Pontifícia Universidade Católica do Rio de Janeiro

Laboratório de Estatística Computacional

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Modelos Estruturais para Séries Temporais

Prof. Cristiano Fernandes

2005

NOTAS DE AULA - I

1. INTRODUÇÃO

1.1 Objetivos e Motivação:

• O objetivo central do curso é a apresentação e discussão dos algoritmos utilizados para estimação, previsão e alisamento dos modelos em espaço de estado gaussianos e lineares. (MEEGL) para séries temporais (ST).

• Os algoritmos a serem estudados são o filtro de Kalman, os algoritmos de suavização e a verossimilhança por decomposição de erro de previsão.

• Ênfase especial será dada a uma classe especial dos MEEGL, os modelos estruturais, onde será abordada a formulação clássica de Harvey et al.

• Nos modelos estruturais uma ST é decomposta em componentes de interesse, tais como tendência, sazonalidade e ciclo. Esquemáticamente:

ST = Tendência + Sazonalidade + Ciclo + Irregular

• Este tipo procedimento tem-se mostrado bastante útil na prática, fornecendo subsídios para a resposta de várias perguntas de interesse na modelagem de séries reais.

• <u>Macroeconomia- I:</u> é usual tentar separar as séries de produção (e.g., PIB, PDB) em duas componentes:

tendência + irregular/ciclo.

• Muitos economistas acreditam que superposta às flutuações de curta duração na atividade econômica, a economia evolui ao longo de um caminho de crescimento, o qual pode ser pensado como a tendência.

• Para facilitar o entendimento desta questão, pode-se pensar a economia como sendo afetada por dois tipos de choques:

> $\Rightarrow$ **choques permanentes** = possuem efeito permanente na produção: aumento de produtividade, desenvolvimento tecnológico, aumento da força de trabalho, aumento do nível educacional,etc.
>
> $\Rightarrow$ **choques transientes** = possuem efeito passageiro na produção: choques fiscais (diminuir impostos) e monetários (taxa de juros), greves, etc.

• Assim, nesta visão, a tendência seria a parte da produção econômica associada com choques permanentes, e seria não estacionária, por construção.

• A parte da produção associada aos choques transitórios seria o ciclo, estacionário por construção.

• A componente cíclica de séries macroeconômicas contém as frequências que possuem período identificado como pertencentes a “ciclos econômicos” típicos.

• Estes períodos se situam entre 6 e 32 trimestres, isto é, entre 1.5 e 8 anos, isto é , com freqüências no intervalo
$2\pi/6 < \omega < 2\pi/3$.

•Portanto as técnicas/modelos de extração da componente cíclica de séries macroeconômicas devem deixar passar freqs. nesta banda.

• Uma vez determinado um modelo de ST que decompõe a série do PIB em tendência e ciclo, pode-se, tentativamente, responder à seguinte questão: qual a componente que explica a maior parte das variações do PIB ?

$\Rightarrow$ se a componente cíclica no PIB não existir ou for desprezível, o governo não deve se preocupar com as políticas de curto prazo (choques fiscais e monetários), devendo-se concentrar nas políticas de longa duração (educação, etc).

<u>Macroeconomia- II</u>: a dinâmica de muitas séries macroeconômicas (produção, desemprego, M1, etc) apresenta variação sazonal, que é um tipo de flutuação já esperada. Muitas vezes é de interesse obter uma estimativa da série “descontada” da sazonalidade, de forma a fornecer uma direção “clara” do movimento da série. Para operacionalizar esta estimação precisamos de um modelo c/ a seguinte decomposição:

tendência + sazonalidade + irregular

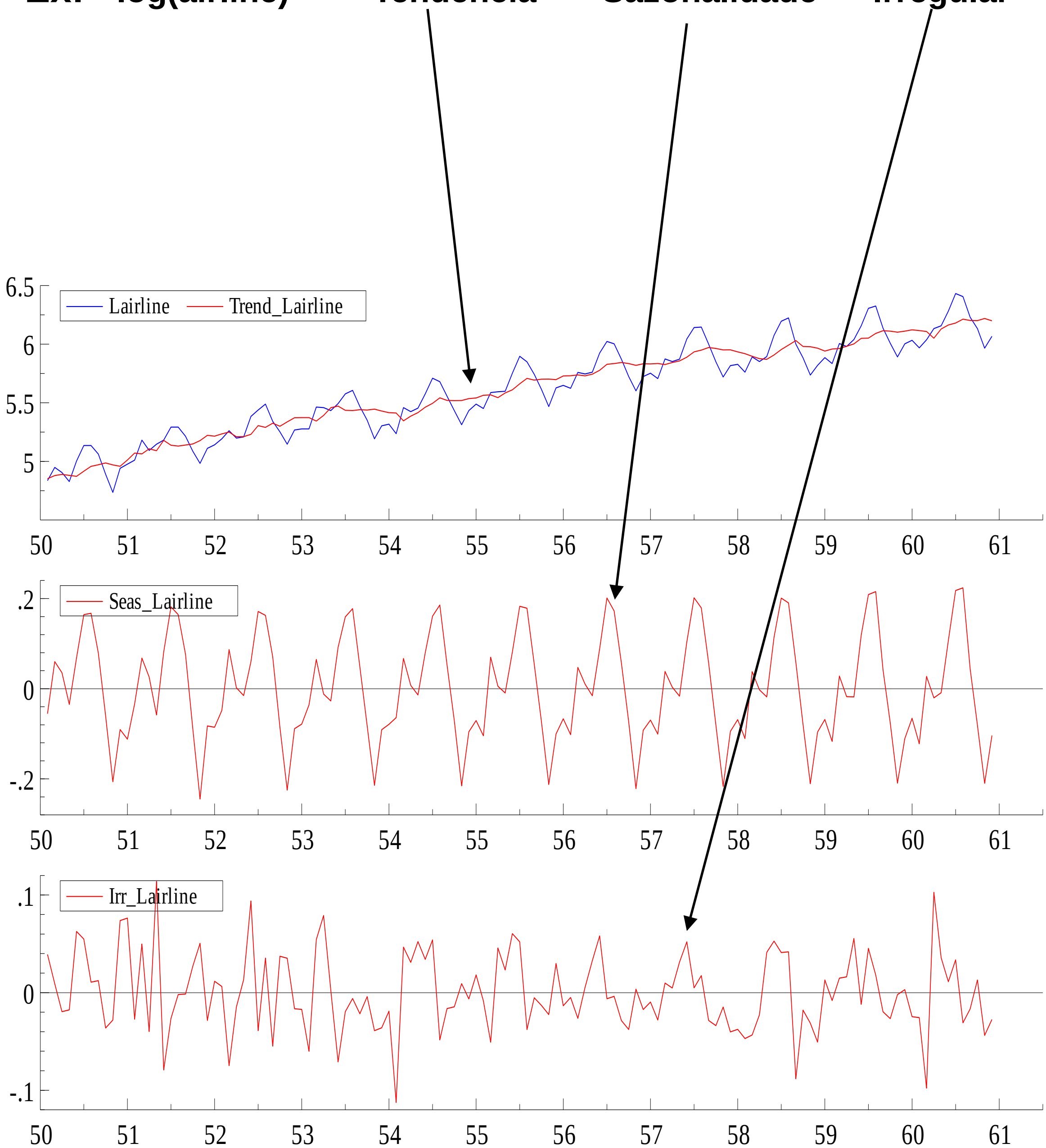
Meteorologia: é sabido que muitas séries de chuvas possuem ciclo. Por exemplo, Fortaleza, c/ ciclo com período entre 11 e 13 anos.

$\Rightarrow$ Para respondermos a perguntas do tipo: ” o próximo ano será de seca severa ?” necessitaríamos de um modelo com a seguinte decomposição:

tendência(?) + ciclo + irregular

• A modelagem de ST de CNO permite-nos construir uma resposta para estes tipos de indagações.

Ex: log(airline) =   Tendência   + Sazonalidade + Irregular



•As componentes de uma ST são também denominadas de fatos estilizados.

• Uma dada ST pode possuir apenas alguns dos fatos estilizados listados.

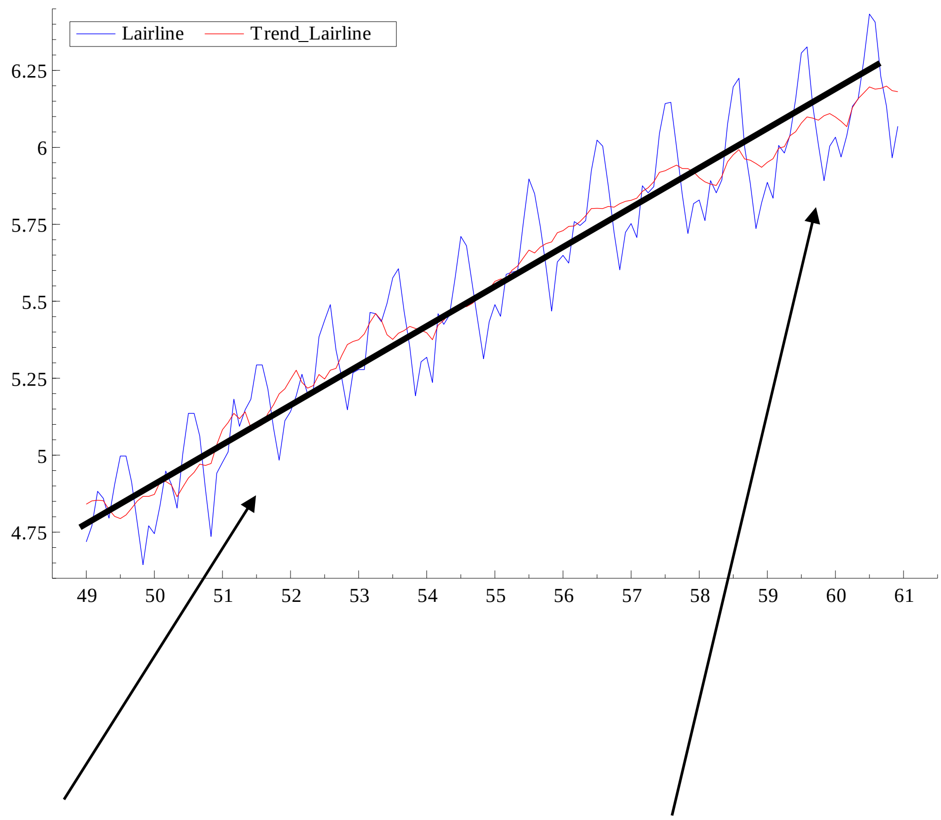
• Estas componentes, em princípio, são estocásticas, ou seja, evoluem probabilisticamente ao longo do tempo. São também conhecidas por componentes locais.

• As componentes estocásticas contrapõe-se às componentes deterministas, cuja forma permanece inalterada ao longo do tempo, sendo por isso denominadas de componentes globais.

