**A generalized linear model with lasso or elastic net regularization （GLMNET）**

GLMNET is a model, which fit a generalized linear model via penalized maximum likelihood. The regularization path is computed for the lasso or elastic net penalty at a grid of values for the regularization parameter lambda. This model can deal with all shapes of data, including vary large sparse data matrices. Fits linear, logistic and multinomial, passion, and Cox regression models.

**Model component:**

**The GLM model**: the GLM functions for different families are different.

For “Gaussian” family is

For the other models are

**The penalty**:

**Objective functions:**

For “Gaussian” family is

+

For the other models are

**: The shrinkage parameter**

The value of has direct relationship with the final result. For instance, in ridge regression, controls the size of the coefficients and amount of regularization.

If , we could obtain the least squares solutions.

If , the estimate parameters we get could tend to 0.

In R, the default value of is a sequence of number. The # of values is 100 and the default number of the smallest value of the depends on the sample size relative to the number of the variables. If #observations> #variables, the default is 0.0001, on contrast, the default is 0.01.

**The penalty:**

* **LASSO Regularization**

**Advantage:**

Owing to the nature of the , the lasso could both continuous shrinkage and automatic variable selection simultaneously.

**Limitation:**

1. When the #predictor> #observations, the lasso selects at most n variables before it saturates, because of the convex optimization problem. Moreover, the lasso is not well defines unless the bound on the of the coefficients is smaller than a certain value.
2. If there is a group of variables among which the pairwise correlations are very high, then the lasso tends to select only one variable form the group and does not care which one is selected.

* **Ridge penalty**

.

**Advantage**:

As a continuous shrinkage method, ridge regression achieves its better prediction performance through a bias- variance trade off.

**Limitation:**

Since it always keeps all the predictors in the model, this regression method could not produce a parsimonious model.

* **Elastic net**

The penalty of elastin net is

,

which is a convex combination of the lasso and ridge penalty.

The elastin net simultaneously does automatic variable selection and continuous shrinkage, and it can select groups of correlated variables.

The objective function turns to

***Lemma***: Given data set (y,X) and(), define an artificial data set () by

Let and. The the elastic net criterion can be written as

According to the lemma, this function could be transformed into an equivalent LASSO problem and then be solved.