# Time Series: M.Phil. Econometrics, Hilary Term 2014
# Solutions for week 4 problem set

Oleg I. Kitov

February 10, 2014

## Question 1

Consider the model

$$X_t = \alpha X_{t-1} + \mu D_t + \epsilon_t, \quad \text{for } t = 1, ..., T, \tag{1}$$

conditionally on $X_0$, where $\epsilon_t$ are independent $N\left(0, \sigma^2\right)$.

(a) Derive the likelihood function.

The joint density for $X_1, \ldots, X_T$ given $X_0$ is (notice that lower case denotes the random variables, while upper case denotes the realizations of the random variables):

$$\begin{aligned}
& \mathsf{f}_{\alpha,\mu,\sigma^2}(x_1, \ldots, x_T | x_0) & (2)\\
= & \text{ [ prediction decomposition ] } = \prod_{t=1}^{T} \mathsf{f}_{\alpha,\mu,\sigma^2}(x_t | x_{t-1}, \ldots, x_0) & (3)\\
= & \text{ [ 1st order proces ] } = \prod_{t=1}^{T} \mathsf{f}_{\alpha,\mu,\sigma^2}(x_t | x_{t-1}) & (4)\\
= & \text{ [ normality ] } = \prod_{t=1}^{T} (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}(x_t - \alpha x_{t-1} - \mu d_t)^2\right\} & (5)\\
= & (2\pi\sigma^2)^{-T/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(x_t - \alpha x_{t-1} - \mu d_t)^2\right\}, & (6)
\end{aligned}$$

leading to log likelihood

$$\ell_{X_1,\ldots,X_T|X_0}(\alpha, \mu, \sigma^2) = -\frac{T}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(X_t - \alpha X_{t-1} - \mu D_t)^2. \tag{7}$$

(b) Derive maximum likelihood estimators for $\alpha$, $\mu$, and $\sigma^2$.

Using the orthogonalisation method from week 3 we can immediately see that the the three estimates are given by:

$$\hat{\alpha} = \frac{\sum_{t=1}^{T} Y_t (Y_{t-1}|D_t)}{\sum_{t=1}^{T} (Y_{t-1}|D_t)^2} \tag{8}$$

$$\hat{\mu} = \frac{\sum_{t=1}^{T} Y_t (D_t|Y_{t-1})}{\sum_{t=1}^{T} (D_t|Y_{t-1})^2} \tag{9}$$

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} (Y_t|Y_{t-1}, D_t)^2 \tag{10}$$

where $(Y_{t-1}|D_t)$ and $D_t|Y_{t-1}$ are the residuals from regressing $Y_{t-1}$ on $D_t$ and vice-versa, and $(Y_t|Y_{t-1}, D_t)$ are the residuals from the full regression. Just to prove my point that this representation is much more beautiful than the direct derivation without orthogonalisation, I will show you the results from the the latter. You would have seen (and hopefully done) this several times throughout the course, but let's do it one more time, just for the fun of it. First, let me rewrite the regression

$$X_t = \alpha X_{t-1} + \mu D_t + \epsilon_t, \tag{11}$$

as follow

$$X_t = Z_t'\theta + \epsilon_t, \tag{12}$$

where

$$\underset{(1\times 2)}{Z_t'} \equiv \begin{bmatrix} X_{t-1} & D_t \end{bmatrix}, \tag{13}$$

and

$$\underset{(2\times 1)}{\theta} \equiv \begin{bmatrix} \alpha \\ \mu \end{bmatrix}. \tag{14}$$

where $\theta$ is a standard way of denoting a vector of model parameters. Hence, the log likelihood becomes

$$\ell_{X_1,\dots,X_T|X_0}(\theta, \sigma^2) = -\frac{T}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(X_t - Z_t'\theta)^2. \tag{15}$$

We will first the scores by taking first derivatives of log likelihood:

$$\frac{\partial}{\partial\theta}\ell(\theta,\sigma^2) = \frac{2}{2\sigma^2}\sum_{t=1}^{T} Z_t(X_t - Z_t'\theta) \tag{16}$$

$$\frac{\partial}{\partial\sigma^2}\ell(\theta,\sigma^2) = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{t=1}^{T}(X_t - Z_t'\theta)^2. \tag{17}$$

