Modelos Dinâmicos Bayesianos

Helio S. Migon Mariane B. Alves

Instituto de Matemática
Universidade Federal do Rio de Janeiro

2018

✓ **Superposição de modelos**

✓ **Fator de desconto**

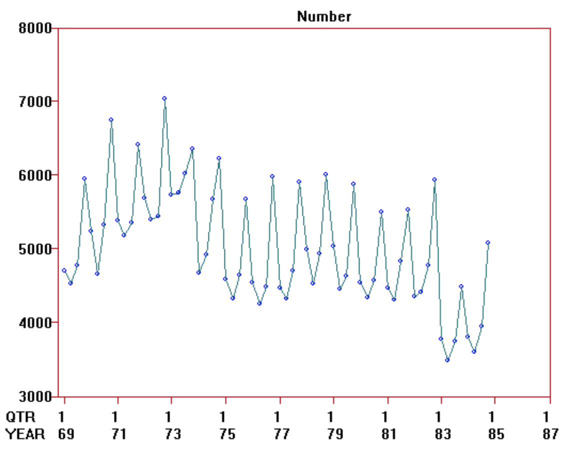
✓ **Regressão dinâmica**

✓ **Sazonalidade**

Superposição de modelos

Muitas séries temporais exibem um comportamento bastante complexo. Ao identificarmos as características mais marcantes, estamos caminhando na direção de formular um modelo. A série de acidentes é um exemplo típico.

Acidentes rodoviários graves (1969:1-1984:4)



- A tendência global parece ser de uma variação suave do nível.
- Se agora nos concentramos na variação em torno desse nível, podemos detectar um comportamento cíclico.

Essa inspeção permitiu identificar os *dois componentes de um modelo:*

- A estrutura dos modelos dinâmicos é apropriada, pois permite que as componentes sejam modeladas separadamente e depois integradas num modelo.

- No caso mais comum de duas componentes: tendência e sazonalidade, estruturamos a equação das observações com dois termos.

$$y_t = y_{Nt} + y_{St} + v_t$$

Cada um dos termos é descrito através de um modelo dinâmico

$$\begin{aligned} y_{Nt} &= F'_{Nt}\theta_{Nt} \\ \theta_{Nt} &= G_{Nt}\theta_{Nt-1} + w_{Nt} \end{aligned}$$

$$\begin{aligned} y_{St} &= F'_{St}\theta_{St} \\ \theta_{St} &= G_{St}\theta_{St-1} + w_{St} \end{aligned}$$

Se agora integramos esses termos, obtemos a equação das observações

$$y_t = \mathbf{F}'\theta_\mathbf{t} + \mathbf{v_t}$$

$$\text{onde } \mathbf{F} = \begin{pmatrix} F_{Nt} \\ F_{St} \end{pmatrix} \text{ e } \theta = \begin{pmatrix} \theta_{Nt} \\ \theta_{St} \end{pmatrix}$$

Similarmente, a equação do sistema (integrado) fica

$$\theta_\mathbf{t} = \mathbf{G}\theta_\mathbf{t-1} + \mathbf{w_t}; \qquad \mathbf{w_t} \sim \mathbf{N}(\mathbf{0}, \mathbf{W_t})$$

$$\text{onde } \mathbf{G_t} = \begin{pmatrix} G_{Nt} & 0 \\ 0 & G_{St} \end{pmatrix} \text{ e } \mathbf{W_t} = \begin{pmatrix} W_{Nt} & 0 \\ 0 & W_{St} \end{pmatrix}$$

Fator de Desconto

- Até agora, não foi discutido o processamento da incerteza relativa a $\mathbf{W}$.
- A abordagem Bayesiana para parâmetros desconhecidos é sempre a mesma: atualização via teorema de Bayes.
- O tratamento dado à variância das observações é analítico. O mesmo não acontece com a variância do sistema.

Felizmente, existe uma solução baseada em **fatores de desconto** que produz uma alternativa aceitável.

- Como já dissemos, o *valor da informação diminui com o tempo*.
- Essa diminuição é controlada pela evolução do sistema, através do aumento da incerteza do sistema.

