Referências Bibliográficas

[1] BOX, G.; JENKINS, G. ; REINSEL, G.. **Time Series Analysis, Forecasting and Control**. Prentice Hall, New Jersey, 1994.

[2] BREIMAN, L.; FRIEDMAN, J.; OLSHEN, R. ; STONE, C. J.. **Classification and Regression Trees**. Belmont Wadsworth Int. Group, New York, 1984.

[3] CHAN, K.; TONG, H.. **On estimating thresholds in autoregressive models**. Journal of Time Series Analysis, 7:179–190, 1986.

[4] FAMA, E.. **Random walks in stock market prices**. Financial Analysts Journal, 51:75–80, 1965.

[5] GREENE, W. H.. **Econometrics Analysis**. Macmillan Publishing Company, New York, 1993.

[6] GRANGER, C.; TERÄSVIRTAL, T.. **Modelling Nonlinear Economic Relationships**. Oxford University Press, Oxford, 1993.

[7] HAMILTON, J.. **Time Series Analysis**. Princeton University Press, Princeton, 1994.

[8] CORRÊA DA ROSA, J.. **Modelos de Regressão com Transição Suave Estruturados por Árvores**. PhD thesis, PUC-Rio, Departmento de Engenharia Elétrica, Rio de Janeiro, 2005.

[9] KUAN, C.-M.; LIM, T.. **Forecasting exchange rates using feedforward and recurrent neural networks**. Journal of Applied Econometrics, 10:347–364, 1994.

[10] LJUNG, G.; BOX, G.. **On a measure of lack of fit in time series models**. Biometrika, 66:265–270, 1979.

[11] LEVENBERG, K.. **A method for the solution of certain problems in last squares**. Quarterly of Applied Mathematics, 2:164–168, 1944.

[12] LUUKKONEN, R.; SAIKKONEN, R. ; TERÄSVIRTA, T.. **Testing linearity against smooth transition autoregressive models**. Biometrika, 75:491–499, 1988.

[13] MEDEIROS, M.; VEIGA, A.. **A flexible coefficient smooth transition time series model**. Working Paper Seiries in Economics ans Finance, 360, 2000.

[14] MARQUARDT, D.. **An algorithm for least-squares estimation of nonlinear parameters**. SIAM Journal on Applied Mathematics, 11:431–441, 1963.

[15] TERÄSVIRTA, T.. **Specification, estimation, and evaluation of smooth transition autoregressive models**. Journal of the American Statistical Association, 89:208–218, 1994.

[16] TONG, H.. **On a threshold model**. In: Chen, C. H., editor, IN PATTERN RECOGNITION AND SIGNAL PROCESSING, volumen 5. Sijhoff Noordhoff, Amsterdam, 1978.

[17] TONG, H.; LIM, K.. **Threshold autoregression, limit cycles and cyclical data**. Journal of the Royal Statistical Society, Series B, Methodological, 42:245–292, 1980.

[18] TONG, H.. **Threshold models in non-linear time series analysis**. In: Verlag, S., editor, SPRINGER LECTURE NOTES IN STATISTICS, volumen 21. Springer Verlag, 1983.

[19] ZADEH, L.. **Fuzzy sets**. Information and Control, 8:338–353, 1965.

[20] VAN DIJK, D.; T.TERÄSVIRTA ; FRANSES, P.. **Smooth transition autoregressive models**. Working Paper Series in Economic and Finance, 380, 2000.

[21] VAN DIJK, D.; FRANSES, P.. **Modelling multiple regimes in the business cycle**. Macroeconomic Dynamics, 3:311–340, 1999.

A
Aproximação de Taylor para a Função Logística

Para resolver o problema de identificação ocorrido no teste ML, sugere-se a aproximação de Taylor de terceira ordem ao redor de $\gamma = 0$ para a função logística usada

$$G(s_t, \gamma, c) = \frac{e^{-\gamma(s_t - c)}}{1 + e^{-\gamma(s_t - c)}}, \quad \gamma > 0 \tag{A-1}$$

onde γ é o parâmetro que determina o grau de suavidade da transição e c é o limiar entre dois regimes.

