Univariate Time Series

Kevin Sheppard
http//www.kevinsheppard.com

Oxford MFE

This version: November 20, 2009

November 26, 2009

Time Series Questions

- What is stationarity and why is it important?
- What is an ARMA process?
- What are the statistical properties of an ARMA process?
- When are ARMAs stationary?
- What is the autocorrelation function (ACF) and partial autocorrelation function (PACF)? What good are they?
- How can I estimate autocorrelations and partial autocorrelations and make inference?
- How does the Box-Jenkins model for time-series analysis work?
- How do I forecast from an ARMA?
- How are forecasts evaluated?
- What is a non-stationary process and a unit root?
- How do I test for unit roots?

Stochastic Processes

- Stochastic process

$$\{y_t\}$$

- Examples:
  - i.i.d. $$y_t \overset{\text{i.i.d.}}{\sim} N(0,1)$$
  - Random Walk $$y_t = y_{t-1} + \epsilon_t$$
  - ARMA(1,1) $$y_t = \phi_1 y_{t-1} + \theta \epsilon_{t-1} + \epsilon_t$$
  - GARCH(1,1) $$y_t \sim N(0, \sigma_t^2)$$ $$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$$
  - Many more...
- Today's focus: ARMA

Autocovariance

Definition (Autocovariance)

The autocovariance of a covariance stationary scalar process $\{y_t\}$ is defined

$$\gamma_s = \mathrm{E}\left[(y_t - \mu)(y_{t-s} - \mu)\right]$$

where $\mu = \mathrm{E}\left[y_t\right]$. Note that $\gamma_0 = \mathrm{E}\left[(y_t - \mu)(y_t - \mu)\right] = \mathrm{V}\left[y_t\right]$.

- Covariance of a process at different points in time
- Otherwise identical to usual covariance

Stationarity

- Key concept
- Stationarity is a statistically meaningful form of regularity
- First type:

Definition (Covariance Stationarity)

A stochastic process $\{y_t\}$ is covariance stationary if

$$
\begin{aligned}
\mathrm{E}\left[y_t\right] &= \mu && \text{for } t = 1, 2, \ldots \\
\mathrm{V}\left[y_t\right] &= \sigma^2 < \infty && \text{for } t = 1, 2, \ldots \\
\mathrm{E}\left[(y_t - \mu)(y_{t-s} - \mu)\right] &= \gamma_s && \text{for } t = 1, 2, \ldots, s = 1, 2, \ldots, t-1
\end{aligned}
$$

- *Unconditional* mean, variance and autocovariance do *not* depend on time

Stationarity

Second type (stronger):

Definition (Strict Stationarity)

A stochastic process $\{y_t\}$ is strictly stationary if the joint distribution of $\{y_t, y_{t+1}, \ldots, y_{t+h}\}$ only depends only on h and not on t.

- *Entire joint distribution* does not depend on time.
- Examples of stationary time series:
  - i.i.d. : Always strict, covariance if $\sigma^2 < \infty$
  - i.i.d. sequence of t_2 random variables, strict only
  - Multivariate normal, both
  - AR(1): $y_t = \phi_1 y_{t-1} + \epsilon_t$, covariance if $|\phi_1| < 1$ and $\mathrm{V}[\epsilon_t] < \infty$, strict is ϵ_t is i.i.d.
  - ARCH(1): $y_t \sim N(0, \sigma_t^2)$,$\sigma_t^2 = \omega + \alpha y_{t-1}^2$ both if $\alpha < 1$.

What processes are not stationary?