Mais especificamente, se a equação do sistema é

$$\theta_t = \mathbf{G}\theta_{t-1} + \omega_t, \quad \omega_t \sim N(\mathbf{0}, \mathbf{W}_t)$$

temos que, para $V(\theta_{t-1}|D_{t-1}) = \mathbf{C}_{t-1}$,

$$\mathbf{R}_t = V(\theta_t|D_{t-1}) = \mathbf{P}_t + \mathbf{W}_t$$

onde $\mathbf{P}_t = V(\mathbf{G}_t\theta_{t-1}|D_{t-1}) = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t'$ e $\mathbf{W}_t = \mathbf{R}_t - \mathbf{P}_t$.

Se definimos δ de tal forma que $\mathbf{R}_t = \mathbf{P}_t/\delta$ podemos interpretar δ como a percentagem de informação que passa de $t-1$ para t e nesse caso

$$\mathbf{W_t} = \mathbf{R_t} - \mathbf{P_t} = \mathbf{P_t}/\delta - \mathbf{P_t} = \mathbf{P_t}(\delta^{-\mathbf{1}} - \mathbf{1}).$$

- No caso de superposição de modelos (ex.: tendência + sazonalidade) podemos extender o raciocínio acima.

- No caso geral de k modelos, poderíamos definir

$$\mathbf{P}_{i,t} = V(\mathbf{G}_{it}\theta_{i,t-1}|D_{t-1}), \quad i = 1, 2, \cdots, k$$

  e utilizar um desconto δ_i, $i = 1, 2, \cdots, k$ para cada bloco de forma que $\mathbf{R}_{i,t} = \mathbf{P}_{i,t}/\delta_i$ e $\mathbf{W}_i = \mathbf{P}_{i,t}(\delta_i^{-1} - 1)$.

- O modelo total obtido pela superposição das k componentes teria então $\mathbf{W}_t = \mathrm{diag}(\mathbf{W}_{1,t}, \cdots, \mathbf{W}_{k,t})$, onde $\mathbf{W}_{i,t}$ seria dado como acima.

O **fator de desconto** δ é a percentagem de informação que passa de um período a outro.

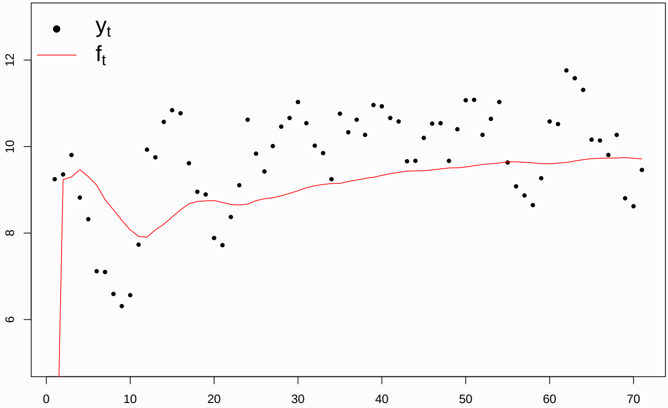
- Valores típicos para sistemas sem variações bruscas se encontram acima de 90%
- A escolha do valor adequado vai depender da aplicação e sugere-se que alguns valores sejam comparados.
- Valores muito próximos não produzem diferenças perceptíveis.
- Valores muito baixos (abaixo de 0,8) tendem a introduzir muita incerteza e produzem limites de incerteza para predição muito grandes.
- Valores muito altos representam um sistema com mudanças muito suaves.
- No limite, quando $\delta = 1$, temos o modelo estático onde não há perda de informação.
- A mesma idéia de desconto pode ser estendida a modelos mais gerais com vários descontos aplicados a partes diferentes do modelo. Essa formulação ficará mais clara quando abordamos superposição de modelos.

Aplicação: Candy

- Ilustramos o uso de fatores de desconto com a série de vendas de um doce (SALES) do arquivo CANDY.DAT.
- As figuras a seguir apresentam as previsões um passo a frente resultantes do ajuste de um modelo de tendência constante com os seguintes fatores de desconto: 1,0; 0,9 e 0,8

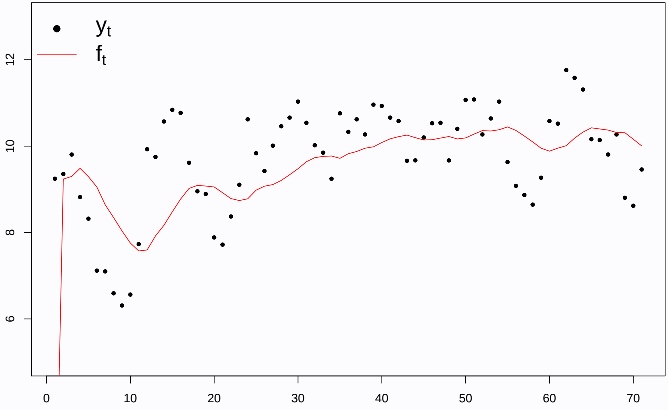
Fator de desconto $\delta = 1$ (estrutura estática)

