Dr. Randall R. Rojas
Department of Economics
UCLA

Economics 144
Economic Forecasting
Spring, 2016

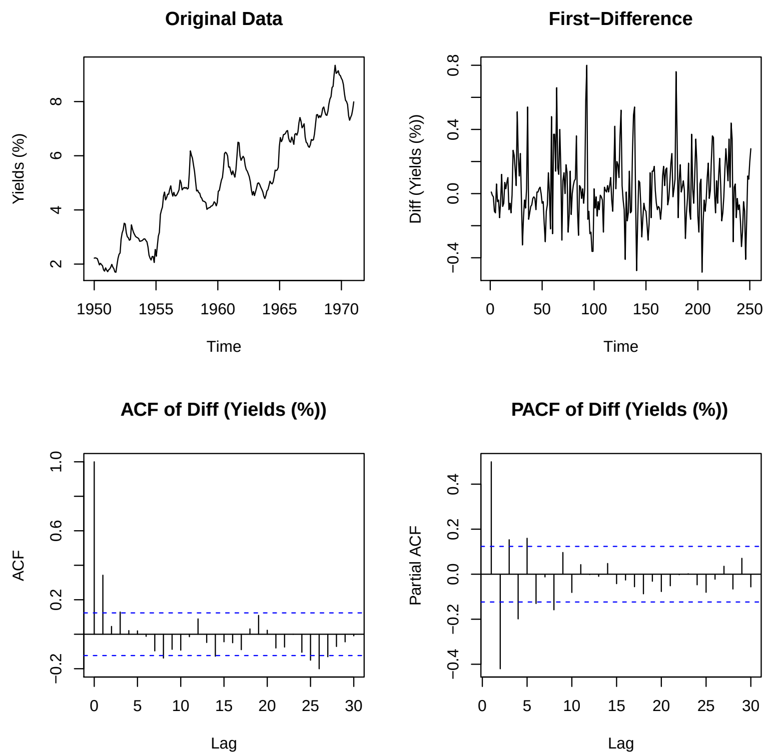
Midterm Exam
May 5, 2016

For full credit on a problem, you need to show all your work and the formula(s) used.

First Name	
Last Name	
UCLA ID #	

Please do not start the exam until instructed to do so.

1.(15%) For this problem we look at the yield of short-term government securities for 21 years (1950-1970, monthly observations) for an unknown European country. The figure below shows the original data, first difference of the data (ΔD), ACF(ΔD), and PACF(ΔD) respectively from left to right, top to bottom. Spikes at zero-lags can be ignored.



(a) Explain what problem(s) are we trying to mitigate by taking the first difference of the original data?

The first difference is typically used to detrend the series in order to make it covariance stationary.

(b) Based on the ACF and PACF plots what model would you propose for this process (i.e., for ΔD) and why? Write down an analytical equation for your proposed model.

Based on the strong spike in the ACF plot at $k = 1$ and the damped oscillatory behavior of the PACF, I would suggest trying an MA(1) process. An alternative might be an MA(3) model (based on the marginally significant spike at $k = 3$) with the terms for only y_{t-1} and y_{t-2}. Therefore the two possibilities would be:
MA(1): $y_t = \mu + \theta_1 \varepsilon_{t-1} + \varepsilon_t$ or MA(3): $y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_3 \varepsilon_{t-3} + \varepsilon_t$

(c) Regardless of the model you suggested in the previous question, estimate an MA(1) model assuming $\widehat{\rho}_1 = 0.4$ and $\mu = 0$.

Recall that for an MA(1) model, $\rho_1 = \theta/(1+\theta^2)$. Therefore, we can solve this equation for θ as a function of $\widehat{\rho}_1$ to get our estimate $\widehat{\theta}_1$: $\widehat{\theta} = \frac{1}{2\widehat{\rho}_1} \pm \sqrt{\left(\frac{1}{2\widehat{\rho}_1}\right)^2 - 1} = \{0.5, 2\}$.
MA(1): $y_t = 0.5\varepsilon_{t-1} + \varepsilon_t$

(d) Using the MA(1) process you estimated in part (c), compute the optimal 1-step-ahead forecast $f_{t,1}$ assuming $\varepsilon_{t=T} = 0.28$.

