

3 Modelos Não-Lineares

3.1. Modelos Não-Lineares Univariados

3.1.1. Modelo TAR

O modelo Auto-regressivo com limiar (TAR – *Threshold Autoregressive*) foi proposto inicialmente por Tong (1978). Um pouco mais trabalhado, foi mais bem desenvolvido por Tong e Lim (1980) e Tong (1983). Conforme os avanços das pesquisas, este modelo se popularizou com inúmeras aplicações em séries temporais não-lineares. As análises tornaram-se interessantes pelo fato deste modelo atribuir um modelo linear diferente para distintas regiões onde se encontram os valores de uma variável determinada variável de transição. Definiu-se que se a variável de transição for uma defasagem da variável endógena, o modelo é, então, denominado modelo auto-regressivo com limiar auto-excitante SETAR (*Self-Exciting Threshold Autoregressive*).

3.1.1.1. Formulação Matemática

Defina y_t como uma série temporal. Esta série segue um processo TAR caso,

$$y_t = \alpha_0 + \sum_{j=1}^p \alpha_j y_{t-j} + \sum_{i=1}^h \left[\lambda_{0i} + \sum_{j=1}^p \lambda_{ji} y_{t-j} \right] I_i(q_t) + \varepsilon_t, \quad (3.1)$$

onde,

os termos $\alpha_0, \alpha_1, \dots, \alpha_p$ e $\lambda_{0i}, \lambda_{1i}, \dots, \lambda_{pi}$ $i = 1, \dots, h$, são os coeficientes reais do modelo;

$\varepsilon_t \sim NID(0, \sigma^2)$;

$I_i(\cdot)$ é uma função indicadora, definida por

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

O modelo pode ser reescrito na forma vetorial,

$$y_t = \alpha' z_t + \sum_{i=1}^h \lambda_i' z_t I_i(q_t) + \varepsilon_t, \quad (3.2)$$

onde,

os coeficientes do modelo são os vetores $\alpha = [\alpha_0, \alpha_1, \dots, \alpha_p]$ e

$\lambda_i = [\lambda_{0i}, \lambda_{1i}, \dots, \lambda_{pi}]$ e, ainda, $z_t = [1, y_{t-1}, y_{t-2}, \dots, y_{t-p}]$.

Esta representação do modelo permite verificar que, dependendo do valor assumido pela variável q_t , o modelo ativa um dos $h+1$ modelos lineares autorregressivos de ordem p , AR(p). Conforme dito, caso $q_t = y_{t-d}$ o modelo TAR é denominado SETAR, e tem a seguinte representação matricial:

$$y_t = \alpha' z_t + \sum_{i=1}^h \lambda_i' z_t I_i(y_{t-d}) + \varepsilon_t, \quad (3.3)$$

onde, o escalar d é conhecido como tamanho do limiar ou parâmetro de defasamento.

3.1.2. Modelo STAR

Uma generalização do modelo SETAR com dois regimes, incorporando uma transição suave entre eles, foi proposta por Chan e Tong (1986). Este modelo foi denominado modelo STAR (*Smooth Threshold Autoregressive*). Para uma revisão, consulte Teräsvirta (1994).

3.1.2.1. Formulação matemática

Considere y_t uma série temporal univariada. A expressão matemática representada pelo modelo com dois regimes é dada por:

$$y_t = \phi_1' x_t (G(s_t; \gamma, c)) + \phi_2' x_t (1 - G(s_t; \gamma, c)) + \varepsilon_t \quad (3.4)$$

onde,

o vetor $\phi_i = (\phi_{i,0}, \phi_{i,1}, \dots, \phi_{i,p})'$; $i = 1, 2$ são os coeficientes dos modelos lineares ligados aos regimes

o vetor $x_t = (1, y_{t-1}, y_{t-2}, \dots, y_{t-p})'$, é formado por 1 na posição inicial indicando o intercepto do modelo e nas demais posições as defasagens da variável endógena

