# Proposal 1

John Pluta

October 7, 2014

In a typical regression problem, the solution to maximize the likelihood of the coefficient estimates is usually found through ordinary least squares (OLS). However, there are two cases where the OLS estimate is not ideal. When the number of predictors, p, approaches or exceeds the number of observations, n, the variability of the coefficient estimates is greatly increased. A similar effect is observed when the predictor variables are highly correlated. Penalized regression refers to a class of techniques that address these issues by placing a constraint on the size of the coefficients, which takes the general form of:

$$\hat{\beta} = (Y - X\beta)'(Y - X\beta) - \phi$$

where $\phi$ is the penalty term that shrinks the coefficients. There are many types of penalized regression, and each is uniquely determined by its penalty term. One method of particular interest is the least absolute shrinkage and selection operator (LASSO), proposed in [1], which is a form of penalized regression that also enforces variable selection by shrinking some terms to exactly zero.

The LASSO estimate takes the form:

$$\hat{\beta} = argmin_{\beta}(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta) - \lambda \sum_{j=1}^{p} |\beta_j|$$

There is no closed form solution to this problem, thus quadratic programming methods are required for its solution. However, because the penalized least-squares solution is non-differentiable, popular methods like Newton-Raphson and gradient ascent will fail. For this project, three alternative algorithms will be studied: a modified gradient-based algorithm [2], the popular LARs algorithm [3], and a grafting algorithm [4]. Algorithms will be written in R and Stata.

## References

[1] Tibshirani, R. 1996. *Regression shrinkage and selection via the lasso.* J. R. Statist Soc. 58(1):267-288.

[2] Kim, J., Kim, Y., Kim, Y. 2008.*A gradient-based optimization algorithm for LASSO.* Journal of Computational and Graphical Statistics. 17(4):994-1009.

[3] Efron, B., Hastie, T., Johnstone, Iain, Tibshirani, R. 2004. *Least angel regression.* The Annals of Statistics. 32(2):407-499.

[4] Perkins, S., Lacker, K., Theiler, J. 2003. *Grafting: Fast, incremetal feature selection by gradient descent in function space.* Journal of Machine Learning Research. 3:1333-1356.