October 2008

NOTES ON PROGRAMS

TRAMO AND SEATS©

SEATS PART

Signal Extracion in ARIMA Times Series

Agustín Maravall

Bank of Spain

In the remote past, unobserved components where estimated using

Deterministic Models

$x_t = p_t + s_t + u_t$

For example:

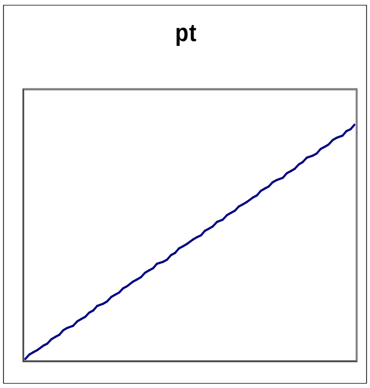
$p_t = a + bt$ linear trend

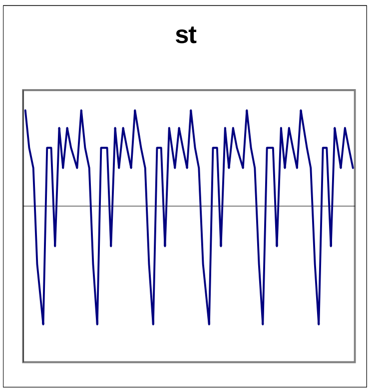
$s_t = \Sigma\beta_i d_{it}$ dummies (d_{1t} = 1 for January, 0 otherwise;…)

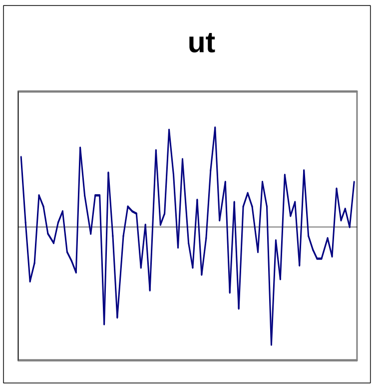
or, equivalenty, sine-cos functions

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

u_t = white noise: niid (0, V_u)







– concept of deterministic: if we know the "true" parameters of the model, the variable can be forecast with no error (ex.: p_t or s_t above)

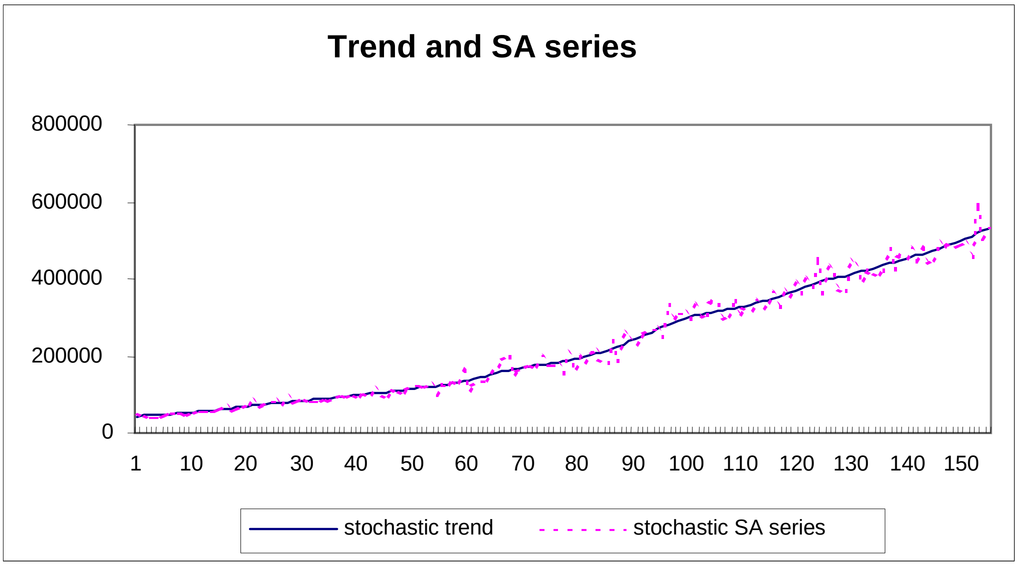
– concept of **white noise**: $[a_t]_1^T$ is w.n. iff $(a_1 \ldots a_T) \sim$ niid (0, Va)

MOST COMMON OBJECTIVE OF SEASONAL ADJUSTMENT:

Better understanding of underlying (still, short-term) evolution of the series.

In so far as highly transitory noise can also distort the picture, it is often helpful to look at:

Trend-Cycle estimation (i.e. the SHORT-TERM TREND)



"Short term" analysis $\equiv$ **at the most** two-year horizon

When main interest is to remove:

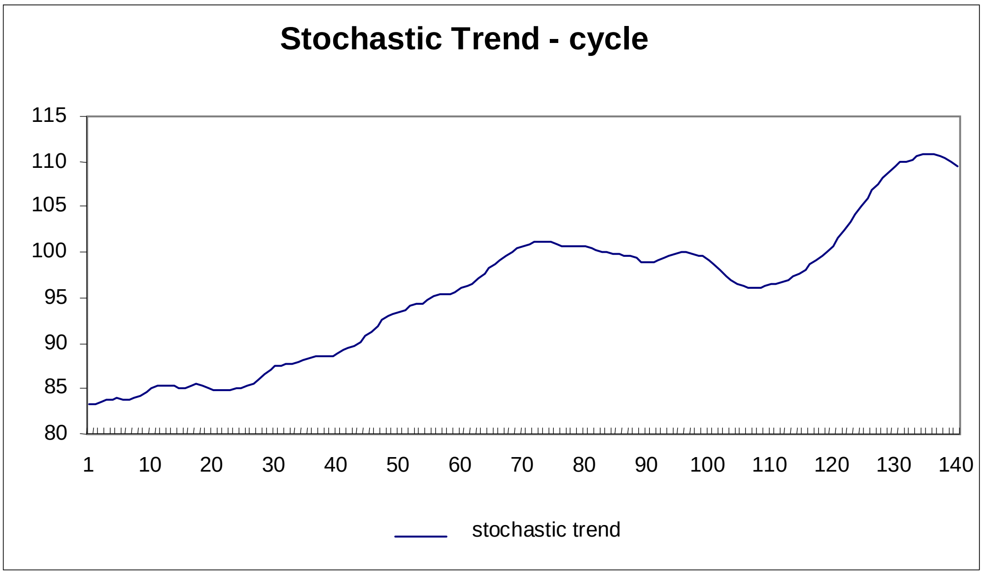
– Seasonal noise,

– (Highly transitory) irregular noise,

so as to read data better in short-term policy, the remaining signal may well contain variation for cyclical frequencies.

In this case,

trend $\rightarrow$ "trend-cycle"



Gradual realization that seasonality evolves in time ("moving seasonality")

[An obvious and basic example: the weather,

one of the main causes of seasonality]

↓

MOVING AVERAGE METHODS

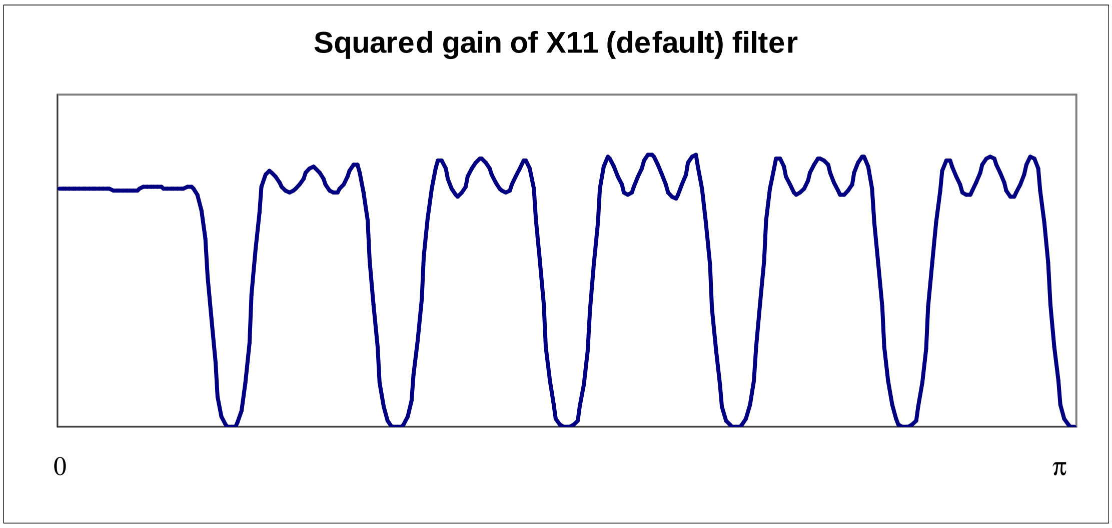
1) Fixed ("band-pass") filters

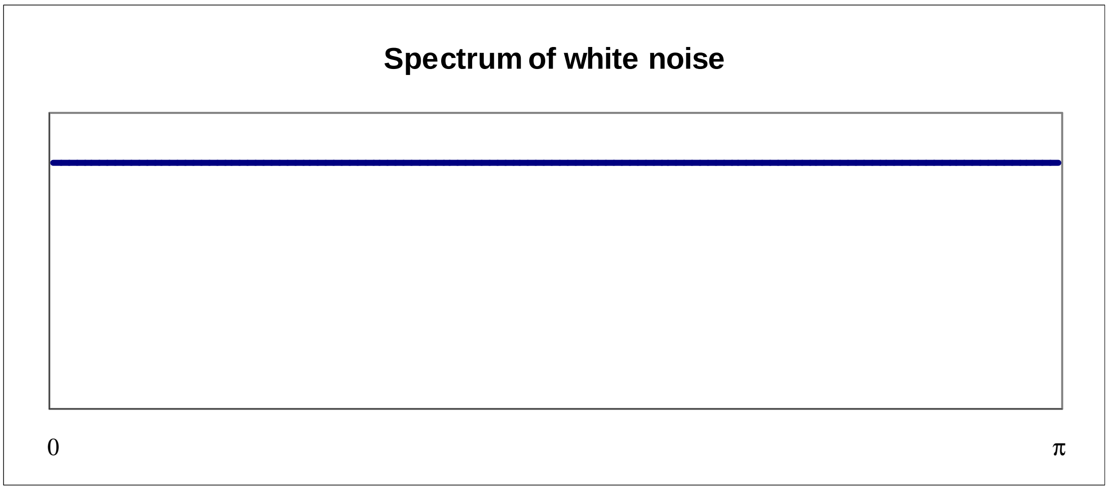
Some limitations:

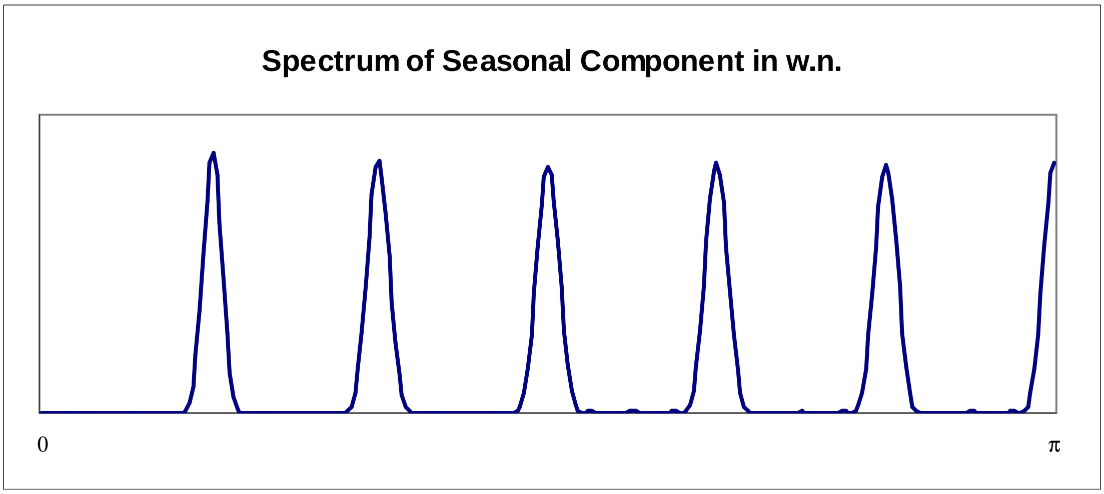
* Spurious results
* Can overadjust

  Can underadjust

....


Squared gain of X11 (default) filter
0
π


Spectrum of white noise
0
π


Spectrum of Seasonal Component in w.n.
0
π

2) An alternative approach:

Use simple stochastic models to capture structure of series. (ARIMA models)

Derive optimal filter

(Signal Extraction)

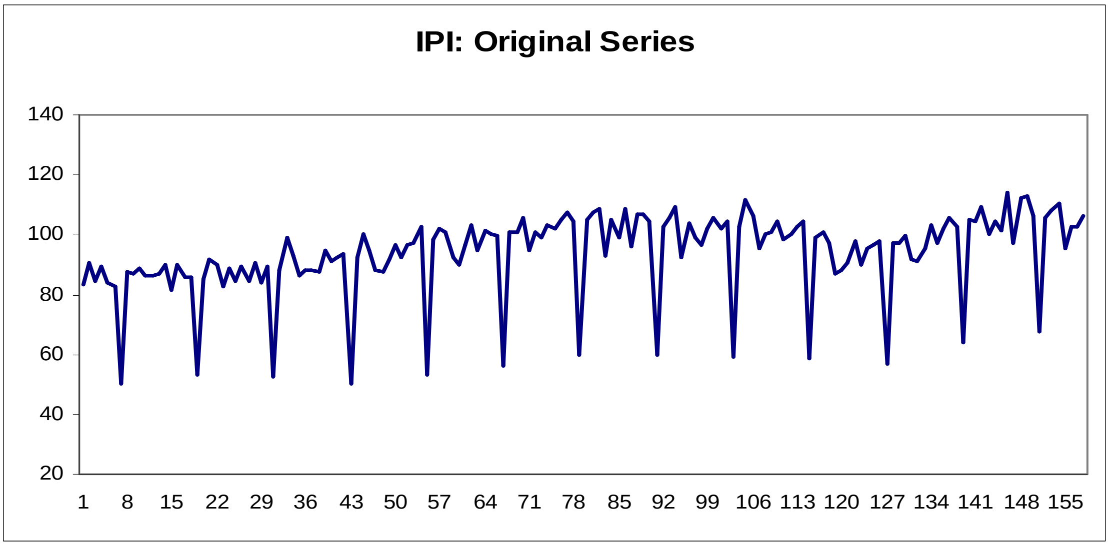
THIS IS OUR APPROACH

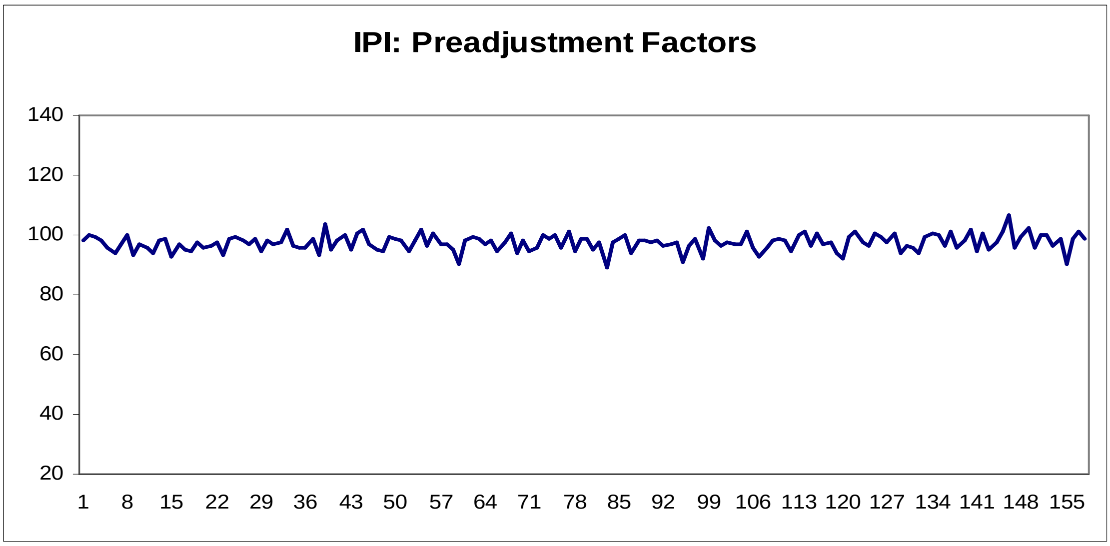
The method permits us to jointly solve many problems of applied interest.

