Solution to Midterm

1.

(a). the characteristic roots is

$$\lambda_1 = 4 \quad \lambda_2 = 9$$

for $\lambda_1 = 4$, we denote the correspondent characteristic vector by $(\lambda_{11}, \lambda_{12})'$, after calculation, $\lambda_{11} = -2 * \lambda_{12}$, add standardization condition $\lambda_{11}^2 + \lambda_{12}^2 = 1$, then the characteristic vector is $(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})'$.

Similarly, for $\lambda_2 = 9$, the correspondent characteristic vector is $(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})'$

(b). The $(1-\alpha)\%$ joint confidence interval is given by

$$\beta \in R^2 \mid (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \leq F_{\alpha(2,n-2)} 2\sigma^2$$

where $\beta = (\beta_1, \beta_2)'$, $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)'$, this is an equation of an ellipse with the center at $(\hat{\beta}_1, \hat{\beta}_2)'$ and the ratio of the lengths of the axes of the ellipse is $\sqrt{\frac{\lambda_1}{\lambda_2}} = \frac{3}{2}$. The axes lengths are different but not too much.

2.

(a). We can write

$$(Y - X\beta) = (Y - X\hat{\beta}) + X(\hat{\beta} - \beta)$$

Hence

$$\begin{aligned} (Y - X\beta)'(Y - X\beta) &= [Y - X\hat{\beta}) + X(\hat{\beta} - \beta]'[Y - X\hat{\beta}) + X(\hat{\beta} - \beta] \\ &= (Y - X\hat{\beta})'(Y - X\hat{\beta}) + (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \\ &+ 2(Y - X\hat{\beta})'X(\hat{\beta} - \beta) \end{aligned}$$

Since $(Y - X\hat{\beta})'X = e'X = 0$, so the proof is done. And it is easy to see that $(\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \geq 0$, therefore (1) implies $(Y - X\beta)'(Y - X\beta) = \sum u_i^2$ has the least value at $\beta = \hat{\beta}_{OLS}$, which is essentially the least squares principle.

(b). For the model $y_i = \beta + u_i$,

$$X = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}' \quad \text{and} \quad Y = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}'$$

then

$$X'X = n, \quad \text{and} \quad \hat{\beta} = (X'X)^{-1}X'Y = \frac{\sum y_i}{n} = \bar{y}$$

therefore,

$$\sum (y_i - \beta)^2 - \sum (y_i - \bar{y})^2 = n(\bar{y} - \beta)^2$$

3.

(a). For our case,

$$\theta = \begin{bmatrix} u \\ \sigma^2 \end{bmatrix}, \qquad h(\theta) = u, \quad \text{and} \quad H(\theta) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$h(\hat{\theta}) = \hat{u} = \bar{y}, \qquad I^{-1}(\hat{\theta}) = \begin{bmatrix} \frac{n}{\hat{\sigma}^2} & 0 \\ 0 & \frac{n}{2\hat{\sigma}^4} \end{bmatrix}$$

where $\hat{\sigma}^2 = \frac{1}{n}\sum(y_i - \bar{y})^2$. Hence,

$$\begin{aligned} W &= h(\hat{\theta})'[H'(\hat{\theta})I(\hat{\theta})^{-1}H(\hat{\theta})]^{-1}h(\hat{\theta}) \\ &= \bar{y}[(1,0)\left[\begin{array}{cc} \frac{\hat{\sigma}^2}{n} & 0 \\ 0 & \frac{2\hat{\sigma}^4}{n} \end{array}\right](1,0)']^{-1}\bar{y} \\ &= \bar{y}^2 \cdot \frac{n}{\hat{\sigma}^2} = \frac{\bar{y}^2}{\hat{\sigma}^2/n} \end{aligned}$$

(b). We can write model (2) as

$$y_i = u + \epsilon_i$$

note that $Y_i = y_i, \quad x_i = 1, \quad \beta = u$ and ϵ_i has the following properties:

$$\epsilon_i \quad \text{follows} \quad IID \;\; N\,(0, \sigma^2) \quad \text{for} \quad i = 1, 2, \cdots, n$$

(c). $H_0: \;\; u = 0$, a special case of $H: \;\; R\beta = \delta$, with $R = 1, \quad \beta = u, \quad$ and $\quad \delta = 0$. We also have X'X=n, hence

$$W = \frac{(\bar{y} - 0)'[1 \cdot n^{-1} \cdot 1]^{-1}](\bar{y} - 0)}{\hat{\sigma}^2} = \frac{n\bar{y}_2}{\hat{\sigma}^2}$$

which is the same as part (a).

