Econometrics II - Pset 2

Mark McAvoy

February 27, 2018

Problem 1

Let $(y_1, y_2, ..., y_T)$ be a sample of size T drawn from an iid $N(\mu, \sigma^2)$ distribution

a) Show that the MLE estimates are given by

$$\hat{\mu} = \frac{\sum_{t=1}^{T} y_t}{T} \tag{1}$$

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{T}(y_t - \hat{\mu})}{T} \tag{2}$$

For an iid $N(\mu, \sigma^2)$ distribution the exact Maximum Likelihood Estimator (MLE) with $\theta = (\mu, \sigma^2)$ is given by:

$$L(\theta) \equiv f_{Y_T, Y_{T-1}, \ldots, Y_1}(y_T, y_{T-1}, \ldots, y_1; \theta) = \prod_{t=1}^{T} f_{Y_t|Y_{t-1}}(y_t|y_{t-1}; \theta) \tag{Δ}$$

By our iid assumption. And where

$$f_{Y_t|Y_{t-1}}(y_t|y_{t-1}; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(y_t - \mu)^2}{2\sigma^2}\right) \tag{$\Delta\Delta$}$$

by our gaussian assumption. Combining equations (Δ) and ($\Delta\Delta$) gives us

$$MLE = (2\pi\sigma^2)^{-T/2} exp\left(-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t - \mu)^2\right)$$

then the log likelihood is

$$\mathcal{L}(\theta) \equiv log(L(\theta)) = -\frac{T}{2} log(2\pi) - \frac{T}{2} log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t - \mu)^2 \tag{☾}$$

$\hat{\mu}$ solves the maximization of $L(\theta)$, equivalently as a maximum is invariant up to linear transformation, $\hat{\mu}$ also solves the maximization problem in $\mathcal{L}(\theta)$. Hence we will find $\hat{\mu}$ and $\hat{\sigma}^2$ from considering $\partial\mathcal{L}(\theta)/\partial\mu$ and $\partial\mathcal{L}(\theta)/\partial\sigma^2$, respectively.

$$\left.\frac{\partial \mathcal{L}(\theta)}{\partial \mu}\right|_{\theta=\hat{\theta}} : -\frac{2}{2\sigma^2}\sum_{t=1}^{T}(y_t-\hat{\mu})(-1)=0$$

$$\Rightarrow \hat{\mu}=\frac{1}{T}\sum_{t=1}^{T} y_t$$

$$\left.\frac{\partial \mathcal{L}(\theta)}{\partial \sigma^2}\right|_{\theta=\hat{\theta}} : -\frac{T}{2}\frac{1}{\hat{\sigma}^2}+\frac{1}{2\hat{\sigma}^4}\sum_{t=1}^{T}(y_t-\hat{\mu})^2=0$$

$$\Rightarrow \hat{\sigma}^2=\frac{1}{T}\sum_{t=1}^{T}(y_t-\hat{\mu})^2$$

b) Show that the matrix $\hat{\mathcal{I}}_{2D}$ defined in Hamilton's equation 5.8.2 is given by

$$\hat{I}_{2D}=\begin{bmatrix} 1/\hat{\sigma}^2 & 0 \\ 0 & 1/(2\hat{\sigma}^4) \end{bmatrix} \tag{3}$$

We estimate the second derivative of th information matrix with

$$\hat{\mathcal{I}}_{2D}=-T^{-1}\left.\frac{\partial^2 \mathcal{L}(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\hat{\theta}}$$

