Seminar 8: ARIMA Forecasting

Multi-Step Forecasts, Pseudo Out-of-Sample Evaluation, and Forecast Uncertainty

Giulio Rossetti*

giuliorossetti94.github.io

March 13, 2026

* email: giulio.rossetti.1@wbs.ac.uk

Disclaimer

Important

Full solutions are available on my.wbs. All exercises are examinable material, not just the ones we covered in the seminars.

Roadmap

Exercise 1: Forecasting with ARMA(1,2)

Exercise 2: Pseudo Out-of-Sample Evaluation

Exercise 3: Forecast Uncertainty from MA(2)

Roadmap

Exercise 1: Forecasting with ARMA(1,2)

Exercise 2: Pseudo Out-of-Sample Evaluation

Exercise 3: Forecast Uncertainty from MA(2)

Why Does This Matter?

Motivation

Producing multi-step ahead forecasts is at the heart of time-series econometrics. Understanding how forecasts are built recursively from an ARMA model is essential for applications in finance and macroeconomics.

Exercise 1: The ARMA(1,2) Process

The process y_t follows an ARMA(1,2):

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \qquad \varepsilon_t \sim N(0, \sigma^2)$$

This is ARMA($p = 1$, $q = 2$): one autoregressive lag, two moving average lags.

Task: Write down the 1-step, 2-step, and 3-step ahead forecasts, standing at time T.

Key rule: When forecasting, replace future ε terms with 0 (their expected value), and replace future y terms with their forecasts.

One-Step Ahead Forecast $(h = 1)$

$T+1$

$$y_{t+1} = \phi y_t + \varepsilon_{t+1} + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}$$

$$E[y_{t+1} | I_t] = E_t[\phi y_t + \varepsilon_{t+1} + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}]$$

$$E_t[y_{t+1}] = \phi y_t + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}$$

Two-Step Ahead Forecast ($h = 2$)

Intuition: Each step builds on the previous forecast (chain rule).

$T+2$

$$y_{t+2} = \phi y_{t+1} + \varepsilon_{t+2} + \theta_1 \varepsilon_{t+1} + \theta_2 \varepsilon_t$$

$$E_t[y_{t+2}] = \phi \underbrace{E_t[y_{t+1}]} + \theta_2 \varepsilon_t$$

$$= \phi^2 y_t + \phi \theta_1 \varepsilon_t + \phi \theta_2 \varepsilon_{t-1} + \theta_2 \varepsilon_t$$

$$= \underbrace{\phi^2 y_t + \varepsilon_t(\phi \theta_1 + \theta_2) + \phi \theta_2 \varepsilon_{t-1}}$$

Three-Step Ahead Forecast ($h = 3$)

Pattern: As h grows, the MA terms wash out and the forecast converges toward the unconditional mean. The AR component $\phi^h y_T$ drives the long-run forecast.

T+3

$$y_{t+3} = \phi y_{t+2} + \varepsilon_{t+3} + \theta_1 \varepsilon_{t+2} + \theta_2 \varepsilon_{t+1}$$

$$E_t[y_{t+3}] = \phi E_t[y_{t+2}]$$

Takeaway: Multi-Step Forecasting

Key Points

- Future shocks ε_{T+h} are replaced with their expectation (zero)
- Future y values are replaced with their forecasts (chain rule)
- MA terms contribute only for $h \leq q$ steps; AR terms decay geometrically
- Forecasts converge to the unconditional mean as $h \to \infty$

Roadmap

Exercise 1: Forecasting with ARMA(1,2)

