

# Case Study 2 — Bayesian Quadratic Regression: Gradient-Based Optimisation and Posterior Variance Collapse

David Ewing (82171165)

12 May 2026

## Abstract

Mean-field Variational Inference (VI) provides a computationally efficient approximation to Bayesian posterior inference, but it can systematically underestimate posterior uncertainty. This failure mode is the primary concern of this report and is referred to as *posterior variance collapse*. In Bayesian regression, posterior variance collapse appears when the variational covariance matrix for the regression coefficients is too small, producing posterior intervals that are narrower than the uncertainty justified by the data and model.

The primary objective of this report is to investigate posterior variance collapse using the gradient-based optimisation methods taught in ENEL445. VI recasts Bayesian posterior computation as an optimisation problem through maximisation of the Evidence Lower Bound (ELBO), making it a natural setting for applying gradient ascent, Newton’s method, and BFGS.

To support this objective, the numerical study is designed as a complete inference-and-optimisation pipeline. Stochastic data are generated from known Bayesian regression models, and a Gibbs sampler is used as a posterior-reference method. The mean-field variational approximation is then formulated and the ELBO is derived as the objective function. This ELBO objective is then solved using the Coordinate Ascent Variational Inference (CAVI) and the gradient-based methods identified above.

The main objective is to determine whether successful ELBO optimisation also produces reliable posterior uncertainty. This distinction is important because posterior variance collapse can remain even when the ELBO has converged. The methods are therefore compared using both optimisation diagnostics, such as convergence behaviour, final ELBO value, step-size behaviour, and computational cost, and posterior-uncertainty diagnostics, especially posterior standard-deviation ratios relative to the Gibbs reference.

For the gradient-based methods, the report states the optimisation variables, constraints, entry conditions, search directions, step-size rules, and exit conditions. The variational parameters include  $\boldsymbol{\mu}_\beta$ ,  $\boldsymbol{\Sigma}_\beta$ ,  $a_e$ , and  $b_e$ , with Cholesky and log-space reparameterisations used to keep the covariance matrix positive definite and the Gamma parameters positive throughout optimisation. For line-search methods, the search direction must satisfy the ascent condition  $\nabla \text{ELBO}(\boldsymbol{\phi}^{(t)})^T \boldsymbol{d}^{(t)} > 0$ . Gradient ascent uses Armijo backtracking, while BFGS uses Wolfe conditions so that the curvature condition  $\boldsymbol{y}_t^T \boldsymbol{s}_t > 0$  is maintained for the inverse-Hessian approximation.

Two numerical test cases are used: a linear Bayesian regression problem and a quadratic Bayesian regression problem. The quadratic case is treated through feature expansion, so the model remains linear in the coefficient vector while including the nonlinear feature  $x^2$ . Code for generating both test cases and applying the optimisation methods is included in the report, with the full implementation structure provided in the appendix.

Awaiting review by Professor Le Yang