

Case Study 3 — 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 Bayesian logistic regression. Unlike the Gaussian regression setting, logistic regression has no conjugate variational family, and the Evidence Lower Bound (ELBO) does not admit a simple closed-form update. The Jaakkola–Jordan (Jaakkola & Jordan, 2000) variational lower bound provides a tractable ELBO for logistic regression via the local variational bound $\log \sigma(\psi) \geq \frac{\psi}{2} - \log 2 \cosh(\xi/2)$, which is tight when $\xi = \sqrt{\mathbb{E}[\psi^2]}$.

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 $q(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with a Cholesky reparameterisation to preserve positive definiteness throughout optimisation.

The reference posterior is obtained from a Pólya–Gamma Gibbs sampler (Polson et al., 2013) using a truncated sum-of-gammas approximation with $T = 50$ terms. Diagnostics include effective sample size (ESS), autocorrelation function (ACF) plots, and computational cost per effective sample.

The model is $y_i \sim \text{Bernoulli}(\sigma(\beta_0 + \beta_1 x_i))$ with $\boldsymbol{\beta}_{\text{true}} = (-1.0, 2.0)$ and $n = 200$ observations. All four VI methods recover posterior means within 0.02 of the Gibbs reference. CAVI converges in fewer than 50 iterations and is several orders of magnitude faster than the Pólya–Gamma sampler, which requires 200 Gamma draws per iteration. The fair timing comparison is each VI method versus the PG Gibbs reference within this notebook; differences in likelihood structure, parameter count, and Gibbs cost mean that cross-notebook timing figures are not controlled experiments.

Awaiting review by Professor Le Yang