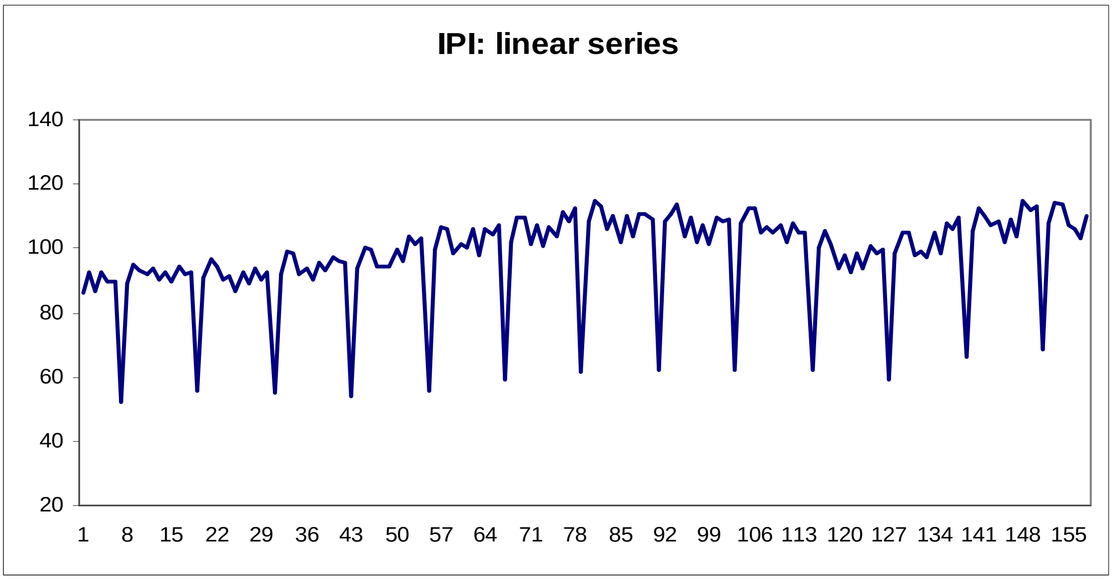
In the most general case:

A series contaminated by outliers, affected by regression variables, subject to deterministic effects (TD, EE, Intervention variable, ...) has been cleaned by TRAMO ("preadjustment").

Then the preadjusted or "linearized" series (the output of the ARIMA model) is decomposed into components by SEATS.


IPI: Original Series
140
120
100
80
60
40
20
1 8 15 22 29 36 43 50 57 64 71 78 85 92 99 106 113 120 127 134 141 148 155


IPI: Preadjustment Factors
140
120
100
80
60
40
20
1 8 15 22 29 36 43 50 57 64 71 78 85 92 99 106 113 120 127 134 141 148 155


IPI: linear series
140
120
100
80
60
40
20
1 8 15 22 29 36 43 50 57 64 71 78 85 92 99 106 113 120 127 134 141 148 155

Use of TRAMO as a PREADJUSTMENT program

Observed series

TRAMO

↓

(Interpolated series)

↓

Stochastic Part:	Deterministic Part:
"Linearized" Series (output of ARIMA)	Regression Effects

↓

SEATS

↓

Trend
+
Seasonal
+
(Transitory)
+
Irregular

AO (Outlier)
TC (Outlier)
LS (Outlier)

Trading Day/Leap Year
Easter Effect
Holidays
Intervention Variables
Regression Variables
(can be assigned to any component)

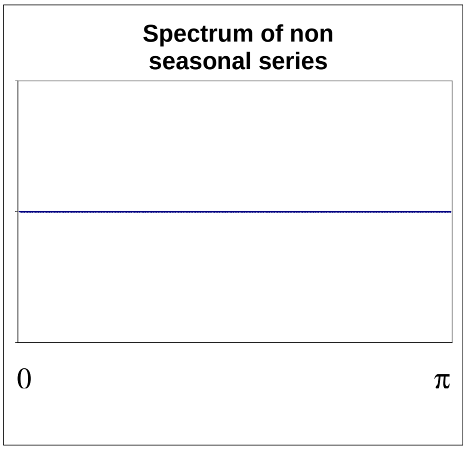
↓ ↓

Final components

BASIC IDEA BEHIND THE FILTERS IN SEATS

a) NONSEASONAL SERIES

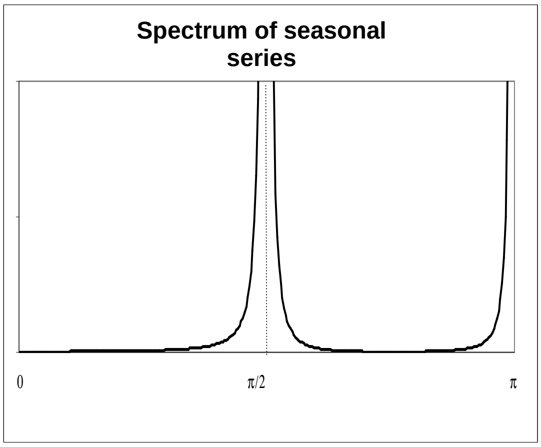
x_t = white noise



filter for SA series:
$\upsilon(B, F) = 1$

b) PURELY SEASONAL SERIES

$S\, x_t = w_t$ (stationary MA)



filter for SA series:
$\upsilon(B, F) = 0$

Conclusion:

SERIES WITH DIFFERENT STOCHASTIC STRUCTURES REQUIRE DIFFERENT FILTERS

What SEATS does:
To taylor the filter to the structure of the series
(in some optimal way)

The decomposition can be

multiplicative:

$$X_t = \text{trend\_cycle x seasonal (x transitory) x irregular}$$

in which case:

- trend-cycle gives level
- others expressed as factors
  (in this case, usually multiplied by 100.
  Thus s_t = 103.7 implies that the seasonal effect for month t is an increase of 3.7 percent for that month series' value.)

or

additive:

$$x_t = \text{trend\_cycle} + \text{seasonal} \, (+ \text{transitory}) + \text{irregular}.$$

Since

$$x_t = \log X_t$$

makes

multiplicative $\rightarrow$ additive (in logs),

we discuss additive decomposition.

The decomposition is of the type:

$$z_t = p_t + s_t + (c_t) + u_t$$

p_t : trend

s_t : seasonal

c_t : transitory

u_t : irregular

or

$$z_t = n_t + s_t$$

n_t : seasonally adjusted series

$$n_t = p_t + (c_t) + u_t$$

<u>Assumption</u>:

COMPONENTS ARE ORTHOGONAL

(what causes seasonal fluctuations -weather, holidays, ...- has little to do with what causes the long-term evolution of the series -productivity, technology, ...)

SEATS allows for the sum of the components to respect the stochastic structure of the observed series.
This stochastic structure is captured with an ARIMA model.
Given the ARIMA model for the observed data:

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

$\phi(B)$: " Full " AR polynomial (includes unit roots)

or:

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

SEATS decomposes z_t in the following manner:

1) Factorize the AR polynomial $\phi(B)$ as in:

$$\phi(B) = \phi_p(B) \times \phi_s(B) \times \phi_c(B)$$

where:

$\phi_p(B)$: trend roots

$\phi_s(B)$: seasonal roots

$\phi_c(B)$: "transitory" roots

(roots are assigned according to their associated frequency)

Assumption: Two different components cannot share the same AR root.

Strictly speaking, the assumption is only needed for UNIT AR roots.
But it simplifies exposition.

2) Express z_t as:

$$\frac{\theta(B)}{\phi(B)} a_t = \frac{\theta_p(B)}{\phi_p(B)} a_{pt} + \frac{\theta_s(B)}{\phi_s(B)} a_{st} + \frac{\theta_c(B)}{\phi_c(B)} a_{ct} + u_t ,$$

with u_t white noise.

Hence, model for components are:

$$\phi_p(B)\, p_t = \theta_p(B)\, a_{pt}$$

$$\phi_s(B)\, s_t = \theta_s(B)\, a_{st}$$

$$\phi_c(B)\, c_t = \theta_c(B)\, a_{ct}$$

$u_t =$ white noise

If the spectra of all components are nonnegative,

the decomposition is called

ADMISSIBLE

Example:
Let the model be

$$(1-.4B)\,\nabla\nabla_{12}\,x_t = \theta(B)\,a_t \quad .$$

Then, $\phi(B) = (1-.4B)\,\nabla\nabla_{12} =$

$$= (1-.4B)\,\nabla^2\,S$$

We know that
* $(1-.4B)$ generates stationary (highly transitory) behavior,
* ∇ (and ∇^2) generates trends,
* S generates seasonality.

Thus the allocation of roots will be

$\phi_p(B) = \nabla^2$

$\phi_s(B) = S$

$\phi_c(B) = (1-.4B)$,

and the series x_t is decomposed as in the "Stochastic Partial Fraction Expansion"

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

$$= \frac{\theta_p(B)}{\nabla^2} a_{pt} + \frac{\theta_s(B)}{S} a_{st} + \frac{\theta_c(B)}{1-.4B} a_{ct} + u_t$$

$$= \quad p_t \quad + \quad s_t \quad + \quad c_t \quad + u_t$$

where

$$\nabla^2 p_t = \theta_p(B)\, a_{pt} \quad ,$$

$$S\ s_t = \theta_s(B)\, a_{st} \quad ,$$

$$(1-.4B)\ c_t = \theta_c(B)\, a_{ct} \quad ,$$

and u_t is white noise. All components are mutually orthogonal.

Notice that components also follow ARIMA-type models and can be interpreted.

SEASONAL COMPONENT

For a deterministic seasonal component, the sum over a year period of the component should be zero,

$$s_t + s_{t-1} + \ldots + s_{t-11} = 0 \qquad \text{(monthly data)}$$

or

$$(1 + B + B^2 + \ldots + B^{11})\, s_t = 0 \quad .$$

In short, if

$$S = 1 + B + B^2 + \ldots + B^{11},$$

$$S\, s_t = 0$$

For "moving" or stochastic seasonality, this condition cannot hold for every t. (Precisely because component is moving.)

But, in any case, the annual sum of s_t should, on average, be zero, and should not depart too much from it.

Thus we may say

$$S\, s_t = a_{st} \quad ,$$

where a_{st} is w.n. , with

$$E\, a_{st} = 0 \ ,$$

$$Var\,(a_{st}) = V_s \ \text{relatively small;}$$

this yields a stochastic component (Harvey-Todd, 1983).

More generally,

for the seasonal component, often:

$$\phi_s(B) = S = 1 + B + B^2 + \ldots + B^{s-1}$$

where s = # of observations / year.

Hence, a model for the seasonal of the type

$$S\, s_t = w_t \;,$$

where w_t is a stationary ARMA model with:

* zero mean
* small variance,

implies

"annual aggregation of the seasonal component will on average be zero, and will not depart too much from it".

A Comment on Stationary Seasonal AR Roots

Assume the ARIMA model for the observed series contains the seasonal AR factor

$$(1 + \phi_s B^s).$$

*	When $|\phi_s| < k$,

k = a preassigned (moderate) value (in SEATS: parameter RMOD = .5 by default),

then factor is assigned to the transitory component.

(A small correlation whose effect disappears, in practice, after one or two years cannot be properly called “seasonality”.)

*	When $\phi_s > k$, (very rarely encountered)

the factor $(1+\phi_s B^s)$ is associated with a stationary 2-year period. It is thus assigned to the transitory component.

*	When $\phi_s < -k$,

the following identity is used

[similar to $1-B^s = (1-B)\,(1+B+\ldots+B^{s-1})$] .

Let ϕ_s denote now $(-\phi_s)$. Then,

$$1-\phi_s B^s = (1-\varphi B)\,(1+\varphi B+\varphi^2 B^2+\ldots+\varphi^{s-1}B^{s-1}),$$

where

$$\varphi = [\phi_s]^{1/s}$$

(Ex.: $\phi_4 = .7 \quad \to \quad \varphi = .915$

$\phi_{12} = .7 \quad \to \quad \varphi = .987$) .

Then,

- the root $(1-\varphi B)$ is assigned to the trend-cycle component.
- the roots of the polynomial

$$\varphi_s(B) = 1 + \varphi B + \varphi^2 B^2 + \ldots + \varphi^{s-1} B^{s-1}$$

are assigned to the seasonal component.

Thus the model for the seasonal component will in general be of the type

$$\varphi_s(B)\ S^{d_s}\ s_t = \theta_s(B)\ a_{st} \quad ,$$

(most often with $\varphi = 0$ and $d_s = 1$), with a_{st} a zero mean, small variance w.n. The model will be balanced (i.e.: total AR order = total MA order).

