a program for a posterior sampler that uses the marginal posterior density of θ and the conditional posterior density of $\boldsymbol{x}$ given θ.
Now, a sample of θ is drawn first from the marginal posterior density of θ using a Metropolis step and then a sample of $\boldsymbol{x}$ is drawn from the conditional posterior density of $\boldsymbol{x}$ given θ.

(f) Compute the posterior medians and 95% posterior intervals for η, u_1, ..., u_Q, and τ. Use four chains where each chain consists of 13000 iterations and the first 3000 are used for burn-in. Draw the posterior and prior densities of τ on the same graph. Present the posterior medians and 95% posterior intervals for η and τ in a table. Present the u_js in a figure, that is, plot the number of the category on the x-axis and the 95% posterior intervals along with a point for the posterior median on the y-axis.