A equação geral da aproximação pelo polinômio de Taylor de terceira ordem ao redor de um valor a é dada por

$$f(x) = f(a) + \left.\frac{df}{dx}\right|_{x=a} \cdot (x-a) + \frac{1}{2!}\left.\frac{d^2f}{dx^2}\right|_{x=a} \cdot (x-a)^2 + \frac{1}{3!}\left.\frac{d^3f}{dx^3}\right|_{x=a} \cdot (x-a)^3 + R_3(x) \tag{A-2}$$

Considerando (A-1), a aproximação fica na forma a seguir

$$\begin{aligned} G(s_t, \gamma, c) =& \left.G(s_t, \gamma, c)\right|_{\gamma=0} + \left.\frac{dG}{d\gamma}\right|_{\gamma=0} (\gamma - 0) + \frac{1}{2!}\left.\frac{d^2G}{d\gamma^2}\right|_{\gamma=0} (\gamma - 0)^2 + \\ & + \frac{1}{3!}\left.\frac{d^3G}{d\gamma^3}\right|_{\gamma=0} (\gamma - 0)^3 + R_3(\gamma) \\ =& \frac{1}{2} + \frac{1}{4}\gamma(-s_t + c) - \frac{1}{48}\gamma^3(-s_t + c)^3 + R_3(\gamma) \end{aligned} \tag{A-3}$$

onde

$$
\begin{aligned}
\left.\frac{dG}{d\gamma}\right|_{\gamma=0} &= \left.\frac{(-s_t+c)\, e^{-\gamma(s_t-c)}}{1+e^{-\gamma(s_t-c)}}\right|_{\gamma=0} - \left.\frac{\left(e^{-\gamma(s_t-c)}\right)^2(-s_t+c)}{\left(1+e^{-\gamma(s_t-c)}\right)^2}\right|_{\gamma=0} = \frac{1}{4}(-s_t+c) \\
\left.\frac{d^2G}{d\gamma^2}\right|_{\gamma=0} &= \left.\frac{(-s_t+c)^2\, e^{-\gamma(s_t-c)}}{1+e^{-\gamma(s_t-c)}}\right|_{\gamma=0} - 3\left.\frac{(-s_t+c)^2\left(e^{-\gamma(s_t-c)}\right)^2}{\left(1+e^{-\gamma(s_t-c)}\right)^2}\right|_{\gamma=0} + \\
&\quad + 2\left.\frac{\left(e^{-\gamma(s_t-c)}\right)^3(-s_t+c)^2}{\left(1+e^{-\gamma(s_t-c)}\right)^3}\right|_{\gamma=0} = 0 \\
\left.\frac{d^3G}{d\gamma^3}\right|_{\gamma=0} &= \left.\frac{(-s_t+c)^3\, e^{-\gamma(s_t-c)}}{1+e^{-\gamma(s_t-c)}}\right|_{\gamma=0} - 7\left.\frac{(-s_t+c)^3\left(e^{-\gamma(s_t-c)}\right)^2}{\left(1+e^{-\gamma(s_t-c)}\right)^2}\right|_{\gamma=0} + \\
&\quad + 12\left.\frac{(-s_t+c)^3\left(e^{-\gamma(s_t-c)}\right)^3}{\left(1+e^{-\gamma(s_t-c)}\right)^3}\right|_{\gamma=0} - 6\left.\frac{\left(e^{-\gamma(s_t-c)}\right)^4(-s_t+c)^3}{\left(1+e^{-\gamma(s_t-c)}\right)^4}\right|_{\gamma=0} \\
&= -\frac{1}{8}\,(-s_t+c)^3
\end{aligned}
\tag{A-4}
$$

B
Algoritmo de Levenberg-Marquardt

O algoritmo de Levenberg-Marquardt [11] e [14], é um algoritmo de otimização não linear usado para o cálculo de minimização de funções. Este algoritmo é usado especialmente em problemas que requerem o uso de mínimos quadrados não lineares. A função a ser minimizada possui a forma

$$f(\boldsymbol{\theta}) = \frac{1}{2}\sum_{i=1}^{n} e_i^2(\boldsymbol{\theta}) \qquad \text{(B-1)}$$

onde $\boldsymbol{\theta}$ é o vetor de parâmetros de interesse e $\mathbf{e} = (e_1, e_2, ..., e_n)'$ são os resíduos.

