625.661 Statistical Models and Regression

Module 10 Discussion Question

H.M. James Hung

In building a regression model for the response variable y with a set of k regressors x_i, i = 1, …, k, let $r_{y,xi}$ denote the observed correlation coefficient between y and x_i, i = 1, …, k. Can the values of $r_{y,xi}$ suggest what subset of the regressors is likely to be selected when any of the subset selection techniques are applied to a set of orthogonal regressors? If so, explain how. [Note: the regressors x_i and x_j are orthogonal if $\sum_{h=1}^{n}(x_{ih}-\bar{x}_i)(x_{jh}-\bar{x}_j)=0$, where n is number of observations].