11th March 2003

NOTES ON PROGRAMS TRAMO AND SEATS©

PART I
Introduction and Brief Review of Applied Time Series Analysis

Agustín Maravall
Bank of Spain

Thanks are due to Victor Gómez, Gianluca Caporello, Fernando Sánchez, and Nieves Morales.

1. INTRODUCTION

In our application,
we center on series observed with a 1, 2, 3, 4, 6, and 12 times a year frequency. The most relevant ones:

MONTHLY and QUARTERLY time series

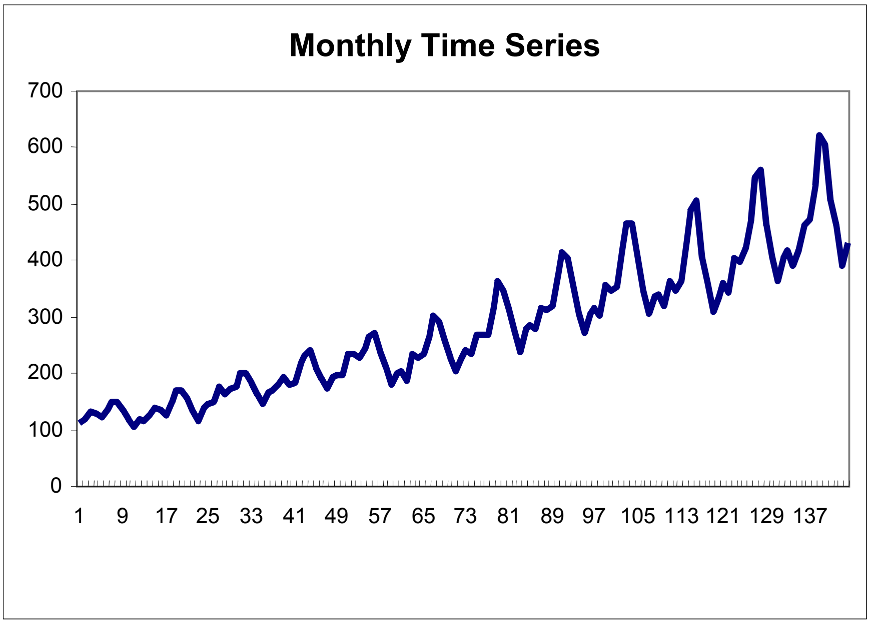
$$\text{TIME SERIES} \equiv [\, x_1, x_2, \ldots, x_T \,]$$

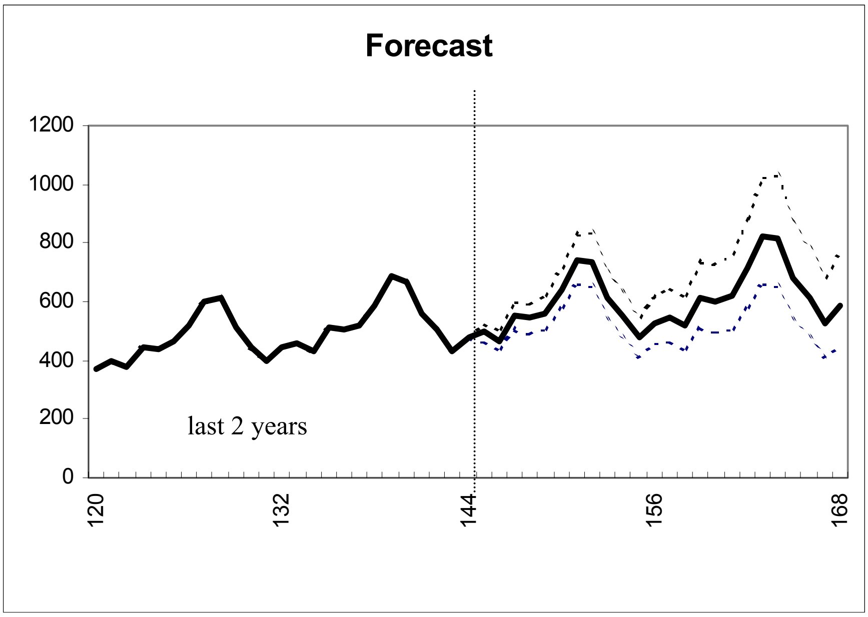
$$12 - 36 \leq T \leq 600 \text{ observations}$$

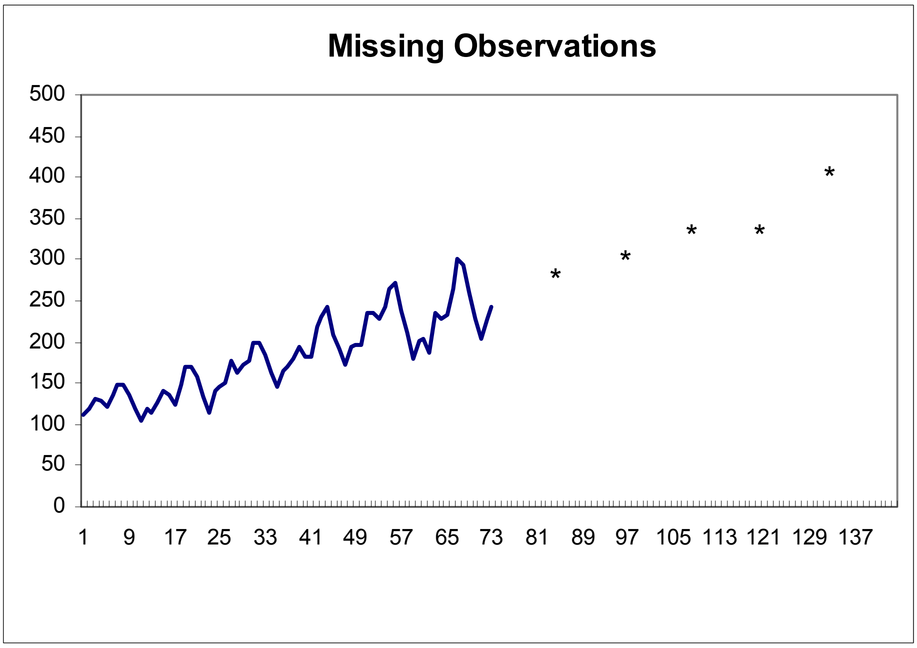
(minimum depends on the frequency of observations and on the type of analysis performed).

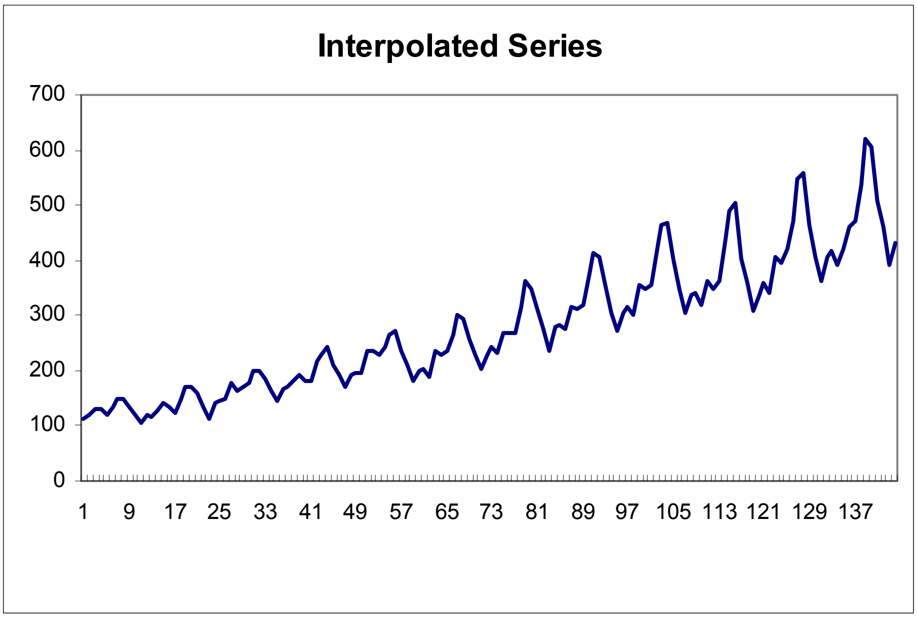
Our interest: SHORT-TERM ANALYSIS

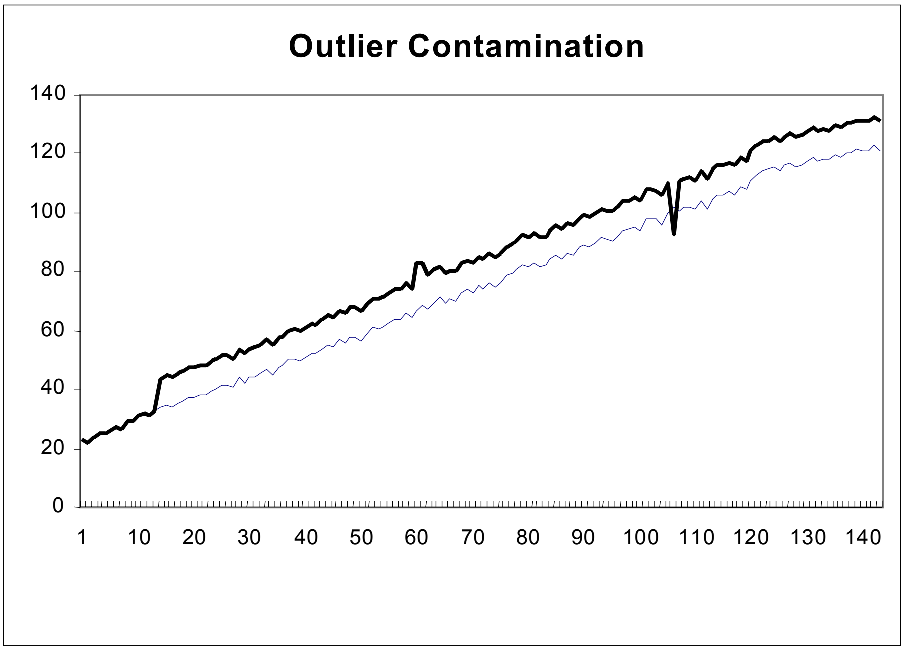
SOME EXAMPLES OF PROBLEMS THAT WE SHALL ADDRESS:

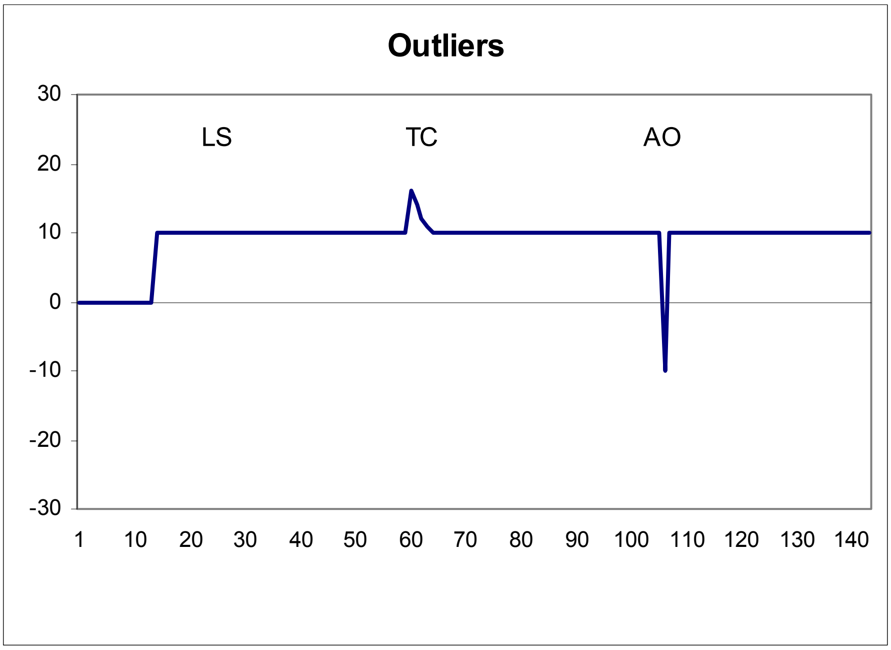

Monthly Time Series
700
600
500
400
300
200
100
0
1 9 17 25 33 41 49 57 65 73 81 89 97 105 113 121 129 137

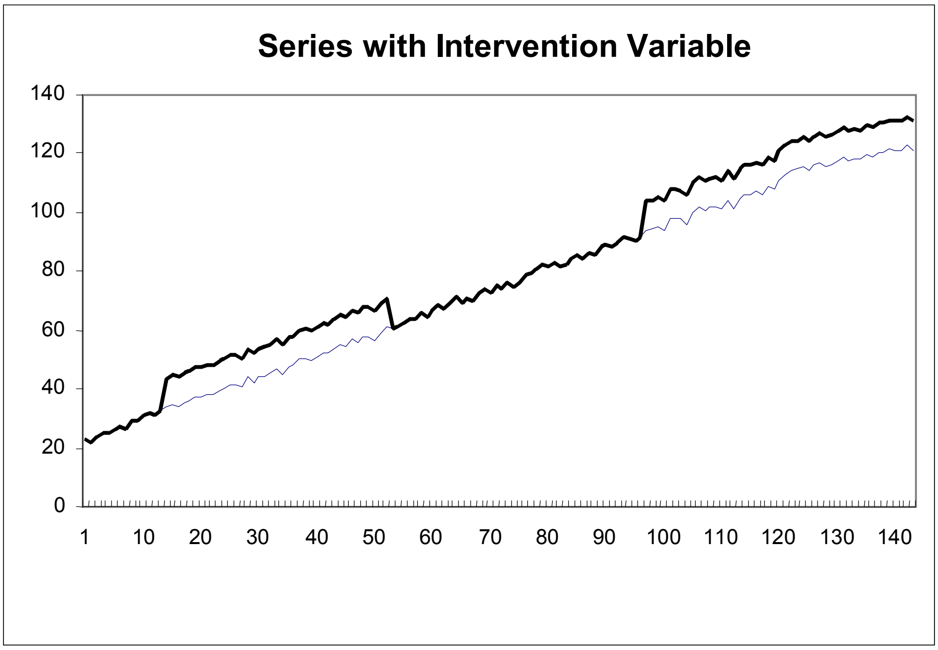

Forecast
1200
1000
800
600
400
200
0
last 2 years
120
132
144
156
168

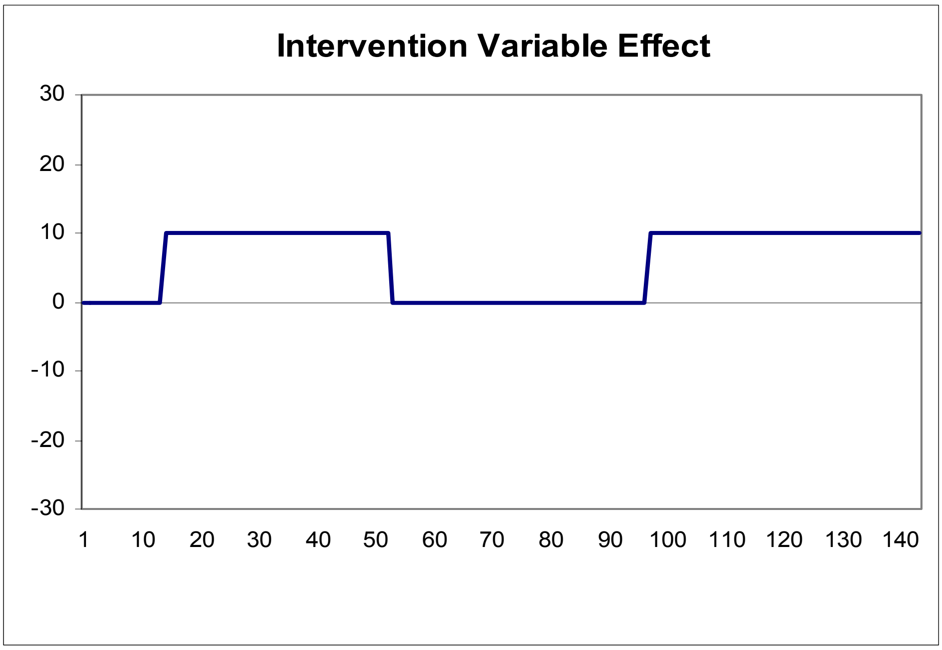

Missing Observations
500
450
400
350
300
250
200
150
100
50
0
1 9 17 25 33 41 49 57 65 73 81 89 97 105 113 121 129 137

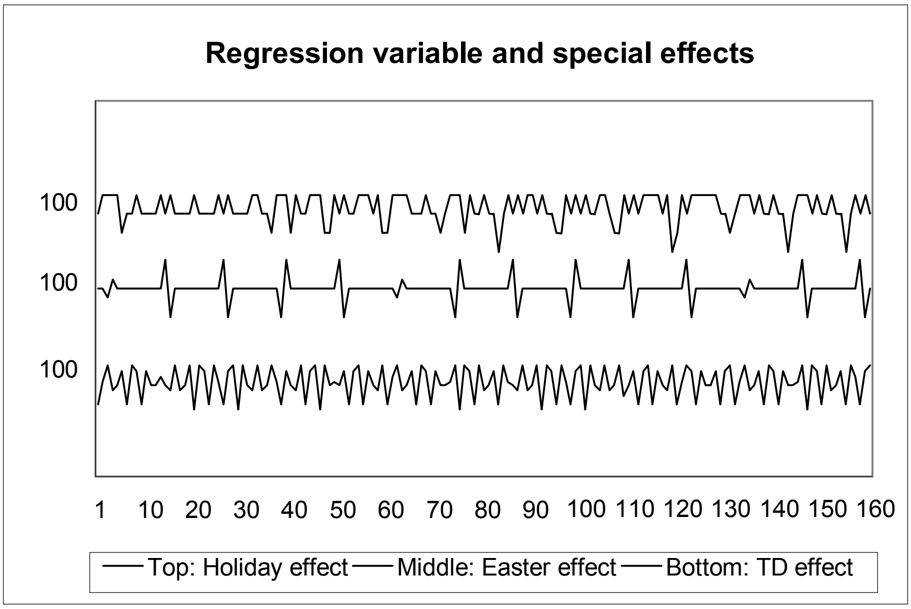

Interpolated Series
700
600
500
400
300
200
100
0
1 9 17 25 33 41 49 57 65 73 81 89 97 105 113 121 129 137