Then the first order conditions for $\theta$ is given by

$$\sum_{t=1}^{T} Z_t(X_t - Z_t'\theta) = 0, \tag{18}$$

hence

$$\hat{\theta} = \left(\sum_{t=1}^{T} Z_t Z_t'\right)^{-1} \sum_{t=1}^{T} Z_t X_t \tag{19}$$

$$= \begin{bmatrix} \sum_{t=1}^{T} X_{t-1}^2 & \sum_{t=1}^{T} X_{t-1}D_t \\ \sum_{t=1}^{T} X_{t-1}D_t & \sum_{t=1}^{T} D_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} X_{t-1}X_t \\ \sum_{t=1}^{T} D_t X_t \end{bmatrix}. \tag{20}$$

In particular,

$$\left(\sum_{t=1}^{T} Z_t Z_t'\right)^{-1} = \left[\begin{array}{cc} \sum_{t=1}^{T} X_{t-1}^2 & \sum_{t=1}^{T} X_{t-1} D_t \\ \sum_{t=1}^{T} X_{t-1} D_t & \sum_{t=1}^{T} D_t^2 \end{array}\right]^{-1} \tag{21}$$

$$= \frac{1}{\left(\sum_{t=1}^{T} X_{t-1}^2\right)\left(\sum_{t=1}^{T} D_t^2\right) - \left(\sum_{t=1}^{T} X_{t-1} D_t^2\right)^2} \tag{22}$$

$$\times \left[\begin{array}{cc} \sum_{t=1}^{T} D_t^2 & -\sum_{t=1}^{T} X_{t-1} D_t \\ -\sum_{t=1}^{T} X_{t-1} D_t & \sum_{t=1}^{T} X_{t-1}^2 \end{array}\right], \tag{23}$$

and

$$\left[\begin{array}{cc} \sum_{t=1}^{T} D_t^2 & -\sum_{t=1}^{T} X_{t-1} D_t \\ -\sum_{t=1}^{T} X_{t-1} D_t & \sum_{t=1}^{T} D_t^2 \end{array}\right] \times \left[\begin{array}{c} \sum_{t=1}^{T} X_{t-1} X_t \\ \sum_{t=1}^{T} D_t X_t \end{array}\right] \tag{24}$$

$$= \left[\begin{array}{c} \sum_{t=1}^{T} D_t^2 \sum_{t=1}^{T} X_{t-1} X_t - \sum_{t=1}^{T} X_{t-1} D_t \sum_{t=1}^{T} D_t X_t \\ \sum_{t=1}^{T} D_t X_t \sum_{t=1}^{T} X_{t-1}^2 - \sum_{t=1}^{T} X_{t-1} D_t \sum_{t=1}^{T} X_{t-1} X_t \end{array}\right]. \tag{25}$$

And just one more step to go:

$$\hat{\theta} = \left[\begin{array}{c} \hat{\alpha} \\ \hat{\mu} \end{array}\right] = \left[\begin{array}{c} \frac{\sum_{t=1}^{T} D_t^2 \sum_{t=1}^{T} X_{t-1} X_t - \sum_{t=1}^{T} X_{t-1} D_t \sum_{t=1}^{T} D_t X_t}{\left(\sum_{t=1}^{T} X_{t-1}^2\right)\left(\sum_{t=1}^{T} D_t^2\right) - \left(\sum_{t=1}^{T} X_{t-1} D_t\right)^2} \\ \frac{\sum_{t=1}^{T} D_t X_t \sum_{t=1}^{T} X_{t-1}^2 - \sum_{t=1}^{T} X_{t-1} D_t \sum_{t=1}^{T} X_{t-1} X_t}{\left(\sum_{t=1}^{T} X_{t-1}^2\right)\left(\sum_{t=1}^{T} D_t^2\right) - \left(\sum_{t=1}^{T} X_{t-1} D_t\right)^2} \end{array}\right] \tag{26}$$

Isn't this pretty! Finally, the first order conditions for $(\sigma^2)$ yields:

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} (X_t - Z_t' \hat{\theta})^2.$$

(c) Derive the likelihood ratio test statistic, $\mathsf{LR}$ say, for the hypothesis $\alpha_2 = 0$.