Previsão 1 passo a frente:



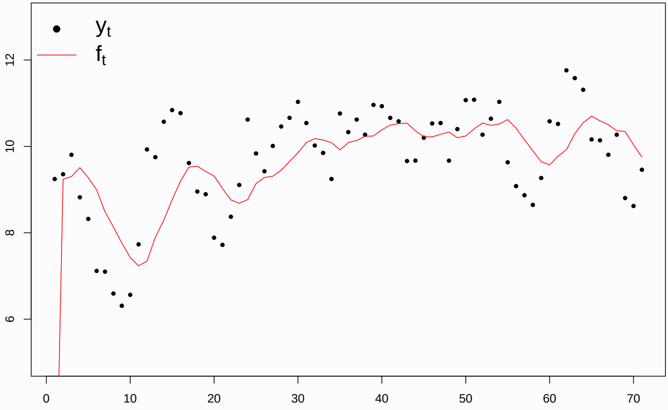
Fator de desconto $\delta = 0.9$ (estrutura estática)

Previsão 1 passo a frente:



Fator de desconto $\delta = 0.8$ (estrutura estática)

Previsão 1 passo a frente:



Regressão Dinâmica

- O modelo genérico usado para introduzir os modelos dinâmicos foi obtido a partir de uma generalização dos modelos de regressão.
- Se a série de *vendas*, (y_t) é explicada pela série de *preços* (x_t) temos:

$$\begin{aligned} y_t &= \mu_t + \beta_t x_t + v_t \\ \mu_t &= \mu_{t-1} + w_{1t} \\ \beta_t &= \beta_{t-1} + w_{1t} \end{aligned}$$

Nada na estrutura acima impede que outras variáveis sejam incluídas como regressores.

No caso de uma série y_t com variáveis explicativas $x_{1t}, \ldots, x_{pt}$, temos

$$\begin{aligned} y_t &= \beta_{0t} + \beta_{1t} x_{1t} + \ldots + \beta_{pt} x_{pt} + v_t \\ \beta_{it} &= \beta_{i,t-1} + w_{it}; \qquad i = 0, 1, \ldots, p \end{aligned}$$

chamado de modelo de regressão dinâmica

A estrutura de modelo dinâmico é evidente com:

- $\mathbf{F_t'} = (\mathbf{1}, \mathbf{x_{1t}}, \ldots, \mathbf{x_{pt}})$
- $\mathbf{G_t} = \mathbf{I_{p+1}}$, a matriz identidade de ordem $p+1$ e
- $\mathbf{w_t'} = (\mathbf{w_{01}}, \mathbf{w_{1t}}, \ldots, \mathbf{w_{pt}})$

MLD de Regressão Múltipla

$$\begin{array}{lll} \text{Eq. observação:} & y_t = \mathbf{F_t'}\theta_\mathbf{t} + \nu_\mathbf{t}, & \nu_t \sim N[0, V_t] \\ \text{Eq. sistema:} & \theta_\mathbf{t} = \theta_\mathbf{t-1} + \omega_\mathbf{t}, & \omega_t \sim N[\mathbf{0}, \mathbf{w_t}] \end{array}$$

