**AR MA**  
Notation: • AR(p) • MA(q) • ARMA(p,q) • ARIMA(p,d,q)

**Autoregressive (AR) Models**

• Often you can forecast a series based solely on the past values of Yt . • We are going to focus on the most basic case – only one lag value of 𝑌𝑡 – called an AR(1) model:

**𝑌𝑡 = 𝜔 + 𝜙𝑌𝑡−1 + 𝑒𝑡**

This relationship between t and t-1 exists for all one time period differences across the data set.

𝑌𝑡 = 𝜔 + 𝜙𝑌𝑡−1 + 𝑒𝑡

**𝑌𝑡−1 = 𝜔 + 𝜙𝑌𝑡−2 + 𝑒𝑡−1**

𝑌𝑡−2 = 𝜔 + 𝜙𝑌𝑡−3 + 𝑒𝑡−2

AR(1): The ACF decreases exponentially as the number of lags increases. • The PACF has a significant spike at the first lag, followed by nothing after.

𝑌𝑡 = 0 + 0.8𝑌𝑡−1 + 𝑒t (first spike is 0.8)

**So the effect of shocks that happened long ago has little effect on the present IF the value for 𝜙 < 1.**

This goes back to our idea of stationarity – the dependence of previous observations declines over time.

There is a pattern for AR(1) models when it comes to stationarity.

• If f =1, then Random Walk and NOT Autoregressive model

• If f >1, then today depends on tomorrow (doesn’t really make sense

Notice that RW affect the Acorrelation plots ( all same length spikes in ACF)

RW\_PACF: Only dependent on previous observation. Perfect correlation ( one spike)

**A time series that is a linear function of 2 past values plus error is called an autoregressive process of order 2 – AR(2).**

AR(2) Model 14 𝑌𝑡 = 𝜔 + 𝜙1𝑌𝑡−1 + 𝜙2𝑌𝑡−2 + 𝑒t

There is a pattern in PACF plots for AR(2) models when it comes to stationarity (2 spikes in PACF).

• The effect of shocks that happened long ago has little effect on the present IF the value for 𝜙1 + 𝜙2 < 1.

A time series that is a linear function of p past values plus error is called an autoregressive process of order p – AR(p).

• More complicated restrictions on fi ’s (software will warn you when this becomes an issue)

AR(p) Model 16 𝑌𝑡 = 𝜔 + 𝜙1𝑌𝑡−1 + 𝜙2𝑌𝑡−2 + ⋯ + 𝜙𝑝𝑌𝑡−𝑝 + 𝑒

The PACF has significant spikes at the lags up to p lags, followed by nothing after

You can also **forecast a series based solely on the past error values**. •

We are going to focus on the most basic case – only one error lag value of 𝑒𝑡 , called an MA(1) model:

𝑌𝑡 = 𝜔 + 𝑒𝑡 − 𝜃𝑒𝑡−1

MA(1) model:

This is true for all observations (each observation is dependent on the error from the previous observation).

• Therefore, for an MA(1) model, individual “shocks” only last for a short time.

• In the MA model, we do not have the restrictions that we did on the AR models (but do want them to be invertible)

𝑌𝑡 = 𝜔 + 𝑒𝑡 − 𝜃𝑒𝑡−1

𝑌𝑡−1 = 𝜔 + 𝑒𝑡−1 − 𝜃𝑒𝑡−2

Coor fun MA(1)• The ACF has a significant spike at the first lag, followed by nothing after. • The PACF decreases exponentially as the number of lags increases.

𝑌𝑡 = 0 + 𝑒𝑡 − 0.8𝑒𝑡−( for first lag spike at -1,..)

