

Redes Neurais

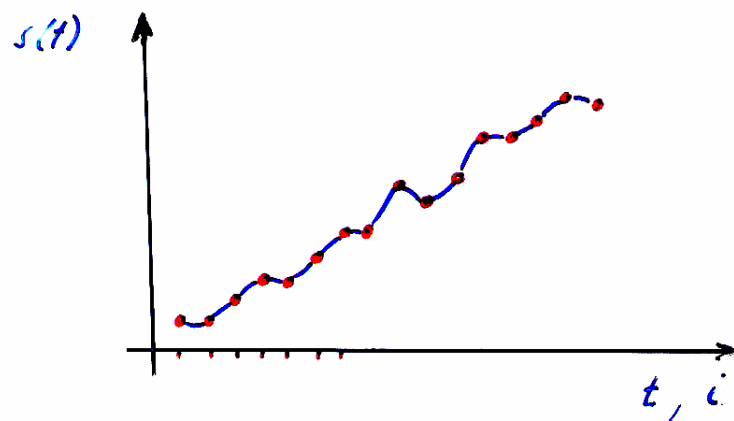
em Séries Temporais

Séries Temporais

$s(t)$ s_i

contínuas $t \in (t_0, t_1)$

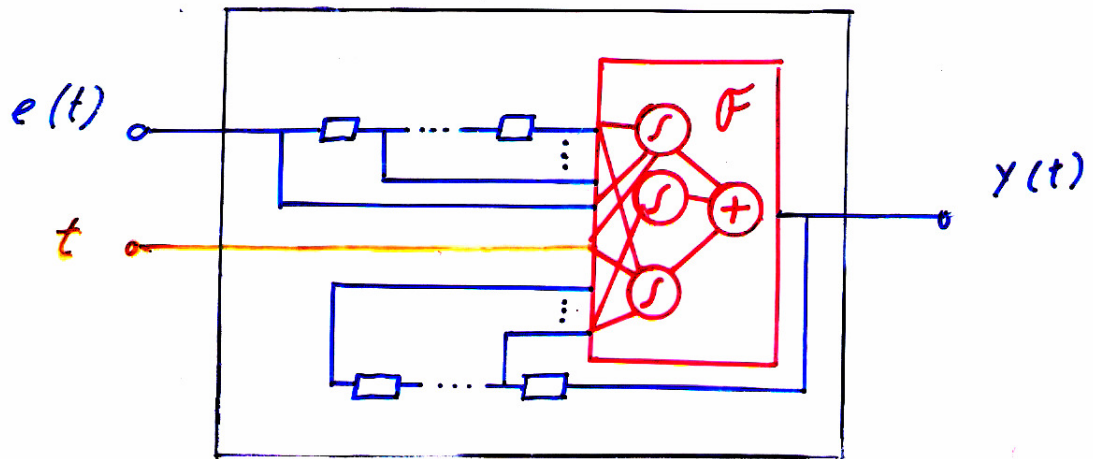
discretas $t, i \quad \{1, 2, 3, \dots, N\}$



discreta	i, t	$s(t), s_i$
	1	.5
	2	.5
	3	.9
	⋮	⋮

Sistemas Dinâmicos

(modelos ARMA e NARMA)



Sistema Linear $\Rightarrow F = \oplus$

Não Linear $\Rightarrow F = RN$, etc.

Sistema Invariante no Tempo

$F = e(t)e(t)$ $\Rightarrow t$ desnecessário

Resposta impulsiva

$e(t) = \delta(t)$ e estado inicial nulo

ou

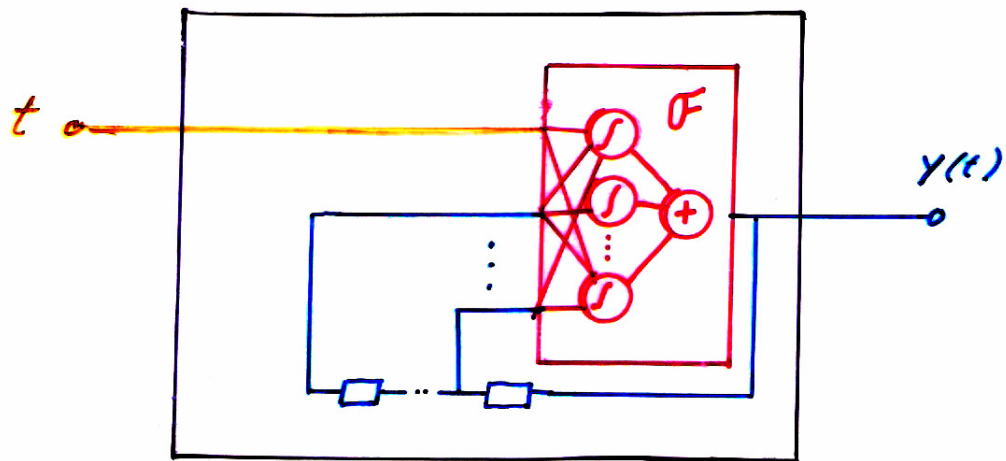
estado inicial $\neq 0$

$e(t) = 0$ desnecessário

Séries Temporais

Como

Resposta impulsiva de um sistema dinâmico



Série estacionária no tempo:

$t \rightarrow$ desnecessário

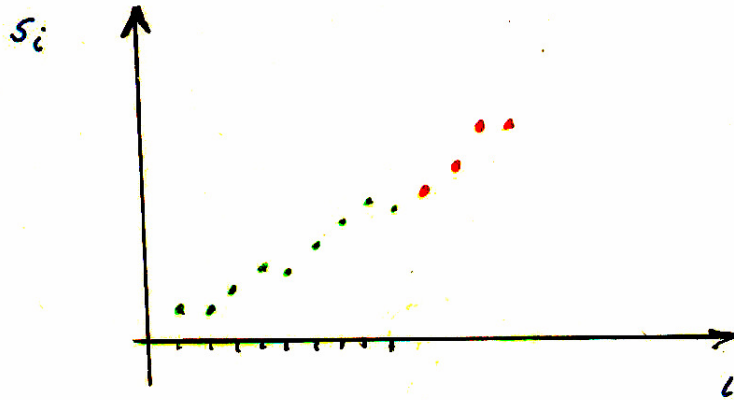
$$\sigma_y^2 = \text{cte}(t)$$

Sistema não dissipativo

↳ componentes periódicas

Previsão

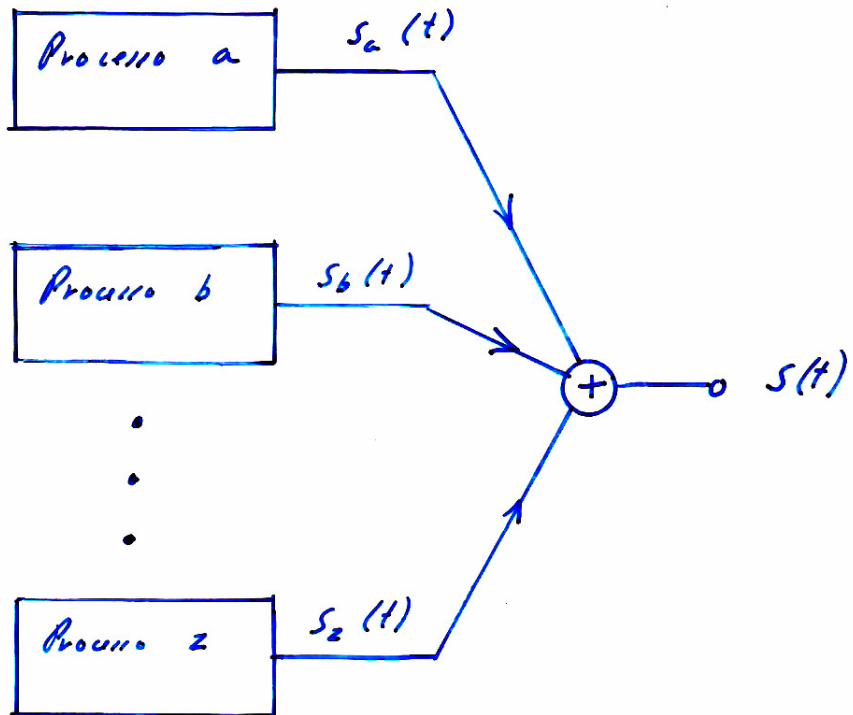
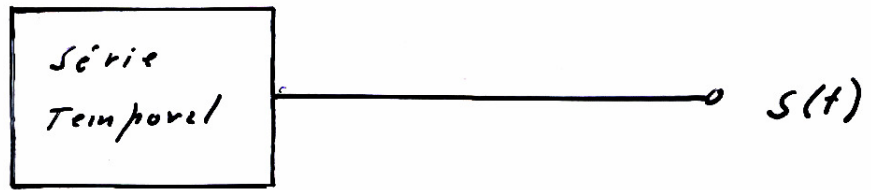
$$s_1 \quad s_2 \quad \dots \quad s_i \quad \xRightarrow{?} \quad \tilde{s}_{i+1} \quad \dots \quad \tilde{s}_{i+k}$$