Where $\mathcal{L}(\theta)$ is given by equation (☾); we have previously found the first derivatives of θ, now let us find the second derivatives:

$$\left.\frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu^2}\right|_{\theta=\hat{\theta}} = -\frac{1}{\hat{\sigma}^2}\sum_{t=1}^{T}(1) = -\frac{T}{\hat{\sigma}^2}$$

$$\Rightarrow -T^{-1}\frac{-T}{\hat{\sigma}^2}=\frac{1}{\hat{\sigma}^2}$$

$$\left.\frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu\partial\sigma^2}\right|_{\hat{\theta}} = -\frac{1}{\sigma^4}\sum_{t=1}^{T}(y_t-\hat{\mu})$$

$$\Rightarrow -T^{-1}\left(-\frac{1}{\sigma^4}\sum_{t=1}^{T}(y_t-\hat{\mu})\right)=\frac{\hat{\mu}-\hat{\mu}}{\sigma^4}=0$$

$$\left.\frac{\partial^2 \mathcal{L}(\theta)}{\partial \sigma^2\partial\mu}\right|_{\hat{\theta}} = -\frac{2}{2\hat{\sigma}^4}\sum_{t=1}^{T}(y_t-\hat{\mu})$$

$$\Rightarrow -T^{-1}\left(-\frac{1}{\sigma^4}\sum_{t=1}^{T}(y_t-\hat{\mu})\right)=\frac{\hat{\mu}-\hat{\mu}}{\sigma^4}=0$$

$$\left.\frac{\partial^2 \mathcal{L}(\theta)}{(\partial\sigma)^2}\right|_{\theta=\hat{\theta}} = \frac{T}{2}\frac{1}{\hat{\sigma}^4} - \frac{2}{2}\frac{1}{\hat{\sigma}^6}\sum_{t=1}^{T}(y_t - \hat{\mu})^2$$

Recalling $\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{\mu})^2$ this becomes

$$\Rightarrow -T^{-1}\left(\frac{T}{2\hat{\sigma}^4} - \frac{T\hat{\sigma}^2}{\hat{\sigma}^6}\right) = \frac{1}{2\hat{\sigma}^4}$$

Putting these values into a matrix with μ the first ordinate and σ^2 the second, will give the desired form of the information matrix

c) Show that the variance covariance matrix based on the second derivative estimate $\hat{\mathcal{I}}_{2D}$ (See Hamilton's eqs. 5.8.1 and 5.8.3) is given by

$$E\left(\hat{\theta} - \theta\right)\left(\hat{\theta} - \theta\right)' = \begin{bmatrix} \hat{\sigma}^2/T & 0 \\ 0 & 2\hat{\sigma}^4/T \end{bmatrix} \tag{4}$$

We could either tediously perform the same algebra as in (b) to show this, or we can can simply observe equations (5.8.1) and (5.8.3) in the text, starting with the former, since:

$$\hat{\theta} \simeq N(\theta_0, T^{-1}\mathcal{I}^{-1})$$

then,

$$\hat{\theta} - \theta_0 \simeq N(0, T^{-1}\mathcal{I}^{-1})$$

such that

$$E\left(\hat{\theta} - \theta\right)\left(\hat{\theta} - \theta\right)' = T^{-1}\mathcal{I}^{-1}$$

plug in our estimate of $\mathcal{I}$ as $\mathcal{I}_{2d}$

$$= T^{-1}\left(\hat{\mathcal{I}}_{2d}\right)^{-1}$$

and use $\hat{\mathcal{I}}_{2d}$ as from (b)

$$= T^{-1}\begin{bmatrix} 1/\hat{\sigma}^2 & 0 \\ 0 & 1/(2\hat{\sigma}^4) \end{bmatrix}^{-1} = T^{-1}\begin{bmatrix} \hat{\sigma}^2 & 0 \\ 0 & (2\hat{\sigma}^4) \end{bmatrix} = \begin{bmatrix} \hat{\sigma}^2/T & 0 \\ 0 & (2\hat{\sigma}^4)/T \end{bmatrix}$$

where we made use of the fact the inverse of a diagonal matrix is found by simply taking the reciprocal of each element

