Markov Switching Models

Profa. Airlane Alencar

Depto de Estatística - IME-USP

www.ime.usp.br/∼lane

Ref: Kim e Nelson (1999) e Hamilton (1990)

Objetivo

Mudança nos parâmetros de um modelo de regressão definindo diferentes regimes.

- Datas conhecidas - Teste de Chow (1960);
- Quandt (1972): Regimes independentes;
- Goldfeld e Quandt (1973) - Regimes Markovianos;
- Hamilton (1989) - Mudança Markoviana no modelo AR(p).

Modelo com mudança de regime

$$y_t = x_t \beta_{S_t} + e_t, \tag{1}$$

em que:

1. x_t é um vetor de variáveis exógenas $1 \times k$;
2. S_t define o regime;
3. $e_t \ iid \ \sim N(\mathbf{0}, \sigma^2_{S_t})$.

Consideremos dois regimes $S_t = 0, 1$.

1. $\beta_{S_t} = \beta_0(1 - S_t) + \beta_1 S_t$;
2. $\sigma^2_{S_t} = \sigma^2_0(1 - S_t) + \sigma^2_1 S_t$.

A função log-verossimilhança é

$$
\begin{aligned}
\ln L &= \sum_{t=1}^{T} lnf(y_t|\widetilde{y}_{t-1}) \\
&= \sum_{t=1}^{T} ln\left[\sum_{S_t=0}^{1} f(y_t|S_t,\widetilde{y}_{t-1})P(S_t|\widetilde{y}_{t-1})\right]
\end{aligned}
$$

Regimes independentes

Probabilidades de cada regime dependentes de Z usando a função de ligação logística:

$$\begin{aligned} P(S_t = 1|\widetilde{y}_{t-1}) &= p_t = \frac{exp(\gamma_0 + Z_t\gamma_1)}{1 + exp(\gamma_0 + Z_t\gamma_1)} \\ P(S_t = 0|\widetilde{y}_{t-1}) &= 1 - p_t = \frac{1}{1 + exp(\gamma_0 + Z_t\gamma_1)} \end{aligned}$$

ou usando alguma outra função de ligação, como por exemplo a probit, que facilita a obtenção da posteriori usando inferência bayesiana.

Transição Markoviana

$$
\begin{aligned}
P(S_t = 1 | S_{t-1} = 1, \widetilde{y}_{t-1}) &= p_t = \frac{exp(\gamma_0 + Z_t\gamma_1)}{1 + exp(\gamma_0 + Z_t\gamma_1)} \\
P(S_t = 1 | S_{t-1} = 0, \widetilde{y}_{t-1}) &= q_t = \frac{exp(\delta_0 + Z_t\delta_1)}{1 + exp(\delta_0 + Z_t\delta_1)}
\end{aligned}
$$

Filtro de Probabilidades

Passo 1

$$P(S_t = j|\widetilde{y}_{t-1}) = \sum_{i=0}^{1} P(S_t = j|S_{t-1} = i)P(S_{t-1} = i|\widetilde{y}_{t-1})$$

Passo 2 - Atualização

$$\begin{aligned} P(S_t = j|\widetilde{y}_t) &= \frac{f(S_t = j, y_t|\widetilde{y}_{t-1})}{f(y_t|\widetilde{y}_{t-1})} = \\ &= \frac{f(y_t|S_t = j, \widetilde{y}_{t-1})P(S_t = j|\widetilde{y}_{t-1})}{\sum_{j=0}^{1} f(y_t|S_t = j, \widetilde{y}_{t-1})P(S_t = j|\widetilde{y}_{t-1})} \end{aligned}$$

Para iniciar o filtro tem que inicializar $P(S_0|\widetilde{y}_0)$, por exemplo usando a probabilidade invariante no caso de cadeia estacionária.