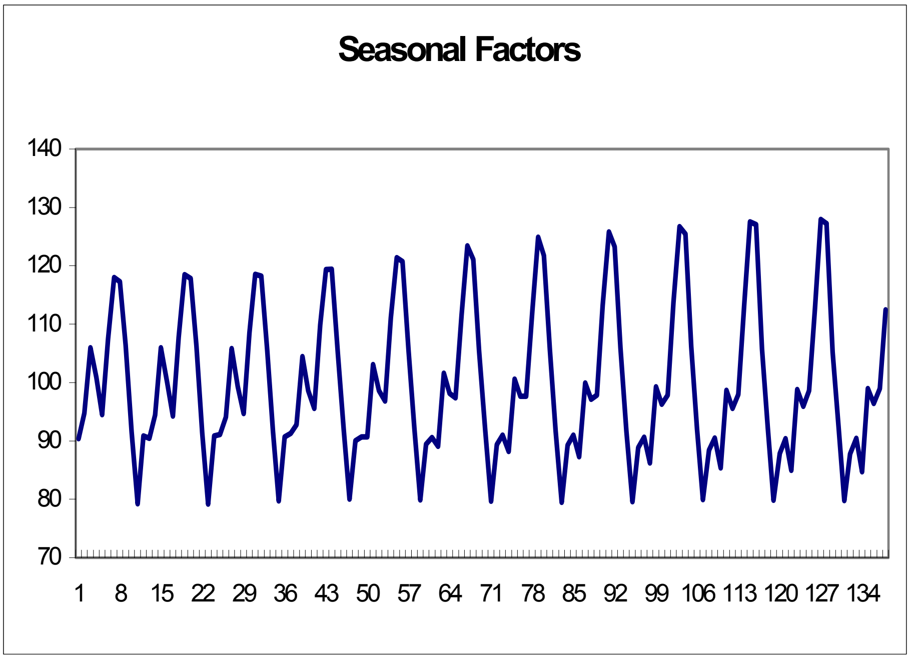

Outlier Contamination
140
120
100
80
60
40
20
0
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140

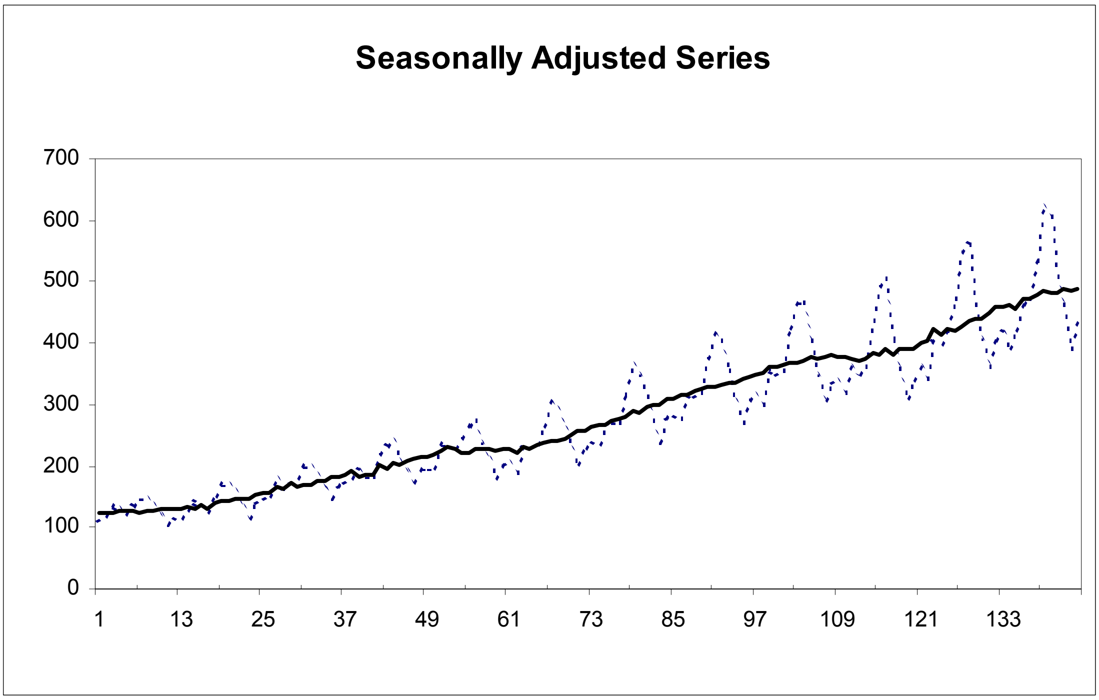

Outliers
30
20
10
0
-10
-20
-30
LS
TC
AO
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140


Series with Intervention Variable
140
120
100
80
60
40
20
0
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140

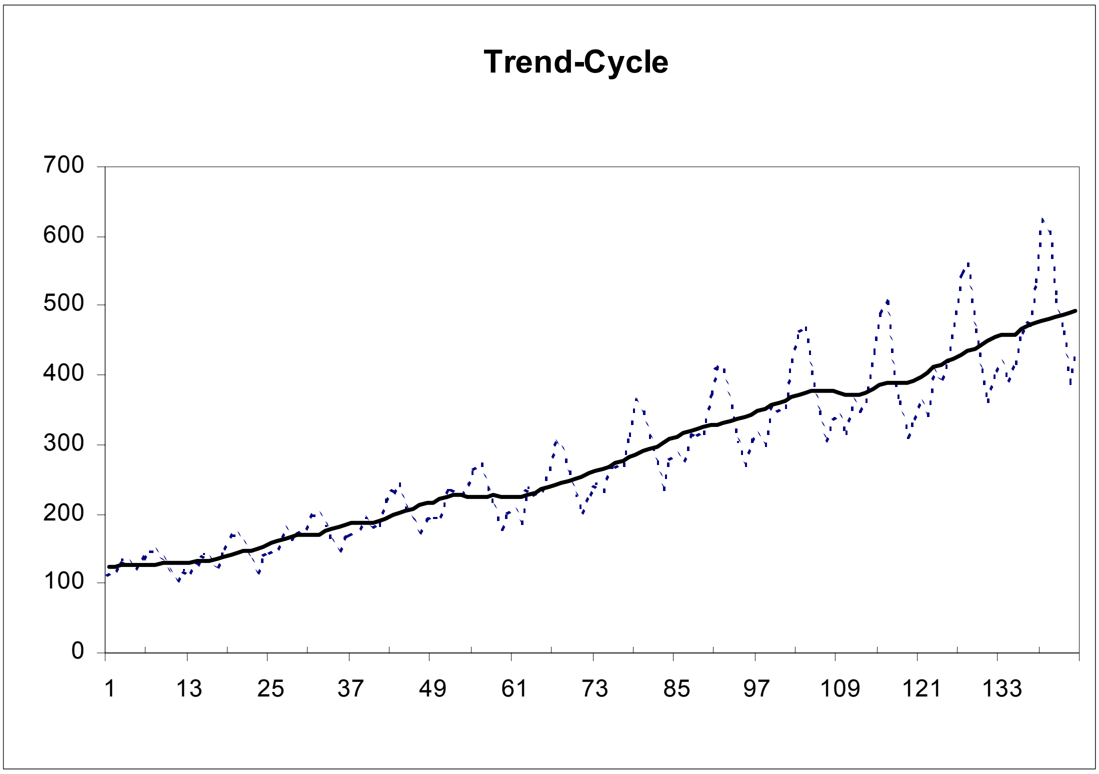

Intervention Variable Effect
30
20
10
0
-10
-20
-30
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140


Regression variable and special effects
100
100
100
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
Top: Holiday effect
Middle: Easter effect
Bottom: TD effect

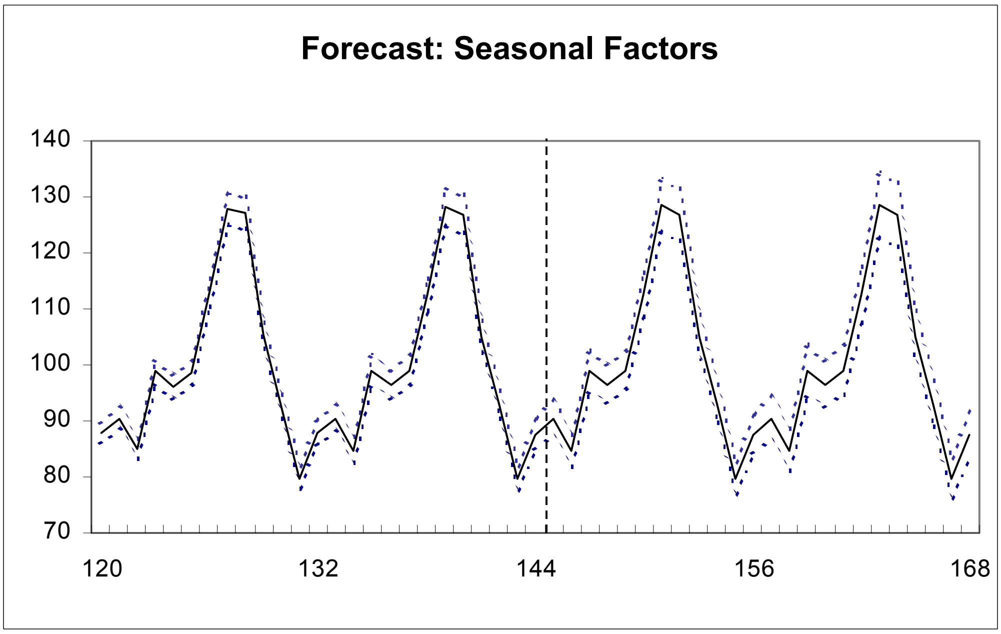

Seasonal Factors
140
130
120
110
100
90
80
70
1
8
15
22
29
36
43
50
57
64
71
78
85
92
99
106
113
120
127
134

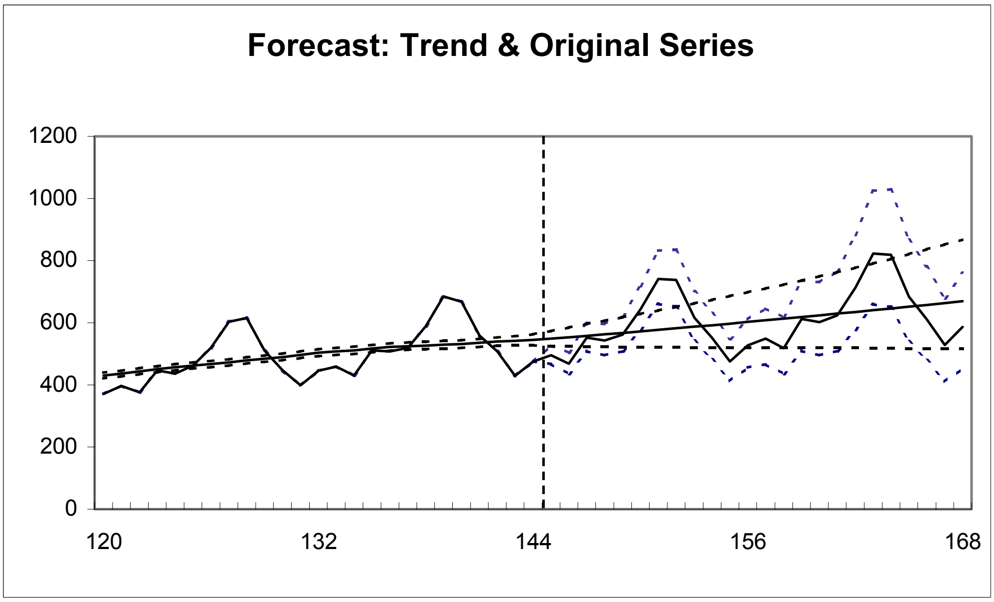


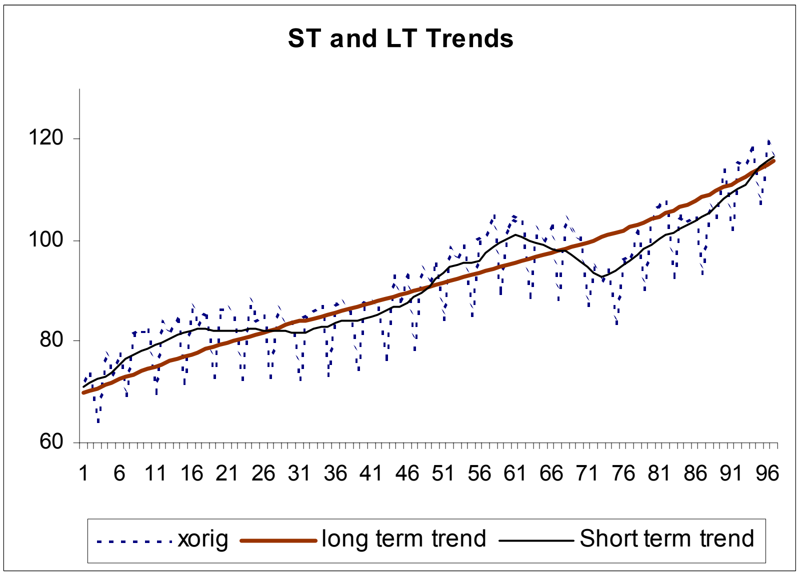
Seasonally Adjusted Series
700
600
500
400
300
200
100
0
1
13
25
37
49
61
73
85
97
109
121
133

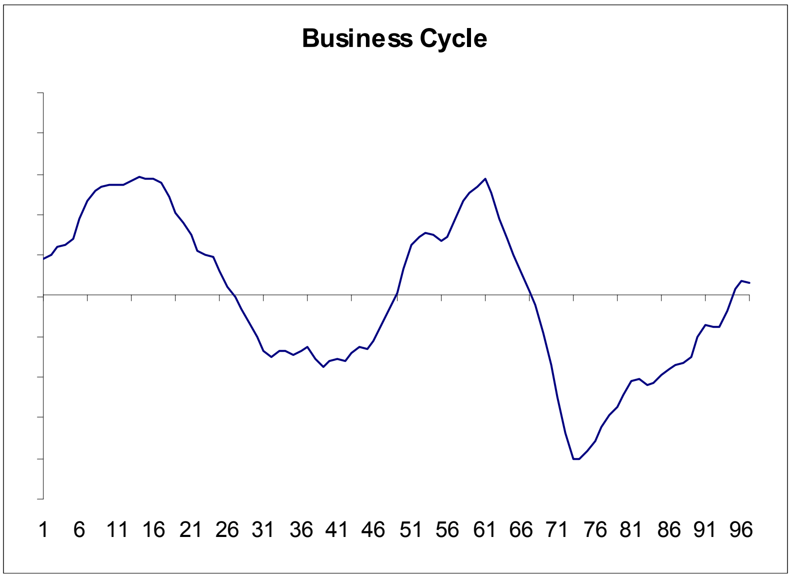


Trend-Cycle
700
600
500
400
300
200
100
0
1
13
25
37
49
61
73
85
97
109
121
133


Forecast: Seasonal Factors
140
130
120
110
100
90
80
70
120
132
144
156
168


Forecast: Trend & Original Series
1200
1000
800
600
400
200
0
120
132
144
156
168


ST and LT Trends
120
100
80
60
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71 76 81 86 91 96
xorig
long term trend
Short term trend


Business Cycle
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71 76 81 86 91 96

Standard ("routine") treatment at present typically solves the previous problems using different procedures, that often have little to do with each other. For ex.:

* Forecasting: ARIMA, EWMA

* Interpolation: Chow-Lin, Denton

* Seasonal adjustment: X11 , X11A

* Preadjustment for trading day, easter effect, holidays...: Regression / Prior correction (For ex., divide by # of working days)…

* Trend extraction: Henderson Moving Averages, HP filter…

* Outliers: Some weighted trimming?
  Robust procedures instead?

* Forecast of a trend: ?
  Often: Fit ARIMA to some trend, and obtain ARIMA forecasts (not recommended)

* SE of $\hat{x}_t^a$? (x_t^a : SA series)
  Important issue (Bach Commitee, Moore Committee,…)

and so on...

We shall present a methodology that

* permits to deal with all those issues jointly, within a unified framework.

* This framework provides OPTIMAL ESTIMATORS (OR FORECASTS) with respect to

  - well-defined STATISTICAL MODELS,

  - well-defined ESTIMATION CRITERION,

  in an EFFICIENT way.

The Model-based approach will facilitate

* <u>interpretation</u>

  For example, the model may specify that the sum of the seasonal component over a 12 consecutive-month period is a zero-mean, small variance, stationary process.
* <u>diagnostics</u>

  The joint distribution of the estimators can be derived, and hence standard tests can be performed.

* <u>inference</u>

  For example, we can obtain optimal forecasts of the rate of growth of the SA series, with the associated SE.

The methodology is based on:

1) Identifying REG-ARIMA models for the observed series.

2) Decomposing the previous model for the series into unobserved components.

3) Obtain the MMSE estimator of the components (or signals). These estimators will be:

E (signal | observations)

Before explaining the methodology, it will be helpful to start with a brief (and informal) review of some applied time series analysis concepts and tools.

2. BRIEF REVIEW OF APPLIED TIME SERIES ANALYSIS

General Framework:

Stochastic process:

$z_t \sim f_t (z_t)$

Time Series: $[z_1, z_2, \ldots , z_T]$

We consider it as a particular (partial) realization of a stochastic process.

Hence, a sample of size 1 for each f_t.

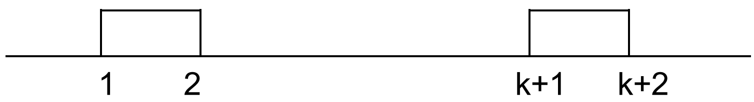
Need to add more structure.

STATIONARITY AND DIFFERENCING

Strong condition. Although few economic series will satisfy it, simple transformations will render them stationary.

Basic condition:

$$f(z_1 \dots, z_T) = f(z_{1+k}, \dots, z_{T+k})$$



In particular, for marginal distribution (T=1)

$f_t(z_t) = f(z_t)$ for every t, hence

$$\left.\begin{aligned} Ez_t &= \mu \\ Vz_t &= V_z \end{aligned}\right\}$$

both:

- are finite

- do not depend on t

In practice, <u>constant variance</u> is achieved through:

- log-level transformation

  +

- outlier correction

Alternatively, one may use

NONLINEAR FILTERS

(ex. : GARCH, Bilinear, Stochastic Volatility models, …)

- Not yet fit for large-scale use.
- For many of these models, point forecasts or point estimator of the series and components obtained with linear model remains approximately optimal.
- The decomposition of NL models still poses some problems.
- Monthly and lower-frequency data seldom display markedly nonlinear structures.
  We shall not consider them.

Roughly:

LOGS are appropriate when the amplitude of the series oscillations is approximately proportional to the level.

Note: LOG transformation has some nice features.

- "Scale" free

- Natural interpretation: $(d \log x = dx / x \cong \nabla x / x)$

variations are expressed as fractions ("per one") of

the level of the series.

Thus, for example,

ARIMA fit to the logs $\rightarrow \sigma_a$ (residuals) = .004

We can say: The serie is forecasted (1 p.a.) with an error equivalent to roughly 4‰ of the level of the series.

