Spring 1997 John Rust
Economics 551b 37 Hillhouse, Rm. 27

PROBLEM SET 3: SOLUTIONS
Classical Methods (I)

QUESTION 1 Since Σ is symmetric, positive definite, there exists Cholesky decomposition P such that

$$\begin{aligned} \Sigma &= PP' \\ \Sigma^{-1} &= P^{-1\prime}P^{-1} \end{aligned}$$

Now we know $k \times 1$ vector $Z = P^{-1}Y \sim N(0, I)$, and each element of Z, $z_i \sim N(0, 1)$ for all $i = 1, \ldots, k$. Therefore,

$$Y'\Sigma^{-1}Y = Y'P^{-1\prime}P^{-1}Y = Z'Z = \sum_{i=1}^{k} z_i^2 \sim \chi^2(k).$$

QUESTION 2 Since

$$\varepsilon = y - X\beta \sim N\left(0, \Omega\right),$$

log likelihood function for β can be written as

$$l(\beta) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|\Omega| - \frac{1}{2}\Big((y - X\beta)'\Omega^{-1}(y - X\beta)\Big).$$

MLE of β, $\hat{\beta}$, satisfies

$$\frac{\partial l(\hat{\beta})}{\partial \beta} = X'\Omega^{-1}(y - X\hat{\beta}) = 0.$$

Therefore

$$X'\Omega^{-1}y = X'\Omega^{-1}X\hat{\beta} \Rightarrow \hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y = \hat{\beta}_{GLS}.$$

QUESTION 3

$$\begin{aligned} \mathcal{I}(\beta) &= E\left[\frac{\partial l(\hat{\beta})}{\partial \beta}\frac{\partial l(\hat{\beta})}{\partial \beta'}\right] \\ &= E\Big[X'\Omega^{-1}(y - X\beta)(y - X\beta)'\Omega^{-1}X\Big] \\ &= X'\Omega^{-1}E[(y - X\beta)(y - X\beta)']\Omega^{-1}X \\ &= X'\Omega^{-1}\Omega\Omega^{-1}X = X'\Omega^{-1}X \end{aligned}$$

Therefore

$$\mathcal{I}(\beta)^{-1} = \Big(X'\Omega^{-1}X\Big)^{-1}.$$

QUESTION 4 Define

$$X_i = \begin{bmatrix} X_{i1} & & 0 \\ & \ddots & \\ 0 & & X_{ip} \end{bmatrix} \quad , X = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Omega & & 0 \\ & \ddots & \\ 0 & & \Omega \end{bmatrix}$$

We know that GLS estimator and its covariance matrix can be written as follows;

$$\begin{aligned} \hat{\beta}_{GLS} &= (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y \\ &= \left[\sum_{i=1}^{N} X_i'\Omega^{-1}X_i\right]^{-1} \sum_{i=1}^{N} X_i'\Omega^{-1}y_i \end{aligned}$$

$$V(\hat{\beta}_{GLS}) = (X'\Sigma^{-1}X)^{-1}$$

Note this is stacked observation by observation. Now OLS estimators equation by equation can be written with matrices stacked by equation by equation as follows;

$$\hat{\beta} = \begin{bmatrix} (X_{(1)}'X_{(1)})^{-1}X_{(1)}{}'y_{(1)} \\ \vdots \\ (X_{(p)}'X_{(p)})^{-1}X_{(p)}{}'y_{(p)} \end{bmatrix}$$