Exemplo - AR(1)

$$
\begin{aligned}
y_t - \mu_{S_t} = \phi(y_{t-1} - \mu_{S_{t-1}}) + e_t, t = 1, \ldots, T \\
e_t \sim N(0, \sigma^2_{S_t}) \\
S_t = 1, \ldots, M.
\end{aligned}
$$

Agora a densidade de y_t depende de S_t e S_{t-1}

$$
f(y_t|\widetilde{y}_{t-1}, S_t, S_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2_{S_t}}} exp\left[-\frac{(y_t - \mu_{S_t} - \phi(y_{t-1} - \mu_{S_{t-1}}))^2}{2\sigma^2_{S_t}}\right]
$$

e

$$
\begin{aligned}
\ln L &= \sum_{t=1}^{T} lnf(y_t|\widetilde{y}_{t-1}) \\
&= \sum_{t=1}^{T} ln\left[\sum_{S_t=1}^{M}\sum_{S_{t-1}=1}^{M} f(y_t|S_t, S_{t-1}, \widetilde{y}_{t-1})P(S_t, S_{t-1}|\widetilde{y}_{t-1})\right]
\end{aligned}
$$

Filtro de Probabilidades

Passo 1

$P(S_t = j, S_{t-1} = i|\widetilde{y}_{t-1}) = P(S_t = j|S_{t-1} = i)P(S_{t-1} = i|\widetilde{y}_{t-1})$

Passo 2 - Atualização

$$
\begin{aligned}
&P(S_t = j, S_{t-1} = i|\widetilde{y}_t) = \\
&= \frac{f(S_t=j,S_{t-1}=i,y_t|\widetilde{y}_{t-1})}{f(y_t|\widetilde{y}_{t-1})} = \\
&= \frac{f(y_t|S_t=j,S_{t-1}=i,\widetilde{y}_{t-1})P(S_t=j,S_{t-1}=i|\widetilde{y}_{t-1})}{\sum_{i=1}^{M}\sum_{j=1}^{M} f(y_t|S_t=j,S_{t-1}=i,\widetilde{y}_{t-1})P(S_t=j,S_{t-1}=i|\widetilde{y}_{t-1})}
\end{aligned}
$$

Para iniciar o filtro tem que inicializar $P(S_0|\widetilde{y}_0)$.

Suavização e EM

A variância assintótica dos estimadores de máxima verossimilhança podem ser obtidos utilizando-se o inverso da matriz informação de Fisher (estimada usando a matriz hessiana no ponto de máximo).

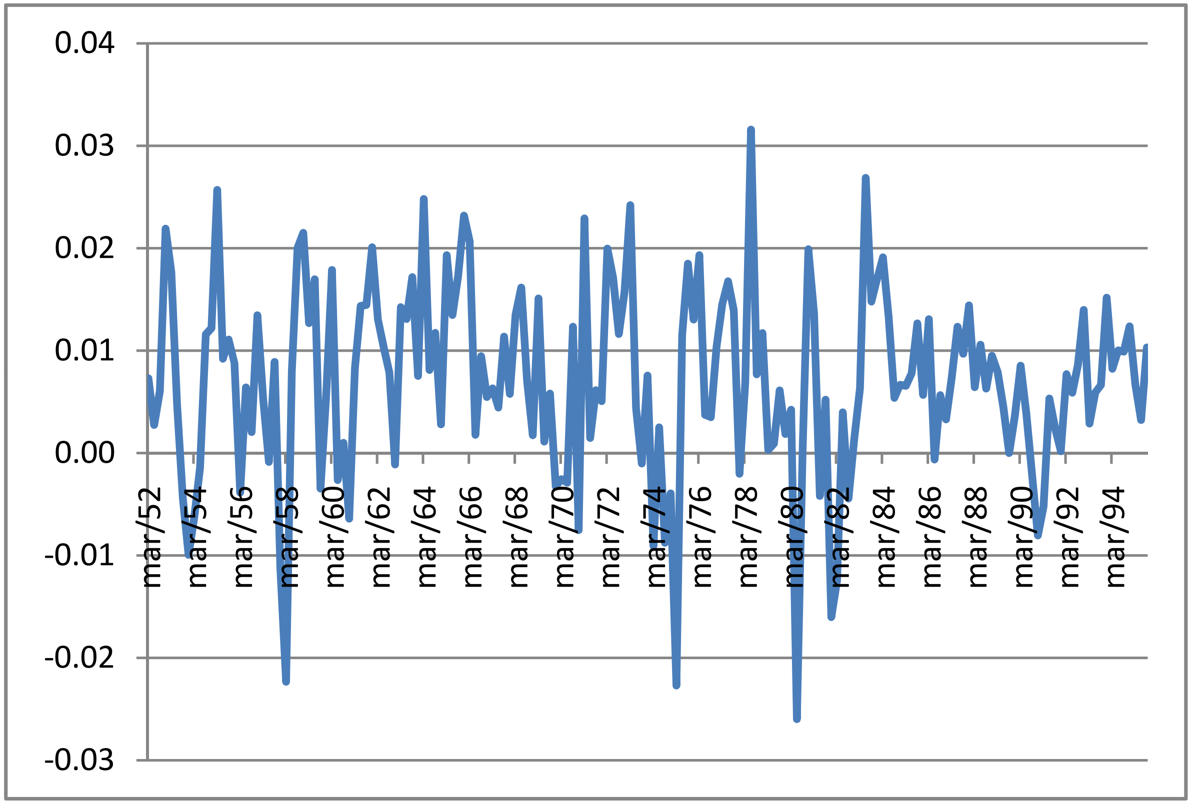
O algoritmo de suavização proposto por Kim permite obter uma aproximação para $P(S_t = j|\widetilde{y}_T)$.

