Macroeconometrics

Univariate Time Series Analysis
Lectures 1-3
Box-Jenkins Approach (1976)
A four-stage approach to pure time series modeling.

1. selecting appropriate transformation (i.e. log transformation)
2. identification
   graphical display, determining integration level (unit root tests),
   using correlogram to determine the orders of the process
3. estimation
   estimation of the parameters, checking the stationarity, invertibility conditions
4. diagnostic checking.
   residual analysis: residuals should follow a WN
Choosing the model which explains the series better:

Compare: Akaike’s Information Criterion (AIC)
        Schwarz’s Information Criterion (SIC)

The model having smaller value of AIC or SIC proposes a better fit.

Akaike’s Information Criterion is

\[ AIC = \ln \hat{\sigma}_k^2 + \frac{n + 2k}{n} \]

where \( \hat{\sigma}_k^2 \) is the sample variance, \( k \) is the number of the parameters in the models and \( n \) is the number of observations.

The value of \( k \) yielding the minimum AIC specifies the best model.

Corrected AIC (AICc) is a modified AIC for eliminating the bias.

\[ AICc = \ln \hat{\sigma}_k^2 + \frac{n + k}{n - k - 2} \]
Schwarz’s Information Criterion (SIC) (Bayesian Information Criterion (BIC))

\[ SIC = \ln \hat{\sigma}_k^2 + \frac{k \ln n}{n} \]

SIC does well getting the correct order in large samples, whereas AIC tends to be superior in small samples where the relative number of parameters is large.

Durbin-Watson Statistics
Detect the serial correlation in error process. It is an informative statistics for the regression estimations.

\[ DW = \frac{\sum_{t=2}^{n} (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n} \varepsilon_t^2} \]

\(\varepsilon_t\) is the residual from the estimated equation.

\(DW \approx 2 - 2\rho\) where \(\rho\) is the first order serial correlation coefficient.

When \(\rho = 0\) and \(DW \approx 2\).

\(DW < 2\) positive serial correlation; \(DW > 2\) negative serial correlation.
INTEGRATED MODELS of order d, I(d)

Random Walk Model: \( X_t = \mu + X_{t-1} + Z_t \)

- \( Z_t \) is a white noise variable and \( \mu \) is drift. The first difference of the series
- \( \Delta X_t = X_t - X_{t-1} = \mu + Z_t \)
- \( E[\Delta X_t] = E[\mu + Z_t] = \mu \)
- \( \text{Var}[\Delta X_t] = \sigma^2 \)

is independent of time
Trend stationary time series are not mean stationary but include a trend. Including a trend component into the regression model

\[ Y_t = a + bt + \beta X_t + \epsilon_t. \]

Differencing this series increases variance of the error term.

Source: Dr. Fuess
**Difference stationary time series** (which are most of economic time series) contain a stochastic trend, differencing results in a stationary time series.

Therefore: a linear process has a **unit root** if 1 is a root of the process's characteristic equation which leads nonstationarity.
Dickey Fuller Test
Consider \( X_t = \varphi X_{t-1} + Z_t \)
Test \( H_0: \varphi = 1 \) (there exists unit root)
\[ H_a: \varphi < 1 \] (the process is stationary)

Augmented Dickey Fuller Test
\[
\Delta X_t = \mu + \beta t + \varphi^* X_{t-1} + \varphi_1 \Delta X_{t-1} + \ldots + \varphi_p \Delta X_{t-p+1} + Z_t
\]
By including lags of the order \( p \) the ADF formulation allows for higher-order autoregressive processes. This means that the lag length \( p \) has to be determined when applying the test. One possible approach is to test down from high orders and examine the t-values on coefficients. An alternative approach is to examine information criteria such as the Akaike information criterion, Bayesian information criterion or the Hannon Quinn criterion.

Test \( H_0: \varphi^* = 0 \) (there exists unit root) \[ \varphi^* = \varphi - 1 \]
\[ H_a: \varphi^* < 0 \] (the trend stationary process)
Critical values are determined by standard t-tables which may not be useful. MacKinnon (1991) developed a simulation model to determine the critical values for arbitrary sample sizes.

**Sequential test procedure:**

1. Start with a relatively high number of lags, such as 10

2. Subsequently, reduce the number of lags until the last coefficient is significant different from zero at 10% level of significance.

3. Compare the models (without drift and trend, with drift, and with drift and trend) by looking at the Akaike criterion. Then choose the model having the lowest Akaike criterion.