4.

(a). Note that $\bar{Y} = \beta_0 + \beta_1\bar{X}_1 + \beta_2\bar{X}_2 + \bar{\epsilon}$, hence

$$Y_i - \bar{Y} = \beta_1(X_{1i} - \bar{X}_1) + \beta_2(X_{i2} - \bar{X}_2) + (\epsilon_i - \bar{\epsilon})$$

then we have

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + (\epsilon_i - \bar{\epsilon})$$

so there is no β_0 in the deviation model.

$$X'X = \left[\begin{array}{cc} 12 & 8 \\ 8 & 12 \end{array}\right] \quad \text{and} \quad X'Y = (10, 8)' \quad \Longrightarrow (X'X)^{-1} = \left[\begin{array}{cc} 0.15 & -0.1 \\ -0.1 & 0.15 \end{array}\right]$$

Hence,

$$(\hat{\beta}_1, \hat{\beta}_2) = (0.7, 0.2) \quad \text{and} \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2 = 4$$

(b). Note that

$$V\left[\begin{array}{c} \hat{\beta}_1 \\ \hat{\beta}_2 \end{array}\right] = \sigma^2\left[\begin{array}{cc} 12 & 8 \\ 8 & 12 \end{array}\right]^{-1} = \sigma^2\left[\begin{array}{cc} \frac{12}{80} & -\frac{8}{80} \\ -\frac{8}{80} & \frac{12}{80} \end{array}\right]$$

and $\hat{\sigma}^2 = \frac{y'y - \hat{\beta}'x'y}{n-3} = 0.07$, hence $V(\hat{\beta_2}) = 0.07 * \frac{12}{80} = 0.0105$, so $Se(\hat{\beta}_2) = \sqrt(0.0105) = 0.102$

$$t_{\beta_2} = \frac{\hat{\beta}_2 - 0}{Se(\hat{\beta}_2)} = 1.96 \; > \; t_{0.05,20}$$

so reject H_0 at 5% significant level.

5.

(a). We have $\hat{\beta} = \beta + (X'X)^{-1}X'\epsilon$ i.e.

$$\sqrt{n}(\hat{\beta} - \beta) = \left(\frac{X'X}{n}\right)^{-1}\frac{X'\epsilon}{\sqrt{n}}$$

Since $\frac{X'\epsilon}{\sqrt{n}} \rightsquigarrow N\,(0, \; \sigma^2 Q)$ and $lim\;\frac{X'X}{n} = Q$, so

$$\sqrt{n}(\hat{\beta} - \beta) \rightsquigarrow \; N\,(0, Q^{-1}\sigma^2)$$

so

$$\sqrt{n}(R\hat{\beta} - R\beta) \rightsquigarrow \ N\left(0, RQ^{-1}R'\sigma^2\right)$$

Under $H_0: \ R\beta = \delta$

$$\sqrt{n}(R\hat{\beta} - \delta) \rightsquigarrow \ N\left(0, RQ^{-1}R'\sigma^2\right)$$

so

$$\frac{R\hat{\beta} - \delta}{\sqrt{\hat{\sigma}^2 R(X'X)^{-1}R'}} = \frac{\sqrt{n}(R\hat{\beta} - \delta)}{\sqrt{\hat{\sigma}^2 R\left(\frac{X'X}{n}\right)^{-1}R'}} \rightsquigarrow \frac{\sqrt{n}(R\hat{\beta} - \delta)}{\sqrt{\hat{\sigma}^2 RQ^{-1}R'}}$$

