Chapter 2: ARMA Models

Seminar

Seminar Outline

1. Multiple Choice Quiz
2. True/False Questions
3. Calculation Exercises
4. Python Exercises
5. Discussion Questions
6. Summary

Quiz 1: Lag Operator

Question: What is the result of applying $(1 - L)^2$ to X_t?

- A. $X_t - X_{t-1}$
- B. $X_t - 2X_{t-1} + X_{t-2}$
- C. $X_t + X_{t-1} + X_{t-2}$
- D. $X_t - X_{t-2}$

Quiz 1: Solution

Answer: B

Explanation:

$$\begin{aligned}(1-L)^2X_t &= (1-2L+L^2)X_t\\ &= X_t - 2LX_t + L^2X_t\\ &= X_t - 2X_{t-1} + X_{t-2}\end{aligned}$$

This is the **second difference** of X_t.

Note: $(1-L)$ is the first difference operator, $(1-L)^2$ is the second difference.

Quiz 2: AR(1) Stationarity

Question: For which value of ϕ is the AR(1) process $X_t = 0.5 + \phi X_{t-1} + \varepsilon_t$ stationary?

A. $\phi = 1.2$
B. $\phi = 1.0$
C. $\phi = -0.8$
D. $\phi = -1.5$

Quiz 2: Solution

Answer: C

Explanation: AR(1) is stationary if and only if $|\phi| < 1$.

Checking each option:

- A. $|\phi| = 1.2 > 1 \rightarrow$ Non-stationary (explosive)
- B. $|\phi| = 1.0 \rightarrow$ Non-stationary (unit root / random walk)
- C. $|\phi| = 0.8 < 1 \rightarrow$ **Stationary**
- D. $|\phi| = 1.5 > 1 \rightarrow$ Non-stationary (explosive)

The stationarity condition requires the root of $1 - \phi z = 0$ to lie outside the unit circle, i.e., $|1/\phi| > 1$, which means $|\phi| < 1$.

Quiz 3: ACF Pattern

Question: You observe the following ACF pattern: significant spike at lag 1, then all other lags are within confidence bands. The PACF shows gradual decay. What model is suggested?

A. AR(1)
B. MA(1)
C. ARMA(1,1)
D. White noise

Quiz 3: Solution

Answer: B

Explanation:

Model	**ACF**	**PACF**
AR(p)	Decays	Cuts off at lag p
MA(q)	Cuts off at lag q	Decays
ARMA	Decays	Decays

The pattern described:

- ACF cuts off after lag 1 $\rightarrow$ suggests MA
- PACF decays $\rightarrow$ confirms MA (not AR)

Therefore, this is an **MA(1)** process.

Quiz 4: MA Invertibility

Question: For the MA(1) process $X_t = \varepsilon_t + 1.5\varepsilon_{t-1}$, is the process invertible?

A. Yes, because MA processes are always invertible
B. Yes, because $1.5 > 0$
C. No, because $|\theta| = 1.5 > 1$
D. No, because MA processes are never invertible

Quiz 4: Solution

Answer: C

Explanation: An MA(1) process $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$ is invertible if $|\theta| < 1$.

Here, $\theta = 1.5$, so $|\theta| = 1.5 > 1 \rightarrow$ **Not invertible**

Key points:

- MA processes are *always stationary* (for finite coefficients)
- But they are *not always invertible*
- Invertibility requires roots of $\theta(z) = 1 + \theta z = 0$ outside unit circle
- Root: $z = -1/\theta = -1/1.5 = -0.67$ is *inside* unit circle

Quiz 5: ARMA Representation

Question: The compact form $\phi(L)X_t = \theta(L)\varepsilon_t$ represents which model?

A. Pure AR model
B. Pure MA model
C. ARMA model
D. None of the above

Quiz 5: Solution

Answer: C

Explanation:

- $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ is the AR polynomial
- $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$ is the MA polynomial

The equation $\phi(L)X_t = \theta(L)\varepsilon_t$ expands to:

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

This is the general **ARMA(p,q)** model.