A time series that is a linear function of q past errors is called a moving average process of order q – called an MA(q). 𝑌𝑡 = 𝜔 + 𝑒𝑡 − 𝜃1𝑒𝑡−1 − 𝜃2𝑒𝑡−2 − ⋯ − 𝜃𝑞𝑒𝑡−q

The ACF has significant spikes at lags up to lag q, followed by nothing after

Any AR(p) model can be rewritten as an MA(∞). • If the MA(q) model is invertible, then this MA(q) model can be rewritten as an AR(∞). • Software should warn you if model is not invertible, if there is no convergence or any other issues….pay attention to the log and any warnings that you encounter when fitting these models. • Depending on how software parameterizes equations, parameters can have different signs.

**WHITE NOISE :**

If we **successfully remove all “correlation” signals, we are left with independent errors**.

TS = signal + Noise

- A white noise **time series have errors that follow a Normal distribution (or bell-shaped) with mean zero and positive, constant variance in which all observations are independent of each oth**er.

**- Autocorrelation and partial autocorrelation functions have a value close to zero at every time point (except for lag of 0).**

The goal of modeling time series is to be left with white noise residuals in the time series.

• If the residuals still have a “significant” dependence structure, then more modeling can typically be done.

How do we know when we are left with white noise at the end of the model? (you already know how to check for normality and constant variance, so we will focus on the dependence structure).

**Ljung Box**

After fitting a model, the Ljung-**Box test may be applied to the original data or the residuals**.

The **null hypothesis is that the series has NO autocorrelation, and the alternative hypothesis is that one or more autocorrelations up to lag m are not zero.**

For non seasonal data until 10 lags

Fitdf =0 raw time series ; Fitdf **= p+q ( AR and MA terms for AR2 its 2),If we have AR4 ( after lag 5 it starts), we have to beyond p+Q number**

 If we have modeled correlation structure in ARIMA , none of P values from Ljung box should be significant, they will be pretty high

Like sanity to check to do for TimeSeries

we do it for residuals ;we need stationary before looking at these residuals

**ARMA Forecasting**

The best part about AR models and MA models is that they are the same thing – approximately

**In certain situations (stationarity), AR models can be represented as an infinite MA model.**

**In certain situations (invertible), MA models can be represented as an infinite AR model**

There is nothing to limit both an AR process and an MA process to be in the model simultaneously. •

These “**mixed” models are typically used to help reduce the number of parameters needed for good estimation in the model.**

𝑌𝑡 = 𝜔 + 𝜙𝑌𝑡−1 + 𝑒𝑡 − 𝜃𝑒𝑡−1

ARMA(p,q) is used to denote mixture models….p indicates the number of autoregressive terms and q represents the number of moving average terms

For example, ARMA(2,3) is the following model:

𝑌𝑡 = 𝜔 + 𝜙1𝑌𝑡−1 + 𝜙2𝑌𝑡−2 + 𝑒𝑡 − 𝜃1𝑒𝑡−1 − 𝜃2𝑒𝑡−2 − 𝜃3𝑒𝑡−3

We also have ARIMA(p,d,q), where p represents the number of autoregressive terms, d represents the number of differences and q represents the number of moving average terms.

Although correlation graphs can potentially help us, they become very complicated with these mixed models. • There are some important things to note: • Characteristics from both are in the correlation functions. • All of the functions tail off exponentially as the lags increase.

ARIMA goal to model correlation term in the data.- We need to have stationary data

Goal is to get simple - so we go mix models

+ or - don't, only struct cares not even greek letters

I-integrated part , differences

Look at ACF and PACF- spikes in data, Based on spikes we know what in model., Try and look at residuals, Try different ones, Can try simpler ones as well., Model that best correlation.

TREND

If you see a visible trend • If there is a trend, the current series is NOT stationary.

• Trending series are not stationary because they do not converge to a mean in the long run.

• One of two things can be happening:

1. The series is stationary ABOUT A REGRESSION LINE (Take away the trend and it is stationary!, This series is stationary)

**Deterministic Trend:** A deterministic trend is what we have done in regression: **𝑌𝑡 = 𝛽0 + 𝛽1𝑡 + 𝜀,** Can also fit quadratic, exponential or any other form of time

2. The series is a Random walk with drift

𝑌𝑡 = 𝜔 + 𝑌𝑡−1 + 𝑒� , w controls the “drift” or the trend (if this is positive, it will “drift” upward; if it is negative, it will “drift downward)

Even if you remove trend line, the resulting residuals are NOT stationary!