where

$$X_{(i)} = \begin{bmatrix} X_{1i} \\ \vdots \\ X_{Ni} \end{bmatrix}.$$

But since

$$\begin{aligned} \hat{\beta} &= \begin{bmatrix} \left[\begin{bmatrix} X_{11} \\ \vdots \\ X_{N1} \end{bmatrix}' \begin{bmatrix} X_{11} \\ \vdots \\ X_{N1} \end{bmatrix}\right]^{-1} \begin{bmatrix} X_{11} \\ \vdots \\ X_{N1} \end{bmatrix}' \begin{bmatrix} y_{11} \\ \vdots \\ y_{N1} \end{bmatrix} \\ \vdots \\ \left[\begin{bmatrix} X_{1p} \\ \vdots \\ X_{Np} \end{bmatrix}' \begin{bmatrix} X_{1p} \\ \vdots \\ X_{Np} \end{bmatrix}\right]^{-1} \begin{bmatrix} X_{1p} \\ \vdots \\ X_{Np} \end{bmatrix}' \begin{bmatrix} y_{1p} \\ \vdots \\ y_{Np} \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} \left[\sum_{i=1}^{N} X_{i1}'X_{i1}\right]^{-1} \sum_{i=1}^{N} X_{i1}'y_{i1} \\ \vdots \\ \left[\sum_{i=1}^{N} X_{ip}'X_{ip}\right]^{-1} \sum_{i=1}^{N} X_{ip}'y_{ip} \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \left[\sum_{i=1}^{N}\begin{bmatrix} X_{i1}'X_{i1} & & 0 \\ & \ddots & \\ 0 & & X_{ip}'X_{ip}\end{bmatrix}\right]^{-1}\sum_{i=1}^{N}\begin{bmatrix} X_{i1}'y_{i1} & & 0 \\ & \ddots & \\ 0 & & X_{ip}'y_{ip}\end{bmatrix} \\
&= \left[\sum_{i=1}^{N}\begin{bmatrix} X_{i1} & & 0 \\ & \ddots & \\ 0 & & X_{ip}\end{bmatrix}'\begin{bmatrix} X_{i1} & & 0 \\ & \ddots & \\ 0 & & X_{ip}\end{bmatrix}\right]^{-1}\sum_{i=1}^{N}\begin{bmatrix} X_{i1} & & 0 \\ & \ddots & \\ 0 & & X_{ip}\end{bmatrix}'\begin{bmatrix} y_{i1} \\ \vdots \\ y_{ip}\end{bmatrix} \\
&= \left[\sum_{i=1}^{N} X_i'X_i\right]^{-1}\sum_{i=1}^{N} X_i'y_i \\
&= \left[\begin{bmatrix} X_1 \\ \vdots \\ X_N\end{bmatrix}'\begin{bmatrix} X_1 \\ \vdots \\ X_N\end{bmatrix}\right]^{-1}\begin{bmatrix} X_1 \\ \vdots \\ X_N\end{bmatrix}'\begin{bmatrix} y_1 \\ \vdots \\ y_N\end{bmatrix} \\
&= (X'X)^{-1}Xy
\end{aligned}
$$

$$
V(\hat{\beta}) = (X'X)^{-1}(X'\Sigma^{-1}X)(X'X)^{-1}
$$

Therefore

$$
V(\hat{\beta}) - V(\hat{\beta}_{GLS}) \geq 0
$$

QUESTION 5

A.

$$
\begin{aligned}
\int_{-\infty}^{\infty} f(\varepsilon)d\varepsilon &= \int_{-\infty}^{\infty} K\exp\left\{-\left|\frac{\varepsilon}{\sqrt{\sigma^2}}\right|\right\} \\
&= \int_{-\infty}^{0} K\exp\left\{\frac{\varepsilon}{\sqrt{\sigma^2}}\right\} + \int_{0}^{\infty} K\exp\left\{-\frac{\varepsilon}{\sqrt{\sigma^2}}\right\} \\
&= K\left[\sqrt{\sigma^2}\exp\left\{\frac{\varepsilon}{\sqrt{\sigma^2}}\right\}\right]_{-\infty}^{0} + K\left[\sqrt{\sigma^2}\exp\left\{-\frac{\varepsilon}{\sqrt{\sigma^2}}\right\}\right]_{0}^{\infty} \\
&= 2K\sqrt{\sigma^2} = 1
\end{aligned}
$$

