Time Series Models
Supervision 1

Samuel Lee

Question 1

Depending on whether the data is more frequent than annual, and if so, whether the data is seasonally adjusted, I might first de-seasonalise the growth rates by keeping the residuals from the regression of GDP growth on a set of seasonal dummies (probably months or quarters). Then, I might want to test whether the de-seasonalised data has a unit root with something like the Augmented Dickey-Fuller or KPSS tests. If there are reasons to suspect that there is a unit root (and the reasons can also come from economic theory), I would take first-differences of the data. I can then fit several ARMA models of varying lag orders to the data, making sure all the models are estimated on the same length of data, and also with or without a trend or intercept. probably through maximum-likelihood estimation. Finally I would select a model using an information criterion, and inspect the autocorrelograms to make sure there's nothing funny going on.

Question 2

For a covariance-stationary process, the autocorrelation function $\rho(\tau)$ is equal to $\frac{\gamma(\tau)}{\gamma(0)}$ where $\gamma(\tau)$ represents the autocovariance at lag τ. If y_t follows an MA(1) process with lag coefficient θ, we have

$$\begin{aligned}
y_t &= \varepsilon_t + \theta\varepsilon_{t-1} \\
\gamma(\tau) &= \text{Cov}[y_t, y_{t-\tau}] = \text{Cov}[\varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_{t-\tau} + \theta\varepsilon_{t-\tau-1}] \\
&= \begin{cases} (1+\theta^2)\sigma^2 & \text{if } \tau = 0 \\ \theta\sigma^2 & \text{if } \tau = \pm 1 \\ 0 & \text{otherwise} \end{cases} \\
\rho(\tau) &= \frac{\gamma(\tau)}{\gamma(0)} = \begin{cases} 1 & \text{if } \tau = 0 \\ \frac{\theta}{1+\theta^2} & \text{if } \tau = \pm 1 \end{cases}
\end{aligned}$$

We can see that if the lag coefficient were $\frac{1}{\theta}$,

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)} = \begin{cases} 1 & \text{if } \tau = 0 \\ \frac{\frac{1}{\theta}}{1+\left(\frac{1}{\theta}\right)^2} & \text{if } \tau = \pm 1 \end{cases}$$

And since $\frac{\frac{1}{\theta}}{1+\left(\frac{1}{\theta}\right)^2} = \frac{\frac{1}{\theta}}{\frac{1+\theta^2}{\theta^2}} = \frac{1}{1+\theta^2}$, the two autocorrelation functions are identical.

Question 3

We have a process y_t comprised of two AR(1) processes $z_{1,t}$ and $z_{2,t}$ with parameters λ_1 and λ_2. We should be able to show that y_t is an ARMA(2,1) process by torturing the expression for $z_{1,t} + z_{2,t}$ until it looks ARMA(2,1).

$$\begin{aligned}
y_t &= z_{1,t} + z_{2,t} \\
y_t &= \lambda_1 z_{1,t-1} + \lambda_2 z_{2,t-1} + \varepsilon_{1,t} + \varepsilon_{2,t} \\
&= \varphi_1 y_{t-1} + (\lambda_1 - \varphi_1) z_{1,t-1} + (\lambda_2 - \varphi_1) z_{2,t-1} + \varepsilon_{1,t} + \varepsilon_{2,t} \\
&= \varphi_1 y_{t-1} + \lambda_1(\lambda_1 - \varphi_1) z_{1,t-2} + \lambda_2(\lambda_2 - \varphi_1) z_{2,t-2} + (\lambda_1 - \varphi_1)\varepsilon_{1,t-1} + (\lambda_2 - \varphi_1)\varepsilon_{t-1} + \varepsilon_{1,t} + \varepsilon_{2,t} \\
&= \varphi_1 y_{t-1} + \varphi_2 y_{t-2} \\
&\quad + \left[\lambda_1(\lambda_1 - \varphi_1) - \varphi_2\right] z_{1,t-2} + \left[\lambda_2(\lambda_2 - \varphi_1) - \varphi_2\right] z_{2,t-2} \\
&\quad + (\lambda_1 - \varphi_1)\varepsilon_{1,t-1} + (\lambda_2 - \varphi_1)\varepsilon_{2,t-1} + \varepsilon_{1,t} + \varepsilon_{2,t}
\end{aligned}$$

We need $\lambda_1(\lambda_1 - \varphi_1) - \varphi_2 = \lambda_2(\lambda_2 - \varphi_1) - \varphi_2 = 0$ to properly separate the AR(2) terms from the unobserved disturbances. (If the coefficients on $z_{1,t-2}$ and $z_{2,t-2}$ were equal to each other but not equal to zero, then their values would just get absorbed into φ_2). As it turns out,

$$\begin{aligned}
\lambda_1(\lambda_1 - \varphi_1) - \varphi_2 &= \lambda_2(\lambda_2 - \varphi_1) - \varphi_2 \\
\lambda_1^2 - \lambda_1\varphi_1 &= \lambda_2^2 - \lambda_2\varphi_1 \\
(\lambda_1 - \lambda_2)\varphi_1 &= \lambda_1^2 - \lambda_2^2 \\
&= (\lambda_1 + \lambda_2)(\lambda_1 - \lambda_2) \\
\varphi_1 &= \lambda_1 + \lambda_2
\end{aligned}$$