TREND

Analogously, we may start with a deterministic trend, say

$$p_t = a + bt$$

We know that

$$\nabla p_t = b,$$

or

$$\nabla^2 p_t = 0$$

We cannot expect a "moving" trend to exactly satisfy the above conditions at every t. Instead, we require that departures from those conditions should, on average, cancel out, and that they should not be too large.

This yields as a possible specification:

$$\nabla p_t = b + a_{pt}$$

with

$$E\, a_{pt} = 0$$

$$Var\,(a_{pt}) = V_p \qquad \text{relatively small}$$

This stochastic trend specification is the well-known

"random walk + drift"

model.

Alternatively, we could use as stochastic specification

$$\nabla^2 p_t = a_{pt}$$

with

$E\ a_{pt} = 0$

$Var\ (a_{pt}) = V_p$ relatively small

This is the so-called

"second-order random walk"

model.

Notice that the 2 stochastic models are different:

$\nabla p_t = b + a_{pt}$ implies a random shock in the slope of the trend

$\nabla^2 p_t = a_{pt}$ implies a random shock in the change of the slope of the trend

More generally, the specification of the stochastic trend will be of the type

$$\nabla^{d_p} p_t = w_t$$

where w_t is a

-zero mean
-stationary

ARMA process.

The model for the trend component can be expressed, in general, as

$$\phi_p(B) \ \nabla^{d_p} \ p_t = \theta_p(B) \ a_{pt} \quad ,$$

with (Maravall, 1993)

- ϕ_p (B) stationary (for example, (1--.8B)),
- d = (0), 1, 2, (3),
- θ_p(B) of low order,
- Var (a_{pt}) = a small fraction of V_a.

The model will also be balanced.

In essence: SEATS finds admissible models for the components

$$\phi_p(B)\,p_t = \theta_p(B)\,a_{pt}$$
$$\phi_s(B)\,s_t = \theta_s(B)\,a_{st}$$
$$\phi_c(B)\,c_t = \theta_c(B)\,a_{ct}$$
$$u_t = \text{w.n.}$$

such that

$$x_t = p_t + s_t + c_t + u_t$$

(Sum of component models $\equiv$ ARIMA for observed series)

At time $t = T$, SEATS PROVIDES

- FOR $t = 1, \ldots, T,$ ⁞ *T+1,…T+FH*

(FH = Forecast Horizon)

the decomposition:

$$x_t = \hat{p}_{t|T} + \hat{s}_{t|T} + \hat{c}_{t|T} + \hat{u}_{t|T}$$

(when $t > T \rightarrow x_t$ is replaced by its ARIMA forecast $\rightarrow \hat{x}_{t|T}$)
where (for ex.)

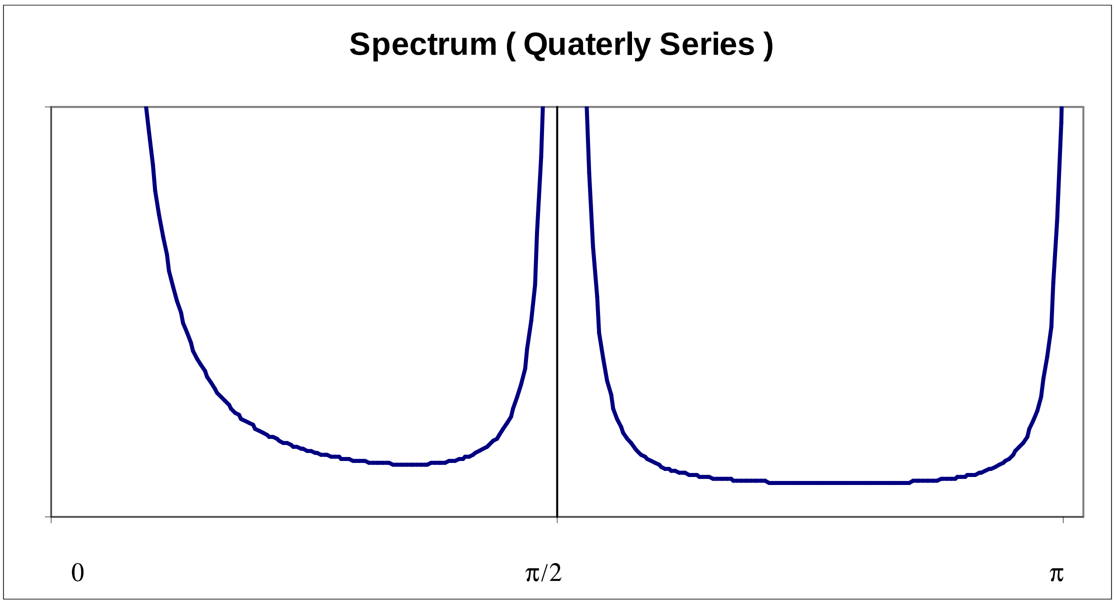
$$\hat{s}_{t|T} = \text{MMSE estimator} = E(s_t \mid x_1 \ldots x_T).$$

SEATS also provides standard error of estimators and forecasts.

ALLOCATION OF AR ROOTS

Ex.: Quartely data

Pseudospectrum:



Trend Roots

Unit AR roots at $\omega = 0$ (i.e., root B=1 in AR polynomial $\rightarrow \quad \phi(1) = 0$).

Also:

Stationary roots for $\omega = 0$ if large enough modulus.

Ex.: (1 - .8B) in AR polynomial.

"Large enough" = above the value of parameter RMOD

Seasonal roots

Seasonal frequencies:

$$\frac{\pi}{2} \;,\; \pi \quad \text{(once-and twice-a year frequencies)}$$

Roots at

$$\omega \in \left[\; \frac{\pi}{2} \pm \varepsilon \; \right]$$

$$\omega \in [\; \pi , \pi - \varepsilon \;]$$

will be treated as seasonal
(ε: controlled by EPSPHI)

Transitory

* AR factors of the type (1 - .4B) or (1 + .4B) (i.e. roots for $\omega = 0$ or $\omega = \pi$ with small moduli, as determined by RMOD)

* AR roots for $\omega \in \left[0 + \varepsilon , \frac{\pi}{2} - \varepsilon \right]$ (range of "cyclical frequency")

  (i.e. between trend and first harmonic)

* AR roots for "intraseasonal" frequencies

* when Q > P : In this case, the SEATS decomposition yields a pure MA (Q - P) component (hence transitory).

Notice that, when Q > P, a transitory component will appear even when there is no AR factor allocated to it.

Irregular

Always white noise

(Convenient for testing)

The TRANSITORY COMPONENT is always stationary, and hence its effect is, by construction, transitory.

It will typically capture short-lived, fairly erratic behavior that is not white noise, sometimes associated with ackward frequencies.

Its separate presence is justified by two considerations:

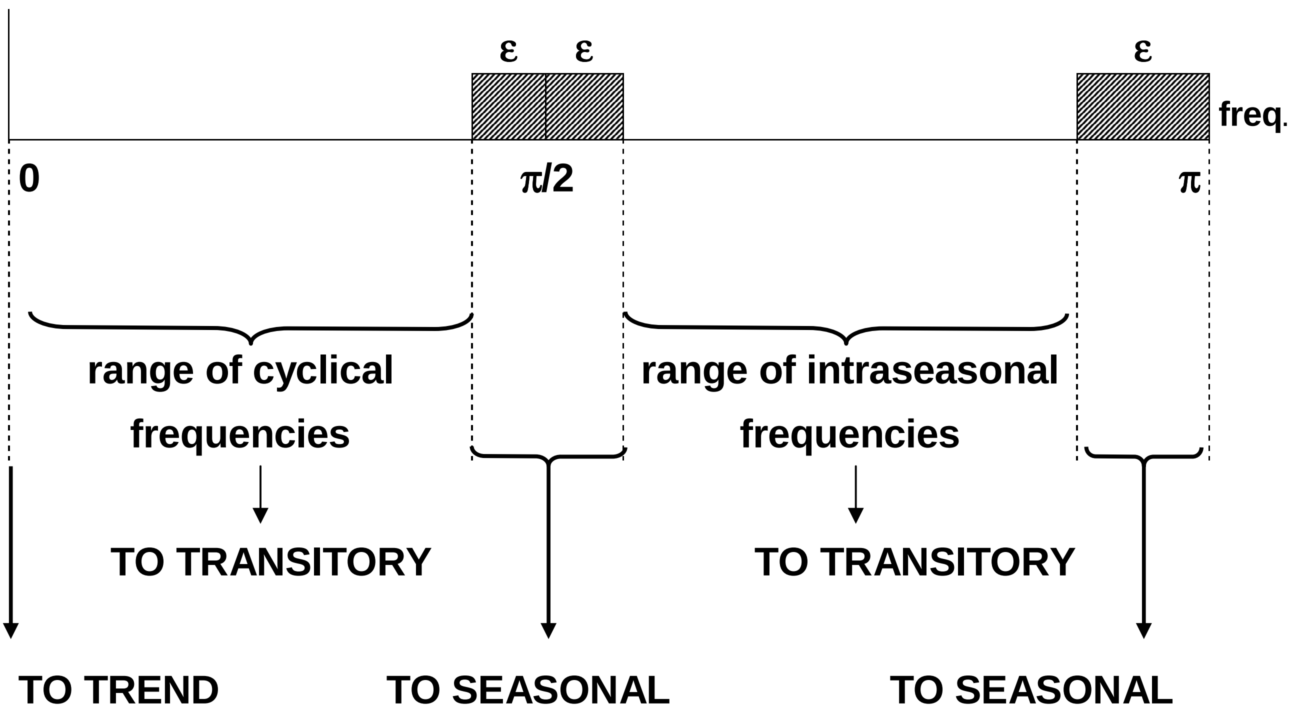
a) The variation it contains should not contaminate the trend or seasonal components. Its removal permits to obtain smoother, more stable trends or seasonals.

b) From the testing and diagnostics point of view, it is desirable to preserve a purely white-noise irregular, computed as a residual.

However, in the final decomposition, it may be convenient to combine the transitory and irregular components into a single "transitory-irregular" component.

ALLOCATION OF AR ROOTS

Example: Quarterly data



An example:

Model for monthly observed series:

$$(1-.78B-.624B^2+.512B^3)\nabla\nabla_{12}x_t=\theta(B)a_t$$

Regular AR polynomial factorizes as:

$$1-.78B-.624B^2+.512B^3=(1+.8B)(1-1.58B+.64B^2)$$

Root of $(1+.8B) \Rightarrow \omega=\pi \quad \Rightarrow \quad \tau=2\pi/\omega=2$ months
hence seasonal root
(6 times-a-year frequency)

Roots of $(1-1.58B+.64B^2) \Rightarrow$
complex root with modulus $r=\sqrt{\phi_2}=.8$;

frequency ω (in radians) = $\arccos(\phi_1/2r)=$

$=.16$ rads;

$$\text{period}=\frac{2\pi}{\omega}=40 \text{ months}.$$

Thus complex root is associated with a 3 1/3 year stationary cycle

$\Rightarrow$ to transitory component

Roots of $\nabla = 1 - B \Rightarrow$ trend

Roots of $\nabla_{12} = 1 - B^{12} =$

$$= (1 - B)(1 + B + \ldots + B^{11}) =$$
$$= \nabla S$$

$\nabla \Rightarrow$ trend

$S \Rightarrow$ seasonal

* Grouping the roots, the series would be decomposed into:

- trend: $\nabla^2 p_t = \theta_p(B) a_{pt}$

- seasonal: $(1 + .8B) S s_t = \theta_s(B) a_{st}$

- transitory: $(1 - 1.58B + .64B^2) c_t = \theta_c(B) a_{ct}$

- irregular: $u_t = w.n.$

* The AR polynomials of the models for the components are determined.

DECOMPOSITION FOR THE "DEFAULT" (AIRLINE) MODEL

$$\nabla\nabla_{12}\, x_t = (1+\theta_1 B)(1+\theta_{12} B)\, a_t$$

models for the components are of the type: $(\nabla\nabla_{12} = \nabla^2 S)$

TRENDCYCLE

$$\nabla^2 p_t = \theta_p(B)\; a_{pt} \quad , \qquad \text{(A)}$$

with order $[\theta_p(B)] = 2$.

SEASONAL

$$(1+B+\ldots+B^{11})\, s_t = \theta_s(B)\, a_{st} \qquad \text{(B)}$$

with order $[\theta_s(B)] = s - 1$

(there is no transitory component)

IRREGULAR

$$u_t \sim \text{ white noise} \qquad \text{(C)}$$

SOME EXAMPLES OF MODEL SPECIFICATION (Monthly series)

A : Basic Structural Model (Harvey-Todd , 1983); ARIMA specifications.

B : ARIMA-Model-Based decomposition of Airline model (Default model TRAMO-SEATS)

C : ARIMA-Model-Based interpretation of X11 (Cleveland, 1975)

	A	B	C
Trend Component	$\nabla^2 p_t = (1+\alpha B) a_{pt}$	$\nabla^2 p_t = (1+\alpha B)(1+B) a_{pt}$	$\nabla^2 p_t = (1+.26B+.30B^2-.32B^3) a_{pt}$
Seasonal Component	$S(B) s_t = a_{st}$	$S(B) s_t = \theta_s(B) a_{st}$ $\theta_s(B)$ of order 11	$S(B) s_t = (1+.26B^{12}) a_{st}$
Irregular Component	w.n.	w.n.	w.n.
Overall Series	$\nabla\nabla_{12} x_t = \theta(B) a_t$ $\theta(B)$ of order 13; 3 parameters	$\nabla\nabla_{12} x_t = (1+\theta_1 B)(1+\theta_{12} B^{12}) a_t$ $\theta(B)$ of order 13; 2 parameters	$\nabla\nabla_{12} x_t = \theta(B) a_t$ $\theta(B)$ of order 14; 0 parameters

See Maravall, 1985.

Model for SEASONALLY ADJUSTED SERIES can be obtained by aggregation

$$n_t \quad = \quad p_t \quad + \quad c_t \quad + \quad u_t \ ,$$

For ex., for default model, since

$$\text{IMA}(2,2) + \text{w.n.} \rightarrow \text{IMA}(2,2) \quad ,$$

$n_t \sim$ IMA (2,2) , say

$$\nabla^2 n_t = \theta_2 (B)\, a_{nt} \quad .$$

Typically one obtains:

$$\theta_2 (B) \approx (1 - .9\, B)\, (1 + \alpha\, B) \quad ,$$

with α of moderate size.

If (1 - .9B) cancels one ∇, the model becomes

$$\nabla\, n_t = (1 + \alpha\, B)\, a_{nt} + k \quad ,$$

with α small.

Hence model for SA series often is not far from the popular

"random walk + drift" model.