a função $G(\cdot)$ é uma função limitada no intervalo $[0, 1]$, aqui determinada como a função logística, dada por:

$$G(s_t; \gamma, c) = \frac{1}{1 + e^{-\gamma(s_t - c)}} \quad (3.5)$$

O vetor de parâmetros $\theta = (\gamma, c)$ dessa função contém os parâmetros de suavidade e locação, respectivamente, com a restrição $\gamma > 0$. O primeiro é o responsável pelo grau de suavidade da função de transição, e o segundo representa o limiar entre os dois regimes. Para o mesmo valor de γ , a distância entre o valor de s_t e c determina o grau de pertinência dos regimes do modelo. Na situação em que $s_t = c$, a observação pertence a ambos os regimes com igual grau de pertinência.

Obtemos o modelo TAR se definirmos a função de transição $G(\cdot)$ como uma função indicadora do tipo:

$$G(\cdot) = \begin{cases} 1, & s_t \leq c \\ 0, & s_t > c \end{cases}$$

Neste caso, o limiar entre os dois regimes é abrupto e determinado por c , o parâmetro de limiar ou locação.

Uma das grandes vantagens na utilização dos modelos de transição suave é a possibilidade de especificar a função de transição de forma a evitar este problema da busca por um limiar “rígido” entre os regimes. A escolha mais comum para a função de transição é a função logística. O modelo com esta função de transição é denominado modelo LSTAR (*Logistic Smooth Transition Autoregressive*).

A fim de experimentar a função logística, fixou-se alguns parâmetros e avaliou-se o seu comportamento através do Gráfico 3.1. O parâmetro de suavidade

γ assumiu os valores do conjunto $\{1, 2.5, 5, 50\}$, para representar diferentes níveis de suavidade da função logística, e o parâmetro de locação c assumiu o valor zero.

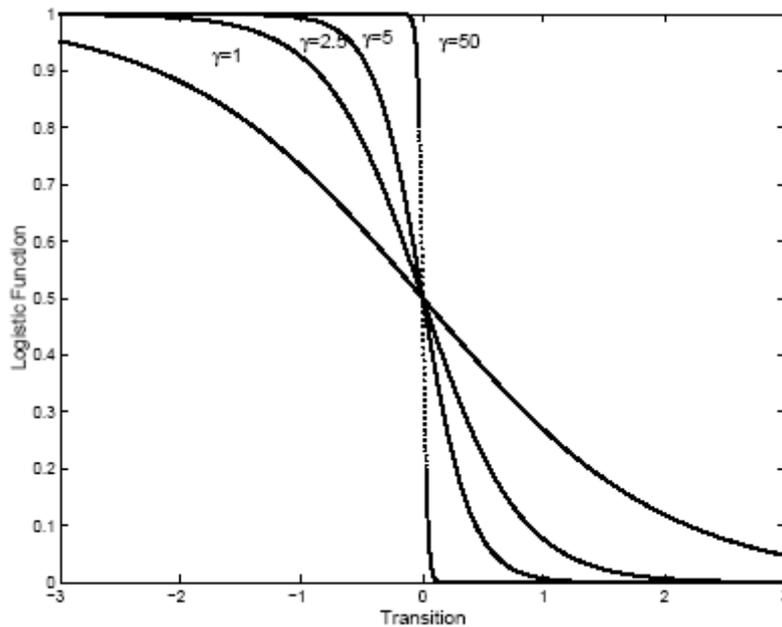


Gráfico 3.1: Função logística com parâmetros fixos

Quando γ tende para zero, a função logística torna-se uma constante igual a 0,5 e o modelo LSTAR se reduz a uma média de dois modelos lineares AR(p). Este comportamento permite concluir que não existe distinção entre os regimes.

Conforme aumentamos o valor do parâmetro de suavidade γ , ou seja, com γ tendendo para infinito, a função logística aproxima-se de uma função do tipo degrau e a transição de um regime para o outro se torna uma transição abrupta. Neste caso, a função logística torna-se uma função indicadora e o modelo é denominado TAR. E ainda, caso a variável de transição seja uma defasagem da variável endógena, $s_t = y_{t-d}$, o modelo é então denominado SETAR (*Self-Exciting Threshold Autoregression*).

