HW 1.

2nd Problem

(a). Consider each term in $E[\frac{1}{m}\Sigma(Y_i' - \hat{\beta}^T X_i')^2]$.

$g(x_j', y_j', \hat{\beta}) = (y_j' - \hat{\beta}^T x_j')^2$. $j = 1, 2, \cdots n$

Conditional mean of $g(x_j', y_j', \hat{\beta})$ on $\hat{\beta}$. since (x_j', y_j') are iid on $\hat{\beta}$ so the conditional distribution of (x_j', y_j') = the conditional dist. of (x_k', y_k') given $\hat{\beta}$.

So $E[g(X_j', Y_j', \hat{\beta}) \mid \hat{\beta} = \beta] = E[g(X_1', Y_1', \beta)] = E[g(X_2', Y_2', \beta)] = \cdots$

$E[(Y_j' - \hat{\beta}^T x_j')^2] = E[E[g(x_1', y_1', \beta)]]$.

So $E[\frac{1}{m}\sum_{i=1}^{m}(Y_i' - \hat{\beta}^T x_i')^2] = E[E[g(x_1', Y_1', \beta)]]$.

$\downarrow$ Iterative Expectation

So $E[\frac{1}{m}\sum_{i=1}^{m}(Y_i' - \hat{\beta}^T X_i')^2] = E[(y_j' - \hat{\beta}^T x_j)^2]$

And we can change m to N.

So $E[\frac{1}{m}\sum_{i=1}^{m}(Y_i' - \hat{\beta}^T X_i')^2] = E[\frac{1}{N}\sum_{i=1}^{N}(Y_i' - \hat{\beta}^T X_i')^2]$. □

(b)

Since $\tilde{\beta}$ is the coeff on test.

And since the training set and test set are iid,

A and B have the same distribution.

(c) $B = \frac{1}{n} \sum_{j=1}^{n} (Y_j' - \tilde{\beta}^T X_j')^2$

Since $\hat{\beta}$ minimizes $RSS = \sum_{i=1}^{n} (Y_i - \beta^T X_i)^2$

So similarly, $\tilde{\beta}$ minimize $\sum_{i=1}^{n} (Y_i' - \beta^T X_i')^2$.

So $\frac{1}{n} \sum_{i=1}^{n} (Y_i' - \tilde{\beta}^T X_i')^2 \le \frac{1}{n} \sum_{i=1}^{n} (Y_i' - \hat{\beta}^T X_i')^2$.

(d) $E\left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{\beta}^T X_i)^2\right] = E(A) = E(B) = E\left[\frac{1}{n} \sum_{i=1}^{n} (Y_i' - \tilde{\beta}^T X_i')^2\right]$

$\le E\left[\frac{1}{n} \sum_{i=1}^{n} (Y_i' - \hat{\beta}^T X_i')^2\right]$.