O algoritmo interativo para a estimação da solução geral de (B-1) é da forma de (B-2) e é chamado de algoritmo do gradiente descendente, pois a rotina de procura pelo valor mínimo é feita com base na direção oposta ao gradiente da função.

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \mu\nabla f, \quad i = 0, 1, ..., i \qquad \text{(B-2)}$$

Na equação (B-2), parte-se de um valor inicial do vetor de parâmetros $\boldsymbol{\theta}_0$ e, em cada interação i, decrementa-se o vetor $\boldsymbol{\theta}_i$, ponderado pelo escalar $\mu, \mu > 0$, na direção contrária ao gradiente.

Como no ponto de mínimo da soma dos erros quadráticos tem-se $\nabla f(\boldsymbol{\theta}) = 0$ e fazendo uso da linearização resultante da aproximação de Taylor do gradiente ao redor do estado atual $\boldsymbol{\theta}_0$ do algoritmo, chega-se a

$$\nabla f = \nabla f(\boldsymbol{\theta}_0) + (\boldsymbol{\theta} - \boldsymbol{\theta}_0)'\nabla^2 f(\boldsymbol{\theta}_0) = 0 \qquad \text{(B-3)}$$

Desenvolvendo (B-3), obtém-se a regra de atualização para o método de Newton.

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - (\nabla^2 f(\boldsymbol{\theta}_i)^{-1})\nabla f(\boldsymbol{\theta}_i) \qquad \text{(B-4)}$$

Fazendo uso da aproximação dos resíduos por funções lineares e considerando-os muito pequenos, pode-se chegar a

$$\nabla f(\boldsymbol{\theta}) = J(\boldsymbol{\theta})'\mathbf{e} \tag{B-5}$$

$$\nabla^2 f(\boldsymbol{\theta}) = J(\boldsymbol{\theta})'J(\boldsymbol{\theta}) \tag{B-6}$$

onde $J(\boldsymbol{\theta})$ é o Jacobiano dos resíduos $\mathbf{e}$ com relação aos parâmetros $\boldsymbol{\theta}$.

O algoritmo de Levenberg-Marquardt é derivado do algoritmo de Newton e apresenta a forma a seguir

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - (\boldsymbol{H} + \mu diag[\boldsymbol{H}])^{-1}\nabla f(\boldsymbol{\theta}_i) \tag{B-7}$$

onde $\boldsymbol{H}$ é a Hessiana aplicada ao ponto $\boldsymbol{\theta}_i$.

C
Testes de Diagnósticos

Este apêndice traz a descrição teórica dos testes de diagnósticos usados para validação do modelo STAR-Tree. Os testes abordados são:

- Resíduos descorrelatados (C.1)
- Não linearidade não remanescente (C.2)

C.1
Resíduos Descorrelatados

Suponha o modelo de regressão da equação (C-1) e o modelo STAR(1) (C-2)

$$\mathbf{y_t} = \mathbf{X_t}\boldsymbol{\theta} + \boldsymbol{\epsilon_t} \quad \text{(C-1)}$$

$$y_t = (\alpha_1 + \beta_1 y_{t-1})G(s_t;\gamma,c) + (\alpha_2 + \beta_2 y_{t-1})(1 - G(s_t;\gamma,c)) + \epsilon_t \quad \text{(C-2)}$$

Fazendo a associação entre as equações (C-1) e (C-2), chega-se a

$$\mathbf{X_t} = \begin{bmatrix} G(s_t;\gamma,c) & y_{t-1}G(s_t;\gamma,c) & 1 - G(s_t;\gamma,c) & y_{t-1}(1 - G(s_t;\gamma,c)) \end{bmatrix}$$

$$\boldsymbol{\theta} = \begin{bmatrix} \alpha_1 \\ \beta_1 \\ \alpha_2 \\ \beta_2 \end{bmatrix} \quad \text{(C-3)}$$

O teste proposto para verificação de correlação serial dos resíduos do modelo foi sugerido por Ljung e Box [10]. O teste é do tipo Multiplicador de Lagrange e a hipótese nula avalia a correlação dos resíduos até a ordem m.

$$\begin{aligned} H_0 &: \rho(1) = \rho(2) = \ldots = \rho(m) = 0 \\ H_1 &: \rho(1) \neq 0 \text{ ou } \rho(2) \neq 0 \text{ ou } \ldots \text{ ou } \rho(m) \neq 0 \end{aligned} \tag{C-4}$$