$f_{t,1} = \mu + \theta\varepsilon_t = 0 + 0.5 \times 0.28 = 0.14.$

(e) Invert the MA(1) process you estimated in part (c), and find its respective autoregressive representation.

For the MA(1) process we know that: $y_t - \mu = \theta(y_{t-1} - \mu) - \theta^2(y_{t-2} - \mu) + \cdots + \varepsilon_t$.
Therefore, the autoregressive representation is: $y_t = 0.5y_{t-1} - 0.25y_{t-2} + \cdots + \varepsilon_t$

2.(15%) Given the time-series below:

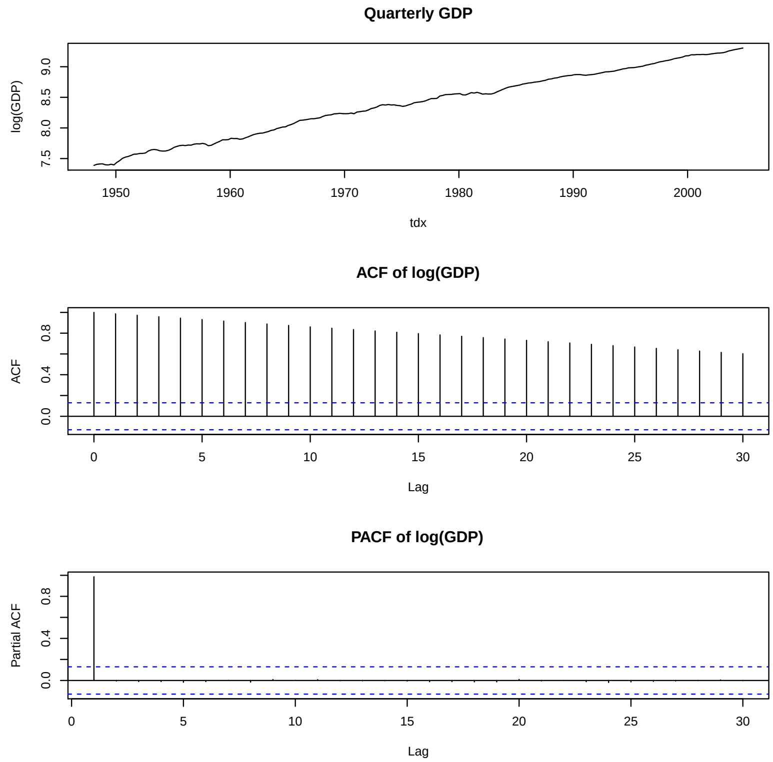
(a) Rewrite the following expression without using the lag operator.
$y_t - y_{t-1} = \left(10 - \frac{2L}{L-3} + \frac{L^2}{L-5}\right) \varepsilon_t$

$$
\begin{aligned}
\rightarrow (L-3)(L-5)(y_t - y_{t-1}) &= \left(10(L-3)(L-5) - 2L(L-5) + L^2(L-3)\right) \varepsilon_t \\
\rightarrow (L^2 - 8L + 15)(y_t - y_{t-1}) &= (L^3 + 5L^2 + 70L + 150)\varepsilon_t \\
\rightarrow 15y_t - 23y_{t-1} + 9y_{t-2} - y_{t-3} &= \varepsilon_{t-3} + 5\varepsilon_{t-2} + 70\varepsilon_{t-1} + 150\varepsilon_t
\end{aligned}
$$

(b) Rewrite the following expression in lag operator form.
$2y_t + 4y_{t-4} - 3y_{t-3} - 1 - \varepsilon_{t-2} + 4\varepsilon_t + 4\varepsilon_{t-2} = 0.$

$$
\rightarrow (2 + 4L^4 - 3L^3)y_t = 1 - (3L^2 + 4)\varepsilon_t
$$

3.(15%) The figure below shows the log series of U.S. quarterly GDP from 1947.I to 2008.IV, and the respective ACF and PACF plots. You can ignore spikes at zero-lags.



(a) What model would you suggest for this process and why? Write down an analytical equation for your proposed model.