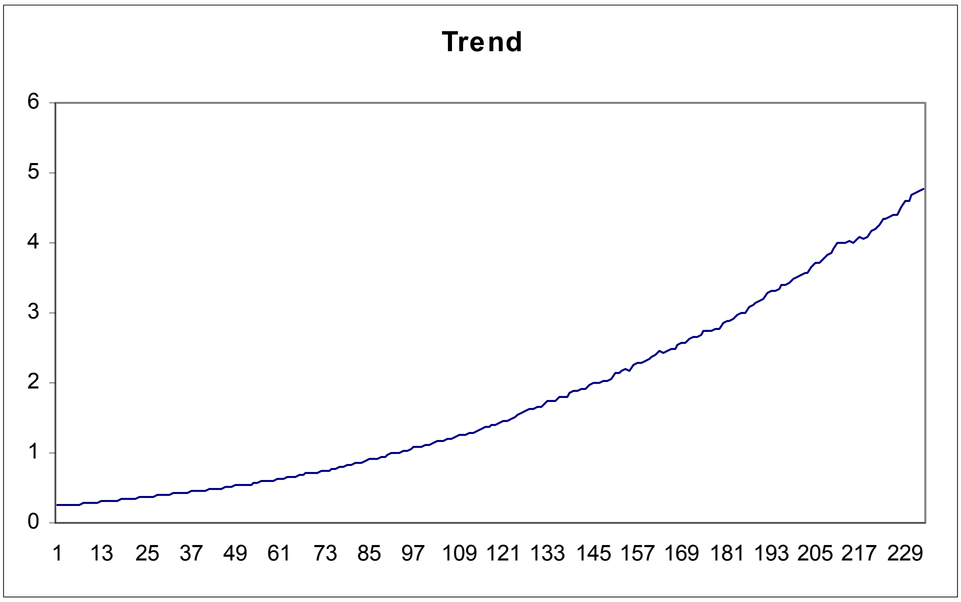
On the negative side, it may induce BIASES (due to the fact that geometric means underestimate aritmetic means).

$\Rightarrow$ Annual mean of original series > mean of SA series (and of trend)

(If biases are large, there are 2 options:
- ad-hoc corrections;
- model the levels.)

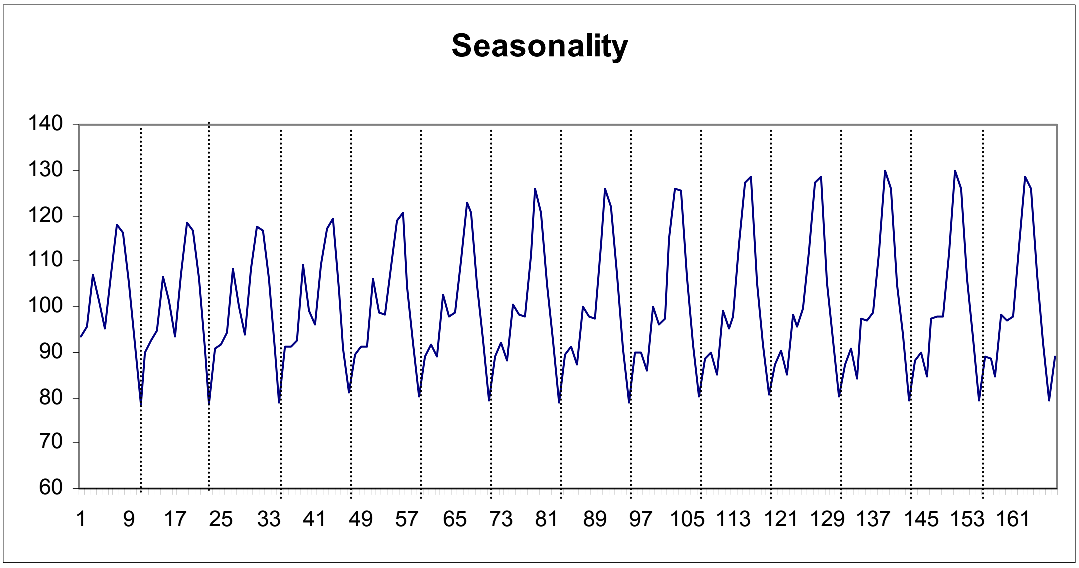
Concerning STATIONARITY IN MEAN, most economic time series display a mean (i.e., a “local level”) that cannot be assumed constant. The two most important reasons:

a) The presence of a trend (or a trend-cycle)



Obviously, the mean of the series in the first years is not the same as the one for the last years.

b) The presence of seasonality:



Obviously, the level of the series depends on the period within the year.

To achieve <u>Constant Mean</u>:

Let:

B: "Backward" operator; $B^j z_t = z_{t-j}$

s = # obs./year (12, if monthly; 4 if quarterly;…)

We use operators:

$\nabla = 1 - B$ (regular difference)

$\nabla_s = 1 - B^s$ (seasonal difference)

$S_s = 1 + B + \ldots B^{s-1}$ (sum over a year)

An important identity : $\boxed{\nabla_s = \nabla S_s}$

* Assume x_t is a linear trend

$$x_t = a + b\ t;$$

then,

$$\nabla x_t = b$$

or

$$\nabla^2 x_t = 0 .$$

In general:

∇^d reduces polynomial of degree d to a constant

(cancels polynomial of degree d-1)

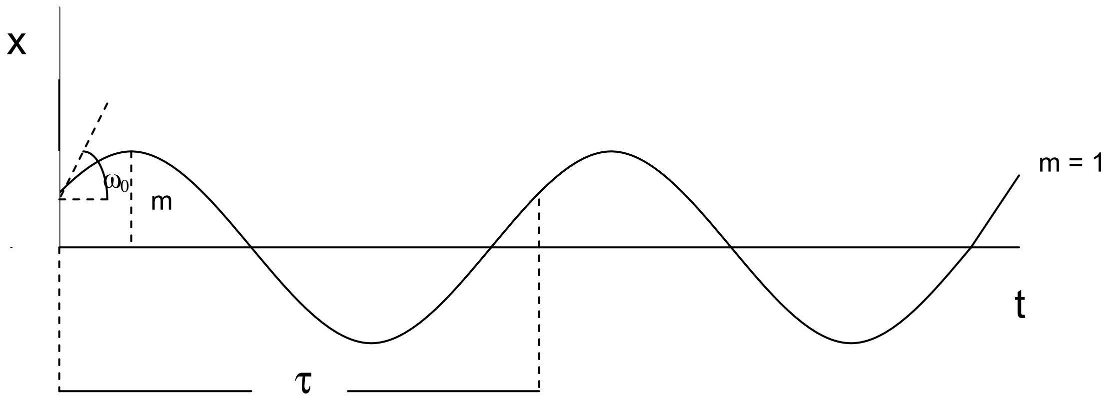
Also: $\nabla_{12} x_t = 12b$

(hence seasonal differencing also affects trend)

Note: **Cosine function**

Basic element in cyclical and seasonal movements: the (deterministic) function

$$x_t = m^t \cos(\omega t + \omega_0) \quad , \qquad t = 0, 1, \ldots \qquad \text{(A)}$$



m = modulus (max. value) ;

ω = Frequency (# of cycles / unit of time), measured in radians $= \frac{2\pi}{\tau}$;

τ = Period (# of units of time needed to complete a full cycle) ;

ω_0 = Phase (angle at $t = 0$)

* Using expression for cos (a + b), (A) above can also be rewritten as

$$x_t = m^t \, (C \cos \omega t + D \sin \omega t) \; .$$

* Recall:

 (A) is solution of 2nd order difference equation (with real coefficients)

$$x_t + \phi_1 \, x_{t-1} + \phi_2 \, x_{t-2} = 0 \; ,$$

when roots are complex

$\Rightarrow$ always, pairs of complex conjugates.

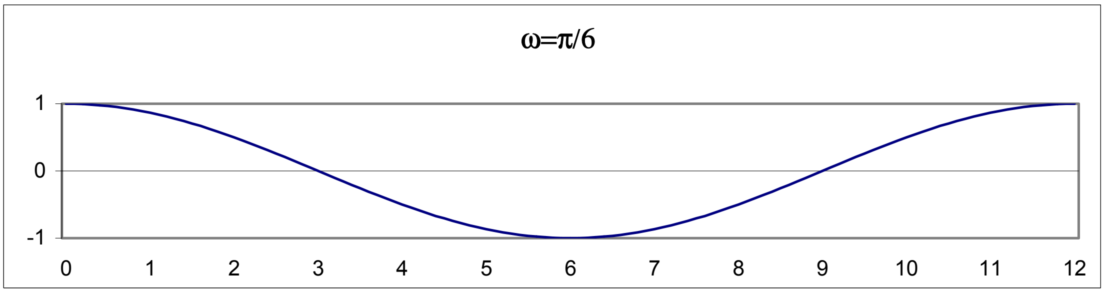
(see below)

Consider a monthly series

* Assume x_t is a sine wave, for ex.

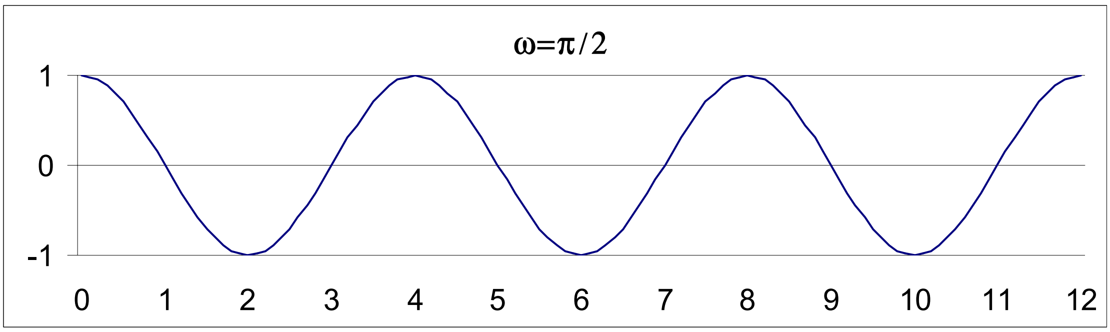
$$x_t = \cos\left(\frac{\pi}{6}\,t\right), \quad \text{then } \nabla_{12}\, x_t = 0$$

given that $$\cos\frac{\pi}{6}(t-12) = \cos\left(\frac{\pi}{6}t - 2\pi\right) = \cos\left(\frac{\pi}{6}t\right)$$



Same thing holds when

$$x_t = \cos\left(\frac{\pi}{2}\,t\right)$$



So: what is the complete solution of

$$\nabla_{12}\, x_t = 0\,?$$

i.e., most general *F*(t) that is cancelled by ∇_{12}

$$\nabla_{12}\, x_t = x_t - x_{t-12} = 0$$

Linear difference equation (homogeneous).

<u>Note:</u> **Homogeneous Linear Difference Equations**

(with real coefficients).

Let equation be

$$x_t + \phi_1\, x_{t-1} + \ldots + \phi_p\, x_{t-p} = 0 \quad , \qquad t = 1, 2, \ldots$$

Replacing x_t by r^t, the **characteristic equation** is obtained, equal to

$$r^p + \phi_1\, r^{p-1} + \ldots + \phi_{p-1}\, r + \phi_p = 0 \quad .$$

This equation has

p roots — some real

p roots — some complex

(always in pairs of complex conjugates:

$r_1 = a + b\,i$

$r_2 = a - b\,i$)

Solution to the difference equation:

$$x_t = \sum_{j=1}^{p} c_j \, r_j^t \quad ,$$

where

r_j : a root of the charact. equat.

c_j: arbitrary constants.

The final way in which the roots are expressed is the following (the c's are always constants, to be determined from the starting conditions).

1) Single real root

$$x_t = c_j \, r_j^t \quad .$$

Notice : if $\left| r_j \right| > 1 \quad \Rightarrow \quad$ root is explosive.

We shall restrict attention to roots with

$$\left| r_j \right| \leq 1 \quad .$$

2) Multiple real roots

order of multiplicity : k + 1 ;

$$x_t = (c_0 + c_1 \, t + \ldots + c_k \, t^k) \; r^t \quad ,$$

where $r_1 = \ldots = r_{k+1} = r$.

3) Single complex root

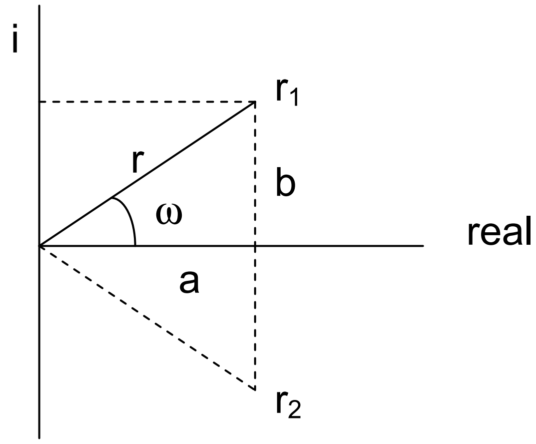
Always as a pair of complex conjugates.

Let

$$r_1 = a + b\,i$$

$$r_2 = a - b\,i$$

be the pair, and let



$$r = (a + b)^{1/2} \quad ,$$

$$\omega = a\cos\left(\frac{a}{r}\right) \quad ,$$

then

$$x_t = c_0\, r^t \cos(\omega\, t + c_1) \quad .$$

(c_0 , c_1: constants)

4) Multiple complex roots

Very rarely encountered.

x_t = A mixture of previous solutions.

Notice that, in all cases, if

$$|r| = 1 \quad ,$$

the solution has a systematic explosive behavior, which is not found in actual economic series.

Thus, we shall assume always

$$|r| \le 1$$

(This assumption is in reality an identification condition.)

In summary:

SOLUTION OF DIFFERENCE EQUATIONS:

x_t = Sum of

* damped exponentials in time ($\to 0$)
* polynomials in time (deterministic trends)
* cosine functions (seasonal and other cycles)

We shall come back often to this result!

Remark

Using the backward operator B, the difference equation can be written as

$$\phi(B)\ x_t = 0 \quad ,$$

where

$$\phi(B) = 1 + \phi_1\ B + \ldots + \phi_p\ B^p \quad .$$

Comparing this last expression with the characteristic equation

* roots of $[\phi(B) = 0] = \left[\frac{1}{r_i}\right]$.

Thus,

In terms of the roots of $\phi(B) = 0$, the condition

$$|r| \le 1$$

becomes

$$|b| \ge 1 \quad ,$$

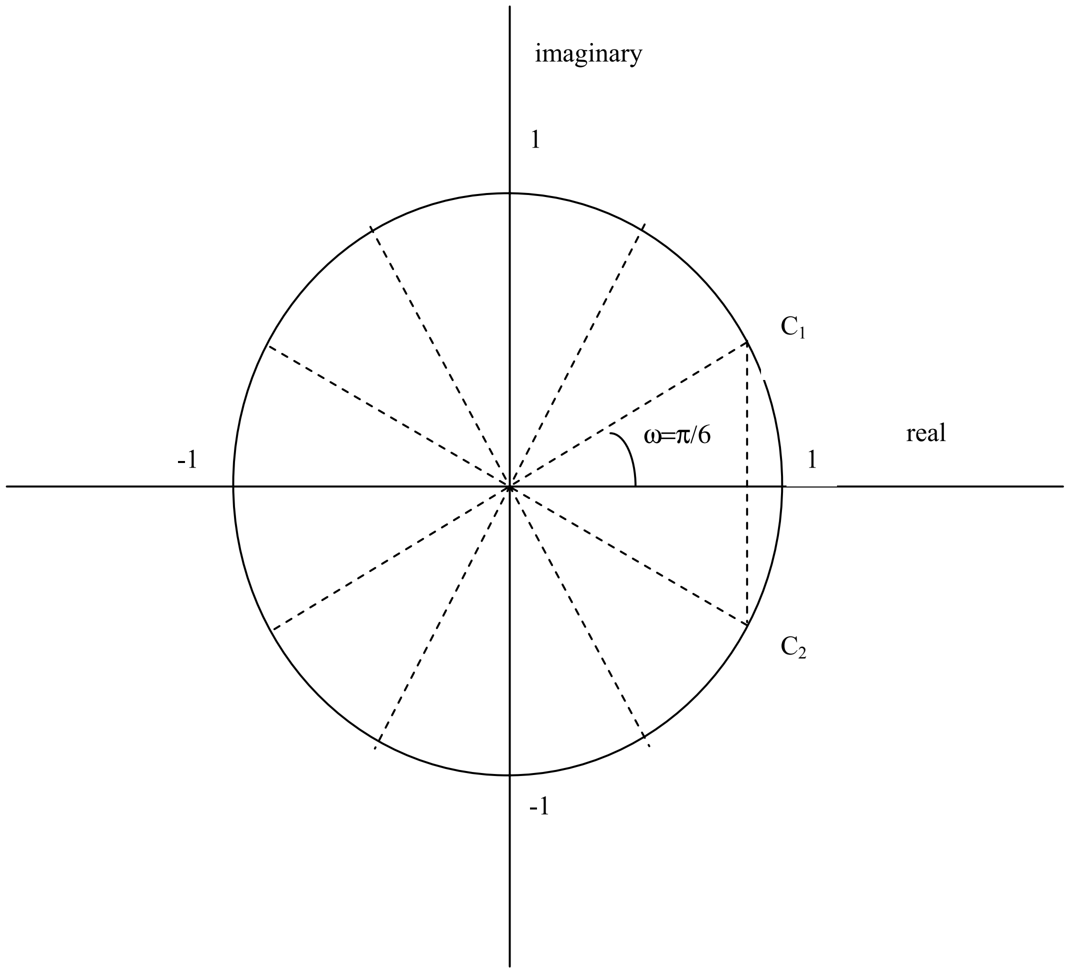
where b is the corresponding root of $\phi(B) = 0$.

Back to the ∇_{12} example.

Thus, for $x_t - x_{t-12} = 0$, the characteristic equation: $r^{12} - 1 = 0$;

or $r = (1)^{1/12}$,

or twelve roots of unit circle



All 12 roots have unit modulus.

Roots are:

* 2 real roots: $r_1 = 1$

$r_2 = -1$

* 10 complex roots,

  in pairs of complex conjugates

Each complex conjugate pair is associated with a frequency

$$\omega = \frac{\pi}{6}$$

$$\omega = 2\,\frac{\pi}{6}$$

$$\omega = 3\,\frac{\pi}{6}$$

$$\omega = 4\,\frac{\pi}{6}$$

$$\omega = 5\,\frac{\pi}{6}$$

Each pair will generate a solution of the type

$$A \cos(\omega t + B)$$

Consider the frequency $\omega = \frac{\pi}{6}$.

