Lecture 4

Time series

Dr. Saman Hussein Mahmood

NONSTATIONARITY

The autoregressive-moving average (ARMA) class of models relies on the assumption that the underlying process is **weakly stationary**, which restricts the mean and variance to be constant and requires the autocovariance to depend only on the time lag. As we have seen, however, many time series are certainly not stationary, for they tend to exhibit time changing means and/or variances.

Definition If d is a nonnegative integer, then {Xt} is an ARIMA(p, d, q) process if Yt : $(1 - B)^d X_t$ is a causal ARMA(p, q) process.

$$\phi^*(B)X_t \equiv \phi(B)(1-B)^d X_t = \theta(B)Z_t, \quad \{Z_t\} \sim \mathrm{WN}\left(0, \sigma^2\right),$$

where φ(z) and θ(z) are polynomials of degrees p and q, respectively, and φ(z)=0 for |z| ≤ 1. The polynomial φ∗(z) has a zero of order d at z 1. The process {Xt} is stationary if and only if d 0, in which case it reduces to an ARMA(p, q) process. Notice that if d ≥ 1, we can add an arbitrary polynomial trend of degree (d -1) to {Xt} without violating the difference equation (6.1.1). ARIMA models are therefore 6.1.1 useful for representing data with trend. It should be noted, however, that ARIMA processes can also be appropriate for modeling series with no trend.

Example

$\{X_t\}$ is an ARIMA(1,1,0) process if for some $\phi \in (-1, 1)$,

$$(1 - \phi B)(1 - B)X_t = Z_t, \quad \{Z_t\} \sim \mathrm{WN}\left(0, \sigma^2\right).$$

We can then write

$$X_t = X_0 + \sum_{j=1}^{t} Y_j, \quad t \geq 1,$$

where

$$Y_t = (1 - B)X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}.$$

When the data shows variation that increases or decreases with the level of the series. And we can do by making more consistent pattern across the data set of the series. Its' not easy to choose from so many methods as each transformation is different from one another and has its own mathematical intuition. In this article we will be covering the Square Root, Logarithmic and Box-Cox Transformations.

Box-Cox transformations

They include both logarithms and power transformations. It depends on the parameter λ, and is defined as-

$$y_t = \begin{cases} \frac{x_t^\lambda - 1}{\lambda}, & \text{if } \lambda \neq 0, \\ \ln x_t, & \text{if } \lambda = 0. \end{cases}$$

$$\begin{aligned}
\lambda &= -1.0, & x_i(\lambda) &= \frac{1}{x_i} \\
\lambda &= -0.5, & x_i(\lambda) &= \frac{1}{\sqrt{x_i}} \\
\lambda &= \ \ 0.0, & x_i(\lambda) &= \ln(x_i) \\
\lambda &= \ \ 0.5, & x_i(\lambda) &= \sqrt{x_i} \\
\lambda &= \ \ 2.0, & x_i(\lambda) &= x_i^2
\end{aligned}$$

To deal with such nonstationarity, we begin by characterizing a time series as the sum of a **nonconstant mean** level plus a random error component:

Unit Roots in Time Series Data

Definition: Unit roots represent non-stationarity, where the mean and variance of a time series are not constant over time.

For example, consider a causal AR(1) process (we assume throughout this section that the noise is Gaussian),

$$x_t = \varphi x_{t-1} + z_t \qquad .....(1)$$

Applying (1 – B) to both sides shows that differencing

$$\text{or} \quad y_t = \varphi y_{t-1} + w_t - w_{t-1}$$

A unit root test provides a way to test whether (1) is a random walk (the null case) as opposed to a causal process (the alternative). That is, it provides a procedure for testing

$$H_0 : \phi = 1 \quad \text{versus} \quad H_1 : |\phi| < 1.$$

1. **Null Hypothesis (H0)**: There exists a unit root in the time series and it is non-stationary. ***Unit root = 1 or δ = 0***
2. **Alternate Hypothesis (H1)**: There exists no unit root in the time series and it is stationary. ***Unit root < 1 or δ < 0***

Condition to reject H0 and accept H1

If the ***test statistic is less than the critical value*** or if the ***p-value is less than a pre-specified significance level (e.g., 0.05)***, then the null hypothesis is rejected and the time series is considered ***stationary***.

Given the estimate $\widehat{\emptyset_T}$ conventional way of testing the null hypothesis would be to construct the t-statistic

$$t_\phi = \frac{\hat{\phi}_T - 1}{\hat{\sigma}_{\hat{\phi}_T}} = \frac{\hat{\phi}_T - 1}{\left(s_T^2 / \sum_{t=1}^{T} x_{t-1}^2\right)^{1/2}}$$

$$\hat{\sigma_{\hat{\phi}_T}} = \left(\frac{s_T^2}{\sum_{t=1}^{T} x_{t-1}^2}\right)^{1/2}$$

Aspect	Unit Circle	Unit Root Tests in Time Series
Definition	A geometric concept representing the complex roots of ARMA processes in the complex plane.	Statistical tests used to evaluate the presence of unit roots and determine time series stationarity.
Nature	Geometric and graphical representation.	Statistical and quantitative analysis.

