Advanced time-series analysis (University of Lund, Economic History Department)

30 Jan-3 February and 26-30 March 2012

Lecture 4 Econometric techniques for stationary series 1: Univariate stochastic models with Box-Jenkins methodology, simple forecasting techniques.

4.a. Box-Jenkins modeling method: building blocks

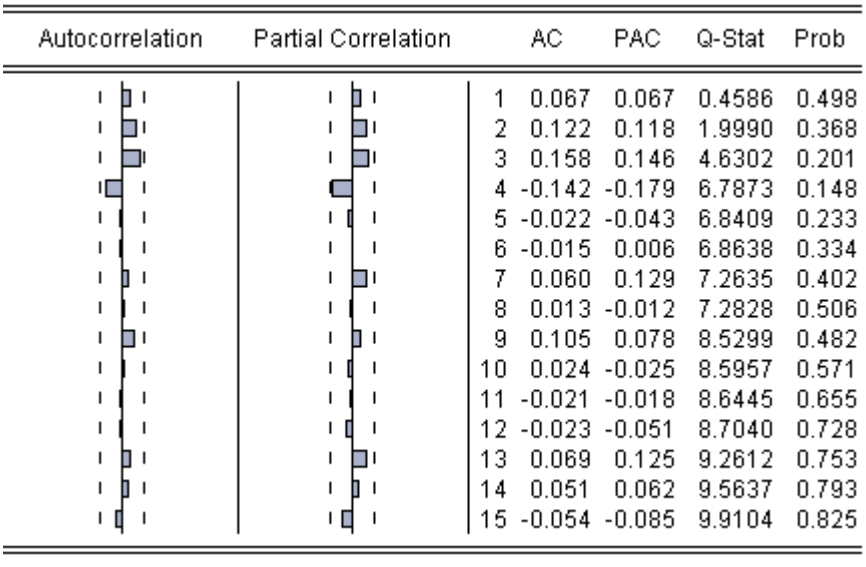
Most of the ideas behind the Box-Jenkins method have implicitly been introduced in the earlier lectures.
The idea is that all covariance stationary time-series can be modeled as the sum of three types of time-series: white noise, autoregressive (AR), and moving average (MA). This technique is often seen as a theoretical, by which it is meant that unlike structural models, the dependent variable is not modeled by a set of exogenous regressors, but by its earlier value. Let us look at these three fundamental building blocks!

White-noise (WN): White noise has the properties of the ideal disturbance (or error) term. It has zero mean, it is homoscedastic and exhibits no serial correlation whatsoever. Time series ε is a white-noise if:

$E(\varepsilon_t) = 0$, $Var(\varepsilon_t) = \sigma^2_{\varepsilon_t} = \sigma^2_{\varepsilon_{t-k}} = \sigma^2_{\varepsilon}$ for all k, and $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$ if k≠0.

The white noise is purely random, it contains no useful information whatsoever. By above definition, a white noise is stationary.

The correlogram of a random standard normal variable (a white-noise)



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.067	0.067	0.4586	0.498
		2	0.122	0.118	1.9990	0.368
		3	0.158	0.146	4.6302	0.201
		4	-0.142	-0.179	6.7873	0.148
		5	-0.022	-0.043	6.8409	0.233
		6	-0.015	0.006	6.8638	0.334
		7	0.060	0.129	7.2635	0.402
		8	0.013	-0.012	7.2828	0.506
		9	0.105	0.078	8.5299	0.482
		10	0.024	-0.025	8.5957	0.571
		11	-0.021	-0.018	8.6445	0.655
		12	-0.023	-0.051	8.7040	0.728
		13	0.069	0.125	9.2612	0.753
		14	0.051	0.062	9.5637	0.793
		15	-0.054	-0.085	9.9104	0.825

Moving average (MA): A moving average process of order q equals a (an optional) constant, white noise term, and the linear combination of the lagged errors from t-1 to t-q.

$y_t = \alpha + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \varepsilon_t$, $\varepsilon_t \sim WN(0, \sigma^2_{\varepsilon})$ an MA(q) process.