Therefore

$$
K = \frac{1}{2\sqrt{\sigma^2}}.
$$

B. Since

$$
f(\varepsilon) = \frac{1}{2\sqrt{\sigma^2}}\exp\left\{-\left|\frac{\varepsilon}{\sqrt{\sigma^2}}\right|\right\},
$$

log likelihood function for β can be written as

$$
l(\beta) = -N\log(2\sqrt{\sigma^2}) - \frac{1}{\sqrt{\sigma^2}}\sum_{i=1}^{N}|y_i - X_i\beta|\,.
$$

Therefore

$$\hat{\beta} = \arg\max l(\beta) \Rightarrow \hat{\beta} = \arg\min \sum_{i=1}^{N} |y_i - X_i\beta| = \hat{\beta}_{LAD}.$$

QUESTION 6 The conditional posterior for β, $p(\beta|\Omega^{-1}, y, X)$, is proportional to $p(\beta)f(y|\beta, \Omega^{-1}, X)$, and the latter is given by:

$$\begin{aligned}
& p(\beta)f(y|\beta,\Omega^{-1},X) \\
= \; & (2\pi)^{-\frac{N}{2}} |A_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\beta-\beta_0)' A_0^{-1}(\beta-\beta_0)\right\} \\
& \times (2\pi)^{-\frac{N}{2}} \left|\Omega^{-1}\right|^{\frac{N}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{N}(y_i - X_i\beta)'\Omega^{-1}(y_i - X_i\beta)\right\} \\
\propto \; & \exp\left\{-\frac{1}{2}\left((\beta-\beta_0)'A_0^{-1}(\beta-\beta_0) + \sum_{i=1}^{N}(y_i - X_i\beta)'\Omega^{-1}(y_i - X_i\beta)\right)\right\} \\
\propto \; & \exp\left\{-\frac{1}{2}\left((\beta-\beta_0)'A_0^{-1}(\beta-\beta_0) + \sum_{i=1}^{N}\beta' X_i'\Omega_i^{-1}X_i\beta - \sum_{i=1}^{N} 2\beta' X_i\Omega^{-1}y_i\right)\right\} \\
\propto \; & \exp\left\{-\frac{1}{2}\left((\beta-\beta_0)'A_0^{-1}(\beta-\beta_0) + \right.\right. \\
& \left.\left.\left(\beta - \left[\sum_{i=1}^{N} X_i'\Omega_i^{-1}X_i\right]^{-1}\sum_{i=1}^{N} X_i\Omega^{-1}y_i\right)' \sum_{i=1}^{N} X_i'\Omega_i^{-1}X_i \left(\beta - \left[\sum_{i=1}^{N} X_i'\Omega_i^{-1}X_i\right]^{-1}\sum_{i=1}^{N} X_i\Omega^{-1}y_i\right)\right)\right\}
\end{aligned}$$

By completing the square,

$$\begin{aligned}
& \beta - \left[\left(A_0^{-1} + \sum_{i=1}^{N} X_i'\Omega_i^{-1}X_i\right)^{-1} A_0^{-1}\beta_0 + \right. \\
& \left. \left(A_0^{-1} + \sum_{i=1}^{N} X_i'\Omega_i^{-1}X_i\right)^{-1} \sum_{i=1}^{N} X_i'\Omega_i^{-1}X_i \left[\sum_{i=1}^{N} X_i'\Omega_i^{-1}X_i\right]^{-1} \sum_{i=1}^{N} X_i\Omega^{-1}y_i\right] \\
= \; & \beta - \left[A_N^{-1}A_0^{-1}\beta_0 + A_N^{-1}\sum_{i=1}^{N} X_i\Omega^{-1}y_i\right] \\
= \; & \beta - \left[A_N^{-1}\left(A_0^{-1}\beta_0 + \sum_{i=1}^{N} X_i\Omega^{-1}y_i\right)\right] \\
= \; & \beta - \hat{\beta}
\end{aligned}$$

$$p(\beta|\Omega^{-1}, y, X) = (2\pi)^{-\frac{N}{2}} |A_N|^{-\frac{1}{2}} \exp\left\{\frac{1}{2}(\beta-\hat{\beta})A_N^{-1}(\beta-\hat{\beta})'\right\}.$$

i.e. the conditional density of β given $\Omega^{-1}, y, X)$ is $N\left(\hat{\beta}, A_N^{-1}\right)$.