Dickey Fuller Test

Dickey Fuller test is a statistical test that is used to check for stationarity in time series. This is a type of unit root test, through which we find if the time series is having any unit root. Presence of unit root makes a time series non-stationary

Augmented Dickey Fuller(ADF) Test

Is an extension of Dickey Fuller test for more complex models than AR(1). The primary difference between the two tests is that the ADF is utilized for a larger sized set of time series models which can be more complicated.

Augmented Dickey Fuller test assumes a AR(p) type time series model and it is represented mathematically as,

$$y_t = \mu + \sum_{i=1}^{p} \varphi_i \, y_{t-1} + \varepsilon_t$$

After we substract yt-1 from both the side, we get:

$$\Delta y_t = \mu + \delta y_{t-1} + \sum_{i=1}^{p} \beta_i \, \Delta y_{t-1} + \varepsilon_t$$

ADF is the same equation as the DF with the only difference being the addition of differencing terms representing a larger time series.

The test statistic formula is:

$$t_{\hat{\beta}_i} = \frac{\hat{\beta}_\iota}{SE(\hat{\beta}_\iota)}$$

Assumptions

The test is conducted under following assumptions:

1. **Null Hypothesis (H0)**: There exists a unit root in the time series and it is non-stationary. ***Unit root = 1 or δ = 0***
2. **Alternate Hypothesis (H1)**: There exists no unit root in the time series and it is stationary. ***Unit root < 1 or δ < 0***

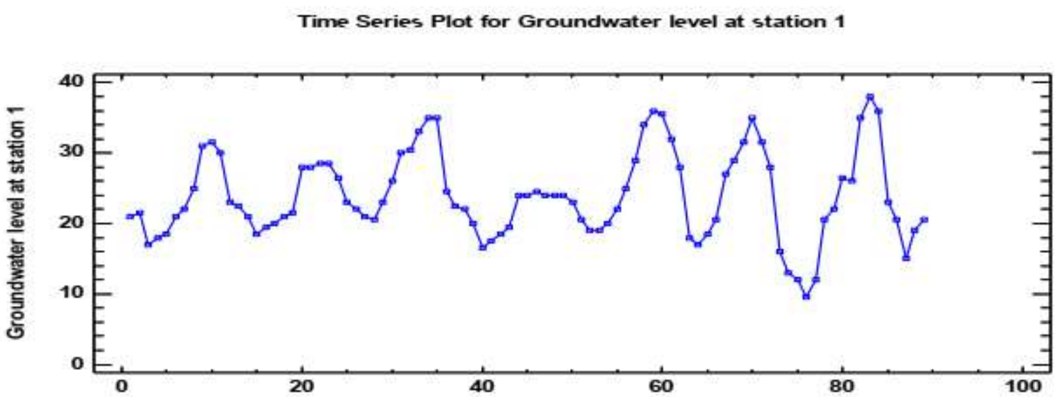
Condition to reject H0 and accept H1

If the ***test statistic is less than the critical value*** or if the ***p-value is less than a pre-specified significance level (e.g., 0.05)***, then the null hypothesis is rejected and the time series is considered ***stationary***.

If the test statistic is greater than the critical value, the null hypothesis cannot be rejected, and the time series is considered ***non-stationary***.

The critical value is found from the Dickey Fuller table (similar to t-table that we use for t-test, we have a table with critical values for Dickey Fuller test).

Example



		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.972437	0.1459
Test critical values:	1% level	-4.065702	
	5% level	-3.461686	
	10% level	-3.157121	

Table (1) explain that the p-value of the Dickey-Fuller test equals (0.1459) and it is greater than (0.05). This result indicates that the data of the time series is **not stationary**

The time series after the first differenced

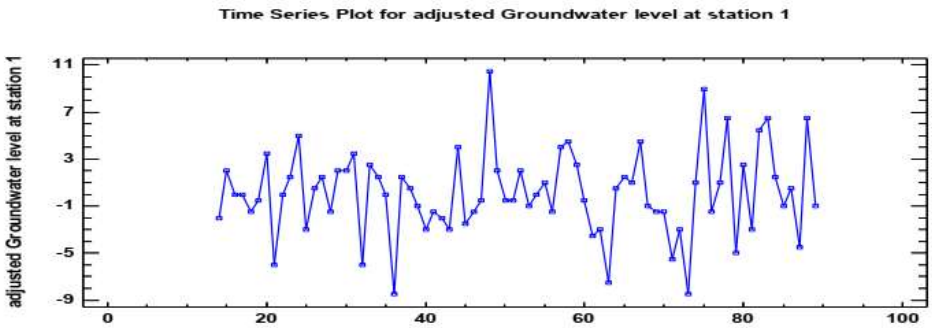


Table 2: shows the results of ADF of the data of the time series of Groundwater level

