Problem Set 7 - Solutions

Applied Statistics and Econometrics II
Spring 2018, NYU
Ercan Karadas

[1] a) Since $E[x_t] = \beta_1 + \beta_2 t$, the mean is not constant, so the process is not stationary. Note that

$$\begin{aligned} x_t - x_{t-1} &= \beta_1 + \beta_2 t + w_t - \beta_1 - \beta_2(t-1) - w_{t-1} \\ &= \beta_2 t + w_t - w_{t-1} \end{aligned}$$

which is clearly stationary. Verify that the mean is β_2 and the autocovariance is 2 for $s = 2$ and -1 for $|s-t| = 1$ and is zero for $|s-t| > 1$.

b) First, write

$$\begin{aligned} E[y-t] &= \frac{1}{2q+1} \sum_{j=-q}^{q} [\beta_1 + \beta_2(t-j)] \\ &= \frac{1}{2q+1} \left[(2q+1)(\beta_1 + \beta_2 t) - \beta_2 \sum_{j=-q}^{q} j \right] \\ &= \beta_1 + \beta_2 t \end{aligned}$$

because the positive and negative terms in the last sum cancel out. To get the covariance write the process as

$$y_t = \sum_{-\infty}^{\infty} a_j w_{t-j},$$

where $a_j = 1$, $j = -q, \dots, 0, \dots, q$ and is zero otherwise. To get the covariance, note that we need

$$\begin{aligned} \gamma_h &= E\left[(y_{t+h} - Ey_{t+h})(y_t - Ey_t)\right] \\ &= (2q+1)^{-2} \sum_j \sum_k a_j a_k E w_{t+h-j} w_{t-k} \\ &= \sigma^2 (2q+1)^{-2} \sum_j \sum_k a_j a_k \gamma_{h+k-j} \\ &= \sum_{j=-\infty}^{\infty} a_{j+h} a_j \end{aligned}$$

where $\delta_{h+k-j} = 1$, $j = k+h$ and is zero otherwise. Writing out the terms in γ_h, for $h = 0, \pm 1, \pm 2, \dots$ we obtain

$$\gamma_h = \frac{\sigma^2(2q+1-|h|)}{(2q+1)^2}$$

for $h = 0, \pm 1, \pm 2, \dots, 2q$ and zero for $|h| > q$.

[**2**] By a computation analogous to that appearing in Example 1.17 in the text, we may obtain

$$\gamma_h = \begin{cases} 6\sigma_w^2 & h = 0 \\ 4\sigma_w^2 & h = \pm 1 \\ \sigma_w^2 & h = \pm 2 \\ 0 & |h| > 2 \end{cases}$$

[**3**] a) Simply substitute $\delta s + \sum_{k=1}^{s}$ for x_s to see that

$$\underbrace{\delta t + \sum_{k=1}^{t} w_k}_{x_t} = \delta + \underbrace{\left(\delta(t-1) + \sum_{k=1}^{t-1} w_k\right)}_{x_{t-1}} + w_t$$

Alternately, the result can be shown by induction.

b) For the mean

$$Ex_t = E\left(\delta t + \sum_{k=1}^{t} w_k\right) = \delta t + \sum_{k=1}^{t} Ew_k = \delta t.$$

For the covariance, without loss of generality, consider the case $s \le t$.

$$\begin{aligned} \gamma(s,t) \equiv \text{Cov}(x_s, x_t) &= E\left[(x_s - \delta s)(x_t - \delta t)\right] \\ &= E\left[\sum_{j=1}^{s} w_j \sum_{k=1}^{t} w_k\right] \\ &= \sum_{j=1}^{s} E(w_j^2) = s\sigma_w^2. \qquad [\text{or } \min(s,t)\sigma_w^2] \end{aligned}$$

c) The series is nonstationary because both the mean function and the autocovariance function depend on time, t.

d) From (b),

$$\rho_x(t-1,t) = \frac{(t-1)\sigma_w^2}{\sqrt{(t-1)\sigma_w^2}\sqrt{t\sigma_w^2}}$$

which yield the result. The implication is that the series tends to change slowly.

e) One possibility is to note that $\nabla x_t = x_t - x_{t-1} = \delta + w$, which is stationary because $\mu_{x,t} = \delta$ and $\gamma_x(t+h,t) = \sigma_w^2 \delta_0(h)$ are both independent of time t, where $\delta_0(h)$ is the delta measure.