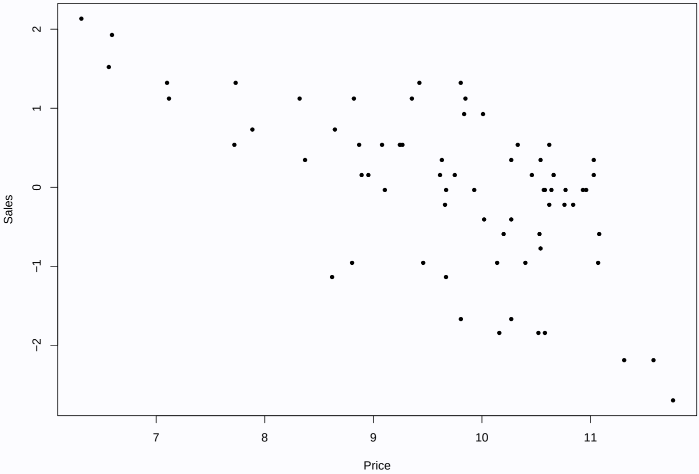
onde

$$\begin{array}{rl} \mathbf{F_t} = (\mathbf{X_{t1}}, \ldots, \mathbf{X_{tn}})' & : \text{vetor de regressoras} \\ X_1, \ldots, X_n & : \text{variáveis independentes} \\ X_{ti} & : \text{valor da } i\text{−ésima variavél } X \text{ no instante } t \\ \theta_\mathbf{t} & : n \times 1 \text{ vetor de parâmetros da regressão} \\ \omega_\mathbf{t} & : \text{matriz da variância de } \theta_\mathbf{t}. \end{array}$$

Aplicação: Candy

- Os dados deste exemplo correspondem as vendas e preços mensais de um doce do arquivo CANDY.DAT.
- Espera-se que série de vendas (SALES) esteja relacionada à série de preços (PRICE).
- Nas figuras a seguir apresentam-se alguns resultados do ajuste de um modelo de regressão dinâmica com tendência estável, utilizando o preço como variável explicativa.

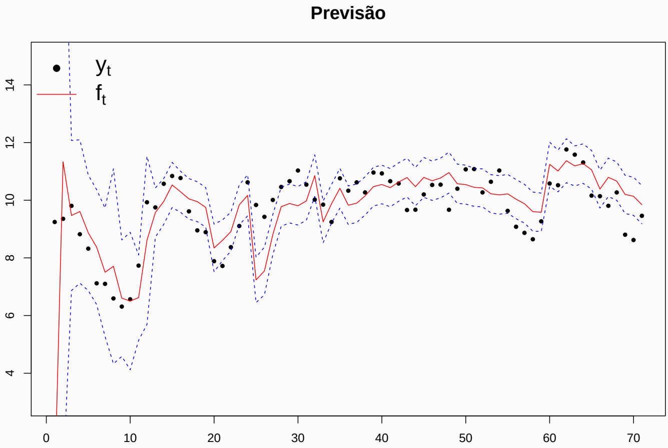
Observa-se um aparente movimento comum das 2 séries.

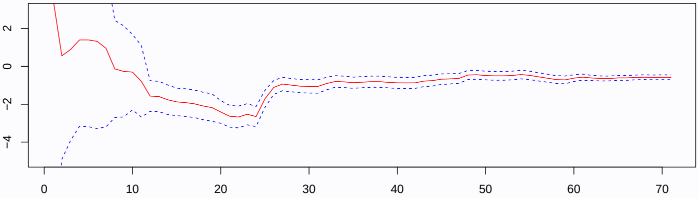


Sales
Price
2
1
0
−1
−2
7
8
9
10
11

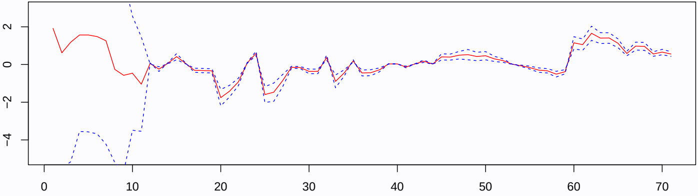
As previsões exibem melhoras consideráveis em comparação com os modelos sem preços.

(Descontos: Nível - 0.9, Covariável - 0.99)





$E[\beta_t|D_t]$

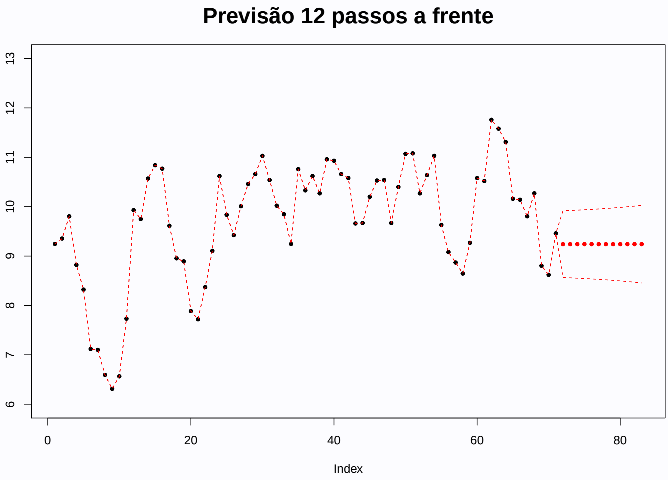


$E[\beta_t|D_t] \, . \, x_t$

Previsão com Regressoras

- Para fazer previsão em modelos com regressoras, *é necessário ter o valores das regressoras ao longo do horizonte de previsão.*
- Normalmente, esses valores também são incertos e o tratamento a ser dado é muito mais complicado.
- Uma alternativa intermediária é fazer previsão sob vários cenários plausíveis.