4. If the value of the test statistic is greater than (or in absolute values lesser than) the critical value, fail to reject the existence of unit root.

**The Phillips-Peron Test:** A nonparametric method of controlling for higher order serial correlation in the series. The test statistic follows a t-distribution asymptotically. MacKinnon table values are also used to test unit root.
SEASONAL MODELS

Seasonal differencing, D:
\[ \Delta_4^1 X_t = X_t - X_{t-4} = (1 - B^4) X_t \]
\[ \Delta_4^2 X_t = \Delta_4(\Delta_4 X_t) = \Delta_4(X_t - X_{t-4}) \]
\[ = \Delta_4 X_t - \Delta_4 X_{t-4} = (X_t - X_{t-4}) - (X_{t-4} - X_{t-8}) \]
\[ = (X_t - 2X_{t-4} + X_{t-8}) = (1 - 2B^4 + B^8) X_t \]

In general
\[ \Delta_s^D X_t = Dth \text{ seasonal difference} \]

Seasonal AR(p) with seasonal variations of s period:
Take p=1, AR(1) is
\[ X_t = \varphi X_{t-1} + Z_t \]
\[ (1 - B\varphi)X_t = Z_t \]

A seasonal variations of s=5-period
\[ (1 - \varphi B)(1 - \lambda B^5) X_t = Z_t \]
\[ (1 - \varphi B - \lambda B^5 + \varphi \lambda B^6) X_t = Z_t \]
\[ X_t = \varphi X_{t-1} + \lambda X_{t-5} + \varphi \lambda X_{t-6} + Z_t \]
Seasonal MA(q) with seasonal variations of s period:
Take p=1, MA(1) is
\[ X_t = Z_t + \theta Z_{t-1} \]
\[ X_t = (1 + \theta B)Z_t \]
A seasonal variations of s=5-period
\[ X_t = (1 + \theta B)(1 + \omega B^5)Z_t \]
\[ X_t = (1 + \theta B + \omega B^5 + \theta \omega B^6)Z_t \]
\[ X_t = Z_t + \theta Z_{t-1} + \omega Z_{t-5} + \theta \omega Z_{t-6} \]

Seasonal ARIMA(p,d,q) with seasonal variations of P and Q and a seasonal difference of D and period of s:
ARIMA(p,d,q)_s(P,D,Q)_s is:
\[ (1 - \varphi_1 B - \ldots - \varphi_p B^p)(1 - B)^d (1 - \lambda B^s)^D X_t = (1 + \theta_1 B + \ldots + \theta_q B^q)(1 + \omega B^s)^D Z_t \]
ARIMA(1,1,1)_s(1,1,0)_4 is:
\[ (1 - \varphi_1 B)(1 - B)(1 - \lambda B^4)X_t = (1 + \theta B)Z_t \]
ARIMA(1,1,1)_s(1,1,1)_12 is:
\[ (1 - \varphi B)(1 - B)(1 - \lambda B^{12})X_t = (1 + \theta B)(1 + \omega B^{12})Z_t \]
FORECASTING

The time index $n$ is called the forecast origin and the positive integer $l$ is the forecast horizon.

Let $\hat{X}_n(l)$ be the forecast of $X_{n+l}$ using the minimum squared error loss function and $F_n$ be the collection of information available at the forecast origin $t$, i.e., $F_n = X_n, X_{n-1}, X_{n-2}, \ldots, X_1$.

Then the forecast $\hat{X}_n(l)$ is chosen such that

$$E\left\{[X_{n+l} - \hat{X}_n(l)]^2 | F_n \right\} \leq \min_g E\left\{[X_{n+l} - g]^2 | F_n \right\}$$

where $g$ is a function of the information available at time $n$.

$\hat{X}_n(l)$ is called the $l$-step ahead forecast of $X$, at the forecast origin $n$.

$l$-period ahead with forecast error $e_n(l)$ and forecast error variance $V[e_n(l)]$ are also calculated.
If an ARMA(p,q) model can be written in form
\[ X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=0}^{q} \theta_j Z_{t-j}; \quad \theta_0 = 1 \]

Then forecast of \( l \)-period ahead for \( l = 1 \) is
\[ \hat{X}_n(1) = E[X_{n+1} | F_n] = \sum_{i=1}^{p} \phi_i X_{n+1-i} + \sum_{j=0}^{q} \theta_j Z_{n+1-j} \]

with a forecast error of \( e_n(1) = X_{n+1} - \hat{X}_n(1) = Z_{n+1} \) and forecast error variance of \( \text{Var}[e_n(1)] = \sigma^2 \).