Para valores no intervalo (0,1) assumidos pela função logística $G(\cdot)$, o modelo LSTAR com dois regimes é definido como uma média ponderada de dois modelos AR(p), onde os pesos das observações são determinados por esses valores da função de transição, $G(s_t; \gamma, c)$ e $(1 - G(s_t; \gamma, c))$.

O modelo STAR citado anteriormente possui 2 regimes. Porém, este pode ser estendido para um número maior de regimes. Neste caso, denomina-se como

MRSTAR (*Multiple Regime Smooth Transition AutoRegression*). Por exemplo, a representação de um modelo MRSTAR de 4 regimes pode ser escrita da seguinte forma:

$$\begin{aligned}
 y_t = & \left[\phi_1' x_t G_1(s_{1t}; \gamma_1, c_1) + \phi_2' x_t (1 - G_1(s_{1t}; \gamma_1, c_1)) \right] * G_2(s_{2t}; \gamma_2, c_2) \\
 & + \left[\phi_3' x_t G_1(s_{1t}; \gamma_1, c_1) + \phi_4' x_t (1 - G_1(s_{1t}; \gamma_1, c_1)) \right] \\
 & * (1 - G_2(s_{2t}; \gamma_2, c_2)) + \varepsilon_t
 \end{aligned} \tag{3.6}$$

Considerando conhecidas as variáveis de transição s_{1t} e s_{2t} , nota-se que os regimes na equação são ponderados por uma composição de funções logísticas ($G_1(\cdot)$ e $G_2(\cdot)$). Essa composição soma a unidade, por isso podem ser vista como funções de pertinência. O conceito de pertinência é largamente utilizado na teoria da Lógica Fuzzy (Zadeh, 1965). Maiores detalhes referentes aos modelos MRSTAR podem ser obtidos em van Dijk e Franses (1999).

3.1.2.2. Especificação do modelo

Esta seção apresenta uma estratégia de especificação do modelo STAR. Estratégia esta, definida como “específica-para-geral”. Este procedimento inicia com um modelo simples e, de acordo com os resultados dos testes estatísticos aplicados, o modelo tem sua complexidade aumentada.

A primeira preocupação refere-se à seleção das variáveis que irá compor o modelo, tanto as variáveis que formarão o vetor z_t quanto aquelas denominadas variáveis de transição, que formam o vetor x_t .

A modelagem STAR parte de um modelo simples e aumenta sua complexidade de acordo com os resultados dos testes aplicados. van Dijk, Terasvirta e Franses (2002) em seu trabalho, propuseram um processo de construção destes modelos, o qual segue um ciclo de modelagem. Os passos são:

1) Especificação de um modelo AR(p)

Diversos modelos lineares são estimados, começando com um modelo AR(1) e aumentando a ordem p do modelo, com $p = 1, 2, \dots, p_{max}$. Aquele modelo que minimizar os critérios de informação AIC (Akaike, 1974) ou BIC (Schwarz, 1978) deve ser selecionado. Todas as

propriedades dos modelos lineares $AR(p)$ devem ser verificadas, incluindo a que se refere aos resíduos do modelo, definindo-os como aproximadamente um ruído branco.

2) Teste da hipótese de linearidade contra uma alternativa da família STAR

Na construção do modelo STAR, o teste de linearidade tem duas funções. A primeira verifica a adequação do modelo linear para descrever os dados. A hipótese nula do teste é a de linearidade. No caso em que esta não for rejeitada, não é necessário estimar um modelo não-linear para os dados. A segunda determina as variáveis que formam o vetor de transição x_t . Tsay (1989) propôs a aplicação do teste de linearidade para cada uma das defasagens da variável endógena (y_{t-d}) e selecionar como variável de transição aquela que apresentar o menor p-valor do teste.