Neste caso todos os valores de preço para o período de predição são zero, portanto, as previsões têm a forma constante do modelo estável.



Modelos com Sazonalidade

1 Modelos sazonais *requerem uma componente periódica* no modelo.

2 A representação mais simples é através de **fatores ou indicadores de cada período no ciclo**. Para dados trimestrais, são usados quatro indicadores.

3 Uma pequena alteração envolve o uso de efeitos indicando a variação sazonal em torno de um nível. Nesse caso, os efeitos estão restritos a ter soma zero.

4 Fatores trimestrais de 100, 140, 80 e 120 equivalem a um nível de 110 e efeitos trimestrais de - 10, 30, - 30 e 10.

5 A última formulação é mais atraente pois permite a separação entre sazonalidade e tendência.

6 A restrição deve ser mantida em todas as afirmações probabilísticas mas é facilmente incorporável ao método de inferência utilizado.

Definições

Seja $g(t)$ qualquer função real definida para os inteiros não-negativos, onde t é o indexador de tempo.

i. g(t) é cíclica ou periédica se, para algum inteiro $p \geq 1$, $g(t + np) = g(t)$, para todo inteiro $t \geq 0$ e $n \geq 0$.

ii. O menor inteiro p para o qual o ítem anterior é válido é chamado de período de $g(.)$.

iii. $g(.)$ exibe um ciclo completo para todo intervalo de p pontos consecutivos: $(t, t + p - 1)$, $\forall t \geq 0$.

iv. Os fatores sazonais de $g(.)$ são os p valores de qualquer ciclo completo: $\theta_j = g(j) \quad j = 0, 1, 2, \cdots , p - 1$. Note que, para $t > 0$, $g(t) = g(j)$ onde j é o resto da divisão de t por p. $(j = p \mid t)$

Definições

v. O vetor de fatores sazonais em t é simplesmente o vetor de fatores sazonais permutados, de modo que o primeiro elemento é aquele para o tempo t

$$\theta_t = (\theta_j, \theta_{j+1}, \cdots, \theta_{p-1}, \theta_0, \cdots, \theta_{j-1})'$$

quando o fator sazonal corrente é θ_j. Em particular, para quaisquer inteiros n e $k = np$,

$$\theta_k = (\theta_0, \cdots, \theta_{p-1})'.$$

vi. Em qualquer ciclo, o tempo correspondente ao fator sazonal θ_j tem um rétulo, M(j). Claramente, os rétulos são cíclicos com período p e $j = p \mid t$ se, e só se, $M(j) = t$.

Caracterização de fatores sazonais cíclicos

- Duas possibilidades: (i) indicadores ou (ii) funções trigonométricas.
- A cada ponto no tempo fica associado um vetor de fatores sazonais.
- Seja um ciclo completo de p períodos. Denote cada fator por θ_1, θ_2, $\cdots$, θ_p. O vetor de fatores sazonais em t é denotado por $\theta_\mathbf{t}$.
- Se o modelo consistir somente da componente sazonal, então o nível da série no tempo t, μ_t é dado pelo primeiro componente de θ_t e pode ser obtido como $\mu_t = \mathbf{E}_p' \theta_t$ onde $\mathbf{E}_p' = (1, \underbrace{0, 0, \cdots, 0}_{\text{(p-1) termos}})$.

Se tempo t corresponde ao j-ésimo per'{ı}odo, 4então $t+1$ corresponderá ao $(j+1)$-ésimo. Permutando o vetor θ_t obtemos $\theta'_{t+1} = (\theta_{j+1}, \cdots, \theta_p, \theta_1, \cdots, \theta_j)$ e $\mu_{t+1} = \mathbf{E}'_p \theta_{t+1}$.

A passagem de θ_t para θ_{t+1} é realizada através de

$$\mathbf{P}_p = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{p-1} \\ 1 & \mathbf{0}' \end{pmatrix},$$

i.é.: $\theta_{t+1} = \mathbf{P}_p \theta_t$ onde $\mathbf{P}_p$ ' tq. $\mathbf{P}_p^{k+np} = \mathbf{P}_p^k$.