The expected value of an MA(q) process is:

$E(y_t) = \alpha_0 + \sum_{j=1}^{q} \beta_j E(\varepsilon_{t-j}) + E(\varepsilon_t) = \alpha_0$, since this does not depend on t, this condition fulfills the requirements of stationarity.

The variance of an MA(q) process is:

$\sigma_{y_t}^2 = \sum_{j=1}^{q} \beta_j^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 = \left(1 + \sum_{j=1}^{q} \beta_j^2\right) \sigma_\varepsilon^2$, that is, as long as q is finite or $\sum_{j=1}^{\infty} \beta_j^2 < \infty$ for all j, this process has a finite variance.

The first-order autocovariance of an MA(1) process is

$$Cov(y_t, y_{t-1}) = E\left((\alpha + \beta_1 \varepsilon_{t-1} + \varepsilon_t)(\alpha + \beta_1 \varepsilon_{t-2} + \varepsilon_{t-1})\right) - E(y_{t-1}) E(y_{t-1}) =$$
$$= \beta_1 E(\varepsilon_{t-1} \varepsilon_{t-1}) = \beta_1 \sigma_\varepsilon^2$$

and zero for higher orders.

The ACF is:

$$\gamma_k = \begin{cases} \dfrac{\beta_1}{(1+\beta_1^2)}, k = 1 \\ 0 \quad , k > 1 \end{cases}$$

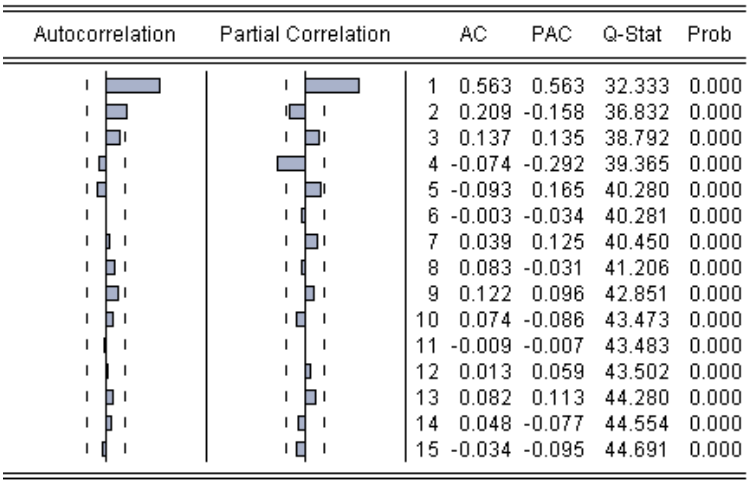
The PACF can be obtained by inverting the MA(1) to an AR(∞) process:

$y_t = (1+\beta L)\varepsilon_t \rightarrow (1+\beta L)^{-1} y_t = \varepsilon_t \rightarrow \sum_{k=o}^{\infty} \left(-\beta_1 L\right)^k y_t = \varepsilon_t$ or

$y_t - \beta_1 y_{t-1} + \beta_1^2 y_{t-2} - \beta_1^3 y_{t-3} + \ldots = \varepsilon_t$ or $y_t = \beta_1 y_{t-1} - \beta_1^2 y_{t-2} + \beta_1^3 y_{t-3} - \ldots + \varepsilon_t$

Where the coefficients are the respective PACF coefficients.

Correlogram of the MA(1) process: $y_t = \varepsilon_t + .8\varepsilon_{t-1}$



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.563	0.563	32.333	0.000
		2	0.209	-0.158	36.832	0.000
		3	0.137	0.135	38.792	0.000
		4	-0.074	-0.292	39.365	0.000
		5	-0.093	0.165	40.280	0.000
		6	-0.003	-0.034	40.281	0.000
		7	0.039	0.125	40.450	0.000
		8	0.083	-0.031	41.206	0.000
		9	0.122	0.096	42.851	0.000
		10	0.074	-0.086	43.473	0.000
		11	-0.009	-0.007	43.483	0.000
		12	0.013	0.059	43.502	0.000
		13	0.082	0.113	44.280	0.000
		14	0.048	-0.077	44.554	0.000
		15	-0.034	-0.095	44.691	0.000

Autoregressive (AR) model: An autoregressive process of order p equals a (an optional) constant, white noise term, and the linear combination of the lagged dependent variable from t-1 to t-p.