A estatística de teste (C-5) é construída a partir da regressão de e_t, os resíduos do modelo estimado, em $\mathbf{X_t}$ e, sob a hipótese nula, segue uma distribuição assintótica χ^2 com m graus de liberdade.

$$Q_{LB} = T(T+2)\sum_{j=1}^{m} \frac{r_j^2}{T-j} \tag{C-5}$$

onde T é o número de observações e $r_j = \frac{\sum_{t=j+1}^{T} e_t e_{t-j}}{\sum_{t=1}^{T} e_t^2}$ é a j–ésima autocorrelação.

A escolha do valor de m apresenta um problema: caso a ordem escolhida seja pequena, o teste pode não detectar a correlação com ordens superiores, e caso a ordem seja grande, correlações relevantes podem ser diluídas por correlações insignificantes de outras ordens.

C.2 Não linearidade não remanescente

O teste utilizado aqui, proposto por van Dijk e Franses [21], é do tipo Multiplicador de Lagrange e confronta a hipótese nula de um modelo LSTAR(C-6) contra um modelo MRSTAR(C-7).

$$y_t = \phi_1' x_t + (\phi_2 - \phi_1)' x_t G(s_t;\gamma,c) + \epsilon_t \tag{C-6}$$

$$\begin{aligned} y_t = {} & [\phi_1' x_t G_1(s_{1t};\gamma_1,c_1) + \phi_2' x_t (1 - G_1(s_{1t};\gamma_1,c_1))] G_2(s_{2t};\gamma_2,c_2) + \\ & + [\phi_3' x_t G_1(s_{1t};\gamma_1,c_1) + \phi_4' x_t (1 - G_1(s_{1t};\gamma_1,c_1))](1 - G_2(s_{2t};\gamma_2,c_2)) + \epsilon_t \end{aligned} \tag{C-7}$$

A hipótese nula deste teste é definida como $H_0 : \gamma_2 = 0$ ou $H_0' : \phi_1 = \phi_3$ e $\phi_2 = \phi_4$. Com isso, deve-se recorrer à aproximação de Taylor de terceira ordem da função $G_2(s_{2t};\gamma_2,c_2)$ ao redor de $\gamma_2 = 0$ para contornar o problema de identificação já exposto. A equação abaixo traz este resultado.

$$\begin{aligned} y_t = & \theta_1' x_t + \theta_2' x_t G_1(s_{1t};\gamma_1, c_1) + \beta_1' x_t s_{2t} + \beta_2' x_t s_{2t}^2 + \beta_3' x_t s_{2t}^3 + \\ & (\beta_4' x_t s_{2t} + \beta_5' x_t s_{2t}^2 + \beta_6' x_t s_{2t}^3) G_1(s_{1t};\gamma_1, c_1) + e_t \end{aligned} \tag{C-8}$$

onde $\beta_i = (\beta_{i,1}, ..., \beta_{i,p})',\ i = 1, ..., 6$ é função de $\phi_i,\ i = 1, ...4$, γ_2 e c_2.

A nova hipótese nula passa a ser $H_0'' : \beta_i = 0, i = 1, ..., 6$. Sob a hipótese nula, tem-se $\theta_1 = \phi_1$, $\theta_2 = \phi_2 - \phi_1$ e $e_t = \epsilon_t$. A estatística de teste é então obtida através das seguintes etapas:

1. Estimação do modelo LSTAR por mínimos quadrados não lineares e obtenção da soma dos quadrados dos resíduos, $SSR_0 = \sum_{t=1}^{T} \hat{\epsilon}_t^2$ sob a hipótese nula.

2. Estimação dos resíduos e_t e obtenção da soma dos quadrados de $\hat{e}_t$ como $SSR_1 = \sum_{t=1}^{T} \hat{e}_t^2$.

3. A estatística de teste sob a hipótese nula seguirá uma distribuição χ^2 com $6(p+1)$ graus de liberdade e terá a forma

$$LM = \frac{T(SSR_0 - SSR_1)}{SSR_0} \tag{C-9}$$

O procedimento é similar para o caso de inclusão de regimes adicionais.