1 ARMA Models

1. The Lag Operator

The lag operator L shifts a time series backward by one period:

$$Ly_t = y_{t-1}, \qquad L^2 y_t = y_{t-2}.$$

Consider an AR(1) model with unconditional mean zero ($\mu = 0$):

$$y_t = \phi y_{t-1} + \varepsilon_t.$$

Using the lag operator, we can rewrite the model as

$$\begin{aligned} y_t - \phi y_{t-1} &= \varepsilon_t, \\ y_t - \phi L y_t &= \varepsilon_t, \\ (1 - \phi L) y_t &= \varepsilon_t. \end{aligned}$$

The lag operator allows us to express time-series models in a more compact form.

2. AR, MA, and ARMA Models

For simplicity, we first consider zero-mean AR, MA, and ARMA models ($\mu = 0$), where ε_t is white noise.

With a nonzero unconditional mean μ, the models become:

$$\begin{aligned} \text{AR(1):} \quad & y_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t, \\ \text{MA(1):} \quad & y_t = \mu + \varepsilon_t - \theta \varepsilon_{t-1}, \\ \text{ARMA(1,1):} \quad & y_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t - \theta \varepsilon_{t-1}. \end{aligned}$$

Autoregressive (AR) Model

An AR model expresses the current value of a time series as a function of its **past values and a current shock**.

$$\underbrace{y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t}_{\text{AR}(p)}$$

For an AR(1):

$$y_t = \phi y_{t-1} + \varepsilon_t.$$

Recall the geometric series formula:

$$\frac{1}{1-r} = 1 + r + r^2 + \cdots, \qquad |r| < 1.$$

We can apply the same algebraic idea to the lag operator L. Although L is not a number, we can still add and multiply lag operators, just as we do with ordinary algebraic expressions.

To see why this works, multiply the expansion by $(1 - \phi L)$:

$$\begin{aligned} &(1 - \phi L)(1 + \phi L + \phi^2 L^2 + \cdots) \\ &\quad = 1 + \phi L + \phi^2 L^2 + \cdots \\ &\qquad - \phi L - \phi^2 L^2 - \phi^3 L^3 - \cdots \\ &\quad = 1. \end{aligned}$$

All terms cancel except the first one! Therefore, we can write

$$\frac{1}{1-\phi L} = 1 + \phi L + \phi^2 L^2 + \cdots .$$

If $|\phi| < 1$, the coefficients decay geometrically, and the stationary AR(1) can be represented as an infinite moving average of shocks:

$$\begin{aligned} y_t &= \frac{1}{1-\phi L}\varepsilon_t \\ &= (1 + \phi L + \phi^2 L^2 + \cdots)\varepsilon_t \\ &= \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \cdots . \end{aligned}$$

Here, $L^j\varepsilon_t = \varepsilon_{t-j}$. The condition $|\phi| < 1$ applies to ϕ, not to L.

Moving Average (MA) Model

An MA model expresses the current value of a time series as a weighted sum of **current and past shocks**.

$$\underbrace{y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}}_{\text{MA}(q)}$$

For an MA(1):

$$y_t = \varepsilon_t - \theta\varepsilon_{t-1}.$$

Using the lag operator:

$$y_t = (1 - \theta L)\varepsilon_t.$$

Why is it called a moving average?

The model combines shocks over a moving time window. Unlike an ordinary average, the weights need not be equal or sum to one.

In an MA(1), a shock affects the current and next periods, but has no effect after that.

Autoregressive Moving Average (ARMA) Model

An ARMA model combines the AR and MA components.

$$\underbrace{y_t = \sum_{i=1}^{p}\phi_i y_{t-i} + \varepsilon_t - \sum_{j=1}^{q}\theta_j\varepsilon_{t-j}}_{\text{ARMA}(p,q)}$$

For an ARMA(1,1):

$$y_t = \phi y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}.$$

Using the lag operator:

$$(1 - \phi L)y_t = (1 - \theta L)\varepsilon_t.$$

Model	**Depends on**	**Example**
AR(1)	Past values of y_t	$y_t = \phi y_{t-1} + \varepsilon_t$
MA(1)	Current and past shocks	$y_t = \varepsilon_t - \theta\varepsilon_{t-1}$
ARMA(1,1)	Past values and shocks	$y_t = \phi y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}$

Key Takeaway

Although an AR model is written in terms of past observations, those observations themselves contain past shocks.

For a stationary AR(1), the effect of a shock gradually dies out over time. For an MA(1), the effect disappears completely after one additional period.

3. Forecasting with an MA(1)

We now allow for a nonzero unconditional mean μ. Consider the general MA(1) model:

$$y_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1},$$

where ε_t is white noise.

Suppose T is the last observed period.

Recall that $E_T[\cdot]$ denotes the conditional expectation given information available at time T. Assuming current and past shocks are known and future shocks have zero conditional mean,

$$E_T[\varepsilon_{T+h}] = 0 \quad (h \geq 1), \qquad E_T[\varepsilon_T] = \varepsilon_T.$$

One-step-ahead forecast

$$\begin{aligned}\widehat{y}_{T+1|T} &= E_T(\mu + \varepsilon_{T+1} - \theta\varepsilon_T) \\ &= \boxed{\mu - \theta\varepsilon_T}.\end{aligned}$$

Two-step-ahead forecast

$$\begin{aligned}\widehat{y}_{T+2|T} &= E_T(\mu + \varepsilon_{T+2} - \theta\varepsilon_{T+1}) \\ &= \boxed{\mu}.\end{aligned}$$

General result

$$\boxed{\widehat{y}_{T+h|T} = \begin{cases}\mu - \theta\varepsilon_T, & h = 1, \\ \mu, & h \geq 2.\end{cases}}$$

Thus, MA(1) forecasts return to the unconditional mean μ after one period. When $\mu = 0$, the forecasts return to zero.

