

## 5

### Modelos não lineares para séries temporais – *STAR(p)*

Nas últimas décadas, os modelos não-lineares começaram a ser mais desenvolvidos e surgiram os modelos de troca de regime tais como o modelo autorregressivo com limiar (*TAR*), modelos autorregressivo com transição suave (*STAR*) e o modelo de troca de regime markoviano (*Markov Switching Models – MSM*) (Krolzig, 1998). Estes modelos encontraram um grande número de aplicações de sucesso na literatura, a idéia básica dos mesmos é atribuir modelos lineares distintos em regiões distintas onde se encontram os valores das séries em estudos. Nos casos do *STAR* e *MSM* as transições entre cada modelo são suaves.

Outro modelo também com grande destaque na literatura é o de Redes Neurais Artificiais (RNA) (Haykin, 2001) e pode-se interpretá-los como uma forma especial de transição suave.

Abaixo é elaborada uma relação dos principais modelos:

- 1) Modelos de múltiplos regimes autorregressivo (*TAR*)
  - 1.1) Modelos de múltiplos regimes autorregressivo com limiar auto-excitante (*SETAR*)
  - 1.2) Modelos de múltiplos regimes autorregressivo com limiar auto-excitante multivariado (*M - SETAR*)
- 2) Modelos autorregressivos com transição suave (*STAR*)
  - 2.1) Modelos autorregressivo com transição suave logística (*LSTAR*)
  - 2.2) Modelos autorregressivo com transição suave exponencial (*ESPAR*)
- 3) Redes neurais artificiais autorregressivo (*AR-ANN*)
- 4) Modelo de troca de regime markoviano (*MSM*)
- 5) Modelos GARCH não lineares (*TARCH* e *EGARCH*)

É importante dizer que, obviamente, existem ainda outros modelos não lineares disponíveis na literatura. Boas referências para estes outros modelos e os acima citados podem ser encontradas em Gooijer & Kumar (1992), Granger & Teräsvirta (1993), Chen & Tsay (1993), Astatkie *et al.* (1997), Medeiros (1998), Lundbergh *et al.* (2000), Van Dijk *et al.* (2000). E para um leitor que não está

muito familiarizado com modelos não lineares, uma introdução a esse tipo de modelo é encontrada em Tiao & Tsay (1994) e Potter (1999).

O modelo TAR, como uma idéia de aproximar uma função não-linear geral por um modelo linear por partes, foi proposto primeiramente por Tong (1978) e obteve forte desenvolvimento por Tong & Lim (1980) e Tong (1983/1990). Em Chan & Tong (1986) é descrito uma generalização dos modelos TAR com uma transição mais suave, tal generalização ficou conhecida como modelos *STAR* (*Smooth Transition Autoregressive*). Posteriormente, Teräsvirta (1994) desenvolveu ainda mais este tipo de modelo. Em 1997, Liu & Li propõem um modelo auto-regressivo com um limiar que considere a heterocedasticidade da variância condicional da série temporal.

A sua popularidade se deve ao fato de sua facilidade de especificação, estimação e interpretação quando comparado com outros modelos não-lineares. Todavia apesar do desenvolvimento de inferência por Hansen (1997) pouco se sabe a respeito das propriedades estatísticas de seus estimadores. A idéia principal dos modelos TAR é descrever o seu processo por partes, ou seja, alterar os parâmetros de um modelo auto-regressivo linear dependendo da região em que se encontrem valores de uma determinada variável, isto é, onde a quebra de cada modelo depende da região a que pertence o valor da variável.

Seja  $y_t$  o processo da série,  $Z_t \in R_i^{(r)}$  a variável de estado unidimensional e  $R_i^{(r)}$ ,  $i = 1, \dots, k$ , uma partição do  $R^{(r)}$ , pode-se então escrever o modelo TAR como uma composição de  $L$  modelos autorregressivos, cada um ativo dentro da região em que se encontra o valor da variável  $Z_t$ , isto é,

$$y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^{k_1} \phi_i^{(1)} y_{t-i} + \varepsilon_t^{(1)} & \text{se } z_t \in \mathfrak{R}_1 \\ \phi_0^{(2)} + \sum_{i=1}^{k_2} \phi_i^{(2)} y_{t-i} + \varepsilon_t^{(2)} & \text{se } z_t \in \mathfrak{R}_2 \\ \vdots & \\ \phi_0^{(L)} + \sum_{i=1}^{k_L} \phi_i^{(L)} y_{t-i} + \varepsilon_t^{(L)} & \text{se } z_t \in \mathfrak{R}_L \end{cases} \quad (5.1)$$

A escolha da variável  $Z_t$  é bem flexível e por esta razão mostra a dinâmica e o grande número de modelos possíveis.