Special cases:

- $\theta(L) = 1$ (no MA): Pure AR
- $\phi(L) = 1$ (no AR): Pure MA

Quiz 6: Information Criteria

Question: When comparing ARMA(1,1) and ARMA(2,1) using BIC, which statement is correct?

- A. Lower BIC always means better forecasts
- B. BIC penalizes complexity less than AIC
- C. The model with lower BIC is preferred
- D. BIC can only compare models with same number of parameters

Quiz 6: Solution

Answer: C

Explanation:

- A is **false**: Lower BIC indicates better in-sample fit relative to complexity, but doesn't guarantee best forecasts
- B is **false**: BIC penalizes complexity *more* than AIC (penalty is $k \ln(n)$ vs $2k$)
- C is **true**: Lower BIC = better balance of fit and parsimony
- D is **false**: BIC is specifically designed to compare models with different numbers of parameters

Formulas:

$$\text{AIC} = -2\ln(\hat{L}) + 2k$$

$$\text{BIC} = -2\ln(\hat{L}) + k\ln(n)$$

Quiz 7: Ljung-Box Test

Question: After fitting an ARMA(2,1) model, you run the Ljung-Box test on residuals and get p-value $= 0.02$. What do you conclude?

A. The model is adequate
B. Residuals are white noise
C. There is significant autocorrelation in residuals
D. The model has too many parameters

Quiz 7: Solution

Answer: C

Explanation: The Ljung-Box test has:

- H_0: Residuals are white noise (no autocorrelation)
- H_1: Residuals have significant autocorrelation

With p-value $= 0.02 < 0.05$:

- We **reject** H_0
- Conclusion: residuals are **not** white noise
- The model is **inadequate** — significant structure remains

Next step: Try a different model (e.g., increase p or q)

Quiz 8: Forecasting

Question: For an AR(1) model with $\phi = 0.6$ and mean $\mu = 10$, what happens to forecasts as horizon $h \to \infty$?

A. Forecasts grow without bound
B. Forecasts converge to 0
C. Forecasts converge to $\mu = 10$
D. Forecasts oscillate forever

Quiz 8: Solution

Answer: C

Explanation: For AR(1), the h-step ahead forecast is:

$$\hat{X}_{n+h|n} = \mu + \phi^h (X_n - \mu)$$

Since $|\phi| = 0.6 < 1$:

$$\lim_{h \to \infty} \phi^h = 0$$

Therefore:

$$\lim_{h \to \infty} \hat{X}_{n+h|n} = \mu + 0 \cdot (X_n - \mu) = \mu = 10$$

Key insight: Long-run forecasts from stationary ARMA models always converge to the unconditional mean.

True/False Questions

Determine if each statement is True or False:

1. An AR(2) process can exhibit pseudo-cyclical behavior.
2. MA processes require a stationarity condition.
3. The PACF of an AR(p) process cuts off after lag p.
4. If AIC selects ARMA(2,1) and BIC selects ARMA(1,1), they cannot both be correct.
5. Forecast confidence intervals narrow as the forecast horizon increases.
6. The Yule-Walker equations can be used to estimate MA parameters.

True/False: Solutions

1. **TRUE**: AR(2) with complex roots shows damped oscillations
2. **FALSE**: MA processes are always stationary; they need *invertibility* condition
3. **TRUE**: This is the key identification feature of AR(p)
4. **FALSE**: Both can be "correct" — they optimize different criteria (AIC favors fit, BIC favors parsimony)
5. **FALSE**: Confidence intervals *widen* as horizon increases (more uncertainty)
6. **FALSE**: Yule-Walker is for AR models only; MA uses MLE

Exercise 1: AR(1) Properties

Problem: Consider the AR(1) process:

$$X_t = 2 + 0.7X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, 9)$$

Calculate:

1. The mean μ
2. The variance $\gamma(0)$
3. The autocovariance $\gamma(1)$ and $\gamma(2)$
4. The autocorrelation $\rho(1)$ and $\rho(2)$

