

Fast marginal likelihood estimation of penalties for group-adaptive elastic net

Mirrelijm M. van Nee¹, Tim van de Brug¹, Mark van de Wiel^{1,2}

¹Epidemiology and Data Science, Amsterdam University Medical Centers, Amsterdam, The Netherlands

²MRC Biostatistics Unit, University of Cambridge, Cambridge, UK

E-mail for correspondence: m.vannee@amsterdamumc.nl

Abstract: Nowadays, clinical research routinely uses omics, such as gene expression, for predicting clinical outcomes or selecting markers. Additionally, so-called co-data are often available, providing complementary information on the covariates, like groups of genes corresponding to pathways. Elastic net is widely used for prediction and covariate selection. Group-adaptive elastic net learns from co-data to improve prediction and selection, by penalising important groups of covariates less than other groups. Existing methods are, however, computationally expensive. Here we present a fast method for marginal likelihood estimation of group-adaptive elastic net penalties for generalised linear models. The method uses a low-dimensional representation of the Taylor approximation of the marginal likelihood and its first derivative for group-adaptive ridge penalties, to efficiently estimate these penalties. Then we show by using asymptotic normality of the linear predictors that the marginal likelihood for elastic net models may be approximated well by the marginal likelihood for ridge models. The ridge group penalties are then transformed to elastic net group penalties by using the variance function. The method allows for overlapping groups and unpenalised variables. We demonstrate the method in a cancer genomics application. The method substantially decreases computation time while outperforming or matching other methods by learning from co-data.

Key words: Clinical prediction; Empirical Bayes; Omics; Penalised generalised linear models; Prior information;