Quando se tem  $Z_t = y_{t-d}$ , o modelo resultante é conhecido como SETAR( $L, k_1, \dots, k_L$ ) – autorregressivo de múltiplos regimes auto-excitante - e a sua formulação é:

$$y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^{k_1} \phi_i^{(1)} y_{t-i} + \varepsilon_t^{(1)} & \text{se } y_{t-d} \in \mathfrak{R}_1 \\ \phi_0^{(2)} + \sum_{i=1}^{k_2} \phi_i^{(2)} y_{t-i} + \varepsilon_t^{(2)} & \text{se } y_{t-d} \in \mathfrak{R}_2 \\ \vdots \\ \phi_0^{(L)} + \sum_{i=1}^{k_L} \phi_i^{(L)} y_{t-i} + \varepsilon_t^{(L)} & \text{se } y_{t-d} \in \mathfrak{R}_L \end{cases} \quad (5.2)$$

Definindo  $\mathfrak{R}_j = (r_{j-1}, r_j)$ , denomina-se de parâmetros limiares o conjunto  $\{r_0, \dots, r_L\}$  e  $d$  o parâmetro de defasamento e os  $\varepsilon_t^{(j)} \sim \text{NID}(0, \sigma^2 \varepsilon^{(j)})$ .

Uma generalização imediata do modelo anterior é o SETARMA  $(L; k_1, \dots, k_L; k'_1, \dots, k'_L)$ , definido por:

$$y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^{k_1} \phi_i^{(1)} + \sum_{i=1}^{k'_1-1} \theta_i^{(1)} \varepsilon_{t-i}^{(1)} & \text{se } y_{t-d} \in \mathfrak{R}_1 \\ \phi_0^{(2)} + \sum_{i=1}^{k_2} \phi_i^{(2)} + \sum_{i=1}^{k'_2-1} \theta_i^{(2)} \varepsilon_{t-i}^{(2)} & \text{se } y_{t-d} \in \mathfrak{R}_2 \\ \phi_0^{(L)} + \sum_{i=1}^{k_L} \phi_i^{(L)} + \sum_{i=1}^{k'_L-1} \theta_i^{(L)} \varepsilon_{t-i}^{(L)} & \text{se } y_{t-d} \in \mathfrak{R}_L \end{cases} \quad (5.3)$$

O modelo SETAR pode ser reescrito em notação vetorial da seguinte forma:

$$y_t = \alpha' z_t + \sum_{i=1}^h \lambda'_i z_t I_i(y_{t-d}) + \varepsilon_t \quad (5.4)$$

Onde  $\alpha' = [\alpha_0, \dots, \alpha_p]$ ,  $\lambda'_i = [\lambda_{0i}, \dots, \lambda_{pi}]$ , e  $z'_t = [1, z'_t] = [1, y_{t-1}, \dots, y_{t-p}]$

Nessa notação foi introduzida uma função que é chamada de indicadora; ela é definida como:

$$I_i(Z_t) = \begin{cases} 1, & \text{se } Z_t \geq r_i \\ 0, & \text{caso contrário} \end{cases}$$

Onde  $Z_t$  é o parâmetro de defasamento ( $y_{t-d}$ ) e  $r_i$  são os parâmetros de limiares linearmente ordenados. Pode-se perceber que neste modelo as mudanças de regime são abruptas, caracterizando descontinuidades. Este modelo está sendo exposto, pois apresenta uma única característica (a descontinuidade) que não é semelhante ao STAR. É importante dizer que essa característica implica em outras diferenças na estratégia de modelagem desses modelos. Para maiores detalhes dessa estratégia em modelos SETAR veja TSAY (1986, 1989 e 1998), Clements & Smith (1997), Clements & Krolzig (1998).