Problem 2

Let $(y_1, y_2, \ldots, y_T)$ be a sample of size T given by the following parametrization:

$$\begin{aligned} y_t &= \alpha + \rho y_{t-1} + u_t \\ u_t &= \psi_1 u_{t-1} + \psi_2 u_{t-1} + \psi_3 u_{t-3} + \varepsilon_t \end{aligned} \tag{5}$$

where $\varepsilon_t \sim iid(0, \sigma^2)$

a) Show that y_t can be alternatively expressed as

$$y_t = \mu + \beta y_{t-1} + \phi_1 \Delta y_{t-1} + \phi_2 \Delta y_{t-2} + \phi_3 \Delta y_{t-3} + \varepsilon_t \tag{6}$$

where $\Delta y_{t-i} = y_{t-i} - y_{t-i-1}$; and (b) find the values for $\phi_1, \phi_2, \phi_3, \mu, \beta$

note from equation the first line of (5) that we can express u_t as

$$\{u_{t-i}\}_{i=0}^{3} = \{y_{t-i} - \alpha - \rho y_{t-i-1}\}_{i=0}^{3} \tag{♁}$$

plug in our values of (♁) into the second line of (5) gives

$$y_t - \alpha - \rho y_{t-1} = \psi_1(y_{t-1} - \alpha - \rho y_{t-2}) + \psi_2(y_{t-2} - \alpha - \rho y_{t-3}) + \psi_3(y_{t-3} - \alpha - \rho y_{t-4}) + \varepsilon_t \tag{☿}$$

Since we want to express the $y's$ in difference form, let us consider the last y-term in (6) first. We solve this by adding zero, and using the method of undetermined coefficients to match the Δy_3 in (☿) with that in equation 6:

$$-\psi_3 \rho y_{t-4}$$

$$= (\psi_3 \rho y_{t-3} - \psi_3 \rho y_{t-3}) - \psi_3 \rho y_{t-4}$$

$$= (-\psi_3 \rho) y_{t-3} + \psi_3 \rho \Delta y_{t-3}$$

such that

$$\phi_3 = \psi_3 \rho$$

Now consider the y_{t-2} and y_{t-3} terms in (☿) and the $-\psi_3 \rho y_{t-3}$ term we added in finding ϕ_3

$$(\psi_2 y_{t-2} - \psi_1 \rho y_{t-2}) + (\psi_3 - \rho\psi_2 - \psi_3 \rho) y_{t-3}$$

$$= (\psi_2 y_{t-2} - \psi_1 \rho y_{t-2}) + (\psi_3 - \rho\psi_2 - \psi_3\rho) y_{t-2} - (\psi_3 - \rho\psi_2 - \psi_3\rho)\Delta y_{t-2}$$

such that

$$\phi_2 = -(\psi_3 - \rho\psi_2 - \psi_3\rho)$$

Continuing in this fashion, we search for ϕ_1 by looking at the values in (☿) and that which we have added in finding ϕ_2

$$(\psi_1 + \rho)y_{t-1} + (\psi_2 - \psi_1\rho)y_{t-2} + (\psi_3 - \rho\psi_2 - \psi_3\rho)y_{t-2}$$

$$= (\psi_1+\rho)y_{t-1}+(\psi_2-\psi_1\rho+\psi_3-\rho\psi_2-\psi_3\rho)y_{t-1}-(\psi_2-\psi_1\rho+\psi_3-\rho\psi_2-\psi_3\rho)\Delta y_{t-1}$$

such that

$$\phi_1 = -(\psi_2 - \psi_1\rho + \psi_3 - \rho\psi_2 - \psi_3\rho)$$

and

$$\beta = (\psi_1 + \rho + \psi_2 - \psi_1\rho + \psi_3 - \rho\psi_2 - \psi_3\rho)$$

with

$$\mu = \alpha(1 - \psi_1 - \psi_2 - \psi_3)$$

Problem 3

Consider the following stationary data generation process for a random variable y_t,

$$y_t = \beta y_{t-1} + e_t \tag{7}$$

with $e_t \sim iid\ N(0, \sigma^2), |\beta| < 1$, and $y_0 \sim N(0, (1-\beta^2))$. The conditional density for observation t is therefore

$$log\ f(y_t|y_{t-1}; \beta\sigma^2) = -\frac{1}{2}log(2\pi) - \frac{1}{2}log(\sigma^2) - \frac{(y_t - \beta y_{t-1})^2}{2\sigma^2} \tag{8}$$

