Models for Non-Stationary Time Series:

ARIMA Processes - Part I

Week VII: Video 19

STAT 485/685, Fall 2020, SFU

Sonja Isberg

Our Roadmap

1. **Key Ideas:**
   - **Intro and fundamental concepts (Ch. 1-2):** Means, autocovariances and autocorrelations of time series, and the concept of stationarity.
   - **Estimating trends (Ch. 3):** Temporarily ignoring the rest of the variability in the series, how do we just estimate the mean trend?
   - **Models for stationary time series (Ch. 4):** These are the bulk of the key time series models that will be useful to us: MA, AR and ARMA models.
   - **Models for non-stationary time series (Ch. 5)**: Differencing (ARIMA models), transformations, and adding non-zero mean terms.
2. **Building a Model:** Model specification (Ch. 6), Parameter estimation (Ch. 7), Model diagnostics (Ch. 8)
3. **Forecasting (Ch. 9)**
4. Other topics, as time permits.

Models for Non-Stationary Time Series: Introduction

In Chapter 4 (Week 5 & 6 videos), we considered only stationary time series (with zero means), and we developed several models for them: MA, AR and ARMA models.

In Chapter 5 (Week 7 videos), we consider what can be done when the process of interest is *not* stationary (or if it has a constant non-zero mean).

Some techniques we will learn:

- Differencing to obtain a stationary series from a non-stationary one ("ARIMA models")
- Taking transformations to obtain a stationary series from a non-stationary one
- Adding a constant term if the mean is non-zero

Video 19 Learning Objectives

By the end of this video, we should be able to:

- Describe the importance of stationarity for an ARMA process
- Define an integrated autoregressive moving average process of orders p and q and degree d, i.e. ARIMA(p,d,q)
- Given a process that looks like an ARMA process, identify whether or not it is stationary. If not, use differencing to obtain a stationary process
- After using differencing to obtain a stationary process, identify the original process as an ARIMA process and identify the values of p, d and q

Why Is Stationarity Important?

Consider the following "AR(1)" model:

$$Y_t = 5Y_{t-1} + e_t$$

This model is not stationary, because: $|\phi| = |5| = 5 \not< 1$.

Why Is Stationarity Important?

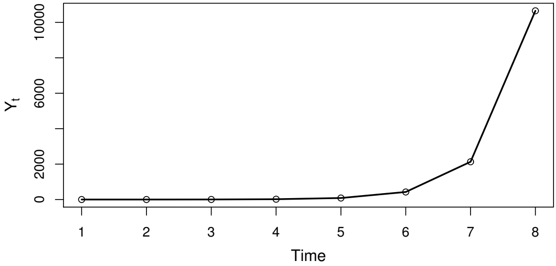
Consider the following "AR(1)" model:

$$Y_t = 5Y_{t-1} + e_t$$

This model is not stationary, because: $|\phi| = |5| = 5 \not< 1$.

In our Chapter 4 videos, we relied on stationarity in our derivations for the autocovariance and autocorrelation functions of the AR models.

Also: How does this "AR(1)" process behave?



Stationarity Through Differencing: Example 1

Recall: The **random walk process** is defined as

$$Y_t = Y_{t-1} + e_t$$

Its mean and autocovariance function are (Video 14, slides 13-14):

$$\mu_t = 0$$
$$\gamma_{t,s} = \min\{t, s\}\,\sigma_e^2$$

So, $\{Y_t\}$ is *not* weakly stationary.

(Note also: This is an “AR(1)” process with $\phi = 1$.)

Stationarity Through Differencing: Example 1 (cont'd)

Recall: The **differenced series** of a process $\{Y_t\}$ defined as

$$\nabla Y_t = Y_t - Y_{t-1}$$

Let $W_t = \nabla Y_t$. Then, for the random walk process:

$$Y_t = Y_{t-1} + e_t$$

$$Y_t - Y_{t-1} = e_t$$

$$W_t = e_t$$

So, $\{W_t\} = \{\nabla Y_t\}$ *is* weakly stationary!

Stationarity Through Differencing: Example 2

Consider the process:

$$Y_t = 1.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

This appears to be an "ARMA(2,1)" process.

Stationarity Through Differencing: Example 2

Consider the process:

$$Y_t = 1.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

This appears to be an "ARMA(2,1)" process.

