Sparse CCA using Lasso

Lykou, Anastasia & Whittaker, Joe
Lancaster University, Mathematics and Statistics Department
Fylde College, LA1 4YF, U.K.

E-mail for correspondence: a.lykou@lancaster.ac.uk

Collinearity among regressors may make ordinary least squares (OLS) estimates unreliable and difficult to interpret. This arises in electronic data recording where typically the number of regressors is larger than the number of observations. It is well known that shrinkage methods, for instance ridge regression, Hoerl and Kennard (1970, 1988), may lead to smaller prediction error. This error may be further be reduced by setting some regression coefficients to zero and just using the resulting subset. More recently, the past decade has seen the development of methods that simultaneously perform shrinkage and selection, including the non negative garotte (Breiman, 1995), the lasso (Tibshirani, 1996) and the elastic net (Zou and Hastie, 2005), all giving sparse predictors.

Here we introduce the lasso with positivity, a restricted version of Tibshirani's lasso. This algorithm gives sparse estimates by considering positivity constraints on the regression coefficients. The positivity constraints may be applied to ensure that the sign of the sparse coefficients is the same as the OLS estimate, and so this simple algorithm still has rather general application. In many cases, the proposed algorithm gives the same estimates as Tibshirani's lasso and lars-lasso algorithm based on the least angle regression (Efron et al., 2004).

We extend the application of the positive lasso to construct a sparse version of canonical correlation analysis (CCA). CCA describes the relationship between two multidimensional variables, by finding linear combinations of the variables with maximal correlation. The canonical variates can be obtained either by solving the eigenvalue equations for the covariance matrix or by using an alternating least squares algorithm. We propose a method of estimating canonical variates by combining the alternating least squares (ALS) algorithm with the positivity constrained lasso or lars-lasso algorithm.

Properties of this method of estimating a sparse CCA and illustrative examples are given in the full version of this paper.

Keywords: *alternating least squares, canonical correlation analysis, Lasso*

References

Breiman, L. (1995). Better subset regression using the nonnegative garotte. Technometrics, *37:373–384.*

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. Annals of Statistics, *32:407–499.*

Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: biased estimation for nonorthogonal problems. Technometrics, *12:55–67.*

Hoerl, A. E. and Kennard, R. W. (1988). Encyclopedia of Statistical Sciences, *volume 8, chapter Ridge regression, pages 129–136. Wiley, New York.*

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. J. Royal. Statist. Soc. B., *58:267–288.*

Zou, J. and Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society Series B, *67:301–320.*