Chapter 9 Discussion Problems

Jake Anderson

ECON 104 – UCLA

Outline

AR Model

An Autoregressive Model An **autoregressive model**, or an **autoregressive process**, is one where a variable y depends on past values of itself. The general representation with p lagged values $\left(y_{t-1}, y_{t-2}, \ldots, y_{t-p}\right)$ is called an autoregressive model (process) of order p, abbreviated as AR(p), and is given by

$$y_t = \delta + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \cdots + \theta_p y_{t-p} + e_t \tag{9.2}$$

For example, an AR(2) model for the unemployment rate series U in Figure 9.2(a) would be $U_t = \delta + \theta_1 U_{t-1} + \theta_2 U_{t-2} + e_t$. AR models can be used to describe the time paths of variables and capture their correlations between current and past values; they are generally used for forecasting. Past values are used to forecast future values.

Intuition: shocks persist because the system carries momentum from one period to the next.

- Academic performance: last quarter's GPA predicts this quarter's GPA due to study habits and cumulative mastery
- Mood levels: your mood this hour influences your mood next hour due to emotional inertia

DL Model

Finite Distributed Lags Suppose that the value of a variable y depends on current and past values of another variable x, up to q periods into the past. We can write this model as

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + e_t \tag{9.1}$$

We can think of (y_t, x_t) as denoting the values for y and x in the current period; x_{t-1} means the value of x in the previous period; x_{t-2} is the value of x two periods ago, and so on. Equations like

Figure: DL Model

Intuition: exposure accumulates and symptoms emerge with lags.

- Air pollition, ad spending, agricultural yields

Autoregressive Distributed Lag Models A more general model that includes both finite distributed lag models and autoregressive models as special cases is the **autoregressive distributed lag** model

$$y_t = \delta + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + \delta_0 x_t + \delta_1 x_{t-1} + \cdots + \delta_q x_{t-q} + e_t \tag{9.3}$$

This model, with p lags of y, the current value x, and q lags of x, is abbreviated as an **ARDL(p, q) model**. The AR component of the name ARDL comes from the regression of y on lagged values of itself; the DL component comes from the distributed lag effect of the lagged x's. For example, an ARDL(2, 1) model relating the unemployment rate U to the growth rate in the economy G would be given by $U_t = \delta + \theta_1 U_{t-1} + \theta_2 U_{t-2} + \delta_0 G_t + \delta_1 G_{t-1} + e_t$. ARDL models can be used for both forecasting and policy analysis. Notice that we have used "δ" with no subscript for the intercept and "δ_s" (δ with a subscript) for the coefficient of x_{t-s}. This notation is a little strange, but it avoids introducing another Greek letter for ARDL models.

Figure: ARDL Model

Outline

Uber Data Science Interview Project

Interview Project: You are given a dataset of logins: "logins.csv". You have 8 hours to do an analysis:

1. Plot these data. Add a best fit line or curve to this graph, and include any relevant metrics/statistics to quantify the quality of fit, and Discuss any significant trends or deviations you observe In the dataset.
2. Repeat this analysis by graphing logins by day of week and by hour of day, noting any interesting findings. Based on what you find, why do you think this is?

The Data

```
login_time
2012-03-01T00:05:55+00:00
2012-03-01T00:06:23+00:00
2012-03-01T00:06:52+00:00
2012-03-01T00:11:23+00:00
2012-03-01T00:12:47+00:00
2012-03-01T00:12:54+00:00
2012-03-01T00:16:14+00:00
2012-03-01T00:17:31+00:00
2012-03-01T00:21:23+00:00
2012-03-01T00:21:26+00:00
2012-03-01T00:22:25+00:00
2012-03-01T00:28:24+00:00
2012-03-01T00:31:21+00:00
2012-03-01T00:32:20+00:00
2012-03-01T00:33:32+00:00
2012-03-01T00:35:21+00:00
2012-03-01T00:38:14+00:00
2012-03-01T00:39:24+00:00
2012-03-01T00:43:12+00:00
2012-03-01T00:46:13+00:00
2012-03-01T00:46:31+00:00
2012-03-01T00:48:03+00:00
```

Figure: Just a bunch of timestamps

The Data

```
login_time
2012-03-01T00:05:55+00:00
2012-03-01T00:06:23+00:00
2012-03-01T00:06:52+00:00
2012-03-01T00:11:23+00:00
2012-03-01T00:12:47+00:00
2012-03-01T00:12:54+00:00
2012-03-01T00:16:14+00:00
2012-03-01T00:17:31+00:00
2012-03-01T00:21:23+00:00
2012-03-01T00:21:26+00:00
2012-03-01T00:22:25+00:00
2012-03-01T00:28:24+00:00
2012-03-01T00:31:21+00:00
2012-03-01T00:32:20+00:00
2012-03-01T00:33:32+00:00
2012-03-01T00:35:21+00:00
2012-03-01T00:38:14+00:00
2012-03-01T00:39:24+00:00
2012-03-01T00:43:12+00:00
2012-03-01T00:46:13+00:00
2012-03-01T00:46:31+00:00
2012-03-01T00:48:03+00:00
```