However, the stationarity conditions are not met:

- The stationarity conditions for an ARMA(p,q) are the same as for the AR(p) model. When $p = 2$, these conditions are:

$$\phi_1 + \phi_2 < 1 \quad \& \quad \phi_2 - \phi_1 < 1 \quad \& \quad |\phi_2| < 1$$

Check: $\phi_1 + \phi_2 = 1.5 - 0.5 = 1 \not< 1$

Stationarity Through Differencing: Example 2

Consider the process:

$$Y_t = 1.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

This appears to be an "ARMA(2,1)" process.

However, the stationarity conditions are not met:

- The stationarity conditions for an ARMA(p,q) are the same as for the AR(p) model. When $p = 2$, these conditions are:

$$\phi_1 + \phi_2 < 1 \quad \& \quad \phi_2 - \phi_1 < 1 \quad \& \quad |\phi_2| < 1$$

Check: $\phi_1 + \phi_2 = 1.5 - 0.5 = 1 \not< 1$

Can we transform this into a stationary process through differencing?

Stationarity Through Differencing: Example 2 (cont'd)

Let $W_t = \nabla Y_t$. Then:

$$Y_t = 1.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

$$Y_t - Y_{t-1} = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

$$W_t = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

$$W_t = 0.5(Y_{t-1} - Y_{t-2}) + e_t - 0.1e_{t-1}$$

$$W_t = 0.5W_{t-1} + e_t - 0.1e_{t-1}$$

This looks like an ARMA(1,1).

Stationarity Through Differencing: Example 2 (cont'd)

Let $W_t = \nabla Y_t$. Then:

$$\begin{aligned} Y_t &= 1.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1} \\ Y_t - Y_{t-1} &= 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1} \\ W_t &= 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1} \\ W_t &= 0.5(Y_{t-1} - Y_{t-2}) + e_t - 0.1e_{t-1} \\ W_t &= 0.5W_{t-1} + e_t - 0.1e_{t-1} \end{aligned}$$

This looks like an ARMA(1,1).

Check the stationarity condition, for $p = 1$: $|\phi| = |0.5| = 0.5 < 1$ ✓

So, $\{W_t\} = \{\nabla Y_t\}$ is a stationary ARMA(1,1) process!

The ARIMA(p,d,q) Process

Definition: A process $\{Y_t\}$ is said to be an **integrated autoregressive moving average process of orders p and q and degree d** (i.e. **ARIMA(p,d,q)**) if:

The d^{th} difference $W_t = \nabla^d Y_t$ is a stationary ARMA(p,q) process.

The ARIMA(p,d,q) Process

Definition: A process $\{Y_t\}$ is said to be an **integrated autoregressive moving average process of orders p and q and degree d** (i.e. **ARIMA(p,d,q)**) if:

The d^{th} difference $W_t = \nabla^d Y_t$ is a stationary ARMA(p,q) process.

In other words:

$$\begin{aligned} W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} \\ + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \end{aligned}$$

for some $\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q$.

The ARIMA(p,d,q) Process

Definition: A process $\{Y_t\}$ is said to be an **integrated autoregressive moving average process of orders p and q and degree d** (i.e. **ARIMA(p,d,q)**) if:

The d^{th} difference $W_t = \nabla^d Y_t$ is a stationary ARMA(p,q) process.

In other words:

$$\begin{aligned} W_t = {} & \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} \\ & + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \end{aligned}$$

for some $\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q$.

In most cases, $d = 1$ or $d = 2$ will suffice.

If there are no AR terms in $\{W_t\}$: $\{Y_t\}$ is ARIMA(0,d,q) = **IMA(d,q)**.

If there are no MA terms in $\{W_t\}$: $\{Y_t\}$ is ARIMA(p,d,0) = **ARI(p,d)**.

Re-Writing Y_t in Difference Equation Form

Suppose $\{Y_t\}$ is an ARIMA(p,1,q) process.

Then: $W_t = \nabla Y_t = Y_t - Y_{t-1}$.

