Case Study 4 — Hierarchical Bayesian Linear Regression: Gradient-Based Optimisation and Posterior Variance Collapse

David Ewing (82171165)

12 May 2026

Abstract

This report applies the gradient-based optimisation methods from ENEL445 to variational inference in hierarchical Bayesian linear regression. Extending the simple Gaussian regression setting, this model introduces $J = 5$ group-level random effects $u_j \sim \mathcal{N}(0, \tau_u^{-1})$ alongside the noise precision τ_e as additional latent variables with Gamma priors. The Evidence Lower Bound (ELBO) for this conjugate model admits closed-form CAVI updates, but also serves as a smooth objective for gradient-based methods, with $D = 19$ unconstrained parameters.

Four optimisation methods are applied to maximise the ELBO: Coordinate Ascent Variational Inference (CAVI), gradient ascent with Armijo backtracking, Newton's method with regularised finite-difference Hessian, and BFGS with Wolfe conditions. The variational family is a mean-field product $q(\boldsymbol{\beta}) \prod_j q(u_j) \cdot q(\tau_e) \cdot q(\tau_u)$ with Gaussian random effects and Gamma precision approximations, and a Cholesky reparameterisation for $\boldsymbol{\Sigma}_\beta$.

The reference posterior is obtained from a conjugate blocked Gibbs sampler using 3 chains of 5000 iterations (500 burn-in each). A striking finding is severe posterior variance collapse for β_0 (SD ratio ≈ 0.13 for CAVI and BFGS), driven by the mean-field factorisation absorbing shared group-level variance into the random effects u_j rather than the population intercept. Newton's method fails to converge to the correct optimum, highlighting the sensitivity of second-order methods to poor Hessian conditioning in high-dimensional hierarchical models.

Awaiting review by Professor Le Yang