Based on the strong spike in the PACF plot at $k = 1$ and the highly persistent but decaying behavior of the ACF, I would suggest trying an AR(1) process with $\phi \approx 1$. Therefore, AR(1): $y_t = c + \phi_1 y_{t-1} + \varepsilon_t$.

(b) Comment on the persistence of the parameter(s) based on the ACF and PACF plots.

Since the ACF spikes out to 30 are still not close to zero, this suggests it will take a lot longer for the ACF to decay to zero, i.e., strong persistence. If we look at the original series, it does not appear to be covariant stationary, hence the comment in (a) about $\phi \approx 1$, which is consistent with strong persistence.

(c) Is this process covariance stationary? Explain your answer.

Given that the original series seems to never return to some 'common value', i.e., never reverts to the mean, then this process does not appear to be covariance stationary.

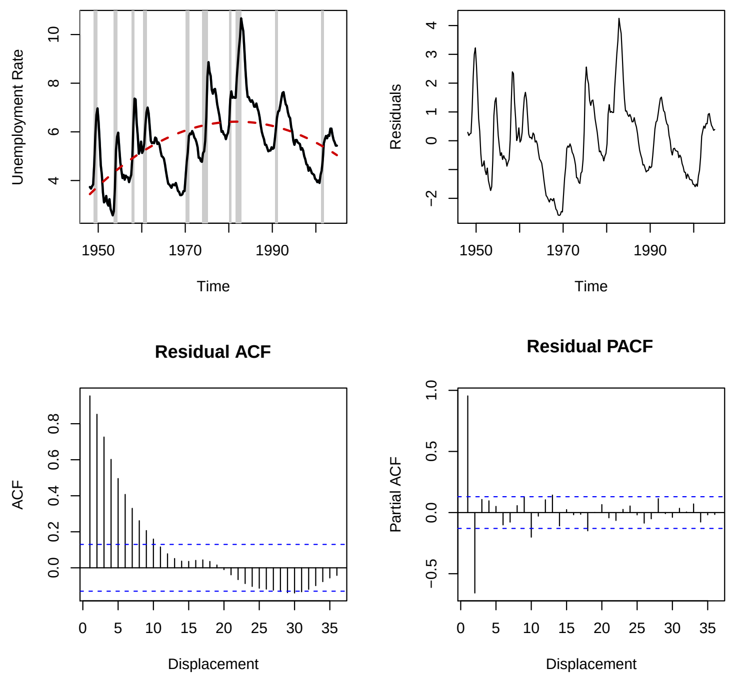
(d) Regardless of the model you suggested in the previous question, estimate an AR(1) model assuming $\widehat{\rho}_1 = 0.98$ and $c = 8.0$.

For an AR(1) process we know that $\rho_1 = \phi$. Therefore, $\widehat{\rho}_1 = \widehat{\phi} = 0.98 \rightarrow$
$y_t = 8 + 0.98 y_{t-1} + \varepsilon_t$.

(e) Using the AR(1) model you estimated in part (d), compute the 1-step-ahead optimal point forecast. Assume $y_{t=T} = 9.31$

$f_{t,1} = c + \widehat{\phi} y_t = 8 + 0.98 \times 9.31 = 17.1238.$

4.(15%) The figure below shows U.S. quarterly unemployment rates from 1947.I to 2008.IV, and the respective ACF and PACF plots of the residuals from a quadratic fit. You can ignore spikes at zero-lags. The recession bands are the vertical grey lines and the dashed-curve is the quadratic fit.



(a) Based on the ACF and PACF of the residuals, can the quadratic fit be improved? If so, how would you improve it and know when you can no longer improve the model fit.

Based on the strong persistence observed in the ACF and spikes in the PACF, these functions are not consistent with white noise. Therefore, we can definitely improve the quadratic fit by including e.g., a seasonal component and maybe even an MA(q) and/or AR(p) process to handle any potential cycles in the data.

(b) Comment on the persistence of the parameter(s) based on the ACF and PACF plots.

The ACF exhibits strong persistence as observed in the slow decay of the high amplitudes out to $k \approx 11$. The pattern is then reversed showing marginally significant spikes around $k \approx 28 - 31$. For the PACF, the spike at $k = 1$ has a large value (~ 0.8), therefore, this would suggest strong persistence as well if it were an AR(1) process.