Therefore:

$$\begin{aligned} W_t = &\phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} \\ &+ e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \end{aligned}$$

$$\begin{aligned} Y_t - Y_{t-1} = &\phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \cdots + \phi_p(Y_{t-p} - Y_{t-p-1}) \\ &+ e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \end{aligned}$$

$$\begin{aligned} Y_t = &(1+\phi_1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} + (\phi_3 - \phi_2)Y_{t-3} + \dots \\ &+ (\phi_p - \phi_{p-1})Y_{t-p} - \phi_p Y_{t-p-1} \\ &+ e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \end{aligned}$$

So, $\{Y_t\}$ can be written as a non-stationary "ARMA($p+1$,q)" process.

This is called the **difference equation form** of the model.

Re-Writing Y_t in Difference Equation Form (cont'd)

Suppose $\{Y_t\}$ is an ARIMA(p,2,q) process.

Then:

$$\begin{aligned} W_t &= \nabla^2 Y_t \\ &= \nabla(\nabla Y_t) \\ &= \nabla(Y_t - Y_{t-1}) \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \end{aligned}$$

Then, in its difference equation form, $\{Y_t\}$ can be written as a non-stationary "ARMA($p+2$,q)" process.

Re-Writing Y_t in Difference Equation Form: Usefulness

We've seen that:

- An ARIMA(p,1,q) process can look like a non-stationary "ARMA($p+1$,q)" process
- An ARIMA(p,2,q) process can look like a non-stationary "ARMA($p+2$,q)" process

Re-Writing Y_t in Difference Equation Form: Usefulness

We've seen that:

- An ARIMA(p,1,q) process can look like a non-stationary "ARMA($p+1$,q)" process
- An ARIMA(p,2,q) process can look like a non-stationary "ARMA($p+2$,q)" process

Therefore:

- If we see a non-stationary ARMA process $\{Y_t\}$, and its *first difference* $\{W_t\} = \{\nabla Y_t\}$ turns out to be stationary, we should expect $\{W_t\}$ to have an AR order that is *one less* than that of the original process
- If we see a non-stationary ARMA process $\{Y_t\}$, and its *second difference* $\{W_t\} = \{\nabla^2 Y_t\}$ turns out to be stationary, we should expect $\{W_t\}$ to have an AR order that is *two less* than that of the original process

ARIMA(p,1,q) Example

Consider the process we saw earlier:

$$Y_t = 1.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

This appears to be an "ARMA(2,1)" process.

However, the stationarity conditions are not met:

- The stationarity conditions for an ARMA(p,q) are the same as for the AR(p) model. When $p = 2$, these conditions are:

$$\phi_1 + \phi_2 < 1 \quad \& \quad \phi_2 - \phi_1 < 1 \quad \& \quad |\phi_2| < 1$$

Check: $\phi_1 + \phi_2 = 1.5 - 0.5 = 1 \not< 1$

ARIMA(p,1,q) Example (cont'd)

Let $W_t = \nabla Y_t$. Then:

$$\begin{aligned}
Y_t &= 1.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1} \\
Y_t - Y_{t-1} &= 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1} \\
W_t &= 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1} \\
W_t &= 0.5(Y_{t-1} - Y_{t-2}) + e_t - 0.1e_{t-1} \\
W_t &= 0.5W_{t-1} + e_t - 0.1e_{t-1}
\end{aligned}$$

ARIMA(p,1,q) Example (cont'd)

Let $W_t = \nabla Y_t$. Then:

$$Y_t = 1.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

$$Y_t - Y_{t-1} = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

$$W_t = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$$

$$W_t = 0.5(Y_{t-1} - Y_{t-2}) + e_t - 0.1e_{t-1}$$

$$W_t = 0.5W_{t-1} + e_t - 0.1e_{t-1}$$

This looks like an ARMA(1,1).

Check the stationarity condition, for $p = 1$: $|\phi| = |0.5| = 0.5 < 1$ ✓

So, $\{W_t\} = \{\nabla Y_t\}$ is a stationary ARMA(1,1) process! (Notice: The AR order went down by one.)

Conclusion: $\{Y_t\}$ is an ARIMA(1,1,1) process!

ARIMA(p,1,q) Example 2

The random walk process is defined as

$$Y_t = Y_{t-1} + e_t$$

It looks like an "AR(1)" (i.e., "ARMA(1,0)") process.