• Estudos empíricos (Nelson & Plosser, 1982) sugerem que, pelo menos para séries macroeconômicas, é mais adequado considerar que as componentes sejam estocásticas (raiz unitária).

• Nos modelos estruturais (ME) as componentes são estimadas recursivamente através de um algoritmo denominado de filtro de Kalman (FK), ou de forma mais completa utilizando os algoritmos de suavização.

Exemplo de séries com componentes determinísticas e estocásticas: log(airline)



tendência determinística **tendência estocástica**

$\Downarrow$

$\ln(y_t) = a + bt$

$\Downarrow$

modelos estruturais

• Observar os seguintes pontos nos modelos de CNO:

$\Rightarrow$ a natureza empírica das definições das componentes:

ex: o tamanho da ST pode determinar o que pode ser reconhecido como ciclo ou tendência;

$\Rightarrow$ o significado e a definição de uma componente depende da definição das outras componentes:

ex: a estimação da tendência necessita de um bom modelo para o ciclo e vice-versa;

$\Rightarrow$ a idéia defendida por alguns de que diferentes componentes presentes em uma ST estariam associadas à diferentes forças causais:

ex: as forças econômicas que impulsionam a tendência do nível da atividade econômica são independentes daquelas que criam as flutuações cíclicas.

$\Rightarrow$ a decomposição não é única, i.e., dado uma ST não- estacionária existem vários procedimentos de decomposição em tendência estocástica e ciclo, e os componentes efetivamente estimados por cada um destes procedimentos produsem diferentes estimativas dos componentes !

• Na literatura estatística contemporânea os modelos/métodos p/ ST baseados em CNO possuem a seguinte cronologia (não exaustiva):

1958- método de amortecimento exponencial de Holt (EWMA);
1960- método de Holt-Winters p/ ST com tendência e saz.;
1963- modelo de Brown;
1976- modelos estruturais Bayesianos de Harrison-Stevens;
1979- modelos de CNO de Nerlove e Carvalho;
1983- modelos estruturais clássicos de Harvey;
1988- modelos de Young;
1990 em diante: modelos não-lineares/não-Gaussianos, utilizando métodos computacionalmente intensivos (MCMC, amostragem por importância, etc) p/ a sua estimação.

2. MODELOS DE TENDÊNCIA

2.1 Definição:

• O que é a tendência de uma ST ? várias definições...

- parte da série que “muda pouco ao longo do tempo”;

- componente de baixa freqüência, que apresenta maior suavidade quando estimada;

- *a trend is a trend is a trend*
but the question is, will it bend ?
will it alter its course
through some unforessen force,
and come to a premature end ?

Sir Alec Cairncross (1969)

• A tendência de ST tem sido estimada através de:

- filtros médias móveis;
- regressão determinista em uma função do tempo;
- modelos estocásticos.

• Inicialmente apresentaremos as principais distinções entre os modelos de tendência determinista e estocástica.

2.2 Classificação dos Processos de Tendência (Econometria)

Determinística: o modelo da tendência é uma função determinista do tempo(TS);

Tendência

Estocástica: o modelo da tendência é uma função estocástica (DS).

•**Tendência Determinista ou processo *trend stationary (TS)*:** definido pela seguinte equação:

$$y_t = \sum_{j=0}^{d} \beta_j \; t^j + \psi(L)a_t,\; a_t \sim \; NID(0,\sigma^2)$$

onde $\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j$

Mostra-se que: $E(y_t) = \sum_{j=0}^{d} \beta_j \; t^j$

$$Var(y_t) = \left(\sum_{j=0}^{\infty} \psi_j^2 \right) \sigma^2, \sum_{j=0}^{\infty} \psi_j^2 < \infty$$

$$\gamma(k) = \sigma^2 (\psi_k \psi_0 + \psi_{k+1}\psi_1 + \psi_{k+2}\psi_2 + \dots)$$

•Caso particular: considere d=1 e a_t ~AR(1):

$$y_t = \beta_0 + \beta_1 t + a_t ,$$
$$a_t = \phi a_{t-1} + \varepsilon_t$$
$$|\phi| < 1, \varepsilon_t \sim NID(0, \sigma^2)$$

• Segue que:

$$E(y_t) = \beta_o + \beta_1 t$$
$$Var(y_t) = \sigma^2 / (1-\phi^2)$$
$$\rho(y_t \; y_{t-k}) = \phi^k$$

• Previsão s-passos à frente:

$$y_t = \beta_o + \beta_1 t + e_t$$
$$y_{t+s} = \beta_o + \beta_1 (t+s) + e_{t+s}$$
$$\hat{y}_{t+s|t} = E(y_{t+s} \,|\, Y_t) = \beta_o + \beta_1 (t+s) + E[e_{t+s} \,|\, Y_t)]$$

$$\hat{y}_{t+s|t} = \beta_o + \beta_1 (t+s) + e_t \phi^s$$
$$= (y_t + \beta_1 s) + e_t (\phi^s - 1)$$

$$Var(e_{t+s|t}) = \sigma^2 (1\text{-}\phi^{2s})/(1\text{-}\phi^2) \qquad s = 1, 2, \ldots$$

• Para retirar a tendência do processo efetuamos uma regressão por MQO: a série s/ tendência será o resíduo da regressão

$$y_t^* = y_t - (\hat{\beta}_0 + \hat{\beta}_1 t) = \hat{a}_t$$

• Se tivéssemos retirado a tendência por primeira diferença:

$$w_t = y_t - y_{t-1} = \beta_1 + a_t - a_{t-1}$$
$$(1-L)(1-\phi L)y_t = \beta_1(1-\phi) + (1-L)\varepsilon_t$$
$$\Leftrightarrow y_t \sim ARIMA(1,1,1)\ com\ \ \theta = 1.$$

• Portanto, a operação de filtrar a tendência por primeiras diferenças não será adequada se a tendência for determinista, (TS) pois implicará em um modelo MA não invertível, o que não é desejável do ponto de vista estatístico (pq ?)