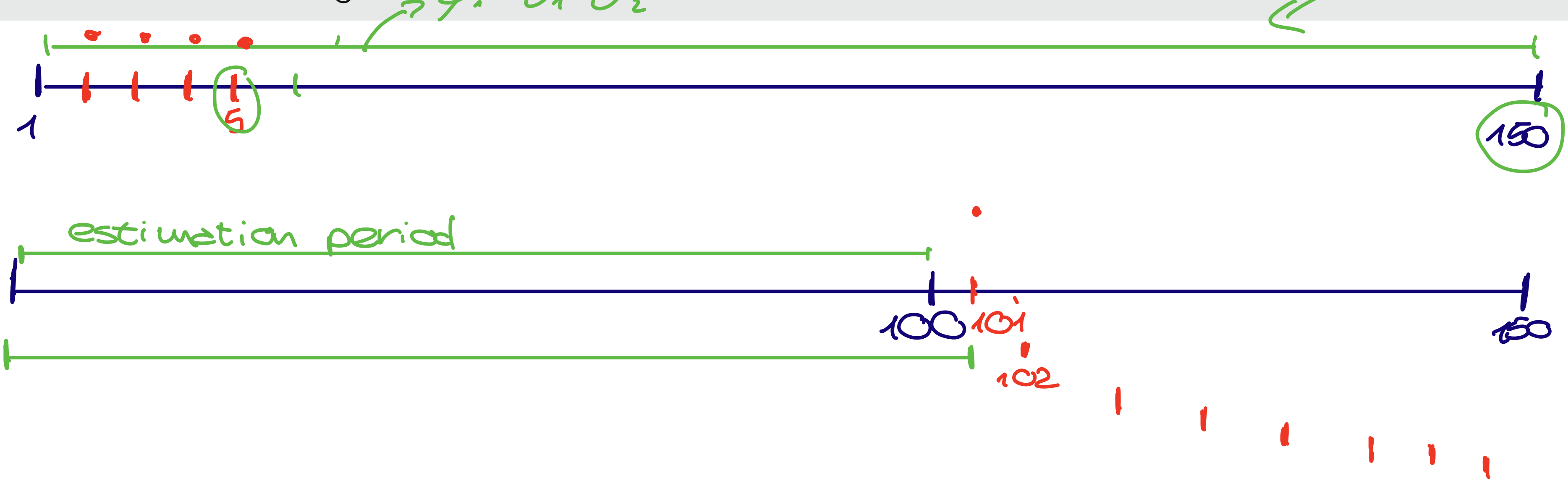
Exercise 2: Pseudo Out-of-Sample Evaluation

Exercise 3: Forecast Uncertainty from MA(2)

Why Does This Matter?

Motivation

In-sample fit can be misleading — a complex model always fits the data better. The pseudo out-of-sample exercise tests whether a model can actually predict the future, by mimicking real-time forecasting.



Exercise 2: The Setup

Goal: Compare a benchmark AR(1) against the ARMA(1,2) model.

Stage 1: Split the data

- Total sample: $t = 1, 2, \ldots, T$
- Initial estimation window: $t = 1, \ldots, R$
- Evaluation period: $t = R + 1, \ldots, T$
- Number of forecasts: $P = T - R$

Example: $T = 150$, $R = 100 \Rightarrow$ 50 forecast origins, 50 forecast errors per model.

Stage 2: Recursive Estimation

Use an expanding window:

1. Estimate both models on $t = 1, \ldots, R$. Forecast $\hat{y}_{R+1}$
2. Re-estimate on $t = 1, \ldots, R + 1$ Forecast $\hat{y}_{R+2}$
3. Continue until estimation window is $t = 1, \ldots, T - 1$. Forecast $\hat{y}_T$

At each forecast origin t, produce h-step ahead forecasts using the same methodology.

Key idea: At each step, we only use information that would have been available at the time — no peeking at future data.

Stage 3: Forecast Error Comparison

Forecast errors: $e_{t+h,t} = y_{t+h} - \hat{y}_{t+h,t}$

Root Mean Square Forecast Error: $RMSE = \sqrt{\frac{1}{P}\sum_{p=1}^{P} e_{t+h,t}^2}$

Comparison: Compute the ratio $RMSE_{\text{ARMA(1,2)}}/RMSE_{\text{AR(1)}}$:

- Ratio < 1: ARMA(1,2) forecasts better
- Ratio > 1: AR(1) benchmark wins

① $\{e_{t+1}, e_{t+2} \ldots\}$ ARMA(1,2)

② $\{e_{t+1}, \ldots\}$ AR(1)

$$\frac{RMSE_{ARMA(1,2)}}{RMSE_{AR(1)}} < 1$$

Formal Testing: The Diebold-Mariano Test

To test whether the difference in forecast accuracy is statistically significant:

Define the loss differential:

$$d_{t+h,t} = e^2_{t+h,t}(\text{ARMA}) - e^2_{t+h,t}(\text{AR})$$

Regress $d_{t+h,t}$ on a constant:

$$d_{t+h,t} = \mu + \epsilon_t, \qquad t = R, \ldots, T-h$$

Test: $H_0 : \mu = 0$ (equal predictive accuracy).

Important: For multi-step forecasts ($h > 1$), forecast errors are serially correlated. Use HAC standard errors (Newey-West) when testing.