- Nonstationary time series:
  - Seasonalities, Diurnality, Hebdomadality: $y_t + \mu + I_{\text{Monday}} + \epsilon_t$
    - $\mathrm{E}[y_t]$ is different on Monday than the rest of the week
  - Trending: $y_t = t + \epsilon_t$
    - $\mathrm{E}[y_t] = t$
  - Random walks: $y_t = y_{t-1} + \epsilon_t$
    - $\mathrm{V}[y_t] = t\sigma^2$
  - Processes with structural breaks: $y_t = \mu_1 + \epsilon_t$ if $t < 1974$, $y_t = \mu_2 + \epsilon_t$, $t \geq 1974$.
    - $\mathrm{E}[y_t] = \mu_1 + (\mu_2 - \mu_1)(1 - I_{t<1974})$

Ergodicity

- Measure of "asymptotic independence"

Theorem (Ergodic Theorem)

If $\{y_t\}$ is ergodic and the r^{th} moment μ_r is finite, then $T^{-1}\sum_{t=1}^{T} y_t^r \xrightarrow{p} \mu_r$

- Asymptotic independence ensures that averages that use points far apart in time converge to their expected value
- Example of a nonergodic process:

$$y_t = \mu + \epsilon_t$$

  - $\mu \sim N(0,1)$ and $\epsilon_t \overset{\text{i.i.d.}}{\sim} N(0,1)$
  - $\mathrm{E}[y_t] = 0$
  - $T^{-1}\sum_{t=1}^{T} y_t \xrightarrow{p} \mu \neq 0$
  - μ has a permanent effect on all y_t

White noise

- White Noise:

Definition (White Noise)

A process $\{\epsilon_t\}$ is known as white noise if

$$
\begin{aligned}
\mathrm{E}[\epsilon_t] &= 0 && \text{for } t = 1, 2, \ldots \\
\mathrm{V}[\epsilon_t] &= \sigma^2 < \infty && \text{for } t = 1, 2, \ldots \\
\mathrm{E}\left[\epsilon_t \epsilon_{t-j}\right] &= 0 && \text{for } t = 1, 2, \ldots, \ j \neq 0
\end{aligned}
$$

White noise

- White Noise:

Definition (White Noise)

A process $\{\epsilon_t\}$ is known as white noise if

$$\begin{aligned} \mathrm{E}\left[\epsilon_t\right] &= 0 && \text{for } t = 1, 2, \ldots \\ \mathrm{V}\left[\epsilon_t\right] = \sigma^2 &< \infty && \text{for } t = 1, 2, \ldots \\ \mathrm{E}\left[\epsilon_t \epsilon_{t-j}\right] &= 0 && \text{for } t = 1, 2, \ldots,\ j \neq 0 \end{aligned}$$

- Basic building block of time series
- Not necessarily independent
  - ARCH(1) process $y_t \sim N(0, \sigma_t^2)$, $\sigma_t^2 = \omega + \alpha y_{t-1}^2$
  - Variance is dependant, mean is not

The information set and the law of iterated expectations

- Information set: $\mathcal{F}_t$
- Contains a lot of information!
  - Every time t *measurable* event
  - Observed variables: prices, returns, GDP, interest rates, FX, sun spot activity
  - Functions of these
  - Excludes variables which are latent: volatility
- Conditional expectation:

$$\mathrm{E}[y_{t+1}|\mathcal{F}_t]$$

- Conditional Variance

$$\mathrm{V}[y_{t+1}|\mathcal{F}_t]$$

  - Shorthand $\mathrm{E}_t[y_{t+1}]$ and $\mathrm{V}_t[y_{t+1}]$
- Law of Iterated Expectation (LIE):

$$E_t[E_{t+1}[y_{t+2}]] = E_t[y_{t+2}]$$

  - Monday's belief about what Tuesday's belief about Wednesday is the same as Monday's belief of Wednesday

Linear Time-series Processes

- Standard tool of time-series analysis
- *Linear* time series process can always be expressed as

$$y_t = y_0 + \sum_{i=0}^{t} \theta_i \epsilon_{t-i}$$

  - Linear in the errors
- Example of non-linear processes
  - GARCH(1,1)

$$y_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$$

  - Threshold Autoregression

$$y_t = \phi_s y_{t-1} + \epsilon_t$$

$$\phi_s = \begin{cases} 1 \text{ if } L < y_t < U \\ .9 \text{ otherwise} \end{cases}$$

ARMA Processes

- Inclusive class of all linear time-series processes

Definition (Autoregressive-Moving Average Process)

An Autoregressive Moving Average process with orders P and Q, abbreviated ARMA(P,Q), has dynamics which follow

$$y_t = \phi_0 + \sum_{p=1}^{P} \phi_p y_{t-p} + \sum_{q=1}^{Q} \theta_q \epsilon_{t-q} + \epsilon_t$$

where ϵ_t is a white noise process with the additional property that $E_{t-1}[\epsilon_t] = 0$.