3) Estimação dos parâmetros do modelo STAR selecionado

Para estimar os parâmetros do modelo, utiliza-se o método de Mínimos Quadrados Ordinários (MQO) para os parâmetros lineares e o método de Mínimos Quadrados Não-Lineares (MQNL) para os parâmetros não-lineares do modelo STAR. Este último, sob a normalidade dos erros, é equivalente ao método de Máxima Verossimilhança (MV).

4) Análise de diagnóstico do modelo

O modelo STAR selecionado e estimado deve apresentar resíduos com boas propriedades. Basicamente, verifica-se a correlação dos resíduos de forma que eles se comportem como ruído branco.

5) Re-especificação do modelo de acordo com os resultados do diagnóstico

No caso de selecionar um modelo STAR que não produza resíduos com boas propriedades, deve-se voltar a etapa 1 e re-especificar um modelo $AR(p)$ e, assim, seguir todas as etapas.

6) Utilização do modelo com fins descritivos ou de previsão

Com todas as etapas acima verificadas, determina-se este modelo como o modelo final e este deve ser utilizado de acordo com os seus fins.

3.2. Modelos Não-Lineares Multivariados

3.2.1. Modelo TVAR

Tsay (1998) estendeu a abordagem dos modelos TAR (*Threshold Autoregressive*) para modelos multivariados, definindo-os como modelos TVAR (*Threshold Vector Autoregressive*). Este modelo teve como motivação uma aplicação no mercado financeiro, onde um ativo foi negociado em dois mercados, simultaneamente.

3.2.1.1. Formulação matemática

Considere que os modelos lineares locais dependam de algumas variáveis exógenas. Seja $y_t = (y_{1t}, \dots, y_{kt})'$ uma série temporal k-dimensional e $x_t = (x_{1t}, \dots, x_{vt})'$ uma série temporal v-dimensional de variáveis exógenas. Seja $-\infty = r_0 < r_1 < \dots < r_{s-1} < r_s = \infty$, então y_t segue um modelo multivariado de limiar com variável de limiar z_t e um lag d dado pela seguinte expressão:

$$y_t = c_j + \sum_{i=1}^p \phi_i^{(j)} y_{t-i} + \sum_{i=1}^q \beta_i^{(j)} x_{t-i} + \epsilon_t^{(j)}, \quad r_{j-1} < z_{t-d} \leq r_j \quad (3.7)$$

onde,

$j = 1, \dots, s$; c_j são vetores de constantes; p e q são inteiros não negativos.

A inovação satisfaz

$$\epsilon_t^{(j)} = \Sigma_j^{1/2} a_t$$

onde,

o primeiro termo é uma matriz positiva definida simétrica;

o segundo é uma seqüência de vetores descorrelatados aleatórios com média zero e matriz de covariância I , a matriz identidade.

O modelo apresentado tem s regimes e é um modelo linear com relação ao espaço de limiar z_{t-d} , mas é não linear no tempo se $s > 1$. Assume-se que a variável de limiar z_t é conhecida, estacionária e apresenta uma distribuição contínua e o lag d , o número de regimes s , e os limiares r_i são desconhecidos.

3.2.1.2. Teste de Linearidade

Primeiramente, Tsay (1998) propôs um teste estatístico para detectar a necessidade de estimar o modelo TVAR ao invés de um modelo linear, isto é, testou $s = 1$ contra $s > 1$. O teste é simples e apresenta um bom desempenho em amostras finitas. Este teste é uma generalização do proposto em Tsay (1989) para o caso univariado, e apresenta uma distribuição assintótica Qui-Quadrado. A generalização também leva em conta a presença de variáveis exógenas e heterocedasticidade condicional.