[**4**] a) For the mean, write

$$Ey_t = E(\exp(x_t)) = \exp\left\{\mu_x + \frac{1}{2}\gamma_x(0)\right\},$$

using the mean equation at $\lambda = 1$.

b) For the autocovariance function, note that

$$\begin{aligned} E(y_{t+h}y_t) &= E(\exp(x_{t+h})\exp(x_t)) \\ &= E(\exp(x_{t+h} + x_t)) \\ &= \exp\{2\mu_x + \gamma_x(0) + \gamma_x(h)\}, \end{aligned}$$

since $x_{t+h} + x_t$ is the sum of two correlated normal random variables and will be normally distributed with mean $2\mu_x$ and variance

$$\gamma_x(0) + \gamma_x(0) + 2\gamma_x(h) = 2(\gamma_x(0) + \gamma_x(h)).$$

For the autocovariance of y_t

$$\begin{aligned} \gamma_h &= E(y_{t+h}y_t) - E(y_{t+h})E(y_t) \\ &= \exp\{2\mu_x + \gamma_x(0) + \gamma_x(h)\} - \left(\exp\left\{\mu_x + \frac{1}{2}\gamma_x(0)\right\}\right)^2 \\ &= \exp\{2\mu_x + \gamma_x(0)\}(\exp\{\gamma_x(h)\} - 1) \end{aligned}$$

[5] a,b) Code for parts (a) and (b) is below. You should have about 1 in 20 ACF values within the bounds, but the values for part (b) will be larger in general than for part (a).

```
wa = rnorm(500,0,1)
wb = rnorm(50,0,1)
par(mfrow=c(2,1))
(acf(wa, 20)) # plot and print results
(acf(wb, 20)) # plot and print results
```

c,d) This is similar to the previous problem. Generate 2 extra observations due to loss of the end points in making the MA.

```
wa = rnorm(502,0,1)
wb = rnorm(52,0,1)
va = filter(wa, sides=2, rep(1,3)/3)
vb = filter(wb, sides=2, rep(1,3)/3)
par(mfrow=c(2,1))
(acf(va, 20, na.action = na.pass))   # plot and print results
(acf(vb, 20, na.action = na.pass))   # plot and print results
```

[6] a) $Ex_t = \beta_0 + \beta_1 t$, the mean depends on t and hence the process in not stationary. (Note that the points will be randomly distributed around a straight line.)

b) Note that $\nabla x_t = \beta_1 + w_t - w_{t-1}$ so that $E(\nabla x_t) = \beta_1$ and

$$\text{cov}(\nabla x_t, \nabla x_{t+h}) = \begin{cases} 2\sigma_w^2 & h = 0 \\ -\sigma_w^2 & h = \pm 1 \\ 0 & |h| > 1 \end{cases}$$

c) Here $\nabla x_t = \beta_1 + y_t - y_{t-1}$ so that $E(\nabla x_t) = \beta_1 + \mu_y - \mu_y = \beta_1$ and

$\text{cov}(\nabla x_t, \nabla x_{t+h}) = \text{cov}(y_{t+h} - y_{t+h-1}, y_t - y_{t-1}) = 2\gamma_y(h) - \gamma_y(h+1) - \gamma_y(h-1)$

which is independent of t.

[7] a) Use induction or insert the solution into the equation, i.e.,

$$\sum_{j=0}^{t} \phi^j w_{t-j} = \phi \left(\sum_{j=0}^{t-1} \phi^j w_{t-1-j} \right) + w_j$$

b) $E[x_t] = \sum_{j=0}^{t} \phi^j E(w_{t-j}) = 0$

c) $\text{Var}[x_t] = \sum_{j=0}^{t} \phi^{2j} \text{Var}(w_{t-j}) = \sigma_w^2 \sum_{j=0}^{t} \phi^{2j} = \frac{\sigma_w^2}{1-\phi^2} \left(1 - \phi^{2(t+1)}\right)$ using the fact that w_t is noise.

d) Notice that

$$\begin{aligned} x_{t+h} &= \sum_{j=0}^{t+h} \phi^j w_{t+h-j} \\ &= \sum_{j=0}^{h-1} \phi^j w_{t+h-j} + \sum_{j=h}^{t+h} \phi^j w_{t+h-j} \\ &= \sum_{j=0}^{h-1} \phi^j w_{t+h-j} + \phi^h \sum_{k=0}^{t} \phi^k w_{t-k} \\ &= \sum_{j=0}^{h-1} \phi^j w_{t+h-j} + \phi^h x_t \end{aligned}$$