How many periods are needed to complete a full circle?

$\Rightarrow$ 12 periods.

Hence frequency implies 1 circle (or cycle) per year.

For frequency $\omega = 2\frac{\pi}{6}$, 6 periods are needed to complete the circle, hence $\omega = \frac{\pi}{3}$ implies that 2 circles per year are completed. For frequency $\omega = \frac{3\pi}{6} = \frac{\pi}{2}$, 4 periods are needed to complete the circle, hence $\omega = \frac{\pi}{2}$ $\Rightarrow$ 3 circles completed per year, and so on.

…..

Finally, for $\omega = 6\frac{\pi}{6} = \pi$, the root is real and equal to r = -1.

For this frequency, a full circle is completed in two periods.

Hence, for monthly data,

r = - 1 $(\omega = \pi)$ $\Rightarrow$ 6 circles per year

Notice that the root r = -1 implies that the factor (1 + B) appears in the factorization of the AR polynomial. (Such is the case when this polynomial is S, or ∇_s .)

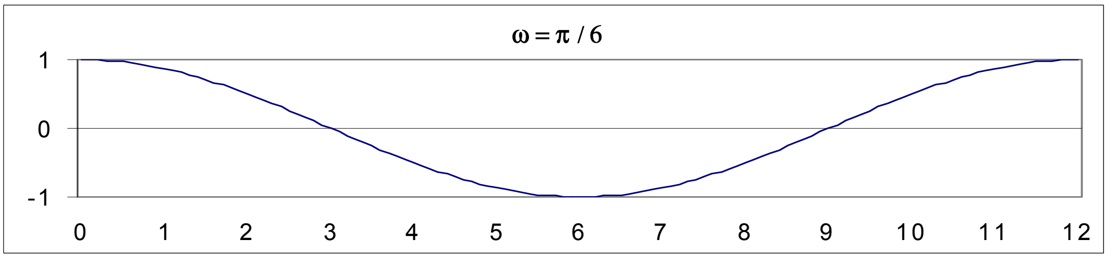
In short:

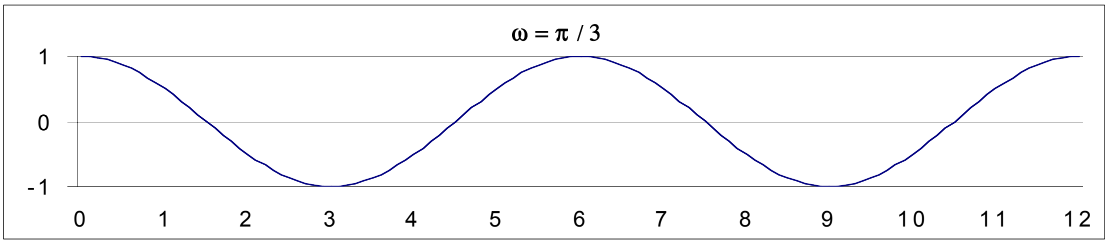
$$X_t = C + \sum_{j=1}^{6} A_j \cos\left(j\frac{\pi}{6} t + B_j \right)$$

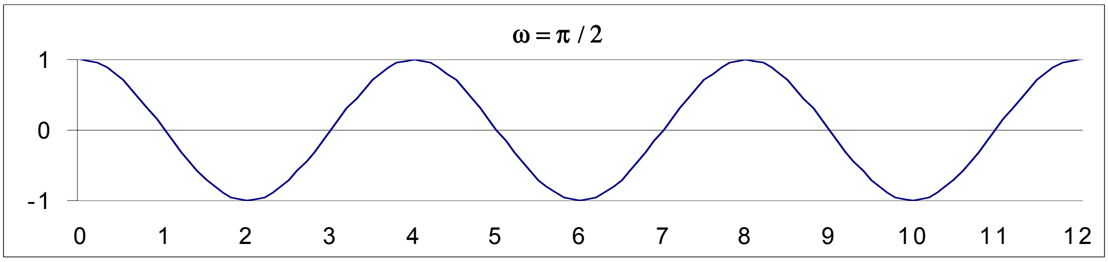
C: constant, associated with zero frequency root B = 1 (i.e., with the factor (1-B))

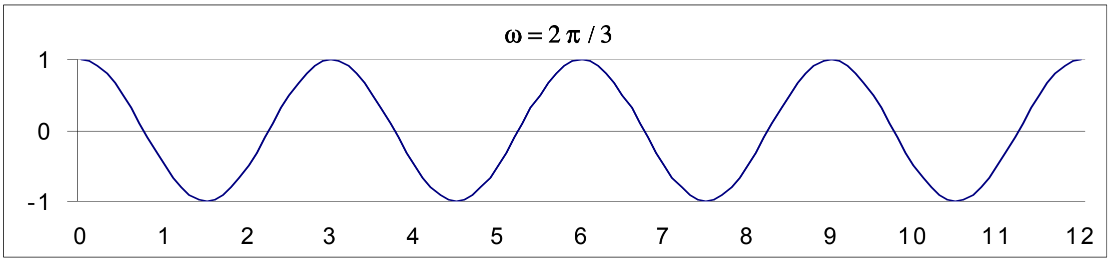
$\left(j\frac{\pi}{6} \right)$: seasonal frequencies.

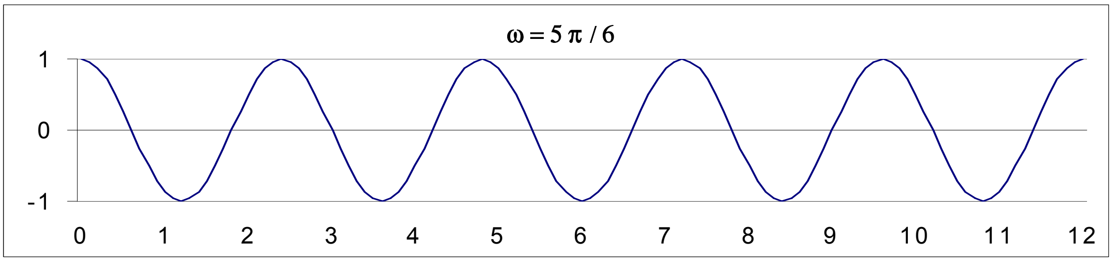
j = 1	once a year	: " Fundamental " frequency
j = 2	twice a year	" harmonics "
...	....	
j = 6	six times a year	

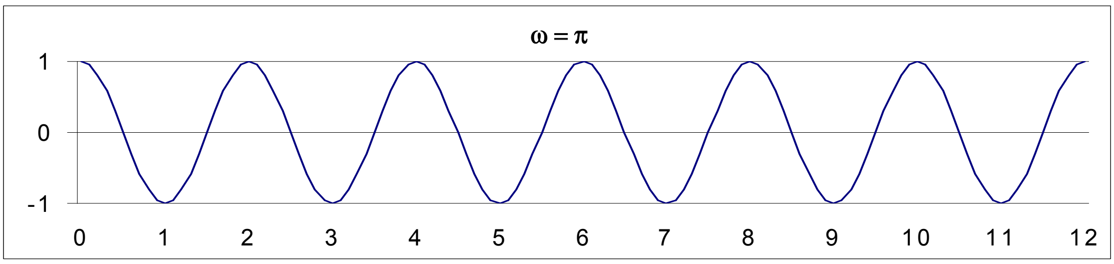

ω = π / 6
1
0
-1
0 1 2 3 4 5 6 7 8 9 10 11 12


ω = π / 3
1
0
-1
0 1 2 3 4 5 6 7 8 9 10 11 12


ω = π / 2
1
0
-1
0 1 2 3 4 5 6 7 8 9 10 11 12


ω = 2 π / 3
1
0
-1
0 1 2 3 4 5 6 7 8 9 10 11 12


ω = 5 π / 6
1
0
-1
0 1 2 3 4 5 6 7 8 9 10 11 12


ω = π
1
0
-1
0 1 2 3 4 5 6 7 8 9 10 11 12

Another way to look at it:

∇_{12} can be factorized as

	AR factors	Frequency
$1-B^{12} =$	$(1-\sqrt{3}\ B+B^2)$ x	Once a year
	$(1-B+B^2)$ x	Twice a year
	$(1+B^2)$ x	3 times a year
	$(1+B+B^2)$ x	4 times a year
	$(1+\sqrt{3}B+B^2)$ x	5 times a year
	$(1+B)$	6 times a year
	$(1-B)$	associated with trend

$= (1-B)S$

1 – B contains a trend root,

S contains the seasonal roots (one real and 5 pairs of complex conjugates).

All of them are "unit roots" (unit in modulus).

Hence, for example,

$$\nabla \nabla_{12} x_t = 0 \ , \qquad (1)$$

since $\nabla \nabla_{12} = \nabla^2 S$,

will cancel

(A) the polynomial $p_t = a + b\,t$ $\left(\nabla^2 p_t = 0\right)$

and

(B) the seasonal cycles for the 1, 2, ..., 6 times-a-year frequencies ($S\, s_t = 0$)

$$s_t = \sum_j A_j \cos(\omega_j t + B_j)$$

$$\omega_j = j \frac{2\pi}{12} \quad , \quad j = 1,2,\ldots, 6.$$

Notice that, when specifying stochastic ARIMA models, we will

not say that $\nabla\nabla_{12}\, x_t$ is exactly = 0, but instead that

$$\nabla\nabla_{12}\, x_t = z_t$$

where z_t is a zero mean, finite variance stationary stochastic process.

That is, $\nabla\nabla_{12}\, x_t$ will on average be zero and will not depart too much from it.

Thus, every period, the functions (A) and (B) will be

"perturbated" by a stochastic input,

so that

$$a, b \rightarrow a^{(t)}, b^{(t)}$$

$$C \rightarrow C^{(t)}$$

$$A_j \rightarrow A_j^{(t)}$$

$$B_j \rightarrow B_j^{(t)}$$

and so on.

Thus, the "MOVING" and ADAPTIVE features of the components.

DISTRIBUTION OF THE STATIONARY SERIES

So, let in general

$$\delta(B) = \nabla^d \nabla_s^D$$

represent all the differences applied for reaching stationary, i.e.,

$z_t = \delta(B)\ x_t$ is stationary.

Implication : $[\ z_1,\ ...,\ z_t\ ]$ will have a well-defined proper joint distribution.

Further, we assume: JOINT NORMALITY

Hence, we consider: LINEAR STATIONARY STOCHASTIC PROCESSES

The time series generated by them will be jointly Normally distributed.

In the multivariate Normal distribution, conditional expectations are linear functions of the observed series. For ex.:

Expectation of a future value:

$E\ (\ z_{t+j}\ |z_1, ..., z_t\ )$: forecast $(\ j > 0\ )$

Expectation of a missing value:

$E\ (z_{(t)}\,|\,z_1 \ldots z_{t-1}\ z_{t+1} \ldots z_T\ )$: interpolator ($z_t$ missing)

Expectation of a signal s_t buried in z_t ($z_t = s_t$ + noise)

$E\ (s_t\ |\ z_1\ ,\ldots,\ z_T)$: signal extraction

all are linear in the observations (i.e., **linear filters**)

Stationarity $\Rightarrow$

$$Ez_t = \mu$$

$$Vz_t = V_z$$

$$\text{Cov}(z_t, z_{t-k}) = \gamma_k$$

γ_k: only depends on $|k|$, the relative distance between observations.
(it does not depend on t)

Thus,

$$(z_1, \ldots, z_T) \sim N_T(\mu, \Sigma)$$

$$\Sigma = \begin{bmatrix} \gamma_z & \gamma_1 & \gamma_2 & \cdots & \gamma_{T-1} \\ & \gamma_z & \gamma_1 & \cdots & \\ & \cdots & \cdots & \cdots & \\ & & & \gamma_z & \gamma_1 \\ (\text{sym.}) & & & & \gamma_z \end{bmatrix}$$

(elements in each diagonal are the same)

More parsimonious representation:

AUTOCOVARIANCE FUNCTION:

A Cov F = γ_k as a function of k

Let $F = B^{-1}$ (i.e., $F\, z_t = z_{t+1}$);

F $\equiv$ "Forward" Operator

AUTOCOVARIANCE GENERATING FUNCTION:

$$\text{ACov G.F.} = \gamma(B, F)$$

$$\gamma(B, F) = \gamma_0 + \gamma_1(B + F) + \gamma_2(B^2 + F^2) + \ldots$$

$$= \gamma_0 + \sum_{j=1}^{\infty} \gamma_j (B^j + F^j)$$

Better (scale free) measure: autocorrelation

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \text{Lag - k} \quad \text{autocorrelation}$$

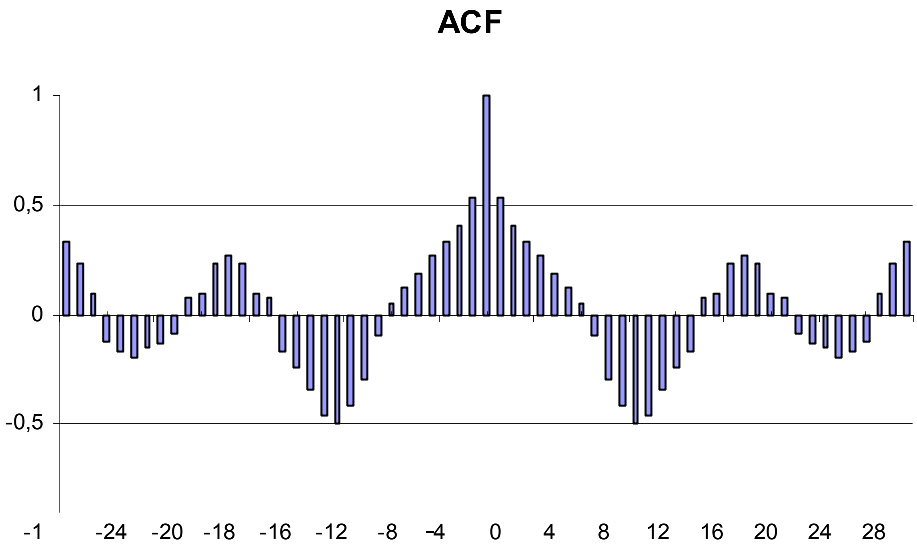
AUTOCORRELATION FUNCTION

$$\text{ACF} \equiv \rho \text{ as function of k}$$

(since symmetric, only needed for k > 0)

AUTOCORRELATION GENERATING FUNCTION

$$\text{ACGF} \equiv \rho(B, F) = 1 + \sum_{j=1}^{\infty} \rho_j (B^j + F^j)$$



Results:

If z_t stationary, then

1) $\rho_0 = 1$,

2) $\rho_j = \rho_{-j}$, (symmetric)

3) $|\rho_k| < 1, \quad k \neq 0$

4) $\rho_k \to 0 \quad \text{as} \quad k \to \infty$

5) $\sum_{k=0}^{\infty} |\rho_k| < \infty$ (Convergence condition)

Note: JOINT NORMALITY implies that

$\mu, \quad \gamma_0, \quad \rho(B, F)$

fully characterize the joint distribution function of $[z_1, ..., z_T]$; they contain all the "sample" information.

ρ_k = Lag-k autocorrelation is a measure of the (linear) dependence between observations distant k periods

Wold Representation

A linear ($\equiv$jointly Normal) stationary purely stochastic process can be expressed as:

$$z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

$$= \sum_{j=0}^{\infty} \psi_j a_{t-j}, \; (\psi_0 = 1),$$

where

$$a_t = \begin{cases} -\text{niid}\,(0, V_a) \equiv \text{white noise variable} \\ -\text{"residuals"} \\ -\text{"innovations"} = 1\text{ p. a. forecast error (known parameters)} \end{cases}$$

$$a_t = z_t - \hat{z}_{t|t-1}$$

($\hat{z}_{t|t-j}$: forecast of z_t made at period t-j)

Hence

z_t = Linear filter applied to innovations. Also called the MA representation of z_t. Filter is one-sided (only past and present innovations) and convergent,

$$\left.\begin{array}{l} \psi_j \to 0 \\ \\ j \to \infty \end{array}\right\}$$

$$\sum_0^{\infty} |\psi_j| < \infty \quad .$$

In short:

$$\boxed{z_t = \psi(B)\, a_t} \qquad \psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$$

Useful result:

Let $\gamma_z(B, F) = \text{A Cov GF}(z_t)$ Then,

$$\boxed{\gamma_z(B, F) = \psi(B)\, \psi(F)\, V_a}$$

In particular, for the variance :

$$\gamma_0 = (1 + \psi_1^2 + \psi_2^2 + \dots)\, V_a$$

ACF: Basic tool in "Time Domain analysis" of a series.

Another important tool:

Spectrum

Basic tool in "Frequency Domain analysis" of a series.

Consider the time series:

$$x_t = [x_1, x_2, \ldots , x_T]$$

We can “exactly explain” the series with the polynomial of degree (T-1):

$$x_t = a_0 + a_1 t + \ldots + a_{T-1} t^{T-1}$$

(Set, succesively, t = 1, 2, …, T, and a linear system of T equations in the T unknowns a_0 , a_1 , ... , a_{T-1} is obtained).

In a similar manner, we can represent (exactly) the T observations $[x_t]$ with sine-cosine functions as follows.

To simplify, assume T is even, so that

$$T = 2q$$

Define the Fundamental Frequency

$$\omega_1 = \frac{2\pi}{T}$$

(i.e., the frequency of one full circle completed in T periods), and its Harmonics:

$$\omega_j = \left(\frac{2\pi}{T}\right) j \quad , \quad j = 1, 2, \ldots, q.$$

Then, express x_t as

$$x_t = \sum_{j=1}^{q} \left(a_j \cos \omega_j t \; + \; b_j \sin \omega_j t \right)$$

Letting $t = 1, 2, \ldots, T$,

a linear system of T equations in the T unknowns : $\left(a_j , b_j \right)$,

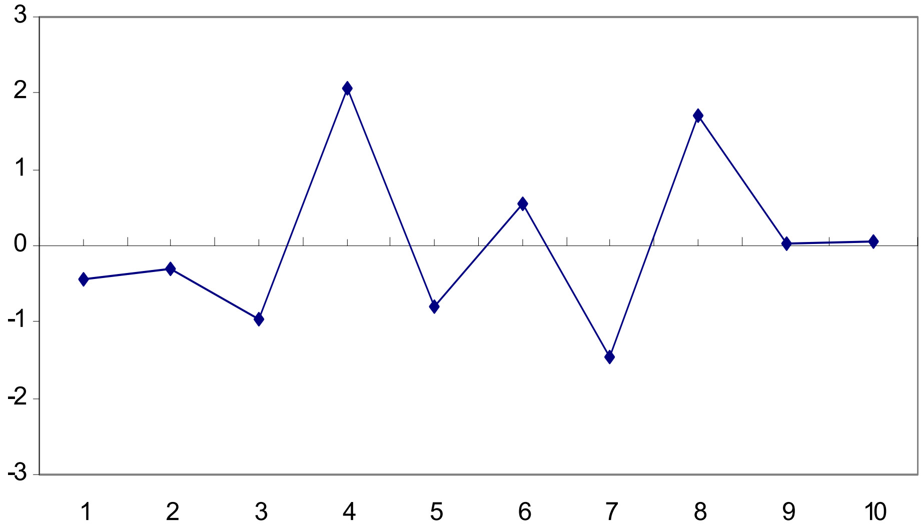
$j = 1, \ldots, q$, is obtained.