(c) Suppose you were only interested in the seasonality dynamics and therefore fit a quarterly-seasonal model to the data. The table below shows the respective model fit summary.

	Estimate	Std. Error	t value	Pr(>|t|)
season1	5.6228	0.2032	27.67	0.0000
season2	5.6386	0.2032	27.75	0.0000
season3	5.6351	0.2032	27.73	0.0000
season4	5.6474	0.2032	27.79	0.0000

Based on the results from this table, what would you conclude about the seasonal factors?

The seasonal factors are all comparable in magnitude. This suggests that this series does not exhibit much seasonal behavior (I might even say that it does not contain any seasonal behavior). Therefore, the behavior we discussed earlier in the ACF and PACF must be do to cycles.

5.(15%) For this problem we look at U.S. quarterly real GNP from 1947 to 1991. The table below shows an OLS estimated model fit to the data based on lagged values of itself. For example, the estimate of L(y,3) corresponds to the coefficient of y_{t-3}, and so on.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0052	0.0011	4.62	0.0000
L(y, 1)	0.3373	0.0772	4.37	0.0000
L(y, 2)	0.1979	0.0811	2.44	0.0157
L(y, 3)	-0.1105	0.0810	-1.36	0.1743
L(y, 4)	-0.0991	0.0774	-1.28	0.2025

(a) Based on the table above, for $\alpha = 0.05$, what model would you propose? Provide an analytical equation for this model.

For $\alpha = 0.05$ the only statistically significant terms are the y-intercept, L(y,1), and L(y,2), therefore I would propose the model: $y_t = 0.0052 + 0.3373y_{t-1} + 0.1979y_{t-2} + \varepsilon_t$.

(b) What do you expect the ACF and PACF plots of the original data (GNP) to look like?

Based on the model in (a), we are looking at an AR(2) process. Therefore, I expect the ACF to decay to zero while the PACF should exhibit two strong (statistically significant) spikes at $k = 1$ and $k = 2$.

(c) Interpret the coefficients of L(y,1) and L(y,2).

L(y,1): For a given GNP value in any quarter, the next quarter GNP value is expected to be 0.3373 higher.
L(y,2): For a given GNP value in any quarter, two quarters later, GNP is expected to be 0.1979 higher.

(d) Comment on the persistence of the parameter(s).

Individually, $\phi_1 = 0.3373$ and $\phi_2 = 0.1979$ would not induce much persistence. If we consider their sum $\phi_1 + \phi_2 = 0.5352$, this is somewhat higher but still far enough from 1 that it would not induce much persistence either.

(e) Based on the table above, does this process appear to be covariance stationary? Explain your answer.

Yes, it does appear to be covariance stationary because this is an AR(2) process with $\phi_1 + \phi_2 = 0.5352 << 1$.

Questions 6-10 are multiple choice. Please select one answer per question only.

6.(5%) At what horizon are the forecasts generated by models selected by the AIC and BIC likely to be most accurate?

(a) The AIC and BIC are designed to select models with good h-step-ahead forecasting performance with h very large.

(b) The AIC and BIC are designed to select models with good 1-step-ahead forecasting performance.
Note: The reason is that both metrics depend on the MSE which is computed from the in-sample data.

(c) The AIC and BIC will always agree on the same model regardless of the forecast horizon.

(d) None of the above.

7.(5%) From the list of four time-series processes below, which ones are expected to exhibit a deterministic cycle?
I. $Y_t = 20.5 \sin(t-1) + \varepsilon_t$
II. $Y_t = 100Y_{t-10} + \varepsilon_t$
III. $Y_t = 0.25Y_{t-1} + 0.25Y_{t-2} + \varepsilon_t$
IV. $Y_t = t + 0.01 + \varepsilon_t$