• De um forma geral, p/ um processo TS genérico, com polinômio de grau d, necessitaríamos de diferenciar a série d vezes, introduzindo assim um processo MA c/ d raízes unitárias

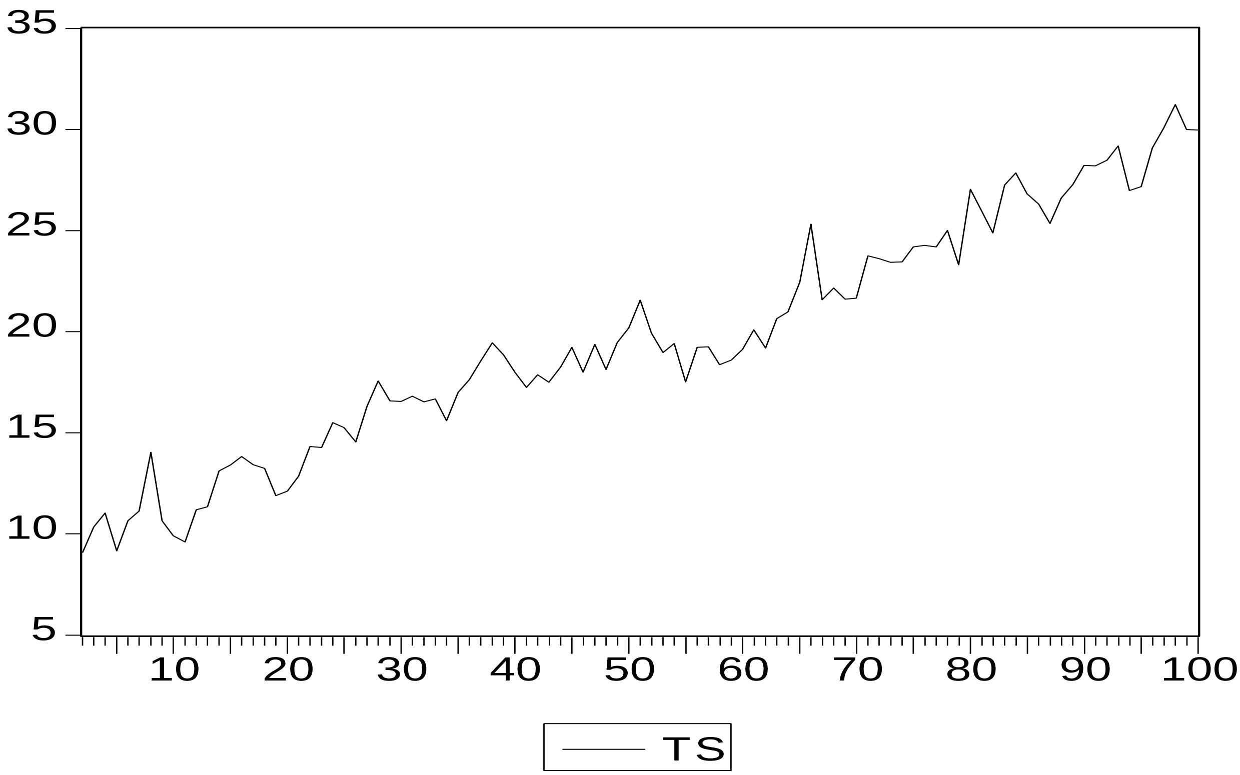
$$y_t = \sum_{j=0}^{d} \beta_j t^j + \frac{\theta(L)}{\phi(L)} a_t \ \ então$$