For a particular periodic component (with frequency ω_j), the "amplitude" (i.e., the height of the peak) is equal to

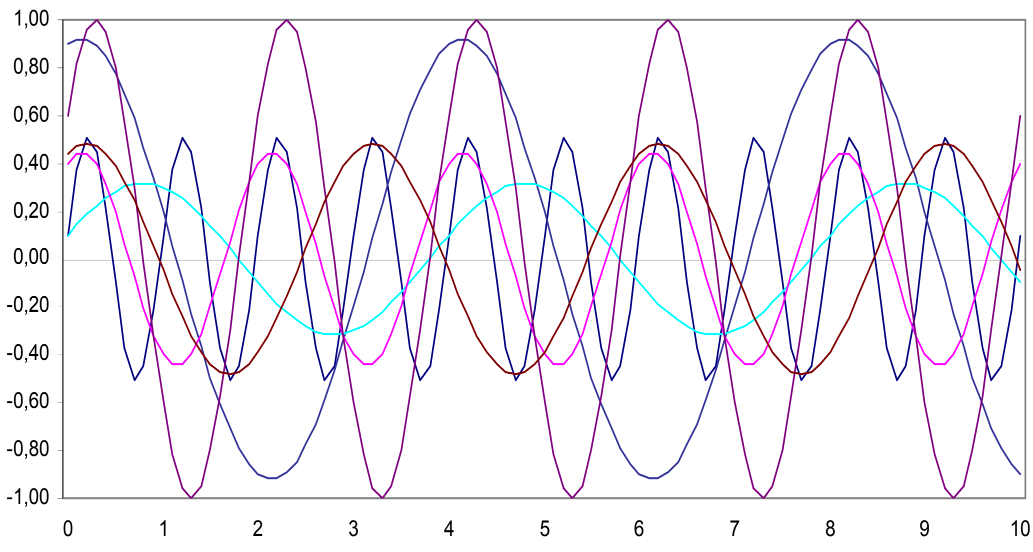
$$A_j^2 = a_j^2 + b_j^2$$

Notice: The bigger this amplitude, the larger will be the contribution of the component in explaining x_t .

Example:



3
2
1
0
-1
-2
-3
1 2 3 4 5 6 7 8 9 10

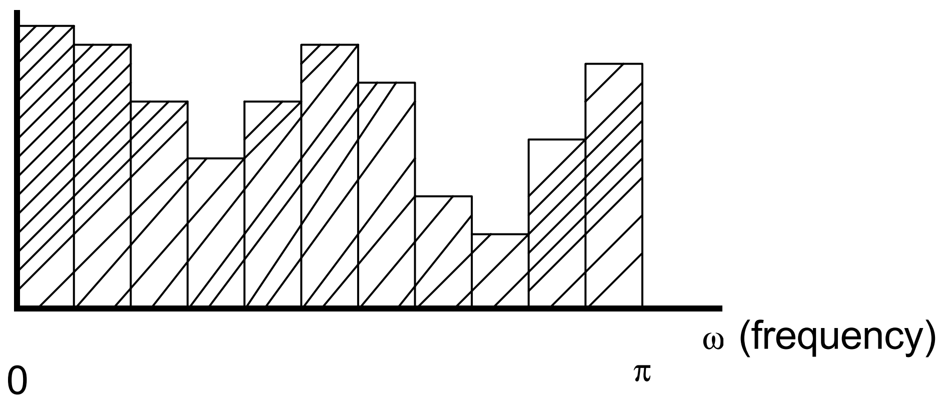


1,00
0,80
0,60
0,40
0,20
0,00
-0,20
-0,40
-0,60
-0,80
-1,00
0 1 2 3 4 5 6 7 8 9 10

In general, group the cosine functions by intervals of frequency

(summing the amplitudes)

HISTOGRAM OF THE DISTRIBUTION BY FREQUENCY



In the same way that:

density function ≡ population counterpart of standard histogram,

spectrum ≡ population counterpart of frequency histogram.

Given that

$$\cos(\alpha) = \cos(\alpha + 2\pi),$$
$$\cos(\alpha) = \cos(-\alpha),$$

spectrum is periodic and symmetric $\Rightarrow$

enough to consider: $0 \leq \omega \leq \pi$

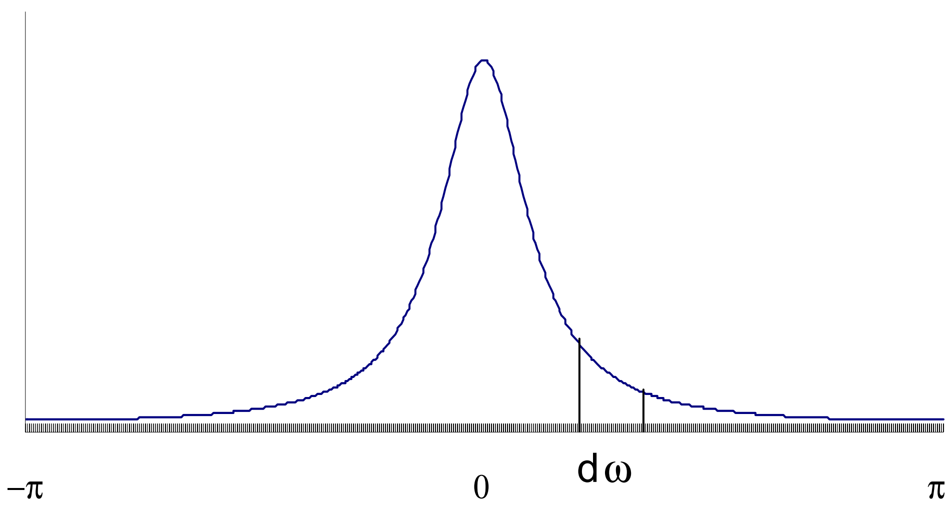
$$\int_0^{2\pi} g(\omega)\, d\omega = \gamma_0 \quad \text{(variance of series)}$$

(or $\int_{-\pi}^{\pi}$)

$\therefore$ Spectrum can be interpreted as a decomposition of variance by intervals of frequency.

(Standarized) it displays properties similar to those of a density function.

SPECTRUM SERIES

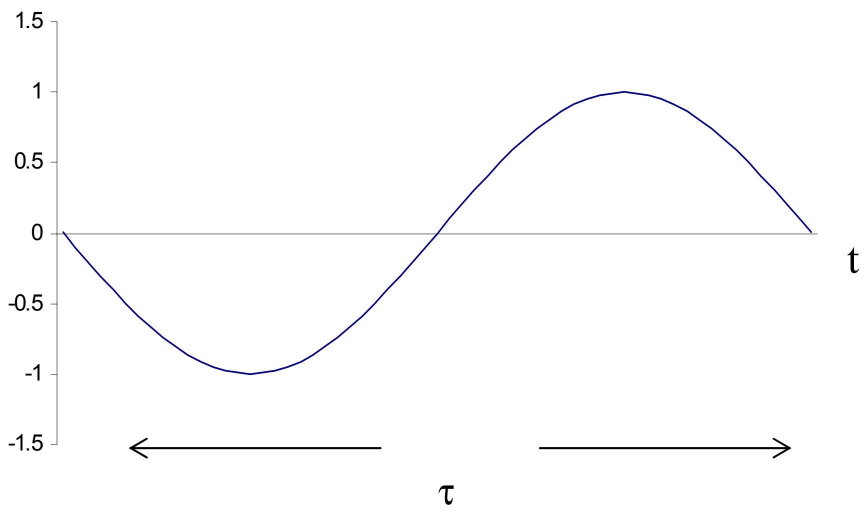


$\omega =$ frequency in radians

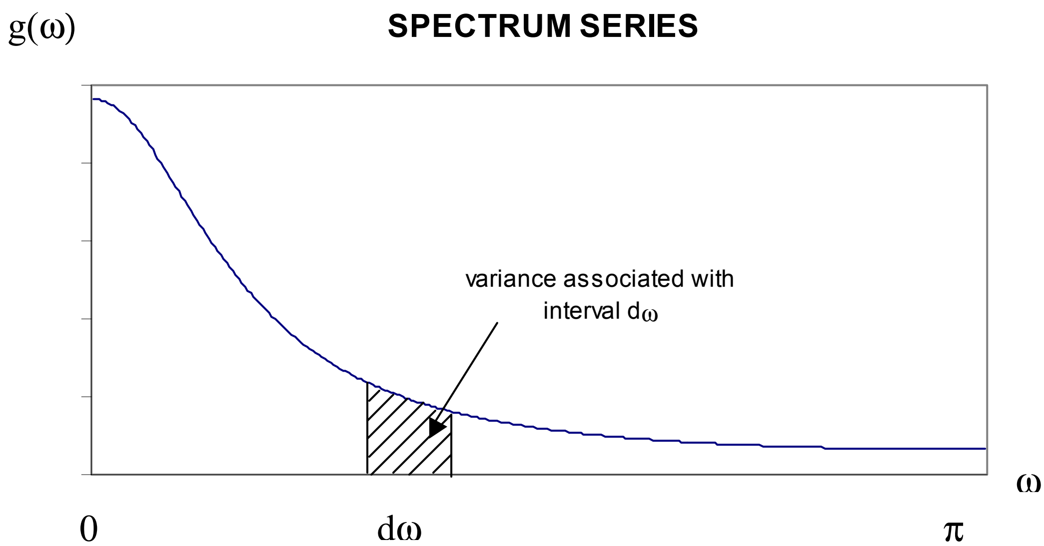
Recall:

$$\omega = \frac{2\pi}{\tau}$$

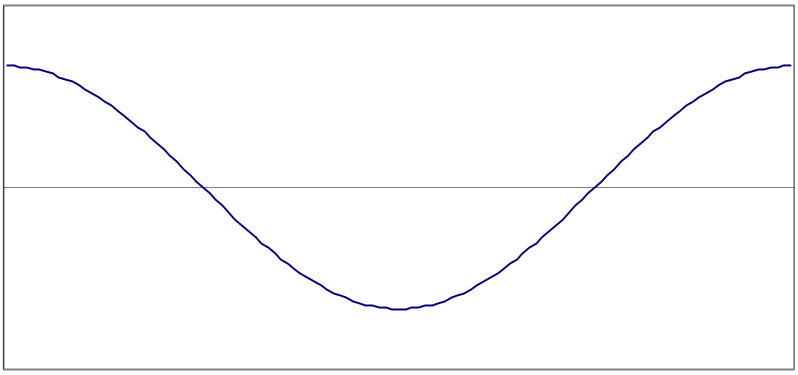
$\tau =$ Period



Spectrum $g(\omega)$: decomposes V_x by frequency.



Consider the once-a-year frequency in monthly data

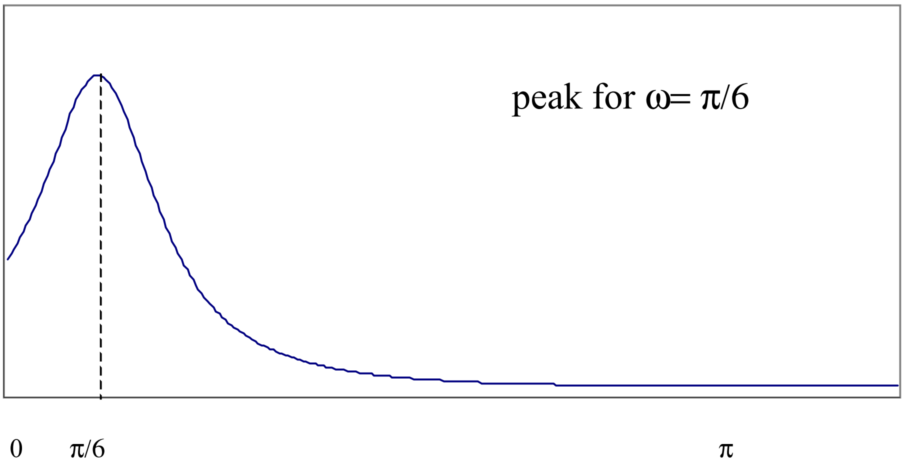


$$\tau = 12 \;\Rightarrow\; \omega = \frac{2\pi}{12} = \frac{\pi}{6}$$

Hence: If series has important seasonal component with that frequency,

in $g_x(\omega)$:

SPECTRUM AR(2)

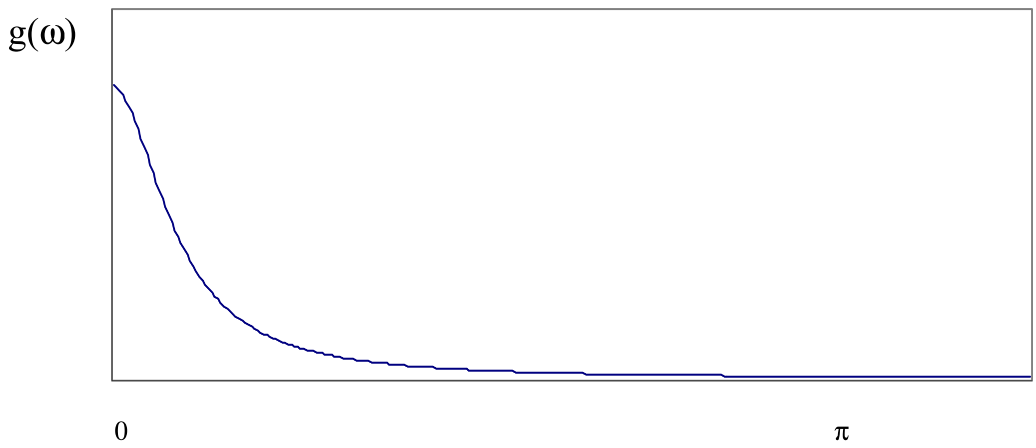


For trend:

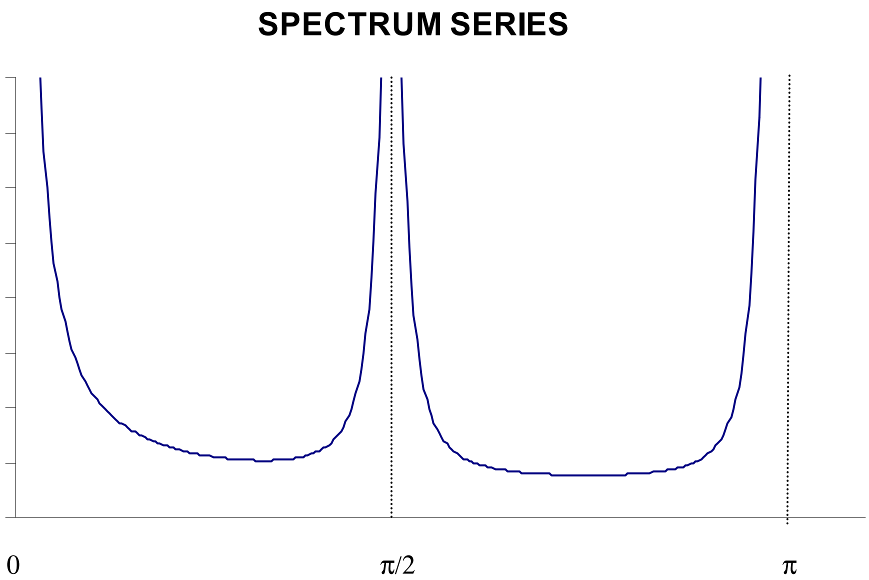
A way to think about trends:

Cycles with period close to ∞ (ex.: cycles with periods 1000 years, 10.000 years,…)

$\tau \to \infty \;\Rightarrow\; \omega \to 0$



If the spectrum of a quarterly series is, for example:
$g_x(\omega)$



Peak for:

$\omega = 0 \longrightarrow$ Trend

$\omega = \frac{\pi}{2} \rightarrow$ Once a year

$\omega = \pi \rightarrow$ Twice a year

} seasonal frequencies

To extract some signal from a series, for example, to S.A. a series:

- remove variation around seasonal frequencies
- leave the rest unchanged.

If n_t = SA series is estimated through

$$\hat{n}_t = c(B,\ F)\, x_t = \ldots + c_2\, x_{t-2} + c_1\, x_{t-1} + c_0\, x_t + c_1\, x_{t+1} + c_2\, x_{t+2} + \ldots$$

where

$$c(B, F) = c_0 + \sum_j c_j \left(B^j + F^j\right)$$

is a symmetric filter, the F.T. of the filter

(recall: $B^j + F^j \rightarrow 2 \cos j\,\omega$)

is

$$\tilde{c}(\omega) = c_0 + 2 \sum_j c_j \cos(j\,\omega)$$

From $\hat{n}_t = c(B,\ F)\ x_t$,

$$g_{\hat{n}}(\omega) = \left[\tilde{c}(\omega)\right]^2\ g_x(\omega)$$

$\tilde{c}(\omega) =$ Gain of filter

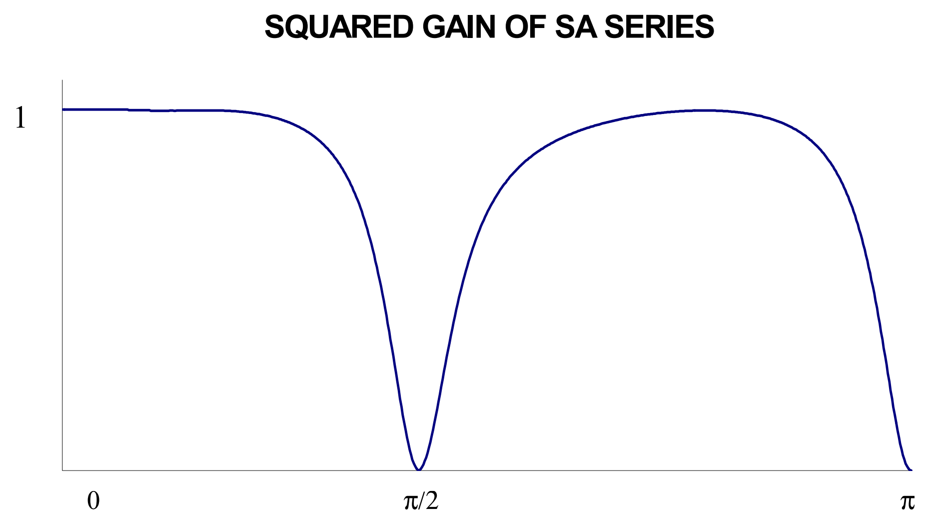
$\left[\tilde{c}(\omega)\right]^2 =$ Squared gain of filter

Hence:

Spectrum of $\hat{n}_t$ = (Squared gain of filter) x (Spectrum of series)

Squared gain: Determines, for each frequency, which proportion of the series variance is passed on to the signal estimator.