Let $\hat{\theta} = (\hat{\beta}, \hat{\sigma}^2)$ be the MLE estimates of $\theta = (\beta, \sigma^2)'$

a) Verify that $\hat{\beta}$ minimizes the sum of squared residuals so that it is the same as the OLS estimator. What assumptions did you have to make regarding $f(y_0|\beta, \sigma^2)$?

Taking the exponent of (8) and then product of (8) over all $t \in \{0, 1, \dots, T\}$ gives us the likelihood $\mathcal{L}(\theta)$,

$$L(\theta) \equiv \prod_{t=0}^{T} f_{Y_T, Y_{T-1}, \dots, Y_0}(y_T, y_{T-1}, \dots, y_0; \theta)$$

conditioning on the first observation (y_0) gives

$$= f_{Y_0}(y_0; \theta) \prod_{t=1}^{T} f_{Y_T, Y_{T-1}, \dots, Y_1|Y_0}(y_T, y_{T-1}, \dots, y_1|y_0; \theta)$$

Since $y_0 \sim N(0, (1-\beta^2))$, its pdf is given by

$$f_{Y_0}(y_0; \theta) = \frac{1}{\sqrt{2\pi(1-\beta^2)}} exp\left(-\frac{y_0^2}{2\sigma^2}\right)$$

Next consider the log of $L(\theta)$ defined as $\mathcal{L}(\theta)$,

$$\mathcal{L}(\theta) = -\frac{1}{2}log(2\pi(1-\beta^2)) + \left(\frac{y_0^2}{2\sigma^2}\right) + -\frac{T}{2}log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t - \beta y_{t-1})^2$$

Then to find $\hat{\beta}$ take the derivative of $\mathcal{L}(\theta)$ with respect to β and evaluate at our estimates of parameters

$$\left.\frac{\partial \mathcal{L}(\theta)}{\partial \beta}\right|_{\theta=\hat{\theta}} : \frac{2\hat{\beta}}{2(1-\hat{\beta}^2)} + \frac{2}{2\sigma^2}\sum_{t=1}^{T}(y_t - \hat{\beta} y_{t-1})(y_{t-1}) = 0$$

Which actually has no nice closed form solution for $\hat{\beta}$, so instead, condition having observed y_0 such that $f_{Y_0}(y_0; \theta) = 1$, then we have the conditional log likelihood

$$\left.\frac{\partial \mathcal{L}(\theta|y_0)}{\partial \beta}\right|_{\theta=\hat{\theta}} : \frac{2}{2\sigma^2}\sum_{t=1}^{T}(y_t - \hat{\beta}y_{t-1})(y_{t-1}) = 0$$

$$\Rightarrow \hat{\beta} = \frac{\sum_{t=1}^{T} y_t y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2}$$

Which is the same as an OLS estimate. To check this, the OLS estimate uses the assumption $E(e_t x_t) = 0$ where x_t are the covariates in a linear regression, here we have $x_t = y_{t-1}$, so multiply (7) by y_{t-1} and take expectations gives

$$E[y_t y_{t-1}] = \beta E[y_{t-1}^2] + E[e_t y_{t-1}]$$

$$\Rightarrow \beta = \frac{E[y_{t-1}^2]}{E[y_t y_{t-1}]}$$

with

$$\hat{\beta}_{OLS} = \frac{1/T\sum_{t=1}^{T} y_t y_{t-1}}{1/T\sum_{t=1}^{T} y_{t-1}^2}$$