Pode ser utilizado o algoritmo EM para realizar a estimação, escrevendo-se a log-verossimilhança completa, ou seja, usando a densidade das observações e as variáveis não observadas que nesse caso são os regimes.

Hamilton - GDP

Hamilton (1989) modelou o crescimento do PIB real como um modelo AR(4) com dois regimes para a média. A seguir, y_t é o log do PIB real.

Δ Log do PIB real



Modelo

$$
\begin{aligned}
(\Delta y_t - \mu_{S_t}) &= \phi_1(\Delta y_{t-1} - \mu_{S_{t-1}}) + \ldots + \phi_4(\Delta y_{t-4} - \mu_{S_{t-4}}) + e_t \\
e_t &\sim N(0, \sigma^2) \\
\mu_{S_t} &= \mu_0(1 - S_t) + \mu_1 S_t
\end{aligned}
$$

$$
P(S_t = 1 | S_{t-1} = 1) = p,\ P(S_t = 0 | S_{t-1} = 0) = q
$$

Modelo estacionário, sujeito a $\phi(B) = (1 - \phi_1 B - \ldots - \phi L^4) = 0$ com raízes fora do círculo unitário.

É possível distinguir dois regimes: recessão e expansão (média negativa e positiva para o crescimento do PIB real)

Maximum likelihood estimates of the Hamilton model (real GDP; 1952:II–1984:IV; 1952:II–1995:III)

	1952:II-1984:IV		1952:II-1995:III			
	(No dummy variable)		(With a dummy variable)			
p	0.9008	(0.0443)	0.9113	(0.0363)	0.9187	(0.0309)
q	0.7606	(0.1206)	0.7658	(0.0857)	0.7668	(0.0863)
ϕ_1	0.0898	(0.1981)	0.0496	(0.1347)	0.0477	(0.1117)
ϕ_2	−0.0186	(0.2082)	−0.0495	(0.1295)	−0.0422	(0.1103)
ϕ_3	−0.1743	(0.1381)	−0.2112	(0.1129)	−0.2095	(0.1008)
ϕ_4	−0.0839	(0.1248)	−0.0953	(0.1140)	−0.0984	(0.0970)
σ	0.7962	(0.0858)	0.6902	(0.0505)	0.6939	(0.0474)
μ_0	−0.2132	(0.2613)	−0.2996	(0.1892)	−0.2328	(0.1895)
μ_1	1.1283	(0.1596)	1.1479	(0.0768)	1.1510	(0.0776)
μ_0^*	—		0.4516	(0.3209)	—	
μ_1^*	—		−0.3346	(0.1340)	−0.3699	(0.1244)
Log likelihood	−175.24		−212.17		−212.99	

Note: Standard errors are in parentheses.