$\Rightarrow$ Análise



Decomposição em
séries mais simples



$$S(t) = S_a(t) + S_b(t) + \dots + S_z(t)$$

Tipos de decomposição

Aditiva

$$s(t) = s_a(t) + s_b(t) + s_c(t) + s_r(t)$$

$$s_r = s - s_a - s_b - s_c$$

Multiplicativa

$$s(t) = s_a(t) \cdot s_b(t) \cdot s_c(t) \cdot s_r(t)$$

lg $\Rightarrow$ aditiva

Mista

$$s(t) = s_a(t) [s_b(t) + s_c(t)] + s_r(t)$$

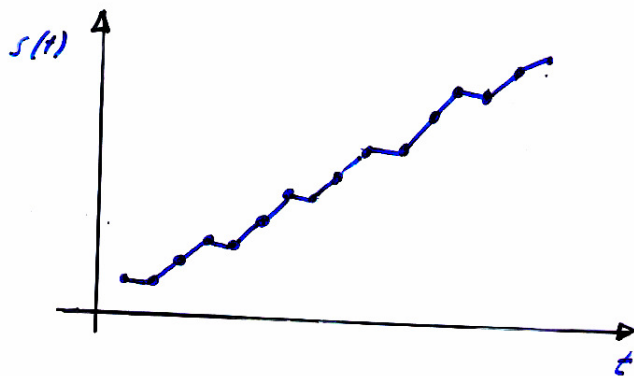
$\Rightarrow$ Decomposição aditiva

$$s_r(t) = s(t) - s_a(t) - s_b(t) - s_c(t)$$

Tipo de Componentes

como determinar?

1ª análise - subjetiva - plotar a série!



Componentes da decomposição

Tendência



$t(t)$

Sazonalidade

período T

$$\sum_{i=1}^T \left[\text{data } \frac{i}{T} \right]$$



Ciclo sazonal

$$a \cos[wt + \theta]$$

Termos não lineares

$n(t)$

ruido

$r(t)$

Quebras e outras irregularidades drásticas!

Série residual

$$s_r(t) = s(t) - \text{tend}(t) - \\ s_z(t) - cs(t)$$

$$= nl(t) + r(t)$$

estacionária no tempo?

ou o mais próximo possível disso!

Séries estacionárias no tempo

$$m_j(t) = c t t_e(t)$$

$$\left. \begin{array}{l} \mu(t) = c t t_e(t) \\ \sigma^2(t) = c t t_e(t) \end{array} \right\} \begin{array}{l} \text{fracamente} \\ \text{estacionária} \end{array}$$

$$m_3(t) = c t t_e(t)$$

$\vdots$

Decomposição:

$$S = S_e + S_{ne}$$

$$\Rightarrow S_1 = S - S_{ne}$$

$$S = S_{ne} + f(i) S_e$$

$$\Rightarrow S_1 = S - S_{ne}$$

$$S_2 = \frac{1}{f(i)} S_1$$

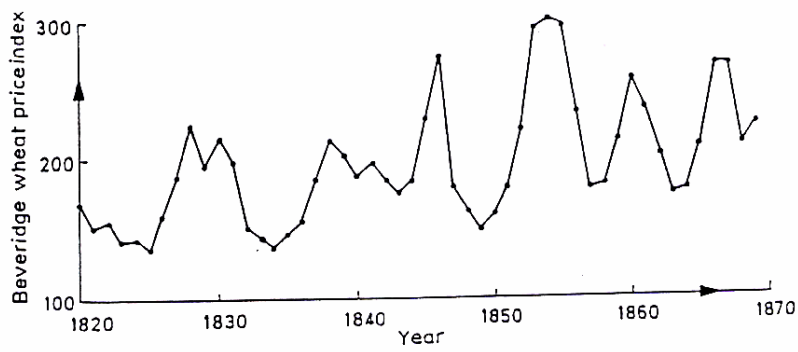


Figure 1.1 Part of the Beveridge wheat price index series.

trigo

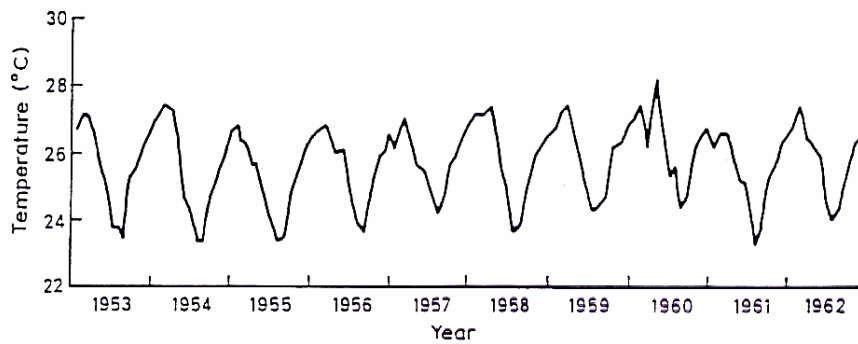


Figure 1.2 Average air temperature at Recife, Brazil, in successive months.

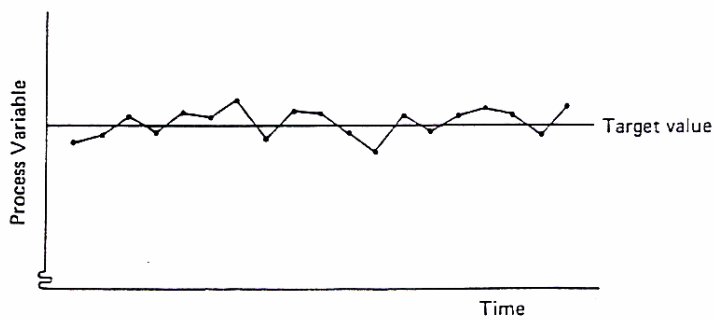


Figure 1.4 A process control chart.

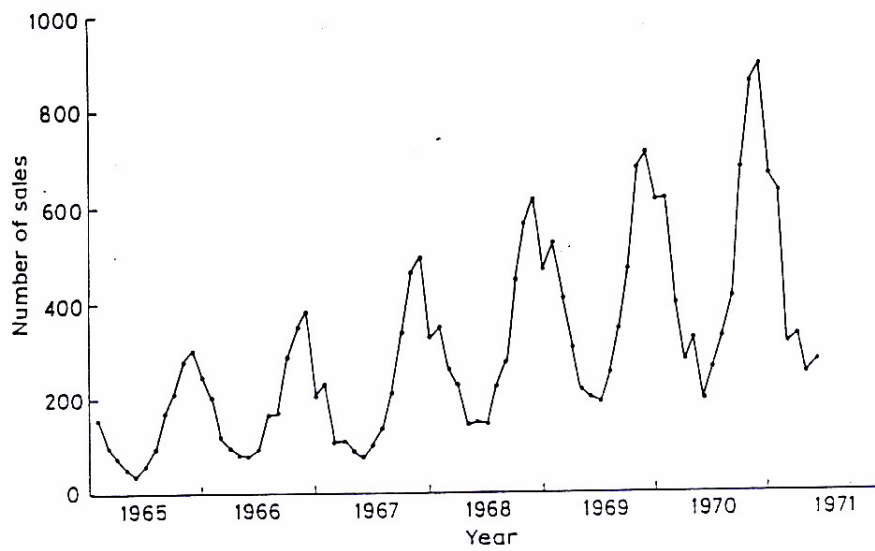


Figure 1.3 Sales of a certain engineering company in successive months.