- ARMA(1,1)

$$y_t = \phi_1 y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

Special cases of ARMA processes: Moving Average

- ARMA family compromises two sub-classes

Definition (Moving Average Process of Order Q)

A Moving Average process of order Q, abbreviated MA(Q), has dynamics which follow

$$y_t = \phi_0 + \sum_{q=1}^{Q} \theta_q \epsilon_{t-q} + \epsilon_t$$

where ϵ_t is white noise series with the additional property that $E_{t-1}[\epsilon_t] = 0$.

Special cases of ARMA processes: Moving Average

- ARMA family compromises two sub-classes

Definition (Moving Average Process of Order Q)

A Moving Average process of order Q, abbreviated MA(Q), has dynamics which follow

$$y_t = \phi_0 + \sum_{q=1}^{Q} \theta_q \epsilon_{t-q} + \epsilon_t$$

where ϵ_t is white noise series with the additional property that $E_{t-1}[\epsilon_t] = 0$.

- 1st order Moving Average (MA(1))

$$y_t = \phi_0 + \theta_1 \epsilon_{t-1} + \epsilon_t$$

- Simplest non-degenerate time series process

Special cases of ARMA processes: Autoregression

- Other sub-class of ARMA

Definition (Autoregressive Process of Order P)

An Autoregressive process of order P, abbreviated AR(P), has dynamics which follow

$$y_t = \phi_0 + \sum_{p=1}^{P} \phi_p y_{t-p} + \epsilon_t$$

where ϵ_t is white noise series with the additional property that $E_{t-1}[\epsilon_t] = 0$.

Special cases of ARMA processes: Autoregression

- Other sub-class of ARMA

Definition (Autoregressive Process of Order P)

An Autoregressive process of order P, abbreviated AR(P), has dynamics which follow

$$y_t = \phi_0 + \sum_{p=1}^{P} \phi_p y_{t-p} + \epsilon_t$$

where ϵ_t is white noise series with the additional property that $E_{t-1}[\epsilon_t] = 0$.

- 1st order Autoregression (AR(1))

$$y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

Moments and Autocovariances

$$y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

- Three key properties
  - *Unconditional* Mean: $\mathrm{E}\left[y_t\right]$
  - *Unconditional* Variance: $\gamma_0 = \mathrm{V}\left[y_t\right]$
  - Autocovariance: $\gamma_s = \mathrm{E}\left[\left(y_t - \mathrm{E}\left[y_t\right]\right)\left(y_{t-s} - \mathrm{E}\left[y_{t-s}\right]\right)\right]$
  - *Conditional* Mean: $\mathrm{E}_t[y_{t+1}] = \mathrm{E}[y_{t+1}|\mathcal{F}_t]$
  - *Conditional* Variance: $\mathrm{V}_t[y_{t+1}] = \mathrm{E}_t[(y_{t+1} - \mathrm{E}_t[y_{t+1}])^2]$

How to work with ARMA processes: AR(1)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

- Use backward substitution (assume $|\phi_1| < 1$)

$$\begin{aligned}
y_t &= \phi_0 + \phi_1 y_{t-1} + \epsilon_t \\
&= \phi_0 + \phi_1(\phi_0 + \phi_1 y_{t-2} + \epsilon_{t-1}) + \epsilon \\
&= \phi_0 + \phi_1\phi_0 + \phi_1^2 y_{t-2} + \phi_1\epsilon_{t-1} + \epsilon \\
&= \phi_0 + \phi_1\phi_0 + \phi_1^2(\phi_0 + \phi_1 y_{t-3} + \epsilon_{t-2}) + \phi_1\epsilon_{t-1} + \epsilon \\
&= \phi_0 \sum_{i=0}^{\infty} \phi_1^i + \sum_{i=0}^{\infty} \phi_1^i \epsilon_{t-i} \\
&= \frac{\phi_0}{1 - \phi_1} + \sum_{i=0}^{\infty} \phi_1^i \epsilon_{t-i}
\end{aligned}$$