The likelihood ratio statistic is given by

$$\mathsf{LR} = -2(\hat{\ell}_R - \hat{\ell}) \tag{27}$$

where $\hat{\ell}$ is the likelihood for the unrestricted model, and $\hat{\ell}_R$ is the likelihood for the restricted model in which the hypothesis $H_0 : \mu = 0$ is imposed. Substituting the estimates from (b) into the (unrestricted) likelihood function we get

$$\hat{\ell}(\hat{\theta}, \widehat{\sigma}_U^2) = -\frac{T}{2} \log(2\pi \widehat{\sigma}_U^2) - \frac{1}{2\widehat{\sigma}_U^2} \sum_{i=1}^{T} \hat{\epsilon}_{Ut}^2 = -\frac{T}{2} \log(2\pi \widehat{\sigma}_U^2) - \frac{T}{2} \tag{28}$$

where

$$\widehat{\sigma}_U^2 = \frac{1}{T} \sum_{t=1}^{T} (X_t - \widehat{\alpha}_U X_{t-1} - \widehat{\mu}_U D_t)^2 \tag{29}$$

The restricted likelihood value is given by

$$\hat{\ell}(\hat{\theta}_R, \widehat{\sigma}_R^2) = -\frac{T}{2} \log(2\pi \widehat{\sigma}_R^2) - \frac{1}{2\widehat{\sigma}_R^2} \sum_{i=1}^{T} \hat{\epsilon}_{Rt}^2 = -\frac{T}{2} \log(2\pi \widehat{\sigma}_R^2) - \frac{T}{2} \tag{30}$$

where

$$\widehat{\sigma}_R^2 = \frac{1}{T}\sum_{t=1}^{T}(X_t - \widehat{\alpha}_R X_{t-1})^2 \tag{31}$$

and

$$\widehat{\alpha}_R = \frac{\sum_{t=1}^T X_t X_{t-1}}{\sum_{t=1}^T X_{t-1}^2} \tag{32}$$

The Likelihood Ratio statistic can therefore be calculated as

$$\begin{aligned}
\mathsf{LR} &= -2(\hat{\ell}_R - \hat{\ell}) && (33)\\
&= -2\left[-\frac{T}{2}\log(2\pi\widehat{\sigma}_R^2) - \frac{T}{2} + \frac{T}{2}\log(2\pi\widehat{\sigma}_U^2) + \frac{T}{2}\right] && (34)\\
&= -2\left[-\frac{T}{2}\log(2\pi\widehat{\sigma}_R^2) + \frac{T}{2}\log(2\pi\widehat{\sigma}_U^2)\right] && (35)\\
&= -T\left[\log\left(\frac{\widehat{\sigma}_U^2}{\widehat{\sigma}_R^2}\right)\right] && (36)\\
&= -T\log\left(1 - \frac{\hat{\sigma}_R^2 - \widehat{\sigma}_U^2}{\hat{\sigma}_R^2}\right) && (37)\\
&= -T\log\left(1 - \hat{\lambda}_{\mu=0}^2\right). && (38)
\end{aligned}$$

where the last equality comes from our derivations in week 3. Recall that

$$\begin{aligned}
\hat{\lambda}_{\mu=0}^2 &= \hat{\text{Corr}}^2\,(Y_t, D_t|Y_{t-1}) && (39)\\
&= \frac{\left(\sum_{t=1}^T Y_t D_t - \frac{\sum_{t=1}^T Y_t Y_{t-1}\sum_{t=1}^T D_t Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2}\right)^2}{\{\sum_{t=1}^T Y_t^2 - \frac{(\sum_{t=1}^T Y_t Y_{t-1})^2}{\sum_{t=1}^T Y_{t-1}^2}\}\{\sum_{t=1}^T D_t^2 - \frac{(\sum_{t=1}^T D_t Y_{t-1})^2}{\sum_{t=1}^T Y_{t-1}^2}\}} && (40)
\end{aligned}$$

(d) What is the link between $\mathsf{LR}$ and the $\mathsf{t}$- and $\mathsf{F}-$statistic?

