**Statistical Learning and Analysis**

**Instructor: Prof. Varun Rai**

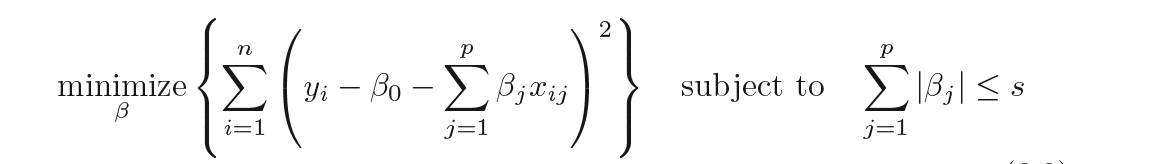
**Assignment 4**

**Question 1:**

Read and summary the subsection on "Another Formulation for Ridge Regression and the Lasso" (pp 220-222).

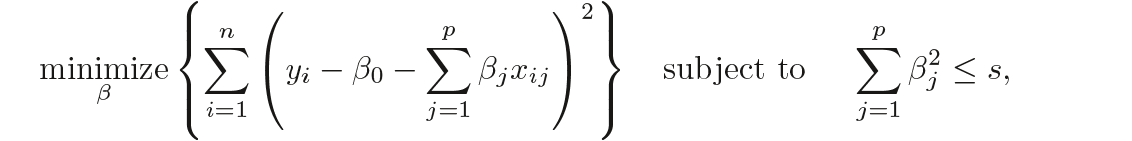
**Another Formulation for Ridge Regression and the Lasso:**

The given equation of Lasso defines how to find coefficients with the least RSS subject to being restricted in a budget ‘s’. So, when value of ‘s’ is large eventually budget is not restrictive and coefficient estimates can be large. Also, when ‘s’ is large, chances of falling least square solution within the budget are high. Find the below equation yields the least square solution,

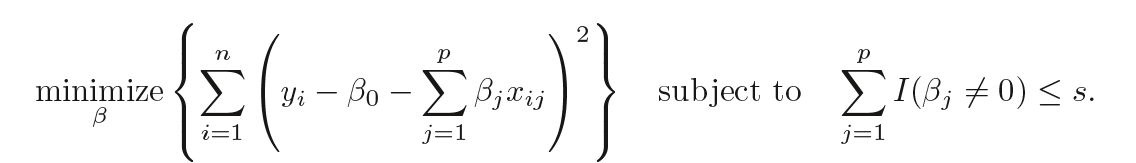


Similarly, if “s” is small then the budget is extremely restrictive, and the result must be small to not violate the budget. By changing values of λ, some values of s we can coefficient estimates for lasso estimates and for some values of λ and s we would get the ridge coefficient estimates. p=2 corresponds to lowest RSS for lasso model |β1| + |β2| ≤ s and for ridge model lowest RSS corresponds to the points which lie within the circle. objective of performing lasso is to find the set of coefficients that have lowest RSS keeping in mind the constraint of budget that to determine how large can summation of βj be. For Ridge, the exact same principle is followed as defined for Lasso above except 

Term should be within the budget. The equation used for Ridge is as follows:



To find the best subset of coefficients to minimize RSS we use:



Here, I(βj != 0) is the indicator variable. The indicator variable takes the value 1 when βj is not equal to zero and zero elsewise. Unfortunately, solving the above is computationally infeasible when p is large, since it requires considering all models containing s predictors. Therefore, we can conclude Lasso and Ridge are the computationally feasible alternatives to this best subset selection.