Week 4

3.1) $E\{[\hat{y}-f(x')]^2\}$

$= E\{[\hat{y}-f(x')+E[\hat{y}]-E[\hat{y}]]^2\}$

$= E\{[(\hat{y}-E[\hat{y}])-(f(x')-E[\hat{y}])]^2\}$

$= E\{(\hat{y}-E[\hat{y}])^2-2(\hat{y}-E[\hat{y}])(f(x')-E[\hat{y}])+(f(x')-E[\hat{y}])^2\}$

$= E\{(\hat{y}-E[\hat{y}])^2\}-2\left(E[\hat{y}]f(x')-\cancel{E[\hat{y}]E[\hat{y}]}-\cancel{E[\hat{y}]f(x')}+\cancel{E[\hat{y}]}E[\hat{y}]\right)$

$+ \cancel{E}\{(f(x')-E(\hat{y}))^2\}$

$= (E[\hat{y}]-f(x'))^2 - var(\hat{y})$ $\square$

feb. 11-08:08

3.3) $\hat{\beta}$: from training

$\tilde{\beta}$: from test set.

$E\left[\frac{1}{n}\sum_{i=1}(y_i-\hat{\beta}^T x_i)^2\right] = E\left[\frac{1}{n}\sum_{i=1}^{n}(y_{n+i}-\tilde{\beta}^T x_{n+i})^2\right]$

$\boxed{\frac{1}{n}\sum_{i=1}^{n}(y_{n+i}-\tilde{\beta}^T x_{n+i})^2 \le \frac{1}{n}\sum_{i=1}^{n}(y_{n+i}-\hat{\beta}^T x_{n+i})^2}$

$\le E\left[\frac{1}{n}\sum_{i=1}^{m}(y_{n+i}-\hat{\beta}^T x_{n+i})^2\right] = E\left[\frac{1}{n}\sum_{i=1}^{m}(y_{n+i}-\hat{\beta}^T x_{n+i})^2\right]$ $\square$

feb. 11-08:29

3.4) $\hat{y}_{-i} = x_i^T \hat{\beta}_{-i}$, $P = X(X^TX)X^T$

$\hat{\beta}_{-i} = (X_{-i}^T X_{-i})^{-1} X_{-i}^T y_{-i}$

$X_{-i}^T X_{-i} = X^TX - x_i x_i^T$, $X_{-i}^T y_{-i} = X^T y - x_i y_i$

From Sherman-Morrison

$(X_{-i}^T X_{-i})^{-1} = (X^TX)^{-1} + \frac{(X^TX)^{-1} x_i x_i^T (X^TX)^{-1}}{1-\underbrace{x_i^T (X^TX)^{-1} x_i}_{P_{ii}}}$

$\hat{\beta}_{-i} = (X_{-i}^T X_{-i})^{-1}(X^T y - x_i y_i)$

$= \underbrace{(X^TX)^{-1}X^Ty}_{\hat{\beta}} - (X^TX)^{-1}x_i y_i + \frac{(X^TX)^{-1}x_i}{1-P_{ii}}\left(\underbrace{x_i^T(X^TX)^{-1}X^Ty}_{\hat{\beta}} - \underbrace{x_i^T(X^TX)^{-1}x_i}_{P_{ii}} y_i\right)$

$= \hat{\beta} - \frac{(X^TX)^{-1}x_i}{1-P_{ii}}\left[(1-P_{ii})y_i - \underbrace{x_i^T\hat{\beta}}_{\hat{y}_i} + P_{ii} y_i\right]$

$= \hat{\beta} - \frac{(X^TX)^{-1}x_i}{1-P_{ii}}(y_i-\hat{y}_i)$

feb. 11-08:39

$y_i - \hat{y}_{-i} = y_i - x_i^T\left(\hat{\beta} - \frac{(X^TX)^{-1}x_i}{1-P_{ii}}[y_i-\hat{y}_i]\right)$

$= y_i - \hat{y}_i + \frac{P_{ii}}{1-P_{ii}}(y_i-\hat{y}_i)$

$= (y_i-\hat{y}_i)/(1-P_{ii})$ $\square$

feb. 11-08:49