Problem 1:

(a) better (b) worse (c) better (d) worse

Problem 2:

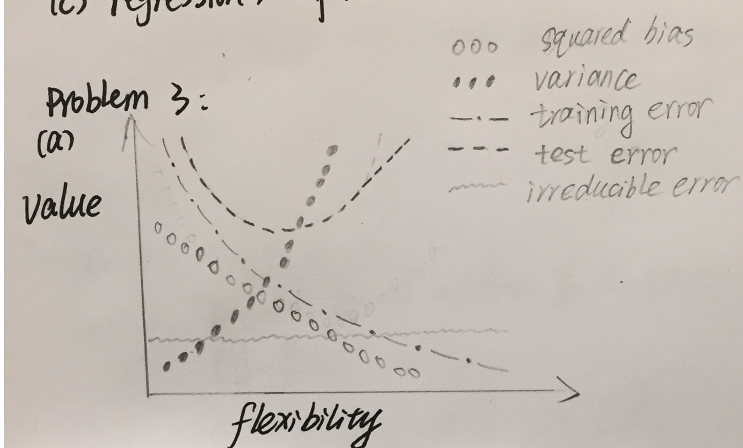
(a) regression, inference, n=500, p=3.

(b) classification, prediction, n=20, p=13

(c) regression, inference, n=52, p=3

Problem 3:

(a)



(b) squared bias: This will keep decreasing because as the model becomes more flexible, it can fit the training data better, hence lower squared bias.

variance: With more flexibility, the model tends to be more unstable, hence a higher variance.

training error: With more flexibility, the model can always fit the training data better, eventually leads to even 0 traing error.

test error: Firstly, with more flexibility, the model can get more trend of the data and such trend generalizes well to the test data, hence a lower test error. But later on with even more flexibility, the model is fitting noise in the training data and such noise does not generalize well to the test data, therefore the test error goes up again.

irreducible error: The is the part of error that can never be reduced and it will remain constant as the variance of ε.

Problem 4:

(a) No. Because we do not know y_0. So there's no way for us to compute $(y_0 - \hat{f}(x_0))^2$

(b) No. bias $= (E\hat{f}(x_0) - f(x_0))^2$. Though we can estimate $E\hat{f}(x_0)$, we don't know f, therefore we can not estimate bias.

(C). Yes. Variance $= E(\hat{f}(x_0) - E\hat{f}(x_0))^2$. we can sample multiple time to get multiple $\hat{f}$, then we average all $\hat{f}(x_0)$ to estimate $E\hat{f}(x_0)$, then we calculate all of the $(\hat{f}(x_0) - E\hat{f}(x_0))^2$ (with $E\hat{f}(x_0)$ replaced by its estimator) and average them, this gives us a estimator of variance.

(d) No. Because MSE = bias + Variance + irreducible error. We can only estimate Variance. The bias and irreducible error sum up and we can not tell how much each part contribute to the total MSE (Not to mention that we cannot estimate MSE)