Since $\hat{\sigma}^2 \rightsquigarrow \sigma^2$ and $lim \ \frac{X'X}{n} = Q$, so

$$\frac{\sqrt{n}(R\hat{\beta} - \delta)}{\sqrt{\hat{\sigma}^2 RQ^{-1}R'}} \rightsquigarrow \ N\left(0, 1\right)$$

therefore

$$\frac{R\hat{\beta} - \delta}{\sqrt{\hat{\sigma}^2 R(X'X)^{-1}R'}} \rightsquigarrow \ N\left(0, 1\right)$$

(b). Under normality assumption

$$\hat{\beta} \sim N\left(\beta, \sigma^2(X'X)^{-1}\right) \quad \text{and} \quad \frac{(n-k)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-k}$$

Under H_0,

$$R\hat{\beta} \sim N\left(R\beta, \sigma^2 R(X'X)^{-1}R'\right) = \ N\left(\delta, \sigma^2 R(X'X)^{-1}R'\right)$$

so

$$\frac{R\hat{\beta} - \delta}{\sqrt{\hat{\sigma}^2 R(X'X)^{-1}R'}} = \frac{\frac{R\hat{\beta} - \delta}{\sqrt{\sigma^2 R(X'X)^{-1}R'}}}{\sqrt{\frac{(n-k)\hat{\sigma}^2}{\sigma^2}/(n-k)}} = \frac{N(0,1)}{\sqrt{\chi^2_{n-k}/(n-k)}}$$

therefore, under normal assumption, the statistic has exact t-distribution, asymptotically the distribution does not change.

6.

(a).

$$L_t = -\frac{1}{2}ln(2\pi) - \frac{1}{2}ln(\sigma^2) - \frac{(y_t - x_t'\beta)^2}{2\sigma^2}$$

$$L = \sum L_t = -\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\sigma^2) - \frac{\sum(y_t - x_t'\beta)^2}{2\sigma^2}$$

$$\hat{\beta} = \sum y_t x_t' (\sum x_t x_t')^{-1} \quad \text{and} \quad \hat{\sigma}^2 = \sum \hat{u}_t^2/n$$

where $\hat{u}_t = y_t - x_t'\hat{\beta}$

(b).

$$A_n(\hat{\theta}) = n^{-1}\sum \begin{bmatrix} -\frac{1}{\hat{\sigma}^2} & -\frac{x_t}{\hat{\sigma}^2} & -\frac{\hat{u}_t}{\hat{\sigma}^4} \\ -\frac{x_t}{\hat{\sigma}^2} & -\frac{x_t^2}{\hat{\sigma}^2} & -\frac{\hat{u}_t x_t}{\hat{\sigma}^4} \\ -\frac{\hat{u}_t}{\hat{\sigma}^4} & -\frac{\hat{u}_t x_t}{\hat{\sigma}^4} & -\frac{\hat{\sigma}^2 - 2\hat{u}_t^2}{2\hat{\sigma}^6} \end{bmatrix}$$

$$B_n(\hat{\theta}) = n^{-1}\sum \begin{bmatrix} \frac{\hat{u}_t}{\hat{\sigma}^4} & * & * \\ \frac{\hat{u}_t^2 x_t}{\hat{\sigma}^4} & \frac{\hat{u}_t^2 x_t^2}{\hat{\sigma}^4} & * \\ \frac{\hat{u}_t^3 - \hat{\sigma}^2\hat{u}_t}{2\hat{\sigma}^6} & \frac{(\hat{u}_t^3 - \hat{\sigma}^2\hat{u}_t)x_t}{2\hat{\sigma}^6} & \frac{\hat{u}_t^4 - 2\hat{\sigma}^2\hat{u}_t^2 + \hat{\sigma}^4}{4\hat{\sigma}^8} \end{bmatrix}$$