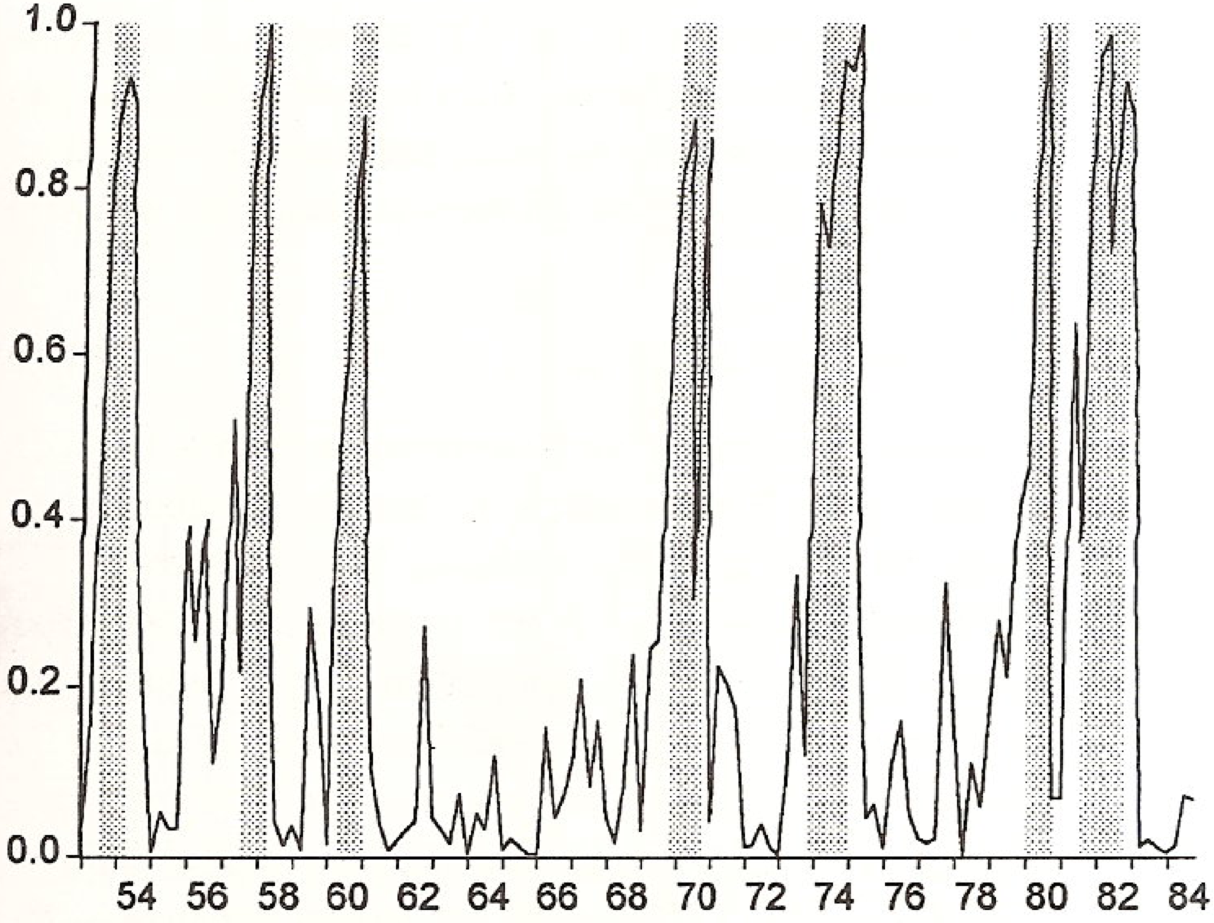


Figure 4.2
Filtered probability of a recession (GDP: 1952:II–1984:IV)

Quando Kim e Nelson incluíram os dados de 1985 a 1995, o modelo não consegue detectar dois regimes. Por isso, foi proposto o modelo para as médias:

$$\mu_{S_t} = (\mu_0 + \mu_0^* S_t)(1 - D_t) + (\mu_1 + \mu_1^* D_t) S_t,$$

com D_t igual a 1 no período 1983:I-1995:III e zero no período anterior.