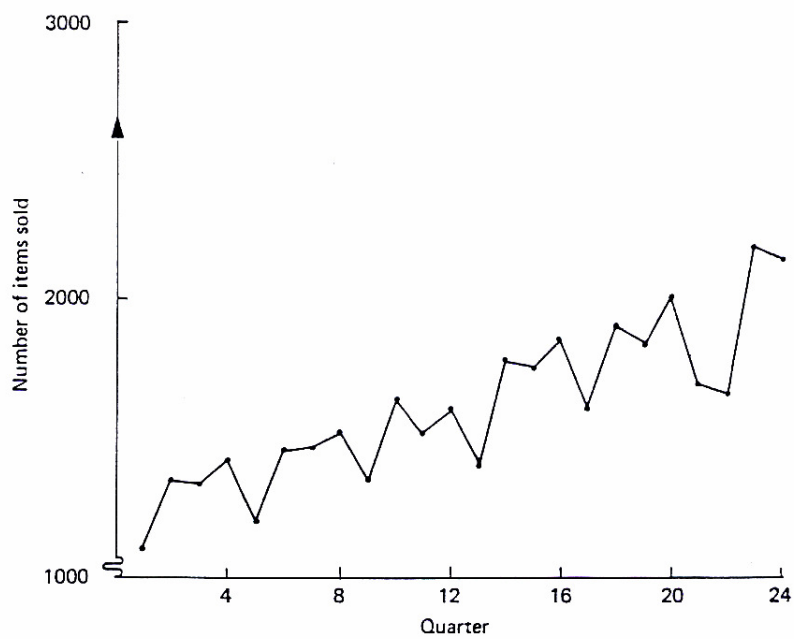
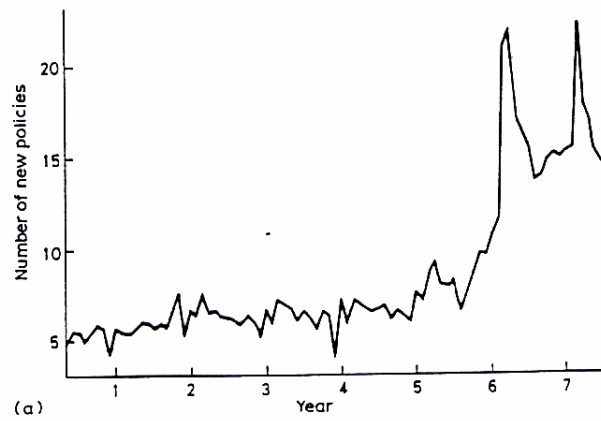
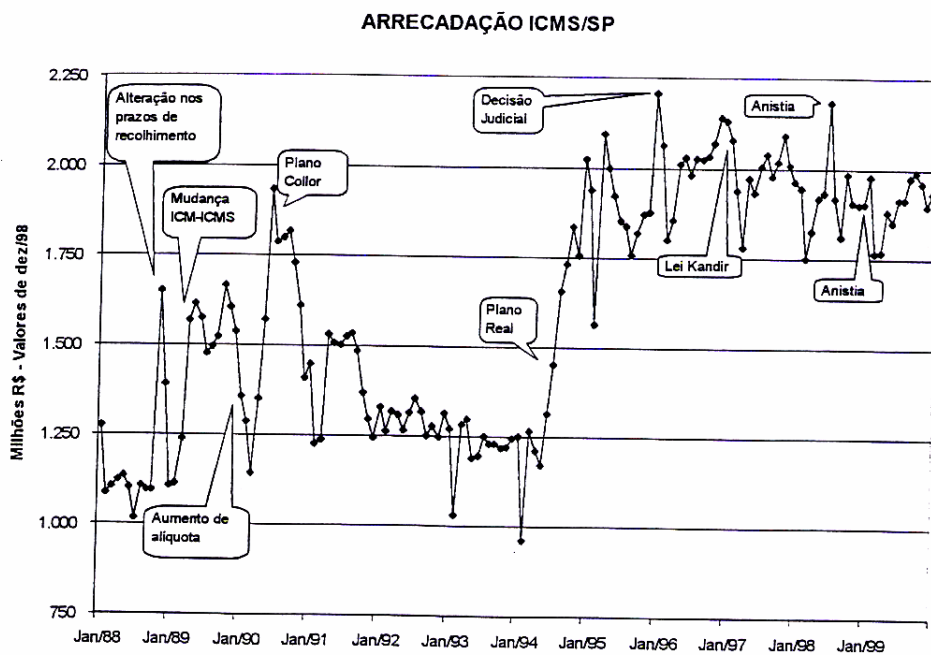


Figure 5.2 Sales of a certain company in successive three-month periods.



.. (a) Discontinuity present: numbers of new insurance policies issued by a particular life office (monthly in hundreds).



Análise no domínio do tempo

Ferramenta: Correlação

Séries $x_1 \quad x_2 \quad x_3 \quad \dots \quad E(x) = 0$

$y_1 \quad y_2 \quad y_3 \quad \dots \quad E(y) = 0$

$$\begin{aligned} r(x, y) &= \frac{\sum_{i=1}^N x_i y_i}{\left[\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i^2 \right]^{1/2}} \\ &= \frac{\frac{1}{N} \sum x_i y_i}{\sigma_x \sigma_y} \end{aligned}$$

Confiança

Sequências randômicas $\mu_r \approx 0$

$r \approx \text{normal}$ $\sigma_r = \frac{1}{\sqrt{N}}$

$$|r| > 2\sigma_r = 2/\sqrt{N}$$

$\Rightarrow 95\%$ confiança

Auto correlação

$$r_k = \frac{\sum_{i=1}^{N-k} x_i x_{i+k}}{\sum_{i=1}^N x_i^2} \frac{N}{N-k} \quad k > 0$$

Correlação cruzada

$$r_k(x, y) = \frac{\sum_{i=1}^{N-k} x_i y_{i+k}}{\left[\sum_{i=1}^{N-k} x_i^2 \sum_{i=k+1}^N y_i^2 \right]^{1/2}}$$

Exemplos

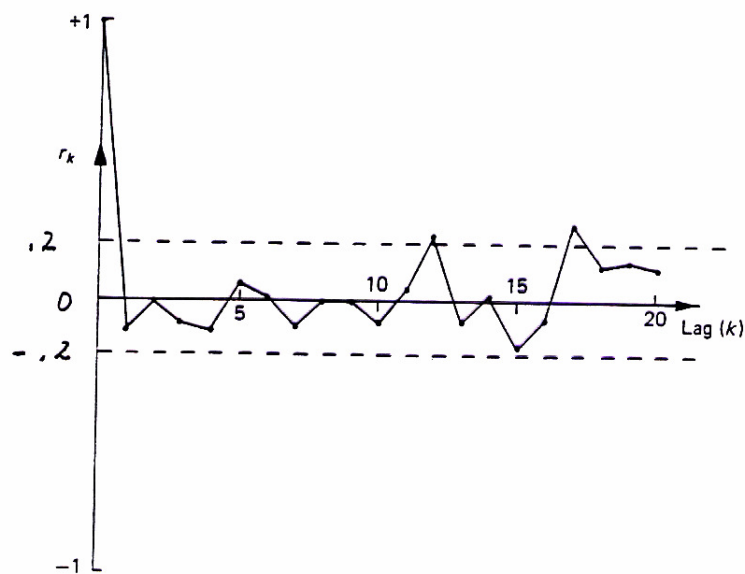


Figure 4.1 The correlogram of 100 'independent' normally distributed observations.

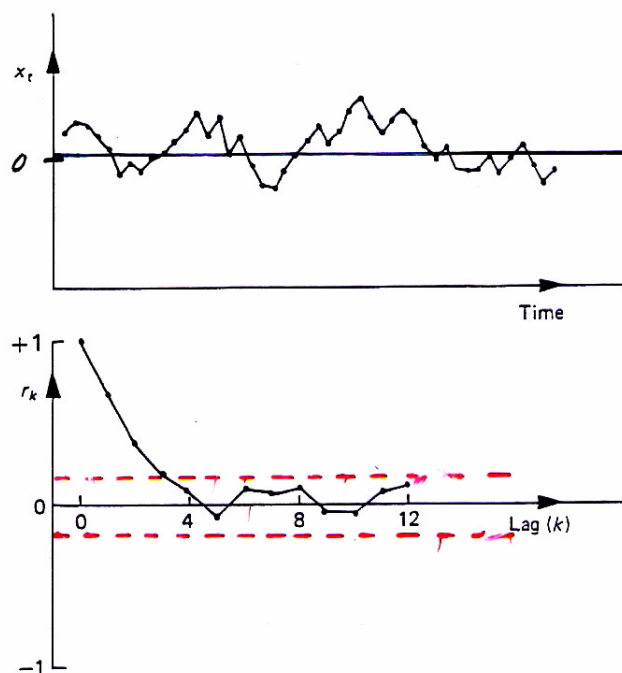


Figure 2.1 A time series showing short-term correlation together with its correlogram.

