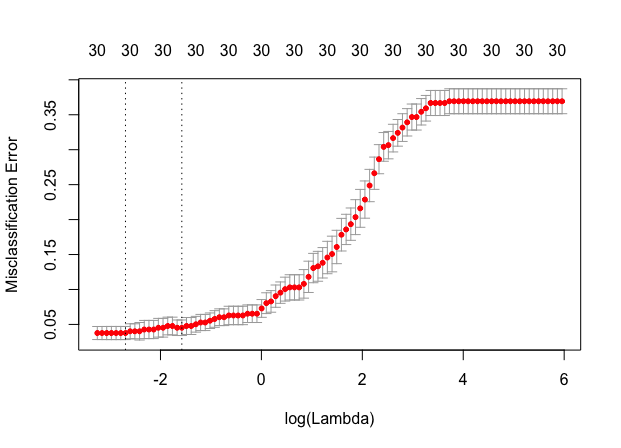
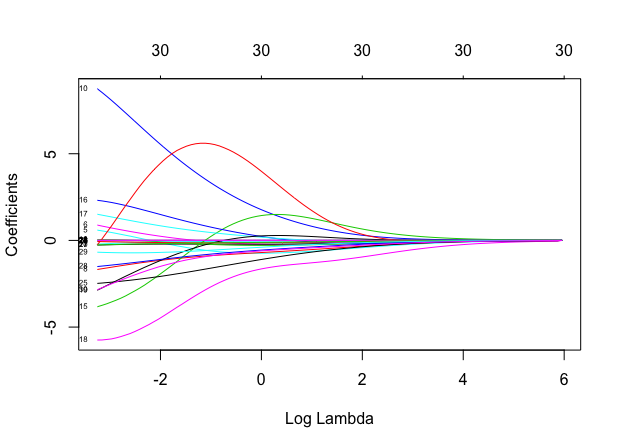
We split the data into 70% training and 30% test and apply data into Tikhonov regularization (Ridge regression) and least absolute shrinkage and selection operator (LASSO) to evaluate the Breast Cancer Wisconsin. The Ridge is a regression method that perform L2 regularization. The Lasso is a regression method that perform L1 regularization and variable selection when the data has huge number of features. After find the best lambda.min and lambda.1se by using 10-fold cross-validation to fit the model, we apply 30% test into each Ridge and Lasso model to predict the mass and find out the best accuracy model.

For the Ridge regression result. The plot coefficients with log lambda (Wilson figures.1) shows the result that when we increase Log Lambda more and more, almost all the variables shrink into coefficients close to zero, but never drop off from the model. Next, to find the best lambda for our Ridge Model, the plot misclassification error with log lambda (Wilson figures.2) indicate that when log lambda around -2 which are the vertical dotted line interval, the model has =low misclassification error. One is Lambda.min 0.0677 which is the value that gives a minimum mean error, and the other is Lambda.1se: 0.207 which is the value that gives one standard error of the minimum. Between Lambda.min and Lambda.1se have amount difference. We apply both Lambda.min and Lambda.1se into our model. Both Lambda model keeps all the variables and shows the same accuracy. After we apply this model to our testing set, this mode indicates that prediction accuracy for the mass is 94.73%, Moreover, the false negatives for the prediction 8.49%.

