

Homework 07
Joseph Blubaugh
jblubau1@tamu.edu
STAT 608-720

1.

- a) X1 and X2 should have a correlation of 1 because they are the same minus a factor of 2.2. Adding X2 to the model will not improve the model so the output of an added variable plot should randomness with a flat trendline. The coefficient alpha should be very close to 0.
- b) Since X1 is so highly correlated with Y, there will be virtually no scatter around the regression line. The added variable plot of X2 should look like a perfectly flat line with not scatter.
- c) It looks like the fitted vs. standardized residuals plot. Since X1 and X2 are independent they are not correlated and sot eh added variable plots should not be flat.

2.

- a) The variance is considered large when it is greater than 5. When $p = .905 \frac{.905}{(1-.905^2)} = 5.000$.
- b)

$$\begin{aligned} s &= \sqrt{\frac{S_{xx}}{n-1}} \\ &= \sqrt{\frac{1}{1}} \\ &= 1 \end{aligned}$$

$$\begin{aligned} Var(\hat{\beta}_j) &= \frac{1}{1-r} \frac{\sigma^2}{(n-1)S_j^2} \\ &= \sigma^2 \end{aligned}$$

3.

- a) The VIF for x1 and x2 are both 1 because there is no corelation between x1 and x2. The correlation matrix is orthogonal and does not depend on y.
- b) The VIF for x1 = 1.375 which is higher than in a) because there is some correlation between x1 and x2
- c) If all variables are completely independent and have no correlation with one another, then the VIF would be 1. Any amount of correlation will make the VIF greater than 1.