We define d_t as the following way:

$$d_t = \begin{bmatrix} d_{t1} \\ d_{t2} \\ d_{t3} \\ d_{t4} \\ d_{t5} \\ d_{t6} \end{bmatrix} = \begin{bmatrix} \frac{(\hat{u}_t^2 - \hat{\sigma}^2)}{\sigma^4} \\ \frac{(\hat{u}_t^2 - \hat{\sigma}^2)x_t}{\sigma^4} \\ \frac{(\hat{u}_t^2 - \hat{\sigma}^2)x_t^2}{\sigma^4} \\ \frac{(\hat{u}_t^3 - 3\hat{u}_t\hat{\sigma}^2)}{2\sigma^6} \\ \frac{(\hat{u}_t^3 - 3\hat{u}_t\hat{\sigma}^2)x_t}{2\sigma^6} \\ \frac{3\hat{\sigma}^4 - 2\hat{\sigma}^2\hat{u}_t^2 + \hat{u}_t^4}{4\hat{\sigma}^8} \end{bmatrix}$$

(c).

$$D_n(\hat{\theta}) = n^{-1}\sum d_t(\hat{\theta}) = \begin{bmatrix} 0 \\ \frac{\sum(\hat{u}_t^2 - \hat{\sigma}^2)x_t}{n\sigma^4} \\ \frac{\sum(\hat{u}_t^2 - \hat{\sigma}^2)x_t^2}{n\sigma^4} \\ \frac{2\sum \hat{u}_t^3}{2n\sigma^6} \\ \frac{2\sum \hat{u}_t^3 x_t}{2n\sigma^6} \\ \frac{2\sum(\hat{u}_t^4 - 3\hat{\sigma}^4)}{4n\sigma^8} \end{bmatrix}$$

(d).

We can write

$$F_{1t}(\hat{\theta}) = \frac{\partial L_t(\hat{\theta})}{\partial \theta} = \begin{bmatrix} \frac{\hat{u}_t}{\hat{\sigma}^2} \\ \frac{\hat{u}_t x_t}{\hat{\sigma}^2} \\ \frac{\hat{u}_t^2 - \hat{\sigma}^2}{2\hat{\sigma}^2} \end{bmatrix}$$

$$E_t^*[d_t F_{1t}'] = \begin{bmatrix} 0 & 0 & m_1 \\ 0 & 0 & m_2 \\ 0 & 0 & m_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad E_t^*[F_{1t}F_{1t}'] = \begin{bmatrix} \sigma^{-2}M & 0 \\ 0 & \frac{1}{2\sigma^4} \end{bmatrix}$$

where $m_1 = lim\, n^{-1}\sum \frac{1}{\sigma^4}$, $m_2 = lim\, n^{-1}\sum \frac{x_t}{\sigma^4}$, and $m_3 = lim\, n^{-1}\sum \frac{x_t^2}{\sigma^4}$

(e).

$$E_t^*(d_t d_t') = plim\, n^{-1}\sum \begin{bmatrix} \frac{2}{\sigma^4} & \frac{2x_t}{\sigma^4} & \frac{2x_t^2}{\sigma^4} & 0 & 0 & 0 \\ \frac{2x_t}{\sigma^4} & \frac{2x_t^2}{\sigma^4} & \frac{2x_t^3}{\sigma^4} & 0 & 0 & 0 \\ \frac{2x_t^2}{\sigma^4} & \frac{2x_t^3}{\sigma^4} & \frac{2x_t^4}{\sigma^4} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{2\sigma^6} & \frac{3x_t}{2\sigma^6} & 0 \\ 0 & 0 & 0 & \frac{3x_t}{2\sigma^6} & \frac{3x_t^2}{2\sigma^6} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{3}{2\sigma^8} \end{bmatrix}$$

Based on the discussion about $F_{1t}(\hat{\theta})$ and $E_t^*[d_t F_{1t}']$ in the above question, so we have

$$E_t^*(d_t F_{1t})[E_t^*[F_{1t}F_{1t}']]^{-1}E_t^*(F_{1t}d_t')$$

$$= \begin{bmatrix} m_1^2 2\sigma^4 & m_1 m_2 2\sigma^4 & m_1 m_3 2\sigma^4 & 0 & 0 & 0 \\ m_1 m_2 \sigma^4 & m_2^2 2\sigma^4 & m_2 m_3 2\sigma^4 & 0 & 0 & 0 \\ m_1 m_3 \sigma^4 & m_2 m_3 \sigma^4 & m_3^2 2\sigma^4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$E_t^*(d_t d_t')$ is block diagonal.