= 1	all the variation is passed
= 0	the frequency is ignored



SPECTRUM OF A LINEAR PROCESS
(AND OF AN ARIMA MODEL)

In general, from ACovGF, the spectrum is easily obtained as its Fourier Transform:

$$\gamma(B,F) = \gamma_0 + \sum \gamma_j (B^j + F^j)$$

$$\left. \begin{array}{r} B \rightarrow e^{-i\omega} \\ B^j + F^j \Rightarrow 2\cos j\omega \end{array} \right\} \quad \text{F.T.}$$

$$0 \leq \omega \leq 2\pi$$

$$g(\omega) = \frac{1}{2\pi}\left[\gamma_0 + 2\sum \gamma_j \cos j\omega\right]$$

For notational simplicity, we shall work with

$$\mathrm{g}(\omega) = 2\pi\, g(\omega) \, ,$$

and avoid the factor 2π.

Ex. 1:

ACF and spectrum of MA (2)

$$x_t = (1 + \theta_1 B + \theta_2 B^2)\, a_t \qquad [V_a = 1]$$

$$\text{ACovGF} = \theta(B)\,\theta(F) \quad ,$$

$$(1 + \theta_1 B + \theta_2 B^2)\ (1 + \theta_1 F + \theta_2 F^2)$$

$= 1 + \theta_1^2 + \theta_2^2$ coeff . of B^0 : γ_0

$+ \theta_1 (1 + \theta_2) [B + F]$ coeff . of B and F : γ_1

$+ \theta_2 [B^2 + F^2]$ coeff. of B^2 and F^2 : γ_2

In short

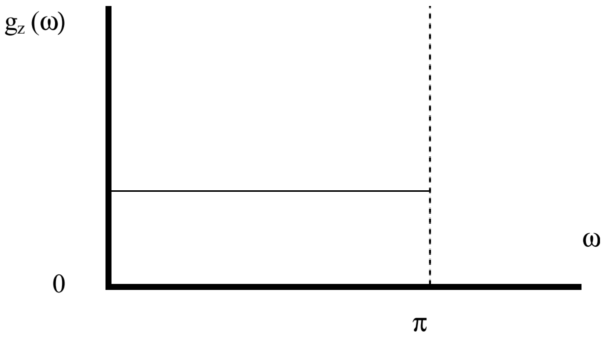
$$B^k + F^k \rightarrow 2 \cos k\, \omega$$

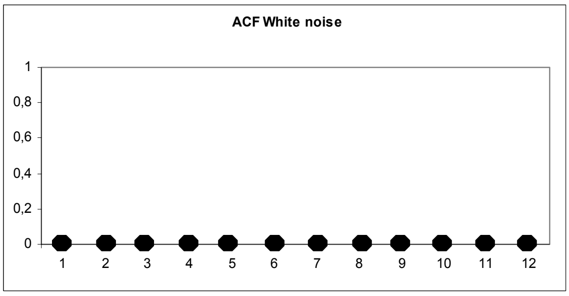
$$g(\omega) = [\gamma_0 + 2\,\gamma_1 \cos \omega + 2\,\gamma_2 \cos 2\omega]\, V_a$$

(γ_0, $2\,\gamma_1$, $2\,\gamma_2$: coeff. of "harmonic" expressions)

Ex. 2:

$z_t \sim$ white noise $\Rightarrow g_z(\omega) = \text{constant}$





Ex. 3:

AR(1)

$$z_t + \phi z_{t-1} = a_t, \qquad |\phi| < 1$$

$$(1 + \phi B) z_t = a_t,$$

$$z_t = \left[\frac{1}{1 + \phi B} \right] a_t$$

$$z_t = \psi(B) a_t \quad \rightarrow \quad \psi(B) = \frac{1}{1 + \phi B} \qquad \text{(Wold representation)}$$

Therefore, ACovGF is: $\psi(B)\,\psi(F)\,V_a$

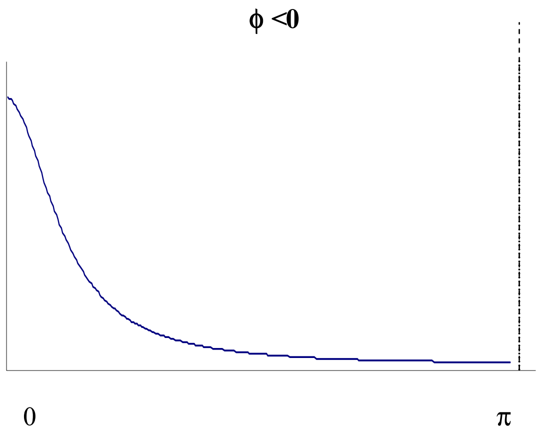
$$\gamma_z(B, F) = \frac{1}{(1 + \phi B)} \frac{1}{(1 + \phi F)} V_a =$$

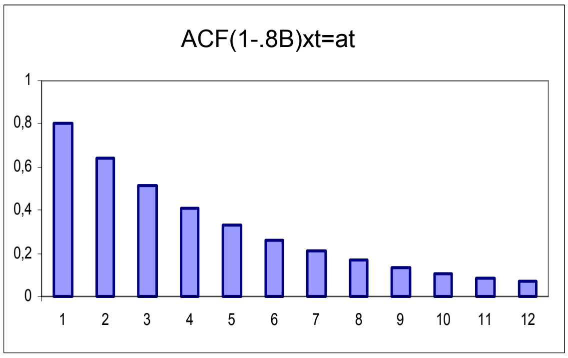
$$= \frac{1}{1 + \phi^2 + \phi(B + F)}$$

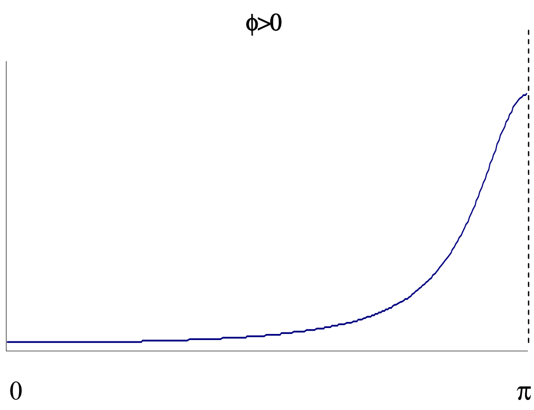
Hence the spectrum is: (B + F $\rightarrow$ 2 cos ω)

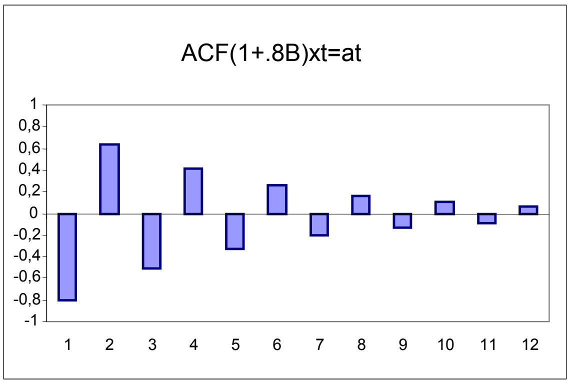
$$g_z(\omega) = \frac{1}{1 + \phi^2 + 2\phi\cos\omega} V_a$$

$g_z(\omega)$









Ex. 4 : PSEUDO-SPECTRUM

As $\phi \rightarrow -1$ the model approaches NONSTATIONARITY. In the limit:

$(1\text{-}B) z_t = a_t \equiv$ RANDOM WALK

$$\nabla z_t = a_t \qquad (v_a = 1)$$

$$z_t = a_t + a_{t-1} + a_{t-2} + a_{t-3} + \dots$$

As $t \rightarrow \infty$

* Mean is not defined $(\text{"}0 \cdot \infty\text{"})$.
* Variance goes to ∞.

As with stationary models, write

$$\psi(B) = \frac{1}{\nabla}$$

"Pseudo" A Cov. GF = $\left(\frac{1}{1-B}\right)\left(\frac{1}{1-F}\right) v_a$

(does not converge)

"Pseudo – spectrum" = F. T. of pseudo-A Cov. GF

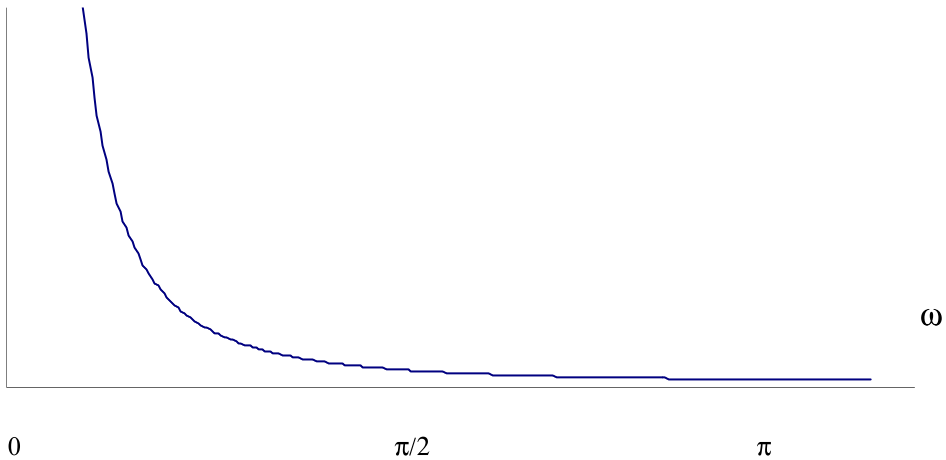
$$F.T. \equiv g_z(\omega) = \frac{1}{2\,(1-\cos\omega)}$$

$\omega = 0 \quad \Rightarrow g \to \infty$

$\int g$ does not converge (variance goes to ∞)

(pseudo) spectrum of a random walk

SPECTRUM SERIES



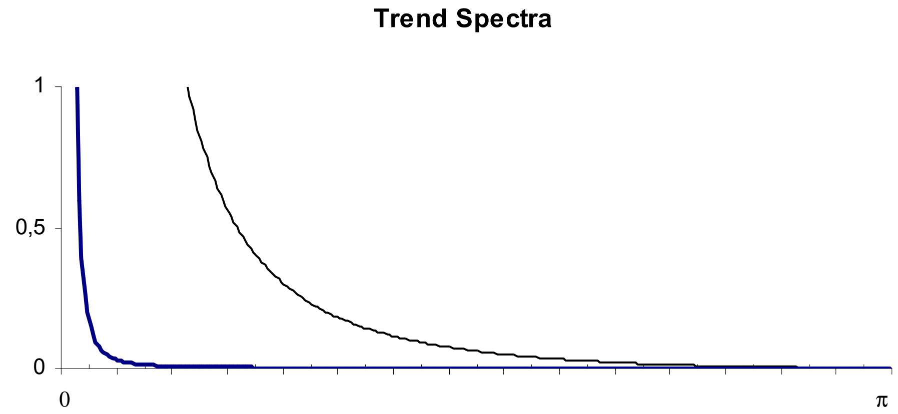
p-spectrum is:

- informative

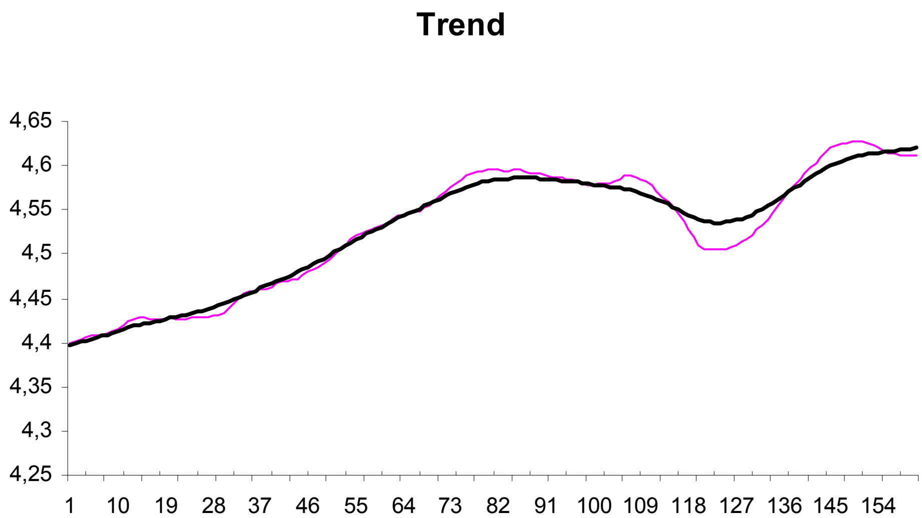
- well-behaved

in a very basic way.

For example, consider the two trend spectra:



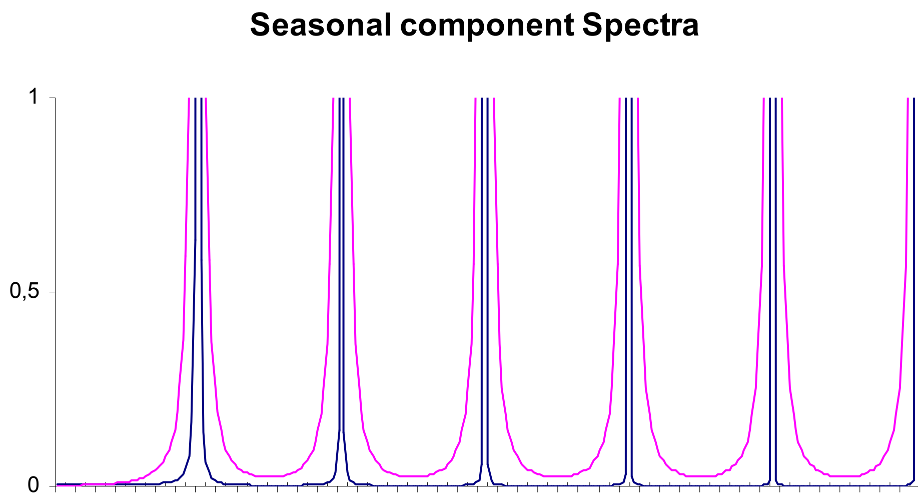
and two associated realizations



The trend that corresponds to the wider spectral peak contains more stochastic variability (i.e., is of a more "moving" nature).

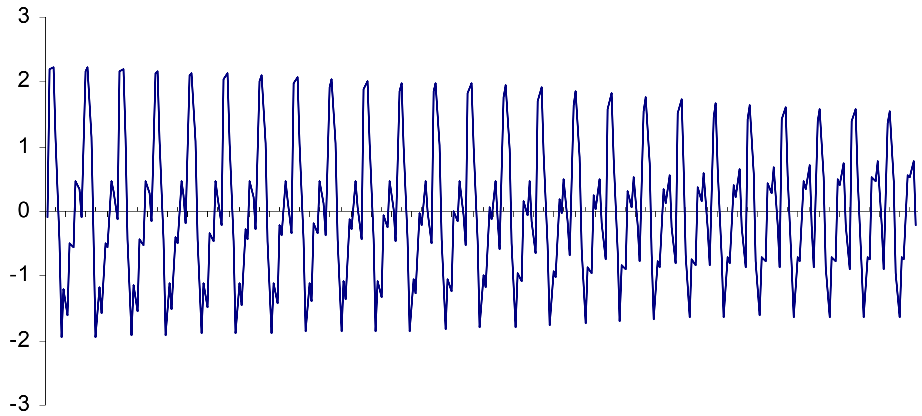
The narrow peak generates a more stable trend.

Similarly, from the two spectra:

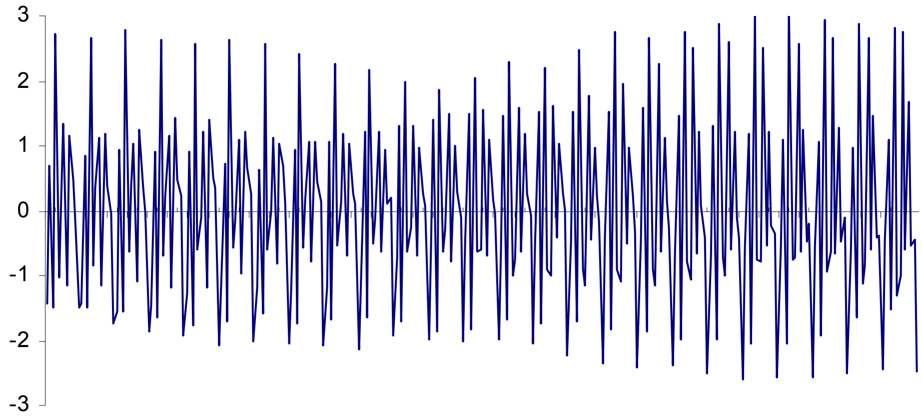


the following two seasonal components are generated

Seasonal Component



Seasonal Component



As was the case with the trend, the narrow spectral peaks produce stable seasonal components.

The wider peaks produce components that change faster (more moving components).

In what follows we shall also refer to the “p-spectrum” simply as the "spectrum".