Exercise 1: Solution

Given: $c = 2$, $\phi = 0.7$, $\sigma^2 = 9$

1. **Mean:**

$$\mu = \frac{c}{1-\phi} = \frac{2}{1-0.7} = \frac{2}{0.3} = 6.67$$

2. **Variance:**

$$\gamma(0) = \frac{\sigma^2}{1-\phi^2} = \frac{9}{1-0.49} = \frac{9}{0.51} = 17.65$$

3. **Autocovariance:**

$$\gamma(1) = \phi \cdot \gamma(0) = 0.7 \times 17.65 = 12.35$$

$$\gamma(2) = \phi^2 \cdot \gamma(0) = 0.49 \times 17.65 = 8.65$$

4. **Autocorrelation:**

$$\rho(1) = \phi = 0.7, \quad \rho(2) = \phi^2 = 0.49$$

Exercise 2: MA(1) Properties

Problem: Consider the MA(1) process:

$$X_t = 5 + \varepsilon_t - 0.4\varepsilon_{t-1}, \quad \varepsilon_t \sim WN(0, 4)$$

Calculate:

1. The mean μ
2. The variance $\gamma(0)$
3. The autocovariance $\gamma(1)$
4. The autocorrelation $\rho(1)$
5. Is this process invertible?

Exercise 2: Solution

Given: $\mu = 5$, $\theta = -0.4$, $\sigma^2 = 4$

1. **Mean:**
$$\mathbb{E}[X_t] = \mu = 5$$
2. **Variance:**
$$\gamma(0) = \sigma^2(1 + \theta^2) = 4(1 + 0.16) = 4 \times 1.16 = 4.64$$
3. **Autocovariance at lag 1:**
$$\gamma(1) = \theta\sigma^2 = -0.4 \times 4 = -1.6$$
4. **Autocorrelation:**
$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{-1.6}{4.64} = -0.345$$
5. **Invertibility:** $|\theta| = 0.4 < 1 \rightarrow$ **Yes, invertible**

Exercise 3: Characteristic Roots

Problem: Consider the AR(2) process:

$$X_t = 0.5X_{t-1} + 0.3X_{t-2} + \varepsilon_t$$

1. Write the characteristic equation
2. Find the characteristic roots
3. Is this process stationary?

Exercise 3: Solution

1. Characteristic equation:

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 = 1 - 0.5z - 0.3z^2 = 0$$

Or: $0.3z^2 + 0.5z - 1 = 0$

2. Roots (using quadratic formula):

$$z = \frac{-0.5 \pm \sqrt{0.25 + 1.2}}{0.6} = \frac{-0.5 \pm 1.204}{0.6}$$

$$z_1 = \frac{0.704}{0.6} = 1.17, \quad z_2 = \frac{-1.704}{0.6} = -2.84$$

3. Stationarity check:
Both roots have $|z| > 1$: $|z_1| = 1.17 > 1$ and $|z_2| = 2.84 > 1$
$\rightarrow$ **Stationary** (roots outside unit circle)

Exercise 4: Forecasting

Problem: You have fit an AR(1) model:

$$X_t = 3 + 0.8X_{t-1} + \varepsilon_t, \quad \sigma^2 = 4$$

Given $X_{100} = 20$, calculate:

1. The 1-step ahead forecast $\hat{X}_{101|100}$
2. The 2-step ahead forecast $\hat{X}_{102|100}$
3. The long-run forecast $\hat{X}_{100+h|100}$ as $h \to \infty$
4. The 95% confidence interval for $\hat{X}_{101|100}$

Exercise 4: Solution

Given: $c = 3$, $\phi = 0.8$, $\sigma^2 = 4$, $X_{100} = 20$
Mean: $\mu = \frac{3}{1-0.8} = 15$