We have already seen this last week, but let's do do it once again using a faster method. First, from the the definition of the $\mathsf{F}-$statistic we get that:

$$\begin{aligned}
\mathsf{F} &= \frac{(RSS_R - RSS_U)}{(T-2)-(T-1)}\Big/\frac{RSS_R}{T-2} && (41)\\
&= \frac{\left(\hat{\sigma}_R^2 - \widehat{\sigma}_U^2\right)/1}{\widehat{\sigma}_U^2/\,(T-2)} && (42)\\
&= (T-2)\left(\frac{1 - \hat{\sigma}_U^2/\widehat{\sigma}_R^2}{\hat{\sigma}_U^2/\widehat{\sigma}_R^2}\right) && (43)\\
&= (T-2)\left(\frac{\hat{\lambda}^2}{1-\hat{\lambda}^2}\right) && (44)
\end{aligned}$$

For the $\mathsf{t}$-statistic we have already seen that

$$\mathsf{t} = \sqrt{T-2}\left(\frac{\hat{\lambda}}{\sqrt{1-\hat{\lambda}^2}}\right) \tag{45}$$

And the following relationship follows immediately:

$$\mathsf{t}_{\mu=0} = \sqrt{\mathsf{F}_{\mu=0}} \tag{46}$$

Now, the LR statistic is given by:

$$LR = -T \log\left(1 - \hat{\lambda}^2\right) \tag{47}$$

And so the LR statistic is related to the $\mathsf{t}$- and $\mathsf{F}-$statistics via the partial correlation term $\mathsf{t}$- and $\mathsf{F}-$statistic. For instance, we can write down:

$$\begin{aligned} \mathsf{t}^2 &= (T-2)\left(\frac{\hat{\lambda}^2}{1-\hat{\lambda}^2}\right) \\ &= (T-2)\left(\frac{1 - exp\left(-\frac{1}{T}LR\right)}{exp\left(-\frac{1}{T}LR\right)}\right) \\ &= (T-2)\left(exp\left(\frac{1}{T}LR\right) - 1\right) \end{aligned}$$

Note that this of course only works for testing hypothesis on a single coefficient. Interpretation of $\hat{\lambda}$ as a ratio of standard error estimates from the regression is the following: by how much does the fit of the regression improve if we add $\mu$ as an additional regressors.

(e) Show that the long-run mean for $X_t$ when $|\alpha| < 1$ has the following form:

- $\mu/\left(1-\alpha\right)$ when $D_t = 1$
- Alternating between $\mu/\left(1-\alpha^2\right)$ for even $t$ and $\alpha\mu/\left(1-\alpha^2\right)$ for odd $t$ when $D_t$is the seasonal dummy.

There are several proof strategies. Here we have a go directly at the recursive equation. Rewrite $Y_t = \alpha Y_{t-1} + \eta_t$ where $\eta_t = \mu D_t + \epsilon_t$. We have

$$Y_t = \sum_{j=0}^{t-1} \alpha^j \eta_{t-j} + \alpha^t Y_0 = \sum_{j=0}^{t-1} \alpha^j \epsilon_{t-j} + \mu \sum_{j=0}^{t-1} \alpha^j D_{t-j} + \alpha^t Y_0 \tag{48}$$

Note that the first and the third terms go to zero as $T$ goes to infinity. Then all we need to do is to compute the second term for different variations of $D_t$. First, suppose $D_t = 1$. Then

$$\mu \sum_{j=0}^{t-1} \alpha^j D_{t-j} = \mu \sum_{j=0}^{t-1} \alpha^j = \mu \frac{1-\alpha^t}{1-\alpha} \to \frac{\mu}{1-\alpha} \tag{49}$$

Thus the long-run mean is $\mu_Y = \mu/(1-\alpha)$.

Alternatively, suppose $D_t = 1_{(t \text{ is even})}$. If $t$ is even then

$$\mu \sum_{j=0}^{t-1} \alpha^j D_{t-j} = \mu(\alpha^0 + \alpha^2 + \alpha^4 \cdots + \alpha^{t-2}) \tag{50}$$

$$= \mu \sum_{j=0}^{t/2-1} \alpha^{2j} = \mu \frac{1-\alpha^t}{1-\alpha^2} \to \frac{\mu}{1-\alpha^2} \tag{51}$$

If, on the other hand, $t$ is odd then

$$\mu \sum_{j=0}^{t-1} \alpha^j D_{t-j} = \mu(\alpha^1 + \alpha^3 \cdots - \alpha^{t-2}) \tag{52}$$

$$= \mu\alpha \sum_{j=0}^{(t-1)/2-1} \alpha^{2j} \tag{53}$$

$$= \mu\alpha \frac{1-\alpha^{t-1}}{1-\alpha^2} \to \frac{\mu\alpha}{1-\alpha^2} \tag{54}$$

Thus the long-run mean is alternating between$\mu/(1-\alpha^2)$ for even $t$ and $\alpha\mu/(1-\alpha^2)$ for odd$t$. This is called a non-centered seasonal dummy. If replacing$D_t$ by$(-1)^t$ we get a centered seasonal dummy, where the long run mean is$(-1)^t\mu/(1+\alpha)$.