Figure: Just a bunch of timestamps

The Data

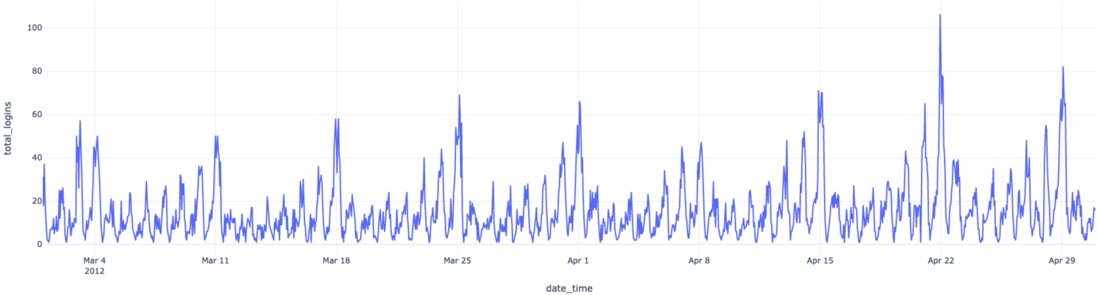


Figure: Very squiggly

The Data

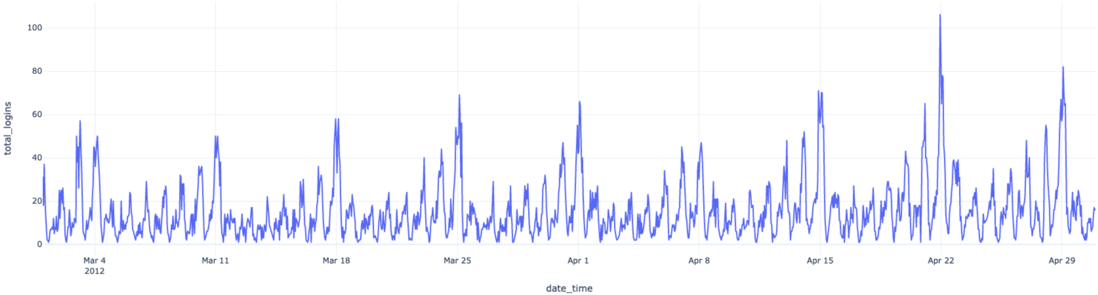
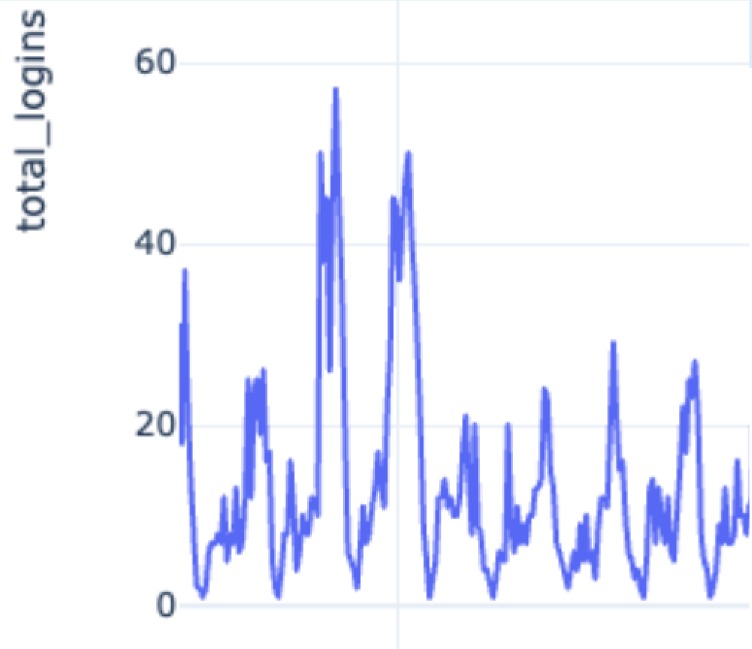


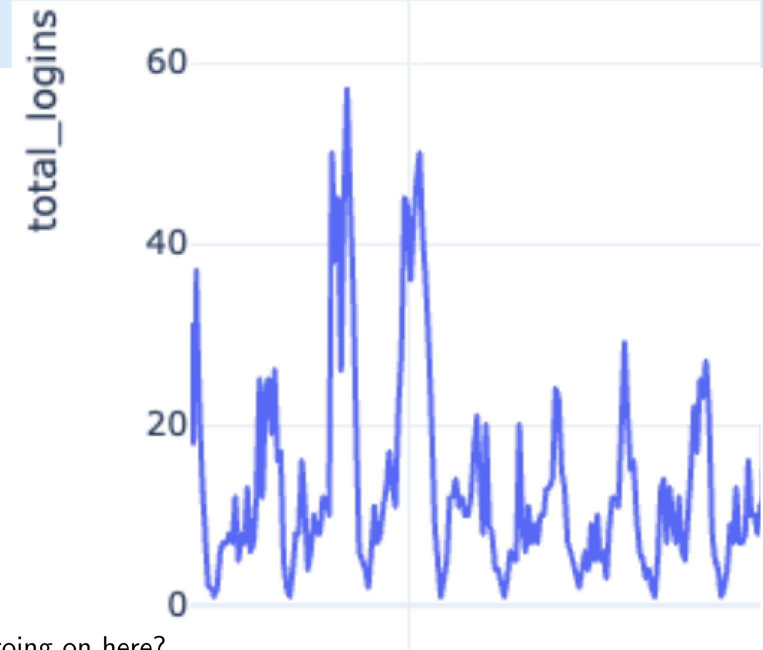
Figure: Very squiggly

$\Longrightarrow$ We need to zoom in a bit...

The Data



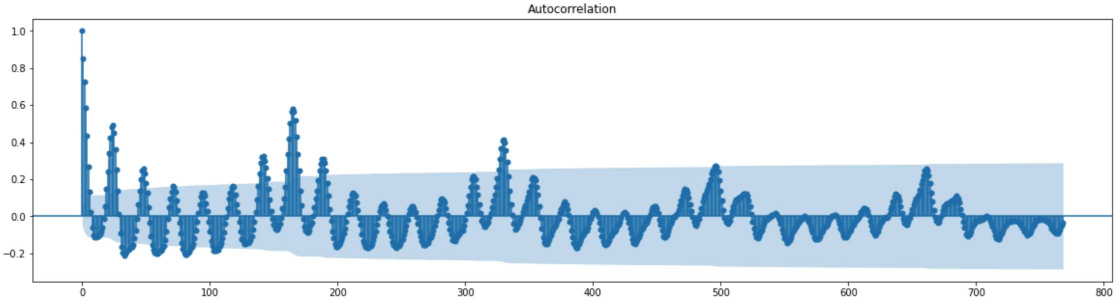
The Data



⟶ What is going on here?

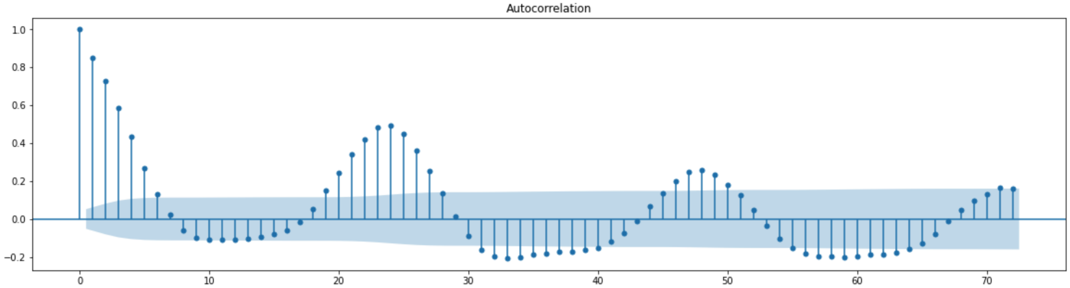
Let's look briefly at autocorrelation:

```
from statsmodels.graphics.tsaplots import plot_acf
plot_acf(df_ts, lags = 24*32, alpha = .05)
```




```
# zoom into 3 days of data
```