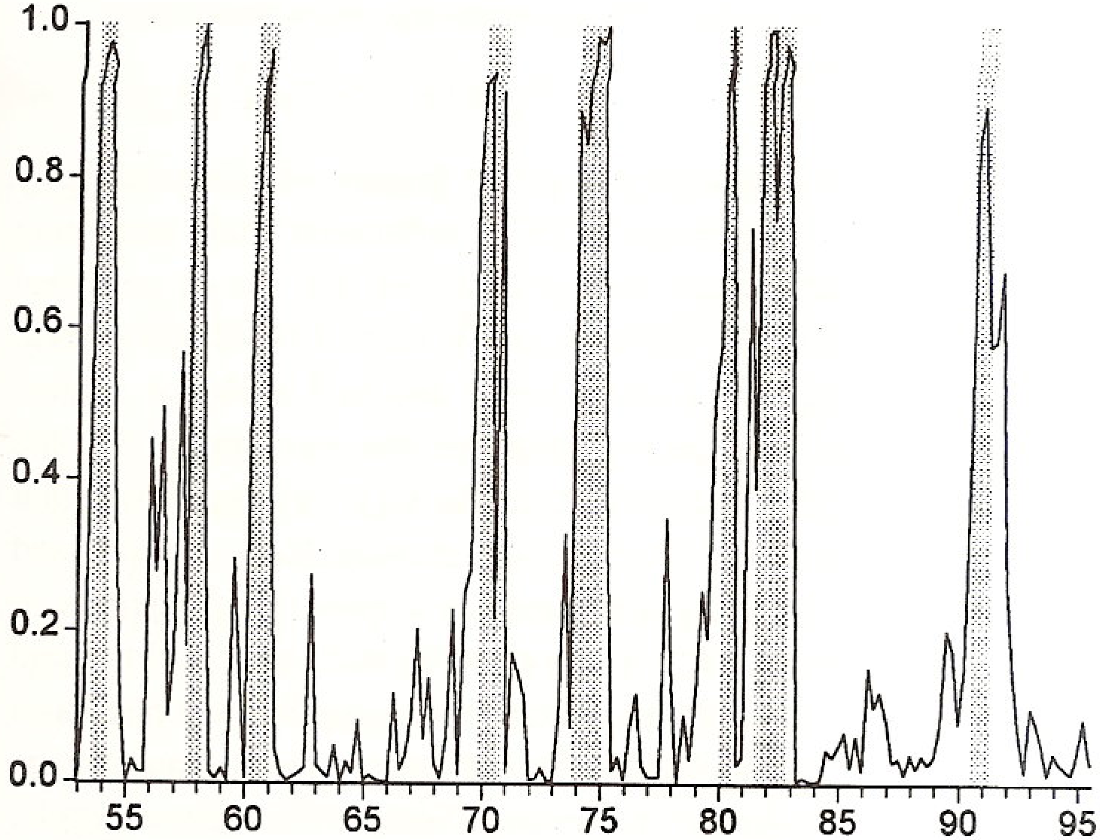


Figure 4.4
Filtered probability of a recession (GDP: 1952:II–1995:III; model with dummy variables)

Modelo Threshold Auto-regressivo

TAR

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t}, se \;\; s_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t}, se \;\; s_{t-k} \geq r \end{cases}$$