Essa formulação leva 'a definição do modelo sazonal

$$\begin{aligned} \text{Eq. das Obs.: } y_t &= \mathbf{E}'_p \theta_t + \nu_t \quad \nu_t \sim N[0, V_t] && (1) \\ \text{Eq. Sist.: } \theta_t &= \mathbf{P}_p \theta_{t-1} + \omega_t \quad \omega_t \sim N[0, \mathbf{W}_t] && (2) \\ (\theta_0 | D_0) &\sim N(\mathbf{m_0}, \mathbf{C}_0) && (3) \end{aligned}$$

Observações

(i) Se $\mathbf{W}_t = c\mathbf{I}_p$, os erros ω_t são não correlacionadas e o problema se reduz à análise de p modelos de 1a. ordem, se a priori $\mathbf{C}_0 = diag(c_1, \cdots, c_p)$. Pois não haverá troca de informação entre diferentes níveis de ciclo sazonal;

(ii) Para obter a função de previsão suponha que $\theta_t = (\theta_1, \cdots, \theta_p)'_t$ e, portanto, $E(\theta_t|D_t) = \mathbf{m}_t = (m_{t,1}, \cdots, m_{t,p})'$. A previsão h-passos-a-frente é dada por

$$\begin{aligned} f_t(h) &= E(y_{t+h}|D_t) = E(\mu_{t+h}|D_t) \\ &= E(\mathbf{E}_p\theta_{t+h}|D_t) = \mathbf{E}_p\mathbf{P}_p^h E(\theta_t|D_t) = \mathbf{E}_p\mathbf{P}_p^h\mathbf{m}_t \\ &= \mathbf{E}_p(m_{t,h+1}, \cdots, m_{t,p}, m_{t,1}, \cdots, m_{t,h})' \\ &= m_{t,h+1} \end{aligned}$$

A modelagem por indicadores sazonais é feita através de um nível médio e **variações sazonais com relação a este nível médio μ_t. Assim os fatores sazonais satisfazem a $1'\theta_t = \sum_{j=1}^p \theta_{t,j} = 0$.**

2) Restrições são preservadas pelo modelo dinâmico com $\mathbf{1}'\omega_t = 0$.

Uma vez garantidas essas restrições podemos redefinir o modelo sazonal como

$$
\begin{array}{rcll}
\text{Eq. Observações: } y_t & = & \mu_t + \theta_{t,j} + \nu_t \quad \nu_t \sim N[0, V_t] & \\
& = & \mu_t + \mathbf{E}_p'\theta_t + \nu_t & \\
& = & \mathbf{F}'\beta_t + \nu_t & \\
\text{Eq. do Sistema : } \mu_t & = & \mu_{t-1} + \omega_{0,t}, \quad \omega_{0,t} \sim N(0, W_{0,t}) & \\
\theta_t & = & \mathbf{P}_p\theta_{t-1} + \omega_t \quad \omega_t \sim N(0, \mathbf{W}_t) & (4) \\
\text{Inf. Inicial : } (\mu_0|D_0) & \sim & N(m_{00}, \mathbf{C}_{00}) \quad (\theta_0|D_0), \sim N(\mathbf{m}_0, \mathbf{C}_0) &
\end{array}
$$

Restrição adicional : $\mathbf{1}'\theta_t = 0, \forall t$.

Supondo novamente que $E[\beta_t|D_t] = (m_{t,0}, m_{t,1}, \cdots, m_{t,p})'$ onde $\sum_{j=1}^{p} m_{t,j} = 0$ temos a função de previsão dada por

$$
\begin{aligned}
f_t(h) &= E(y_{t+h}|D_t) \\
&= E(\mu_{t+h} + \mathbf{E}_p'\theta_{t+h}|D_t) \\
&= E(\mu_{t+h}|D_t) + \mathbf{E}_p' E(\theta_{t+h}|D_t) \\
&= E(\mu_t + \sum_{j=1}^{h} \omega_{0,t+j}|D_t) + \mathbf{E}_p' E(\mathbf{P}_p^h \theta_t + \sum_{j=1}^{h} \mathbf{P}_p^{h-j}\omega_{t+j}|D_t) \\
&= E(\mu_t|D_t) + \mathbf{E}_p' E(\mathbf{P}_p^h \theta_t|D_t) \\
&= m_{t,0} + m_{t,h+1}
\end{aligned}
$$