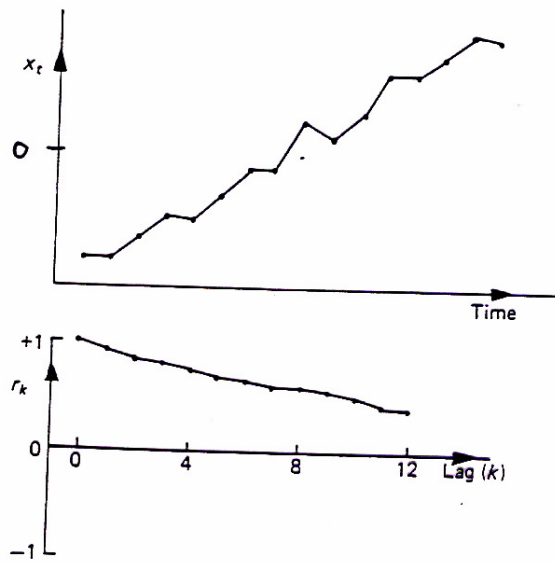


Figure 2.3 A non-stationary time series together with its correlogram.

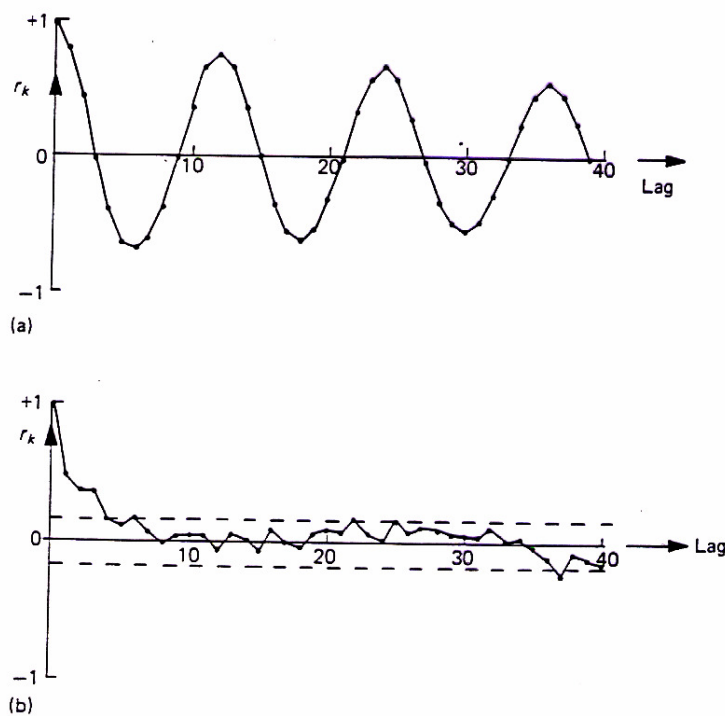


Figure 2.4 The correlogram of monthly observations on air temperature at Recife: (a) for the raw data; (b) for the seasonally adjusted data. The dotted lines in (b) are at $\pm 2/\sqrt{N}$. Values outside these lines are significantly different from zero.

Correlação parcial

$$x_i = \alpha_1 x_{i-1} + \dots + \alpha_p x_{i-p} + \text{res}$$

$$\mu = 0$$

$$p=1$$

$$\alpha_1 = r_1$$

$$p=2$$

$$\alpha_1 = r_1 (1 - r_2) / (1 - r_1^2)$$

$$\alpha_2 = (r_2 - r_1^2) / (1 - r_1^2)$$

$$p$$

$$\underline{R} = \begin{bmatrix} 1 & r_1 & r_2 & \dots & r_{p-1} \\ r_1 & 1 & r_1 & \dots & r_{p-2} \\ & & & & \\ r_{p-1} & r_{p-2} & & & 1 \end{bmatrix}$$

$$[R] \underline{\alpha} = r] \quad \therefore \underline{\alpha} = \dots$$

Análise no domínio da frequência

Ferramenta: Análise de Fourier,

Espectrograma, Periodograma

$$s(t) \quad t = 1, \dots, N \quad \Delta t = 1$$

$$\begin{aligned} s(t) &= a_0 + \sum_{i=1}^{\frac{N}{2}-1} \left[a_i \cos 2\pi \frac{i}{N} t + b_i \sin 2\pi \frac{i}{N} t \right] + \frac{a_N}{2} \cos \pi t \\ &= a_0 + \sum_{i=1}^{\frac{N}{2}-1} R_i \cos \left(2\pi \frac{i}{N} t + \theta_i \right) + \frac{a_N}{2} \cos \pi t \end{aligned}$$

$$\text{Frequências: } f_i = \frac{i}{N} = \frac{1}{N}, \frac{2}{N}, \dots, \frac{N}{2} \cdot \frac{1}{N}$$

$$\text{Resolução: } \Delta f = \frac{1}{N}$$

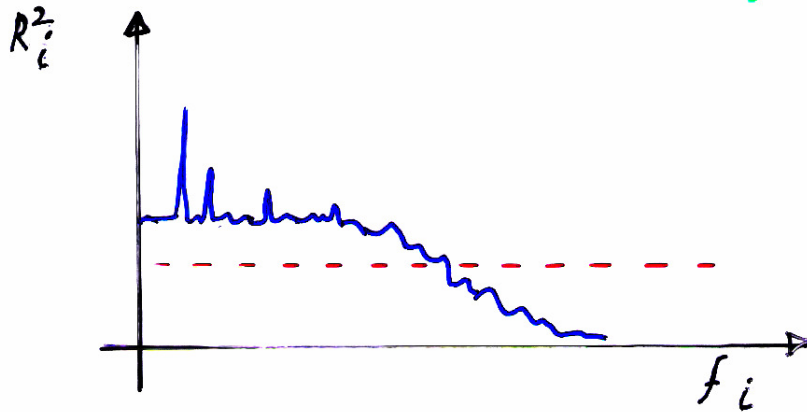
"Energia"

$$P(f_i) = R_i^2 = a_i^2 + b_i^2$$

Espectrograma

$$R_i^2 = a_i^2 + b_i^2 \quad i=1, \dots, \frac{N}{2}-1$$

$$f_i = \frac{1}{N}, \dots, \frac{1}{2} - \frac{1}{N}$$



$$\text{Resolução} \quad \Delta f = 1/N$$

Confiança

x_i randomica (ruído branco)

$$E(x) = 0 \quad \text{Pot}(x) = \sigma^2$$

$$\Rightarrow \frac{N R_p^2}{2 \sigma^2} = \chi_2^2 \quad \chi_2^2(.95) = 5.99$$

$$R_p^2 > \frac{12}{N} \sigma^2 \quad \Rightarrow \quad 95\% \text{ confiança}$$

Exemplos

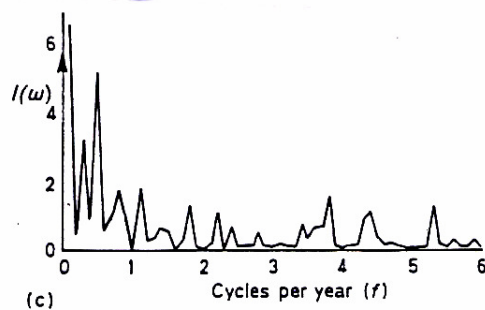
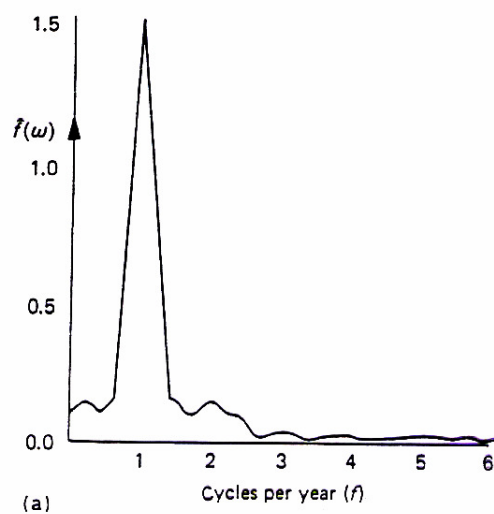
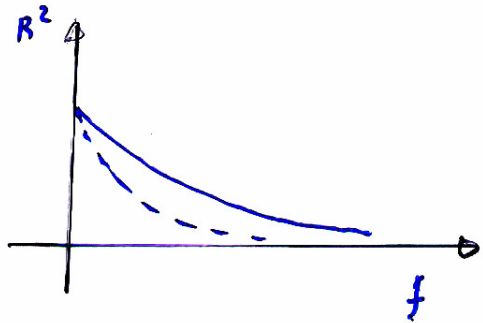
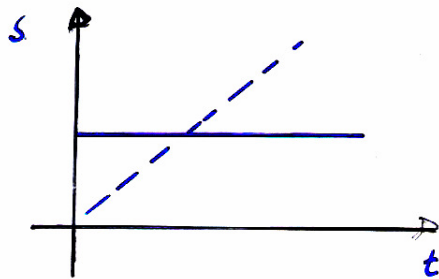


Figure 7.5 Spectra for average monthly air temperature readings at Recife, (a) for the raw data; (b) for the seasonally adjusted data using the Tukey window with $M=24$; (c) the periodogram of the seasonally adjusted data is shown for comparison.