$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \varepsilon_t \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ an AR(p) process.

The expected value of an AR(p) process is:

$E(y_t) = \alpha_0 + \sum_{i=1}^{p} \alpha_i E(y_{t-i}) = \frac{\alpha_0}{1 - \sum_{i=1}^{p} \alpha_i}$, provided $\sum_{i=1}^{p} \alpha_i < 1$, the expected value of an AR(p) process is finite.

The variance of an AR(p) process is:

$\sigma_{y_t}^2 = \sum_{i=1}^{p} \alpha_i^2 \sigma_{y_{t-i}}^2 + \sigma_\varepsilon^2 \rightarrow \sigma_{y_t}^2 = \frac{\sigma_\varepsilon^2}{1 - \sum_{i=1}^{p} \alpha_i^2}$, that is, as long as $\sum_{i=1}^{p} \alpha_i^2 < 1$, this process has a finite positive variance.

The autocovariance of an AR(1) process is easily obtained by inverting it to an infinite MA process:

$(1 - \alpha_1 L) y_t = \alpha_0 + \varepsilon_t \rightarrow$

$y_t = (1 - \alpha_1 L)^{-1} (\alpha_0 + \varepsilon_t) = (1 + \alpha_1 L + \alpha_1^2 L^2 + \alpha_1^3 L^3 + \ldots)(\alpha_0 + \varepsilon_t) = y_0 + \sum_{i=0}^{t-1} \alpha_0 \alpha_1^i + \sum_{i=0}^{t-1} \varepsilon_{t-i} \alpha_1^i$

Front his it is obvious that the dynamic multipliers are α_1^k for all lags k. The covariance is:

$\gamma_k = \alpha_1 \sigma_y^2$

$\gamma_0 = \sigma_y^2 = \sigma_\varepsilon^2 \sum_{i=0}^{t-1} \alpha_1^{2i} = \frac{\sigma_\varepsilon^2}{1 - \alpha_1^2}$

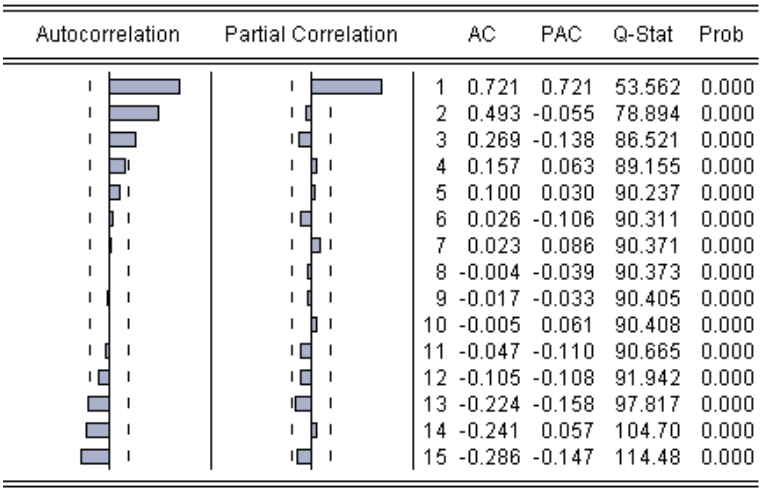
The ACF is:

$ACF_k = \alpha_1^k$

The PACF is obvious:

$PACF(k) = \begin{cases} q, & \text{if } k = 1 \\ 0, & \text{if } k > 1 \end{cases}$