Exercise 1: Forecasting with an MA(1)

Suppose y_t follows

$$y_t = \varepsilon_t - 0.5\varepsilon_{t-1},$$

where ε_t is white noise with mean zero.

At time T, the most recent innovation is

$$\varepsilon_T = 2.$$

1. Compute the one-step-ahead forecast $\widehat{y}_{T+1|T}$.

2. Compute the two-step-ahead forecast $\widehat{y}_{T+2|T}$.

3. What happens to the forecast as the forecast horizon becomes large? Explain briefly.

4. How does the forecasting behavior of an MA(1) differ from that of a stationary AR(1)?

2 Forecast Plan

Due: Monday, October 19, at noon The Forecast Plan should be approximately one page, plus a time-series plot.

Requirements

1. Identify the variable you plan to forecast (e.g., the unemployment rate).
2. Specify the agency that publishes the data (e.g., the Bureau of Labor Statistics).
3. State the date of the observation you plan to forecast (e.g., November 2026).
4. Specify where and when the forecasted observation will be released (e.g., the BLS website, December 4, 2026).
5. Include a time-series plot of the historical data. Indicate the data source (e.g., FRED) and add a horizontal line at the sample mean.

Important: The forecasted observation must be released by December 16, 2026.

Note: You do not need to choose or estimate a forecasting model for the **Forecast Plan.**

Example: Finding Data on FRED

FRED (`https://fred.stlouisfed.org`) provides economic time-series data.
Some examples include:

Variable	FRED ID	Frequency
Unemployment Rate	UNRATE	Monthly
Consumer Price Index	CPIAUCSL	Monthly
Nonfarm Payroll Employment	PAYEMS	Monthly
Industrial Production	INDPRO	Monthly
Real GDP	GDPC1	Quarterly

Remember to check the original data-producing agency and the official release calendar.

Example: Plotting the Unemployment Rate in Stata

Stata Access UW–Madison students can download Stata using their NetID through this link.
Sample Code Download `UNRATE.csv` from FRED and save it in your Stata working directory.

```
* ============================================================
* Step 1: Import data
* ============================================================

* Clear all existing data from Stata's memory
clear all

* Set the working directory (replace with your own folder path)
cd "YOUR_FOLDER_PATH"

* Import the unemployment rate data from the CSV file
import delimited "UNRATE.csv", clear


* ============================================================
* Step 2: Create a monthly date variable
* ============================================================

* Convert the date string into a Stata monthly date
gen mdate = monthly(substr(observation_date, 1, 7), "YM")

* Display dates in year-month format (e.g., 2026m10)
format mdate %tm

* Declare monthly time-series data (required for tsline and lag operators)
tsset mdate


* ============================================================
* Step 3: Calculate the sample mean
* ============================================================

* Calculate summary statistics for the unemployment rate
summarize unrate
```

```
* Store the sample mean in a local macro named avg
local avg = r(mean)


* ============================================================
* Step 4: Plot the data and sample mean
* ============================================================

* Plot the unemployment rate with a horizontal line at its sample mean
tsline unrate, ///
    yline(‘avg’, lcolor(red) lpattern(dash)) ///
    title("U.S. Unemployment Rate") ///
    ytitle("Percent") ///
    legend(off)


* ============================================================
* Step 5: Export the figure
* ============================================================

* Save the graph as a PNG file in the working directory
graph export "unemployment.png", replace
```

Answers

Exercise 1 Forecasting with an MA(1).

Suppose

$$y_t = \varepsilon_t - 0.5\varepsilon_{t-1},$$

where $E(\varepsilon_t) = 0$, and $\varepsilon_T = 2$.

1. Compute the one-step-ahead forecast $\widehat{y}_{T+1|T}$.

   From the MA(1) model,

   $$y_{T+1} = \varepsilon_{T+1} - 0.5\varepsilon_T.$$

   Since the future shock has conditional mean zero,

   $$\widehat{y}_{T+1|T} = -0.5\varepsilon_T = -0.5(2) = -1.$$

2. Compute the two-step-ahead forecast $\widehat{y}_{T+2|T}$.

   From the MA(1) model,

   $$y_{T+2} = \varepsilon_{T+2} - 0.5\varepsilon_{T+1}.$$

   Both shocks are in the future, so

   $$\widehat{y}_{T+2|T} = 0.$$

3. What happens to the forecast as the forecast horizon becomes large? Explain briefly.

   For an MA(1), the forecast is

   $$\widehat{y}_{T+h|T} = 0, \qquad h > 1.$$

   Thus, the forecast returns to the unconditional mean of zero after one period.

4. How does the forecasting behavior of an MA(1) differ from that of a stationary AR(1)?

   For a stationary AR(1),

   $$\widehat{y}_{T+h|T} = \mu + \phi^h(y_T - \mu).$$

   Since $|\phi| < 1$, the forecast approaches the unconditional mean as h increases.

   In contrast, for an MA(1), the forecast returns to the unconditional mean after one period.

   Thus, AR(1) forecasts generally exhibit gradual mean reversion, while MA(1) forecasts return to the mean after the one-step-ahead forecast.