Portanto, trocando  $I_r$  por uma função de “amortecimento”  $F$  o modelo se torna o modelo autorregressivo com transição suave (*STAR*). A escolha de  $F$  é bem flexível. Por exemplo, pode-se trocar  $I_r(x)$  pela função de distribuição da normal padrão, ou então, pela função logística dada por:

$$\left( \frac{1}{1 + e^{-\gamma(x-c)}} \right)$$

Daí o  $L$  antes do *STAR*, significando modelo autorregressivo com transição suave logística ou *LSTAR*. A figura 7 é o gráfico da função logística para diferentes valores do  $\gamma$ .

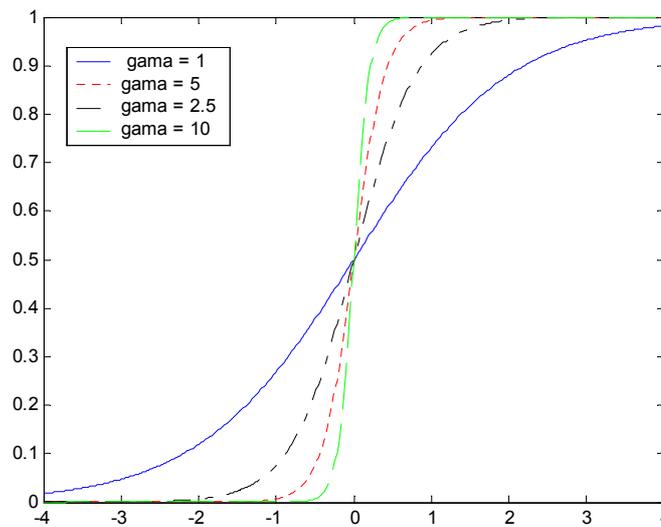


Figura 1 – Função Logística para diferentes valores do  $\gamma$

Nesta pesquisa utilizou-se a função logística como a função de transição. A única condição imposta a esta função é de que seja contínua e não decrescente para  $\gamma > 0$ . Esta função tem que ser sempre contínua, entretanto, não existem restrições para que ela seja decrescente desde que o  $\gamma < 0$ , caso contrário, o modelo é não identificável.

Pode-se perceber, analisando a figura 7, que quando  $\gamma \rightarrow \infty$  o modelo *LSTAR* converge para o *SETAR* (Tong 1990, p. 183).

A formulação geral vetorial do modelo *LSTAR* é mostrada abaixo,

$$Y_t = G(Z_t, x_t; \Psi) + \varepsilon_t = \alpha' Z_t + \lambda' Z_t f(\gamma(x_t - c)) + \varepsilon_t, \quad (5.5)$$

onde

$$f(\gamma(x_t - c)) = \frac{1}{1 + e^{-\gamma(x_t - c)}} \quad (5.6)$$

Onde  $G(\cdot)$  é uma função não linear de  $z_t$  e  $x_t$ , definida pelo vetor de parâmetros  $\psi$ .  $\alpha$  e  $\lambda$  são vetores de parâmetros e  $z_t$  é um vetor de variáveis. Os erros são considerados NID com média 0 e variância  $\sigma^2$ . Note que é possível utilizar outra variável para explicar a dinâmica do modelo, isto é, se utilizou  $x_t$  ao invés de  $y_{t-d}$  sem perda de generalização.

## 5.1.

### **Especificação, teste de linearidade, estimação dos parâmetros e validação para modelos *LSTAR***

Os modelos *LSTAR* podem seguir o mesmo ciclo de modelagem da metodologia Box & Jenkins. Ou seja, especificação, estimação dos parâmetros e validação do modelo. É importante dizer que para muitos outros modelos não lineares isso não seria possível. Pode-se citar artigos de Teräsvirta (1994) e Eitrheim e Teräsvirta (1996) para o *LSTAR*.

#### 5.1.1.

##### **Estacionariedade**

Um ponto muito importante nesta formulação é a necessidade de que o processo gerador dos dados seja ergódico. Entretanto, como isso é uma condição muito forte para uma série temporal, a modelagem pode ser aplicada com resultados satisfatórios, desde que a distribuição dos dados seja estacionária de segunda ordem. Para isso, deve-se aplicar testes de estacionariedade na série temporal em estudo para saber se a série é estacionária. Os testes utilizados para tal serão o de *Dickey Fuller* aumentado (ADF) e o *Phillips Perron* (PP). Entretanto, não existem garantias de que se trabalhando com uma série não estacionária, o processo se tornaria instável, ou seja, a imposição de que a série seja estacionária é uma grande controvérsia neste tipo de modelo em atividades práticas, como é o caso desta dissertação. No próximo capítulo serão apresentados todos os resultados de todos os procedimentos que serão abordados neste capítulo para a série do CMO.