Exemplo

{Para dados trimestrais, os modelos dinâmicos utilizam quatro indicadores. A passagem do tempo faz com que eles experimentem uma rotação. Assim,

$$\theta_{t-1} = \begin{pmatrix} trim4 \\ trim1 \\ trim2 \\ trim3 \end{pmatrix}$$

Essa rotação pode ser efetuada pela matriz de evolução

$$G = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

O modelo é completado por uma equação de observações que considera apenas a primeira componente do vetor paramétrico, ou seja, $F_t' = (1, 0, 0, 0)$. A extensão para um ciclo de p períodos é análoga.

Modelagem por harmônicos

- Uma outra modelagem de padrões cíclicos pode ser feita usando **funções trigonométricas**.
- A função $cos(\omega(t-1))$ é periódica com período $\frac{2\pi}{\omega}$. Se $\omega = \pi/6$, o período é 12 e o máximo ocorre para $t = 1$.
- Dados mensais com ciclo anual podem ser concisamente modelados via

$$y_t = a_t \cos\left(\frac{\pi(t-1)}{6}\right) + \upsilon_t$$

onde a_t é um parâmetro que controla a amplitude e o máximo ocorre em *janeiro*. Observe a redução drástica na dimensão do vetor paramétrico de 11 para 1.

- Defasagens no ponto de máximo do ciclo podem ser acomodadas com um parâmetro extra segundo

$$y_t = a_t \cos\left(\frac{\pi(t-1)}{6}\right) + b_t \sin\left(\frac{\pi(t-1)}{6}\right) + \upsilon_t$$

A formulação dinâmica dessa função harmônica utiliza 2 parâmetros,

$$F_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, G_t = \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix}$$

onde $\omega = \pi/6$.
Esse modelo descreve um ciclo segundo uma função *coseno*.

- Padrões cíclicos mais complicados podem ser modelados com a inclusão de formas harmônicas de freqüência maior. A função $\cos(2\omega(t-1))$ é similar porém completa 2 ciclos durante um período de tempo $p = 2\pi/\omega$ ou $\omega = 2\pi/p$.
- O resultado fundamental aqui informa que qualquer padrão cíclico de período p pode ser reproduzido com a soma de, no máximo, $p/2$ harmônicos de períodos $p/j, j = 1, \ldots, [p/2]$ harmônicos.
- A vantagem desse resultado reside em que podemos fazer economia no número de parâmetros utilizados e conseqüentemente aumentar nossa capacidade de aprendizado sobre o sistema e melhorar nossas previsões.

Aplicação: Candy

- A série de vendas do arquivo CANDY exibe um comportamento cíclico que até agora não havia sido tratado.
- Nas figuras a seguir apresentam-se os resultados da análise da série CANDY.DAT considerando um modelo de tendência constante e um regressor (preço), incluindo a componente sazonal representada de forma livre e com harmônicos.

Candy: Modelagem com Harmônicos

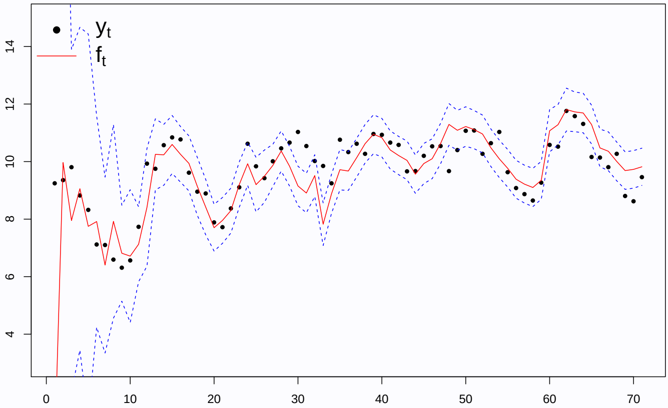
- Como os dados são mensais, temos que o período tem tamanho $p = 12$, havendo portanto até 6 harmônicos de períodos. O primeiro ou fundamental, de período $12 = 12/1$; o segundo, de período $6 = 12/2 \ldots$ até o último, de período $2 = 12/6$.

- A diminuição da dimensão do vetor paramétrico é importante pois embora não altere as previsões pontuais, diminui a incerteza e facilita a monitoração da performance do modelo. Além disso, uma modelagem mais parcimoniosa acelera o tempo de processamento da análise.