## Question 2

Consider a $2 \times 2$ block diagonal matrices:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad L = \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix}$$

Let $A$ be symmetric so $A_{12} = A_{21}'$. Define the matrices:

$$A_{11\cdot 2} = A_{11} - A_{12}A_{22}^{-1}A_{21},. \quad A_{22\cdot 1} = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

(a) Compute $LAL'$.

Using the fact that $A$ is symmetric, the matrix $LAL'$ can be constructed as follows:

$$LAL' = \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} \tag{55}$$

$$= \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} \tag{56}$$

$$= \begin{bmatrix} A_{11\cdot 2} & 0 \\ 0 & A_{22} \end{bmatrix} \tag{57}$$

where $A_{11\cdot 2} = A_{11} - A_{12}A_{22}^{-1}A_{21}$. This is the orthogonalized matrix. We will see this later. Also, notice that since $LAL'$ is block diagonal, its inverse can be obtained by simply inverting the elements on the diagonal. Hence,

$$(LAL')^{-1} = \begin{bmatrix} A_{11\cdot 2}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix} \tag{58}$$

(b) Compute $A^{-1}$ using the formula $A^{-1} = L'(LAL')^{-1}L$.

The matrix $A^{-1} = L'(LAL')^{-1}L$ is

$$L'(LAL')^{-1}L = \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} \begin{bmatrix} A_{11\cdot 2}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix} \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \quad (59)$$

$$= \begin{bmatrix} A_{11\cdot 2}^{-1} & 0 \\ -A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} & A_{22}^{-1} \end{bmatrix} \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \quad (60)$$

$$= \begin{bmatrix} A_{11\cdot 2}^{-1} & -A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \end{bmatrix} \quad (61)$$

(c) Check your answer: Do you get $AA^{-1} = A^{-1}A = I$.

Manipulate the result from part (b) to get:

$$\begin{aligned}
AA^{-1} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} A_{11\cdot 2}^{-1} & -A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \end{bmatrix} \\
&= \begin{bmatrix} A_{11}A_{11\cdot 2}^{-1} - A_{12}A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} & -A_{11}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} + A_{12}\left(A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1}\right) \\ A_{21}A_{11\cdot 2}^{-1} - A_{22}A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} & -A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} + A_{22}\left(A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1}\right) \end{bmatrix} \\
&= \begin{bmatrix} \left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)A_{11\cdot 2}^{-1} & -A_{11}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} + A_{12}\left(A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1}\right) \\ A_{21}A_{11\cdot 2}^{-1} - A_{21}A_{11\cdot 2}^{-1} & -A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} + I + A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \end{bmatrix} \\
&= \begin{bmatrix} A_{11\cdot 2}A_{11\cdot 2}^{-1} & A_{12}A_{22}^{-1} + \left(-A_{11} + A_{12}A_{22}^{-1}A_{21}\right)A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \\
&= \begin{bmatrix} I & A_{12}A_{22}^{-1} - A_{11\cdot 2}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \\
&= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}
\end{aligned}$$

And similarly for $A^{-1}A = I$.

(d) Why is $A_{22\cdot 1}^{-1} = A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1}$?

Let

$$\tilde{L} = \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}, \quad (62)$$

and compute $A^{-1} = \tilde{L}'(\tilde{L}A\tilde{L}')^{-1}\tilde{L}$. First notice that

$$\tilde{L}A\tilde{L}' = \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \quad (63)$$

$$= \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \quad (64)$$

$$= \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}, \quad (65)$$

so that

$$(\tilde{L}A\tilde{L}')^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & \left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1} \end{bmatrix}, \quad (66)$$

and

$$\tilde{L}'(\tilde{L}A\tilde{L}')^{-1}\tilde{L} = \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & \left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \tag{67}$$

$$= \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1} \\ 0 & \left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \tag{68}$$

$$= \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{21}\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1} \\ -\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1}A_{21}A_{11}^{-1} & \left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1} \end{bmatrix} \tag{69}$$