- $\lim_{s\to\infty} \sum_{i=0}^{s} \phi_1^i = 1/(1 - \phi_1)$
- AR(1) is actually an MA(∞)

Properties of an AR(1)

$$
\begin{aligned}
\mathrm{E}[y_t] &= \mathrm{E}\left[\frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i \epsilon_{t-i}\right] \\
&= \frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i \mathrm{E}\left[\epsilon_{t-i}\right] \\
&= \frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i 0 \\
&= \frac{\phi_0}{1-\phi_1}
\end{aligned}
$$

Properties of an AR(1)

$$
\begin{aligned}
\mathrm{E}[y_t] &= \mathrm{E}\left[\frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i \epsilon_{t-i}\right] \\
&= \frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i \mathrm{E}\left[\epsilon_{t-i}\right] \\
&= \frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i 0 \\
&= \frac{\phi_0}{1-\phi_1}
\end{aligned}
$$

- In general AR(P): $\mathrm{E}[y_t] = \frac{\phi_0}{1-\phi_1-\phi_2-\ldots-\phi_P}$
- Only sensible if $\phi_1 + \phi_2 + \ldots + \phi_P < 1$

Properties of an AR(1)

$$
\begin{aligned}
\mathrm{E}[y_t] &= \mathrm{E}\left[\frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i \epsilon_{t-i}\right] \\
&= \frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i \mathrm{E}\left[\epsilon_{t-i}\right] \\
&= \frac{\phi_0}{1-\phi_1} + \sum_{i=0}^{t} \phi_1^i 0 \\
&= \frac{\phi_0}{1-\phi_1}
\end{aligned}
$$

- In general AR(P): $\mathrm{E}[y_t] = \frac{\phi_0}{1-\phi_1-\phi_2-\ldots-\phi_P}$
- Only sensible if $\phi_1 + \phi_2 + \ldots + \phi_P < 1$
- Variance can be shown in same manner
  - AR(1): $\mathrm{V}[y_t] = \frac{\sigma^2}{1-\phi_1^2}$
  - AR(P): $\mathrm{V}[y_t] = \frac{\sigma^2}{1-\rho_1\phi_1-\rho_2\phi_2-\ldots-\rho_P\phi_P}$
    - ρs are autocorrelations

Autocovariance of an AR(1)

$$
\begin{aligned}
\mathrm{E}\left[(y_t - \mathrm{E}[y_t])(y_{t-s} - \mathrm{E}[y_{t-s}])\right] &= \mathrm{E}\left[\left(\sum_{i=0}^{\infty}\phi_1^i \epsilon_{t-i}\right)\left(\sum_{i=0}^{\infty}\phi_1^i \epsilon_{t-s-i}\right)\right] \\
&= \phi_1^s \frac{\sigma^2}{1-\phi_1^2}
\end{aligned}
$$

- Full details in notes

$$
\gamma_s = \phi_1^s \frac{\sigma^2}{1-\phi_1^2}
$$

- Autocovariance declines geometrically with the lag length
- Requires $\phi_1^2 < 1$ to exist
  - Same condition as the mean

Stationarity of ARMA processes

- Primarily interested in covariance stationarity
- Stationarity depends on parameters of *AR* portion
- AR(0) or finite order MA: always stationary

Stationarity of ARMA processes

- Primarily interested in covariance stationarity
- Stationarity depends on parameters of *AR* portion
- AR(0) or finite order MA: always stationary
- AR(1) or ARMA(1,Q): $y_t = \phi_1 y_{t-1} + \text{MA} + \epsilon_t$
  - $|\phi_1| < 1$

Stationarity of ARMA processes

- Primarily interested in covariance stationarity
- Stationarity depends on parameters of *AR* portion
- AR(0) or finite order MA: always stationary
- AR(1) or ARMA(1,Q): $y_t = \phi_1 y_{t-1} + \text{MA} + \epsilon_t$
  - $|\phi_1| < 1$
- AR(P) or ARMA(P,Q) $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_P y_{t-P} + \text{MA} + \epsilon_t$
- Rewrite $y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \ldots - \phi_P y_{t-P} = \text{MA} + \epsilon_t$
- Each to determine using the characteristic equation and corresponding characteristic roots