Alternately, just iterate x_{t+h} back h time units.
Since x_t involves the w_s for $s \leq t$,

$$\text{Cov}(x_t, x_{t+h}) = \text{Cov}\left(\sum_{j=0}^{h-1} \phi^j w_{t+h-j} + \phi^h x_t, x_t \right) = \phi^h \text{Var}(x_t)$$

e) x_t is not stationary because the (co)variance depends on time t.

f) As $t \to \infty$, $\text{Var}(x_t) \to \frac{\sigma_w^2}{1-\phi^2}$ by part (c), and hence by (d), the autocovariance function is independent of t.

g) Generate more than n observations, for example, generate $n + n_0$ observations, where n_0 is fairly large (like 50), and discard the first n_0.

h) Write

$$x_t = \phi^t x_0 + \sum_{j=0}^{t-1} \phi^j w_{t-j} \, .$$

Now,

$$\begin{aligned}\mathrm{Var}(x_t) &= \frac{\phi^{2t}}{1-\phi^2}\sigma_w^2 + \sigma_w^2\sum_{j=0}^{t-1}\phi^{2j} \\ &= \sigma_w^2\left[\frac{\phi^{2t}}{1-\phi^2} + \frac{1-\phi^{2t}}{1-\phi^2}\right] \\ &= \sigma_w^2\frac{1}{1-\phi^2}\end{aligned}$$

which is independent of t.

[8] a) Write this as $(1-0.3L)(1-0.5L)x_t = (1-0.3L)w_t$ and reduce to $(1-0.5L)x_t = w_t$. Hence the process is a causal and invertible AR(1): $x_t = 0.5x_{t-1} + w_t$.

b) The AR polynomial is $1 - 1z + 0.5z^2$ which has complex roots $1 \pm i$ outside the unit circle (note that $|1 \pm i|^2 = 2$). The MA polynomial is $1 - z$ which has root unity. Thus the process is a causal but not invertible ARMA(2,1).

[9] a) There's not much to do in verifying the calculations... almost everything is done in the example.

b) The ACF distinguishes the MA(1) case but not the ARMA(1,1) or AR(1) cases, which look similar to each other (see the figure below).

```
u1 = ARMAacf(ar=.6, ma=.9, lag.max=10)
u2 = ARMAacf(ar=.6, ma=0, lag.max=10)
u3 = ARMAacf(ar=0,ma=.9, lag.max=10)
plot(u1[-1], type="o", col=1, ylab="ACF", xlab="Lag")
lines(u2[-1], type="o", col=2)
lines(u3[-1], type="o", col=4)
legend("topright", c("ARMA","AR","MA"), col=c(1,2,4), lty=1)
```

c) R code:

```
ar = arima.sim(list(order=c(1,0,0), ar=.6), n=100)
ma = arima.sim(list(order=c(0,0,1), ma=.9), n=100)
arma = arima.sim(list(order=c(1,0,1), ar=.6, ma=.9), n=100)
acf2(ar)
acf2(ma)
acf2(arma)
```

The results should be close to Table 3.1 with randomness taken into consideration.

[10] a) u_t is strictly stationary if $|\alpha| < 1$ and cov. stationary if further $E[v_T^2] < \infty$. u_t is an AR(1) process, so that the ACF is given by $\rho_v(j) = \alpha^j$. Then

$$\lim_{T\to\infty} \mathrm{Var}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t\right].$$