Where we find equivalence between $\hat{\beta}_{MLE}$ and $\hat{\beta}_{OLS}$ by simply multiplying by T/T in $\hat{\beta}_{MLE}$

We had to make the assumption y_0 was observed before any other $y_t, t \in \{1, 2, \ldots, T\}$ in our conditioning.

b) Write down the exact likelihood function (which includes y_0

Well I'm not sure how we could have considered (a) without having considered this, espcially since it is what we started with, but I'll rewrite this here:

$$L(\theta) \equiv \prod_{t=0}^{T} f_{Y_T, Y_{T-1}, \ldots, Y_0}(y_T, y_{T-1}, \ldots, y_0; \theta)$$

$$= f_{Y_0}(y_0;\theta)\prod_{t=1}^{T} f_{Y_T, Y_{T-1}, \ldots, Y_1|Y_0}(y_T, y_{T-1}, \ldots, y_1|y_0; \theta)$$

$$= \frac{1}{\sqrt{2\pi(1-\beta^2)}} exp\left(-\frac{y_0^2}{2\sigma^2}\right)\left(\frac{T}{\sqrt{2\pi\sigma^2}}\right)\prod_{t=1}^{T} exp\left(-\frac{(y_t - \beta y_{t-1})^2}{2\sigma^2}\right)$$

c) Let

$$Q_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} log\ f(y_t|y_{t-1}; \beta, \sigma^2) \tag{9}$$

Show that

$$Q_T(\hat{\theta}) = -\frac{1}{2}log(2\pi) - \frac{1}{2} - \frac{1}{2}log\left(\frac{SSR}{T}\right) \tag{10}$$

where $SSR = \sum_{t=1}^{T}(y_t - \hat{\beta}y_{t-1})^2$. ***Hint: show that*** $\hat{\sigma}^2 = \frac{SSR}{T}$

We have already found $\hat{\beta}_{MLE}$ in (b), so lets next find $\hat{\sigma}^2_{MLE}$

$$\left.\frac{\partial\mathcal{L}(\theta|y_0)}{\partial\sigma^2}\right|_{\theta=\hat{\theta}} : -\frac{T}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\sum_{t=1}^{T}(y_t - \hat{\beta}y_{t-1})^2 = 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{\beta}y_{t-1})^2 = \frac{SSR}{T}$$

Next, lets recall

$$f(y_t|y_{t-1};\beta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}exp\left(-\frac{(y_t - \beta y_{t-1})^2}{2\sigma^2}\right)$$

so simply plug in $\hat{\beta}$ for β and $\hat{\sigma}^2$ for σ^2 which gives

$$f(y_t|y_{t-1};\hat{\beta},\hat{\sigma}^)) = \frac{1}{\sqrt{2\pi SSR/T}}exp\left(-\frac{(y_t - \hat{\beta}y_{t-1})^2}{2SSR/T}\right)$$

then

$$Q_T(\hat{\theta}) = \frac{1}{T}\left(-\frac{T}{2}log(2\pi) - \frac{T}{2}log\left(\frac{SSR}{T}\right) - \frac{1}{2SSR}\sum_{t=1}^{T}(y_t - \hat{\beta}y_{t-1})^2\right)$$

$$Q_T(\hat{\theta}) = -\frac{1}{2}log(2\pi) - \frac{1}{2}log\left(\frac{SSR}{T}\right) - \frac{1}{2SSR}\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{\beta}y_{t-1})^2$$

$$Q_T(\hat{\theta}) = -\frac{1}{2}log(2\pi) - \frac{1}{2}log\left(\frac{SSR}{T}\right) - \frac{1}{2SSR}SSR$$

$$Q_T(\hat{\theta}) = -\frac{1}{2}log(2\pi) - \frac{1}{2}log\left(\frac{SSR}{T}\right) - \frac{1}{2}$$