#### 5.1.2.

##### **Especificação**

A primeira análise a ser feita diz respeito à especificação. Primeiramente, é necessária a escolha da ordem da autorregressão. Para isso, propõe-se estimar modelos AR de ordem 1 até ordem  $p$  e selecionar a ordem que minimize algum

critério (Teräsvirta, 1994), como por exemplo o BIC (Schwarz, 1978) ou o AIC (Akaike, 1974) definidos como:

$$BIC = T \ln(\hat{\sigma}) + 2K, \quad (5.7)$$

$$AIC = T \ln(\hat{\sigma}) + K \ln(T)$$

Todos esses dois critérios são estruturalmente similares desde que envolvem uma estimativa da função do log da verossimilhança esperada do modelo em consideração e um termo de penalidade que depende diretamente ou indiretamente do número de parâmetros estimados do modelo e do número de observações.

Existem outros métodos como a utilização da função de autocorrelação parcial ou, então, utilização de métodos não paramétricos, entretanto, este último método requer uma enorme quantidade de observações e o custo computacional é muito elevado.

### 5.1.3. Teste de linearidade

O próximo passo é saber se o processo é não linear, pois se o processo for linear, não é necessário à aplicação de uma modelagem não linear para o problema tendo em vista que os modelos lineares desenvolvidos até os dias de hoje seriam mais que suficientes para solucionar o problema. Além disso, o teste serve para escolher a variável de transição caso utilize-se um valor defasado da própria série temporal. Efetua-se o teste em várias variáveis (ordens) candidatas e escolhe-se aquela que minimizar o p-valor do teste.

O teste implementado nesta dissertação foi o proposto por Teräsvirta (1994). Consiste em um teste do tipo LM (*Lagrange Multiplier*) e se baseia em regressões auxiliares para construir a estatística do teste. Para maiores detalhes da formulação consulte também Luukkonen, Saikkonen e Teräsvirta (1988) e sobre testes do tipo LM veja Breusch e Pagan (1980).

Considere o modelo LSTAR em (5.5) e o reescreva como

$$Y_t = \alpha' Z_t + \lambda' Z_t f^*(\gamma(x_t - c)) + \varepsilon_t, \quad (5.8)$$

onde

$$f^*(\gamma(x_t - c)) = \frac{1}{1 + e^{-\gamma(x_t - c)}} - \frac{1}{2} \quad (5.9)$$

Subtraindo-se  $\frac{1}{2}$  da função logística é útil somente na derivação do teste de linearidade onde simplifica a notação, mas não afeta o argumento. Os modelos estimados neste trabalho não possuem este termo.

Note que sob a hipótese nula de que  $\gamma = 0$  a função  $f^* = 0$ . Luukkonen, Saikkonen e Teräsvirta (1988) sugerem aproximar a função  $f^*$  por uma expansão de Taylor de primeira ordem em torno de  $\gamma = 0$ , que é,

$$T_1(x_t; \gamma, c) \approx f^*(x_t; 0, c) + \gamma \left. \frac{\partial f^*(x_t; \gamma, c)}{\partial \gamma} \right|_{\gamma=0} = \frac{1}{4} \gamma (x_t - c), \quad (5.10)$$

onde usou-se o fato de que  $f^*(x_t; 0, c) = 0$ . Depois de substituir  $f_1^*(.)$  para  $T_1(.)$  em (5.8) e reordenando os termos, tem-se o modelo de regressão auxiliar

$$Y_t = \beta_{0,0} + \beta'_1 \tilde{Z}_t + \beta'_2 \tilde{Z}_t x_t + \eta_t, \quad (5.11)$$