Remark:

Also we could aggregate the transitory and the irregular to yield a stationary (transitory- irregular) component

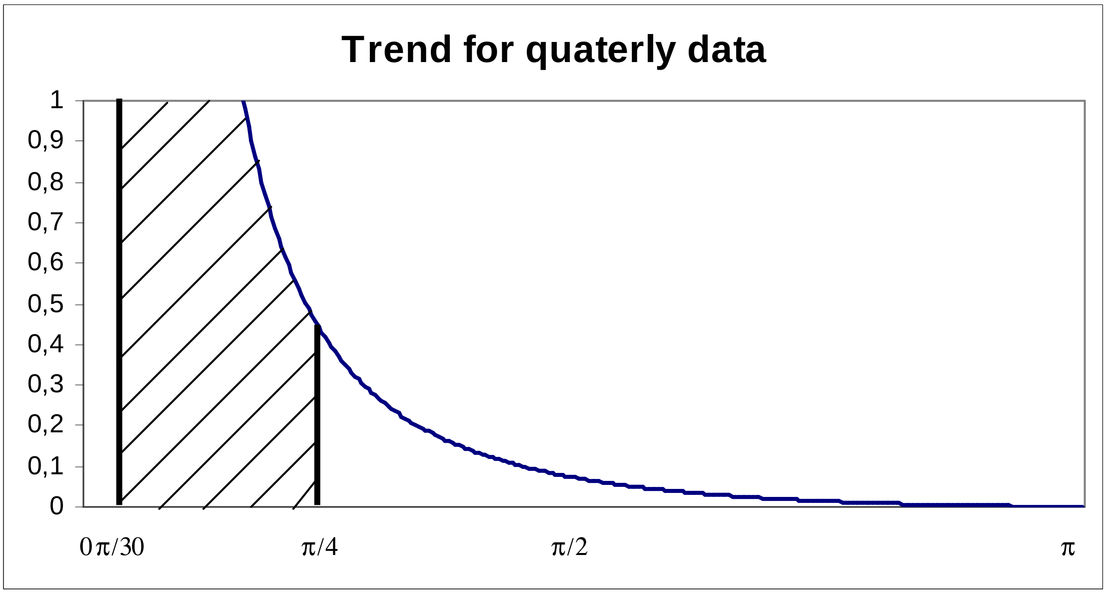
$v_t = c_t + u_t$

If c_t is ARMA (p_c, q_c), then
v_t is ARMA (p_v, q_v) with
$p_v = p_c$
$q_v = \max(p_c, q_c)$

However, a word of caution:
Transitory + Irreg. = Stationary deviations from SA and detrended series

But trend is "short-term" trend
(i.e., a trend for short-term analysis)
and may contain variation for cyclical frequencies. More properly called “trend-cycle”.

Ex: Quarterly data:

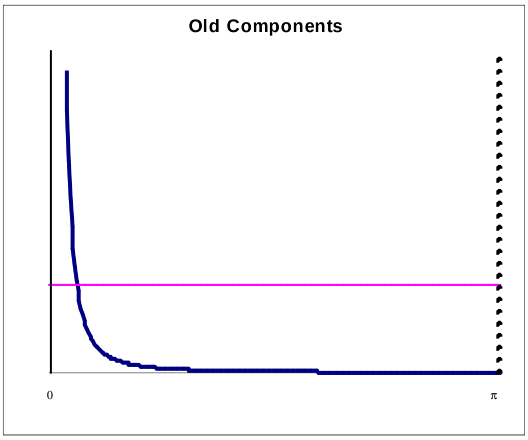


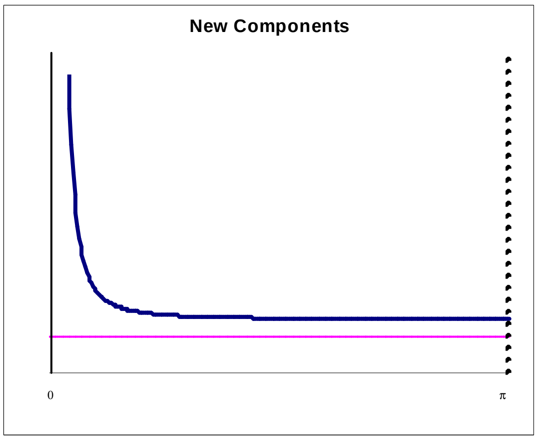
Thus "transitory-irregular component" in SEATS is not meant to be interpreted as the economic "business cycle".

<u>Note</u>: The trend-cycle of SEATS can be decomposed in a second run of SEATS into a "long-term trend" plus a "business cycle" component (Kaiser and Maravall, 2001).

IDENTIFICATION PROBLEM

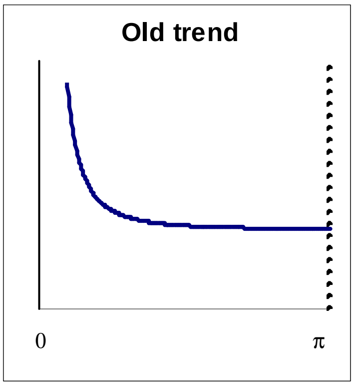
Example:

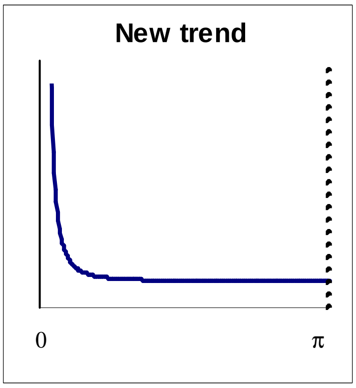


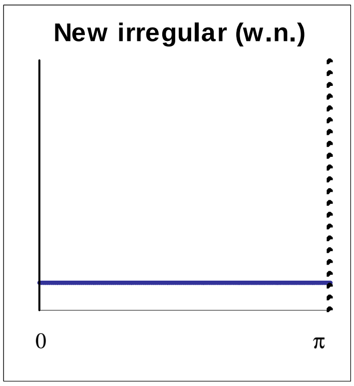


They both yield identical aggregate.

Alternatively, if p_t is invertible, we can remove some noise and add it to the irregular:







old trend = new trend + w.n.

In gral:

Can exchange noise among invertible components.

Hence:

* UNDERIDENTIFICATION problem:

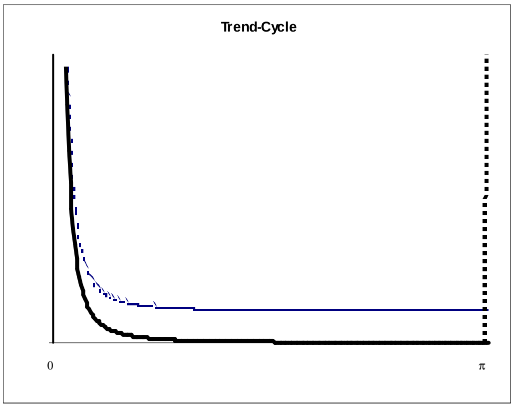
- There are ∞ models that yield the same aggregate
- They only differ in the relative allocation of white noise to the components.

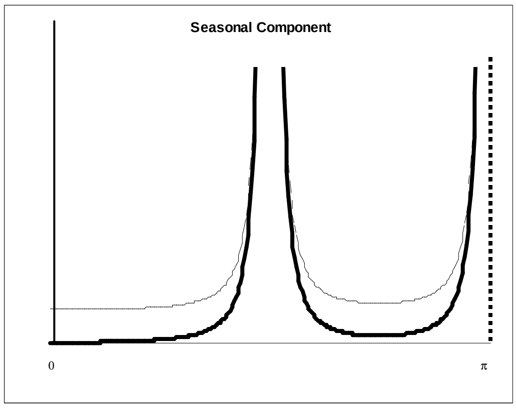
SEATS follows solution of Pierce, Box-Hillmer - Tiao, and Burman:

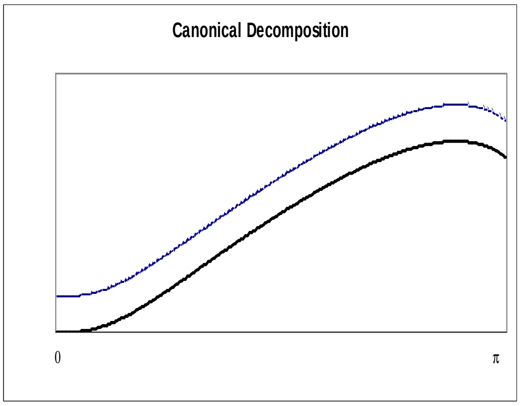
THROW ALL WHITE NOISE TO THE (WHITE-NOISE) IRREGULAR COMPONENT

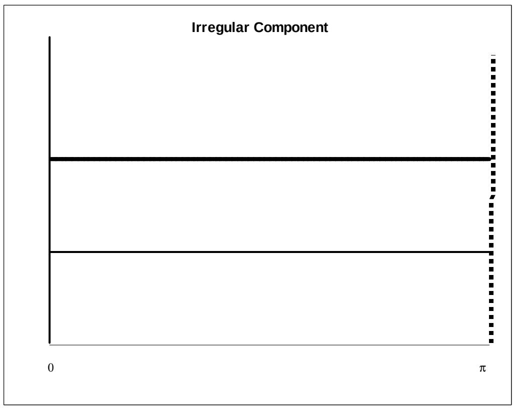
(⇒ MAXIMIZE THE VARIANCE OF IRREGULAR)

CANONICAL SOLUTION







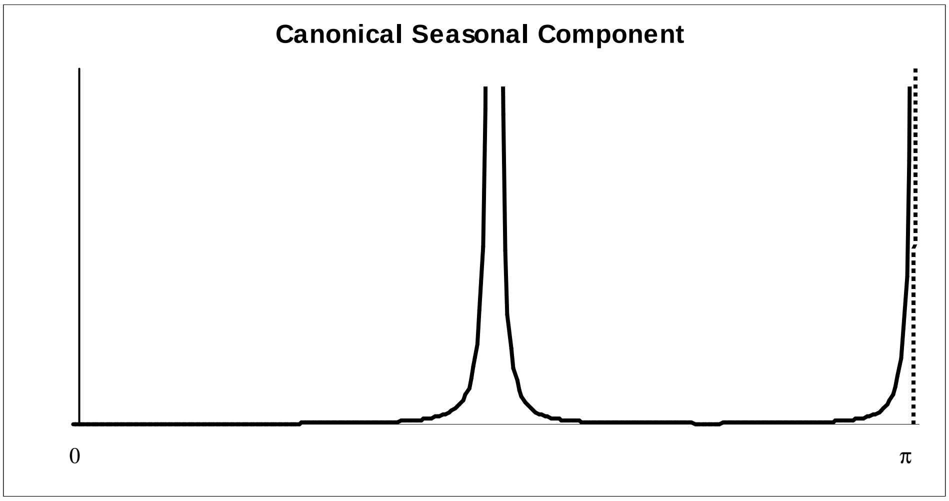


—— : canonical component
—— : any other admissible decomposition

PROPERTIES OF CANONICAL DECOMPOSITION

- Maximizes Var (u_t)
- Makes other components noninvertible:
  - they display a spectral zero;
  - they contain a unit MA root.

<u>Ex. Seasonal</u>:



$g_s(0) = 0 \Rightarrow \theta_s(B)$ contains factor $(1 - B)$

Two important <u>properties</u> of the canonical decomposition (Hillmer-Tiao)

(1) Let p'_t be the trend-cycle component in any admissible decomposition.

It can always be decomposed as

$$p'_t = p_t + e_t, \qquad \text{(B)}$$

where:

- p_t, e_t are $\perp$
- p_t is canonical trend-cycle
- e_t is w.n. with $V_e \geq 0$.

(p can be replaced by s or c)

Hence: For an observed ARIMA model, the canonical decomposition provides the "cleanest" signal.

(2) Canonical decomposition minimizes

Var (p-innovations) in components (except for u_t)

Since the p-innov. is the source of the stochastic behavior of component,

min. Var (p-innov.) $\Rightarrow$ most stable components
(compatible with observed ARIMA)

Notice that:

- if there is an admissible decomposition,
  there is a canonical decomposition.

  - Given any admissible decomposition, the canonical one can be obtained trivially.

Remark:

Sometimes, observed ARIMA model does not accept an admissible decomposition.

Ex: Airline (default) model

$$\nabla\nabla_{12}\, x_t = (1+\theta_1 B)(1+\theta_{12} B^{12})\, a_t$$

for $\theta_{12} > \bar{\theta}(\theta_1) > 0$, (a case seldom found),

the spectrum of u_t becomes negative

SEATS modifies the model until a reasonable decomposable approximation is found.

ESTIMATION OF THE COMPONENTS

In brief:

Assume, first, an ∞ realization.

$$X = [\, x_{-\infty} \dots x_t \dots x_{\infty} \,]$$

MMSE estimator of s_t: $(F = B^{-1}$; $F^j x_t = x_{t+j})$

$$\hat{s}_t = E(s_t \mid X) =$$

$$= \left[\upsilon_0 + \sum_{j=1}^{\infty} \upsilon_j (B^j + F^j) \right] x_t$$

$$= \upsilon_0 + \upsilon_1 (x_{t+1} + x_{t-1}) + \upsilon_2 (x_{t+2} + x_{t-2}) + \dots$$

$$= \upsilon(B, F)\, x_t$$

$\upsilon(B, F) \equiv$ Wiener-Kolmogorov filter.

- Convergent;
- Symmetric and centered;
- Adapts to the series;

WK FILTER

Easy algorithm to obtain it:

Assume we wish to estimate a signal, given by the model:

$$\phi_s (B)\, s_t = \theta_s (B)\, a_{st} \qquad a_{st} \sim \text{w.n}\,(0, V_s)$$

in series given by model:

$$\phi (B)\, x_t = \theta (B)\, a_t \qquad a_t \sim \text{w.n}\,(0, V_a)$$

as in

$$x_t = s_t + r_t\,,$$

$$r_t = \text{"rest"}$$

[Notice: $\phi (B) = \phi_s (B)\, \phi_r (B)$]

Write:

$$s_t = \Psi_s (B)\, a_{st} \;; \qquad \Psi_s (B) = \frac{\theta_s (B)}{\phi_s (B)} \;;$$

$$x_t = \Psi (B)\, a_t \;; \qquad \Psi (B) = \frac{\theta (B)}{\phi (B)} \;;$$

Then, for a doubly ∞ realization, the MMSE estimator of s_t is given by the WK filter

$$\hat{s}_t = \underbrace{\left[\frac{V_s}{V_a} \frac{\Psi_s(B)\,\Psi_s(F)}{\Psi(B)\,\Psi(F)} \right]}_{\text{WK}-\text{filter} = \nu(B,F)} x_t \quad .$$

Thus, in order to estimate the signal, once the ARIMA model for x_t has been identified, only the model for the signal is needed. (The other components can be ignored).

