STATS 305A H.W #3

Adhitya Venkatesh

1. $Y = X\beta + \varepsilon$, $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2 I$

$$\theta_{OLS} = c^T \beta = c^T (X^T X)^{-1} X^T y$$

~~Suppose~~ Let another unbiased est. be

$$\hat{\theta} = a^T y$$

Now let $D^T = a^T - c^T (X^T X)^{-1} X^T$

Then, $E[a^T y] = E\left(c^T (X^T X)^{-1} X^T y + D^T y\right)$

But as $a^T y$ unbiased

$$\Rightarrow E[a^T y] = E[c^T \hat{\beta}] = E(c^T \hat{\beta}) + E(D^T y)$$

$$\left(E(D^T y)\right) = \left(c^T \beta + D^T X \beta\right)$$

$$\Rightarrow D^T X = 0 \Rightarrow X^T D = 0$$

$$\therefore Var(\hat{\theta}) = a^T Var(y)\, a = \sigma^2 a^T a$$

$= \sigma^2 \cancel{(C^T B}$

$\sigma^2 (C^T (X^T X)^{-1} X^T + D^T)(C^T (X^T X)^{-1} X^T + D^T)^T$

$$= \sigma^2 \Big(C^T (X^T X)^{-1} X^T X (X^T X)^{-1} C + C^T (X^T X)^{-1} X^T D$$
$$+ D^T X (X^T X)^{-1} C + D^T D \Big)$$

As $X^T D = D^T X = 0$,

$$= \sigma^2 (C^T (X^T X)^{-1} C + D^T D)$$

$$= \sigma^2 (Var(\theta_{LS}) + M) \quad (\because M \geq 0)$$

$$\leq Var(\theta_{LS})$$

$\therefore Var(\theta_{LS}) \leq Var(\hat{\theta})$ $\forall$ unbiased estim. $\hat{\theta}$.

4. a) We proceed to show independence via mathematical induction:

<u>Base case</u>: $(n=2)$ $\left(V_n = \frac{(n-1)S_n^2}{\sigma^2}\right)$

$$\bar{X}_2 = \frac{X_1 + X_2}{2}, \quad V_2 = \frac{1}{\sigma^2}\left\{(X_1 - \bar{X}_2)^2 + (X_2 - \bar{X}_2)^2\right\}$$

$$\Rightarrow \quad X_1 = \bar{X}_2 \pm \frac{\sigma}{\sqrt{2}}\sqrt{V_2} = \bar{X}_2 + \frac{\sigma}{\sqrt{2}}\sqrt{V_2}$$

$$X_2 = \bar{X}_2 \mp \frac{\sigma}{\sqrt{2}}\sqrt{V_2} = \bar{X}_2 - \frac{\sigma}{\sqrt{2}}\sqrt{V_2}$$

$$\therefore \hat{f}(\bar{x}_2, v_2) = J_{(x_1, x_2) \to (\bar{x}_2, v_2)} \cdot f(x_1, x_2)$$

where $f, \hat{f}$ represent pdfs and J the Jacobian from change in variables.

$$= 2\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^2 \exp\left[-\left(\frac{1}{2\sigma^2}\right)\left\{(x_1-\mu)^2+(x_2-\mu)^2\right\}\right]\cdot$$

$$\left(\frac{\sigma}{\sqrt{2}}\cdot\frac{1}{\sqrt{V_2}}\right)$$

$$= 2\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^2 \exp\left[-\frac{1}{2\sigma^2}\left(\sigma^2 V_2 + 2(\bar{x}_2-\mu)^2\right)\right]\cdot$$

$$\left(\frac{\sigma}{\sqrt{2V_2}}\right)$$

$$= \frac{1}{\sqrt{2\pi}(\sigma/\sqrt{2})}\exp\left(\frac{(\bar{x}_2-\mu)^2}{2\left(\frac{\sigma^2}{2}\right)}\right)\cdot\frac{1}{2^{1/2}\Gamma(1/2)}V_2^{-\frac{1}{2}}\exp\left[-\frac{V_2}{2}\right]$$

$$= g(\bar{x}_2 \mid \mu,\sigma^2)\cdot h(V_2 \mid \mu,\sigma^2)$$

$\Rightarrow$ $\bar{X}_2$ and V_2 independent

and $\bar{X}_2 \sim N(\mu, \sigma^2/2)$

$V_2 \sim \chi^2_1$

<u>General Case</u>: Now we assume the inductive hypothesis is true for $k=n$ (i.e. $\bar{X}_n$ and V_n indep. and dist. accordingly) and proceed to show it must be true for $k=n+1$:

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad \hat{V}_n \sim \chi^2_{n-1}$$

Using the fact that $\bar{X}_n$, $\hat{V}_n$, and X_{n+1} are mutually independent. Defining $Z_{n+1} = X_{n+1}$, we transform

$$(\bar{X}_n, \hat{V}_n, X_{n+1}) \to (\bar{X}_{n+1}, \hat{V}_{n+1}, Z_{n+1})$$

via Jacobian. Upon performing the calculation, the result conforms the inductive hypothesis.

b.)

From the previous part,

$$\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$S_n^2 \sim \frac{\sigma^2 \chi^2_{n-1}}{n-1}$$

$$\Rightarrow Var(S_n^2) = \frac{\sigma^4}{(n-1)^2} Var(\chi^2_{n-1})$$

$$= \boxed{\frac{2\sigma^4}{n-1}}$$

c) For confidence-level $= \alpha$, we have interval

$$\left(\frac{(n-1)S_n^2}{(\chi^2_{n-1})^{-1}_{\frac{1-\alpha}{2}}}, \frac{(n-1)S_n^2}{(\chi^2_{n-1})^{-1}_{\frac{1+\alpha}{2}}} \right)$$

where $(\chi^2_{n-1})^{-1}_{\beta} =$ inverse chi-sq with $(n-1)$ deg. of freedom, at β.