onde  $\tilde{Z}_t = (y_{t-1}, \dots, y_{t-p})'$  e  $\beta_j = (\beta_{1,j}, \dots, \beta_{p,j})$ ,  $j = 0, 1$ . Está criado, então, um simples teste de linearidade. A hipótese nula pode ser definida como  $H_0: \beta'_j = 0$ . Entretanto, os parâmetros  $\beta'_j$  não dependem de  $\lambda_0$ . Logo, quando a única componente não linear em (5.8) for intercepto e  $\tilde{Z}_t x_t = 0$ , o teste não tem potência. Para contornar essa situação, se usa uma expansão de terceira ordem, ou seja,

$$T_3(x_t; \gamma, c) \approx \frac{1}{4} \gamma (x_t - c) + \frac{1}{48} \gamma^3 (x_t - c)^3, \quad (5.12)$$

onde usou-se o fato de que a segunda derivada de  $f^*(x_t; \gamma, c)$  com respeito a  $\gamma$  calculado em  $\gamma = 0$  é igual a zero. Usando essa aproximação o modelo auxiliar é

$$Y_t = \beta_{0,0} + \beta'_1 \tilde{Z}_t + \beta'_2 \tilde{Z}_t x_t + \beta'_3 \tilde{Z}_t x_t^2 + \beta'_4 \tilde{Z}_t x_t^3 + \eta_t, \quad (5.13)$$

Onde  $\beta_{0,0}$  e  $\beta_j$ ,  $j = 1, 2, 3$ , novamente são funções dos parâmetros  $\alpha$ ,  $\lambda$ ,  $\gamma$ , e  $c$ . Inspeccionando-se as relações entre os parâmetros se demonstra que a hipótese nula  $H_0: \gamma = 0$  agora corresponde a  $H_0: \beta_1 = \beta_2 = \beta_3 = 0$ , que pode ser testado por um teste LM. Sob a hipótese nula, a estatística de teste assintótica segue uma distribuição  $\chi^2$  com 3p graus de liberdade.

Para pequenas amostras, a recomendação é que se use a distribuição F para os testes do tipo LM (Granger e Teräsvirta, 1993, capítulo 7). A versão F da estatística de teste baseado em (5.13) pode ser calculado da seguinte forma:

$$(1) \text{ Regrida } y_t \text{ em } Z_t \text{ e calcule } SSR_0 = \sum_{t=1}^n \hat{\varepsilon}_t^2$$

(2) Regrida  $\hat{\varepsilon}_t$  em  $x_t$  e  $\tilde{Z}_t x_t^j$ ,  $j = 1, 2, 3$ , e calcule a soma dos resíduos quadráticos  $SSR_1 = \sum_{t=1}^n \hat{v}_t^2$

(3) Calcule a estatística do teste como

$$LM_F = \frac{(SSR_0 - SSR_1) / 3p}{SSR_1 / (n - 4p - 1)}$$

onde  $n$  é o tamanho da amostra. Sob a hipótese nula, essa estatística segue uma distribuição F com  $3p$  e  $n - 4p - 1$  graus de liberdade. Lembrando que, caso a variável de transição seja um valor defasado da própria série temporal, o valor do  $d$  é escolhido para a ordem do teste que minimiza o p-valor do teste.

#### 5.1.4. Estimação dos parâmetros

Realizado o teste de linearidade, passa-se à fase de estimação dos parâmetros. Nesta fase o método utilizado é o dos mínimos quadrados não lineares ou por máxima verossimilhança condicional (quando  $\varepsilon_t \sim \text{NID}(0, \sigma^2)$ , os métodos são coincidentes). Logo o vetor de parâmetros  $\theta$  do modelo é estimado como:

$$\hat{\Psi} = \arg \min_{\Psi} Q_t(\theta) = \arg \min_{\Psi} \sum_{t=1}^T (y_t - G(Z_t, x_t; \theta))^2 \quad (5.14)$$

Sob certas condições de regularidade (que são discutidas em White e Domowitz, 1984; Gallant, 1987; Pötscher e Prucha, 1997; dentre outros), os estimadores são consistentes e assintoticamente normais, isto é:

$$\sqrt{T}(\hat{\theta} - \theta_0) \rightarrow N(0, C), \quad (5.15)$$

onde  $\theta_0$  é o vetor verdadeiro de parâmetros. A matriz de covariância dos estimadores  $C$  pode ser consistentemente estimada, aplicando-se os conceitos desenvolvidos em Davidson e MacKinnon (1993, capítulo 5), por

$$\hat{C} = \hat{\sigma}^2 (\hat{H}' \hat{H})^{-1}, \quad (5.16)$$

onde  $\hat{\sigma}^2$  é a variância estimada dos resíduos e  $\hat{H}$  é uma matriz com  $T$  linhas dadas por  $\nabla G(Z_t, x_t; \hat{\theta})$ .