The characteristic equation

Definition (Characteristic Equation)

Let y_t follow a P$^{\text{th}}$ order linear difference equation

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_P y_{t-P} + x_t$$

which can be rewritten as

$$\begin{aligned} y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \ldots - \phi_P y_{t-P} &= \phi_0 + x_t \\ (1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_P L^P) y_t &= \phi_0 + x_t \end{aligned}$$

The characteristic equation of this process is

$$z^P - \phi_1 z^{P-1} - \phi_2 z^{P-2} - \ldots - \phi_{P-1} z - \phi_P = 0$$

- Key is in the forming of the characteristic equation and its roots
- L is known as "lag operator"

Characteristic roots

Definition (Characteristic Root)

Let

$$z^P - \phi_1 z^{P-1} - \phi_2 z^{P-2} - \ldots - \phi_{P-1} z - \phi_P = 0$$

be the characteristic polynomial associated with some P^{th} order linear difference equation. The P characteristic roots, $c_1, c_2, \ldots, c_P$ are defined as the solution to this polynomial

$$(z - c_1)(z - c_2) \ldots (z - c_P) = 0$$

- The roots are c_1, c_2, ..., c_P
- AR(P) or ARMA(P,Q) is covariance stationary if $|c_j| < 1$ for all j
- If complex, $|c_j| = |a_j + b_j i| = \sqrt{a^2 + b^2}$ (complex modulus)

Characteristic roots example

- Difficult to determine by inspection
- Example 1:

$$y_t = .1y_{t-1} + .7y_{t-2} + .2y_{t-3} + \epsilon_t$$

Characteristic roots example

- Difficult to determine by inspection
- Example 1:

$$y_t = .1y_{t-1} + .7y_{t-2} + .2y_{t-3} + \epsilon_t$$

  - Characteristic equation

$$z^3 - .1z^2 - .7z^1 - .2$$

  - Roots: 1,-.5, -.4, so nonstationary
- Example 2:

$$y_t = 1.7y_{t-1} - .72y_{t-2} + \epsilon_t$$

Characteristic roots example

- Difficult to determine by inspection
- Example 1:

$$y_t = .1y_{t-1} + .7y_{t-2} + .2y_{t-3} + \epsilon_t$$

  - Characteristic equation

$$z^3 - .1z^2 - .7z^1 - .2$$

  - Roots: 1,-.5, -.4, so nonstationary
- Example 2:

$$y_t = 1.7y_{t-1} - .72y_{t-2} + \epsilon_t$$

  - Characteristic equation

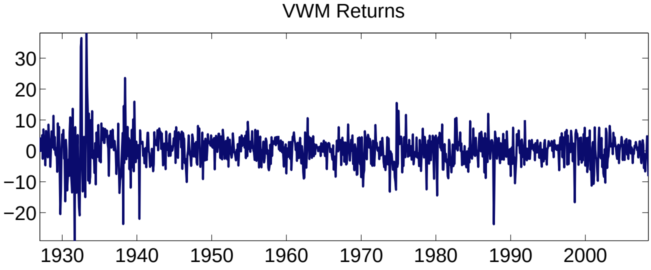
$$z^2 - 1.7z^1 + .72$$

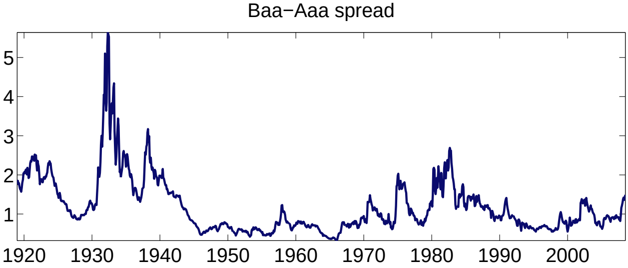
  - Roots: .9, .8, so stationary

Data

- VWM
  - Identical to before
  - Monthly
  - January 1927 though July 2008
  - CRSP
- Moody's Baa-Aaa corporate bond spread
  - Called the default spread
  - Monthly
  - January 1919 though September 2008
  - FRED: St. Louis FED