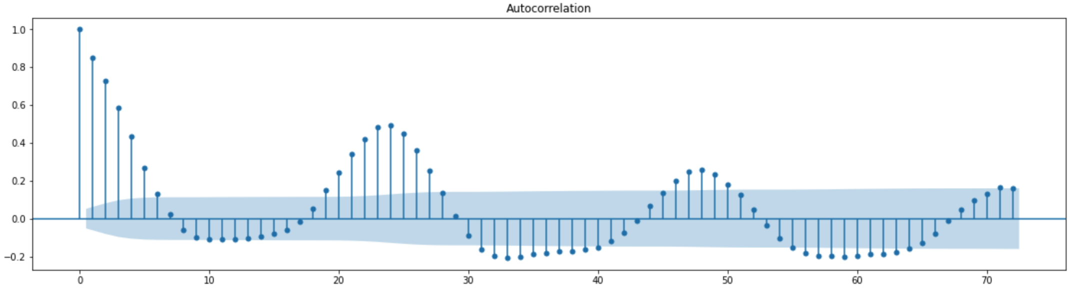
Autocorrelation



We can see predictors at the alpha=.05 level include the following 6 hours with decreasing strength, and the hours nearby 24 hours in the future, with a similar but diminishing trend for 1 day, 2 days, and 3 days afterward. Zooming out to our first figure, we can see that the trend returns most markedly for 7 days later, with echoes into 14 and 21 days.

Figure: Autocorrelation at the hourly level

Autocorrelation



We can see predictors at the alpha=.05 level include the following 6 hours with decreasing strength, and the hours nearby 24 hours in the future, with a similar but diminishing trend for 1 day, 2 days, and 3 days afterward. Zooming out to our first figure, we can see that the trend returns most markedly for 7 days later, with echoes into 14 and 21 days.

Figure: Autocorrelation at the hourly level

Daily Volumes

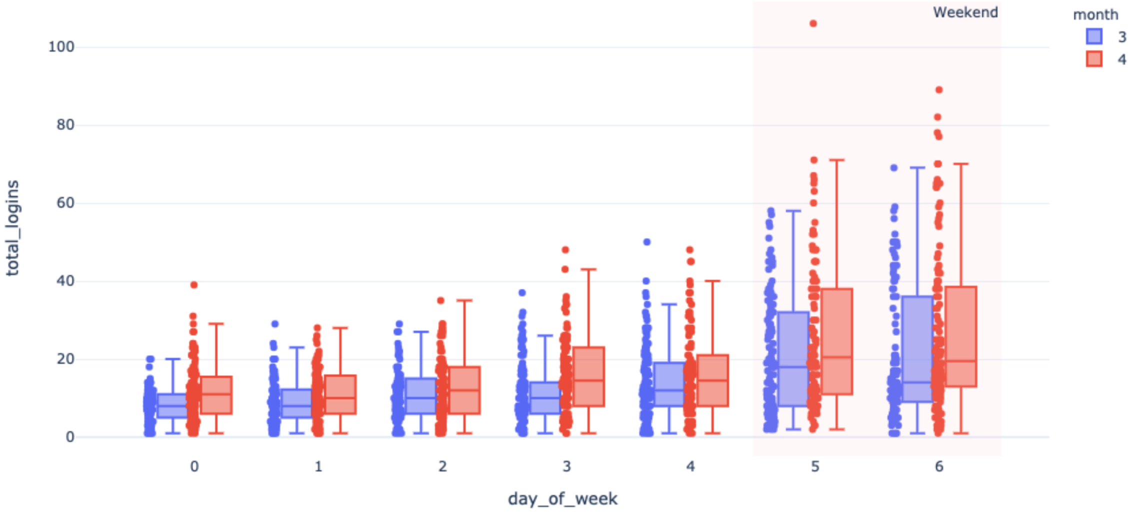


Figure: Autocorrelation by day of week

Daily Volumes

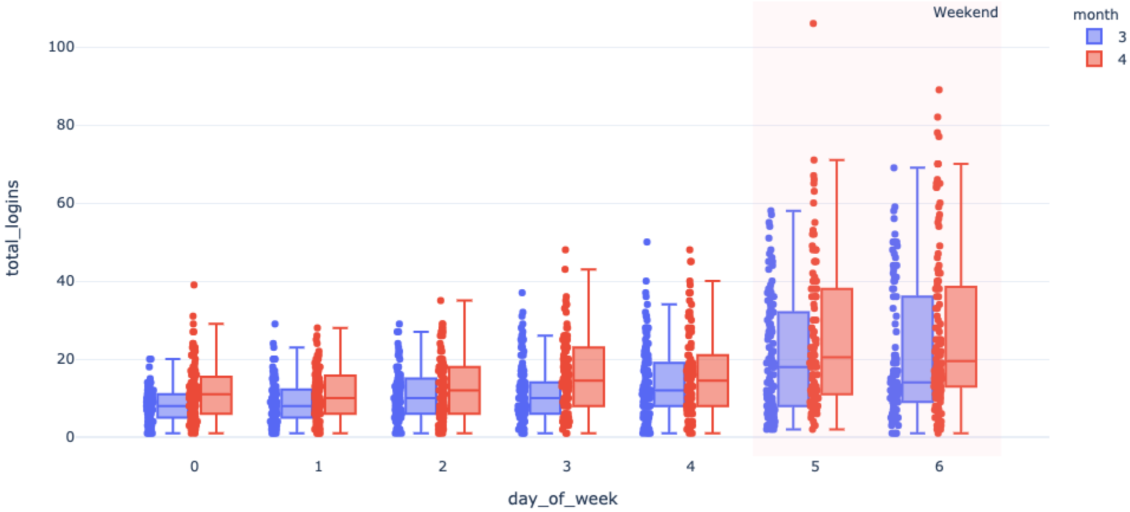
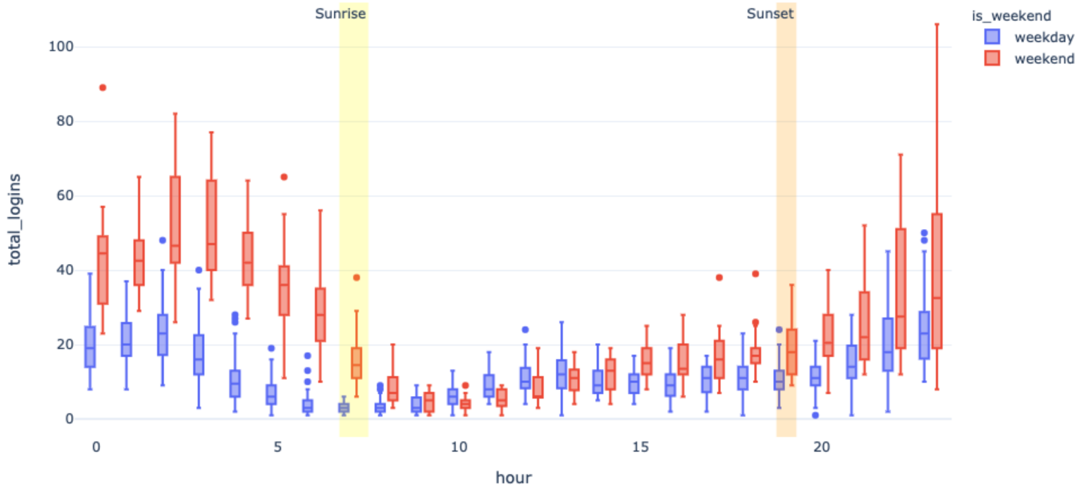
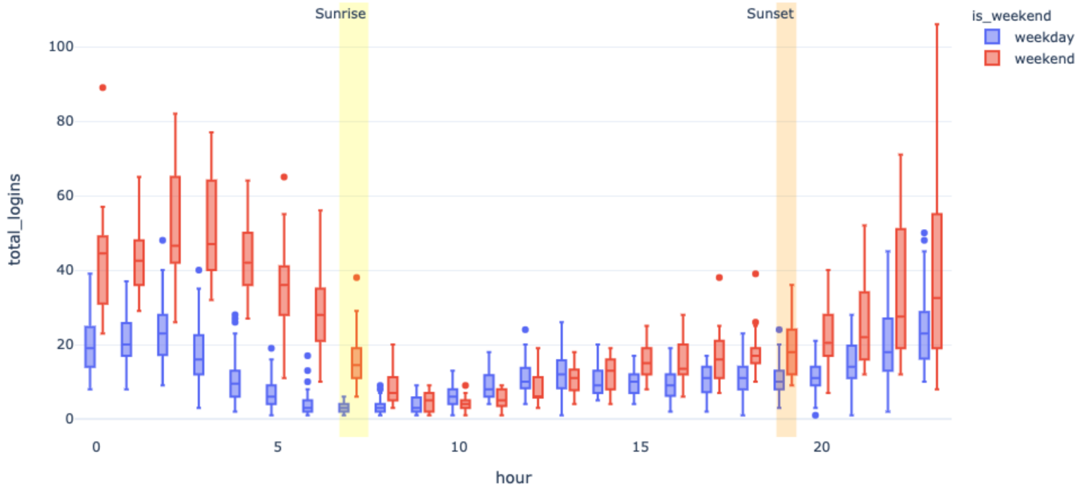