$$\nabla^d y_t = \theta_0 + \nabla^d [\theta(L)/\phi(L)]a_t$$

$$\theta_0 = d!^* \beta_d .$$

- Exemplo de um processo TS sintético

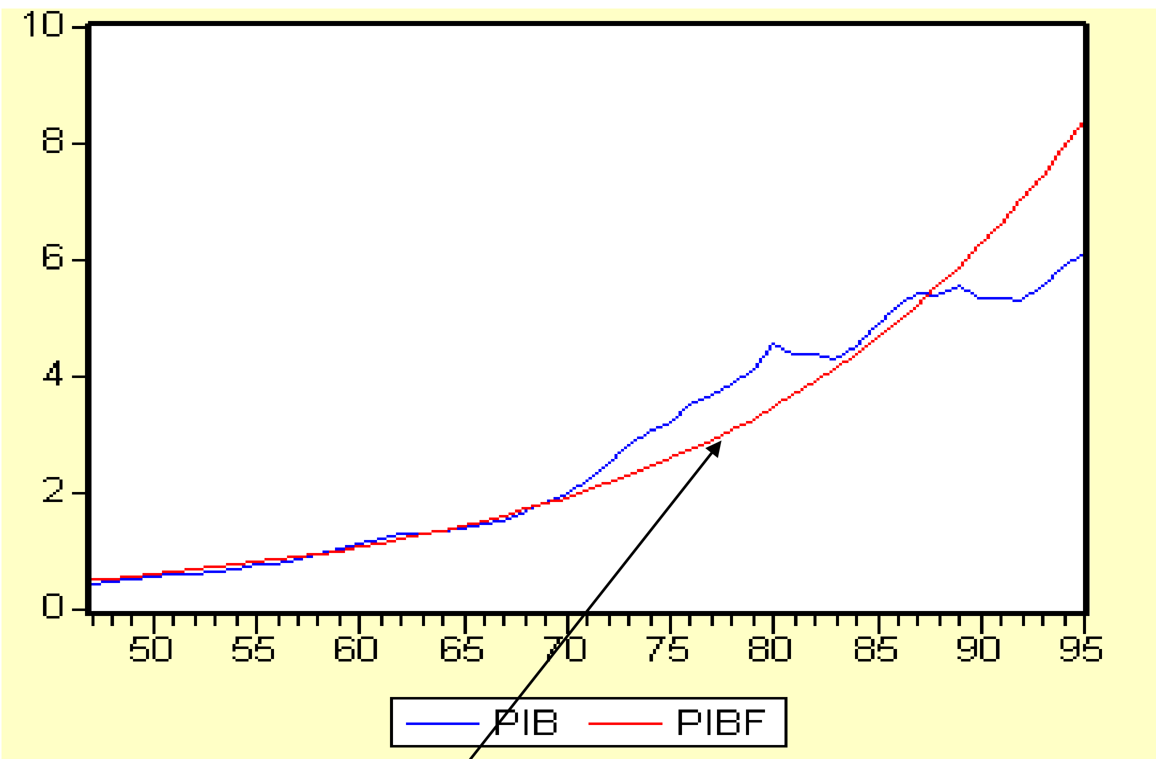
$$y_t = 10 + 0.2t + a_t$$
$$a_t = 0.6a_{t-1} + e_t, e_t \sim NID(0,1)$$



Exemplo de um processo TS ajustado à uma série com tendência estocástica

PIB anual brasileiro (1947-1995), preços de 1980

PIB = a exp(bt) exp(a_t), a_t~ N(0,σ^2)
ln(PIB) = - 0.711 + 0.0591 t
t (-16.770) (38.799)
R^2= 0,97



tendência determinística

• Os processos TS foram utilizados em muitos trabalhos nas décadas de 70 e 80 p/ estimar a componente cíclica de agregados macro-econômicos:

$$y_t = a + bt + a_t, a_t \sim ARMA(p,q)$$

• Após estimado por MQO, o ciclo seria identificado com o resíduo do modelo.

• Problemas neste procedimento:

- a componente cíclica fica super-dimensionada, podendo gerar ciclos espúrios;

- no processo TS os choques possuem efeito transiente, i.e, somente afetam o processo no tempo t.

- como o impacto das inovações tecnológicas na tendência possui efeito permanente, este só poderá ser capturado por uma tendência estocástica, ou DS.

•**Tendência Estocástica ou processos *difference stationary (DS):***

i. **passeio aleatório (random walk):** exemplo canônico de um processo não estacionário estocástico. Ex: série de preços de ativos na bolsa. Seja ε_t~NID(0,σ^2).

$$y_t = y_{t-1} + \varepsilon_t \ \ (AR \ \ com \ \ \phi = 1)$$

$$y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i$$

• Donde segue que:

$$E(y_t) = 0, \ se \ \ y_0 = 0.$$

$$\rho(y_t \ y_{t-k}) = \left(1 - \frac{k}{t}\right)^{1/2}$$

$$Var(y_t) = t\,\sigma^2$$

• Para retirar a tendência de um processo DS efetuamos um número de diferenciações adequadas:

$$\Delta y_t = \varepsilon_t .$$

• Para melhor entendermos a natureza do processo de diferenciação, um pouco mais de formalismo.

- Operadores:

  - L : operador de atraso $L^k y_t = y_{t-k.}$

  - 1^a diferença: $\Delta = 1 - L$: operador de primeira diferença

$$\Delta y_t = (1 - L)\, y_t = y_t - y_{t-1}$$

  - d diferenças: $\Delta^d = (1 - L)^d$, $d = 1,2,3,...$

d = a ordem de integração da série: número de diferenciações para tornar a série estacionária.

se $y_t \sim I(d) \therefore \Delta^d y_t$ é estacionária

Ex: - Ibovespa ~ I(1)
- retornos Ibovespa ~ I(0)

ii. **passeio aleatório c/ *drift*:** Ex: PIB, M1 ($ em circulação), preços no mercado financeiro.

$$y_t = a_0 + y_{t-1} + \varepsilon_t$$

$$y_t = y_0 + a_0 t + \sum_{i=1}^{t} \varepsilon_i$$

$$E(y_t) = y_0 + a_0 t \qquad \rho(y_t\ y_{t-k}) = \left(1 - \frac{k}{t}\right)^{1/2}$$

$$Var(y_t) = t\sigma^2.$$

- Vamos agora projetar a equação s passos à frente

$$y_{t+s} = y_0 + a_0(t+s) + \sum_{i=1}^{t+s} \varepsilon_i$$

$$y_{t+s} = \left(y_0 + a_0 t + \sum_{i=1}^{t} \varepsilon_i\right) + a_0 s + \sum_{i=t+1}^{t+s} \varepsilon_i$$

$$y_{t+s} = y_t + a_0 s + \sum_{i=t+1}^{t+s} \varepsilon_i$$

• Assim sendo num processo de passeio aleatório, que é um processo de tendência DS, os choques possuem efeito permanente, pois:

$$\frac{\partial y_{t+s}}{\partial \varepsilon_i} = 1, \quad i = 1,2, \ldots, t + s - 1$$

• Portanto um choque no passado não é esquecido no futuro, sendo seu efeito permanente. Contraste este resultado com o de um processo AR(1).