Plots of the VWM and the Default Spread





Autocorrelations and the ACF

- Autocorrelations are a key element of model building

Definition (Autocorrelation)

The autocorrelation of a scalar process is defined

$$\rho_s = \frac{\gamma_s}{\gamma_0}$$

where $\gamma_s = \mathrm{E}\left[(y_t - \mu)(y_{t-s} - \mu)\right]$

- Measures the correlation of a process at different points in time
- AR(1):

$$\rho_s = \phi_1^s$$

- One of two possibilities
  - Decay geometrically if $0 < \phi_1 < 1$
  - Oscillate and decay $-1 < \phi_1 < 0$

Partial Autocorrelations (PACF)

- Partial Autocorrelation is the other key element of model building
- More complicated than autocorrelations:
- Regression interpretation of sth partial autocorrelation:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_{s-1} y_{t-s+1} + \varphi_s y_{t-s} + \epsilon_t$$

- φ_s is the sth partial autocorrelation
  - Population (not sample) value of φ_s
- AR(1):

$$\varphi_s = \begin{cases} \phi_1^s \text{ for s=} -1, 0, 1 \\ 0 \text{ otherwise} \end{cases}$$

- Partial autocorrelation function maps the parameters of a process to the sth autocorrelation, $\varphi(s)$

Using the ACF and PACF to categorize processes

- ACF and PACF are useful when choosing models

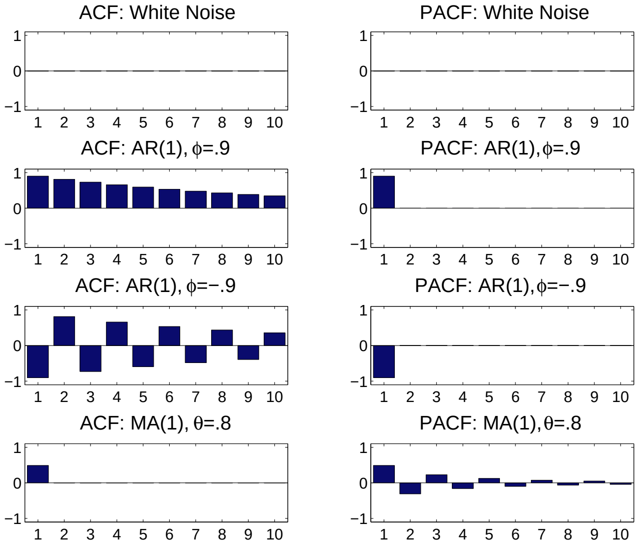
Process	ACF	PACF
White Noise	All 0	All 0
AR(1)	$\rho_s = \phi_1^s$	0 beyond lag 2
AR(P)	Decays toward zero exponentially	Non-zero through lag P, 0 thereafter
MA(1)	$\rho_1 \neq 0$, $\rho_s = 0$, $s > 0$	Decays toward zero exponentially
MA(Q)	$\rho_s \neq 0$ $s \leq Q$, $\rho_s = 0$, $s > Q$	Decays toward zero exponentially, possible oscillating
ARMA(P,Q)	Exponential Decay	Exponential Decay

Using the ACF and PACF to categorize processes

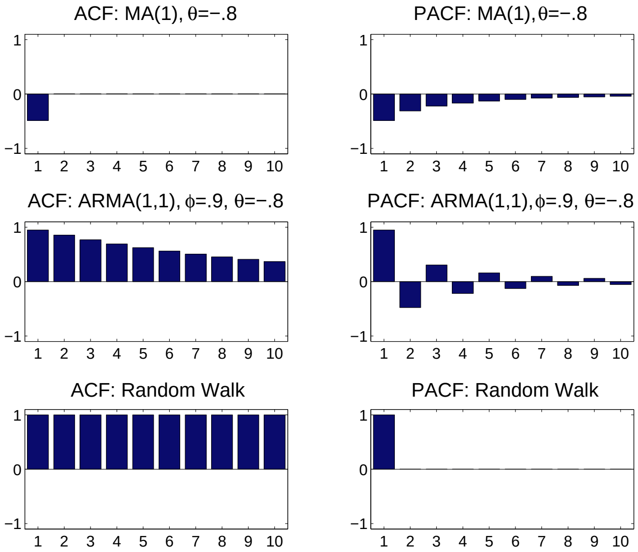
- ACF and PACF are useful when choosing models