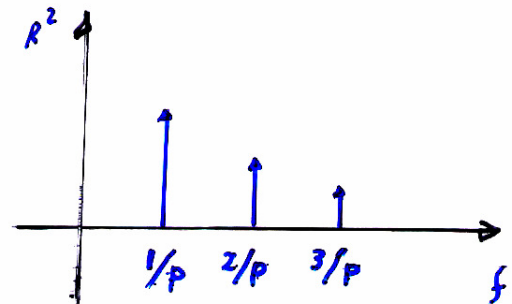
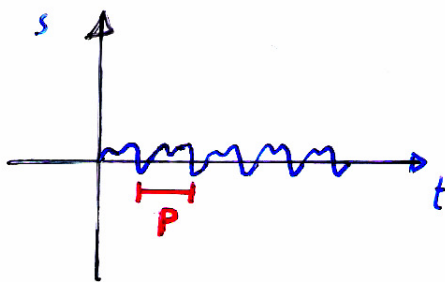
A B C D
M K J L

Espectros Típicos

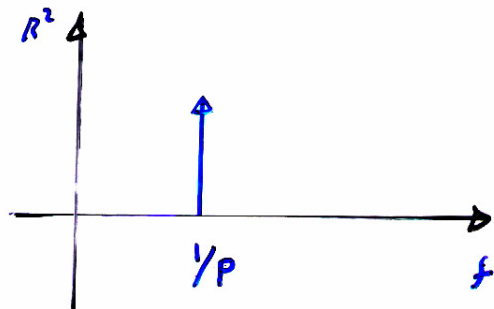
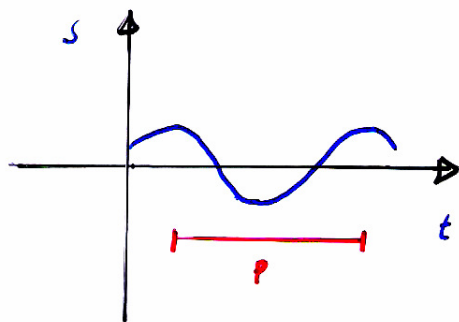
Média, Tendência



Sazonalidade



Ciclo Senoidal

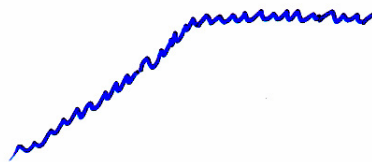


Decompondo a série

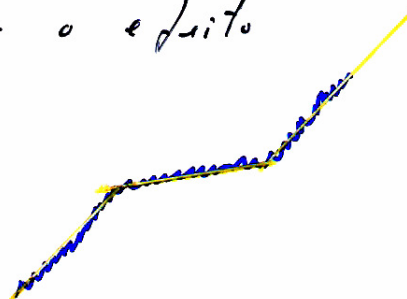
1 - Quebras & outros efeitos drásticos

Observar o plot da série

Restringir as regiões de trabalho



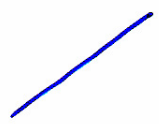
Corrigir o efeito

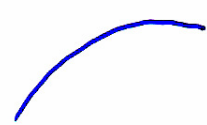


2 - Retirada da Tendência

Tipo? observar o plot da série!

ordem 0 $t_d(t) = a_0$ 

ordem 1 $t_d(t) = a_0 + a_1 t$ 

ordem 2 $t_d(t) = a_0 + a_1 t + a_2 t^2$ 

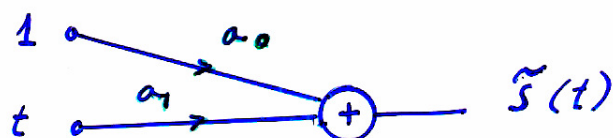
log

exp $t_d(t) = a_0 e^{a_1 t}$

Cálculo: ajuste por e.m.q.

$$\varepsilon(t) = s(t) - t_d(t)$$

$$F(a_0, a_1) = E[\varepsilon^2(t)] = \frac{1}{N} \sum_{t=1}^N (s(t) - (a_0 + a_1 t))^2$$



Tendência linear : solução analítica

$$s \approx \tilde{s} = t_g(t) = a_0 + a_1 t$$

$$F = E(s - \tilde{s})^2 = E(s - (a_0 + a_1 t))^2$$

$$\frac{\partial F}{\partial a_0} = -2 [E(s) - a_1 E(t) - a_0] = 0 \quad (1)$$

$$\frac{\partial F}{\partial a_1} = 2 [a_1 E(t^2) - E(ts) + a_0 E(t)] = 0 \quad (2)$$

$\Rightarrow$

$$a_1 = \frac{E(ts) - E(t)E(s)}{E(t^2) - [E(t)]^2}$$

$$a_0 = E(s) - a_1 E(t)$$

Retirada da tendência

$$s_r(t) = s(t) - t_g(t)$$

linear: $s_r(t) = s(t) - (a_0 + a_1 t)$

Treatamento via diferenciação

1. Processos com média móvel

$$s(t) = s_1(t) + a_0$$

SUAVE E NÃO LINEARMENTE
variante no tempo

$$s'(t) = \frac{ds}{dt} = \frac{d}{dt} s_1(t) + \epsilon$$

Tempo discreto

$$s'(t) \approx s(t+1) - s(t)$$



série à prever, SEM MÉDIA MÓVEL

(com mais ruído em HF)

Reconstrução (Taylor):

$$s(t + \Delta t) \approx s(t) + s'(t) \Delta t$$

$$s(t+1) \approx s(t) + \tilde{s}'(t)$$

↑
real

↑
previsto

Treatamento via diferenciação

1. Processos com média móvel

$$s(t) = s_1(t) + a_0$$



SUAVE E NÃO LINEARMENTE
variante no tempo

$$s'(t) = \frac{ds}{dt} = \frac{d}{dt} s_1(t) + \epsilon$$

Tempo discreto

$$s'(t) \approx s(t+1) - s(t)$$



série à prever, SEM MÉDIA MÓVEL

(com mais ruído em HF)

Reconstrução (Taylor):

$$s(t + \Delta t) \approx s(t) + s'(t) \Delta t$$

$$s(t+1) \approx s(t) + \tilde{s}'(t)$$

↑ real

↑ previsto

Correlograma

+

Espectrograma

+

Fenomenologia

$\Rightarrow$

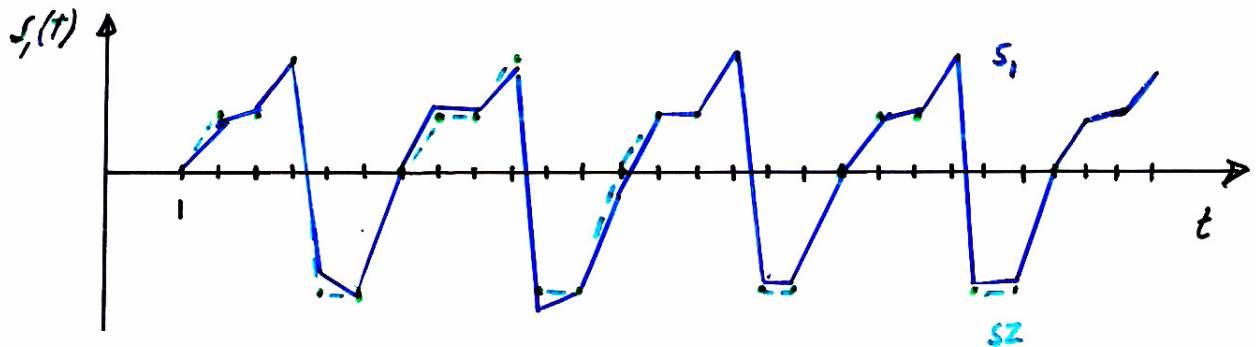
Sazonalidade

Componentes Senoidais

3 - Sazonalidade

Observar a forma temporal

Observar o espectrograma /
correlograma



Período $P = 6$

períodos $L = \text{Int}(N/P)$

Padrão

$$SZ(i) = \frac{1}{L} \sum_{k=0}^{L-1} s_i(i+kP) \quad i=1, \dots, P$$

$$SZ(P) \Rightarrow SZ(0)$$

Sinal

$$s_i(t) \approx SZ\left(\text{Resto } \frac{t}{P}\right)$$

sazonalidade constante

Retirando a sazonalidade

$$S_2(t) = S_1(t) - S_2\left(\text{Reto}\left(\frac{t}{p}\right)\right)$$

Variância "móvel"