Random Walk with drift is NOT stationary if you remove trend line!! Will need to take differences

**Dickey-Fuller Test – Trend**

**null phi= 1 Random walk with drift, phi<1 Deterministic trend, NOT Stochastic trend**

When an obvious trend exists • The series is NOT stationary. •

Need to determine if it is a deterministic trend OR a random walk with drift

**• If it is a deterministic trend, fit a regression line and then use residuals to model AR and MA terms (part of ARIMAX)**

**• If it is a random walk with drift (stochastic), take first difference**

**(** If we look at residuals (y-Yhat) stationary about regression line, we model residuals

 We model regression line,X variable is time,Can also fit for quad,exp,…

If residuals are staionary we can fit AR and MA terms on those residuals., Basically we are focusing on residuals from regression

**Random walk with drift**: Its not same omega- its drift ( drift up +or down-),Bigger trend, larger omega

Trying to fit regression will not help here, even if we remove drift we still have random walk.

How can we tell?Random walk with drft/ stationary about regression line **Type 3 test: Hypothesis same - phi =1 and |phi| <1**

**If we fail to reject atleast one P value - we are going to take diff, If we reject all Hypothesis - its stationary about regression line.**

Stochastic - random walk with drift, Deterministic- I have functional form, like a line.

 I have a trend , I have to deal it first before dealing it with ACF or PACF, We should do Augumented as many lags like before.)

**Missing Values**

Many different ways to interpolate

**Fit a function between points (for example: linear, spline)**

**Last observation carried forward (locf)**

**Weighted moving average**

**Summary statistics (mean or median of the series)**

**Random sample (assume Uniform between two values) • Many more**…

Explore correlation plots (ACF, PACF) to see what patterns there are …VERY circular strategy

What to explore correlations on:

If series is stationary, just use series ggAcf(y.ts)

• If series is a random walk, use differences: • diff.y =diff(y.ts) • ggAcf(diff.y)

• If series is a random walk with drift, use difference: • diff.y =diff(y.ts) • ggAcf(diff.y) (

• If series is stationary about a line, use residuals: • resid.y =lm(y ~time)$resid • ggAcf(resid.y)( If you have a stationary distribution about a regression line, then you will fit a regression line and then use the residuals to model the dependencie)

**Finding mixed models**

* If its stationary - easy look at ACF and PACF to determine which type of correlation.
* If it’s a random walk we are going to see differences
* To look at spikes and model AR and MA terms.

Cleaned data, imputed missing values, determined random data with drift, we take differences.

Focus on modelling AR and MA terms

Here its AR 1 model on differences so its is AR(1,1,0),So it tells R to take diff first, then AR 1

Then look at residuals, if not okay go change AR , MA terms

-After removing all correlation

We look at Ljung box test,We look at residuals ( normal, varaince..),We can fit on validation data set.

(when diff more than once?, like after taking first diff, do ADF test, if its still random walk, we take diff of diff)

Not go beyond 2, make a note of it in ur report :) if you want u can

**Automatic search**

There are a couple of different sets of techniques used for model identification for stationary models.

1. Plotting Patterns – ACF, PACF 2. Automatic Selection Techniques (R and Python):
2. • auto.arima Function
3. 3. Automatic Selection Techniques (SAS..self study): • Minimum Information Criterion – MINIC • Smallest Canonical Correlation – SCAN • Extended Sample Autocorrelation Function – ESAC

If there is a trend, test to see if it is a deterministic trend or random walk with drift.

• **If series has a deterministic trend, fit regression and then use automatic search on residuals**

**• Otherwise, send series through automatic procedure (it will fit a difference if there is a trend)**

**• If there is no trend, you can send series through automatic search.**