2. The conditional posterior for Ω^{-1}, $p(\Omega^{-1}|\beta, y, X)$, is prortional to $p(\Omega^{-1})f(y|\beta, \Omega^{-1}, X)$, which is given by:

$$p(\Omega^{-1})f(y|\beta, \Omega^{-1}, X)$$

$$
\begin{aligned}
&= \frac{|R_0|^{-\frac{\rho_0}{2}} \left|\Omega^{-1}\right|^{\frac{\rho_0-(p-1)}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}(R_0^{-1}\Omega^{-1})\right\}}{2^{\frac{\rho_0 p}{2}}\pi^{-\frac{p(p-1)}{4}}\Pi_{i=1}^{p}, \ \left(\frac{\rho_0+1-i}{2}\right)} \\
&\quad \times(2\pi)^{-\frac{N}{2}}\left|\Omega^{-1}\right|^{\frac{N}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{N}(y_i - X_i\beta)'\Omega^{-1}(y_i - X_i\beta)\right\} \\
&\propto \left|\Omega^{-1}\right|^{\frac{\rho_0-(p-1)}{2}} \left|\Omega^{-1}\right|^{\frac{N}{2}} \exp\left\{-\frac{1}{2}\left(\mathrm{tr}(R_0^{-1}\Omega^{-1}) + \sum_{i=1}^{N}\mathrm{tr}(y_i - X_i\beta)'\Omega^{-1}(y_i - X_i\beta)\right)\right\} \\
&= \left|\Omega^{-1}\right|^{\frac{N+\rho_0-(p-1)}{2}} \exp\left\{-\frac{1}{2}\left(\mathrm{tr}(\Omega^{-1}R_0^{-1}) + \mathrm{tr}\left[\Omega^{-1}\sum_{i=1}^{N}(y_i - X_i\beta)(y_i - X_i\beta)'\right]\right)\right\} \\
&= \left|\Omega^{-1}\right|^{\frac{N+\rho_0-(p-1)}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Omega^{-1}R_0^{-1} + \Omega^{-1}\sum_{i=1}^{N}(y_i - X_i\beta)(y_i - X_i\beta)'\right]\right\} \\
&= \left|\Omega^{-1}\right|^{\frac{N+\rho_0-(p-1)}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Omega^{-1}\left(R_0^{-1} + \sum_{i=1}^{N}(y_i - X_i\beta)(y_i - X_i\beta)'\right)\right]\right\} \\
&= \left|\Omega^{-1}\right|^{\frac{N+\rho_0-(p-1)}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Omega^{-1}R_N^{-1}\right]\right\}
\end{aligned}
$$

Therefore

$$
p(\Omega^{-1}|\beta, y, X) = \frac{|R_N|^{-\frac{(N+\rho_0)}{2}} \left|\Omega^{-1}\right|^{\frac{N+\rho_0-(p-1)}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}(R_N^{-1}\Omega^{-1})\right\}}{2^{\frac{(N+\rho_0)p}{2}}\pi^{-\frac{p(p-1)}{4}}\Pi_{i=1}^{p}, \ \left(\frac{(N+\rho_0)+1-i}{2}\right)}
$$

i.e. the density of Ω^{-1} given (β, y, X) is $W\left(N+\rho_0, R_N\right)$.