Correlogram of the AR(1) process: $y_t = 0.7 y_{t-1} + \varepsilon_t$



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.721	0.721	53.562	0.000
		2	0.493	-0.055	78.894	0.000
		3	0.269	-0.138	86.521	0.000
		4	0.157	0.063	89.155	0.000
		5	0.100	0.030	90.237	0.000
		6	0.026	-0.106	90.311	0.000
		7	0.023	0.086	90.371	0.000
		8	-0.004	-0.039	90.373	0.000
		9	-0.017	-0.033	90.405	0.000
		10	-0.005	0.061	90.408	0.000
		11	-0.047	-0.110	90.665	0.000
		12	-0.105	-0.108	91.942	0.000
		13	-0.224	-0.158	97.817	0.000
		14	-0.241	0.057	104.70	0.000
		15	-0.286	-0.147	114.48	0.000

Higher order AR or MA processes:

Not as if you ever needed this to do by hand, you can guess how the ACF or PACF of a higher order system would look by factorization:

$\left(1-\alpha_1 L-\alpha_1^2 L^2\right) y_t = \varepsilon_t$ this is an AR(2) process, which can be factorized into:

$$\left(1-\alpha_1 L-\alpha_1^2 L^2\right) y_t = \left(1-\lambda_1 L\right)\left(1-\lambda_2 L\right) y_t = \varepsilon_t$$

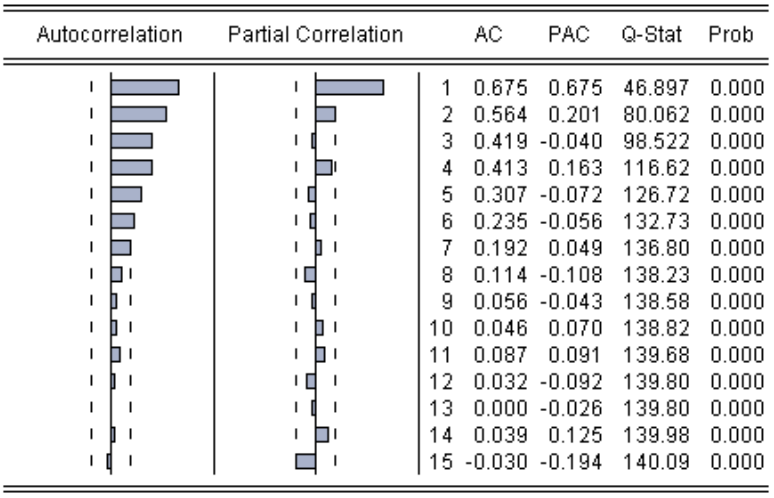
Where λ_1 and λ_2 are the characteristic roots we already know.

$$\left(1-\lambda_2 L\right) y_t = (1+\lambda_1 L+\lambda_1^2 L^2+\ldots)\varepsilon_t \rightarrow$$

$$y_t = (1+\lambda_1 L+\lambda_1^2 L^2+\ldots)(1+\lambda_2 L+\lambda_2^2 L^2+\ldots)\varepsilon_t = \left(1+\theta L+\theta^2 L^2+\ldots\right)\varepsilon_t$$

So actually what you have learned for first order processes will do for higher order systems.

Correlogram of the AR(2) process: $y_t = 0.5 y_{t-1} + 0.2 y_{t-1} + \varepsilon_t$



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.675	0.675	46.897	0.000
		2	0.564	0.201	80.062	0.000
		3	0.419	-0.040	98.522	0.000
		4	0.413	0.163	116.62	0.000
		5	0.307	-0.072	126.72	0.000
		6	0.235	-0.056	132.73	0.000
		7	0.192	0.049	136.80	0.000
		8	0.114	-0.108	138.23	0.000
		9	0.056	-0.043	138.58	0.000
		10	0.046	0.070	138.82	0.000
		11	0.087	0.091	139.68	0.000
		12	0.032	-0.092	139.80	0.000
		13	0.000	-0.026	139.80	0.000
		14	0.039	0.125	139.98	0.000
		15	-0.030	-0.194	140.09	0.000

<u>Autoregressive Moving Average (ARMA) model:</u>

A combination of the AR and MA processes:

$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_i \varepsilon_{t-i} + \varepsilon_t$, $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ this is an ARMA(p,q) model.