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.943665	0.0006
Test critical values:	1% level	-4.066981	
	5% level	-3.462292	
	10% level	-3.157475	

The p-value of the Dickey-Fuller test equals (0.0006) and it is less than (0.05). This result indicates that the data of the time series of monthly Groundwater level is stationary

The Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test

KPSS is another test for checking the stationarity of a time series. The null and alternate hypothesis for the KPSS test are opposite that of the ADF test.

Null Hypothesis: The process is trend stationary.

Alternate Hypothesis: The series has a unit root (series is not stationary).

A function is created to carry out the KPSS test on a time series.

Seasonal ARMA Let us assume that there is seasonality in the data, but no trend. Then we could model the data as

$$X_t = s_t + Y_t, \quad ....(4.1)$$

where Y_t is a stationary process. The seasonality component is such that

$$s_t = s_{t-h},$$

where h denotes the length of the period and

$$\sum_{k=1}^{h} s_k = 0.$$

To remove the seasonal effect from the data by varying the lag h. We have introduced the lag-h operator

$$\nabla_h X_t = X_t - X_{t-h} = X_t - B^h X_t = (1 - B^h) X_t,$$

which, for (4.1), gives

$$\nabla_h X_t = s_t + Y_t - s_{t-h} - Y_{t-h} = \nabla_h Y_t.$$

Hence, this operation removes the seasonality effect. This fact leads to introducingthe **seasonal ARMA model**, denoted by $ARMA(P, Q)_h$, which is of the form

$$\Phi(B^h) X_t = \Theta(B^h) Z_t,$$

Where and

$$\Phi(B^h) = 1 - \Phi_1 B^h - \Phi_2 B^{2h} - \ldots - \Phi_P B^{Ph},$$

$$\Theta(B^h) = 1 + \Theta_1 B^h + \Theta_2 B^{2h} + \ldots + \Theta_Q B^{Qh}$$

are, respectively, the seasonal AR operator and the seasonal MA operator, with seasonal period of length *h*.

Remark . Analogously to $ARMA(p, q)$, the $ARMA(P, Q)_h$ model is causal only when the roots of $\Phi(z^h)$ lie outside the unit circle, and it is invertible only when the roots of $\Theta(z^h)$ lie outside the unit circle.

Example . Seasonal $ARMA(1, 1)_{12}$.

Such a model can be written as

$$(1 - \Phi B^{12}) X_t = (1 + \Theta B^{12}) Z_t,$$

or

$$X_t - \Phi X_{t-12} = Z_t + \Theta Z_{t-12},$$

When written as

$$X_t = \Phi X_{t-12} + Z_t + \Theta Z_{t-12},$$

and compared to $ARMA(1, 1)$

$$X_t = \varphi X_{t-1} + Z_t + \vartheta Z_{t-1}$$

we see that the seasonal ARMA presents the series in terms of its past values at lag equal to the length of the period (**here h=12**), while the non-seasonal ARMA does it in terms of its past values at lag 1. Seasonal ARMA incorporates the seasonality into the model.

Similarly as for the non-seasonal ARMA, here too, we require $|\Phi| < 1$ for the causality and $|\Theta| < 1$ for invertibility of the model.

***Example .* ACF of $MA(1)_{12}$**

A seasonal MA model with the period length $h = 12$ can be written as

$$X_t = Z_t + \Theta Z_{t-12}.$$

It is a zero mean stationary model and it is easy to calculate its autocovariance, namely

$$\begin{aligned}
\gamma(\tau) &= \operatorname{cov}[Z_t + \Theta Z_{t-12}, Z_{t+\tau} + \Theta Z_{t+\tau-12}] \\
&= \mathrm{E}[(Z_t + \Theta Z_{t-12})(Z_{t+\tau} + \Theta Z_{t+\tau-12})] \\
&= \mathrm{E}(Z_t Z_{t+\tau}) + \Theta\,\mathrm{E}(Z_t Z_{t+\tau-12}) + \Theta\,\mathrm{E}(Z_{t-12} Z_{t+\tau}) + \Theta^2\,\mathrm{E}(Z_{t-12} Z_{t+\tau-12}) \\
&= \begin{cases} (1+\Theta^2)\sigma^2 & \text{for } \tau = 0, \\ \Theta\sigma^2 & \text{for } \tau = \pm 12, \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Thus, the only non-zero correlations are $\rho(0) = 1$ and

$$\rho(\pm 12) = \frac{\Theta}{1+\Theta^2},$$

which is of the same form as $\rho(\pm 1)$ for a non-seasonal MA(1).