Ex. 5 :   AR (2)

$(1+\phi_1 B+\phi_2 B^2)\ z_t = a_t$

$1+\phi_1 B+\phi_2 B^2 = 0$

Either:

- 2 real roots (each ~ AR(1))
- Complex conjugate root, $\rightarrow$ $z_t = r\cos(\omega t)+\cdots$

  with

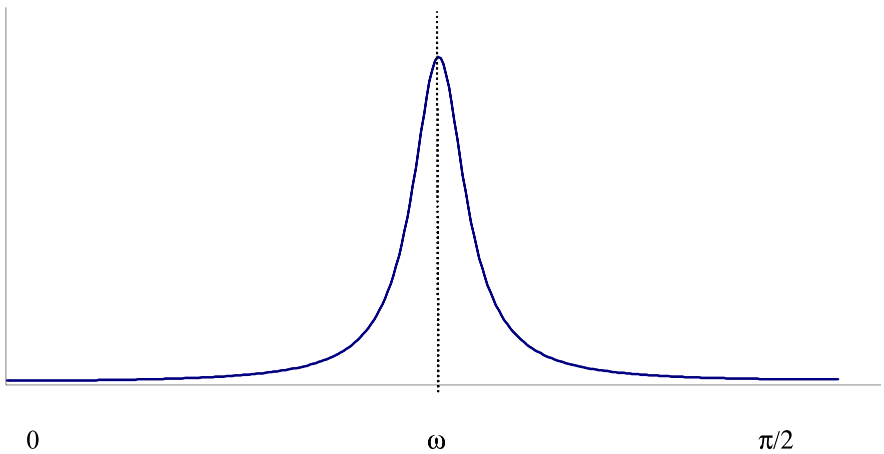
  * modulus $r = \sqrt{\phi_2}$

  * frequency $\omega = \arcos\left(\dfrac{\phi_1}{2r}\right)$ (in rads.)

  period $\tau = \dfrac{2\pi}{\omega}$

In this case, spectrum shows a peak for ω

SPECTRUM SERIES



General Hint: Useful way to look at AR(p) :

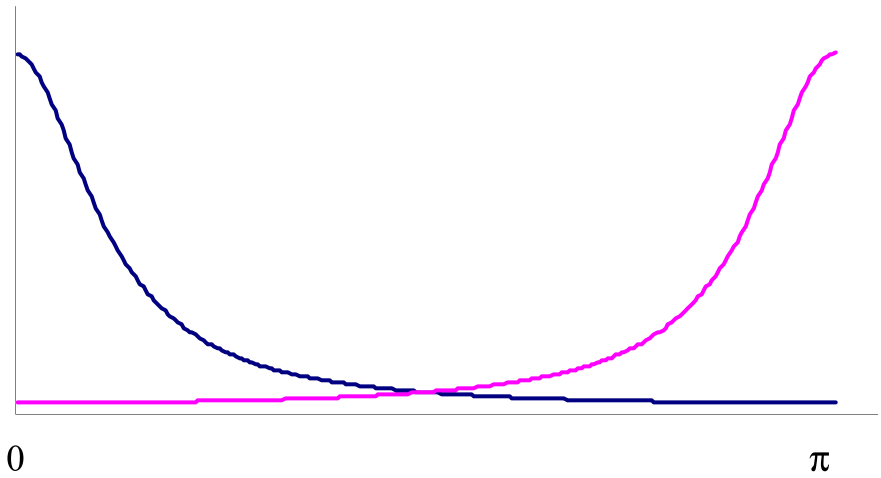
Factorize it as:

Real roots x Complex roots

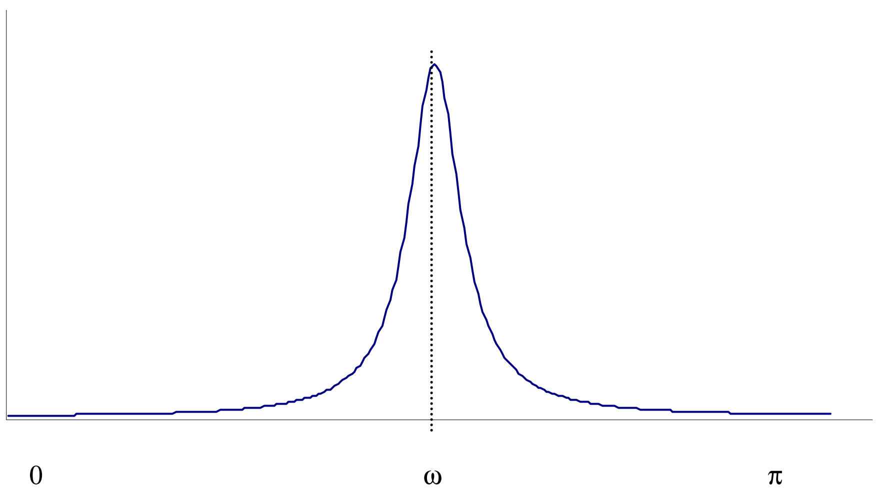
↓ ↓

each one : AR(1) each pair : AR(2)

SPECTRUM SERIES: Real roots AR(1)



SPECTRUM SERIES: Complex roots AR(2)



Ex. 6 : SEASONAL SERIES

If there is seasonal nonstationarity,

unit roots show up as ∞ for seasonal frequencies.

For example, consider a quarterly time series x_t such that

$$\nabla_4 \, x_t = z_t \quad ,$$

z_t : stationary process.

Then,

$$x_t = \frac{1}{\nabla_4} \, z_t = \frac{1}{\nabla S} \, z_t =$$

$$= \frac{1}{\nabla} \; \frac{1}{1+B+B^2+B^3} \, z_t =$$

$$= \frac{1}{\nabla} \; \frac{1}{(1+B)\,(1+B^2)} \, z_t$$

∇ : root associated with $\omega = 0$

$1+B^2$: root associated with $\omega = \pi/2$

$1+B$: root associated with $\omega = \pi$

$$\text{p ACGF}(x) = \frac{1}{(1-B)\,(1-F)} \; \frac{1}{(1+B)\,(1+F)} \; \frac{1}{(1+B^2)\,(1+F^2)} \; \gamma_z(B, F) \, .$$

Operating and using $B^j + F^j = 2 \cos j\omega$, the p-spectrum is

$$g_x(\omega) = \frac{1}{2(1-\cos\omega)} \quad . \qquad \infty \text{ for } \omega = 0$$

$$\frac{1}{2(1+\cos\omega)} \quad . \qquad \infty \text{ for } \omega = \pi$$

(seasonal freq.)

$$\frac{1}{2(1+\cos 2\omega)} \quad . \qquad \infty \text{ for } \omega = \frac{\pi}{2}$$

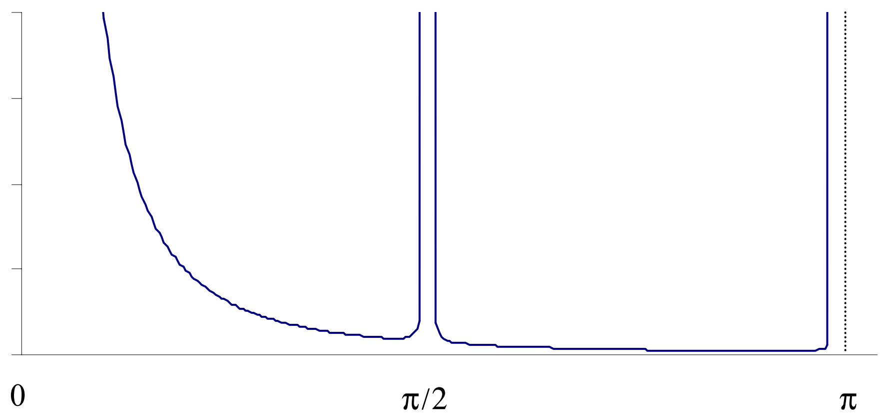
(seasonal freq.)

$\gamma_z(\omega)$ bounded.

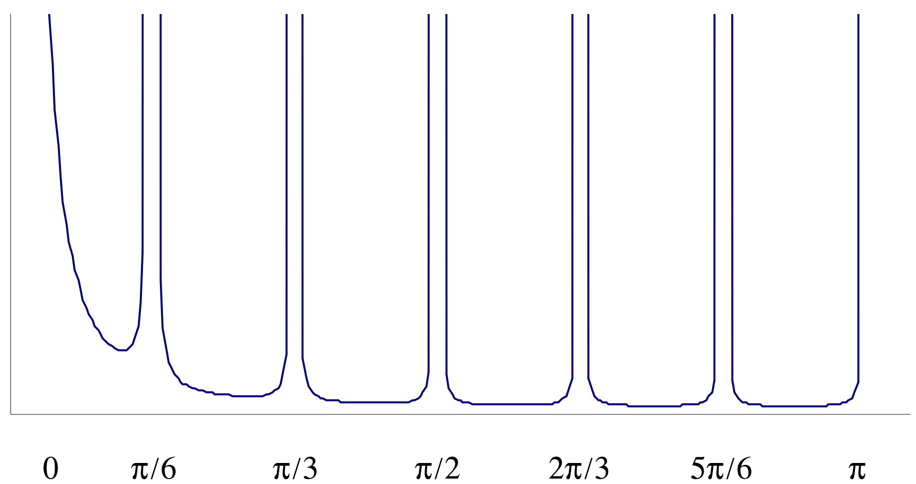
Unit AR roots dominate the spectrum of the series.

Thus, a "standard" series with trend and seasonality (both NS) will display a spectrum of the type:

SPECTRUM QUARTERLY SERIES



SPECTRUM MONTHLY SERIES



ARIMA models

Back to the Wold general representation of a purely stationary series:

$z_t = \psi(B)\, a_t$

Problem: In general $\psi(B)$ of degree $\infty \ldots$

Thus we use rational approximation:

$$\psi(B) = \frac{\theta(B)}{\phi(B)}$$

$\theta(B) =$ finite degree q

$\phi(B) =$ finite degree p

Therefore,

$$z_t = \frac{\theta(B)}{\phi(B)} a_t \ ,$$

or:

$$\phi(B)\, z_t = \theta(B)\, a_t$$

Autoregressive Moving-Average Models: ARMA models

AR (p) polynomial: $1 + \phi_1 B + \ldots + \phi_p B^p$
MA (q) polynomial: $1 + \theta_1 B + \ldots + \theta_q B^q$

$$(1 + \phi_1 B + \ldots + \phi_p B^p) z_t = (1 + \theta_1 B + \ldots + \theta_q B^q) a_t$$

ARMA (p,q) model.

Let $z_t = \delta(B) x_t$ [$\delta(B) \equiv$ stationary transformation]

If $\delta(B) = \nabla^d$

$x_t \sim$ ARIMA (p, d, q) model

I $\equiv$ integrated (of order d) , $x_t \sim I(d)$

Model:

$$\phi(B)\ \delta(B) x_t = \theta(B)\ a_t$$

$\phi(B)$: stationary
$\theta(B)$: invertible
$\delta(B)$: unit roots (Nonstationary roots)

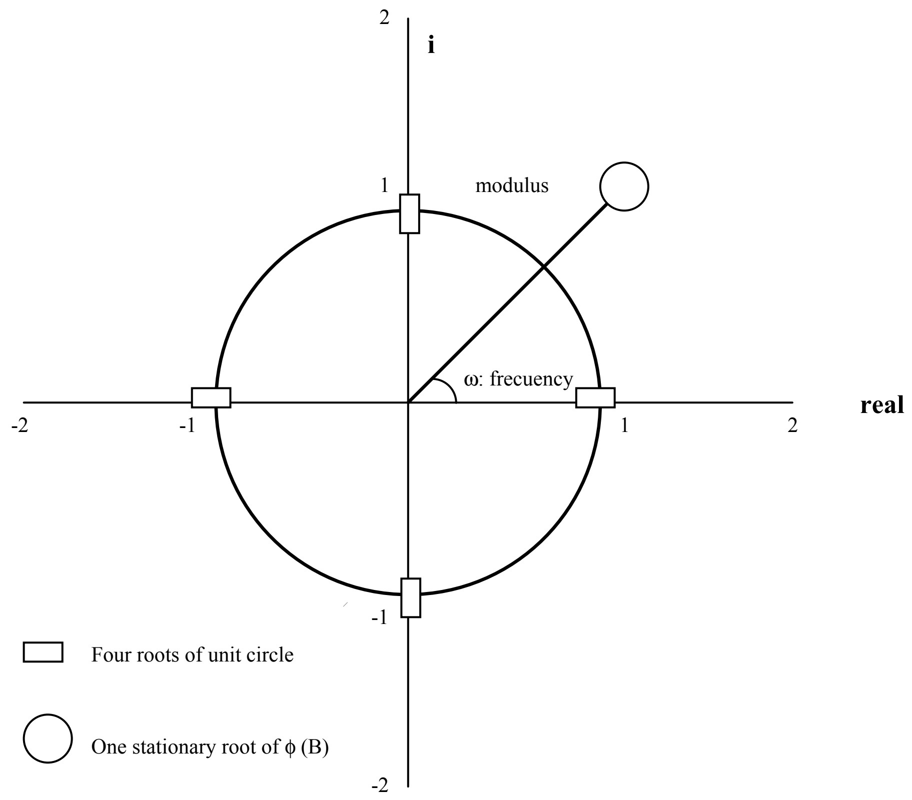
<u>Stationarity</u> of ARMA Models:

Roots of $\phi(B)=0$ lie outside the Unit Circle

UNIT CIRCLE = circle in the complex plane with radius = 1

Let $B_1, \ldots, B_p \equiv$ roots of ϕ (B) = 0.

Stationarity implies that



that is: moduli of the roots $B_1 \cdots B_p$ of $\phi(B)=0$ are > 1

Ex:

a) AR(1) : $x_1 + \phi\, x_{t-1} = a_t$

$(1+\phi\ B)\ x_t = a_t$

* root of $1+\phi B = 0 \Rightarrow B = -\frac{1}{\phi}$

* root outside U.C. when

$\text{mod}\,(B) = \left|-\frac{1}{\phi}\right| > 1 \Rightarrow \ |\phi| < 1$

b) Stationarity region for AR(2) :

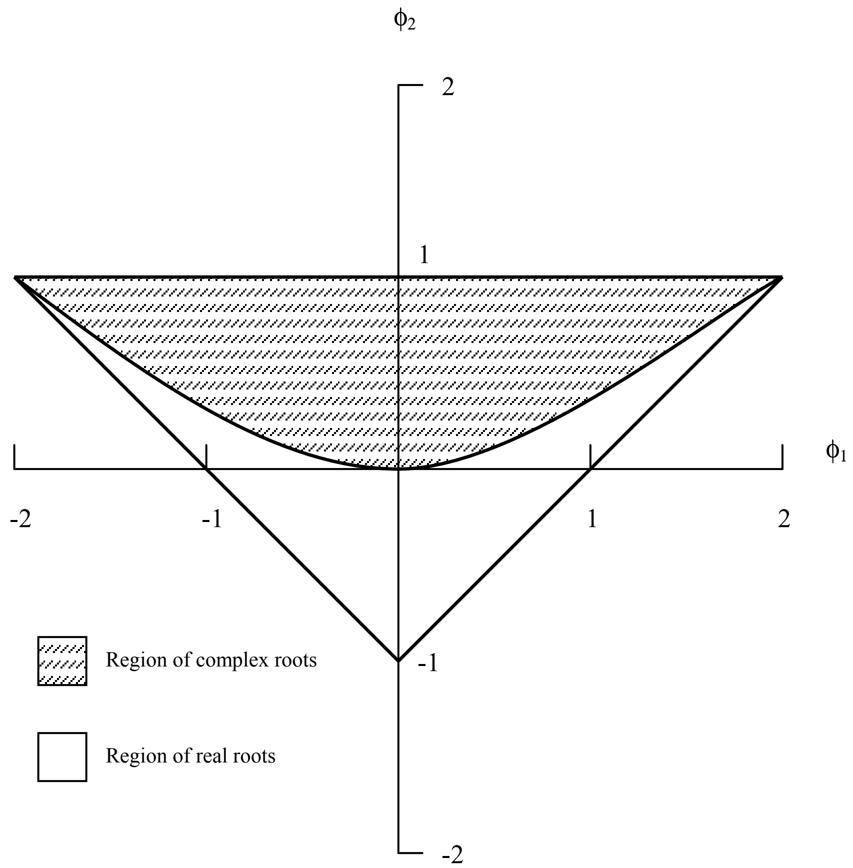
$x_t + \phi_1\, x_{t-1} + \phi_2\ x_{t-2} = a_t$

Conditions for roots of

$1+\phi_1 B+\phi_2 B^2 = 0$

to be > 1 in moduli.

Useful diagram:



The stationary region is the region inside the triangle.

The shaded area is the region of complex roots (periodic behavior).

If z_t is stationary, $\phi(B)^{-1}$ converges, and hence we can write

$z_t = [\ \phi(B)^{-1}\ \theta(B)\ ]\ a_t$.

- The series accepts a convergent MA representation,
- Its ACF converges to zero.

Invertibility

Roots of $\theta(B) = 0$ lie outside the Unit Circle.
Thus $\theta(B)^{-1}$ converges and we can write

$$[\theta(B)^{-1}\ \phi(B)]\ z_t = a_t$$

If z_t is invertible,
- it accepts a convergent AR representation

Remark:

- A unit AR root induces nonstationarity,

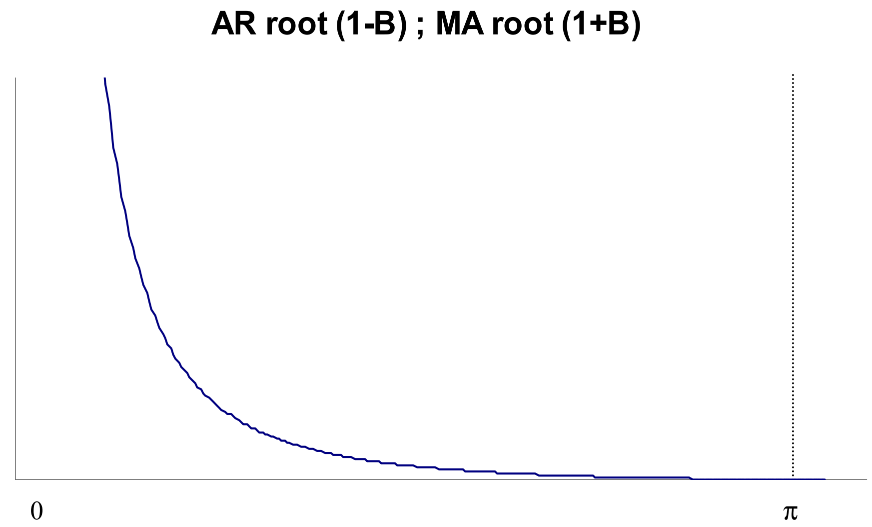
$$g_x(\omega\ \text{assoc. with U.R.}) \rightarrow \infty$$

- A unit MA root induces noninvertibility,

$$g_x(\omega\ \text{assoc. with U.R.}) = 0$$

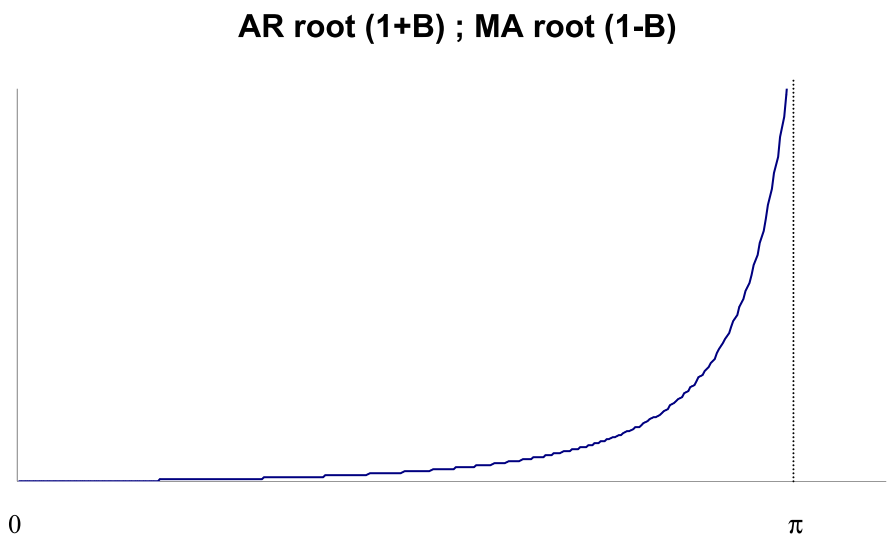
Model:

$$(1-B)\ x_t = (1+B)\,a_t$$



Model:

$$(1+B)\ x_t = (1-B)\,a_t$$



TWO USEFUL ALTERNATIVE REPRESENTATIONS:

ARMA model:

$$\phi(B)\,a_t = \theta(B)\,a_t$$

(a) $$x_t = \frac{\theta(B)}{\phi(B)}\,a_t = (1 + \psi_1 B + \psi_2 B^2 + \dots\ )\,a_t$$

or

$$\boxed{x_t = \psi(B)\,a_t}$$

↑ "**Psi**" - weights (ψ - weights)

* If x_t is stationary:

ψ - weights $\rightarrow 0$

ARMA could be approximated by finite MA

(b) Alternatively,

$$\frac{\phi(B)}{\theta(B)} x_t = a_t \text{ , or}$$

$$(1 + \pi_1 B + \pi_2 B^2 + \dots) x_t = a_t \text{ ,}$$

and in compact form,

$$\boxed{\pi(B)\, x_t = a_t}$$

↑ "**Pi**" - weights (π - weights)

* If x_t is invertible,

π - weights $\rightarrow 0$

ARMA could be approximated by finite AR

For seasonal data, often the general multiplicative ARIMA model is used:

$$\phi(B)\ \Phi(B^s)\ \nabla^d\ \nabla_s^D\ x_t = \theta(B)\ \Theta\ (B^s)\ a_t$$

$\phi(B) \equiv$ regular AR pol.