Now, by definition

$$A^{-1} = L'(LAL')^{-1}L = \tilde{L}'(\tilde{L}A\tilde{L}')^{-1}\tilde{L}, \tag{70}$$

and compare this with the expression from (b)

$$L'(LAL')^{-1}L = \begin{bmatrix} A_{11\cdot 2}^{-1} & -A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \end{bmatrix} \tag{71}$$

hence,

$$A_{22}^{-1} + A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} = \left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1} = A_{22\cdot 1}^{-1}. \tag{72}$$

(e) Consider the regression $Y_i = \beta X_i + u_i$, $i = 1, \ldots, n$. Define $A_{11} = \sum_{i=1}^n Y_i^2$, $A_{12} = \sum_{i=1}^n Y_i X_i$, $A_{22} = \sum_{i=1}^n X_i^2$. Interpret the matrices $L$ and $LAL'$.

In this notation we get:

$$L = \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & -\left(\sum_{i=1}^n Y_i X_i\right)\left(\sum_{i=1}^n X_i^2\right)^{-1} \\ 0 & I \end{bmatrix} \tag{73}$$

so that $A_{12}A_{22}^{-1}$ is the OLS estimator of $\beta$. Moreover, if we use matrix notation that $A_{11} = \sum_{i=1}^n Y_i^2 = Y'Y$ and correspondingly for $A_{12} = \sum_{i=1}^n Y_i X_i = Y'X$ and $A_{22} = \sum_{i=1}^n X_i^2 = X'X$, we see that:

$$LAL' = \begin{bmatrix} A_{11\cdot 2} & 0 \\ 0 & A_{22} \end{bmatrix} \tag{74}$$

$$\begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{bmatrix} \tag{75}$$

$$= \begin{bmatrix} \sum_{i=1}^n Y_i^2 - \left(\sum_{i=1}^n Y_i X_i\right)\left(\sum_{i=1}^n X_i^2\right)^{-1}\left(\sum_{i=1}^n X_i Y_i\right) & 0 \\ 0 & \sum_{i=1}^n X_i^2 \end{bmatrix} \tag{76}$$

$$= \begin{bmatrix} Y'Y - Y'X\left(X'X\right)^{-1}X'Y & 0 \\ 0 & X'X \end{bmatrix} \tag{77}$$

$$= \begin{bmatrix} Y'\left(I - X\left(X'X\right)^{-1}X'\right)Y & 0 \\ 0 & X'X \end{bmatrix} \tag{78}$$

$$= \begin{bmatrix} Y'MMMY & 0 \\ 0 & X'X \end{bmatrix} \tag{79}$$

$$= \begin{bmatrix} e'Me & 0 \\ 0 & X'X \end{bmatrix} \tag{80}$$

where in the last two lines we used the fact that $M^n = M = M' = I - X\left(X'X\right)^{-1}X'$ and that $e = MY$ are the residuals. So $A_{11\cdot 2}$ yields the residual variance.

## Question 3

$$Y_t = \alpha Y_{t-1} + \mu + \epsilon_t \quad t = 1, \dots, T. \tag{81}$$

The least squares estimator for $\alpha$ is, with $\overline{Y}_- = T^{-1} \sum_{t=1}^{T} Y_{t-1}$, equal to

$$\widehat{\alpha} = \frac{\sum_{t=1}^{T} Y_t (Y_{t-1} - \overline{Y}_-)}{\sum_{t=1}^{T} (Y_{t-1} - \overline{Y}_-)^2} \tag{82}$$

Find the asymptotic distribution of $\hat{\alpha}$ when $|\alpha| < 1$ in the following steps.

(a) Argue that $Y_t = X_t + \mu_Y$ where $X_t = \alpha X_{t-1} + \epsilon_t$
Make the substitution on both sides of the equation for $Y_t$

$$\begin{aligned} Y_t = X_t + \mu_Y &= \alpha \left( X_{t-1} + \mu_Y \right) + \mu + \epsilon_t & (83) \\ X_t &= \alpha X_{t-1} + (\alpha \mu_Y - \mu_Y + \mu) + \epsilon_t & (84) \\ &= \alpha X_{t-1} + \epsilon_t & (85) \end{aligned}$$

where $\mu_Y = \mu/(1 - \alpha)$.

