STAT5703 HW3 Exercise 4

Wen Fan(wf2255), Banruo Xie(bx2168), Hanjun Li(hl3339)

Exercise 4.

Question 1.

$$var(\hat{\epsilon}_i) = var[Y_i - \hat{Y}_i] = E[(Y_i - \hat{Y}_i)^2] - E([Y_i - \hat{Y}_i])^2$$
$$var(\delta_i) = var[\hat{Y}_i - E[Y_i]] = E[(\hat{Y}_i - E[Y_i])^2] - E([\hat{Y}_i - E[Y_i]])^2$$
$$var(\delta_i) - var(\hat{\epsilon}_i) = E[(\hat{Y}_i - (E[Y_i])^2] - E[(Y_i - \hat{Y}_i)^2] = E[(\hat{Y}_i - E[Y_i])^2] - E[RSS(\hat{Y}_i)]$$
$$E[RSS(\hat{Y}_i] + \sum var(\delta_i) - \sum var(\hat{\epsilon}_i) = \sum E[(\hat{Y}_i - (E[Y_i])^2] = E[\sum(\hat{Y}_i - E[Y_i])^2] = \sigma^2\Gamma$$

Question 2.

to prove $E[\frac{1}{\sigma^2}RSS(\hat{Y}) + 2tr(S) - n] = \Gamma$, we only need to prove $\sigma^2 E[RSS(\hat{Y}) + 2tr(S) - n] = \sigma^2\Gamma$, $\sigma^2(2tr(S) - n) = \sum var(\delta_i) - \sum var(\hat{\epsilon}_i)$
$= E[(\delta - \delta_\mu)^T(\delta - \delta_\mu)] - E[(\hat{\epsilon} - \hat{\epsilon}_\mu)^T(\hat{\epsilon} - \hat{\epsilon}_\mu)]\backslash$

$$\hat{\epsilon} = Y - \hat{Y} = (I - S)Y$$
$$\hat{\epsilon}_\mu = (I - S)E[Y] = (I - S)\mu$$
$$\delta = \hat{Y} - E[Y] = \hat{Y} - \mu = SY - \mu$$
$$\delta_\mu = SEY - \mu = (S - I)\mu$$

Therefore,

$$\hat{\epsilon} - \hat{\epsilon}_\mu = (I - S)Y - (I - S)\mu = (I - S)(Y - \mu)$$
$$\delta - \delta_\mu = S(Y - \mu)$$

Now, we can go back to $E[(\delta - \delta_\mu)^T(\delta - \delta_\mu)] - E[(\hat{\epsilon} - \hat{\epsilon}_\mu)^T(\hat{\epsilon} - \hat{\epsilon}_\mu)]$
$= E[(Y - \mu)^T S^T S(Y - \mu)] - E[(Y - \mu)^T(I - S)^T(I - S)(Y - \mu)]$
$= E[(Y - \mu)^T(-I + S + S^T)(Y - \mu)]$
$= tr[(-I + S + S^T)cov(Y - \mu)] + (E[Y - \mu])^T(-I + S + S^T)E[Y - \mu]$
$= \sigma^2 tr[-I + S + S^T] = \sigma^2(-n + 2tr[s])$ proved.

Question 3.

$re(S) = tr(X^T X(X^T X)^{-1}) = tr(I_P) = p$ therefore $C_p = \frac{1}{\sigma^2}RSS(\hat{Y}) + 2p - n$ Note that AIC = 2p-2l and $l = -\frac{1}{2}(n\log\sigma^2 + \frac{1}{\sigma^2}RSS(\hat{Y}) + C$ So $AIC = \frac{1}{\sigma^2}RSS(\hat{Y}) + 2p - n\log\sigma^2 + C$ is similar to C_P if we treat σ^2, n to constant.

Question 4.

We have $AIC(\hat{\beta}_{q+1}) - AIC(\hat{\beta}) = 2 - 2(l_{q+1} - l_q)$ where $l_q = -\frac{1}{2}(n\log(RSS(\hat{\beta}_q) + n - nlogn)$ Therefore $P(AIC(\hat{\beta}_{q+1}) - AIC(\hat{\beta})) < 0)$
$= P(2 + nlog\frac{RSS(\hat{\beta}_{q+1})}{RSS(\hat{\beta}_q)} < 0)$
$= P(nlog(1 - X_1^2/n) < -2)$ where n$\rightarrow \infty$ $P(nlog(1 - X_1^2/n) < -2) = P(X_1^2 > 2) > 0$

Question 5.

$BIC = -2l(\hat{\theta}) + plogn$, So $P(BIC(\hat{\beta}_{q+1}) - BIC(\hat{\beta}_q) < 0) =$
$P(logn + nlog\frac{RSS(\hat{\beta}_{q+1})}{RSS(\hat{\beta}_q)} < 0)$ where n$\rightarrow \infty$ $P(logn + nlog\frac{RSS(\hat{\beta}_{q+1})}{RSS(\hat{\beta}_q)} < 0) = P(logn - X_1^2 < 0)$
$P(X_1^2 > logn) = 0$