• Previsão s passos à frente

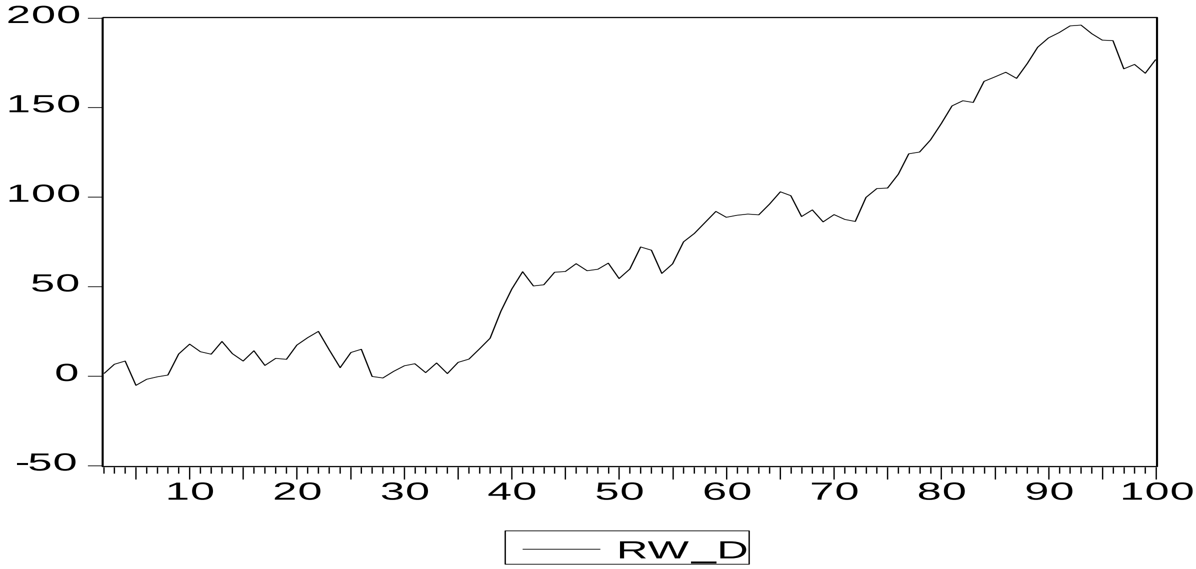
$$\hat{y}_{t+s|t} = E(y_{t+s} \mid Y_t) = y_t + a_0 s$$

$$Var(e_{t+s/t}) = Var\left(\sum_{i=t+1}^{t+s} \varepsilon_i\right) = s\sigma^2 \qquad s = 1, 2, \ldots$$

• Exemplos de séries sintéticas:

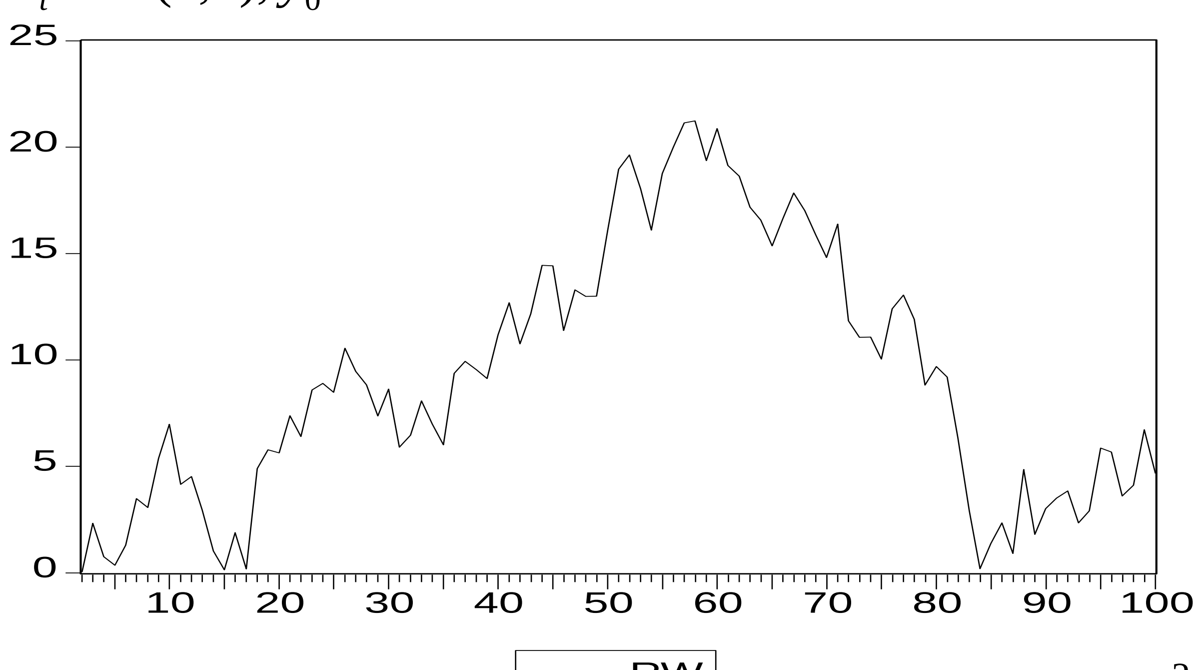
$$y_t = 2 + y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim N(0{,}49), y_0 = 1$$