SETAR = Self-exciting TAR

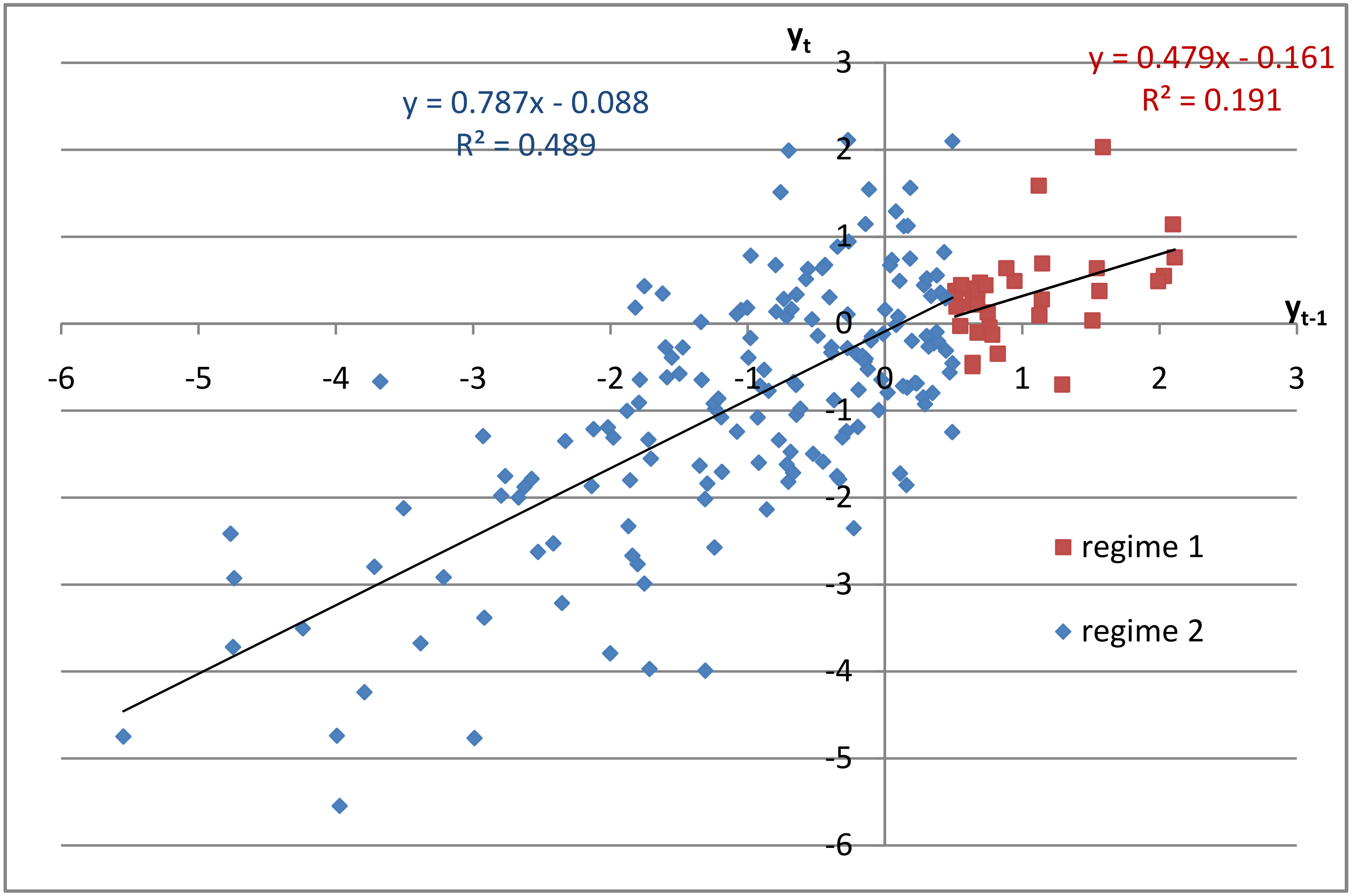
$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t}, se \;\; y_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t}, se \;\; y_{t-k} \geq r \end{cases}$$

Tem que estimar $\mu_1, \mu_2, \phi_1, \phi_2, k, r$ e as variâncias de u_{it}.

- Estimação: máxima verossinilhança e r, k estimados por grid search.
- Pode ter mais que dois regimes e diferentes variáveis para definir os regimes.
- Pode ser usado algum critério tipo AIC para escolher o melhor modelo.