Uma questão muito importante na estimação é a escolha dos valores iniciais dos parâmetros a serem estimados. O algoritmo de otimização é muito sensível a essa escolha.

A metodologia proposta é a seguinte: para valores fixos dos parâmetros da função de transição  $\gamma$  e  $c$ , o modelo *STAR* é linear nos parâmetros autorregressivos  $\alpha$  e  $\lambda$ . Então, condicional a  $\gamma$  e  $c$ , estimativas de  $\theta = (\alpha', \lambda')$  podem ser obtidas por mínimos quadrados ordinários (MQO) como

$$\hat{\theta}(\gamma, c) = (Z_t(\gamma, c)Z_t(\gamma, c)')^{-1}(Z_t(\gamma, c)y_t), \quad (5.17)$$

onde  $Z_t(\gamma, c) = [\tilde{1}, \tilde{y}_{t-1}, \dots, \tilde{y}_{t-p}]$  e a notação  $\hat{\theta}(\gamma, c)$  é usada para indicar que a estimativa de  $\theta$  é condicional a  $\gamma$  e  $c$ . Os correspondentes resíduos podem ser calculados como  $\hat{\varepsilon}_t = y_t - \hat{\theta}(\gamma, c)'Z_t(\gamma, c)$  com variância associada de  $\hat{\sigma}^2(\gamma, c) = n^{-1} \sum_{t=1}^n \hat{\varepsilon}_t^2(\gamma, c)$ . E para se encontrar valores de  $\gamma$  e  $c$  usou-se uma busca em grade bidimensional e selecionou-se os valores que apresentaram menor estimativa para a variância dos resíduos  $\hat{\sigma}^2(\gamma, c)$ .

Uma importante recomendação é de que se usem algoritmos como o Broyden-Fletcher-Goldfarb-Shanno (BFGS) ou Levenberg-Marquadt. Como não é o objetivo do trabalho demonstrar a formulação matemática dos algoritmos de otimização, para maiores detalhes veja Bertsekas (1995). Um outro ponto que merece ser mencionado é quanto à escolha da busca local a ser usada para seleção do tamanho do passo. Na modelagem da série do CMO utilizar-se-á, em um primeiro momento, o algoritmo de Levenberg-Marquadt com interpolação cúbica, entretanto, a interpolação quadrática também é recomendada.

Por último, salienta-se que é notoriamente difícil de se obter uma estimativa precisa do parâmetro de suavidade  $\gamma$ . Uma das razões desse fato é de que o formato da função logística (5.6) muda muito pouco quando o valor do parâmetro é “alto” (Franses e Dijk, 2000). Conseqüentemente, para se obter uma estimativa mais precisa de  $\gamma$  é preciso um número suficiente de dados na vizinhança do parâmetro  $c$ . Como isso não é tipicamente o que acontece, a estimativa de  $\gamma$  é, geralmente, imprecisa e quase sempre insignificante se julgada pela estatística  $t$  do parâmetro. Esse problema da estimação é discutido em um contexto mais geral em Bates e Watts (1988, p. 87). O ponto mais importante de se considerar é que essa insignificância da estimativa do  $\gamma$  não deve ser interpretada como uma evidência contra a presença de não linearidade do tipo *STAR*.

### 5.1.5. Ferramentas de Avaliação

Após a estimação do modelo faz-se necessário à aplicação de testes para avaliar o modelo proposto. Entretanto, é evidenciado pela literatura que os testes de avaliação tradicionais não funcionam tão bem para modelos não lineares. Em amostras pequenas e moderadas é extremamente recomendável o uso de testes baseados no princípio do multiplicador de Lagrange para avaliar o modelo. Os testes aqui utilizados são os propostos por Medeiros (2000). Estes testes foram utilizados, pois, estudos de simulação, comprovaram que os testes apresentam a potência esperada para pequenas amostras. Os testes são de constância dos parâmetros, independência serial dos resíduos e homocedasticidade.