$$\sigma^2(i) = \frac{1}{L} \sum_{k=0}^{L-1} S_2^2(i + kP)$$

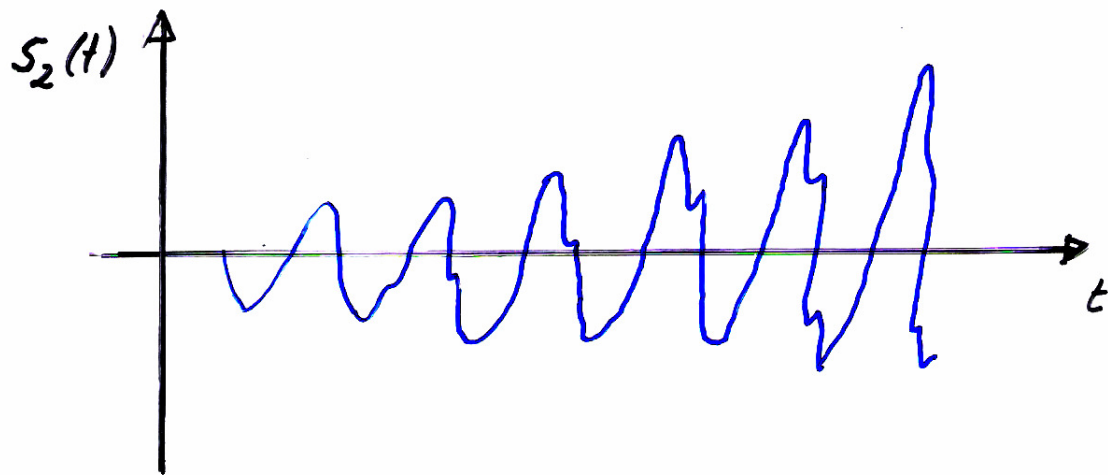
$$i = 1, \dots, P \quad \sigma^2(P) = \sigma^2(0)$$

$$\sigma^2(t) = \sigma^2\left(\text{Reto}\left(\frac{t}{p}\right)\right)$$

Série estacionária, $\mu = 0$ $\sigma^2 = c/t_c = 1$

$$S_3(t) = S_2(t) / \sigma(t)$$

sazonalidade com amplitude
 variavel com o ciclo



$$s_2(t) \approx \underset{\substack{\uparrow \\ \sim a_0 + a_1 t}}{f(t)} \cdot \underset{\substack{\uparrow \\ \text{ctte}(t)}}{s_3(t)}$$

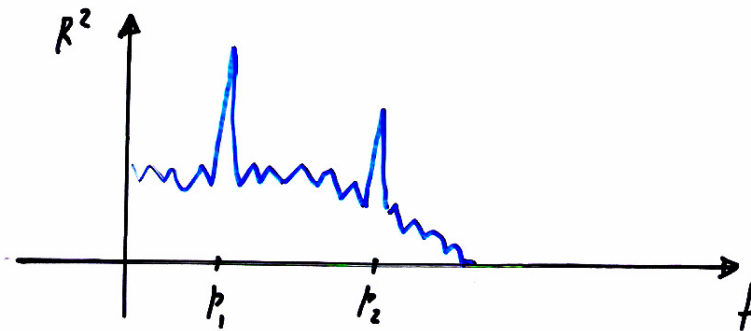
$$\lg s_2(t) = \underset{\substack{\uparrow \\ \text{tendencia}}}{\lg f(t)} + \underset{\substack{\uparrow \\ \text{sazonalidade} \\ \text{ctte}(t)}}{\lg s_3(t)}$$

Extraindo a sazonalidade

$$s_2(t) = s_1(t) - sz\left(a_1 t_0 \frac{t}{p}\right)$$

4 - Cycles senoïdaux

Observer le spectrogramme



$$cs_1(t) = a_{p_1} \cos \frac{2\pi p_1}{N} t + b_{p_1} \sin \frac{2\pi p_1}{N} t$$

Retirer les cycles senoïdaux

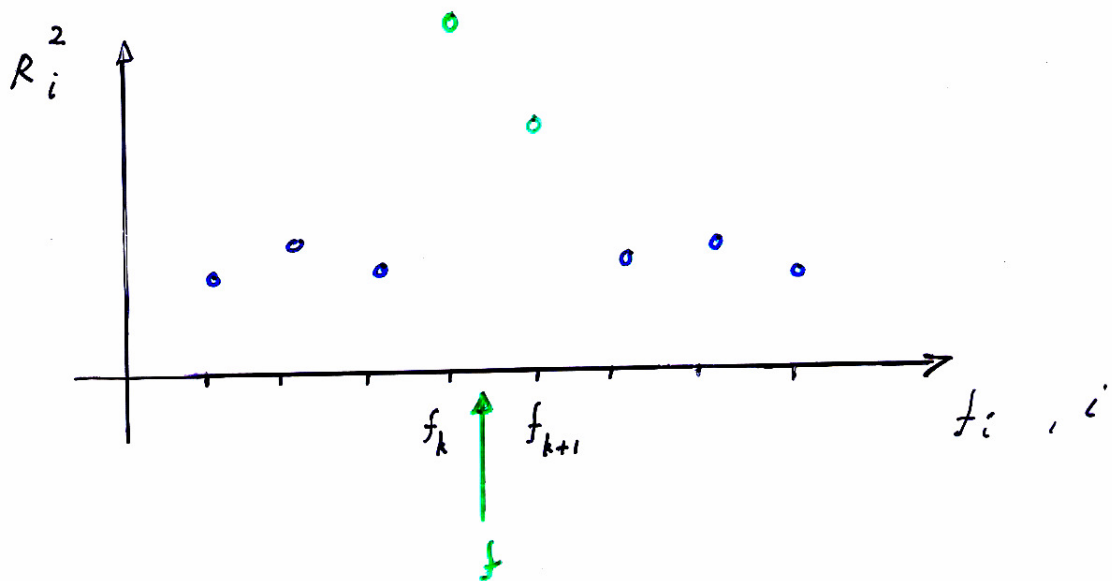
$$s_3(t) = s_2(t) - cs_1(t) - cs_2(t)$$

Precisão

A precisão da FFT pode ser insuficiente, especialmente para frequências muito altas / muito baixas

$$\Delta f = \frac{1}{N}$$

$$f_k < f < f_{k+1} = f_k + \frac{1}{N}$$

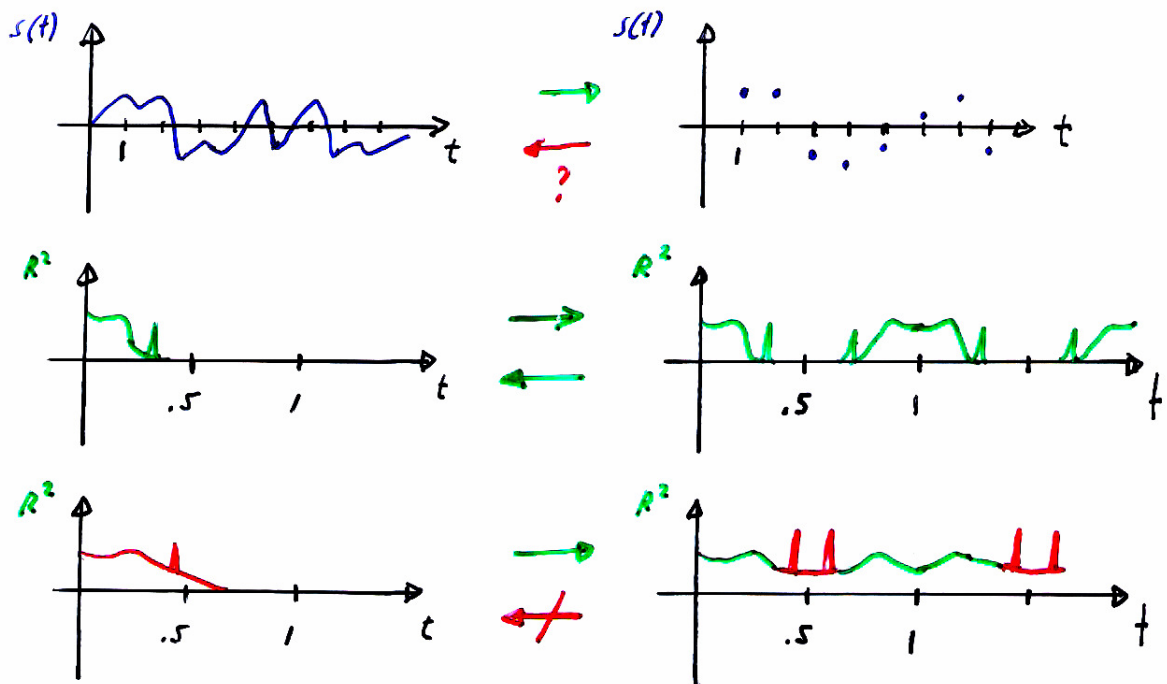


Sub a mortregem

$$\Delta T = 1 \quad f_s = \frac{1}{\Delta T} = 1$$

$$\Delta f = \frac{1}{N} \quad f^* = .5$$

Repetição de espectro



Possíveis problemas na freq. alta

Check: Verificar se a componente

decai no espectrograma de s_2

Solução

Otimizar no domínio do tempo, usando emq e gradiente descendente, p/ ex.