[Note: if series is stationary, WK filter is equal to

$$\nu(B,F) = \frac{\text{ACGF}\,(s_t)}{\text{ACGF}\,(x_t)} \quad]$$

Expressing the Ψ - polynomials as functions of the θ - and ϕ - polynomials, after cancelation of roots, one obtains:

$$\nu(B,F) = \frac{V_s}{V_a} \frac{\theta_s(B)\,\phi_r(B)}{\theta(B)} \frac{\theta_s(F)\,\phi_r(F)}{\theta(F)} \quad ,$$

Hence, the filter is

- Symmetric
- Centered
- Convergent (invertibility of θ (B))
- Bounded

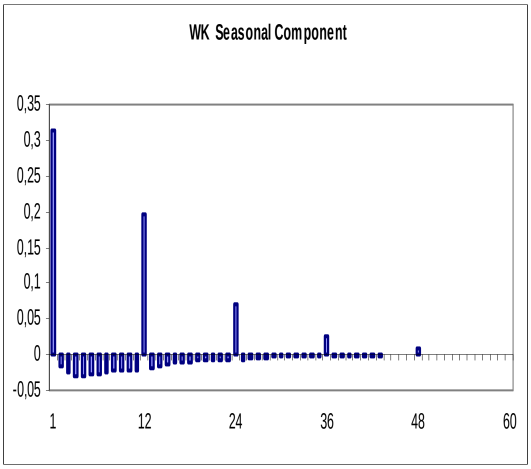
or:

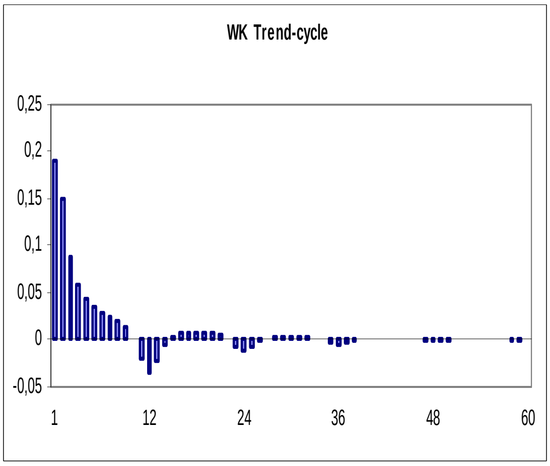
WK filter to estimate s_t is $\equiv$ ACGF of the ARMA model

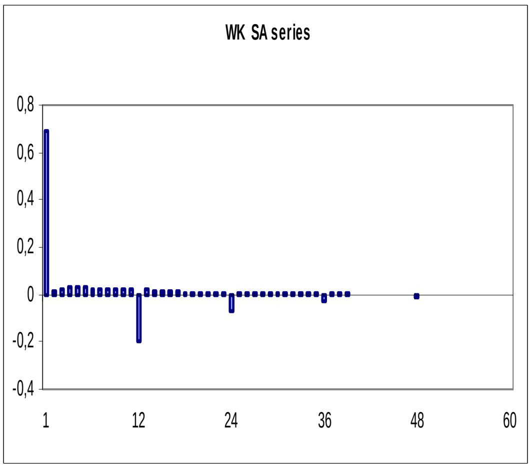
$$\theta(B)\, y_t = \left[\theta_s(B)\,\phi_r(B)\right] a_{yt} \ ;$$

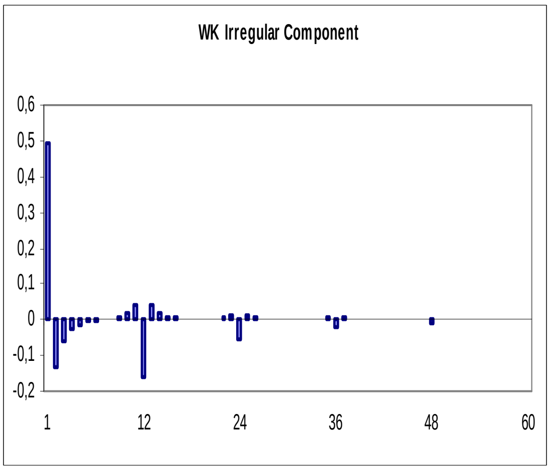
$$a_{yt} = \text{wn}\left(0, \frac{V_s}{V_a}\right) \quad ,$$

a stationary model.


WK Seasonal Component


WK Trend-cycle


WK SA series


WK Irregular Component

Example:

Estimate trend-cycle

$$\nabla^2 p_t = \theta_p (B) a_{pt}$$

in series with model

$$\nabla\nabla_4 x_t = \theta(B) a_t \ ,$$

as in

$$x_t = p_t + r_t \ .$$

We have:

$$\Psi_p(B) = \frac{\theta_p(B)}{\nabla^2}$$

$$\Psi(B) = \frac{\theta(B)}{\nabla\nabla_4} = \frac{\theta(B)}{\nabla^2 S}$$

and

$$\nu_p(B,F) = k_p \frac{\theta_p(B)\, S}{\theta(B)} \ \frac{\theta_p(F)\, \bar{S}}{\theta(F)}$$

$$(\bar{S} = 1 + F + F^2 + F^3)$$

The filter $\nu_p(B,F)$ is the ACF of the model

$$\theta (B) \ y_t = \left[\theta_p \ (B) \ \ S\right] \ a_{yt} \quad , \qquad a_{yt} \sim \ \text{w.n} \ (0, k_p)$$

Invertibility of $\theta(B)$ guarantees that the filter $\nu(B,F)$ is convergent in B and F.

CONVERGENCE of the filter implies that it can always be truncated. Thus,

for large enough series, the estimator

$$\hat{s}_t = \nu(B,F)\ x_t$$

can be assumed for the middle years of the sample.

For ex.,

if data spans 20 years, for most series the full filter can be assumed for the central 10 - 14 years.

This estimator

$$\hat{s}_t = \text{FINAL or HISTORICAL ESTIMATOR}$$

We look next at its structure.

FINAL OR HISTORICAL ESTIMATOR

We have:

$$\hat{s}_t = \nu_s(B,F)\, x_t \quad , \qquad (1)$$

from which one obtains:

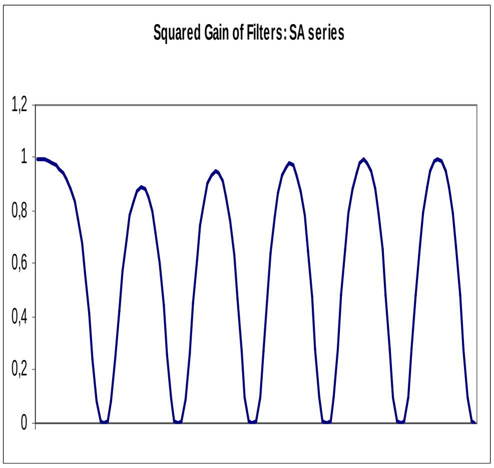
$$g_{\hat{s}}(\omega) = [G_s(\omega)]^2\, g_x(\omega)$$

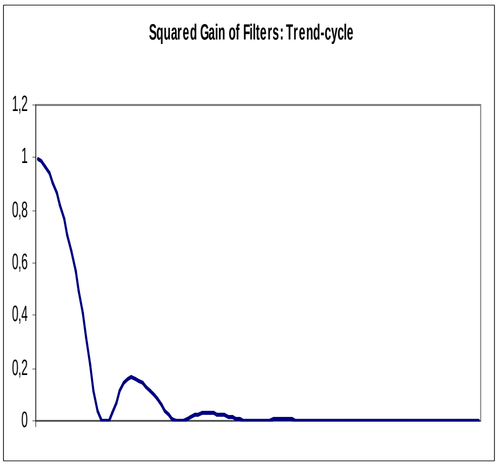
where

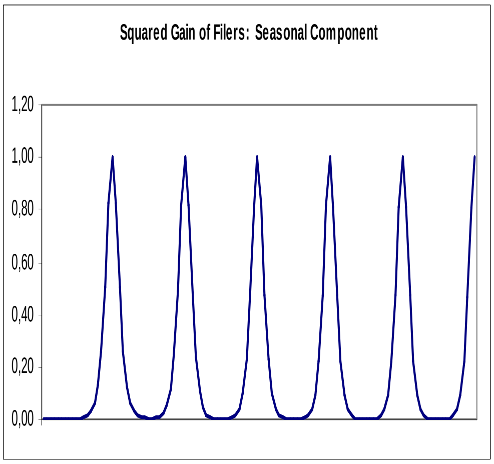
$$G_s(\omega) = [\tilde{\nu}_s(\omega)]^2$$

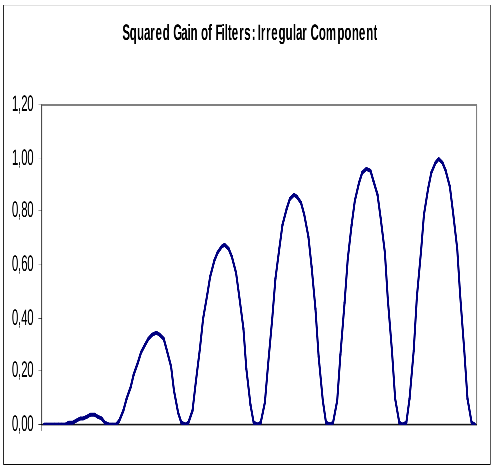
Squared Gain of filter

it determines which frequencies will contribute to the signal (that is, it filters the spectrum of the series by frecuencies).


Squared Gain of Filters: SA series
1,2
1
0,8
0,6
0,4
0,2
0


Squared Gain of Filters: Trend-cycle
1,2
1
0,8
0,6
0,4
0,2
0


Squared Gain of Filers: Seasonal Component
1,20
1,00
0,80
0,60
0,40
0,20
0,00


Squared Gain of Filters: Irregular Component
1,20
1,00
0,80
0,60
0,40
0,20
0,00

TESTING

PRESENCE / ABSENCE OF SEASONALITY;

DETERMINISTIC / STOCHASTIC SEASONALITY.

Absence or presence of seasonality:

Determined in AMI.

However:

Given that concept of seasonality somewhat implies NS, AMI in TRAMO-SEATS is slightly biased towards seasonal differencing.

Thus, on occasion, when a model of the type

$$\left[\phi(B)\ \nabla^{d}\right] \nabla_{12}\ x_t = \theta(B)\ (1 - .98\, B^{12})\ a_t \qquad \text{(D)}$$

is obtained, it may be because of seasonal overdifferencing of the model

$$\phi(B)\ \nabla^{d}\ x_t = \theta(B)\ a_t + \mu\ ,$$

a model that has no seasonality.

It can also be the result of the presence of deterministic seasonality

$$\phi(B)\ \nabla^{d}\ x_t = \theta(B)\ a_t + \mu + \sum_{i=1}^{11} \beta_i\ d_{it}\ , \qquad \text{(E)}$$

where d_{it} is a monthly seasonal dummy.

In both cases, superconsistency of $\hat{\theta}_{12}$ will yield a value close to -1.

To distinguish between the two cases, a simple F-test (easily performed in TRAMO) yields good results.
(More on this issue latter.)

However,
in the case in which there is highly stable seasonality in the series, the stochastic specification (D) is maintained, instead of the dummy-variable specification. Both are very close, and the starting values lost in (D) are compensated by the 12 additional parameters in (E) (μ plus 11 dummies). Yet (D) implicitly allows the μ and β parameters in (E) to evolve –if need be– very slowly, and the stable stochastic specification is likely to outperform the dummy-seasonal specification.

Thus, no special treatment for stable (deterministic) seasonality is needed.

It will be picked up well with the multiplicative structure

$$(\quad)\ \nabla_{12}\ x_t = (\quad)\ (1-.99B^{12})\ a_t\ .$$

TESTING FOR UNDER/OVER ADJUSTMENT

Underestimation of seasonality $\Rightarrow$ Excess Variance in SA series

Overestimation of seasonality $\Rightarrow$ Variance of SA series is too small.

In SEATS, the following comparison is performed.

The variance of the stationary transformation of the SA series and of the seasonal component are obtained for

- the theoretical value of the optimal estimator:

$$V(\hat{s}_t)$$

- the empirical value obtained for actual estimator:

$$\hat{V}(\hat{s}_t)$$

(Bartlet's approximation for $SD(\hat{V})$ yields

$$SD(\hat{V}_{\hat{s}}) = \gamma_0 \left[\frac{2}{T} \left(1 + 2 \sum_{j=1}^{m} \rho_j^2 \right) \right]^{\frac{1}{2}}$$).

Then:

$$H_0 : \quad \hat{V}(\hat{s}_t) = V(\hat{s}_t) \quad .$$

When

$$\hat{V} > \text{(significantly) } V \Rightarrow \text{ overestimation of seasonality;}$$

when

$$\hat{V} < \text{(significantly) } V \Rightarrow \text{ underestimation of seasonality.}$$

For ex.:

$$\hat{V}_{\hat{s}} = .067$$

$$\hat{V}_{\hat{s}} = .100 \qquad (SD = .010)$$

$\Rightarrow$ "EVIDENCE OF OVERESTIMATION OF SEASONALITY".