Figure: Autocorrelation by day of week

Hourly Volumes



Hourly Volumes



Model Goodness of Fit

```
for i in range(1,11):
    mod = ARMA(df_ts, order=(i,0))
    res  = mod.fit()
    print(f"for order {i}: ({round(res.aic,3)},{round(res.bic,3)})")

# Output:
for order 1: (9655.73,9671.539)
for order 2: (9657.425,9678.503)
for order 3: (9635.993,9662.341)
for order 4: (9610.851,9642.469)
for order 5: (9570.926,9607.813)  <--- minimum values for each
for order 6: (9571.644,9613.801)
for order 7: (9573.385,9620.811)
for order 8: (9574.881,9627.577)
for order 9: (9573.209,9631.174)
for order 10: (9572.891,9636.126)
```

Data vs Forecast

```
res = mod.fit()
res.plot_predict(start = '2012-03-01', end='2012-03-07')
plt.show()
```

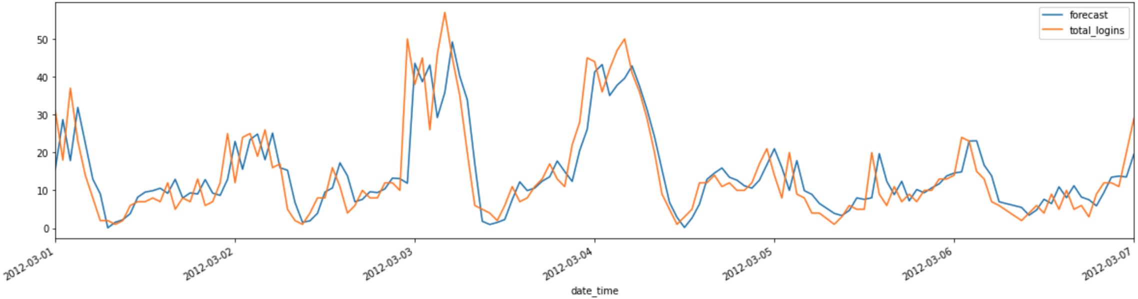


Figure: Data vs forecast

Outline

Problem 9.11: GDP Growth Autocorrelations

Using 250 quarterly observations on U.S. GDP growth (G) from 1947Q2 to 2009Q3, we calculate the following quantities:

$$\sum_{t}^{250}(G_t - \bar{G})^2 = 333.8558 \qquad \sum_{t}^{250}(G_t - \bar{G})(G_{t-1} - \bar{G}) = 162.9753$$

$$\sum_{t}^{250}(G_t - \bar{G})(G_{t-2} - \bar{G}) = 112.4882 \qquad \sum_{t}^{250}(G_t - \bar{G})(G_{t-3} - \bar{G}) = 30.5802$$

Problem 9.11 (a): Compute Autocorrelations

(a) Compute the first three autocorrelations r_1, r_2, and r_3 for G. Test whether each one is significantly different from zero at a 5% significance level.

Problem 9.11 (a): Compute Autocorrelations

(a) Compute the first three autocorrelations r_1, r_2, and r_3 for G. Test whether each one is significantly different from zero at a 5% significance level.

Solution:

Recall the formula for the sample autocorrelation at lag k:

$$r_k = \frac{\sum_t (G_t - \bar{G})(G_{t-k} - \bar{G})}{\sum_t (G_t - \bar{G})^2}$$

Problem 9.11 (a): Compute Autocorrelations

(a) Compute the first three autocorrelations r_1, r_2, and r_3 for G. Test whether each one is significantly different from zero at a 5% significance level.