(b) Argue that

$$\widehat{\alpha} = \alpha + \frac{\sum_{t=1}^{T} \epsilon_t (X_{t-1} - \overline{X}_-)}{\sum_{t=1}^{T} (X_{t-1} - \overline{X}_-)^2} \tag{86}$$

Replace $(Y_{t-1} - \overline{Y}_-)$ by $(X_{t-1} - \overline{X}_-)$ and $Y_t$ by ;$X_t + \mu_Y = \alpha X_{t-1} + \epsilon_t + \mu_Y$ to get

$$\begin{aligned} \widehat{\alpha} &= \frac{\sum_{t=1}^{T} Y_t (Y_{t-1} - \overline{Y}_-)}{\sum_{t=1}^{T} (Y_{t-1} - \overline{Y}_-)^2} & (87) \\ &= \frac{\sum_{t=1}^{T} (\alpha X_{t-1} + \epsilon_t + \mu_Y)(X_{t-1} - \overline{X}_-)}{\sum_{t=1}^{T} (X_{t-1} - \overline{X}_-)^2} & (88) \\ &= \frac{\alpha \sum_{t=1}^{T} X_{t-1} (X_{t-1} - \overline{X}_-) + \sum_{t=1}^{T} \left( \epsilon_t + \mu_Y \right) \left( X_{t-1} - \overline{X}_- \right)}{\sum_{t=1}^{T} (X_{t-1} - \overline{X}_-)^2} & (89) \end{aligned}$$

We need to make two observations here. First note that $\sum_{t=1}^{T} (X_{t-1} - \overline{X}_-) = 0$:

$$\begin{aligned} \sum_{t=1}^{T} (X_{t-1} - \overline{X}_-) &= \sum_{t=1}^{T} X_{t-1} - \sum_{t=1}^{T} \overline{X}_- & (90) \\ &= T\overline{X}_- - T\overline{X}_- & (91) \\ &= 0 & (92) \end{aligned}$$

Secondly, $\sum_{t=1}^{T} X_{t-1}(X_{t-1} - \overline{X}_-) = \sum_{t=1}^{T} (X_{t-1} - \overline{X}_-)^2$, which is slightly more involved:

$$\begin{aligned}
\sum_{t=1}^{T} X_{t-1}(X_{t-1}-\overline{X}_{-}) &= \sum_{t=1}^{T} X_{t-1}^2 - \sum_{t=1}^{T} X_{t-1}\overline{X}_{-} \\
&= \sum_{t=1}^{T}\left(X_{t-1}^2 - X_{t-1}\frac{1}{T}\sum_{t=1}^{T} X_{t-1}\right) \\
&= \sum_{t=1}^{T} X_{t-1}^2 - \frac{1}{T}\sum_{t=1}^{T} X_{t-1}\left(2\sum_{t=1}^{T} X_{t-1} - \sum_{t=1}^{T} X_{t-1}\right) \\
&= \sum_{t=1}^{T} X_{t-1}^2 - \frac{2}{T}\sum_{t=1}^{T} X_{t-1}\sum_{t=1}^{T} X_{t-1} + \frac{1}{T}\left(\sum_{t=1}^{T} X_{t-1}\right)^2 \\
&= \sum_{t=1}^{T} X_{t-1}^2 - 2\sum_{t=1}^{T} X_{t-1}\left(\frac{1}{T}\sum_{t=1}^{T} X_{t-1}\right) + \overline{X}_{-}^2 \\
&= \sum_{t=1}^{T} X_{t-1}^2 - 2\sum_{t=1}^{T} X_{t-1}\overline{X}_{-} + \overline{X}_{-}^2 \\
&= \sum_{t=1}^{T} \left(X_{t-1}-\overline{X}_{-}\right)^2
\end{aligned}$$

it follows that

$$\widehat{\alpha} = \alpha + \frac{\sum_{t=1}^{T} \epsilon_t (X_{t-1}-\overline{X}_{-})}{\sum_{t=1}^{T}(X_{t-1}-\overline{X}_{-})^2} \tag{93}$$

(c) Argue that

$$\widehat{\alpha} - \alpha = \frac{\sum_{t=1}^{T} \epsilon_t X_{t-1} - T\overline{\epsilon}\overline{X}_{-}}{\sum_{t=1}^{T} X_{t-1}^2 - T(\overline{X}_{-})^2} \tag{94}$$