The \( l \)-step ahead forecast \( X_n(l) \) is
\[ X_n(l) = E[X_{n+l} | F_n] = \hat{X}_n(1) = E[X_{n+1} | F_n] = \sum_{i=1}^{p} \phi_i \hat{X}_n(l-i) + \sum_{j=0}^{q} \theta_j Z_{n}(l-i) \]

where \( \hat{X}_n(l-i) = X_{n+l-i} \) if \( l-i \leq 0 \) and \( X_n(l-i) = Z_{n+l-i} \) if \( (l-i) \leq 0 \).

\( e_n(l) = X_{n+l} - \hat{X}_n(l) \)

\( \text{Var}[e_n(l)] = \text{Var}[X_{n+l} - \hat{X}_n(l)] \)
(1-\(\alpha\))x10\% Prediction Limits

\[ X_n(l) \pm z_{\alpha/2} \left\{ V[e_n(l)] \right\}^{1/2} \]

Example: AR(1) Given \(X_n, X_{n-1}, \cdots\) predict \(X_{n+l}\)

\[ X_t - \mu = \phi(X_{t-1} - \mu) + Z_t \]

\(l = 1\)

\[ X_n(1) = E[X_{n+1} \mid X_n \cdots] \]

\[ = E[\mu + \phi(X_n - \mu) + Z_{n+1} \mid X_n, X_{n-1} \cdots] \]

\[ = \mu + \phi(X_n - \mu) \]

\(l = 2\)

\[ X_n(2) = E[X_{n+2} \mid X_n, X_{n-1} \cdots] \]

\[ = E[(\mu + \phi(X_{n+1} - \mu) + Z_{n+2}) \mid X_n, X_{n-1} \cdots] \]

\[ = E[\mu \mid X_n, X_{n-1} \cdots] + \phi \ E[(X_{n+1} - \mu) \mid X_n \cdots] + E[Z_{n+2} \mid X_n \cdots] \]

\[ = \mu + \phi[E[X_{n+1} \mid X_n \cdots] - \mu] \]

\[ = \mu + \phi[X_n(1) - \mu] \]

\[ = \mu + \phi[\mu + \phi(X_n - \mu) - \mu] \]

\[ X_n(2) = \mu + \phi^2(X_n - \mu) \]
l-step ahead prediction

\[ X_n(l) = E(X_{n+l} \mid X_n \cdots) \]

\[ = E(\mu + \varphi(X_{n+l-1} - \mu) + Z_{n+l} \mid X_n, X_{n-1} \cdots) \]

\[ = \mu + \varphi[X_n(l-1) - \mu] \]

\[ = \mu + \varphi'(X_n - \mu) \]

\[ e_n(1) = X_{n+1} - X_n(1) \]

\[ = \{ \mu + \varphi(X_n - \mu) + Z_{n+1} \} - [\mu + \varphi(X_n - \mu)] \]

\[ e_n(1) = Z_{n+1} + 1 \]

\[ e_n(2) = X_{n+2} - X_n(2) \]

\[ = \{ \mu + \varphi(X_{n+1} - \mu) + Z_{n+2} \} - [\mu + \varphi^2(X_n - \mu)] \]

\[ = Z_{n+2} + \varphi[(X_{n+1} - \mu) - \varphi(X_n - \mu)] \]

\[ = Z_{n+2} + \varphi Z_{n+1} \]

\[ e_n(l) = Z_{n+l} + \varphi Z_{n+l-1} + \cdots + \varphi^{l-1} Z_{n+1} \]

\[ V[e_n(l)] = \sigma^2(1 + \varphi^2 + \cdots + \varphi^{2(l-1)}) \]

\[ V[e_n(l)] = \sigma^2 \left( \frac{1 - \varphi^{2l}}{1 - \varphi^2} \right) \]
Efficiency Measures of Forecast

Forecast error statistics are:
Root Mean Squared Error (RMSE)
Mean Absolute Error (MAE)
Mean Absolute Percentage Error (MAPE)
Theil Inequality coefficient (TIC)

RMSE and MAE depend on scale of the dependent variable. Relative measures to compare forecasts for the same series for different models → smaller error better fit

MAPE and TIC are scale invariant.
0<TIC<1. Better fit when TIC→zero