Ratemaking application of Bayesian LASSO with conjugate hyperprior

April 1, 2018

Himchan Jeong[†] Emiliano A. Valdez[‡]

Department of Mathematics, 341 Mansfield Road
University of Connecticut, Storrs, Connecticut 06269-1009

Abstract

The generalized linear model (GLM) is a well developed statistical model widely used in actuarial practice for insurance ratemaking, risk classification, and even reserving. Recently, there has been an explosion of data mining techniques to refine statistical models for better variable selection procedure and for improved prediction accuracy. Such techniques include the increased interest in regularization techniques, or penalized likelihood, to achieve these goals. In this paper, we explore the idea of Least Absolute Selection and Shrinkage Operator (LASSO) in a Bayesian framework within a dependent frequency-severity model as a refinement to the dependent GLM developed by Garrido et al. (2016). The LASSO technique is a penalized least squares procedure developed by Tibshirani (1996) and is extended to a Bayesian interpretation framework by Park and Casella (2008). We show that a new penalty function emerges if we further theoretically extend the Bayesian LASSO using conjugate hyperprior distributional assumptions. While this result has the ease of implementation for variable selection and prediction, we recognize that the use of least squares has been poorly viewed in insurance ratemaking. Instead however, we modify the setting to a penalized dependent GLM within this extended Bayesian LASSO framework. Within such framework, the regression estimates are derived by optimizing a penalized likelihood assuming a hyperprior distribution for the L_1 penalty parameter λ. This has the advantage of avoiding the use of ex-post cross-validation to determine the optimal λ. We calibrated our proposed model using an auto insurance dataset from a Singapore insurance company where we have observed claim counts and amounts from a portfolio of policyholders.

Keywords: Dependent frequency-severity model, GLM, Bayesian LASSO, penalized likelihood, variable selection, data mining.

References

Garrido, J., Genest, C., and Schulz, J. (2016). Generalized linear models for dependent frequency and severity of insurance claims. *Insurance: Mathematics and Economics*, 70:205–215.

Park, T. and Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–288.

[†]himchan.jeong@uconn.edu

[‡]Presenter: emiliano.valdez@uconn.edu, 10th Conference in Actuarial Science & Finance on Samos, Greece