(a) I and II

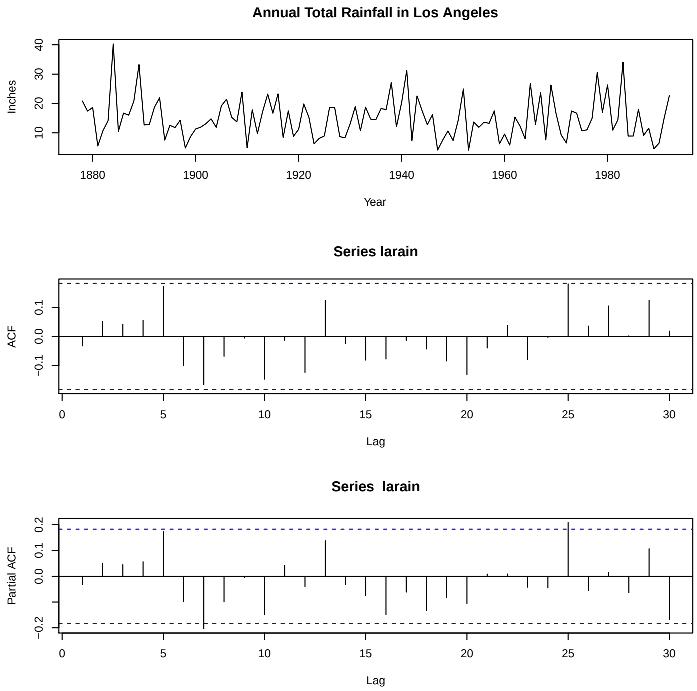
(b) I and III

(c) II and III

(d) I and IV
These are the only two processes that depend explicitly on t.

(e) II and IV

8.(5%) The figure below shows total annual rainfall (in inches) for Los Angeles County from 1878 to 1992 as well as the respective ACF and PACF. For the ACF and PACF values at all lags you can assume the respective p-values (from the Ljung-Box test) are all greater than 0.2. The figure suggests that the series is:



(a) consistent with a white-noise process.
Since none of the spikes in the ACF and PACF are statistically significant, and there is no apparent structure in their respective distributions across al lags, this process seems consistent with white noise.

(b) consistent with an AR(25) process.

(c) consistent with an MA(25) process.

(d) consistent with an MA(5) process.

(e) consistent with an AR(7) process.

9.(5%) Compute the $h = 512$-step-ahead forecast for the following process: $Y_t = 20 + 0.75Y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim WN(0, \sigma^2 = 4)$.

(a) $f_{t,512} = 11.43$

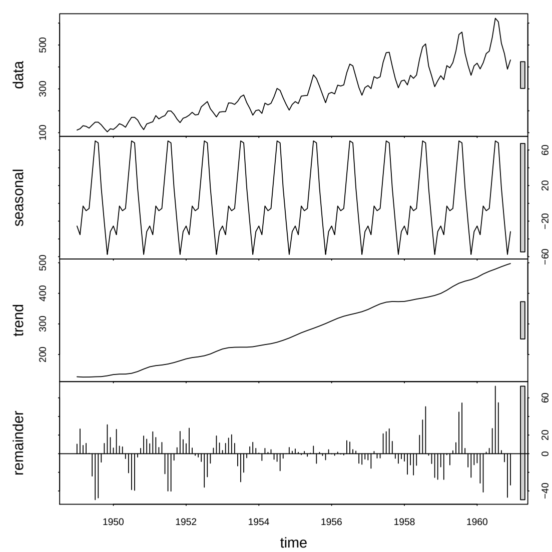
(b) $f_{t,512} = 20.00$

(c) $f_{t,512} \to \infty$

(d) $f_{t,512} = 0$

(e) $f_{t,512} = 80.00$
Recall that for an AR(1) process as $h \to \infty$, $f_{t,h} = c/(1 - \phi) = 20/(1 - 0.75) = 80$.

10.(5%) The figure below shows an 'stl' plot of monthly totals of international airline passengers from 1949 to 1960. Which one of the statements below is true based on the the figure below?



(a) The figure suggests that only when we detrend the series, we are left with white noise residuals.

(b) The figure suggests that only when we detrend and seasonally adjust the series, we are left with white noise residuals.

(c) The figure suggests that there is a strong seasonality component in the data but no cycles.

(d) The figure suggests that there is some dynamic structure in the data in the form of cycles even after detrending and seasonally adjusting the series.