However, its stationarity condition is not met: $|\phi| = |1| = 1 \not< 1$

ARIMA(p,1,q) Example 2

The random walk process is defined as

$$Y_t = Y_{t-1} + e_t$$

It looks like an "AR(1)" (i.e., "ARMA(1,0)") process.

However, its stationarity condition is not met: $|\phi| = |1| = 1 \not< 1$

Let $W_t = \nabla Y_t$. Then:

$$\begin{aligned} Y_t &= Y_{t-1} + e_t \\ Y_t - Y_{t-1} &= e_t \\ W_t &= e_t \end{aligned}$$

ARIMA(p,1,q) Example 2

The random walk process is defined as

$$Y_t = Y_{t-1} + e_t$$

It looks like an "AR(1)" (i.e., "ARMA(1,0)") process.

However, its stationarity condition is not met: $|\phi| = |1| = 1 \not< 1$

Let $W_t = \nabla Y_t$. Then:

$$Y_t = Y_{t-1} + e_t$$

$$Y_t - Y_{t-1} = e_t$$

$$W_t = e_t$$

This is just a white noise process, i.e. ARMA(0,0). It is stationary.

So, $\{W_t\} = \{\nabla Y_t\}$ is a stationary ARMA(0,0) process! (Notice: The AR order went down by one.)

Conclusion: $\{Y_t\}$ is an ARIMA(0,1,0) process!

ARIMA(p,2,q) Example

Consider the following process:

$$Y_t = 2Y_{t-1} - Y_{t-2} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

This appears to be an "ARMA(2,2)" process.

However, the stationarity conditions are not met:

- Conditions for $p = 2$:

$$\phi_1 + \phi_2 < 1 \quad \& \quad \phi_2 - \phi_1 < 1 \quad \& \quad |\phi_2| < 1$$

- Check: $\phi_1 + \phi_2 = 2 - 1 = 1 \not< 1$

ARIMA(p,2,q) Example (cont'd)

Let $W_t = \nabla Y_t$. Then:

$$Y_t = 2Y_{t-1} - Y_{t-2} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$Y_t - Y_{t-1} = Y_{t-1} - Y_{t-2} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$W_t = (Y_{t-1} - Y_{t-2}) + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$W_t = W_{t-1} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

ARIMA(p,2,q) Example (cont'd)

Let $W_t = \nabla Y_t$. Then:

$$Y_t = 2Y_{t-1} - Y_{t-2} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$Y_t - Y_{t-1} = Y_{t-1} - Y_{t-2} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$W_t = (Y_{t-1} - Y_{t-2}) + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$W_t = W_{t-1} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

This looks like an ARMA(1,2).

Check the stationarity condition, for $p = 1$: $|\phi| = |1| = 1 \nless 1$

So, $\{\nabla Y_t\}$ is not stationary, either.

ARIMA(p,2,q) Example (cont'd)

Let $W_t' = \nabla^2 Y_t = \nabla W_t$. Then:

$$W_t = W_{t-1} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$W_t - W_{t-1} = e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$W_t' = e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

ARIMA(p,2,q) Example (cont'd)

Let $W_t' = \nabla^2 Y_t = \nabla W_t$. Then:

$$W_t = W_{t-1} + e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$W_t - W_{t-1} = e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

$$W_t' = e_t - 0.4e_{t-1} - 0.2e_{t-2}$$

This looks like an MA(2), i.e. ARMA(0,2). It is stationary.

So, $\{W_t'\} = \{\nabla^2 Y_t\}$ is a stationary ARMA(0,2) process! (Notice: The AR order went down by two from the original process.)

Conclusion: $\{Y_t\}$ is an ARIMA(0,2,2) process! This is also referred to as an IMA(2,2) process.

Final Comments

That's all for now!

In this video, we've seen some examples of how a non-stationary process can be transformed into a stationary process through differencing.

We've also defined the integrated autoregressive moving average process of orders p and q and degree d, i.e. ARIMA(p,d,q).

Finally, we've seen some examples of how to identify an ARIMA process, and obtain stationarity through differencing.

Coming Up Next: Some more properties and examples of ARIMA processes.

Thank you!

References:

[1] Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R.* Springer Science and Business Media.

[2] Chan, K. S., & Ripley, B. (2020). TSA: Time Series Analysis. R package version 1.2.1.