Problem 4

Consider the following model for y_t:

$$y_t = x_t'\beta + u_t \tag{11}$$

$$u_t = \rho u_{t-1} + \varepsilon_t \tag{12}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. Note that the econometrician's task is to jointly estimate all elements of the parameter vector θ, where $\theta = (\beta, \rho, \sigma^2)$. We have discussed in class two potential approaches for estimating all elements of the vector θ. The first approach relies on the iterative Cochrane Orcutt procedure, which we have discussed in class. The alternative approach, which we have also discussed in class, is conditional MLE.

a) Write a MATLAB routine to estimate the model specified in (11)-(12) via conditional MLE and confirm that the estimates you obtain are in line with the results from the Cochrane Orcutt procedure. Experiment with different sample sizes in your simulations. What do you notice as the sample gets larger?

We begin with our data generating process (DGP), we initialize $\rho = 0.4$ in equation (12) and ε_t an $(T \times 1)$ vector of $N(0,1)$ i.i.d variables. Using this process and x_t constructed in the same way as ε_t (s.t. we have only one covariate and no intercept), we construct the process for y_t; note we also set the seed for reproducibility. We take $T = 1000$ and truncate the first 100 observations in our DGP (to remove our initialization effect).

Maximum Likelihood Estimation:

As $\{u_t\}$ is an AR(1) process (with $\rho \in (0,1)$) and $\varepsilon_t \sim WN(0, \sigma^2)$, we know the covariance matrix takes the form:

$$E(\mathbf{uu}') = \frac{\sigma^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix} = \sigma^2 \mathbf{V} \tag{♂}$$

We can see $\mathbf{V}$ is a $(T \times T)$ symmetric and positive definite, then there exists a nonsingular $(T \times T)$ matrix $\mathbf{L}$ such that

$$\mathbf{V}^{-1} = \mathbf{L}'\mathbf{L}$$

In our example $\mathbf{L}$ takes the form:

$$\mathbf{L} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \tag{⚨}$$

Which is lower triangular. Using Generalized Least Squares procedure, transform our variables by pre-multiplying by $\mathbf{L}$, and taking conditional MLE (i.e. omitting the first observation in $\{y, x, u\}$ gives us our log-likelehood function,

$$L(\beta, \rho, \sigma^2) = -\frac{T-1}{2} log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^{T} ((y_t - x_t'\beta) - \rho(y_{t-1} - x_{t-1}'\beta))^2$$

We use "fminunc" to maximize this across $\{\beta, \sigma\}$, with initial values taken from $U(0,1)$ for $\{\beta, \rho, \sigma\}$ values are given in our table 1 at the end of this section.

Cochrane-Orcutt Procedure
We solve this procedure in two ways.

Method 1:
We construct our $\mathbf{V}$ matrix given in equation (♂), and use Cholesky decomposition on $\mathbf{V}$ to check that $\mathbf{L}$ matches that in (♁), it does almost .. it has the points $(1, 1)$ and (T, T) swapped, we couldn't figure out a way to replace these so we hope it has a marginal effect and continue regardless. (Step 1): To perform the procedure we initialize with a guess $\rho = r^{(1)} = 0$. (Step 2): We use GLS to transform to

$$\{\mathbf{y}, \mathbf{x}, \mathbf{u}\} \rightarrow \{L\mathbf{y}, L\mathbf{x}, L\mathbf{u}\}$$

(Step 3): where OLS will give us $\hat{\beta}$ in

$$\hat{\beta}^{(i)} = (L\mathbf{x}'L\mathbf{x})^{-1}L\mathbf{x}'L\mathbf{y}$$

(Step 4): We find our new residuals from this

$$\mathbf{u} = \mathbf{y} - \mathbf{x}'\hat{\beta}$$

(Step 5): Using this β we guess our next estimate for ρ

$$r^{(i+1)} = (\mathbf{u}_{t-1}'\mathbf{u}_{t-1}^{-1}(\mathbf{u}_{t-1}'\mathbf{u}_t)$$

