MATH1318 TIME SERIES ANALYSIS

- `MA(q) = ACF`
- `AR(p) = PACF`
  AR→
  ↓MA
- PACF→AR(p) ACF→MA(q) SARIMA(pdq)x(PDQ)s
- **MA** Moving average = series moves back and forward across mean

Fundamentals

trend, seasonality, MA/AR, intervention point, variance

Parsimony smallest number of parameters. Adequately represent ts

Random walk neighbouring time points are more correlated than those distant from each other

Stationary probability laws that govern the behaviour of the process do not change over time

Trend

stochastic + correlation between nearby time points with increasing variance over time

deterministic mean function determined beforehand

interpreting regression output

stochastic component (X_t) is white noise

depend on normality of X_t

$\mathbf{R^2}$ high but not too close to 1 = satisfactory fit

Residuals looking for

time-series random/stochastic = no trend

histogram normally distributed

QQ straight line = normal

ACF white noise

Shapiro Wilk H_0: stochastic component of model is normally distributed

Notes from questions

- ? Time series plot *can* be used to identify seasonality in a time series
- ? Seasonality for quarterly values occurs when observations 12 months apart are

related

- ? Having seasonality in time series data does not imply a significant correlation

Correlate with what?

Forecasting

In which situations do we use the model to forecast future values

- ?◀? When the assumptions of the model reasonably well satisfied and the model fits the data well
- ? When the residuals are normally distributed
- When the R-squared value is greater than 90% - NOTE: 99% might overfit

For a **random walk series**:

- ? Variance of the series increases with time
- ? The neighbouring time points are more correlated than those distant from each other
- ?◀? Mean of the series is equal to zero

A model fits the data if:

- ? high value for the residual standard deviation.
- ?◀? normally distributed white noise series for the stochastic component
- ? high negative value for the coefficient of determination
- ? significant regression coefficients

Normality of Residuals [tricked me]

- ? Normal QQ plot of the raw series
- ? Histogram of the transformed series
- ? Time series plot of residuals

estimator of μ is the sample mean for a constant mean model

autoregressive process

- ? The series has a strong autocorrelation between the neighbouring values
- ? PACF of AR(1) process has a positive or negative spike at lag 1 depending on the sign of coefficient then cuts off
- ? ACF of AR(2) process cuts off after lag 2

ACF and PACF

- ? The ACF and PACF are used to find candidate models in practice
- ? The ACF and PACF can be *difficult to calculate* for some data sets
- ? If applied correctly, the ACF and PACF will NOT always deliver unique model selections

◀? **In ARIMA models ‘I’ stands for - Integrated.**

Non-stationary time series steps:

1. Apply a logarithm transformation
2. Compute the first difference
3. Specify model parameters `p and q`

Box-Cox transformation

- ? It is also referred as power transformations
- ? Lambda = 1 implies no transformation
- ? Lambda = 0 log transformation
- ?◀? Lambda can only have positive values
- ? A precise estimate of lambda is usually not warranted

ARIMA(p,d,q) models

- ? For financial time series that the optimal value of `d` could be more than 0
- ? An ARIMA(p,1,q) model estimated on a series of logs of prices is equivalent to an ARIMA(p,0,q) model estimated on a set of continuously compounded returns

ACF plot

- ? We use the ACF to observe the main characteristics of ARMA models
- ? The sample ACF of a white noise series, all autocorrelations should be insignificant at all lags
- ?◀? In the sample ACF of an AR(1) series, all autocorrelations are positive
- ? In the sample ACF of an AR(2) series, there are exponential decays if the roots of the AR equation are real and a damped sine wave if the roots are complex

◀? **EACF** = Extended autocorrelation function

The ACF and PACF plots of a stationary series respectively. [got it wrong]

ARCH more likely to violate non-negative constraints

GARCH(1,1) model will usually be sufficient to capture all of the dependence in the conditional variance

Independence of the noise term of the model

- ◀? sample autocorrelation function of the residuals

Diagnostics checking + over-parameterise the AR(1) model = ◀? ARMA(1,1) & AR(2)

◀? **SARIMA(p,d,q)(P,D,Q)m** Seasonal elements shown by `P,D,Q,m.` - m = seasonality

ACF and PACF (raw series)

- There are strong correlations at lags 12, 24, 36, and so on
- There are significant seasonal autocorrelations in the series
- ◀? ARMA(0,1) and ARMA(0,2) can be considered as candidate models for this series
- ACF and PACF indicates non-stationary series

ACF	tl;dr	Notes
white noise		ACF and PACF insignificant at all lags
`AR(1)` && `0<ϕ<1`	all positive	exponential decays depending on the sign of `ϕ`
`AR(1)` && `−1<ϕ<0`	alternating pattern	starts with negative value
`AR(2)`	decay/sine	equation roots real→exponential decays roots complex→a damped sine wave
`AR(p)`	decay/sine	series tails off as [1] depending on the roots of autocorrelation equation. roots complex→a damped sine wave
`MA(q)`		after lag q insignificant
`ARMA(1,1)`	same as `AR(1)`	exponential decays. Beyond ρ1 `ARMA(1,1)` same pattern as `AR(1)`
! `ARMA(p,q)`	like `AR(p)`	tail off after `lag q` like an `AR(p)`
PACF		
`AR(1)`	`lag 1`	significant +/- spike at lag 1 depending on the sign of ϕ, then cut off
`AR(p)`	`lag q`	`lags 1,2,…,p` will be significant and than they will vanish after lag p
`MA(1)`	`lag 1`	tails off after `lag 1` in one of two forms depending on the sign of `θ`

ACF	tl;dr	Notes
`MA(2)`	decay/sine	equation roots real: exponential decays roots complex: a damped sine wave
`MA(q)`	decay	series tails off as [1:1] depending on the roots of ACF
`ARMA(1,1)`		similar to `MA(1)` & `AR(1)`
`ARMA(p,q)`		contains `MA(q)` process as a special case, [1:2]

---

1. a mixture of exponential decays or damped sine waves ↩? ↩? ↩?