Takeaway: Pseudo Out-of-Sample

Key Points

1. Split data into estimation $(1, \ldots, R)$ and evaluation $(R + 1, \ldots, T)$
2. Recursively re-estimate and forecast with an expanding window
3. Compare RMSE (or MAE) across competing models
4. Use the Diebold-Mariano test for formal statistical comparison
5. Check forecast unbiasedness: regress errors on a constant, test if mean is zero

Roadmap

Exercise 1: Forecasting with ARMA(1,2)

Exercise 2: Pseudo Out-of-Sample Evaluation

Exercise 3: Forecast Uncertainty from MA(2)

Why Does This Matter?

Motivation

A point forecast alone is incomplete. We need confidence intervals to quantify uncertainty.

The MA form is particularly useful because it directly reveals the forecast error structure.

Exercise 3: The MA(2) Process

Consider the MA(2):

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \qquad \varepsilon_t \overset{\text{i.i.d.}}{\sim} (0, \sigma_\varepsilon^2)$$

T+1

$$y_{t+1} = \varepsilon_{t+1} + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}$$

$$E[y_{t+1} \mid I_t] = \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}$$

$$e_{t+1,t} = y_{t+1} - E[y_{t+1} \mid I_t] = \varepsilon_{t+1} + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} - \theta_1 \varepsilon_t - \theta_2 \varepsilon_{t-1} = \varepsilon_{t+1}$$

$$var_t(e_{t+1,t}) = var_t(\varepsilon_{t+1}) = \sigma_\varepsilon^2$$

Two-Step Ahead Forecast Uncertainty

2-step ahead:

For $h > 2$**:** $\hat{y}_{T+h|T} = 0$ and $\sigma_h^2 = \sigma^2(1 + \theta_1^2 + \theta_2^2) = \text{Var}(y_t)$.

Intuition: Beyond the MA order, the forecast is just the unconditional mean, and the uncertainty equals the unconditional variance.

$$y_{t+2} = \varepsilon_{t+2} + \theta_1 \varepsilon_{t+1} + \theta_2 \varepsilon_t$$

$$E_t[y_{t+2}] = \theta_2 \varepsilon_t$$

$$e_{t+2|t} = y_{t+2} - E_t[y_{t+2}] = \varepsilon_{t+2} + \theta_1 \varepsilon_{t+1}$$

$$\text{var}_t(e_{t+2|t}) = \text{var}(\varepsilon_{t+2} + \theta_1 \varepsilon_{t+1}) = \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 = \sigma_\varepsilon^2(1+\theta_1^2)$$

T+3

$$y_{t+3} = \varepsilon_{t+3} + \theta_1 \varepsilon_{t+2} + \theta_2 \varepsilon_{t+1}$$

$$E_t[y_{t+3}] = E_t[\varepsilon_{t+3} + \theta_1 \varepsilon_{t+2} + \theta_2 \varepsilon_{t+1}] = 0$$ = unconditional mean MA(2)

$e_{t+3|t} = y_{t+3}$ ...

$\text{var}(e_{t+3,t}) = \sigma_\varepsilon^2 + \theta_1^2\sigma_\varepsilon^2 + \theta_2^2\sigma_\varepsilon^2$

$= \sigma_\varepsilon^2(1 + \theta_1^2 + \theta_2^2)$ = unconditional variance MA(2)

Summary

Key Takeaways

1. **Multi-step forecasts:** Use the chain rule — replace future unknowns with forecasts or zero. MA terms wash out after q steps.
2. **Pseudo OOS:** Split data, recursively re-estimate, compare RMSE. Use Diebold-Mariano for formal testing.
3. **Forecast uncertainty:** The MA representation reveals forecast error variance at each horizon. Intervals widen with the horizon until they reach the unconditional variance.

Key Formulas

Concept	**Formula**
ARMA(1,2)	$y_t = \phi y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$
1-step forecast	$\hat{y}_{T+1\|T} = \phi y_T + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1}$
RMSE	$\sqrt{\frac{1}{P} \sum e_{t+h,t}^2}$
DM loss differential	$d_{t+h,t} = e_{\text{model}}^2 - e_{\text{benchmark}}^2$
MA(2) 1-step var	$\sigma_1^2 = \sigma^2$
MA(2) 2-step var	$\sigma_2^2 = \sigma^2(1 + \theta_1^2)$
95% forecast interval	$\hat{y}_{T+h\|T} \pm 1.96\sigma_h$