$$F = \frac{1}{N} \sum_{t=1}^N \left[s_i(t) - \left(a \cos 2\pi f t + b \sin 2\pi f t \right) \right]^2$$

$$\Rightarrow a, b \text{ e } f$$

ótimos

Filtros de frequência (fora ^{com} linear)

$\Rightarrow$ adoçamento

Obs:

1 - A retirada de uma componente é válida se a potência da série diminuir, i.e.

$$s_4 = s_3 - c$$

é válida se $\frac{1}{N} \leq s_3^2 > \frac{1}{N} \leq s_4^2$

2 - Normalmente após a retirada de uma componente é necessário (conveniente) restaurar a média nula e recalcular a variância

$$s_4 = s_3 - c$$

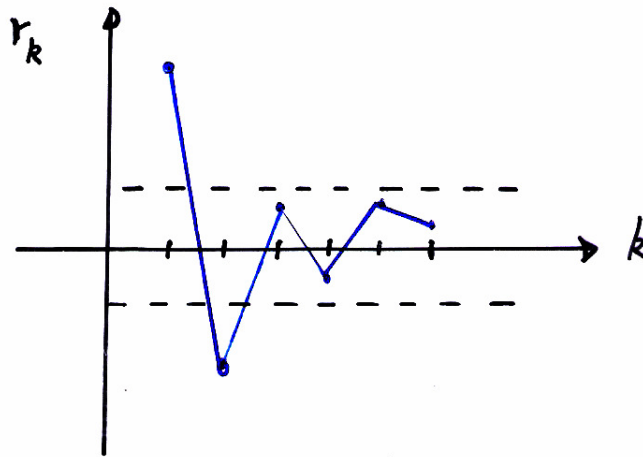
$$s_4 \rightarrow k (s_4 - \bar{s}_4)$$

5 - Termos não Lineares $\Rightarrow$ RN's

$$s_3(t) = n(t) + r(t)$$

Correlações significativas

Autocorrelação



$$k = 1 \text{ e } 2 \quad s_3(t-1) \quad s_3(t-2)$$

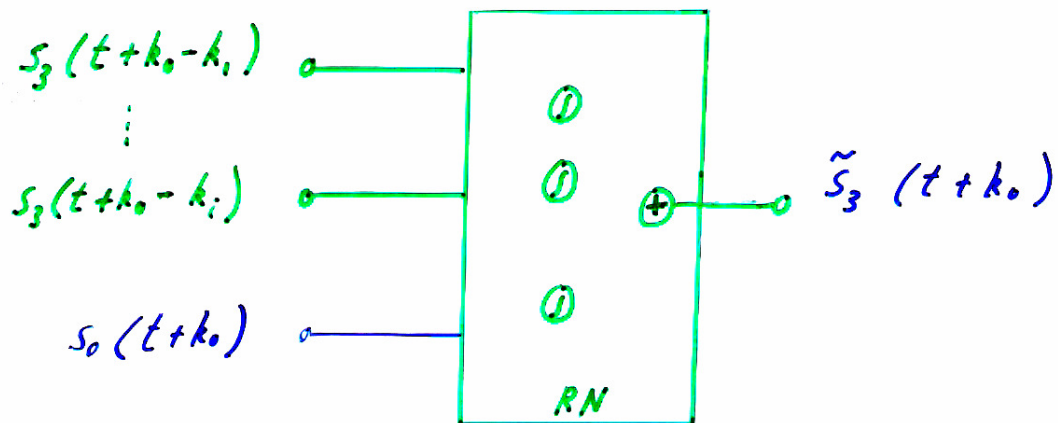
Correlação com amplitude

$$s(t) \approx s_0(t) = \text{tend}(t) + c s(t) + s_z(t)$$

$$r [s_3(t), s_0(t)]$$

não linearidade dependente da amplitude

Preditor para $s_3(t+k_0)$



* k_i com correlação significativa

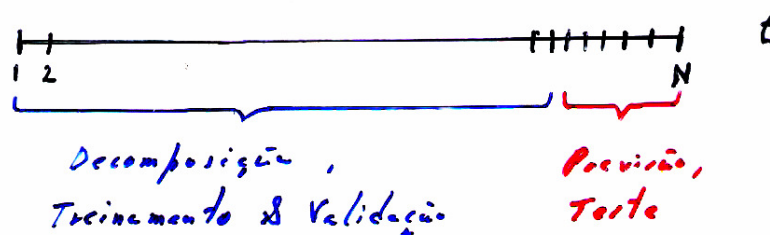
e tal que $k_i \geq k_0$

$s(t+k_0-k_i) \rightarrow$ correlações

$s_0(t+k_0) \rightarrow$ não linearidade

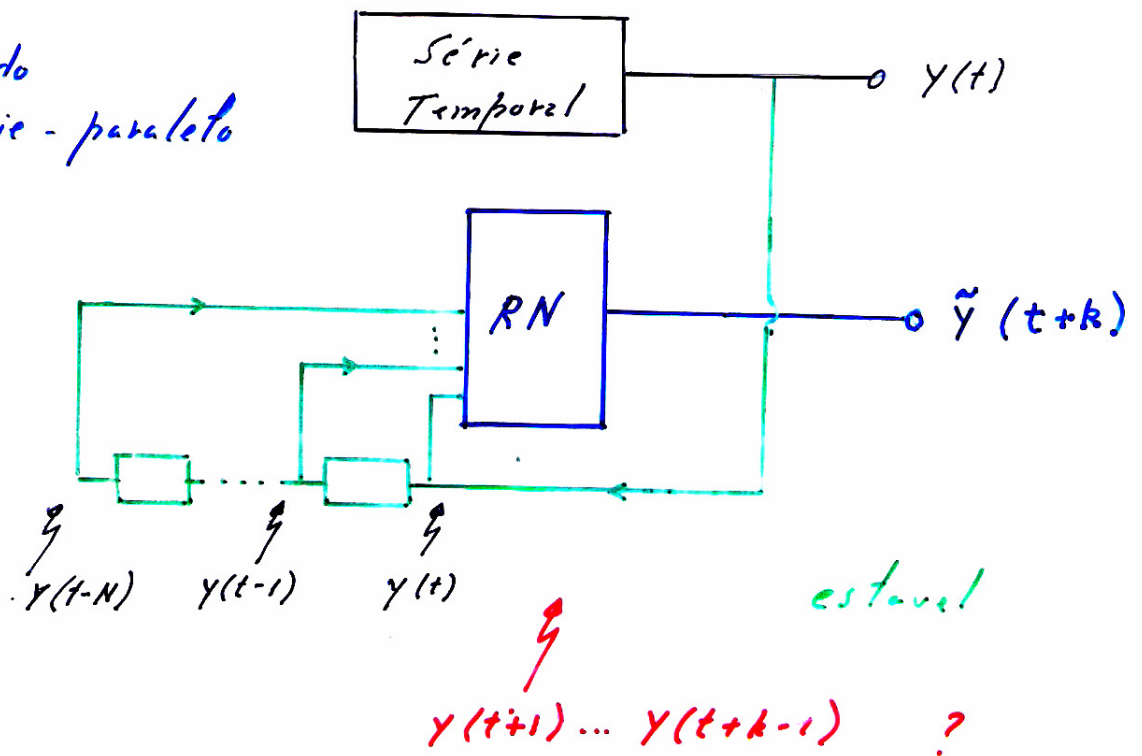
dependente da amplitude

Pares (e, s)