Em seguida, o autor considerou o teste LM com a hipótese nula que y_t é linear contra a hipótese alternativa que y_t segue um modelo TVAR definido anteriormente. O teste LM usa a variável de limiar para construir uma regressão arranjada. Este regressão baseada no crescimento da ordem da variável de limiar z_{t-d} é:

$$y'_{t(i)+d} = X'_{t(i)+d} \Phi + \epsilon'_{t(i)+d}, \quad i = 1, \dots, n - h. \quad (3.8)$$

onde $t(i)$ é o índice temporal de $z_{(i)}$.

É importante notar que a dinâmica da série y_t não mudou. O que mudou foi a ordem que cada dado entra na regressão, isto é, a ordem das linhas, se víssemos a regressão em um contexto matricial. A idéia do teste é simples: se y_t é linear, então o estimador de mínimos quadrados recursivo da regressão arranjada é consistente, logo os resíduos previstos são aproximadamente um ruído branco. Conseqüentemente, os resíduos previstos são descorrelatados do regressor $X_{t(i)+d}$. Por outro lado, se y_t seguir um modelo de limiar, os resíduos previstos

não serão ruído branco, pois o estimador de mínimos quadrados será viesado. Neste caso, os resíduos previstos serão correlatados com o regressor $\mathbf{X}_{t(i)+d}$.

Após a aplicação do teste o autor descreve um procedimento de construção do modelo incluindo a estimação de s e dos limiares. O método de estimação aplicado é o de mínimos quadrados condicional e a seleção do modelo é realizada com base no critério de informação de Akaike.

3.2.1.3. Estimação

Considerando estimação por mínimos quadrados condicional e assumindo que p , q e s são conhecidos, e que a variável de limiar z_t é dada, escrevemos o modelo para o caso de $s = 2$ da seguinte forma:

$$\mathbf{y}_t = \begin{cases} \mathbf{X}'_t \Phi_1 + \Sigma_1^{1/2} \mathbf{a}_t & \text{if } z_{t-d} \leq r_1 \\ \mathbf{X}'_t \Phi_2 + \Sigma_2^{1/2} \mathbf{a}_t & \text{if } z_{t-d} > r_1 \end{cases}$$

Os parâmetros do modelo são

$$(\Phi_1, \Phi_2, \Sigma_1, \Sigma_2, r_1, d)$$

e a sua estimação pode ser obtida em dois passos. Primeiro, para um dado d e r_1 , o modelo acima nada mais é do que duas regressões lineares multivariadas separadas, cujas estimativas de mínimos quadrados dos seus parâmetros são um resultado conhecido:

$$\hat{\Phi}_i(r_1, d) = \left(\sum_t^{(i)} \mathbf{X}_t \mathbf{X}'_t \right)^{-1} \left(\sum_t^{(i)} \mathbf{X}_t \mathbf{y}'_t \right), \quad \hat{\Sigma}_i(r_1, d) = \frac{\sum_t^{(i)} (\mathbf{y}_t - \mathbf{X}'_t \hat{\Phi}_i) (\mathbf{y}_t - \mathbf{X}'_t \hat{\Phi}_i)'}{n_i - k}$$

Define-se a soma do quadrado dos resíduos como

$$S(r_1, d) = S_1(r_1, d) + S_2(r_1, d)$$

onde $S_i(r_1, d)$ é o traço de

$$(n_i - k) \hat{\Sigma}_i(r_1, d)$$

No passo 2, as estimativas de mínimos quadrados condicional de d e r_1 são obtidas fazendo

$$(\hat{r}_1, \hat{d}) = \operatorname{argmin}_{r_1, d} S(r_1, d)$$

Os estimadores de mínimos quadrados condicionais são estimadores consistentes dos coeficientes, do lag e dos limiares e da matriz de covariâncias.

3.2.1.4. Identificação do modelo

O problema de identificação e especificação de um modelo de limiar multivariado envolve a seleção de muitos parâmetros. Os problemas mais difíceis são: a identificação da variável de limiar e a especificação do número de regimes. A identificação de s pode levar em consideração experiências passadas e informações a priori sobre o conjunto de dados, ou a complexidade computacional pode restringir s a um número pequeno.