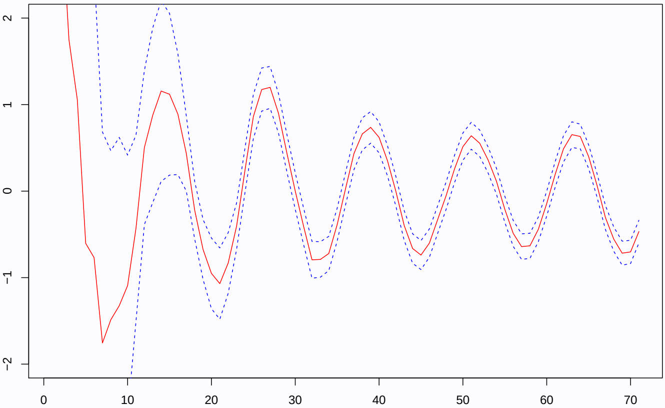
- Esta modelagem da sazonalidade permite que o período do ciclo sazonal seja diferente do período natural da série e que o modelo escolhido contenha apenas o harmônico fundamental, ou seja, uma única função senoidal.

Candy: Modelagem com 1 harmônico

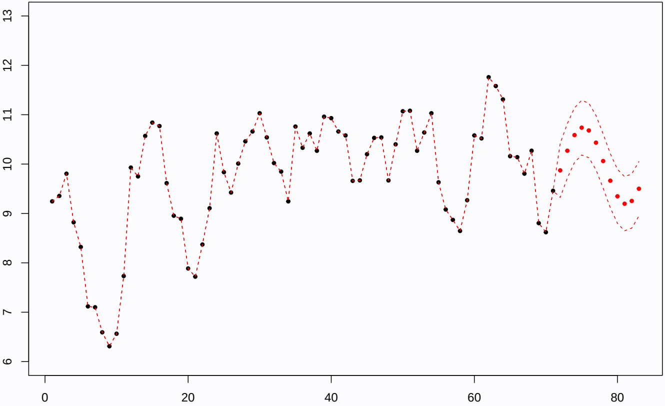
Candy: Modelagem com 1 harmônico



Estimativa da sazonalidade com limites de incerteza: primeiro harmônico



Previsão com limites de incerteza: 1 harmônico



Distribuição Normal Multivariada

O vetor p-dimensional $x = (x_1, x_2, \cdots, x_p)'$ tem distribuição normal p-variada com média μ e variˆancia Σ, $x \sim N(\mu, \Sigma)$, se tem densidade :

$$f_N(x; \mu, \Sigma) = (2\pi)^{p/2} \mid \Sigma \mid^{1/2} exp\left\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right\}$$

onde $\mid A \mid$ – determinante da matriz A. A normal padronizada é obtida quando $\mu = 0$ e $\Sigma = I_p$, identidade de ordem p. Nesse caso, as componentes $x_i' s$ são normais padronizadas independentes. A normal univariada é o caso particular onde $p = 1$.

Propriedades importantes

Transformações lineares: se A é uma matriz $r \times p$ e b é um vetor r-dimensional então

$$y = Ax + b \sim N(A\mu + b, A\Sigma A') \tag{5}$$

Distribuições Marginais: se o vetor x é dividido em 2 blocos x_1 contendo os primeiros r componentes de x e x_2, os outros $p - r$ então procedendo a mesma partição em μ e Σ na forma

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad e \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \tag{6}$$

obtém-se que

$$x_i \sim N(\mu_i, \Sigma_{ii}), \quad i = 1, 2$$

Distribuições condicionais:

$$x_1 \mid x_2 \sim N(\mu_{1.2}, \Sigma_{11.2})$$

Se $x_1 \mid x_2 \sim N(\mu_1 + B_1(x_2 - \mu_2), B_2)$ e $x_2 \sim N(\mu_2, \Sigma_{22})$ então

$$
\begin{aligned}
x &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N(\mu, \Sigma) \qquad \text{com} \\
\mu &= \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{e} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}
\end{aligned}
$$

onde $\Sigma_{11} = B_2 + B_1 \Sigma_{22} B_1'$ e $\Sigma_{21}' = \Sigma_{12} = B_1 \Sigma_{22}$.

em Formas quadráticas : $(x - \mu)' \Sigma^{-1} (x - \mu) \sim \chi_p^2$.