Process	ACF	PACF
White Noise	All 0	All 0
AR(1)	$\rho_s = \phi_1^s$	0 beyond lag 2
AR(P)	Decays toward zero exponentially	Non-zero through lag P, 0 thereafter
MA(1)	$\rho_1 \neq 0$, $\rho_s = 0$, $s > 0$	Decays toward zero exponentially
MA(Q)	$\rho_s \neq 0$ $s \leq Q$, $\rho_s = 0$, $s > Q$	Decays toward zero exponentially, possible oscillating
ARMA(P,Q)	Exponential Decay	Exponential Decay

Autocorrelation for ARMA processes



Autocorrelation for ARMA processes



Sample ACF and PACF

- Sample autocorrelations

$$\hat{\rho}_s = \frac{\sum_{t=s+1}^{T} y_t^* y_{t-s}^*}{\sum_{t=1}^{T} y_t^{*2}} = \frac{\hat{\gamma}_s}{\hat{\gamma}_0}$$

- $y_t^* = y_t - \bar{y}$ where $\bar{y} = T^{-1} \sum_{t=1}^{T} y_t$
- Some prefer the small-sample-size corrected version

$$\hat{\rho}_s = \frac{\sum_{t=s+1}^{T} y_t^* y_{t-s}^*}{\sqrt{\sum_{t=s+1}^{T} y_t^{*2} \sum_{t=1}^{T-s} y_t^{*2}}}.$$

Sample ACF and PACF

- Sample autocorrelations

$$\hat{\rho}_s = \frac{\sum_{t=s+1}^{T} y_t^* y_{t-s}^*}{\sum_{t=1}^{T} y_t^{*2}} = \frac{\hat{\gamma}_s}{\hat{\gamma}_0}$$

- $y_t^* = y_t - \bar{y}$ where $\bar{y} = T^{-1} \sum_{t=1}^{T} y_t$
- Some prefer the small-sample-size corrected version

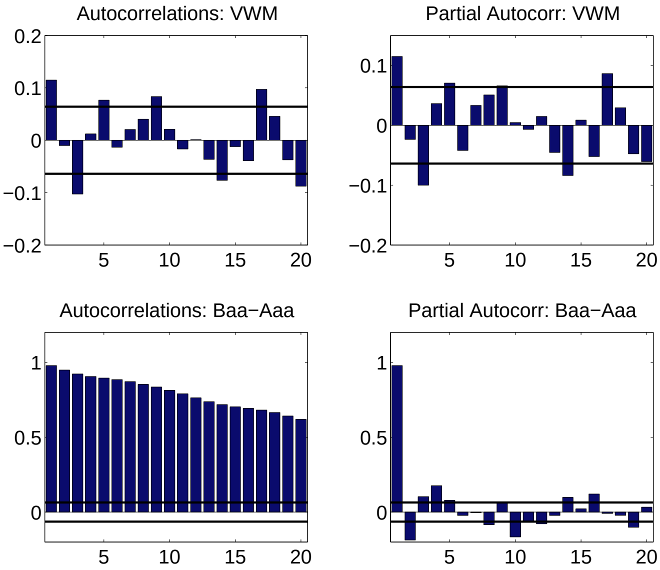
$$\hat{\rho}_s = \frac{\sum_{t=s+1}^{T} y_t^* y_{t-s}^*}{\sqrt{\sum_{t=s+1}^{T} y_t^{*2} \sum_{t=1}^{T-s} y_t^{*2}}}.$$

- Sample partial autocorrelations
  - Run regression to estimate $\hat{\varphi}_s$

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \varphi_s y_{t-s} + \epsilon_t$$

- More efficient ways to compute PACF using Yule-Walker equations (see notes)

