Financial Econometrics
MT Week 8 Assignment

December 2020

Note: Please complete and submit this assignment by **14:00 Monday 9th week**

1. Suppose you have 5 X variables $(X_1, \ldots, X_5)$ and you produce the in-sample SSE values and cross-validated SSE_{xv} in the table below.

   (a) Which model is selected using Best Subset Regression? Explain your reasoning.

   (b) Which model is selected using Forward Stepwise Regression? Explain your reasoning.

Reg	SSE	SSE_{xv}	Reg	SSE	SSE_{xv}
X_1	7.21	11.94	X_1, X_2, X_3	5.10	11.57
X_2	8.15	10.55	X_1, X_2, X_4	6.24	13.78
X_3	10.57	15.40	X_1, X_2, X_5	6.68	15.83
X_4	10.37	15.10	X_1, X_3, X_4	6.47	14.24
X_5	7.31	13.96	X_1, X_3, X_5	4.57	12.54
X_1, X_2	6.96	12.67	X_1, X_4, X_5	5.86	14.79
X_1, X_3	6.54	11.70	X_2, X_3, X_4	7.46	16.19
X_1, X_4	6.80	12.77	X_2, X_3, X_5	4.54	12.37
X_1, X_5	6.72	14.74	X_2, X_4, X_5	6.60	15.75
X_2, X_3	7.52	12.72	X_3, X_4, X_5	6.00	15.42
X_2, X_4	8.10	14.00	X_1, X_2, X_3, X_4	4.99	14.32
X_2, X_5	7.10	14.48	X_1, X_2, X_3, X_5	3.60	13.04
X_3, X_4	10.30	18.15	X_1, X_2, X_4, X_5	5.68	16.13
X_3, X_5	6.00	13.04	X_1, X_3, X_4, X_5	4.44	14.61
X_4, X_5	7.04	14.94	X_2, X_3, X_4, X_5	4.53	15.08
			X_1, X_2, X_3, X_4, X_5	3.44	15.26

2. The OLS $\hat{\beta}$ is defined $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Suppose

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

   where ρ is a value between -1 and 1. Further, suppose

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

   which are like covariances.

(a) When $\rho = 0$, how do the estimated β coefficients depend on c_1 and c_2?

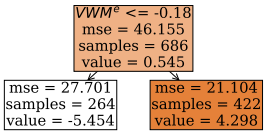
(b) What happens when $\rho \neq 0$? Hint: You need to use the expression for the inverse of a 2 by 2 matrix, and then multiply this by $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$.

(c) Now suppose you estimate a Ridge regression so that

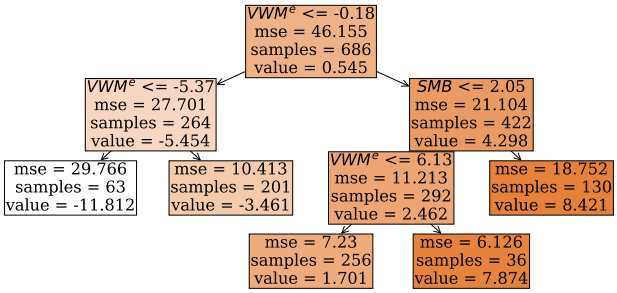
$$\mathbf{X}'\mathbf{X} + \lambda \mathbf{I}_2 = \begin{bmatrix} 1+\lambda & \rho \\ \rho & 1+\lambda \end{bmatrix}.$$

How does $\lambda > 0$ affect the way in which the covariance terms c_1 and c_2 are shared across the two $\hat{\beta}$ estimates?

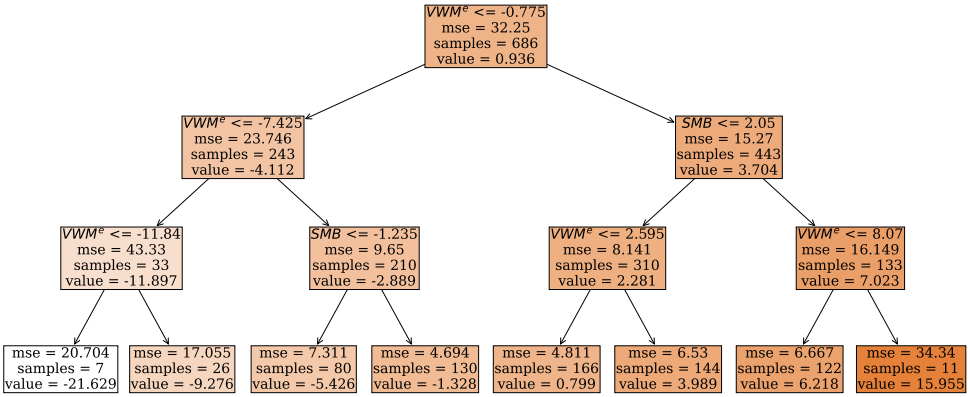
3. Describe the regression model and the estimated parameters for the three trees visualized below. The regressors are the Value-Weighted Market, and the Size, Value and Momentum factors.



Tree A



Tree B



Tree C