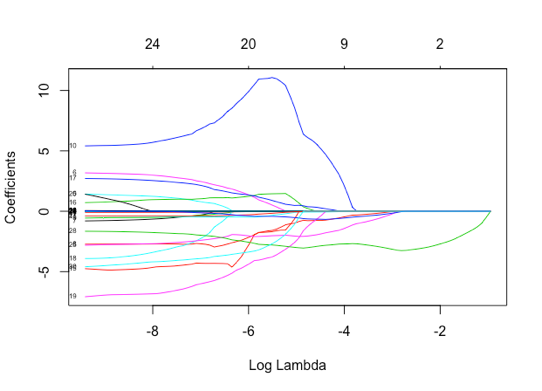
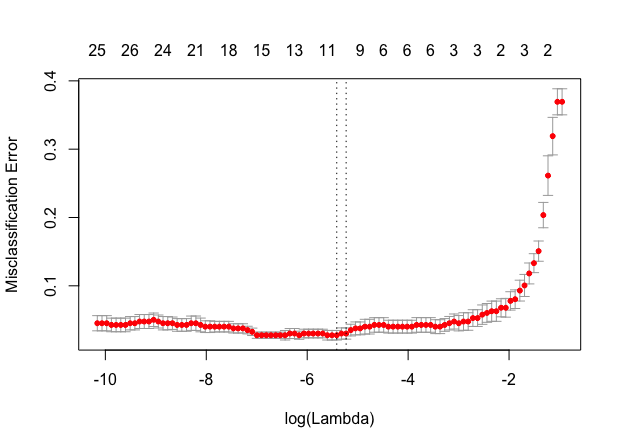


Wilson fig.1. Lambda coefficients plot Wilson fig.2. Lambda misclassification error plot

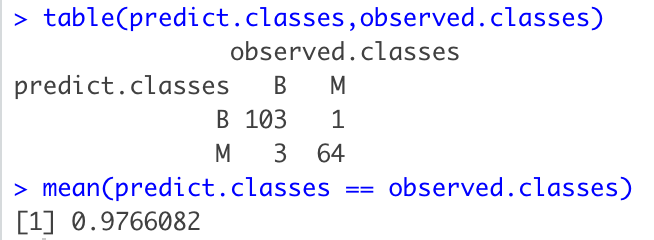
For the Lasso regression result. The plot coefficients with log lambda (Wilson figures.3) shows the result that when we increase Log Lambda more and more, more variables shrink into coefficients to zero. Moreover, we found that when the log lambda is -2, the only concave.points\_worst variable still stay significant. To find the best lambda for our Lasso Model, the plot misclassification error with log lambda (Wilson figures.4) indicate that when log lambda around -5.3 which are the vertical dotted line interval, the model has good prediction and acceptable numbers of variables. One is Lambda.min 0.0045 which is the value that gives a minimum mean error, and the other is Lambda.1se: 0.0054 which is the value that gives one standard error of the minimum. We apply both Lambda.min and Lambda.1se into our model. Lambda.1se indicate less dimension and better accuracy, so we pick Lambda.1se model for our model. Lasso Lambda.1se shrink all variables to 10 significant variables. fractal dimension in se, smoothness in worst, and concave.points in worst have high coefficient values for the models.

Lasso Model:

After we apply this model to our testing set, this mode indicates that prediction accuracy for the mass is 97.66%, Moreover, the false negatives for the prediction 0.94%. Finally, we compare Ridge regression and Lasso regression. We maintain that Lasso model for breast cancer has higher prediction accuracy and not overfitting model. Furthermore, Lasso model has much lower false negatives prediction. We think is important for the not predict malignant to benign.

Wilson fig.3. Lambda coefficients plot Wilson fig.4. Lambda misclassification error plot