ANOTHER REPRESENTATION OF INTEREST:

THE ESTIMATOR AS A FILTER APPLIED TO THE INNOVATIONS IN THE OBSERVED SERIES

$$\hat{s}_t = \nu_s(B,F)\, x_t$$

using $x_t = \dfrac{\theta(B)}{\phi(B)}\, a_t$

$\Rightarrow$

$$\hat{s}_t = \left[\nu_s(B,F)\,\frac{\theta(B)}{\phi(B)}\right] a_t = \left[\,\xi_s(B,F)\,\right] a_t$$

$\xi_s(B,F)$ = "PSIE-weights" (easy to obtain: Maravall, 1994)

$$\xi_s(B,F) = \underbrace{\ldots + \xi_j\, B^j + \ldots + \xi_1\, B + \xi_0 +}_{\text{applies to prior and concurrent innovations}}$$

$$\underbrace{+\, \xi_{-1}\, F + \ldots + \xi_{-j}\, F^j + \ldots}_{\text{applies to "future" innovations (posterior to t)}}$$

ξ_j = contribution of a_{t-j} to $\hat{s}_t$

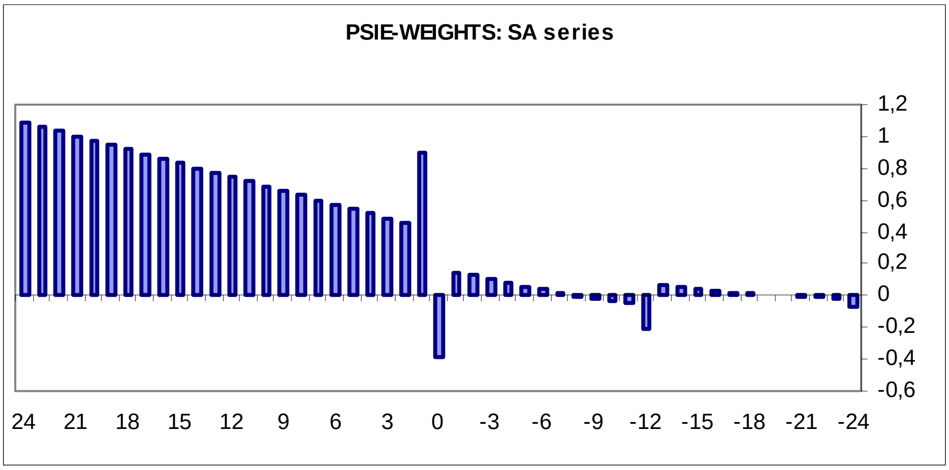
ξ_{-j} = contribution of a_{t+j} to $\hat{s}_t$

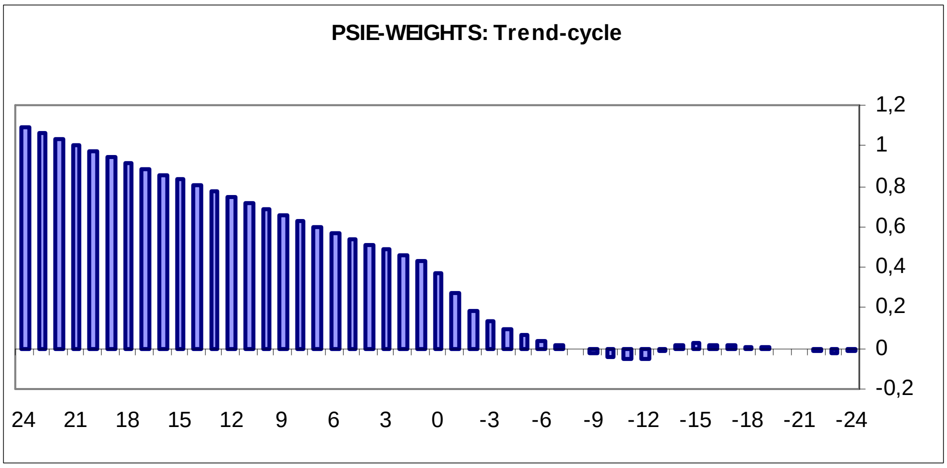
Note

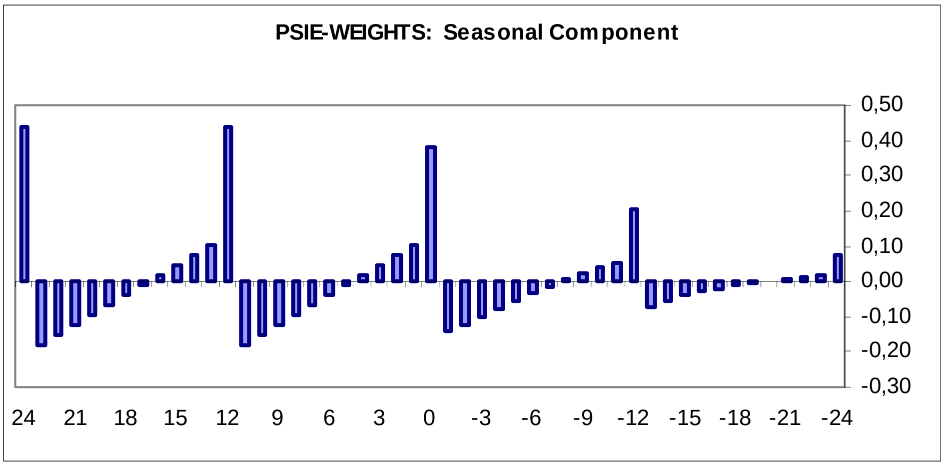
ξ_s **(B,F)** is: asymmetric

non-convergent in B (unless series is stationary)

convergent in F (always)







As we shall see later, the filter is important to analyse **revisions** and **convergence** of the estimator, as well as **SE of preliminary estimators**.

Example:

$\nabla^2 p_t = \theta_p(B)\, a_{pt}$: SIGNAL

$\nabla\nabla_4 x_t = \theta(B)\, a_t$: SERIES

Then:

$$\hat{p}_t = \nu_p(B,F)\, x_t \quad ,$$

$$\nu_p(B,F) = \frac{\theta_p(B)\, S}{\theta(B)} \frac{\theta_p(F)\, \bar{S}}{\theta(F)}\, k_p \quad ,$$

where $k_p = V_p / V_a$,

$\bar{S} = 1 + F + F^2 + F^3$. Thus

$$\hat{p}_t = \nu_p(B,F) \frac{\theta(B)}{\nabla^2 S}\, a_t =$$

$$= \Bigg[\underbrace{k_p \frac{\theta_p(B)}{\nabla^2}}_{\text{part in B}} \quad \underbrace{\frac{\theta_p(F)\, \bar{S}}{\theta(F)}}_{\text{part in F}} \Bigg] a_t$$

Notice:

$$\nabla^2 \hat{p}_t = \theta_p(B) \left[\frac{\theta_p(F)\, \bar{S}}{\theta(F)} \right] a'_t$$

with $\mathrm{Var}(a'_t) = k_p^2 V_a$. Somewhat different from model for SIGNAL above.

Previous remark brings a point of general interest:

<u>MODEL FOR COMPONENT versus MODEL FOR ESTIMATOR</u>

We have

MODEL FOR SERIES

$$\phi(B)\ x_t = \theta(B)\ a_t \qquad \text{("observed")}$$

MODEL FOR COMPONENT (two components)

$$x_t = s_t + n_t\ , \qquad \lfloor \phi(B) = \phi_s(B)\ \phi_n(B) \rfloor$$

$$\phi_s(B)\ s_t = \theta_s(B)\ a_{st}$$

MMSE estimator for s_t (doubtly infinite realization):

$$\hat{s}_t = k_s \left[\frac{\theta_s(B)\ \phi_n(B)}{\theta(B)} \quad \frac{\theta_s(F)\ \phi_n(F)}{\theta(F)} \right] x_t$$

it is found:

MODEL FOR ESTIMATOR

$$\phi_s(B)\ \hat{s}_t = \theta_s(B) \left[\frac{\theta_s(F)\ \phi_n(F)}{\theta(F)} \right] a'_t$$

$(a'_t = k_s\ a_t)$

Comparison of the model for the component with that of the estimator shows the effects induced by the estimation filter.

	Stationary transformation	Stationary model	
		Part in B	Part in F
Component	$\phi_s(B)\, s_t =$	$\theta_s(B)$	—
Estimator	$\phi_s(B)\, \hat{s}_t =$	$\theta_s(B)$	$\dfrac{\theta_s(F)\, \phi_n(F)}{\theta(F)}$

Component and estimator share

- the stationarity-inducing transformation (in particular, the differencing)
- the (stationary) part in B

Difference: estimator includes a part in F (reflecting the 2-sided character of the filter). This part is a **convergent** polynomial in F.

Component and estimator will have different ACF and spectrum.

The different model structures of component and estimator have some implications of applied relevance.

One implication

It is (close to) standard practice to build models on seasonally adjusted data.

This is based on the belief that, by removing seasonality, model dimensions can be reduced.

This belief is wrong.

Example: DEFAULT MODEL

$$\nabla\nabla_{12}\ x_t = (1+\theta_1 B)(1+\theta_{12} B^{12})\ a_t$$

Decomposes into:

$$\left.\begin{aligned} \nabla^2\ n_t &= \theta_n(B)\ a_{nt} \\ S\ s_t &= \theta_s(B)\ a_{st} \end{aligned}\right\}$$

The model for the estimator of the SA series has ACF of model:

$$(1+\theta_1 B)(1+\theta_{12} B^{12})\ \nabla^2\ \hat{n}_t =$$

$$= \theta_n(B)\quad \theta_n(B)\ S\ a_t,$$

an ARIMA (13,2,15) model.

Set, for instance,

$$\theta_1 = -.4$$

$$\theta_{12} = -.6 .$$

The MA expansions (or ψ-weights) of the stationary transformation of

- the original series
- the seasonally adjusted series,

are the following:

LAG	ORIGINAL SERIES	SA-SERIES (ESTIM.)
0	1	1.00
1	- 0.4	-1.33
2	-	0.38
3	-	-
4	-	-
5	-	-
6	-	-
7	-	-
8	-	-
9	-	-
10	-	-
11	-	-
12	- 0.6	-0.40
13	0.24	0.53
14	-	-0.15
15	-	-
16	-	0.37
17	-	0.15
18	-	0.06
19	-	0.02
20	-	0.01
21	-	-
22	-	-
23	-	-
24	-	-0.24
25	-	0.32
26	-	-0.09
27	-	-
28	-	-
29	-	-
30	-	-
31	-	-
32	-	-
33	-	-
34	-	-
35	-	-
36	-	-0.14
37	-	0.19
38	-	-0.05

Hence: * Model for SA series: MORE COMPLEX

* **No reduction in dimension if SA series is used.**

This is a reason to avoid modelling SA series

$\downarrow$

they will have coefficients for seasonal and large lags.

A second implication:

Broadly, the difference between theoretical component and estimator is the following

Component: $\phi_s(B)\, s_t = \theta_s(B)\, a_{st}$

Estimator: $\phi_s(B)\, \hat{s}_t = \theta_s(B)\, \alpha_s(F)\, a_t$

Difference: $\alpha_s(F) = \dfrac{\theta_s(F)\, \phi_n(F)}{\theta(F)}$

When $\theta_s(F)$ or $\phi_n(F)$ contain unit roots, given that these roots will appear in the MA part of the estimator, the estimator will not be invertible.

When n_t (what is removed in order to obtain s_t) is NON-STATIONARY $\rightarrow$ $\phi_n(B)$ will contain unit roots.

Example: Default (Airline) model.

Component models:

$\nabla^2 n_t = \theta_2(B)\, a_{nt}$

$S\, s_t = \theta_{11}(B)\, a_{st}$

Thus $\phi_s(B) = S$.

$S \equiv 11$ unit root.

Therefore, $\hat{n}_t$ will be NI because of these unit roots.

Recall: the presence of seasonality (in general) $\Rightarrow$ unit AR roots in model for seasonal.

Consequence:

In gral, for $\hat{n}_t$

* No convergent AR representation (nor VAR representation) exists.

AVOID USING AR MODELS TO MODEL SA SERIES

General result:

$$x_t = m_t + n_t$$

$\Rightarrow$ $\hat{m}_t$ is Noninvertible if

n_t is Nonstationary (Maravall, 1995)

Hence in a standard

trend + seasonal + irregular

decomposition, with NS trend and NS seasonality, all three : $\hat{p}_t, \hat{s}_t$, and $\hat{u}_t$ will be NI.

Noninvertibility of the estimators
(and hence previous implications)
is a fairly general property of SA and detrending methods (including X11)

Another important applied result:

We saw

$$g_{\hat{u}}(\omega) = [G_u(\omega)]^2 g_x(\omega) \quad ,$$

where u denotes now any of the components, and

$$G_u(\omega) = \frac{g_u(\omega)}{g_x(\omega)} \quad \equiv \text{ Gain of filter } .$$

Thus

$$g_{\hat{u}}(\omega) = \left[\frac{g_u(\omega)}{g_x(\omega)} \right] g_u(\omega)$$

Since $\left[\frac{g_u(\omega)}{g_x(\omega)} \right] \leq 1$,

for all components:

spectrum of component ≥ spectrum of estimator.

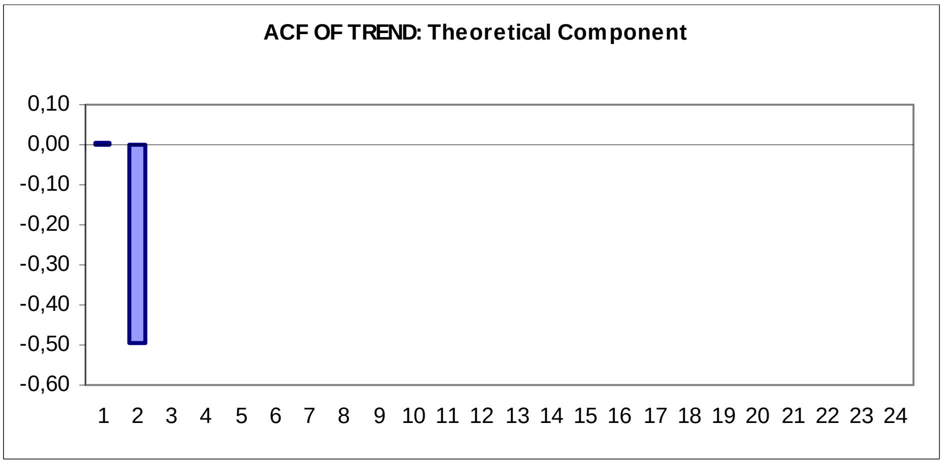
In particular, for the stationary transformation,

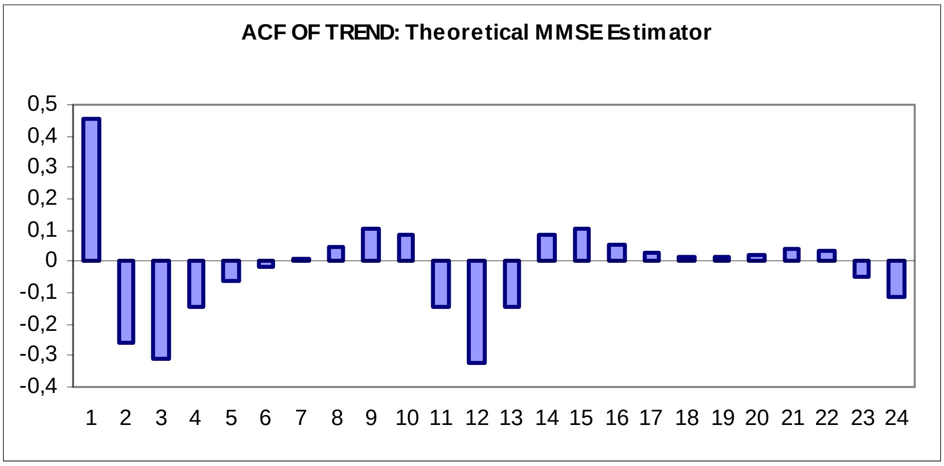
$$\boxed{\text{Var}(\text{comp.}) \geq \text{Var}(\text{estimator}) \quad .}$$

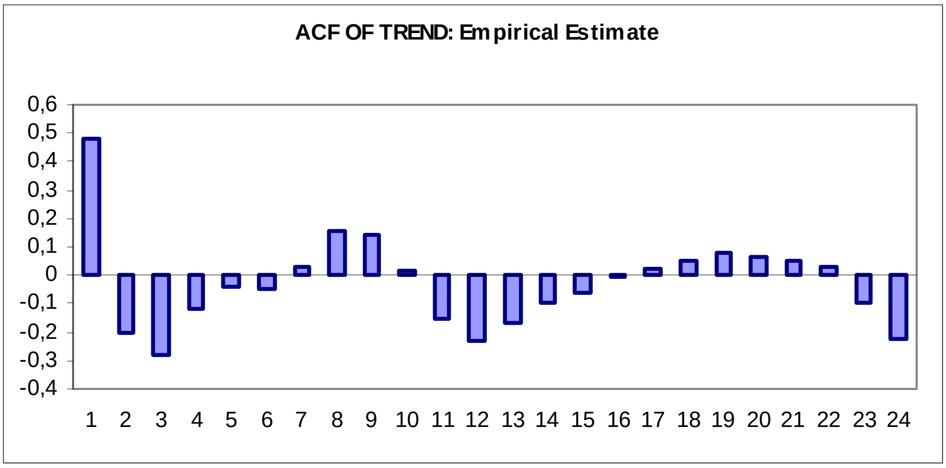
When some other component is present, the estimator will always **underestimate** the stochastic variance of the component (bias towards “stability”).