Antes disso, vale dizer que todos estes testes partem de uma mesma idéia, isto é, assume-se uma igualdade na hipótese nula, porém, essa hipótese não é identificável, pois a teoria assintótica para testes do tipo LM ou razão de verossimilhanças não está disponível e a solução para esse problema é a aproximação da função em questão por uma expansão de Taylor de primeira ordem (até terceira ordem, dependendo do problema) em torno do parâmetro em estudo igual a zero. Daí aproxima-se o logaritmo da função de verossimilhança normal em uma vizinhança de  $H_0$  para a observação  $t$  e sempre se ignorando o resto da expansão. Para continuar derivando o teste do tipo LM, estimadores consistentes para as derivadas parciais do logaritmo da função de verossimilhança devem ser calculados e, a partir deles, criar a estatística do teste.

A seguir, apresenta-se as etapas para a condução dos três testes, lembrando que não é o intuito do trabalho apresentar todos estes procedimentos formalmente, que são, de uma certa forma, complicados. Portanto, para detalhes da formulação dos testes veja Medeiros (2000).

#### 5.1.5.1 Constância dos parâmetros

- (1) Estime o modelo (5.5) assumindo constância dos parâmetros e calcule os resíduos  $\hat{\varepsilon}_t$ . Regrida  $\hat{\varepsilon}_t$  em  $\nabla G(Z_t, x_t; \hat{\theta})$  e calcule

$$SSR_0 = \sum_{t=1}^T \tilde{\varepsilon}_t^2 .$$

- (2) Regrida  $\tilde{\varepsilon}_t$  em  $\nabla G(Z_t, x_t; \hat{\theta})$  e  $\hat{v}_t = [tZ_t', tZ_t' \hat{F}'(\gamma(x_t - c))]$ . Calcule  $SSR_1 = \sum_{t=1}^T \tilde{v}_t^2$
- (3) Calcule a estatística F do teste como  $LM_F^{cp} = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - n - m)}$ , onde T é o número de observações, n é o número de elementos de  $\nabla G(Z_t, x_t; \hat{\theta})$  e  $m = 2p + 2$ .

Sob  $H_0$ ,  $LM_F^{cp}$ , é assintoticamente distribuída como uma  $\chi^2$  com m graus de liberdade e, para pequenas amostras, possui uma distribuição F com m e T-n-m graus de liberdade.

### 5.1.5.2. Independência serial

- (1) Estime o modelo (5.5) sob a hipótese de erros descorrelacionados e calcule  $\hat{\varepsilon}_t$ . Ortogonalize os resíduos regredindo-os em  $\nabla G(Z_t, x_t; \hat{\theta})$  e calcule  $SSR_0 = \sum_{t=1}^T \tilde{\varepsilon}_t^2$ .
- (2) Regrida  $\tilde{\varepsilon}_t$  em  $\hat{v}_t = [\hat{\varepsilon}_{t-1}, \dots, \hat{\varepsilon}_{t-r}]$  e  $\nabla G(Z_t, x_t; \hat{\theta})$ . Calcule  $SSR_1 = \sum_{t=1}^T \tilde{v}_t^2$ .
- (3) Calcule a estatística F do teste como  $LM_F^{is} = \frac{(SSR_0 - SSR_1)/r}{SSR_1/(T - n - r)}$ , onde T é o número de observações, n é a dimensão de  $\nabla G(Z_t, x_t; \hat{\theta})$  e r é o defasamento dos erros que se quer testar.

Sob  $H_0$ ,  $LM_F^{is}$ , é assintoticamente distribuída como uma  $\chi^2$  com r graus de liberdade e, para pequenas amostras, possui uma distribuição F com r e T-n-r graus de liberdade.

### 5.1.5.3. Homocedasticidade

- (1) Estime o modelo (5.5) supondo homocedasticidade e calcule  $\hat{\varepsilon}_t$ . Ortogonalize os resíduos regredindo-os em  $\nabla G(Z_t, x_t; \hat{\theta})$  e

calcule  $SSR_0 = \sum_{t=1}^T \left( \frac{\tilde{\varepsilon}_t^2}{\hat{\sigma}_\varepsilon^2} - 1 \right)^2$ , onde  $\hat{\sigma}_\varepsilon^2$  é a variância incondicional dos erros da regressão.

