Online Appendix 1. Types of analyses of time series econometric models

DECOMPOSITION (CLASSICAL ANALYSIS)

The classical analysis of time series consists of considering them in a nonrandom way and presupposing that the realization of the series can be conceived as originating from the aggregation of four effects or components (some may not exist): secular trend (T), cyclical variation (C), seasonal variation (S), and erratic variation (E).

Two models of aggregation of these effects are usually considered:

- additive: Yt = Tt + Ct + St + Et
- multiplicative: Yt = Tt.Ct.St. Et (easily convertible to additive, by taking logarithms)

Secular trend: This is the general component of the series and can be considered as the overall movement of the series in the long term, usually obtained or described by fitting a mathematical function or by moving averages or exponential smoothing.

Cyclical variations: These are periodic oscillations that occur with a frequency of more than one year and are usually due to the alternation of periods of economic prosperity (peaks) with periods of depression (troughs).

Seasonal variations: fluctuations with a periodicity of less than one year and recognizable every year, which are usually related to the weather or the behaviour of economic agents when the time of year changes.

Erratic irregular or residual variation: which would reflect the variability in the behaviour of the series that is due to small, unpredictable causes.

Calculate Moving Average

K impaired	$k \text{ impar: } y_t = \frac{1}{k}\sum^{m} x_{t+i},\ m = \frac{k-1}{2}$
K paired	$k \text{ par: } y_t = \frac{1}{2k}\left(x_{t-k/2} + 2\sum_{i=-k/2+1}^{k/2-1} x_{t+i} + x_{t+k/2}\right)$

HOLT-WINTERS MODEL

The Holt Winters method is used to forecast the behaviour of a time series based on previously obtained data. The method is based on an iterative algorithm that at each time (month or week) makes a forecast of the behaviour of the series based on weighted averages of the previous data.

- $L_t = \alpha (Y_t/S_{t-p}) + (1 - \alpha) [L_{t-1} + T_{t-1}]$
- $T_t = \gamma [L_t - L_{t-1}] + (1 - \gamma) T_{t-1}$
- $S_t = \delta (Y_t/L_t) + (1 - \delta) S_{t-p}$
- $\hat{Y}_t = (L_{t-1} + T_{t-1}) S_{t-p}$

Were:

Lt the level at time t, α is the weighting for the level; Tt the trend at time t; γ the weighting for the trend; St the seasonal component at time t; δ the weighting for the seasonal component; p seasonal period; Yt the value of the data at time t; ^ Yt the fitted value, or one-period-ahead forecast, at time t

ARIMA MODEL

The extrapolation forecasts of a univariate ARIMA model were calculated for a time series Y [t] (for t = 1, 2..., T). The user can specify a cut-off period K, which implies that the ARIMA model is estimated based on Y [t] for t = 1, 2..., TK and such that the extrapolation forecast F [t] for t = T -K + 1..., T is calculated and compared with the actual values that were dropped: several extrapolation forecast statistics (MPE, RMSE, MAPE...) are calculated. In addition, the following probabilities P (F [t]> Y [t-1]), P (F [t]> Y [ts]) and P (F [t]> Y [TK]) are calculated.

Given time series data X_t where t is an integer index and the X_t are real numbers, an $\mathrm{ARMA}(p', q)$ model is given by

$$X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{p'} X_{t-p'} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q},$$

or equivalently by

$$\left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right) X_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t$$

where L is the lag operator, the α_i are the parameters of the autoregressive part of the model, the θ_i are the parameters of the moving average part and the ε_t are error terms. The error terms ε_t are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean.

Assume now that the polynomial $\left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right)$ has a unit root (a factor $(1 - L)$) of multiplicity d. Then it can be rewritten as:

$$\left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right) = \left(1 - \sum_{i=1}^{p'-d} \varphi_i L^i\right) (1 - L)^d.$$

An ARIMA(p,d,q) process expresses this polynomial factorisation property with $p=p'-d$, and is given by:

$$\left(1 - \sum_{i=1}^{p} \varphi_i L^i\right) (1 - L)^d X_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t$$

and thus can be thought as a particular case of an ARMA($p+d$,q) process having the autoregressive polynomial with d unit roots. (For this reason, no process that is accurately described by an ARIMA model with $d > 0$ is wide-sense stationary.)

The above can be generalized as follows.

$$\left(1 - \sum_{i=1}^{p} \varphi_i L^i\right) (1 - L)^d X_t = \delta + \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t.$$

This defines an ARIMA(p,d,q) process with **drift** $\frac{\delta}{1 - \sum \varphi_i}$.