Computational problems

- SVD can filter out linear dependencies by revealing which components have negligible singular values Σ_i in

$$X = U \Sigma V^T = U \begin{bmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & \ddots & \\ & & & \Sigma_p \end{bmatrix} V^T$$

The values $\Sigma_i \approx 0$, $1 \leq i \leq p$ correspond to features with linear dependencies.

3. $$Y = \begin{cases} A_i = \theta + \varepsilon_i & i = 1, \dots, n \quad (n) \\ B_i = \theta - \phi + \varepsilon_i & i = n+1, \dots, n+m \quad (m) \\ \cancel{\quad} & (m) \\ C_i = \phi - \theta + \varepsilon_i & i = n+m+1, \dots, n+2m \end{cases}$$

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_{n+2m} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \\ -1 & 1 \\ \vdots & \vdots \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \theta \\ \phi \end{bmatrix} + \vec{\varepsilon}$$

$$X^T X = \begin{array}{c} \begin{array}{ccc} n & m & m \end{array} \\ \begin{bmatrix} (1 \dots 1) & (1 \dots 1) & (-1 \dots -1) \\ (0 \dots 0) & (-1 \dots -1) & (1 \dots 1) \end{bmatrix} \end{array} \begin{bmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \\ \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} & \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} \\ \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} & \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} n + 2m & 0 \\ 0 & 2m \end{bmatrix}$$

$$\begin{bmatrix} \hat{\beta} \end{bmatrix} = (X^T X)^{-1} X^T \begin{bmatrix} Y_1 \\ \vdots \\ Y_{2m+n} \end{bmatrix}$$

$$\begin{bmatrix} \hat{\theta} \\ \hat{\phi} \end{bmatrix} = \frac{1}{2m(n+2m)} \begin{bmatrix} 2m & 0 \\ 0 & n+2m \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & -1 \\ -1 & 1 \\ \vdots & \vdots \end{bmatrix}^T \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \\ Y_{n+1} \\ \vdots \\ Y_{2m+n} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{n+2m} & 0 \\ 0 & \frac{1}{2m} \end{bmatrix} \begin{bmatrix} \sum A_i + \sum B_i - \sum C_i \\ \sum C_i - \sum B_i \end{bmatrix}$$

$$= \begin{bmatrix} \left(\sum_{i=1}^{n+m} Y_i - \sum_{j=n+m+1}^{n+2m} Y_j \right) / n+2m \\ \left(\sum_{j=n+m+1}^{n+2m} Y_j - \sum_{k=n+1}^{n+m} Y_k \right) / 2m \end{bmatrix}$$

$$\hat{\theta} = \left(\sum_{i=1}^{n+m} Y_i - \sum_{j=n+m+1}^{n+2m} Y_j \right) \Big/ (2m+n)$$

$$\hat{\phi} \quad Var(\hat{\phi}) = Var\left(\sum_{i=n+1}^{n+m} Y_i \right) + Var\left(\sum_{j=n+m+1}^{n+2m} Y_j \right.$$

$$= \frac{2m\sigma^2}{4m^2} = \frac{\sigma^2}{2m}$$

$$Cov(\hat{\phi}, \hat{\theta}) = Cov\left(\frac{\sum A + \sum B - \sum C}{n+2m}, \frac{\sum C - \sum B}{2m} \right)$$

$$= \frac{1}{2m(n+2m)} \Big(Cov\left(\sum A, \sum C\right) - Cov\left(\sum A, \sum B\right) + 2Cov\left(\sum B, \sum C\right) - Var\left(\sum B\right) - Var\left(\sum C\right) \Big)$$

$$Var(\hat{\phi} - \hat{\theta}) = \cancel{Var\left(2\sum B + \sum A - \sum C\right.}$$

$\cancel{Var(\hat{\phi})}$

$$Var(\hat{\phi} - \hat{\theta}) = Var\left(\frac{\sum B - \sum C}{2m} - \frac{\sum A - \sum B}{m+n}\right)$$

$$= \frac{1}{4m^2}\left(Var(\sum B) + Var(\sum C)\right) + \frac{1}{(m+n)^2}\left(Var(\sum A) + Var(\sum B)\right)$$

$$= \frac{m\sigma^2}{2m^2} + \frac{\sigma^2}{m+n} = \frac{\sigma^2}{m} + \frac{\sigma^2}{m+n} = \boxed{\frac{2m+n}{m(m+n)}\sigma^2}$$

$$Var(\hat{\phi} - \hat{\theta}) = Var(\hat{\phi}) + Var(\hat{\theta}) - 2cov(\hat{\phi}, \hat{\theta})$$

$$= \frac{\sigma^2}{2m} + \frac{\sigma^2}{2m+n} - 2cov(\hat{\theta}, \hat{\phi})$$

$$\Rightarrow \boxed{cov(\hat{\theta}, \hat{\phi}) = \frac{\sigma^2(4m+n)}{4m(2m+n)} - \frac{Var(\hat{\phi} - \hat{\theta})}{2}}$$

$$\hat{\sigma}^2 = \frac{RSS}{n+2m-2} = \boxed{\frac{\sum_{i=1}^{n+2m}(Y_i - \hat{Y}_i)^2}{n+2m-2}}$$