ISLR 2.4 Exercises

1. For each of parts (a) through (d), indicate whether we would generally expect the performance of a flexible statistical learning method to be better or worse than an inflexible method. Justify your answer.

(a)The sample size *n* is extremely large, and the number of predictors *p* is small.

Yes, a flexible statistical learning method would perform better because it can get more information from the large sample. The large sample also reduces the risk of overfitting problem.

(b)The number of predictors *p* is extremely large, and the number of observations *n* is small.

In this scenario, the overfitting is often the concern. The learning results from multiple predictors and small sample samples may be not right. The inflexible method would be appropriate way to fit a small sample size.

(c)The relationship between the predictors and response is highly non-linear.

When the response and predictors show some non-linearized relationship, the inflexible methods will be not appropriate, so we should use a flexible method.

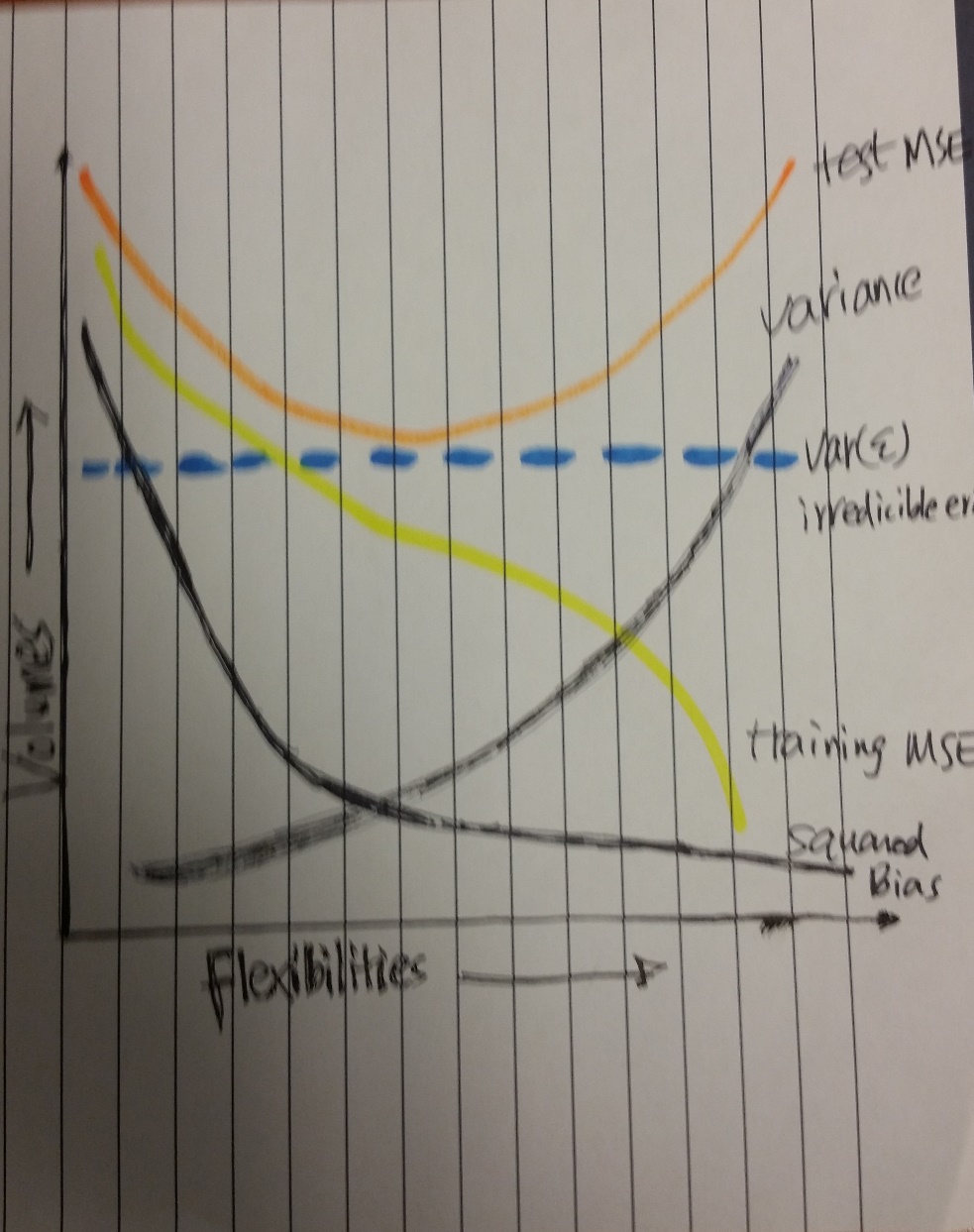
(d) The variance of the error terms, i.e. σ2 = Var(\_), is extremely high.

The high variance of error indicates the samples are highly heterogenetic. So we should consider use an inflexible method.

3. We now revisit the bias-variance decomposition.

(a) Provide a sketch of typical (squared) bias, variance, training error, test error, and Bayes (or irreducible) error curves, on a single plot, as we go from less flexible statistical learning methods

towards more flexible approaches.



(b) Explain why each of the five curves has the shape displayed in

part (a).

In my sketch, the training MSE-mean squared error (yellow line) usually declines as flexibility increases. As flexibility increases the fitted curve fits the observed data more closely.

The test mean squared error-SE (orange line) usually declines at the beginning as flexibility increases but at some point it levels off and then starts to increase again (U-shape). As when a fitted curve yields a small training MSE but a large test MSE we are actually overfitting the data.

The squared bias (dark line) decreases monotonically and the variance (dark line) increases monotonically. It is a general rule that the more flexible methods will increase the variance but decrease the bias.

The irreducible error (Var(ε)) (blue line) is a constant so it is a parallel line, because it come from the specific dataset and does not alter by modelling.

8. See markdown file.