Solution:
Recall the formula for the sample autocorrelation at lag k:

$$r_k = \frac{\sum_t (G_t - \bar{G})(G_{t-k} - \bar{G})}{\sum_t (G_t - \bar{G})^2}$$

Computing each:

$$r_1 = \frac{162.9753}{333.8558} = 0.4882$$
$$r_2 = \frac{112.4882}{333.8558} = 0.3369$$
$$r_3 = \frac{30.5802}{333.8558} = 0.0916$$

Problem 9.11 (a): Significance Testing

Testing for significance:

Under $H_0 : r_k = 0$, the test statistic is:

$$\hat{t}_k = \sqrt{T} \cdot r_k \sim N(0, 1)$$

Problem 9.11 (a): Significance Testing

Testing for significance:
Under $H_0 : r_k = 0$, the test statistic is:

$$\hat{t}_k = \sqrt{T} \cdot r_k \sim N(0, 1)$$

With $T = 250$, we reject H_0 if

$$|r_k| > \frac{z_{0.025}}{\sqrt{T}} = \frac{1.96}{\sqrt{250}} = 0.124$$

Problem 9.11 (a): Significance Testing

Testing for significance:
Under $H_0 : r_k = 0$, the test statistic is:

$$\hat{t}_k = \sqrt{T} \cdot r_k \sim N(0,1)$$

With $T = 250$, we reject H_0 if

$$|r_k| > \frac{z_{0.025}}{\sqrt{T}} = \frac{1.96}{\sqrt{250}} = 0.124$$

Results:

- $|r_1| = 0.488 > 0.124 \quad \Rightarrow$ **Reject** H_0 (significant)
- $|r_2| = 0.337 > 0.124 \quad \Rightarrow$ **Reject** H_0 (significant)
- $|r_3| = 0.092 < 0.124 \quad \Rightarrow$ **Fail to reject** H_0 (not significant)

Problem 9.11 (a): Significance Testing

Testing for significance:
Under $H_0 : r_k = 0$, the test statistic is:

$$\hat{t}_k = \sqrt{T} \cdot r_k \sim N(0, 1)$$

With $T = 250$, we reject H_0 if

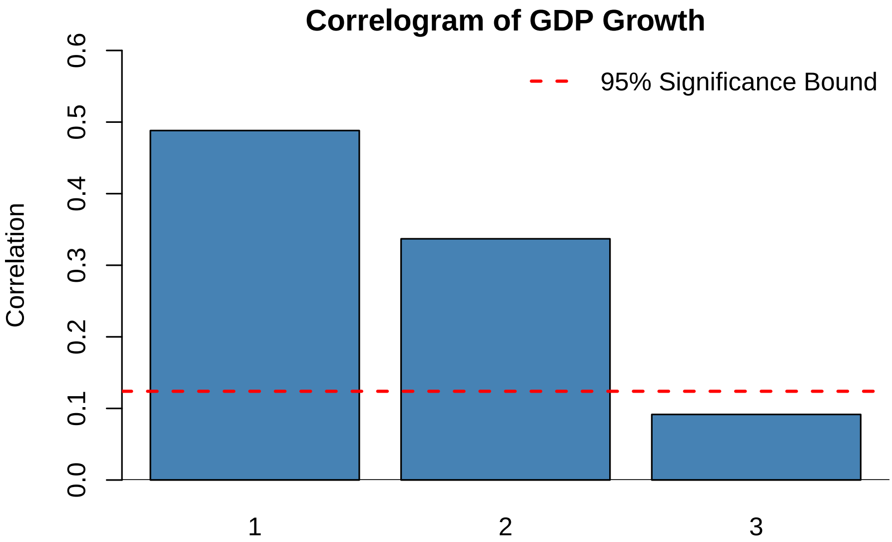
$$|r_k| > \frac{z_{0.025}}{\sqrt{T}} = \frac{1.96}{\sqrt{250}} = 0.124$$

Results:

- $|r_1| = 0.488 > 0.124$ $\Rightarrow$ **Reject** H_0 (significant)
- $|r_2| = 0.337 > 0.124$ $\Rightarrow$ **Reject** H_0 (significant)
- $|r_3| = 0.092 < 0.124$ $\Rightarrow$ **Fail to reject** H_0 (not significant)

Translate into English: GDP growth is autocorrelated at 1 and 2 quarter horizons, but not at 3 quarters.

Problem 9.11 (a): Correlogram



Problem 9.11 (b): AR(1) Model Estimation

(b) Given that

$$\sum_{t}^{250}(G_t - \bar{G}_{-1})^2 = 333.1119, \quad \sum_{t}^{250}(G_t - \bar{G}_{-1})(G_{t-1} - \bar{G}_1) = 162.974$$

where $\bar{G}_1 = 1.662249$ and $\bar{G}_{-1} = 1.664257$, find the least squares estimates of δ and θ_1 in the AR(1) model:

$$G_t = \delta + \theta_1 G_{t-1} + e_t$$

Problem 9.11 (b): AR(1) Model Estimation

(b) Given that

$$\sum_t^{250}(G_t - \bar{G}_{-1})^2 = 333.1119, \quad \sum_t^{250}(G_t - \bar{G}_{-1})(G_{t-1} - \bar{G}_1) = 162.974$$

where $\bar{G}_1 = 1.662249$ and $\bar{G}_{-1} = 1.664257$, find the least squares estimates of δ and θ_1 in the AR(1) model:

$$G_t = \delta + \theta_1 G_{t-1} + e_t$$

Solution:
Using OLS formulas:

$$\hat{\theta}_1 = \frac{\text{Cov}(G_t, G_{t-1})}{\text{Var}(G_{t-1})} = \frac{162.974}{333.1119} = 0.4892$$

$$\hat{\delta} = \bar{G}_1 - \hat{\theta}_1 \bar{G}_{-1} = 1.662249 - 0.4892 \times 1.664257 = 0.8480$$

Problem 9.11 (b): Why $\hat{\theta}_1 \neq r_1$?

Question: Why is $\hat{\theta}_1 = 0.4892$ slightly different from $r_1 = 0.4882$?

Problem 9.11 (b): Why $\hat{\theta}_1 \neq r_1$?

Question: Why is $\hat{\theta}_1 = 0.4892$ slightly different from $r_1 = 0.4882$?

Answer:

- The denominator for r_1 uses $\sum(G_t - \bar{G})^2 = 333.8558$
- The denominator for $\hat{\theta}_1$ uses $\sum(G_{t-1} - \bar{G}_{t-1})^2 = 333.1119$

Problem 9.11 (b): Why $\hat{\theta}_1 \neq r_1$?

Question: Why is $\hat{\theta}_1 = 0.4892$ slightly different from $r_1 = 0.4882$?

Answer:

- The denominator for r_1 uses $\sum(G_t - \bar{G})^2 = 333.8558$
- The denominator for $\hat{\theta}_1$ uses $\sum(G_{t-1} - \bar{G}_{t-1})^2 = 333.1119$

Key differences:

1. r_1 uses the *same* mean $\bar{G}$ for both G_t and G_{t-1}
2. $\hat{\theta}_1$ uses *different* means: $\bar{G}_1$ for G_t and $\bar{G}_{t-1}$ for G_{t-1}
3. The summation in the denominator for r_1 has one more term (why?)