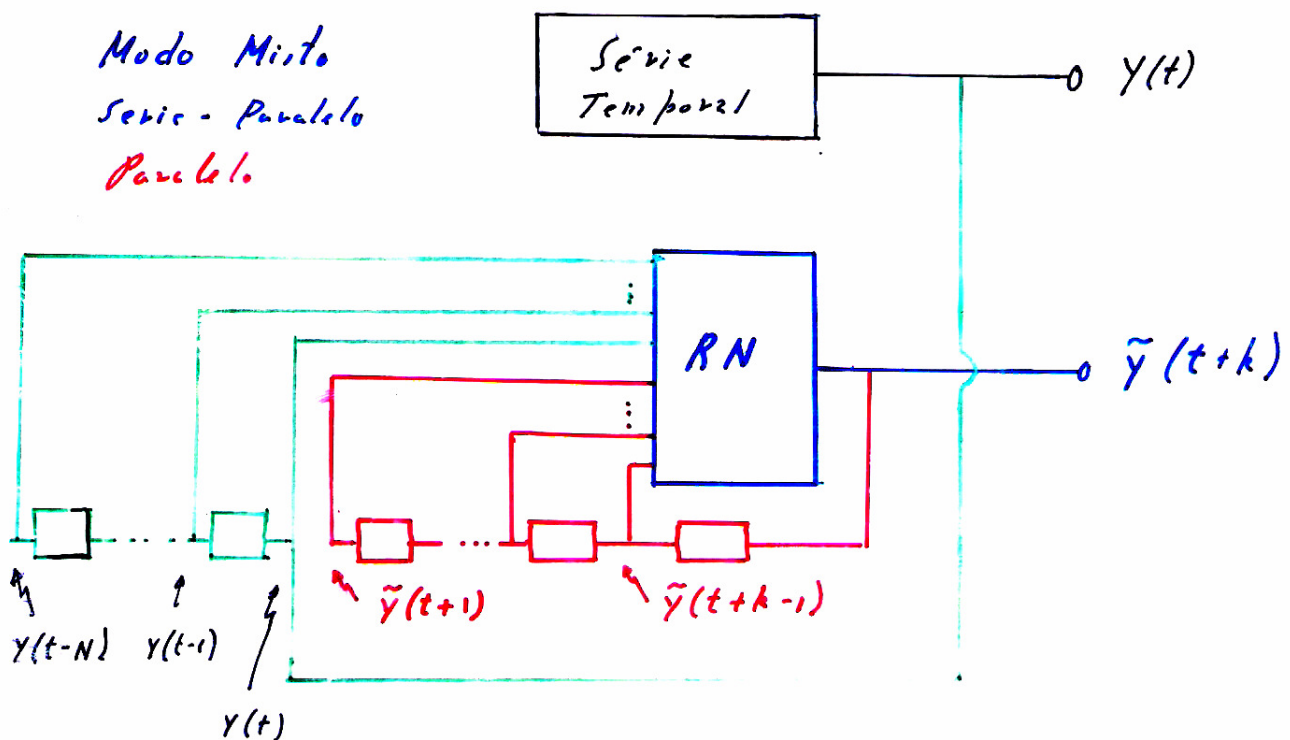


Operação do Preditor

Modo
série - paralelo



Modo Misto
Série - Paralelo
Paralelo.



6 - Recomposição da função

$$\begin{aligned} \tilde{S}(t+k_0) = & \tau(t+k_0) + \\ & CS(t+k_0) + \\ & SZ(t+k_0) + \\ & \tilde{S}_3(t+k_0) \end{aligned} \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \begin{array}{l} \text{clínico} \\ \\ \\ RN \end{array}$$

Uso de Series Auxiliares

série principal $s(t)$

séries auxiliares $f(t)$, $g(t)$

Pretratamento da série principal

$$s_3(t) = s(t) - t(t) - cs(t) - s_2(t)$$

Pretratamento das séries secundarias

$$f_3(t) = f(t) - t_s(t) - cs_s(t) - s_2(t)$$

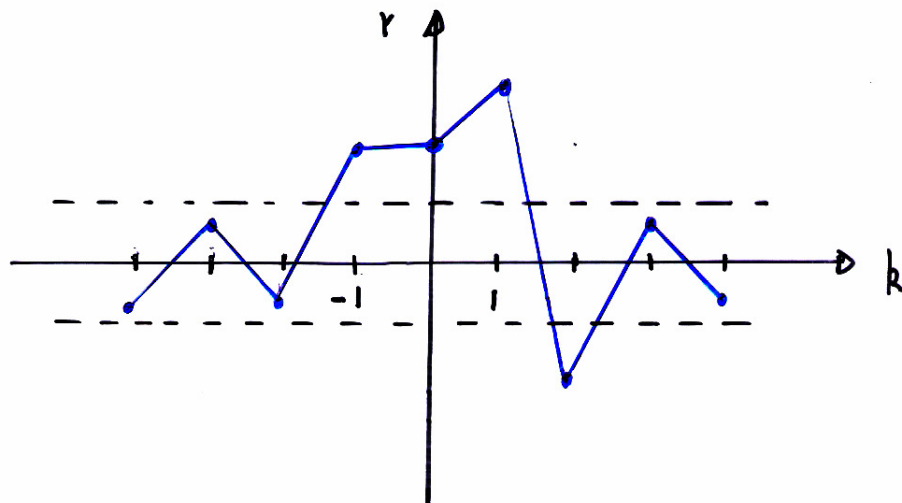
Retirar tendencia $t_s(t)$ do mesmo
tipo que $t(t)$, se existir

Retirar ciclos senoidais $cs_s(t)$
das mesmas frequencias das
de $cs(t)$, se existirem

Retirar sazonalidade $s_2(t)$ com
 mesmo período da de $s_2(t)$, se existir

Verificar correlações significativas

$$r[f_3(t), s(t+k)]$$



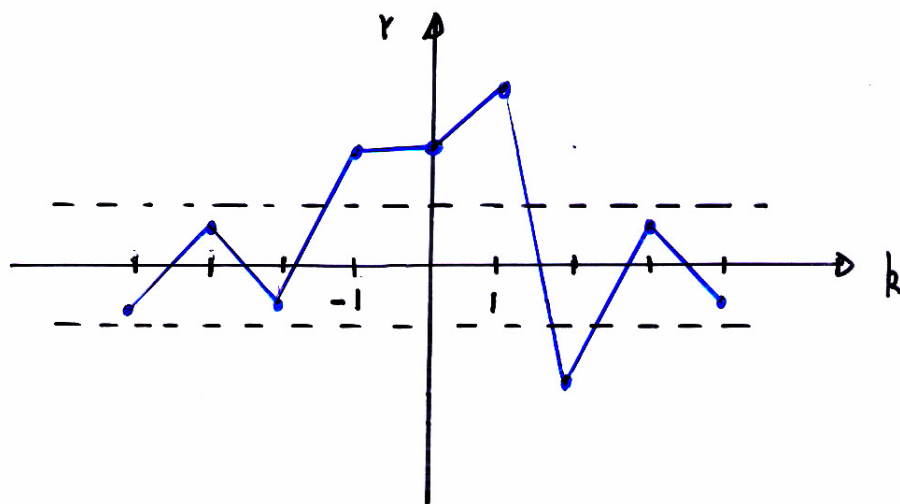
significativas $k = 1, 2$

$k = 0, -1$ simultâneo,
 antecipativo

Retirar sazonalidade $s_2(t)$ com
mesmo período da de $s_2(t)$, se existir

Verificar correlações significativas

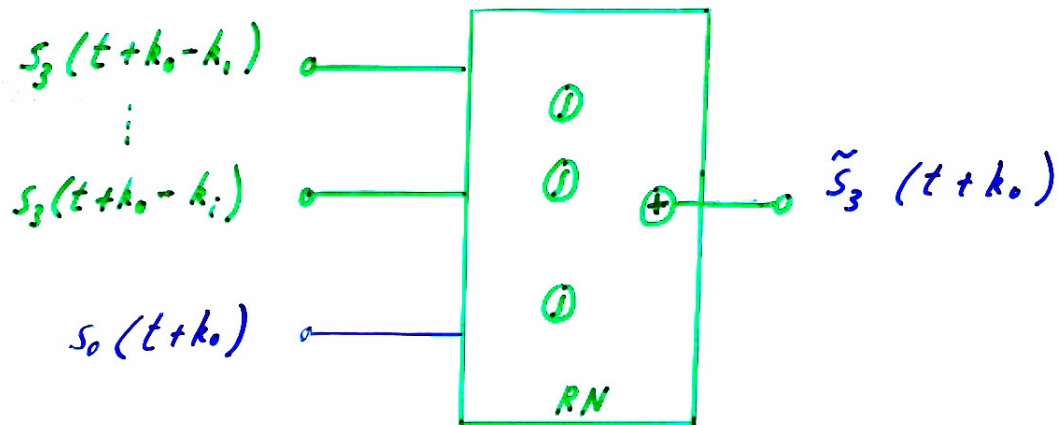
$$r[f_3(t), s(t+k)]$$



significativas $k = 1, 2$

$k = 0, -1$ simultâneo,
antecipativo

Preditor para $s_3(t+k_0)$



* k_i com correlação significativa

e tal que $k_i \geq k_0$

$s(t+k_0-k_i) \rightarrow$ correlações

$s_0(t+k_0) \rightarrow$ não linearidade

dependentes da amplitude

Pares (e, s)

