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.

We want a mean-field approximation to the posterior

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

on the form

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

Solution

For κ, the variational approximation is on the form

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

with

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

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

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

with

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

where

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