QUESTION 7

$$
\begin{aligned}
\frac{\partial U(v_1, \ldots, v_D, X)}{\partial v_d} &= \frac{\partial}{\partial v_d}\int \max_{d=1,\ldots,D}\left[v_d + \varepsilon_d\right] f(\varepsilon|X)d\varepsilon \\
&= \int \frac{\partial}{\partial v_d}\max_{d=1,\ldots,D}\left[v_d + \varepsilon_d\right] f(\varepsilon|X)d\varepsilon \\
&= \int I\{\delta(\varepsilon) = d\} f(\varepsilon|X)d\varepsilon \\
&= P\{d|X\}.
\end{aligned}
$$

QUESTION 8 From Mood-Graybill-Boes(1974,p542), the mean of extreme value distribution with CDF

$$
F(x) = \exp\left\{-\exp\left\{-\frac{(x-\alpha)}{\sigma}\right\}\right\}
$$

is $\alpha + \sigma\gamma$. We know

$$
\begin{aligned}
P\left\{\varepsilon_d \le x\right\} &= F(x) = \exp\left\{-\exp\left\{-x\right\}\right\} \\
P\left\{\varepsilon_d \le x - v_d\right\} &= F(x) = \exp\left\{-\exp\left\{-(x - v_d)\right\}\right\}.
\end{aligned}
$$

First, we derive CDF of $\max_{d=1,\ldots,D}[\upsilon_d + \varepsilon_d]$.

$$
\begin{aligned}
& P\left\{\max_{d=1,\ldots,D}[\upsilon_d + \varepsilon_d] \le x\right\} \\
= \; & F(x) \\
= \; & P\{\upsilon_1 + \varepsilon_1 \le x, \ldots, \upsilon_D + \varepsilon_D \le x\} \\
= \; & P\{\upsilon_1 + \varepsilon_1 \le x\} \cdots P\{\upsilon_D + \varepsilon_D \le x\} \\
& \text{(by independence)} \\
= \; & P\{\varepsilon_1 \le x - \upsilon_1\} \cdots P\{\varepsilon_D \le x - \upsilon_D\} \\
= \; & \exp\{-\exp\{-(x - \upsilon_1)\}\} \cdots \exp\{-\exp\{-(x - \upsilon_D)\}\} \\
= \; & \exp\left\{-\sum_{d=1}^{D} \exp\{-(x - \upsilon_d)\}\right\} \\
= \; & \exp\left\{-\exp\left[\log\left(\sum_{d=1}^{D} \exp\{-(x - \upsilon_d)\}\right)\right]\right\} \\
= \; & \exp\left\{-\exp\left[\log\left(\exp(-x)\sum_{d=1}^{D} \exp(\upsilon_d)\right)\right]\right\} \\
= \; & \exp\left\{-\exp\left[-x + \log\left(\sum_{d=1}^{D} \exp(\upsilon_d)\right)\right]\right\} \\
= \; & \exp\left\{-\exp\left[-\left(x - \log\left(\sum_{d=1}^{D} \exp(\upsilon_d)\right)\right)\right]\right\}.
\end{aligned}
$$

Since this is CDF of extreme value distribution, the extreme value family is max-stable, and the mean of $\max_{d=1,\ldots,D}[\upsilon_d + \varepsilon_d]$ can be written as

$$
\begin{aligned}
E\left[\max_{d=1,\ldots,D}[\upsilon_d + \varepsilon_d]\right] &= U(\upsilon_1, \ldots, \upsilon_D, X) \\
&= \alpha + \sigma\gamma \\
&= \log\left(\sum_{d=1}^{D} \exp(\upsilon_d)\right) + \gamma.
\end{aligned}
$$

By Williams-Daly-Zachary Theorem,

$$
\begin{aligned}
P\{d|X\} &= \frac{\partial U(\upsilon_1, \ldots, \upsilon_D, X)}{\partial \upsilon_d} \\
&= \frac{\exp(\upsilon_d)}{\sum_{d'=1}^{D} \exp(\upsilon_{d'})}.
\end{aligned}
$$