(f).

$$V(\theta) = \begin{bmatrix} \frac{1}{\sigma^4}\begin{bmatrix} 1 & u_1 & u_2 \\ u_1 & u_2 & u_3 \\ u_2 & u_3 & u_4 \end{bmatrix} & 0 & 0 \\ 0 & \frac{3}{2\sigma^6}\begin{bmatrix} 1 & u_1 \\ u_1 & u_2 \end{bmatrix} & 0 \\ 0 & 0 & \frac{3}{2\sigma^8} \end{bmatrix} - \begin{bmatrix} \frac{2}{\sigma^4}\begin{bmatrix} 1 & u_1 & u_2u_1 \\ u_1 & u_1u_1 & u_1u_2 \\ u_1u_1 & u_1u_2 & u_2u_3 \end{bmatrix} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so $T_n \rightsquigarrow \chi^2_5$, and

$$\Longrightarrow T_{1n} = [\sum(\hat{u}_t^2 - \hat{\sigma}^2)\xi'][\sum \xi_t\xi_t']^{-1}[\sum(\hat{u}_t^2 - \hat{\sigma}^2)\xi]/2\sigma^4 \rightsquigarrow \chi^2_2$$

ξ_t is $p(p+1)/2$ vector consisting of the lower triangular elements of $x_t x_t' - n^{-1}\sum x_t x_t'$

$$T_{2n} \rightsquigarrow \chi^2_2 \quad \text{and} \quad T_{3n} \rightsquigarrow \chi^2_1$$

(g). T_{1n}: Heteroskedasity test
T_{2n} and T_{3n} : J-B normality test.

7.
(a). The likelihood function is given by

$$l(\theta) = nLogL(\theta) = n\ log\ |2\sigma| - \frac{1}{\sigma}\sum |y_i - x_i'\beta|$$

$$\hat{\beta}^{MLE} = argmax_\beta l(\theta) = argmin_\beta \sum |y_i - x_i'\beta|$$

(b). When ϵ_i follows normal distribution, $\hat{\beta}_{ols}$ is appropriate. If the distribution is a Laplace distribution, then $\hat{\beta}_{ols}$ will be inefficient and affected by outlying observation. However, $\hat{\beta}_{ols}$ is easier to compute than $\hat{\beta}_{MLE}$.

8.
(a). By integrating w.r.t. y_i

$$log\ f(y_i) = k_1 + \int -\frac{y_i}{c_0 + c_2 y_i^2} dy_i$$

where k_1 is a constant. It is equivalent to:

$$f(y_i) = k_2 e^{\int -\frac{y_i}{c_0 + c_2 y_i^2} dy_i}$$

where k_2 is a constant, we denote $\phi(y_i) = e^{\int -\frac{y_i}{c_0 + c_2 y_i^2} dy_i}$, since $f(y_i)$ is a p.d.f, we should have

$$\int f(y_i) dy_i = 1 \Rightarrow k_2 = \frac{1}{\int \phi(y_i) dy_i}$$

therefore

$$f(y_i) = \frac{\phi(y_i)}{\int \phi(y_i) dy_i}$$

The log-likelihood fn. is

$$l(\theta) = \sum ln\phi(y_i) - n\ ln \int \phi(y_i) dy_i$$

(b). If we put $c_2 = 0$, that is to test Normality of y_i by testing $H_0 :\ \ c_2 = 0$, and LM test for normality will be based on $\frac{\partial l(\theta)}{\partial c_2}|_{c_2=0}$

(c). The test will be a test for kurtosis. For normal distribution, $kurtosis = 3$. The LM test will check whether the sample kurtosis is close to 3. Since we test only one restriction, this test statistic follows χ^2_1 asymptotically.

(d). No, this test basically assumes that the underlying density is symmetric. So for asymmetric density, this test is not suggested.