$\Phi\ (B^s) \equiv$ seasonal AR pol. (in B^s)

$\theta(B) \equiv$ regular MA pol.

$\Theta(B^s) \equiv$ seasonal MA pol. (in B^s)

where: $\phi(B)$ and $\Phi(B^s)$ are stationary.

$\theta(B)$ and $\Theta\ (B^s)$ are invertible.

Two important points:

- PARSIMONY (few parameters);

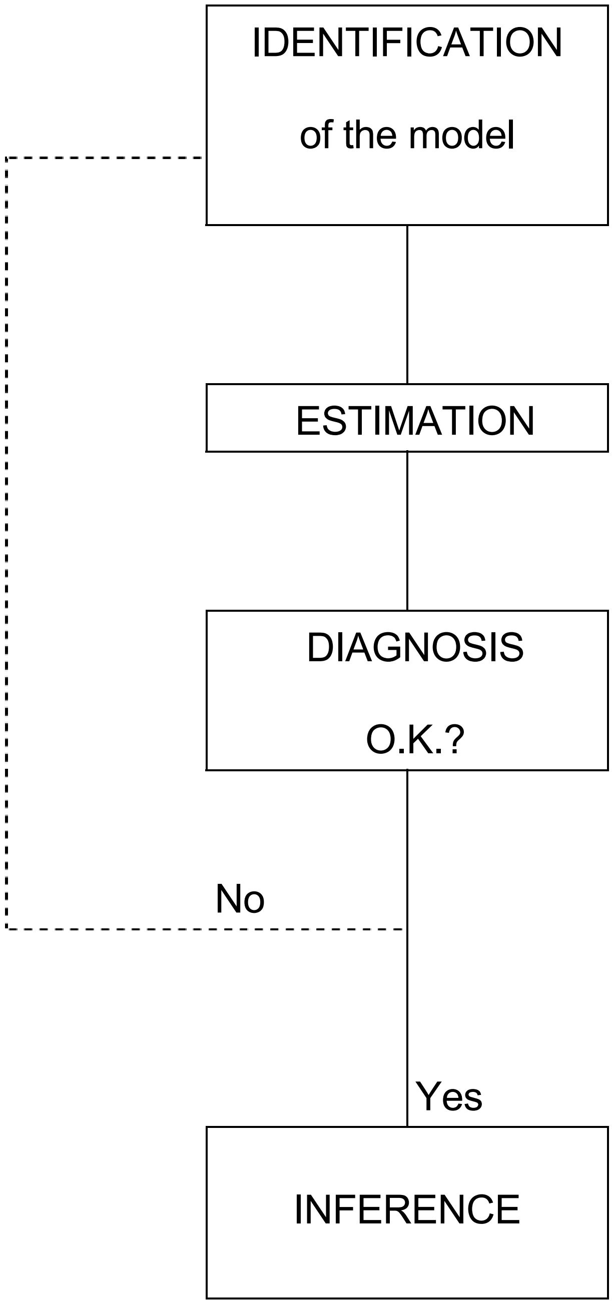
- short-term use (max. horizon: 1 – 2 years)

(it may affect the order of differencing: roughly,

* short-term use favors differencing

* long-term use does not.)

"Box-Jenkins"-type approach:



IDENTIFICATION OF THE ARIMA MODEL

One has to determine:

a) Degrees of differencing

b) Orders p and q of ARMA

a) Traditional criterion: "fast enough convergence of ACF". As we shall see, unit roots are easily detected through estimation (Tiao, Tsay)

b) Main idea: to "match" ACF of some known ARMA.

Basic traits of | ACF of ARMA (p, q) | :

$$(1 + \phi_1 B + \ldots + \phi_p B^p)\, x_t = (1 + \theta_1 B + \ldots + \theta_q B^q)\, a_t$$

ACF:

$\rho_k \equiv$ Lag-k autocorrelation of x_t, as a function of k:

It will display

* q starting conditions,

after which the AR difference equation

* $\rho_k + \phi_1 \rho_{k-1} + \ldots + \phi_p \rho_{k-p} = 0$

holds. Hence, for $k > q$, ρ_k is solution of

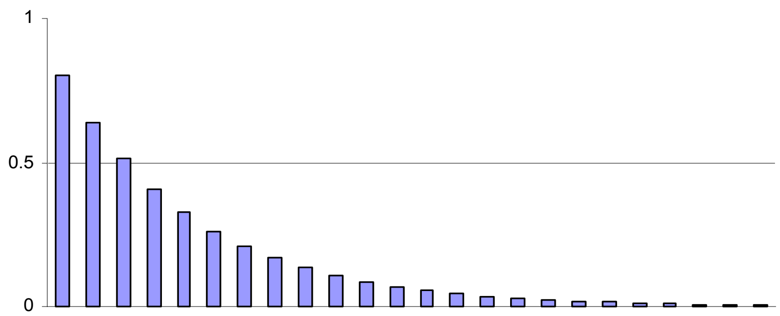
$$\boxed{\phi(B)\, \rho_k = 0}$$

where B operates on k.

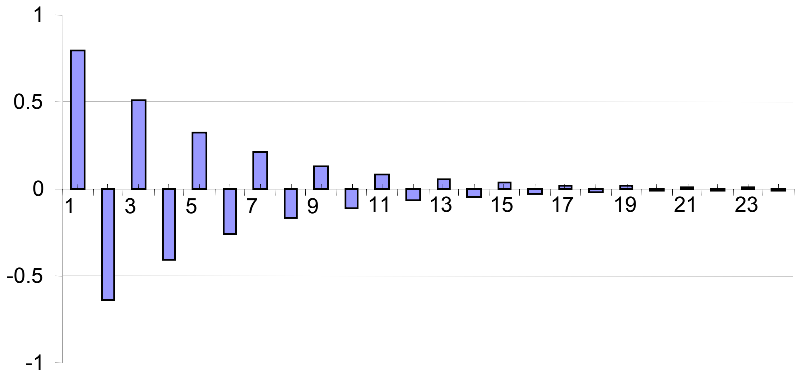
Thus: "Eventual ACF"

$$\rho_k = \sum [(\text{pol. in } t) + \text{cosine func.}]$$

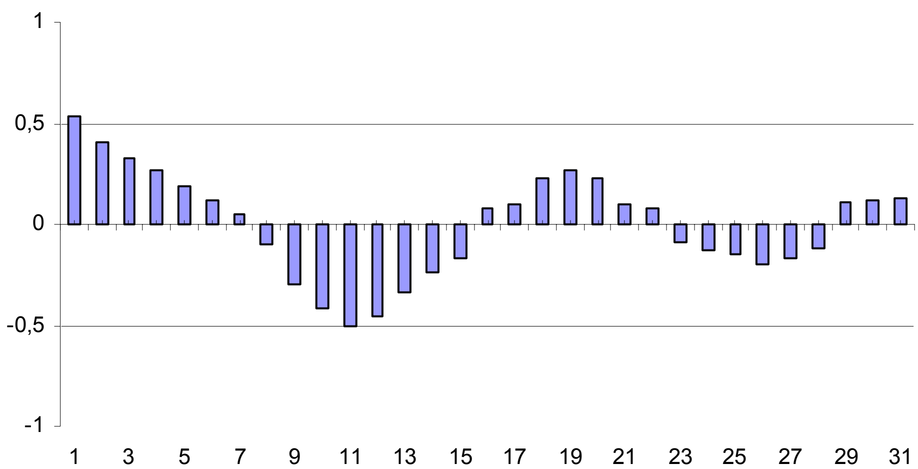
ACF (1-.8B)xt=at



ACF (1+.8B)xt=at



ACF AR(2) complex roots



At present, "identification" uses more efficient procedures

(as will be seen in TRAMO)

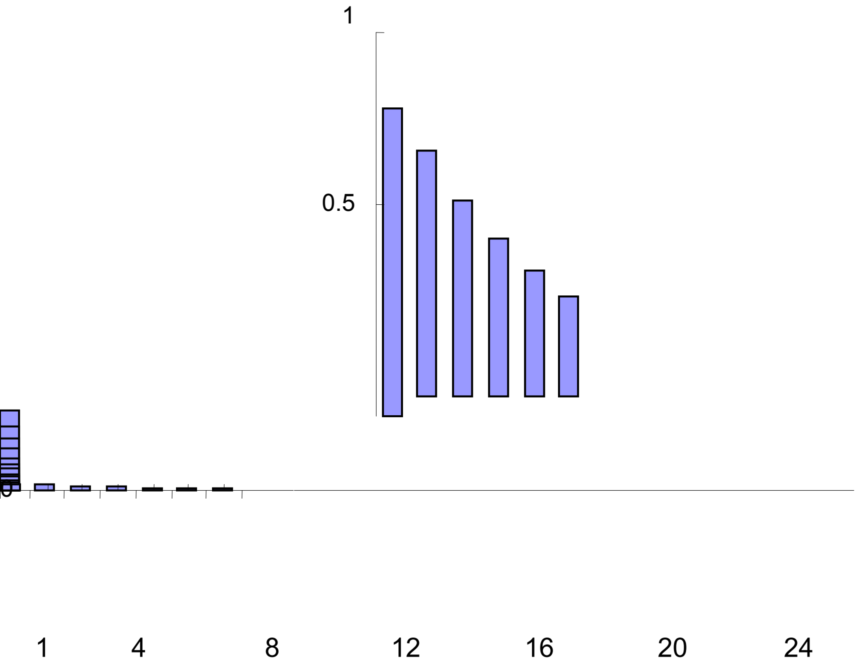
Notes:

a) Often, more than one model may seem reasonable,
hence always some room for the analyst experience or purpose.

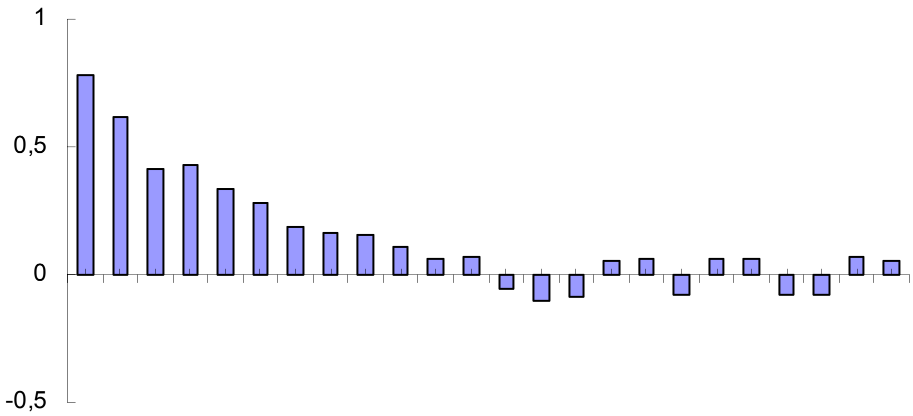
b) In practice we do not know the ACF and autocorrelations have to be estimated.

Estimation can induce large (spurious) covariances that have a distorting effect on the sample ACF, which may fail to damp out according to expectations.

THEORETICAL ACF



SAMPLE ACF



PARAMETER ESTIMATION (more later)

Rough intuition : $[\, x_1 \dots x_T \,]$

Assume:

$$x_t = a_t + \theta a_{t-1} \qquad \text{MA (1),} \quad |\theta| < 1$$

or

$$a_t = x_t - \theta a_{t-1}$$

Conditional on a_0: Set $\theta = \theta_0$

Compute sequentially:

$$a_1^0 = x_1 - \theta_0 a_0$$

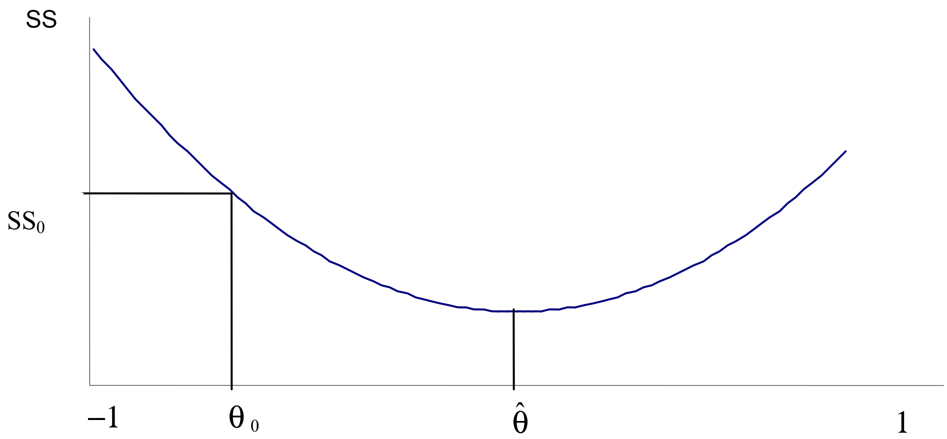
$$a_2^0 = x_2 - \theta_0 a_1^0$$

………………..

$$a_T^0 = x_T - \theta_0 a_{T-1}^0$$

$$SS_0 = \sum_1^T \left(a_j^0\right)^2$$

Varying θ:



$\hat{\theta}$ = min. SS

Notice:

$a_t = x_t - \theta a_{t-1}$

$a_{t-1} = x_{t-1} - \theta a_{t-2}$

$a_{t-2} = x_{t-2} - \theta a_{t-3}$

…..

yields

$$a_T = x_T - \theta x_{T-1} + \theta^2 x_{T-2} - \ldots + (-\theta)^{T-1} x_1 + (-\theta)^T x_0$$

1) HIGHLY NON LINEAR FUNCTION OF PARAMETERS ($\therefore$ NL optimization)

2) Effect of starting conditions
IMPORTANCE OF INVERTIBILITY $(|\theta|<1)$

DIAGNOSTICS (more later)

* In-sample
* out-of-sample

Mostly : residual-based diagnostics.

$\hat{a}_t$ ~ Niid (0, V_a)

INFERENCE (more later)

Example : **Forecasting** (parameters known)

Notation: $\hat{x}_{t+k|t} \equiv$ Forecast of x_{t+k} made in period t.

$$x_t = \hat{x}_{t|t-1} + a_t$$

$$a_t \equiv \text{1-period-ahead forecast error}$$

$$= x_t - \hat{x}_{t|t-1} \quad (\equiv \text{innovation})$$

These a_t 's are the ones of the Wold representation, and of the ARIMA model.

Forecast:

$\hat{x}_{t+k|t} = \text{MMSE}_t(x_{t+k}) =$ (under our assumptions) =

$= E(x_{t+k} | x_1 \ldots x_t)$

Computation with Kalman filter (later).

<u>Forecast function</u> :

$\hat{x}_{t+k|t}$ as a function of k.

For ARMA (p, q) :

Forecast function:

* q starting conditions,

after which the AR difference equation

$$\hat{x}_{t+k|t} + \phi_1 \hat{x}_{t+k-1|t} + \ldots + \phi_p \hat{x}_{t+k-p|t} = 0$$

holds. Hence $\hat{x}_{t+k|t}$ is solution of

$$\boxed{\phi(B)\hat{x}_{t+k|t} = 0}$$

where B operates on k.

Note:

Eventual Forecast Function and ACF are solution of the same AR finite difference equation

(by looking at the correlation between present and past, we know the correlation between present and future...)

USEFUL WAY TO LOOK AT THE FORECAST:

Use ψ - weights:

$$x_{t+k} = a_{t+k} + \Psi_1 a_{t+k-1} + \dots$$

$$+ \Psi_{k-1} a_{t+1} + \Psi_k a_t + \Psi_{k+1} \ a_{t-1} + \dots \quad ;$$

since

$E_t \, a_{t+k} = 0 \quad (k > 0)$: future forecast errors are unknown,

$E_t \, a_{t+k} = a_{t+k} \ (k < 0)$: past forecast errors are known,

$$\hat{x}_{t+k|t} = E_t \, x_{t+k} = \psi_k a_t + \psi_{k+1} \ a_{t-1} + \dots \quad ;$$

a linear combination of past and present innovations

Thus, <u>FORECAST ERROR</u>

$$e_{t+k|t} = x_{t+k} - \hat{x}_{t+k|t}$$

$$= \underbrace{a_{t+k} + \psi_1 a_{t+k-1} + \ldots + \psi_{k-1} a_{t+1}}$$

MA (k-1) of " future " innovations.

From this, distributions are easily derived.

Example: Simplest one

$k = 1$

$e_{t+1|t} \sim N(0, V_a)$,

or, for a vector of 1-period-ahead forecast errors,

$e_{t+1|t} \sim N(0, V_a I) \ldots$