and it is trivial to work out that $\varphi_2 = -\lambda_1\lambda_2$ if the coefficients on $z_{1,t-2}$ and $z_{2,t-2}$ are zero. This means we can express y_t as a sum of an AR(2) process with parameters $\lambda_1 + \lambda_2$ and $-\lambda_1\lambda_2$, and a remainder that is a linear combination of the current and lagged shocks to $z_{1,t}$ and $z_{2,t}$. We just have to check that the remainder is an MA(1) process. We have

$$\begin{aligned}
u_t &= (\lambda_1 - \varphi_1)\varepsilon_{1,t-1} + (\lambda_2 - \varphi_1)\varepsilon_{2,t-1} + \varepsilon_{1,t} + \varepsilon_{2,t} \\
&= \varepsilon_{1,t} - \lambda_2\varepsilon_{1,t-1} + \varepsilon_{2,t} - \lambda_1\varepsilon_{2,t-1} \\
\operatorname{Cov}[u_t, u_{t-\tau}] &= \operatorname{Cov}[\varepsilon_{1,t} - \lambda_2\varepsilon_{1,t-1} + \varepsilon_{2,t} - \lambda_1\varepsilon_{2,t-1}, \varepsilon_{1,t-\tau} - \lambda_2\varepsilon_{1,t-\tau-1} + \varepsilon_{2,t-\tau} - \lambda_1\varepsilon_{2,t-\tau-1}] \\
\gamma(\tau) &= \begin{cases} (1+\lambda_2^2)\sigma_1^2 + (1+\lambda_1^2)\sigma_2^2 & \text{if } \tau = 0 \\ -\lambda_2\sigma_1^2 - \lambda_1\sigma_2^2 & \text{if } |\tau| = 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

where $\gamma(\tau)$ is the autocovariance function for u_t. This looks like the autocovariance function of an MA(1) process. If we had the additional information that the disturbances were Gaussian, we could conclude that the process was MA(1), since the mean and autocovariance function together fully characterizes a Gaussian process. In any case, if u_t is MA(1), then we can recover possible values for its MA(1) parameter by solving

$$\begin{aligned}
\frac{\theta}{1+\theta^2} &= \frac{\gamma(1)}{\gamma(0)} \\
&= -\frac{\lambda_2\sigma_1^2 + \lambda_1\sigma_2^2}{(1+\lambda_2^2)\sigma_1^2 + (1+\lambda_1^2)\sigma_2^2} \\
\theta^2 + \frac{(1+\lambda_2^2)\sigma_1^2 + (1+\lambda_1^2)\sigma_2^2}{\lambda_2\sigma_1^2 + \lambda_1\sigma_2^2}\theta + 1 &= 0
\end{aligned}$$

The above is a quadratic equation in θ, and it has real roots if the discriminant is positive:

$$\begin{aligned}
\left[\frac{(1+\lambda_2^2)\sigma_1^2 + (1+\lambda_1^2)\sigma_2^2}{\lambda_2\sigma_1^2 + \lambda_1\sigma_2^2}\right]^2 - 4 &> 0 \\
\left|\frac{(1+\lambda_2^2)\sigma_1^2 + (1+\lambda_1^2)\sigma_2^2}{\lambda_2\sigma_1^2 + \lambda_1\sigma_2^2}\right| &> 2 \\
(1+\lambda_2^2)\sigma_1^2 + (1+\lambda_1^2)\sigma_2^2 &> 2\left|\lambda_2\sigma_1^2 + \lambda_1\sigma_2^2\right| \\
\left(\lambda_2^2 + 2\left|\lambda_2\right| + 1\right)\sigma_1^2 + \left(\lambda_1^2 + 2\left|\lambda_1\right| + 1\right)\sigma_2^2 &> 0 \\
\left(\left|\lambda_2\right| + 1\right)^2\sigma_1^2 + \left(\left|\lambda_1\right| + 1\right)^2\sigma_2^2 &> 0
\end{aligned}$$

which is always satisfied. So we have shown that the real roots exist, and curious minds can solve for the possible values of θ by simply applying the quadratic formula. We might be interested in verifying if the two roots identified are reciprocals of one another, given that we have shown that the autocorrelation function of an MA(1) process with parameter θ is the same as that of an MA(1) process with parameter $\frac{1}{\theta}$.