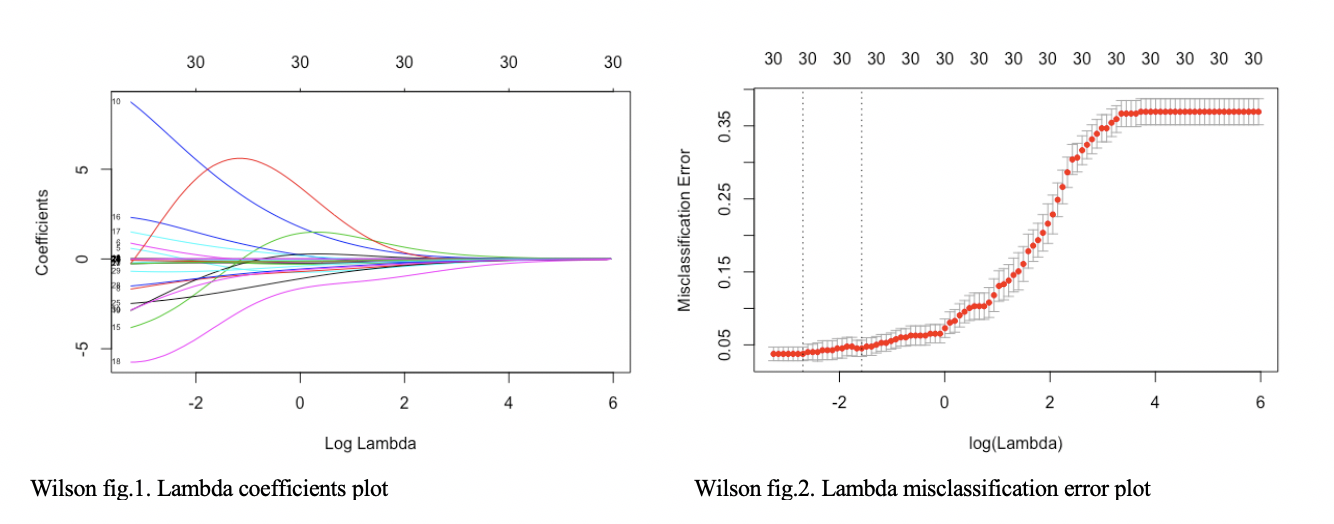


**Wilson Regularization Method**

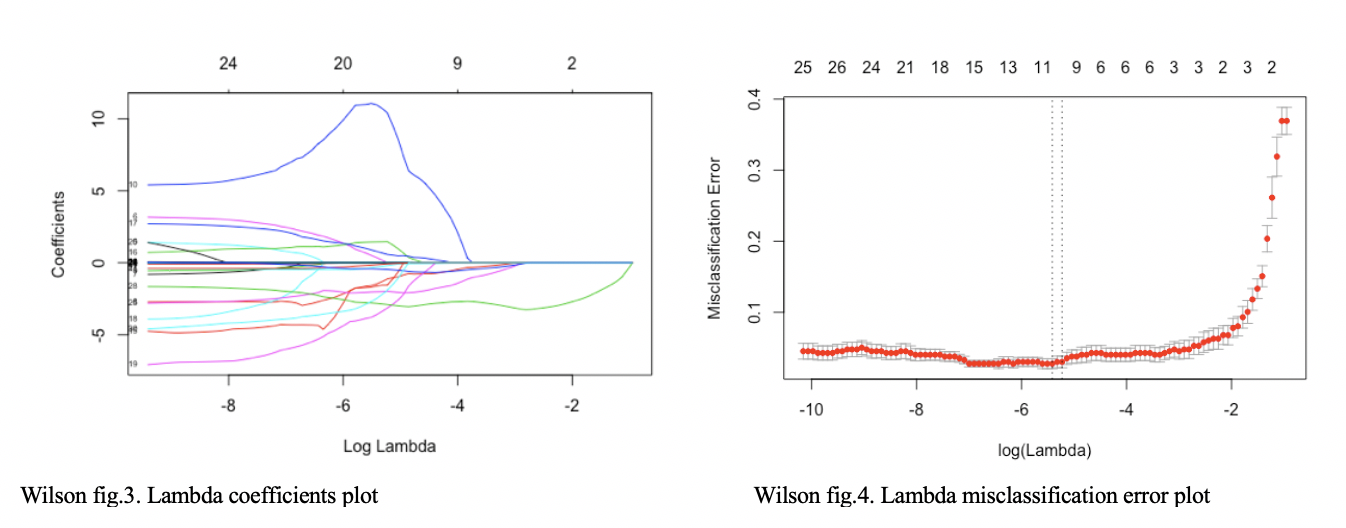
We split the data into 70% training and 30% test and apply data into Tikhonov regularization (Ridge regression) and least absolute shrinkage and selection operator (LASSO) to evaluate the Breast Cancer Wisconsin. The Ridge is a regression method that performs L2 regularization. The Lasso is a regression method that performs L1 regularization and variable selection when the data has a huge number of features. After finding the best lambda.min and lambda.1se by using 10-fold cross-validation to fit the model, we apply 30% test into each Ridge and Lasso model to predict the mass and find out the best accuracy model.