(2) Regrida  $\left( \frac{\tilde{\varepsilon}_t^2}{\hat{\sigma}_\varepsilon^2} - 1 \right)$  em  $\tilde{x}_t = [1, x_t]'$ . Calcule  $SSR_1 = \sum_{t=1}^T \tilde{v}_t^2$ .

(3) Calcule a estatística F do teste como  $LM_F^h = \frac{(SSR_0 - SSR_1) / d}{SSR_1 / (T - 1 - d)}$ ,

onde T é o número de observações e d é a ordem do defasamento de  $x_t$ .

Sob  $H_0$ ,  $LM_F^h$  é assintoticamente distribuída como uma  $\chi^2$  com d graus de liberdade e, para pequenas amostras, possui uma distribuição F com d e T-1-d graus de liberdade.

### 5.1.6. Previsões utilizando o modelo LSTAR

A não linearidade dos modelos STAR torna as previsões com mais de um passo à frente muito mais complicadas do que nos modelos lineares. Para se fazer previsões um passo à frente, não existe diferença entre o STAR e um modelo AR. Para detalhes de como fazer as previsões consulte em Lundbergh e Teräsvirta (2000). A próxima discussão vai se basear no procedimento geral para se obter previsões com mais de um passo à frente.

Considere o modelo (5.5). Segue que  $E(y_{t+1}|y_t) = f(Z_{t+1}; \Psi)$ , que é a previsão não viciada de  $y_{t+1}$  feita no instante t dada toda informação passada  $Y_t$  até este instante. Nesse caso, a informação relevante está contida em  $Z_{t+1} = (1, y_t, y_{t-1}, \dots, y_{t-(p-1)})'$ . Denota-se essa previsão como  $y_{t+1|t}^f$ . Fazer previsões dois períodos à frente já não é tão simples assim porque obter  $E(y_{t+2}|Y_t)$  é mais complicado. Tem-se que:

$$y_{t+2|t}^f = E(y_{t+2} | Y_t) = E\{g(Z_{t+2}^f; \Psi) + \varepsilon_{t+2} | Y_t\} = E\{g(Z_{t+2}^f; \Psi) | Y_t\} \quad (5.18)$$

onde  $Z_{t+2|t}^f = (1, y_{t+1|t}^f + \varepsilon_{t+1}, y_t, \dots, y_{t-(p-2)})'$ . A expressão exata da equação acima seria

$$y_{t+2|t}^f = E\{g(Z_{t+2|t}^f; \Psi) | Y_t\} = \int_{-\infty}^{\infty} g(Z_{t+2|t}^f; \Psi) d\Phi(z) dz \quad (5.19)$$

onde  $\Phi(z)$  é a função de distribuição acumulada de  $\varepsilon_{t+1}$ . Para se obter a previsão seria necessário integração numérica e integrais múltiplas quando se

quer previsões para longos horizontes, ou seja, a dimensão da integral aumenta com o horizonte. Computacionalmente, é mais plausível se obter previsões recursivas sem integração numérica. Uma maneira simples seria ignorar o termo de erro  $\varepsilon_{t+1}$  e apenas usar o “esqueleto”. Uma outra maneira de se encontrar as previsões é através de simulações utilizando Monte Carlo e/ou Bootstrap.

Nesta dissertação utilizou-se como ferramenta para se obter previsões a simulação de Monte Carlo. É um método simples de simulação que pode ser descrito da seguinte forma: para o modelo  $y_t = F(y_{t-1}; \theta) + \varepsilon_t$ , a previsão k-passos-à-frente é definida como

$$\hat{y}_{t+k|t} = \frac{1}{N} \sum_{i=1}^N \hat{y}_{t+k|t}^{(i)}, \quad (5.20)$$

onde N é o número de repetições e

$$\hat{y}_{t+k|t}^{(i)} = F(\hat{y}_{t+k-1|t}; \theta) + \xi_{t+k|t}^{(i)}, \quad (5.21)$$

Onde  $\xi_{t+k|t}^{(i)}$  é um número aleatório amostrado de uma distribuição normal com a mesma média e desvio padrão do que os resíduos estimados do modelo.