(Step 6): We continue by returning to (Step 1) and plugging in our new guess for ρ, we continue until

$$|\beta^{(i)} - \beta^{(i+1)}| < \varepsilon \;\;\&\;\; |r^{(i)} - r^{(i+1)}| < \varepsilon$$

Lastly we find our sample standard deviation ($\hat{\sigma}$) from

$$\hat{\sigma}^2 = \frac{1}{T-1} \sum_{i=1}^{T} (Ly_t - \beta Lx_t)^2$$

Values of this procedure we include in the $C - O - 1$ row

Method 2:
This method is much simpler, we use our process $\{x, y\}$ and put this in the

`cochrane_orcutt`

function the professor posted, these values are given in the $C-O-2$ row

Figure 1: Table 1: Parameters in various procedures

Process	β	ρ	σ
MLE	0.99402	0.43681	1.0137
C-O-1	0.99486	0.43679	1.0146
C-O-2	0.99402	0.43677	1.0137

Values across the three method are close! If we use a smaller sample size, such as our initial test size of $T = 10$ we have close-ish values; but these are much closer for larger T.

b) Next write a Matlab routine that implements the Cochrane Orcutt estima- tion procedure for the case where the residuals follow an AR(2) process, that is:

$$y_t = x_t'\beta + u_t \tag{13}$$

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-1} + \varepsilon_t \tag{14}$$

with $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. Note that in this case the vector of unknown parameters to estimate becomes $\theta = \{\beta, \rho_1, \rho_2, \sigma^2\}$

We need to construct $\mathbf{V}$, so we need to find the autocovariances in our process, i.e. $\{\gamma_0, \gamma_1, \ldots, \gamma_{T-1}\}$. Let $\mathbf{F}$ be the usual coefficient matrix representation of the AR(2) process eqn (14). It can be shown that the $(p \times 1)$ $\{\gamma_1, \gamma_1, \ldots, \gamma_{p-1}\}'$ is given by the first p elements of the first column of the $(p^2 \times p^2)$ matrix $\sigma^2[\mathbf{I}_{p^2} - (\mathbf{F} \otimes \mathbf{F})]^{-1}$; where $\otimes$ is the Kronecker product. Yule-Walker equations help us find that for an AR(p) process autocovarainces are of the form

$$\gamma_j = \begin{cases} \phi_1\gamma_1 + \phi_2\gamma_2 + \cdots + \phi_p\gamma_p + \sigma^2 & \text{for j = 0} \\ \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} + \cdots + \phi_p\gamma_{j-p} & \text{for j = 1,2,\ldots} \end{cases} \tag{15}$$

Since we have found $\{\gamma_0, \gamma_1\}$ we can find any other up to γ_{T-1}. Now that we have $\mathbf{V}$ we find $\mathbf{L}$ as we did before, and carry out our analysis as before. The specific means can be seen in our Matlab code.

c) Test that your Cochrane-Orcutt code is working properly by generating an artificial process matching the new DGP in (13)-(14) and confirming that your code is able to retrieve the parameter values that you used to simulate your artificial process.

We compare our Cochrane-Orcutt procedure with an MLE/GLS procedure, where our likelihood function takes the form:

$$L(\beta,\rho_1,\rho_2,\sigma^2) = -\frac{T-1}{2}log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=3}^{T}((y_t - x_t'\beta) - \rho_1(y_{t-1} - x_{t-1}'\beta) - \rho_2(y_{t-2} - x_{t-2}'\beta))^2$$

Using the same procedure as in (a) gives us:

Figure 2: Table 1: Parameters in various procedures

Process	β	ρ_1	ρ_2	σ^2
MLE	1.0036	0.44855	0.35824	1.0131
C-O	1.0042	0.44853	0.35827	1.0152

Quite close again! And close to the parameter values of $\{\beta,\rho_1,\rho_2,\sigma^2\} = \{1.0, 0.4, 0.4, 1.0\}$ we used to simulate the process.