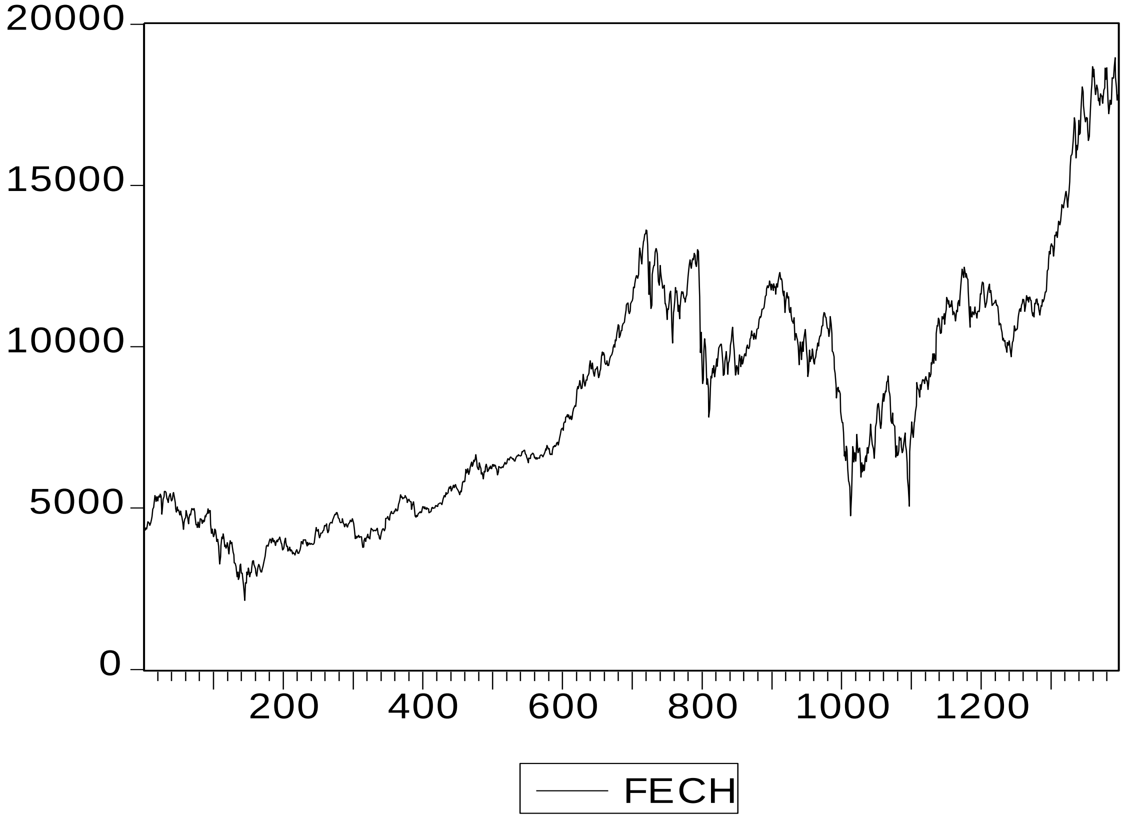


$$y_t = y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim N(0{,}4), y_0 = 1$$



• Exemplo de série real descrita por um processo TS: cotação de fechamento do Ibovespa (01/08/94 à 31/03/2000)



Ajuste de um modelo AR(1) c/drift

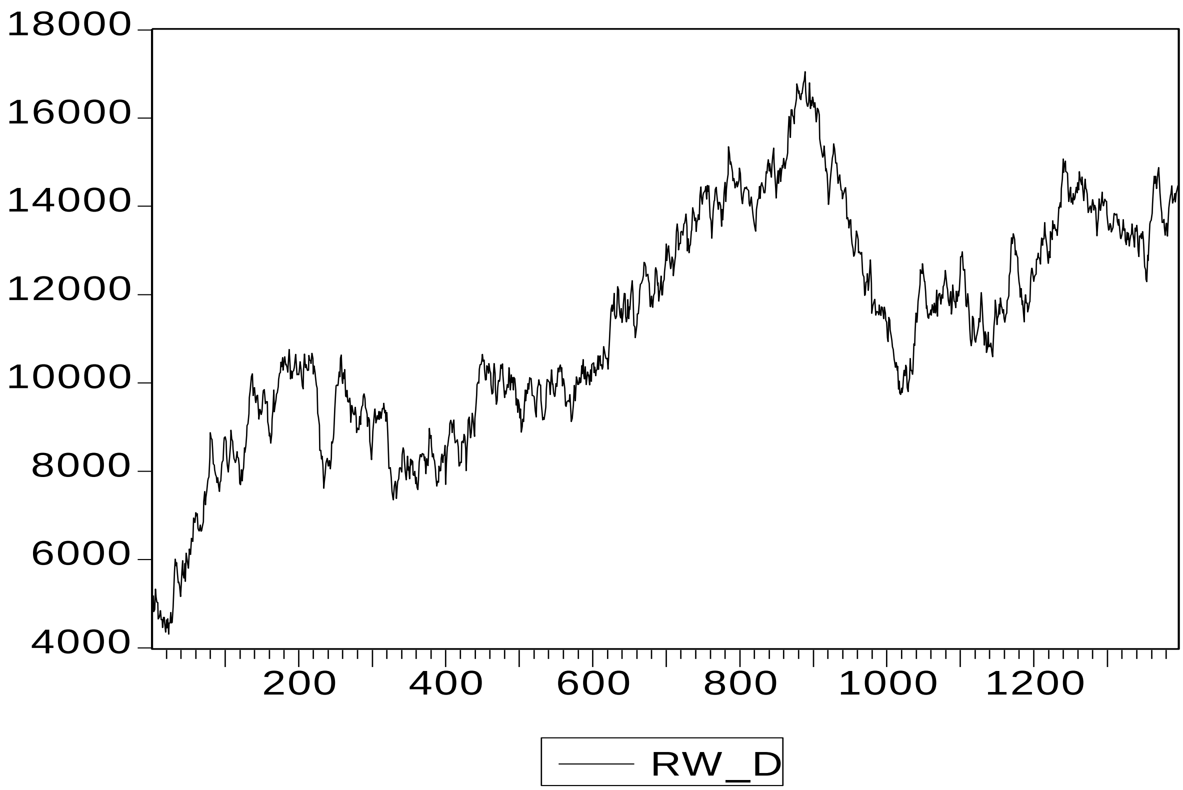
$Ibov_t' = 9.96 + 0.9998\ Ibov_{t-1} \quad \sigma' = 230{,}86$