**Wilson Regularization**

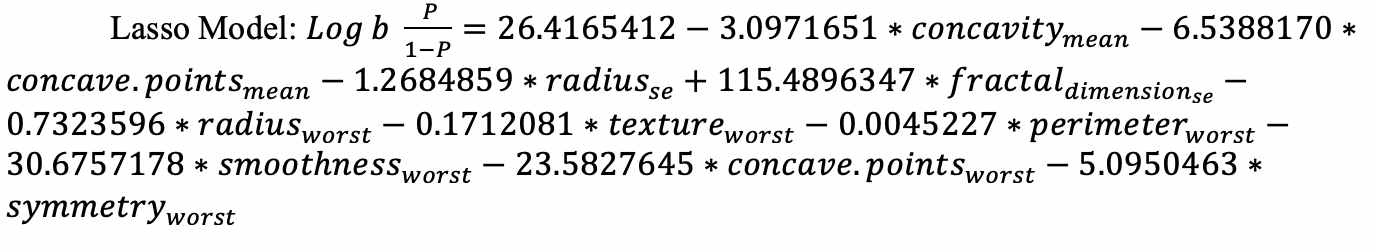
For the Ridge regression result. The plot coefficients with log lambda (Wilson figures.1) shows the result that when we increase Log Lambda more and more, almost all the variables shrink into coefficients close to zero, but never drop off from the model. Next, to find the best lambda for our Ridge Model, the plot misclassification error with log lambda (Wilson figures.2) indicate that when log lambda around -2 which are the vertical dotted line interval, the model has =low misclassification error. One is Lambda.min 0.0677 which is the value that gives a minimum mean error, and the other is Lambda.1se: 0.207 which is the value that gives one standard error of the minimum. Between Lambda.min and Lambda.1se have amount difference. We apply both Lambda.min and Lambda.1se into our model. Both Lambda model keeps all the variables and shows the same accuracy. After we apply this model to our testing set, this mode indicates that prediction accuracy for the mass is 94.73%, Moreover, the false negatives for the prediction 8.49%.



For the Lasso regression result. The plot coefficients with log lambda (Wilson figures.3) shows the result that when we increase Log Lambda more and more, more variables shrink into coefficients to zero. Moreover, we found that when the log lambda is -2, the only concave.points\_worst variable still stay significant. To find the best lambda for our Lasso Model, the plot misclassification error with log lambda (Wilson figures.4) indicate that when log lambda around -5.3 which are the vertical dotted line interval, the model has good prediction and acceptable numbers of variables. One is Lambda.min 0.0045 which is the value that gives a minimum mean error, and the other is Lambda.1se: 0.0054 which is the value that gives one standard error of the minimum.



We apply both Lambda.min and Lambda.1se into our model. Lambda.1se indicate less dimension and better accuracy, so we pick Lambda.1se model for our model. Lasso Lambda.1se shrink all variables to 10 significant variables. fractal dimension in se, smoothness in worst, and concave.points in worst have high coefficient values for the models.



After we apply this model to our testing set, this mode indicates that prediction accuracy for the mass is 97.66%, Moreover, the false negatives for the prediction 0.94%. Finally, we compare Ridge regression and Lasso regression. We maintain that Lasso model for  breast cancer has higher prediction accuracy and not overfitting model. Furthermore, Lasso model has much lower false negatives prediction. We think is important for the not predict malignant to benign.

In our Breast Cancer Wisconsin data set, diagnosis variable, 357 observation in benign and 212 observation in malignant, has higher frequency of Benign cases, therefore dataset might be biased for prediction analysis. Moreover, Variances are not equally distributed between the two cases, which also may lead to some discrepancies .Finally, small dataset which only 569 observation, so we run a risk of the small sample being unusual just by chance