Standard expression for variance shows

$$\begin{aligned}
\widehat{\alpha} - \alpha &= \frac{\sum_{t=1}^{T} \epsilon_t (X_{t-1}-\overline{X}_{-})}{\sum_{t=1}^{T}(X_{t-1}-\overline{X}_{-})^2} && (95) \\
&= \frac{\sum_{t=1}^{T} \epsilon_t X_{t-1} - \sum_{t=1}^{T} \epsilon_t \overline{X}_{-}}{\sum_{t=1}^{T} X_{t-1}(X_{t-1}-\overline{X}_{-})} && (96) \\
&= \frac{\sum_{t=1}^{T} \epsilon_t X_{t-1} - T\overline{\epsilon}\overline{X}_{-}}{\sum_{t=1}^{T} X_{t-1}^2 - \sum_{t=1}^{T} X_{t-1}\frac{1}{T}\sum_{t=1}^{T} X_{t-1}} && (97) \\
&= \frac{\sum_{t=1}^{T} \epsilon_t X_{t-1} - T\overline{\epsilon}\overline{X}_{-}}{\sum_{t=1}^{T} X_{t-1}^2 - T\left(\overline{X}_{-}\right)^2} && (98)
\end{aligned}$$

(d) We have to find the limiting distributions of the four components in the equation for $\widehat{\alpha} - \alpha$. We will use the following results:

(i) $\sqrt{T}\overline{\epsilon} = \sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\epsilon_{t-1} \xrightarrow{D} \mathsf{N}\left(0,\sigma^2\right)$, as the distribution of the residuals.

(ii) $T^{-1}\sum_{t=1}^{T} X_{t-1}^2 \xrightarrow{\mathsf{P}} \frac{\sigma^2}{(1-\alpha)^2} = \sigma_X^2$, this is the result from Theorem 4.5 in the notes.

(iii) $\sqrt{T}\overline{X}_- = \sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T} X_{t-1}\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_{t-1} \xrightarrow{\mathsf{D}} \mathsf{N}\left(0, \frac{\sigma^2}{(1-\alpha)^2}\right) = \mathsf{N}\left(0, \sigma_X^2\right)$, this is the result that follows from Theorem 4.6 in the notes.

(iv) $T^{-1/2}\sum_{t=1}^{T} X_{t-1}\epsilon_t \xrightarrow{\mathsf{D}} \mathsf{N}\left(0, \frac{\sigma^4}{(1-\alpha)^2}\right) = \mathsf{N}\left(0, \sigma_X^2\sigma^2\right)$, which is also from Theorem 4.6.

Putting all these results together in the above formula for $\hat{\alpha} - \alpha$, we see that:

$$\sqrt{T}(\widehat{\alpha} - \alpha) = \sqrt{T}\left(\frac{\sum_{t=1}^{T}\epsilon_t X_{t-1} - T\overline{\epsilon}\overline{X}_-}{\sum_{t=1}^{T} X_{t-1}^2 - T\left(\overline{X}_-\right)^2}\right) \tag{99}$$

$$= \frac{T^{-1/2}\left(\sum_{t=1}^{T}\epsilon_t X_{t-1}\right) - T^{-1/2}\left(T^{1/2}\overline{\epsilon}\right)\left(T^{1/2}\overline{X}_-\right)}{T^{-1}\left(\sum_{t=1}^{T} X_{t-1}^2\right) - T^{-1}\left(T^{1/2}\overline{X}_-\right)^2} \tag{100}$$

$$\xrightarrow{\mathsf{D}} \frac{\mathsf{N}\left(0, \sigma_X^2\sigma^2\right) - 0 \times \mathsf{N}\left(0, \sigma^2\right) \times \mathsf{N}\left(0, \frac{\sigma^2}{(1-\alpha)^2}\right)}{\sigma_X^2 - 0 \times \left\{\mathsf{N}\left(0, \frac{\sigma^2}{(1-\alpha)^2}\right)\right\}^2} = \mathsf{N}\left(0, \frac{\sigma^2}{\sigma_X^2}\right) \tag{101}$$

This is the same limit as for the dynamic model without intercept in (4.21). The reason is that the stationary autoregressive zero intercept component $X_t$ and the intercept are asymptotically orthogonal (uncorrelated), see (4.31).