1. **One-step forecast:**

$$\hat{X}_{101|100} = c + \phi X_{100} = 3 + 0.8 \times 20 = 19$$

2. **Two-step forecast:**

$$\hat{X}_{102|100} = c + \phi \hat{X}_{101|100} = 3 + 0.8 \times 19 = 18.2$$

3. **Long-run forecast:**

$$\lim_{h \to \infty} \hat{X}_{100+h|100} = \mu = 15$$

4. **95% CI for 1-step:**

$$\text{MSFE}(1) = \sigma^2 = 4, \quad \sqrt{\text{MSFE}(1)} = 2$$

$$CI : 19 \pm 1.96 \times 2 = [15.08, 22.92]$$

Python Exercise 1: Simulate and Fit AR(1)

Task:

1. Simulate 500 observations from AR(1) with $\phi = 0.7$
2. Plot the series and ACF/PACF
3. Fit an AR(1) model and check if $\hat{\phi} \approx 0.7$
4. Examine residual diagnostics

Hint code:

```
np.random.seed(42)
n = 500
phi = 0.7
x = np.zeros(n)
for t in range(1, n):
  x[t] = phi * x[t-1] + np.random.randn()
```

Python Exercise 2: Model Selection

Task:

1. Load a real time series (e.g., stock returns)
2. Check stationarity using ADF test
3. Compare AIC/BIC for ARMA(1,0), ARMA(0,1), ARMA(1,1), ARMA(2,1)
4. Select the best model
5. Generate forecasts with confidence intervals

Key functions:

- `adfuller()` for stationarity test
- `ARIMA(data, order=(p,0,q)).fit()` for fitting
- `results.aic`, `results.bic` for criteria
- `results.get_forecast(h)` for predictions

Python Exercise 3: Diagnostic Checking

Task: After fitting a model, perform complete diagnostics:

1. Plot residuals over time
2. Plot ACF of residuals
3. Create Q-Q plot
4. Run Ljung-Box test
5. Check if AR/MA roots are outside unit circle

Key functions:

- `results.resid` for residuals
- `plot_acf(resid)` for ACF plot
- `stats.probplot(resid)` for Q-Q plot
- `acorr_ljungbox(resid)` for portmanteau test
- `results.arroots`, `results.maroots` for roots

Discussion 1: Model Selection

Scenario: You're modeling monthly inflation rates. After checking stationarity (passed), you find:

- ACF: significant at lags 1, 2, 3, then decays
- PACF: significant at lags 1, 2, then cuts off
- AIC selects ARMA(2,3)
- BIC selects ARMA(2,0) = AR(2)

Questions:

1. What does the ACF/PACF pattern suggest?
2. Why do AIC and BIC disagree?
3. Which model would you choose and why?
4. What additional checks would you perform?

Discussion 2: Forecast Evaluation

Scenario: You fit an ARMA(1,1) model to daily stock returns. The in-sample fit looks good (Ljung-Box p-value $= 0.45$), but out-of-sample RMSE is worse than a simple random walk forecast.

Questions:

1. Is this surprising? Why or why not?
2. What does this tell us about stock return predictability?
3. Should you conclude the ARMA model is useless?
4. What alternatives might you consider?

Hint: Think about the Efficient Market Hypothesis and what ARMA captures vs. what it doesn't (e.g., volatility clustering).

Discussion 3: Real-World Application

Scenario: A central bank economist asks you to forecast quarterly GDP growth for policy planning.

Questions:

1. What preliminary analysis would you do before fitting ARMA?
2. GDP is often non-stationary — how would you handle this?
3. Would you use AIC or BIC for model selection? Why?
4. How would you communicate forecast uncertainty to policymakers?
5. What limitations of ARMA models should you mention?

Key Takeaways from Today's Seminar

1. **AR models:** Current value depends on past values
   - Stationarity: $|\phi| < 1$ for AR(1)
   - PACF cuts off at lag p
2. **MA models:** Current value depends on past shocks
   - Always stationary; invertibility: $|\theta| < 1$ for MA(1)
   - ACF cuts off at lag q
3. **Model selection:** Use ACF/PACF patterns + information criteria
4. **Diagnostics:** Residuals must be white noise (Ljung-Box test)
5. **Forecasting:** Point forecasts converge to mean; uncertainty grows

Next Seminar: ARIMA and Seasonal Models