Problem 9.11 (b): Why $\hat{\theta}_1 \neq r_1$?

Question: Why is $\hat{\theta}_1 = 0.4892$ slightly different from $r_1 = 0.4882$?

Answer:

- The denominator for r_1 uses $\sum(G_t - \bar{G})^2 = 333.8558$
- The denominator for $\hat{\theta}_1$ uses $\sum(G_{t-1} - \bar{G}_{t-1})^2 = 333.1119$

Key differences:

1. r_1 uses the *same* mean $\bar{G}$ for both G_t and G_{t-1}
2. $\hat{\theta}_1$ uses *different* means: $\bar{G}_1$ for G_t and $\bar{G}_{t-1}$ for G_{t-1}
3. The summation in the denominator for r_1 has one more term (why?)

The lagged series drops the first observation.

$$\bar{G}_t = \frac{1}{T}\sum_{t=1}^{T} G_t, \qquad \bar{G}_{t-1} = \frac{1}{T-1}\sum_{t=2}^{T} G_{t-1}$$

$\implies$ In large samples, these differences become negligible, but they exist in finite samples.

Outline

Problem 9.13: ARDL to Infinite Lag Representation

Consider the infinite distributed lag (IDL) representation:

$$y_t = \alpha + \sum_{s=0}^{\infty} \beta_s x_{t-s} + e_t$$

for the ARDL model:

$$y_t = \delta + \theta_1 y_{t-1} + \theta_3 y_{t-3} + \delta_1 x_{t-1} + v_t$$

Problem 9.13: ARDL to Infinite Lag Representation

Consider the infinite distributed lag (IDL) representation:

$$y_t = \alpha + \sum_{s=0}^{\infty} \beta_s x_{t-s} + e_t$$

for the ARDL model:

$$y_t = \delta + \theta_1 y_{t-1} + \theta_3 y_{t-3} + \delta_1 x_{t-1} + v_t$$

Goal: Express the ARDL parameters in terms of the IDL parameters. i.e. find the connection between the parameters of the two models.

Problem 9.13 (a): Derive the Lag Weights

(a) Show that:

$$\alpha = \frac{\delta}{1 - \theta_1 - \theta_3}$$
$$\beta_0 = 0$$
$$\beta_1 = \delta_1$$
$$\beta_2 = \theta_1 \beta_1$$
$$\beta_3 = \theta_1 \beta_2$$
$$\beta_s = \theta_1 \beta_{s-1} + \theta_3 \beta_{s-3} \quad \text{for } s \geq 4$$

Problem 9.13 (a): Step 1 – Write the IDL Representation

:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \cdots$$

Problem 9.13 (a): Step 1 – Write the IDL Representation

:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \cdots$$

Shift this equation to get the lagged y's we need:

$$y_{t-1} = \alpha + \beta_0 x_{t-1} + \beta_1 x_{t-2} + \beta_2 x_{t-3} + \beta_3 x_{t-4} + \cdots$$

$$y_{t-3} = \alpha + \beta_0 x_{t-3} + \beta_1 x_{t-4} + \beta_2 x_{t-5} + \beta_3 x_{t-6} + \cdots$$

Problem 9.13 (a): Step 2 – Plug Into the ARDL

The ARDL is:

$$y_t = \delta + \theta_1 y_{t-1} + \theta_3 y_{t-3} + \delta_1 x_{t-1} + v_t$$

Problem 9.13 (a): Step 2 – Plug Into the ARDL

The ARDL is:

$$y_t = \delta + \theta_1 y_{t-1} + \theta_3 y_{t-3} + \delta_1 x_{t-1} + v_t$$

Substitute our IDL expressions for y_{t-1} and y_{t-3}:

$$\begin{aligned} y_t = \delta &+ \theta_1\Big(\alpha + \beta_0 x_{t-1} + \beta_1 x_{t-2} + \beta_2 x_{t-3} + \beta_3 x_{t-4} + \cdots\Big) \\ &+ \theta_3\Big(\alpha + \beta_0 x_{t-3} + \beta_1 x_{t-4} + \beta_2 x_{t-5} + \beta_3 x_{t-6} + \cdots\Big) \\ &+ \delta_1 x_{t-1} + v_t \end{aligned}$$

Problem 9.13 (a): Step 2 – Plug Into the ARDL

The ARDL is:

$$y_t = \delta + \theta_1 y_{t-1} + \theta_3 y_{t-3} + \delta_1 x_{t-1} + v_t$$

Substitute our IDL expressions for y_{t-1} and y_{t-3}:

$$\begin{aligned} y_t = \delta + \theta_1 &\Big(\alpha + \beta_0 x_{t-1} + \beta_1 x_{t-2} + \beta_2 x_{t-3} + \beta_3 x_{t-4} + \cdots\Big) \\ &+ \theta_3\Big(\alpha + \beta_0 x_{t-3} + \beta_1 x_{t-4} + \beta_2 x_{t-5} + \beta_3 x_{t-6} + \cdots\Big) \\ &+ \delta_1 x_{t-1} + v_t \end{aligned}$$

Line them up to see it clearly:

$$\begin{aligned} y_t = \delta + \theta_1 \Big(\alpha + \beta_0 x_{t-1} + \beta_1 x_{t-2} + &\beta_2 x_{t-3} + \beta_3 x_{t-4} + \cdots\Big) \\ +\theta_3\Big(\alpha + &\beta_0 x_{t-3} + \beta_1 x_{t-4} + \beta_2 x_{t-5} + \beta_3 x_{t-6} + \cdots\Big) \\ + \delta_1 x_{t-1} + v_t \end{aligned}$$

Problem 9.13 (a): Step 3 – Distribute θ_1 and θ_3

Expanding:

$$\begin{aligned} y_t = \delta &+ (\theta_1\alpha) + (\theta_3\alpha) \\ &+ (\theta_1\beta_0)x_{t-1} + (\theta_1\beta_1)x_{t-2} + (\theta_1\beta_2)x_{t-3} + (\theta_1\beta_3)x_{t-4} + \cdots \\ &+ (\theta_3\beta_0)x_{t-3} + (\theta_3\beta_1)x_{t-4} + (\theta_3\beta_2)x_{t-5} + (\theta_3\beta_3)x_{t-6} + \cdots \\ &+ \delta_1 x_{t-1} + v_t \end{aligned}$$