Using lag operator:

$(1-\alpha_1 L-\alpha_2 L^2-\ldots-\alpha_p L^p) y_t = \alpha_0 + (1+\beta_1 L+\beta_2 L^2+\ldots+\beta_q L^q)\varepsilon_t$ or

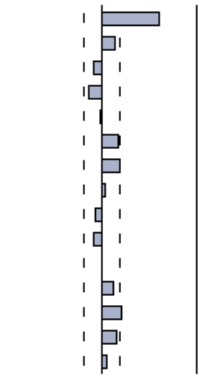
$\alpha(L) y_t = \alpha_0 + \beta(L)\varepsilon_t$ or

$$y_t = \alpha_0 + \frac{\beta(L)}{\alpha(L)}\varepsilon_t$$

Both the ACF and the PACF are dominated by the exponential decay.

Correlogram of the ARMA(1,1) process: $y_t = 0.5 y_{t-1} + 0.2\varepsilon_{t-1} + \varepsilon_t$

Date: 01/27/12 Time: 17:52
Sample: 0001 0100
Included observations: 100



		AC	PAC	Q-Stat	Prob
1		0.627	0.627	40.466	0.000
2		0.148	-0.402	42.760	0.000
3		-0.087	0.069	43.550	0.000
4		-0.144	-0.079	45.757	0.000
5		-0.021	0.190	45.805	0.000
6		0.182	0.133	49.391	0.000
7		0.204	-0.104	53.963	0.000
8		0.036	-0.105	54.104	0.000
9		-0.080	0.068	54.820	0.000
10		-0.084	0.025	55.612	0.000
11		-0.007	0.056	55.618	0.000
12		0.131	0.074	57.598	0.000
13		0.217	0.044	63.119	0.000
14		0.170	0.026	66.561	0.000
15		0.065	0.016	67.069	0.000

4.b. Estimation

The process is simple: if we have a covariance stationary time series, it will work fine.

Step 1.: Check if your series are stationary with a unit root test. If your series are stationary, go ahead. If not, then take first difference and check for stationarity again. The ARMA models can only be fitted on covariance stationary series! If your variable was non-stationary, the resulting Box-Jenkins type model is referred to as ARIMA(*p*,*d*,*q*) where *d* denotes the order of integration for the variable.

Step 2: You will need to estimate an initial model: this requires that you look at the correlogram of the stationary variable. Nevertheless, with real data you will rarely be able to immediately find the right model. Possibly you take a good guess based on the correlogram and estimate a first model.

Step 3: You check if the assumptions of the model are fulfilled: these require two things: a. that the residual is a white noise process. You should look at the residual correlogram and especially the Q test. You should go on with modeling until no serial correlation is to be found in your residual.
b. Several other models can lead to serially uncorrelated residuals, but you need to find a parsimonious model (that fit without unnecessary explanatory variables). This may require that you fit several similar models and choose the one with the best information criteria (Schwarz or Akaike).

Ideally you have a small portfolio of models that all lead to serially uncorrelated residual, but one of them will have the best fit. That is the right model.

An ARIMA(1,1,1) model can be written with lag operators as follows:

$$(1-L)(1-\alpha_1 L)\,y_t = (1+\beta_1 L)\,\varepsilon_t$$

4.c. Why all this?

You may ask yourself why such an in economic sense atheoretical mosel is useful to you?
Well, there are several cases when you need to rely on a Box-Jenkins method:

1. It is good for forecasting, actually for short-run forecasting it is almost unbeatable even by complex structural models.
2. Understanding this method helps you to learn a lot about time-series.
3. Sometimes you do not want to explain the expected value of the variable but you are more interested in its variance (volatility): now an ARMA is a good first step. See: ARCH, GARCH
4. You can use it for policy evaluation (intervention analysis). This means that you model a process and use dummy variables to capture the effect of some change or event.