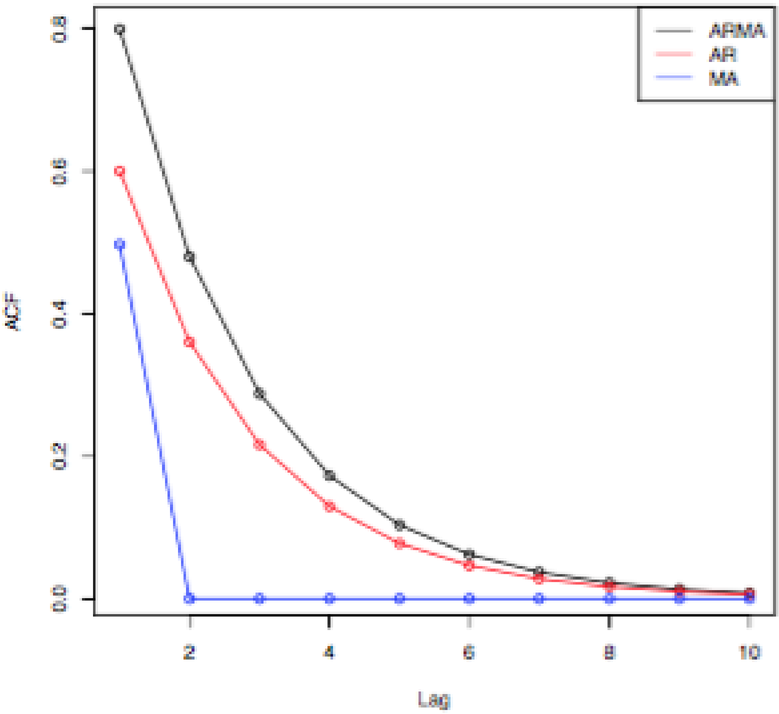

Figure 1: Figure for Problem 9

$$\lim_{T\to\infty} \text{Var}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t\right] = \sigma_u^2 \sum_{-\infty}^{\infty} \rho_v(j)$$
$$= \frac{\sigma_v^2}{1-\alpha^2}\left(2\sum_{0}^{\infty}\alpha^j - 1\right)$$
$$= \frac{\sigma_v^2}{1-\alpha^2}\left(\frac{2}{1-\alpha} - 1\right)$$
$$= \frac{\sigma_v^2}{(1-\alpha^2)^2}$$

b) The OLS is consistent when $\alpha \neq 0$ (or $\sigma_u^2 \neq 0$) because

$$E(y_{t-1}y_t) = E\left[\left([\mathbf{z}'_{t-1}\, y_{t-2}]\boldsymbol{\beta} + u_{t-1}\right)(\alpha u_{t-1} + v_t)\right]$$
$$= E\left[(y_{t-2}\boldsymbol{\beta}_y + u_{t-1})(\alpha u_{t-1} + v_t)\right]$$
$$= \alpha\boldsymbol{\beta}_y E(y_{t-2}u_{t-1}) + E(u_{t-1}(\alpha u_{t-1} + v_t))$$
$$= E(y_{t-1}u_t) + \alpha\sigma_u^2$$
$$= \frac{\alpha\sigma_u^2}{1-\alpha\boldsymbol{\beta}_y}\,.$$

c) A GMM estimate can be defined for instruments $\mathbf{x}_t := \left(\mathbf{z}_t^*, \mathbf{z}_{t-1}^{*'}\right)'$, where $\mathbf{z}_{t-1}^*$ is not including the intercept (which guarantees that $E(v_t) = 0$ for all t), since $E[\mathbf{x}_t u_t] = 0$, by means of

$$\hat{\boldsymbol{\beta}}_{T,GMM} = \hat{\boldsymbol{\beta}}_T(\hat{\mathbf{W}}_T)$$
$$= \left(\sum_{t=1}^{T}[\mathbf{z}_t\, y_{t-1}]'\mathbf{z}_t\mathbf{x}_t'\hat{\mathbf{W}}_T\sum_{t=1}^{T}\mathbf{x}_t[\mathbf{z}_t'\, y_{t-1}]\right)^{-1}\sum_{t=1}^{T}[\mathbf{z}_t\, y_{t-1}]'\mathbf{x}_t'\hat{\mathbf{W}}_T\sum_{t=1}^{T}\mathbf{x}_t y_{t-1}$$

for a particular weighting matrix $\hat{\mathbf{W}}_T \to \mathbf{W} > 0$.
If the $\mathbf{z}_t$ are strictly exogenous we could use an enlarged set of IV's for y_{t-1}, with leads and lags of $\mathbf{z}_t$, and also we could include some lags of y_t such as $y_{t-2}, y_{t-3}, \ldots$.
The complication in the asymptotics resides in the fact that the sequence $\mathbf{x}_t u_t$ need not be uncorrelated, so Newey-West type of asymptotic variances show up.

d) The model is equivalent to the model

$$y_t^* = [\mathbf{z}_t^{*'} \quad y_{t-1}^*]\boldsymbol{\beta} + u_t, \qquad t = 2, 3, \ldots, T,$$

where the error term v_t is IID and independent of the regressors, so OLS is consistent under the appropriate rank condition on the second moment matrix of $[\mathbf{z}_t^{*'} \quad y_{t-1}^*]$, $E\left([\mathbf{z}_t^{*'} \quad y_{t-1}^*]'[\mathbf{z}_t^{*'} \quad y_{t-1}^*]\right)$. In this case the asymptotic covariance matrix of the OLS estimates depend on $E\left([\mathbf{z}_t^{*'} \quad y_{t-1}^*]'[\mathbf{z}_t^{*'} \quad y_{t-1}^*]v_t^2\right)$.