Question 4

We could apply a Beveridge-Nelson decomposition; we've estimated that

$$\Delta y_t = 0.002 + 0.633\Delta y_{t-1} + \varepsilon_t - 0.568\varepsilon_{t-1}$$

and by rearranging, we can decompose the series into two parts:

$$\begin{aligned}
\Delta y_t &= 0.002 + 0.633\Delta y_{t-1} + \varepsilon_t - 0.568\varepsilon_{t-1} \\
\Delta y_t - 0.00545 &= 0.633(\Delta y_{t-1} - 0.00545) + \varepsilon_t - 0.568\varepsilon_t \\
&= \frac{1-0.568L}{1-0.633L}\varepsilon_t \\
\Delta y_t &= 0.00545 + \frac{1-0.568L}{1-0.633L}\varepsilon_t \\
&= 0.00545 + \frac{1-0.568}{1-0.633}\varepsilon_t + \left(\frac{1-0.568L}{1-0.633L} - \frac{1-0.568}{1-0.633}\right)\varepsilon_t
\end{aligned}$$

Just a little bit more...

$$
\begin{aligned}
(1-L)y_t &= 0.00545 + \frac{1-0.568}{1-0.633}\varepsilon_t + \left(\frac{(1-0.568L)(1-0.633)}{(1-0.633L)(1-0.633)} - \frac{(1-0.568)(1-0.633L)}{(1-0.633L)(1-0.633)}\right)\varepsilon_t \\
&= 0.00545 + 2.57\varepsilon_t + \left(\frac{(1-0.568L)(1-0.633)-(1-0.568)(1-0.633L)}{(1-0.633L)(1-0.633)}\right)\varepsilon_t \\
&= 0.00545 + 2.57\varepsilon_t + \\
&\left(\frac{1-0.633-0.568L+0.633\times 0.568L-1+0.633L+0.568-0.568\times 0.633L}{(1-0.633L)(1-0.633)}\right)\varepsilon_t \\
&= 0.00545 + 2.57\varepsilon_t + \left(\frac{-0.633(1-L)+0.568(1-L)}{(1-0.633L)(1-0.633)}\right)\varepsilon_t \\
&= 0.00545 + 2.57\varepsilon_t - \frac{0.065}{(1-0.633L)(1-0.633)}(1-L)\varepsilon_t
\end{aligned}
$$

Finally,

$$
\begin{aligned}
y_t &= \frac{0.00545}{1-L} + 2.57\frac{\varepsilon_t}{1-L} - 0.177\times\frac{\varepsilon_t}{1-0.633L} \\
&= \underbrace{0.00545t}_{\text{Deterministic trend}} + y_0 + \underbrace{2.57\sum_{i=0}^{t}\varepsilon_i}_{\text{Stochastic trend}} - \underbrace{0.177\frac{\varepsilon_t}{1-0.633L}}_{\text{Stationary AR(1) cycle}}
\end{aligned}
$$

Question 5

i.

This is essentially the regression one would run to perform an Augmented Dickey-Fuller test with intercept and trend. The null hypothesis that demand has a unit root is rejected if the Dickey-Fuller statistic, the estimated coefficient on dem_{t-1} divided by its standard error, is significantly negative relative to the Dickey-Fuller distribution. In this case the Dickey-Fuller statistic is $\frac{-0.285923}{0.12928} \approx -2.211657$. For a sample size of 50, the 10% critical value for a model run with a trend and intercept is -3.18, which means the null hypothesis is not rejected, and the evidence is not strong enough to conclude that there demand is not I(1).

ii.

The data is quarterly, so estimating fourth differences will take care of both the nonstationarity and any time-invariant seasonal fixed effects. The regressor is essentially year-on-year changes in electricity demand.

iii.

The DW statistic is very close to 2, which means there is little evidence of autocorrelation in the first order in the residuals. (The DW statistic is approximately equal to $2(1-\hat{\rho})$ where $\hat{\rho}$ is the regression of the estimated residuals on their lagged values. $DW \approx 2$ implies $\hat{\rho} \approx 0$.)

iv.

The time trend is negative, which could be a reflection of some secular decline in the growth of electricity demand. Reasons can include slowing population growth (and therefore slowing growth in electricity consumption), greater environmental consciousness, or energy-saving technological growth.

v.

An autocorrelation value of -0.5 at the fourth lag is actually what one should expect if the unforecastable disturbances to the level of electricity demand are independent and identically distributed. When we take the fourth difference, we end up with a composite error term $u_t = \varepsilon_t - \varepsilon_{t-4}$, and the autocorrelation at the fourth lag is

$$\begin{aligned}
\rho(4) &= \frac{\mathrm{Cov}[u_t, u_{t-4}]}{Var[u_t]} \\
&= \frac{\mathrm{Cov}[\varepsilon_t - \varepsilon_{t-4}, \varepsilon_{t-4} - \varepsilon_{t-8}]}{Var[\varepsilon_t - \varepsilon_{t-4}]} \\
&= \frac{-\sigma^2}{2\sigma^2} = -0.5
\end{aligned}$$

If it is true that the disturbances to the levels (and not the first-differences) of electricity demand are i.i.d, then the DW statistic of 0 is to be expected since the Durbin-Watson test only picks up autocorrelations in the residuals of order 1.

vi.

I have no idea. But we can see that the R^2 is much higher in model II, which somewhat makes sense since the quarter-on-quarter changes in demand are much more easy to forecast when there are seasonal effects.

vii.

Forecast long enough, and you get negative forecasts of electricity demand (provided you have a starting point since the model is estimated in first-differences). And the uncertainty of the forecast doesn't increase past a certain point, even though your predicted changes tend toward negative infinity. Part of this is because the forecasts ignore parameter uncertainty.