Case Study 5 — Hierarchical Bayesian Logistic 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 logistic regression. This model combines two sources of analytical difficulty from previous case studies: the non-conjugate logistic likelihood (Case Study 3) and the hierarchical random-effects structure (Case Study 4). Group random effects $u_j \sim \mathcal{N}(0, \tau_u^{-1})$ augment the fixed effects $\boldsymbol{\beta}$, and the Jaakkola–Jordan (Jaakkola & Jordan, 2000) bound is required to retain a tractable ELBO. The unconstrained parameter dimension is $D = 17$.

Four optimisation methods are applied: CAVI, gradient ascent with Armijo backtracking, Newton's method, and BFGS. The reference posterior is provided by a Pólya–Gamma blocked Gibbs sampler (Polson et al., 2013) with $T = 50$ truncation terms, 3 chains $\times$ 2000 iterations.

A key finding is pronounced posterior variance collapse in both β_0 and β_1 (SD ratios ≈ 0.46), reflecting the mean-field VI family's failure to capture the posterior correlation between the fixed effects and the group random effects. CAVI and BFGS converge to the same ELBO optimum 1050$\times$ and 81$\times$ faster than the PG Gibbs sampler respectively. Newton's method again fails to find the correct optimum, converging to a suboptimal point with markedly worse ELBO.

Awaiting review by Professor Le Yang