Problem 9.13 (a): Step 4 – Group by Powers of x

Rearranging by collecting like lags of x:

$$\begin{aligned} y_t = &\big(\delta + \theta_1\alpha + \theta_3\alpha\big) \\ &+ \big(\delta_1 + \theta_1\beta_0\big)x_{t-1} \\ &+ \big(\theta_1\beta_1\big)x_{t-2} \\ &+ \big(\theta_1\beta_2 + \theta_3\beta_0\big)x_{t-3} \\ &+ \big(\theta_1\beta_3 + \theta_3\beta_1\big)x_{t-4} \\ &+ \big(\theta_1\beta_4 + \theta_3\beta_2\big)x_{t-5} + \cdots + v_t \end{aligned}$$

Problem 9.13 (a): Step 5 – Match Coefficients

Compare to the IDL target:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \cdots + e_t$$

Problem 9.13 (a): Step 5 – Match Coefficients

Compare to the IDL target:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \cdots + e_t$$

Matching term-by-term:

- **Constant:** $\alpha = \delta + \theta_1 \alpha + \theta_3 \alpha \quad \Rightarrow \quad \boxed{\alpha = \dfrac{\delta}{1 - \theta_1 - \theta_3}}$

Problem 9.13 (a): Step 5 – Match Coefficients

Compare to the IDL target:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \cdots + e_t$$

Matching term-by-term:

- **Constant:** $\alpha = \delta + \theta_1 \alpha + \theta_3 \alpha \quad \Rightarrow \quad \boxed{\alpha = \dfrac{\delta}{1 - \theta_1 - \theta_3}}$
- x_t**:** $\beta_0 = 0$ (no x_t appears on the ARDL RHS)

Problem 9.13 (a): Step 5 – Match Coefficients

Compare to the IDL target:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \cdots + e_t$$

Matching term-by-term:

- **Constant:** $\alpha = \delta + \theta_1 \alpha + \theta_3 \alpha \quad \Rightarrow \quad \boxed{\alpha = \dfrac{\delta}{1 - \theta_1 - \theta_3}}$
- x_t: $\beta_0 = 0 \quad$ (no x_t appears on the ARDL RHS)
- x_{t-1}: $\beta_1 = \delta_1 + \theta_1 \beta_0 = \boxed{\delta_1}$

Problem 9.13 (a): Step 5 – Match Coefficients

Compare to the IDL target:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \cdots + e_t$$

Matching term-by-term:

- **Constant:** $\alpha = \delta + \theta_1 \alpha + \theta_3 \alpha \quad \Rightarrow \quad \boxed{\alpha = \dfrac{\delta}{1 - \theta_1 - \theta_3}}$
- x_t: $\beta_0 = 0 \quad$ (no x_t appears on the ARDL RHS)
- x_{t-1}: $\beta_1 = \delta_1 + \theta_1 \beta_0 = \boxed{\delta_1}$
- x_{t-2}: $\beta_2 = \boxed{\theta_1 \beta_1}$

Problem 9.13 (a): Step 5 – Match Coefficients

Compare to the IDL target:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + \cdots + e_t$$

Matching term-by-term:

- **Constant:** $\alpha = \delta + \theta_1 \alpha + \theta_3 \alpha \quad \Rightarrow \quad \boxed{\alpha = \dfrac{\delta}{1 - \theta_1 - \theta_3}}$
- x_t: $\beta_0 = 0$ (no x_t appears on the ARDL RHS)
- x_{t-1}: $\beta_1 = \delta_1 + \theta_1 \beta_0 = \boxed{\delta_1}$
- x_{t-2}: $\beta_2 = \boxed{\theta_1 \beta_1}$
- x_{t-3}: $\beta_3 = \theta_1 \beta_2 + \theta_3 \beta_0 = \boxed{\theta_1 \beta_2}$

Problem 9.13 (a): The General Recursion

For $s \geq 4$, the pattern becomes:

$$\boxed{\beta_s = \theta_1 \beta_{s-1} + \theta_3 \beta_{s-3}}$$

Problem 9.13 (a): The General Recursion

For $s \geq 4$, the pattern becomes:

$$\boxed{\beta_s = \theta_1 \beta_{s-1} + \theta_3 \beta_{s-3}}$$

Summary:

$$\begin{aligned}
\alpha &= \frac{\delta}{1 - \theta_1 - \theta_3} \\
\beta_0 &= 0 \\
\beta_1 &= \delta_1 \\
\beta_2 &= \theta_1 \beta_1 \\
\beta_3 &= \theta_1 \beta_2 \\
\beta_s &= \theta_1 \beta_{s-1} + \theta_3 \beta_{s-3} \quad \text{for } s \geq 4
\end{aligned}$$

Problem 9.13 (b): Estimated Phillips Curve

Part (b): Using quarterly data on U.S. inflation (INF) and change in unemployment (DU) from 1955Q2 to 2016Q1, we estimate:

$$\begin{array}{llllll} \widehat{INF}_t = & 0.094+ & 0.564 \cdot INF_{t-1}+ & 0.333 \cdot INF_{t-3} & - 0.300 \cdot DU_{t-1} \\ \text{(se)} & (0.049) & (0.051) & (0.052) & (0.084) \end{array}$$

with $SSE = 48.857$.

Problem 9.13 (b): Estimated Phillips Curve

Part (b): Using quarterly data on U.S. inflation (INF) and change in unemployment (DU) from 1955Q2 to 2016Q1, we estimate:

$$\begin{array}{lllll} \widehat{INF}_t = & 0.094 + & 0.564 \cdot INF_{t-1} + & 0.333 \cdot INF_{t-3} & - 0.300 \cdot DU_{t-1} \\ \text{(se)} & (0.049) & (0.051) & (0.052) & (0.084) \end{array}$$

with $SSE = 48.857$.

Mapping to our ARDL notation:

- $\theta_1 = 0.564$
- $\theta_3 = 0.333$
- $\delta_1 = -0.300$

Problem 9.13 (c): Computing Lag Weights

- (c) Using the results in part (a), find estimates of the first 12 lag weights in the infinite lag representation.

Problem 9.13 (c): Computing Lag Weights

- Using the results in part (a), find estimates of the first 12 lag weights in the infinite lag representation.