Autocorrelation for the VWM and Default Spread



Testing autocorrelations and partial ACs

- Inference on autocorrelations:

$$
\begin{aligned}
\mathrm{V}[\hat{\rho}_s] &= T^{-1} && \text{for } s = 1 \\
&= T^{-1}\left(1 + 2\sum_{j=1}^{s-1} \hat{\rho}_j^2\right) && \text{for } s > 1
\end{aligned}
$$

- Standard t-stats

$$\frac{\hat{\rho}_s}{\sqrt{\mathrm{V}[\hat{\rho}_s]}} \overset{A}{\sim} N(0, 1).$$

Testing autocorrelations and partial ACs

- Inference on autocorrelations:

$$\begin{aligned} \mathrm{V}[\hat{\rho}_s] &= T^{-1} && \text{for } s = 1 \\ &= T^{-1}(1 + 2\sum_{j=1}^{s-1} \hat{\rho}_j^2) && \text{for } s > 1 \end{aligned}$$

- Standard t-stats

$$\frac{\hat{\rho}_s}{\sqrt{\mathrm{V}[\hat{\rho}_s]}} \overset{A}{\sim} N(0, 1).$$

- Inference on partial autocorrelations:

$$\mathrm{V}[\hat{\varphi}_s] \approx T^{-1}$$

- Standard t-stats

$$T^{\frac{1}{2}} \hat{\varphi}_s \overset{A}{\sim} N(0, 1)$$

Testing multiple autocorrelations

- Testing multiple autocorrelations: Ljung-Box Q

$$Q = T(T+2)\sum_{k=1}^{s}\frac{\hat{\rho}_k^2}{T-k} \sim \chi_s^2$$

- Note: Not heteroskedasticity robust, use LM test for serial correlations

Definition (LM test for serial correlation)

Under the null, $\mathrm{E}[y_t^* y_{t-j}^*] = 0$ for $1 \leq j \leq s$. The LM-test for serial correlation is constructed by defining the score vector $\mathbf{s}_t = y_t^* \left[y_{t-1}^* \, y_{t-2}^* \, \ldots y_{t-s}^*\right]'$,

$$LM = T\bar{\mathbf{s}}'\hat{\mathbf{S}}\bar{\mathbf{s}} \xrightarrow{d} \chi_s^2$$

where $\bar{\mathbf{s}} = T^{-1}\sum_{t=1}^{T}\mathbf{s}_t$ and $\hat{\mathbf{S}} = T^{-1}\sum_{t=1}^{T}\mathbf{s}_t\mathbf{s}_t'$.

Model building the Box-Jenkins way

- Model building is similar to cross-section regression
- Can use same techniques
  - General to Specific or Specific to General
  - Information criteria: AIC, BIC
- Box-Jenkins is dominant methodology, 2-steps
  - Identification: Use ACF and PACF to choose model
  - Estimation: Estimate model and do diagnostic checks
- Two principles
  - Parsimony
  - Invertibility

Strategies

- General to Specific
  - Fit largest specification
  - Drop regressor with largest p-val
  - Refit
  - Stop if all p-vals indicate significance using a size of α
    - α is the econometrician's choice
- Specific to General
  - Fit all specifications with a single variable
  - Retail variable with smallest p-val
  - Extend this model adding on additional variables one at a time
  - Stop if the p-vals of all excluded variables are larger than α

Information Criteria

- Information Criteria
  - Akaike Information Criterion (AIC)

$$AIC = \ln \hat{\sigma}^2 + 2\frac{K}{T}$$

  - Schwartz (Bayesian) Information Criterion (SIC/BIC)

$$BIC = \ln \hat{\sigma}^2 + K\frac{\ln T}{T}$$

- Both have versions suitable for likelihood based estimation
- Reward for better fit: Reduce $\ln \hat{\sigma}^2$
- Penalty for more parameters: $2\frac{K}{T}$ or $K\frac{\ln T}{T}$
- Choose model with smallest IC
  - AIC has fixed penalty $\Rightarrow$ inclusion of extraneous variables
  - BIC has larger penalty if $\ln T > 2$ ($T > 7$)