Dados simulados de

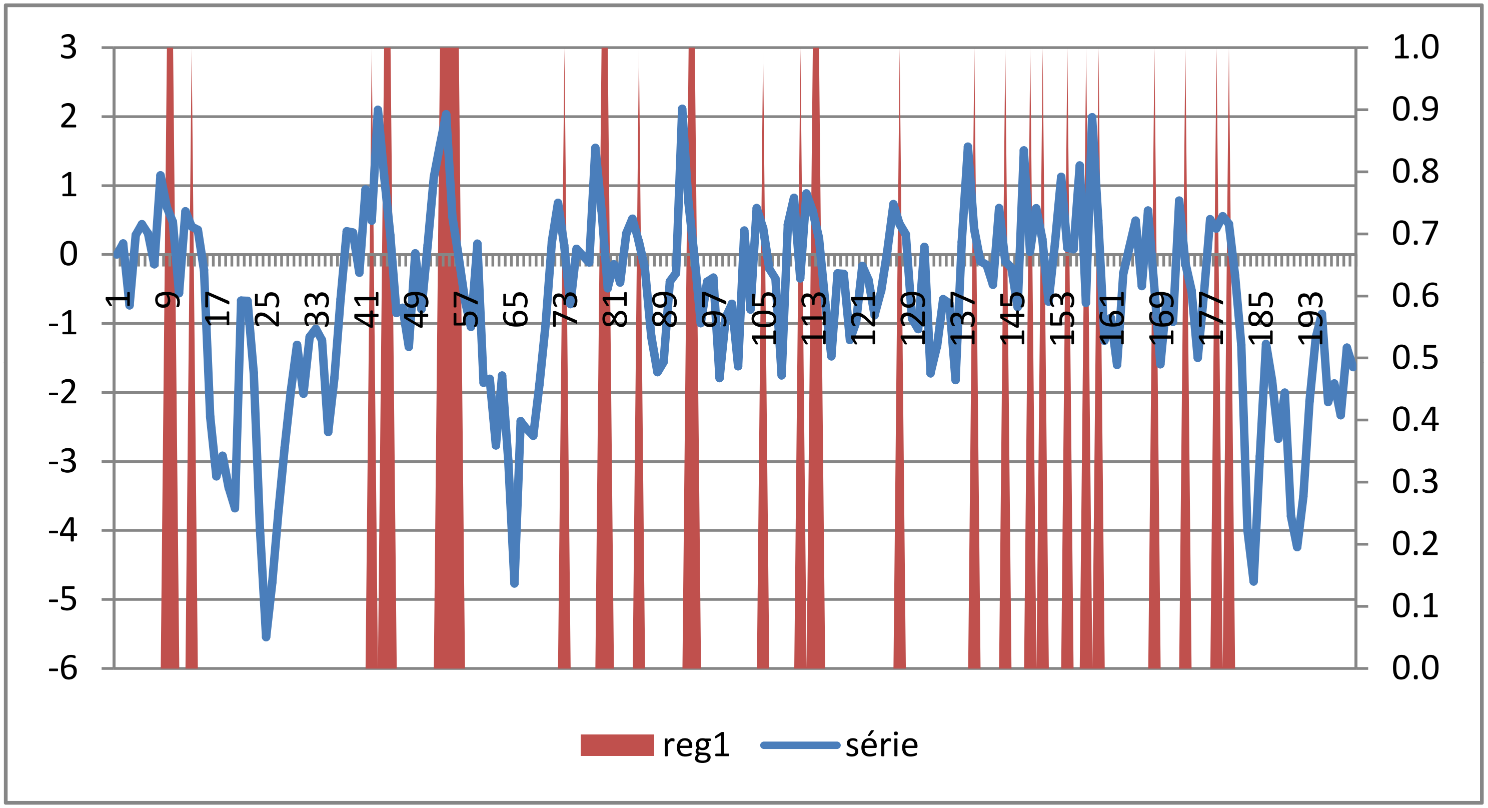
$$y_t = \begin{cases} 0,2y_{t-1} + 0,5e_t, se \;\; y_{t-1} > 0,5 \\ 0,8y_{t-1} + e_t, se \;\; y_{t-1} \leq 0,5 \end{cases}$$



y = 0.787x - 0.088
R² = 0.489
y = 0.479x - 0.161
R² = 0.191
yt
yt-1
regime 1
regime 2

Valores verdadeiros e estimativas obtidas pelo método de Máxima Verossimilhança

parâmetros	verdadeiro	estimativa
ϕ_1	0.20	0.35
ϕ_q	0.80	0.82
σ_1^2	0.50	0.51
σ_2^2	1.00	1.01
r	0.50	0.52



3
2
1
0
-1
-2
-3
-4
-5
-6
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1
9
17
25
33
41
49
57
65
73
81
89
97
105
113
121
129
137
145
153
161
169
177
185
193
reg1
série

Referências

[1] Kim, C. J. e Nelson, C. R. (1999). State-space models with regime switching. MIT Press.

[2] Hamilton, J. D. (1994). *Time series analysis*. Princeton, N.J.: Princeton University Press.

[3] Tong, H. (1990). Non-Linear Time Series: A Dynamical System Approach. New York: Oxford University Press.