Solution:

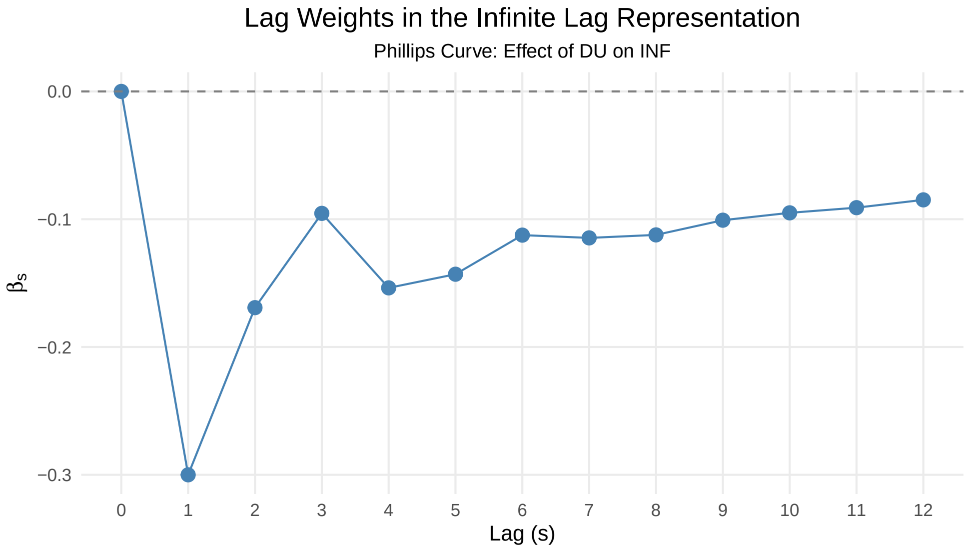
Using $\theta_1 = 0.564$, $\theta_3 = 0.333$, $\delta_1 = -0.300$:

$$\beta_0 = 0$$
$$\beta_1 = -0.300$$
$$\beta_2 = 0.564 \times (-0.300) = -0.169$$
$$\beta_3 = 0.564 \times (-0.169) = -0.095$$

For $s \geq 4$:

$$\beta_4 = 0.564\beta_3 + 0.333\beta_1 = -0.154$$
$$\beta_5 = 0.564\beta_4 + 0.333\beta_2 = -0.143$$
$$\vdots$$

Problem 9.13 (c): Lag Weights Graph



Problem 9.13 (d): Constant Inflation Rate

(d) What rate of inflation is consistent with a constant unemployment rate (where $DU = 0$ in all periods)?

Problem 9.13 (d): Constant Inflation Rate

- (d) What rate of inflation is consistent with a constant unemployment rate (where $DU = 0$ in all periods)?

Solution:
With constant inflation: $INF_t = INF_{t-1} = INF_{t-3} = INF$
And with $DU = 0$, the model becomes:

$$INF = 0.094 + 0.564 \cdot INF + 0.333 \cdot INF$$

Problem 9.13 (d): Constant Inflation Rate

(d) What rate of inflation is consistent with a constant unemployment rate (where $DU = 0$ in all periods)?

Solution:
With constant inflation: $INF_t = INF_{t-1} = INF_{t-3} = INF$
And with $DU = 0$, the model becomes:

$$INF = 0.094 + 0.564 \cdot INF + 0.333 \cdot INF$$

Solving:

$$INF(1 - 0.564 - 0.333) = 0.094$$

$$INF = \frac{0.094}{0.103} = 0.913$$

Problem 9.13 (d): Constant Inflation Rate

(d) What rate of inflation is consistent with a constant unemployment rate (where $DU = 0$ in all periods)?

Solution:
With constant inflation: $INF_t = INF_{t-1} = INF_{t-3} = INF$
And with $DU = 0$, the model becomes:

$$INF = 0.094 + 0.564 \cdot INF + 0.333 \cdot INF$$

Solving:

$$INF(1 - 0.564 - 0.333) = 0.094$$

$$INF = \frac{0.094}{0.103} = 0.913$$

Interpretation: A constant unemployment rate is consistent with approximately 0.91% quarterly inflation (about 3.6% annual).

Problem 9.13 (e): Testing for AR(3) Errors

- (e) The SSE from regressing $\hat{u}_t$ on a constant, INF_{t-1}, INF_{t-3}, DU_{t-1}, $\hat{e}_{t-1}$, and $\hat{e}_{t-3}$ is 47.619. Using a 5% significance level, test the hypothesis that the errors follow an AR(3) process.

Problem 9.13 (e): Testing for AR(3) Errors

- (e) The SSE from regressing $\hat{u}_t$ on a constant, INF_{t-1}, INF_{t-3}, DU_{t-1}, $\hat{e}_{t-1}$, and $\hat{e}_{t-3}$ is 47.619. Using a 5% significance level, test the hypothesis that the errors follow an AR(3) process.

Solution:
This is an LM test for serial correlation. The test statistic is:

$$LM = N \times R^2$$

where R^2 comes from the auxiliary regression.

Problem 9.13 (e): Testing for AR(3) Errors

- (e) The SSE from regressing $\hat{u}_t$ on a constant, INF_{t-1}, INF_{t-3}, DU_{t-1}, $\hat{e}_{t-1}$, and $\hat{e}_{t-3}$ is 47.619. Using a 5% significance level, test the hypothesis that the errors follow an AR(3) process.

Solution:
This is an LM test for serial correlation. The test statistic is:

$$LM = N \times R^2$$

where R^2 comes from the auxiliary regression.

$$R^2 = 1 - \frac{SSE_{aux}}{SSE_{original}} = 1 - \frac{47.619}{48.857} = 0.0253$$

$$LM = 241 \times 0.0253 = 6.11$$

Problem 9.13 (e): LM Test Conclusion

Under H_0 (no AR(3) errors), $LM \sim \chi^2_2$ (2 df for two lagged error terms).

Problem 9.13 (e): LM Test Conclusion

Under H_0 (no AR(3) errors), $LM \sim \chi^2_2$ (2 df for two lagged error terms).

Critical value: $\chi^2_{0.95,2} = 5.99$

Problem 9.13 (e): LM Test Conclusion

Under H_0 (no AR(3) errors), $LM \sim \chi^2_2$ (2 df for two lagged error terms).

Critical value: $\chi^2_{0.95,2} = 5.99$

Since $LM = 6.11 > 5.99$, we **reject** H_0 at the 5% level.

Problem 9.13 (e): LM Test Conclusion

Under H_0 (no AR(3) errors), $LM \sim \chi^2_2$ (2 df for two lagged error terms).

Critical value: $\chi^2_{0.95,2} = 5.99$

Since $LM = 6.11 > 5.99$, we **reject** H_0 at the 5% level.

Implication: There is evidence that the errors *do not* follow the AR(3) process implied by the ARDL structure. This suggests:

- The IDL representation may not be fully adequate
- Additional lags or different specification may be needed
- Standard errors from the ARDL may need adjustment

Thank you!

jakeanderson@g.ucla.edu