Example 2: Bayesian Linear Regression

Bayesian linear regression model

We have the model

$$\begin{aligned} y_i \mid \boldsymbol{\beta} &\sim \mathcal{N}(\boldsymbol{x}_i^T\boldsymbol{\beta}, \phi^{-1}), \quad i = 1, \dots, n; \\ \boldsymbol{\beta} \mid \kappa &\sim \mathcal{N}(0, \kappa^{-1}\mathrm{I}); \\ \kappa &\sim \mathrm{Gamma}(a_0, b_0); \end{aligned}$$

where $\phi = 1/\sigma^2$ is the precision parameter, which we assume is known, $\boldsymbol{x}_i,\ i = 1, \dots, n$ are known covariates, $\boldsymbol{\beta} \in \mathbb{R}^p$ includes the intercept, and is unknown, and I is the identity matrix. Assume a_0 and b_0 are known.

Exercise

We want a mean-field approximation to the posterior

$$p(\boldsymbol{\beta}, \kappa \mid \boldsymbol{y})$$

on the form

$$q(\boldsymbol{\beta}, \kappa) = q(\boldsymbol{\beta})q(\kappa).$$

First, derive the CAVI updates. Then implement the algorithm, and run on simulated data with one covariate $x_{i,1} \sim \mathcal{N}(0, 1)$, $n = 50$, $\phi = 0.5$, $\beta_0 = -1, \beta_1 = 2$, $a_0 = b_0 = 0.001$. Assess convergence by the variational factors, or derive the ELBO to assess convergence. Visualize the resulting bivariate Gaussian approximation for the intercept β_0 and coefficient β_1.

Solution

For κ, the variational approximation is on the form

$$q(\kappa) \propto \mathrm{Gamma}(a_n, b_n);$$

with

$$a_n = a_0 + \frac{p}{2}, \quad b_n = b_0 + \frac{\mathbb{E}_{q_{\boldsymbol{\beta}}}[\boldsymbol{\beta}^T\boldsymbol{\beta}]}{2}.$$

For $\boldsymbol{\beta}$, the variational approximation is on the form

$$q(\boldsymbol{\beta}) \propto \mathcal{N}(\boldsymbol{m}_n, \boldsymbol{S}_n);$$

with

$$\boldsymbol{S}_n = (\mathbb{E}_{q_\kappa}[\kappa]\mathrm{I} + \phi \boldsymbol{X}^T \boldsymbol{X})^{-1}, \quad \boldsymbol{m}_n = \phi \boldsymbol{S}_n \boldsymbol{X}^T \boldsymbol{y};$$

where

$$\mathbb{E}_{q_\kappa}[\kappa] = \frac{a_n}{b_n}, \quad \mathbb{E}_{q_{\boldsymbol{\beta}}}[\boldsymbol{\beta}^T \boldsymbol{\beta}] = \boldsymbol{m}_n^T \boldsymbol{m}_n + \mathrm{tr}(\boldsymbol{S}_n).$$

R code is given in a separate file.