***Example.* ACF of $AR(1)_h$**

Using the techniques for calculating ACVF and ACF of the non-seasonal AR(1) we obtain

Homework

$$\gamma(\tau) = \begin{cases} \frac{\sigma^2}{1-\Phi^2} & \text{for } \tau = 0, \\ \frac{\sigma^2\Phi^k}{1-\Phi^2} & \text{for } \tau = \pm hk,\ k = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

This give the ACF similar to the ACF of a non-seasonal AR(1), namely

$$\rho(\tau) = \begin{cases} 1 & \text{for } \tau = 0, \\ \Phi^k & \text{for } \tau = \pm hk,\ k = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

The following table summarizes the behaviour of the ACF and PACF of the causaland invertible seasonal ARMA models

	$AR(P)_h$	$MA(Q)_h$	$ARMA(P, Q)_h$
ACF	Tails off at lags kh,	Cuts off after lag Qh	Tails off at lags kh
PACF	Cuts off after lag Ph	Tails off at lags kh	Tails off at lags kh

where h is the length of the seasonal period, $k = 1, 2, \ldots$ and the values of ACFand PACF are zero at non-seasonal lags $\tau \neq kh$.

Mixed Seasonal ARMA

When we combine seasonal and non-seasonal operators we obtain a model

$$\Phi(B^h)\varphi(B)X_t = \Theta(B^h)\vartheta(B)Z_t,$$

which is called **mixed seasonal ARMA** and it is denoted by

$$ARMA(p, q) \times (P, Q)_h.$$

We can subtract the effect of the season (say month) using the backshift operator Bh to obtain seasonal stationarity

$X_t - X_{t-h} = (1 - B^h)X_t.$

This is a seasonal difference of order 1. In general we define a seasonal difference of order D as

$$\nabla_h^D X_t = (1 - B^h)^D X_t,$$

where $D = 1, 2, \ldots$. Usually $D = 1$ is sufficient to obtain seasonal stationarity. This leads to a very general **seasonal autoregressive integrated moving average(SARIMA)** model written as follows

$$\Phi(B^h)\phi(B)\nabla_h^D \nabla^d X_t = \alpha + \Theta(B^h)\theta(B)Z_t,$$

Example. The model $ARIMA(0, 1, 1)\times(0, 1, 1)_{12}$ with $\alpha = 0$ is often appliedfor various economic data. Using formula (7.6) we obtain

$$(1 - B^{12})(1 - B)X_t = (1 + \Theta B^{12})(1 + \vartheta B)Z_t$$

or, when expanded, we get the following form

$$(1 - B - B^{12} + B^{13})X_t = (1 + \vartheta B + \Theta B^{12} + \Theta\vartheta B^{13})Z_t$$

Autoregressive Fractionally Integrated Moving Average (ARFMA)

The class of ARIMA processes may be extended to model this type of long-range persistence by relaxing the restriction to just integer values of d, so allowing fractional differencing within the class of AR-fractionally integrated-MA (ARFIMA) processes , and is made operational by considering the binomial series expansion of ∇^d for any real d >-1 :

$$\nabla^d = (1-B)^d = \sum_{k=0}^{\infty} \frac{d!}{(d-k)!k!}(-B)^k$$
$$= 1 - dB + \frac{d(d-1)}{2!}B^2 - \frac{d(d-1)(d-2)}{3!} + \cdots$$

A. When "d" is a positive value in the range (0, 1), it indicates a long-memory process with persistent autocorrelation, suggesting that past values have a lasting influence on future values.
B. When "d" is zero, it corresponds to a stationary time series without long-memory properties
C. When "d" is a negative value in the range (-1, 0), it signifies anti-persistence, where past values have a temporary effect on future values, and the autocorrelation decays rapidly.

Table: Comparison of ARFMA, ARMA, and ARIMA Models

Aspect	ARFMA	ARMA	ARIMA
Components	AR, MA, Fractional Int.	AR, MA	AR, MA, Differencing
Long Memory	Suitable for long memory.	Less effective for LM	May not capture LM
Complexity	More complex	Simpler	Intermediate complexity
Applications	Various, long memory	Stationary data	Stationary with some LM

The standard range for the parameter "d" is typically unrestricted, meaning that "d" can take any real value. However, the more common and widely used range for "d" is (-1, 1), which includes all real numbers between -1 and 1 but excludes -1 and 1 themselves.