Teste de Raiz unitária : ADF Test Statistic -1.586000
1% Critical Value -3.9698
5% Critical Value -3.4155
10% Critical Value -3.1296

Realização do processo TS com parâmetros análogos ao do modelo estimado para a série do Ibovespa

$$y_t = 10 + y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim N(0{,}231), y_0 = 5000$$



Usando uma regressão TS para estimar um processo DS: os resultados de Nelson & Kang (1984):

- **Experimento:**

$\Rightarrow$ gerou a série através de um DS:

$$y_t = y_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim N(0, \sigma^2)$$

$\Rightarrow$ estimou por um TS:

$$y_t = a + bt + u_t$$
$$u_t \sim N(0, \sigma_u^2)$$

- **Resultados:**

i. se gerarmos uma ST RW , e regredirmos essa ST no tempo, usando MQO, o R^2 será de aprox. 0.44, independente do tamanho da série (T). Este resultado é totalmente espúrio pois o RW não depende explicitamente do tempo.

ii. se o processo for um RW + drift, o R^2 da regressão no tempo será ainda maior, e crescerá com T, alcançando o valor limite de 1 !

iii. os resíduos de uma regressão de um RW no tempo tendem a apresentar autocorrelações significativas p/ vários lags, com comportamento pseudo-cíclico de período aprox. 2/3 T! Estas autocorrelações são totalmente espúrias, e dependem de T.

Ex: n=100, r=1000: r(1)=0.88, r(2)=0.77, r(3)=0.68

$\Rightarrow$ podem sugerir um processo AR/ciclo (possivelmente não estacionário) p/ a série sem tendência ! O ciclo assim detectado é pura conseqüência da metodologia estatística em uso, não possuindo uma justificativa fundamental (efeito de Slutzky). Este efeito é também observado em outras metodologias de separação de componentes.

Obs: se um dado processo é DS, então a eliminação da tendência por regressão no tempo, não produzirá um processo estacionário. Isto é mais facilmente detectado para um processo RW + drift:

$$y_t = a_0 + y_{t-1} + \varepsilon_t$$

$$y_t = y_0 + a_0 t + \sum_{i=1}^{t} \varepsilon_i \text{. Retirando a tendência:}$$

$$z_t = y_t - a_0 t = y_0 + \sum_{i=1}^{t} \varepsilon_i \text{, onde}$$

$$E(z_t) = y_0, \quad \text{mas}$$

$$\mathrm{Var}(z_t) = t\sigma^2.$$

• Ou seja, os resíduos resultantes do ajuste de um processo DS por uma tendência determinística, não serão estacionários. Mais uma razão para que a eliminação da tendência em processos DS seja realizada através de diferenciação.

iv. o uso das estatísticas t's para testar a presença de dependência explícita de um RW no tempo, também produzirá resultados espúrios:

$\Rightarrow$ o teste rejeitará a hipótese de não relação com o tempo em 87% dos casos, para uma amostra de tamanho 100 e valor nominal de 5%, qdo na verdade não existe nenhuma relação !

• Em princípio, a identificação de uma tendência como TS ou DS pode ser realizada através de:

i. um teste de RU, embora, geralmente, estes testes não possuam poder elevado.

ii. ajustando um modelo estrutural com tendência estocástica e observando, pelos valores das variâncias dos erros, se a tendência é do tipo determinística. (a ser visto).

• O teste de raiz unitária para distinguir tendências DS de TS é realizado utilizando eqs. do tipo (teste ADF tipo iii):

$$y_t = \alpha + \rho\, y_{t-1} + \beta\, t + \varepsilon_t$$

$$\Delta y_t = \alpha + \gamma y_{t-1} + \beta\, t + \varepsilon_t, \quad \gamma = \rho - 1$$

Ho: $\gamma = 0, \beta = 0$ tendência DS
Ha: $\gamma < 0, \beta \neq 0$ tendência TS

Ho: $\gamma = 0, \beta \neq 0$ tendência mista DS e TS

• Na literatura econométrica/estatística existem vários procedimentos que produzem decomposições para ST univariadas em tendência estocástica e ciclo, cabendo citar entre eles:

- decomposição de Beveridge & Nelson (B&N) = apresenta correlação -1 entre tendência e ciclo.

- decomposição dos modelos estruturais= apresenta correlação 0 entre tendência e ciclo.

• A decomposição de B&N será apresentada nas próximas notas, mas não será cobrada como conteúdo deste curso.