The loss of variance counterpart is the appearance of crosscovariances between components' estimators. (As shall be seen later, these crosscovariances can also be modelled.)

In a particular application, to see if the empirical estimates agree with the model, their variances and ACF should be compared to those of the model for the estimator, not to those of the model for the component.


ACF OF TREND: Theoretical Component
0,10
0,00
-0,10
-0,20
-0,30
-0,40
-0,50
-0,60
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24


ACF OF TREND: Theoretical MMSE Estimator
0,5
0,4
0,3
0,2
0,1
0
-0,1
-0,2
-0,3
-0,4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24


ACF OF TREND: Empirical Estimate
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
-0,2
-0,3
-0,4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

JOINT DISTRIBUTION OF THE ESTIMATORS

From the models for the estimators, the variances, ACFs, spectra, and so on, can be obtained for their stationary transformations (ST).

We can further obtain the

(THEORETICAL) CROSS-COVARIANCE FUNCTION

between any pair of estimators (ST).

Therefore:

- Given our Normality assumption, the joint distribution of the (ST of the) estimators is known;
- it is fully determined by the "observed" ARIMA;
- this knowledge permits us to devise simple tests having to do with issues related to the decomposition obtained in a particular application.

The properties of the estimators have been derived for the case of an ∞ realization.

Since WK-filter is convergent (in B and in F), in practice it could be approx. with a finite (2-sided) filter.

Estimators that are obtained with the full filter:

FINAL (OR HISTORICAL) ESTIMATOR

The time it takes for the filter to converge depends on:

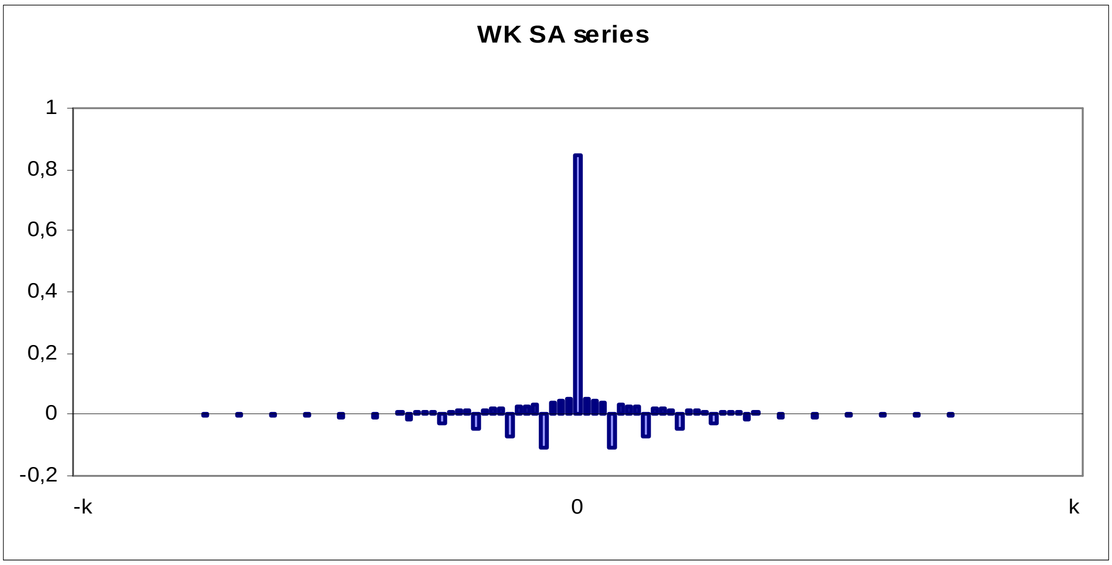
- stochastic structure of the series,
- stochastic structure of the component.

Ex. Monthly series with seasonality.
For many, 3 years of revisions are enough.
(Most are completed in 5 years)
Assuming 3 years, if series has 180 observations,
for the central 108 the component estimator could be considered final.

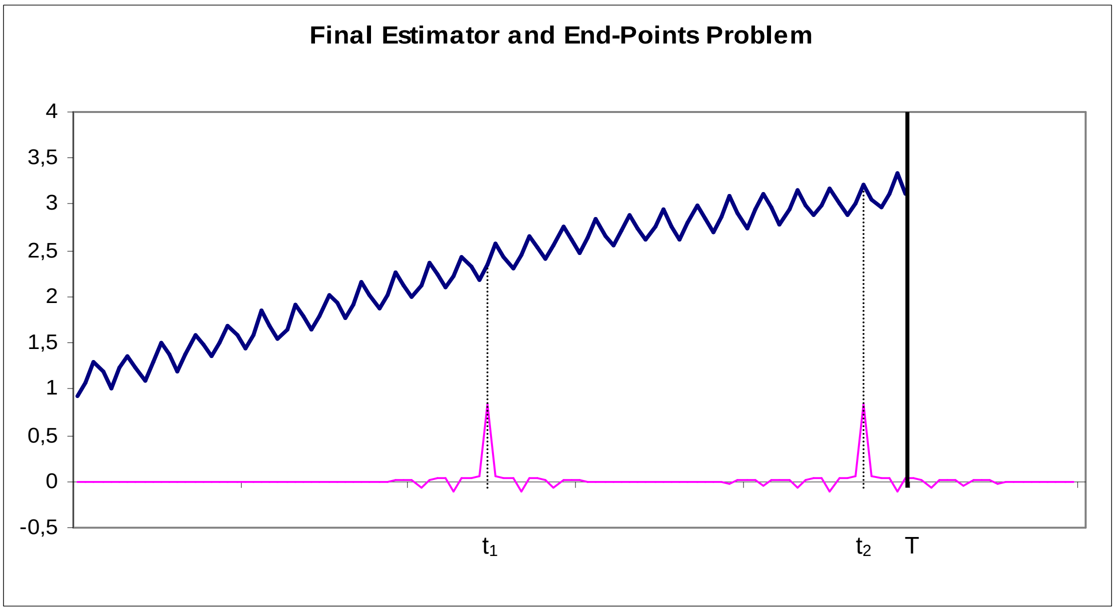
FINITE SERIES ; PRELIMINARY ESTIMATION

Observed series: $[x_1, x_2, \ldots, x_T]$

Consider WK filter to estimate SA series (n_t)



To obtain $\hat{n}_{t_1|T}$ there is no problem: filter has converged to that for final estimator.



However, the filter cannot be used to obtain $\hat{n}_{t_2|T}$ because convergence of the filter requires future observations, not yet available.

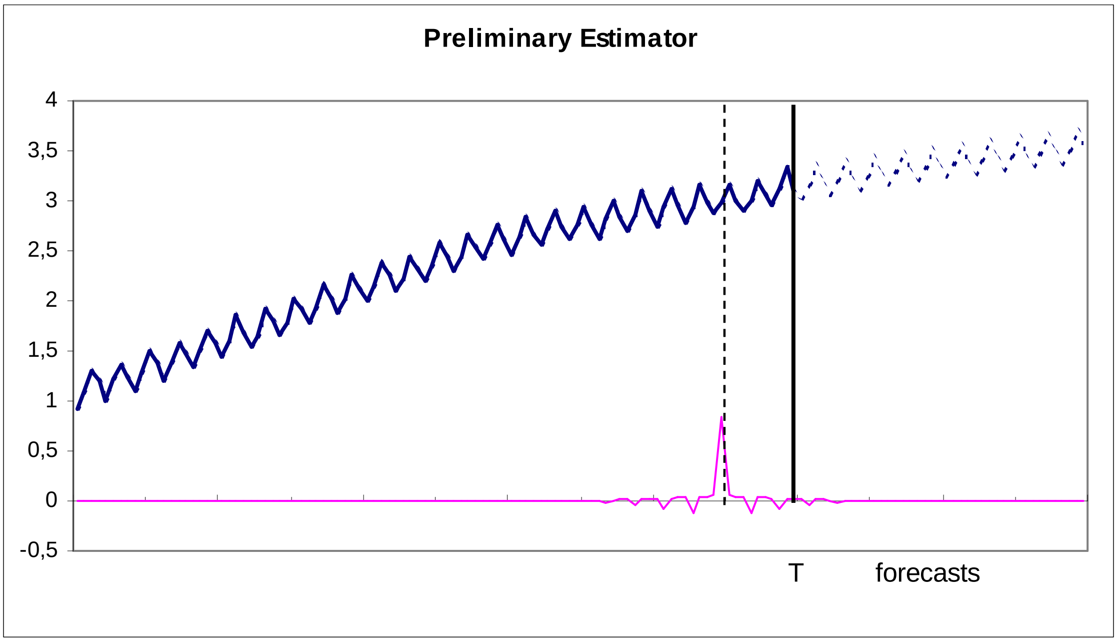
(same problem near the beginning of the series)

way to proceed **(Preliminary Estimators)**

1) EXTEND SERIES WITH FORECASTS AND BACKCASTS (ARIMA ONES)

2) APPLY FILTER TO EXTENDED SERIES

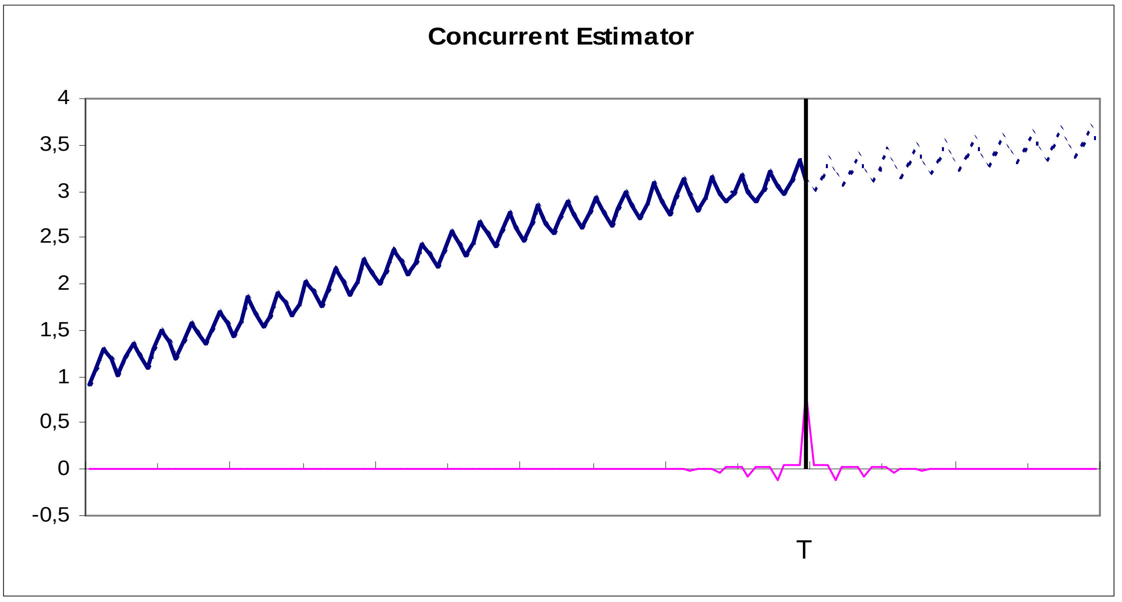
(Cleveland and Tiao, JASA, 1976)



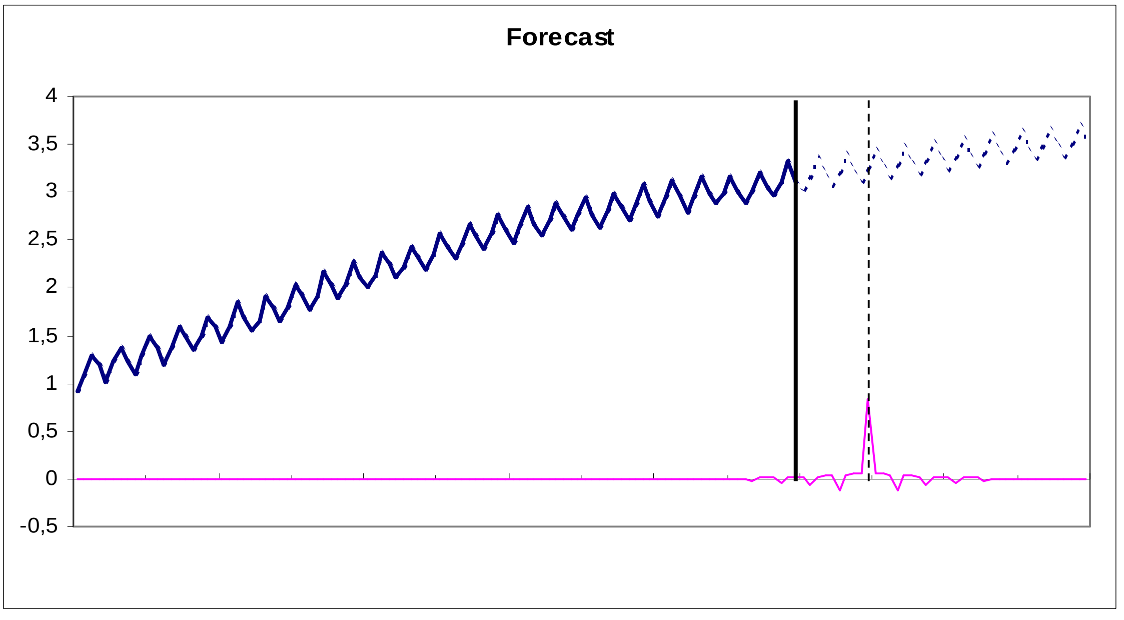
In this way, a PRELIMINARY ESTIMATOR is obtained.

As new observations become available, it will be revised, until the final estimator is obtained.

The most important preliminary estimator: $\hat{n}_{T|T}$



Forecasts of n are obtained in the same way as preliminary estimators, simply by extending the series further.



In SEATS, for monthly data:

24-months-ahead forecasts are computed for the series and components.

In general, if MQ is the number of observations per year, the number of forecasts computed is max (2MQ, 8).

FINAL ESTIMATOR:

$$\hat{s}_t = \nu_s(B,F)\, x_t =$$
$$= \ldots + \nu_1 x_{t-1} + \nu_0 x_t + \nu_1 x_{t+1} + \ldots$$

PRELIMINARY ESTIMATOR:

Obtained by replacing observations not yet available with forecasts.

Let

$\hat{s}_{t|T}$: Estimator of s_t when last observation is x_T

For example:

$$\hat{s}_{t|t-1} = \ldots + \nu_1 x_{t-1} + \nu_0 \underbrace{\underset{\substack{\uparrow \\ \text{1 p.a.f.}}}{\hat{x}_{t|t-1}} + \nu_1 \underset{\substack{\uparrow \\ \text{2 p.a.f.}}}{\hat{x}_{t+1|t-1}}}_{\substack{\text{obtained as} \\ \text{ARIMA forecasts}}} + \ldots$$

In summary, for finite realization:

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

Preliminary Estimator:

$$\hat{s}_{t|T} = \upsilon_s(B, F)\,\hat{x}^e_{t|T}$$

$\hat{x}^e_{t|T} =$ [backcasts, observations, forecasts]

i.e., series extended with forecasts and backcasts

No need for long extensions:

Burman-Wilson algorithm: only a few forecasts and backcasts are needed. (Typically, about 2 years)

Note: Preliminary estimator will imply an asymmetric filter, and will be subject thus to a phase effect.

As new observations become available:

$$\hat{s}_{t|T} \rightarrow \hat{s}_{t|T+1} \rightarrow \cdots \quad \hat{s}_{t|T+k} \rightarrow \cdots$$

the estimator of s_t is revised.

As $k \rightarrow \infty$, (in practice, “large enough”)

$\hat{s}_{t|T+k} \rightarrow \hat{s}_t$ (the "final" or "historical" estimator)

(In practice,

$\hat{s}_t \equiv$ Historical or Final Estimator is valid for central years of the series)

STRUCTURE OF THE SA SERIES AVAILABLE AT TIME T:

For a particular realization $[x_1, x_2, \ldots, x_T]$,

what we have is a sequence of estimators:

$\cdots \hat{s}_{T-100|T} \cdots \hat{s}_{T-j|T} \cdots \hat{s}_{T|T}$

$\hat{s}_{T-100|T}$ $\equiv$ FINAL EST. $= \hat{s}_{T-100}$

$\hat{s}_{T-j|T}$ $\equiv$ PRELIMINARY EST. (j covering a few years)

$\hat{s}_{T|T}$ $\equiv$ CONCURRENT EST.

Each one of these estimators is the output of a different model.
(Each j $\Rightarrow$ a different model.)

Therefore, SA series is a mixture of realizations with different underlying models.

Thus: SA series available at some point in time is nonlinear
($\cong$ time-varying parameters model).

* Another reason to avoid using SA in modeling.

PRELIMINARY ESTIMATORS AND REVISIONS

Preliminary estimators and revisions are implied by the use of

TWO-SIDED FILTER

$$\hat{s}_t = \ldots + \upsilon_1 x_{t-1} + \upsilon_0 x_t + \upsilon_1 x_{t+1} + \upsilon_2 x_{t+2} + \ldots$$

Two-sided filter is

- necessary to avoid phase effects;

- implied by MMSE ("optimal") estimation in model-based approach.

Starting with concurrent estimator:

Observations: $[x_1, \ldots, x_t]$

$$\hat{s}_{t|t} = \ldots + \upsilon_0 x_t + \upsilon_1 \hat{x}_{t+1|t} + \upsilon_2 \hat{x}_{t+2|t} + \upsilon_3 \hat{x}_{t+3|t} + \ldots$$

↑ 1 p.a. forecast ↑ 2 p.a. forecast

Notice that:

Given that all forecasts are l.c. of x_{t-j} (j = 0, 1, 2, …), $\hat{s}_{t|t}$ is implicitely obtained with a one-sided filter.

New observation (x_{t+1}) arrives.

New revised estimator (1-period revision)

$$\hat{s}_{t|t+1} = \dots + \nu_0 x_t + \nu_1 x_{t+1} + \nu_2 \hat{x}_{t+2|t+1} + \nu_3 \hat{x}_{t+3|t+1} + \cdots$$

new observation

updated forecast

Likewise, when x_{t+2} becomes available, the 2-period revision of the concurrent estimator will be given by

$$\hat{s}_{t|t+2} = \dots + \nu_1 x_{t-1} + \nu_0 x_t + \nu_1 x_{t+1} + \nu_2 x_{t+2} + \nu_3 \hat{x}_{t+3|t+2} + \dots$$

and so on.

Of course, to revise series is always disturbing and an inconvenience.

But it is due to the fact that knowledge of the future helps us to understand the present (a very basic fact of life!).

To suppress revisions is

- to ignore relevant information,
- to distort our measurements.

Revisions:

1-period revision:

$$r_{t|t}^{(1)} = \hat{s}_{t|t+1} - \hat{s}_{t|t} = \nu_1 (x_{t+1} - \hat{x}_{t+1|t}) + \nu_2 (\hat{x}_{t+2|t+1} - \hat{x}_{t+2|t}) + \ldots =$$

$$= \xi_1 \ a_{t+1} \qquad (= w.n.)$$

(ξ_1 = a constant) .

2-period revision:

$$r_{t|t}^{(2)} = \hat{s}_{t|t+2} - \hat{s}_{t|t} = \nu_1 (x_{t+1} - \hat{x}_{t+1|t}) + \nu_2 (x_{t+2} - \hat{x}_{t+2|t}) +$$

$$+ \nu_3 (\hat{x}_{t+3|t+2} - \hat{x}_{t+3|t}) + \ldots =$$

$$= \xi_1 \ a_{t+1} + \xi_2 \ a_{t+2}$$

$$= MA(1)$$

……………………………………..

k-period revision:

$$r_{t|t}^{(k)} = \hat{s}_{t|t+k} - \hat{s}_{t|t} = \xi_1 \ a_{t+1} + \ldots + \xi_k \ a_{t+k} =$$

$$= MA(k-1)$$

For the <u>full revision in the concurrent estimator</u> :

$$r_t = \hat{s}_t - \hat{s}_{t|t} = \upsilon_1 (x_{t+1} - \hat{x}_{t+1|t}) +$$

$$+ \upsilon_2 (x_{t+2} - \hat{x}_{t+2|t}) + \ldots$$

$$= \upsilon_1 e_t (1) + \upsilon_2 e_t (2) + \ldots$$

where:

$e_t (j)$: j-th-period-ahead forecast error of the series

[$e_t (1) = a_{t+1}$

$e_t (2) = a_{t+2} + \psi_1 a_{t+1}$

……………………………].

Hence

$$r_t = \sum_{j=1}^{\infty} \upsilon_j \; e_t \; (j)$$

depends on:

- forecast errors

- weights of the WK filter

Thus:

- interest in "small" forecast errors (X11 $\rightarrow$ X11 ARIMA)

- but revision still depends on the υ_j's,

  WHICH DEPEND, in turn, ON THE STOCHASTIC STRUCTURE OF THE SERIES
  (i.e., the ARIMA model).

For some series, the revision can be large;
for other series, they may be small.

Also, for some series the revision will last long;
for others it will disappear fast.

THUS, FOR A GIVEN SERIES, THERE IS AN APPROPRIATE AMOUNT OF REVISION.

THE REVISION SHOULD NOT BE LARGER THAN THAT, NOR SHOULD IT BE SMALLER.

Two features of the revision process are of relevance:

- the size of the revision

- the duration of the revision process.

Often one finds there is a trade-off between them.

ERROR IN THE ESTIMATOR OF A COMPONENT

In the context of Seasonal Adjustment, concern with the error made when measuring seasonality has been periodically expressed (Bach et al. 1976; Moore et al. 1981; Bank of England 1992). This need is especially left for key variables that are (explicitly or implicitly) being subject to some type of targeting (e.g., a monetary aggregate or a consumer price index). In these cases, intrayear monitoring and policy reaction is based on the SA series (e.g., see Maravall 1988).

We are concerned with the precision of the

- concurrent estimator and forecasts
- first revisions
- final estimator
- some rates of growth.

The associated MSE are straightforward to obtain.

From the previous discussion, it is clear that the error will be different for

a) final estimator, $\hat{s}_t$

b) preliminary estimator, $\hat{s}_{t|t+k}$; $k = \pm 1, \pm 2, \ldots$,

(which also includes forecasts).

Total estimation error in the estimator $\hat{s}_{t|T}$

$$\varepsilon_{t|T} = s_t - \hat{s}_{t|T} \quad ,$$

it can be expressed as the sum of the two errors,

$$\varepsilon_{t|T} = (s_t - \hat{s}_t) + (\hat{s}_t - \hat{s}_{t|T}) \quad ,$$

where the first error

$$e_t = s_t - \hat{s}_t$$

is **the error in the final estimator**, and the second error

$$r_{t|T} = \hat{s}_t - \hat{s}_{t|T}$$

denotes the revision in the estimator $\hat{s}_{t|T}$.

* e_t and $r_{t|T}$ are orthogonal (Pierce 1979)

thus, for example, $V(\varepsilon_{t|T}) = V(e_t) + V(r_{t|T})$.

REVISION ERROR: Size and Convergence

Express, as before, component as filter of innovations in series:

$$\hat{s}_t = \upsilon_s(B,F)\, x_t$$

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

or

$$\hat{s}_t = \xi_s(B,F)\, a_t$$

* Divergent in B

* Convergent in F

Write:

$$\hat{s}_t = \xi_s(B)^{-}\, a_t + \xi_s(F)^{+}\, a_{t+1}$$

The filter $\xi_s(B,F)$ can be easily computed (Maravall, Journal of Forecasting, 94).

For a <u>concurrent estimator</u>:

$\xi_s(B)^- a_t$: Effect of starting conditions and present and past innovations in series.

$\xi_s(F)^+ a_{t+1}$: Effect of future innovations.

Taking conditional expectations at time t ,

$$\hat{s}_{t|t} = \xi_s(B)^- a_t$$

the revision is given by

$$r_t = \hat{s}_t - \hat{s}_{t|t} \quad ; \text{ or}$$

$$\boxed{r_t = \xi_s(F)^+ a_{t+1}}$$

a zero-mean convergent one-sided (stationary) MA.

HISTORICAL (OR FINAL) ESTIMATION ERROR

Because of its stochastic nature, the historical estimator $\hat{s}_t$ contains an estimation error

$$e_t = (s_t - \hat{s}_t)$$

* "unobservable"
* finite variance
* can derive distribution.

In particular, e_t has ACF and spectrum of the model (Pierce, 80)

$$\theta(B)\, e_t = \left[\theta_s(B)\, \theta_n(B)\right] a_{et} \quad ,$$

$$a_{et} \sim \text{wn}\left(0, \frac{V_s\, V_n}{V_a}\right)$$

ERROR ANALYSIS: SOME APPLICATIONS

A. From knowledge of the models for the different types of estimation errors, one can build standard

TESTS FOR THE SIGNIFICANCE OF SEASONALITY

in a particular application, such as, for example,

$$H_0 : \quad \underline{\hat{s}} = \underline{0} \ ,$$

where $\underline{\hat{s}}$ is a vector of estimators with known covariance matrix. Notice that it may be possible to detect significant seasonality with the final estimator, yet the forecasts of the seasonal component for the next year may be worthless.

B. Proper **intrayear monitoring** of the economy is greatly facilitated.

For example: Assume an increase in unemployment of 10000 persons in last month, as measured with (the concurrent estimator of) the SA series. We can easily test for whether this increase is significantly different from zero.

Naturally, economic policies based on some (explicit or implicit) annual (or biannual, …) targeting, where **short-term control** is typically based on the SA series, the variance of the estimation error of seasonality can be used to build **confidence intervals around the estimated SA series**. In this way, the question "are we growing too much?" (or "too little") can be answered in a more rigorous manner.

Historical Estimation Error: never known.
The best we can do: Historical Estimator.

Hence, from applied point of view, perhaps CI should only consider the Revision Error.
("How far can I be from my eventual best measurement?")

C. USE IN MODEL ESPECIFICATION

The possibility of deriving properties of the components can be of help in the choice of the proper model.

It is often the case that several ARIMA specifications seem about equally acceptable from the fitting and out-of-sample forecasting criteria. In these cases, one can look at the decompositions implied by the "sample equivalent" models, and select the one that offers the most appealing decomposition. Some important criteria that can be used in the comparison are the following:

- *Stability of the components*

  One may wish to remove a seasonal component as stable as possible. Thus one may seek the decomposition with

$$\min \left[\mathrm{Var}(a_{st}) \right].$$

- *Precission of the estimator*

  (better detection of seasonality…)

- *SMALLER REVISIONS*

  (and fast convergence) .

(examples in Bank of Spain web site).

AMB used as "fixed filter": SEATS by default (RSA = 0).
A remark on the *DEFAULT MODEL*

"Airline Model" (Box-Jenkins, 1970)

$$\nabla\nabla_{12}\ x_t = (1+\theta_1 B)(1+\theta_{12} B^{12})\, a_t\ +\ \mu$$

The annual difference of the monthly growth (rate-of-growth if in logs) is a stationary process, with constant variance

* Parameters have "structural" interpretation

$\theta_1 \rightarrow$ stability of trend

$\theta_{12} \rightarrow$ stability of seasonal

(values close to -1 produce stable components)

$\sigma_a \rightarrow$ overall unpredictability

* Often found

* Displays very well-behaved filters

* Can encompass many models (in a fairly robust way)

Ex.:

- deterministic trend: $\theta_1 \to -.99$

- deterministic seasonal: $\theta_{12} \to -.99$

- even white noise ! (θ_1 and $\theta_{12} \to -.99$)

• No need for the dilema: Deterministic vs. Stochastic.

• Overdifferencing does little damage.

• Good idempotency properties.

• Good for pretesting.