Assumindo que z_t e s são dados, o autor usa o critério AIC para selecionar um modelo. Dados p , q , d e s , o critério AIC do modelo de limiar multivariado é dado por:

$$AIC(p, q, d, s) = \sum_{j=1}^s [2 \ln(L_j(p, q, d, s)) + 2k(kp + vq + 1)] \quad (3.9)$$

onde $L_j(p, q, d, s)$ é a função de verossimilhança do regime j avaliado na estimativa de máxima verossimilhança de $c_j, \phi_i^{(j)}, \beta_i^{(j)}$.

3.2.2. Modelo STVAR

O modelo STVAR (*Smooth Transition Vector Autoregressive*) é a versão multivariada do modelo STAR descrito na seção 3.1.2. Este modelo é severamente utilizado para modelar vetores de séries temporais, citando aqui o campo da Macroeconomia.

3.2.2.1. Formulação matemática

Considere $Y_t = (y_{1t}, y_{2t}, \dots, y_{Kt})'$ como um vetor ($K \times 1$) de séries temporais. Uma analogia K -dimensional da expressão matemática representada pelo modelo STAR com dois regimes é dada por:

$$Y_t = \Phi_1' X_t (G(s_t; \gamma, c)) + \Phi_2' X_t (1 - G(s_t; \gamma, c)) + \varepsilon_t \quad (3.10)$$

onde,

os vetores $\Phi_{i,0}$, $i=1,2$ são vetores ($K \times 1$) dos coeficientes interceptos ligados aos regimes;

a matriz $\Phi_i = (\Phi_{i,1}, \dots, \Phi_{i,p})'$ $i=1,2$ tem dimensão ($K \times K$) e é formada pelos coeficientes dos modelos lineares ligados aos regimes;

o vetor $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})'$ é o vetor k -dimensional de ruído branco com média zero e matriz de variância-covariância positiva definida Σ_ε ;

a matriz $X_t = (1, \dots, 1, Y_{t-1}, Y_{t-2}, \dots, Y_{t-p})'$, é formada por 1 na posição inicial indicando o intercepto do modelo e nas demais posições as defasagens das variáveis endógenas;

a função $G(\cdot)$ é a função logística.

Observe que no modelo STVAR os regimes são comuns às K variáveis, no sentido de que uma mesma função de transição determina o regime e a troca de regimes de todas as K equações do modelo.

3.2.2.2. Teste de Linearidade

Para realizar os testes de linearidade enfrentamos o mesmo problema do caso univariado. Isto é, o STVAR contém parâmetros que não são identificáveis sob a hipótese nula. Para solucionar o problema de identificação os autores usam uma aproximação de Taylor adequada para a função de transição. Por exemplo no caso da função logística utiliza-se a aproximação de Taylor de terceira ordem em torno de $\gamma = 0$, resultando em um modelo re-parametrizado:

$$y_t = B_{0,0} + B_0(L)y_{t-1} + B_1(L)y_{t-1}s_t + B_2(L)y_{t-1}s_t^2 + B_3(L)y_{t-1}s_t^3 + e_t \quad (3.11)$$

Desta forma, a hipótese nula original é equivalente à de que $B_i = 0, i = 1,2,3$. A estatística do teste de multiplicador de Lagrange (LM) resultante tem uma distribuição assintótica qui-quadrada com $3pk^2$ graus de liberdade sob a hipótese nula.

3.2.2.3. Estimação

Quando a linearidade é rejeitada e a variável e a função de transição foram selecionadas, os parâmetros do modelo STVAR podem ser estimados através de mínimos quadrados não lineares (MQNL). Sob algumas condições de regularidade, os estimadores são consistentes e com distribuição assintoticamente Normal.

3.2.2.4. Adequação

Como proposto em Eitrheim e Teräsvirta (1996), três testes são realizados com o objetivo de checar se o modelo estimado é adequado. Testa-se se os resíduos apresentam auto-correlação, se os dados ainda apresentam alguma não linearidade e se os parâmetros são constantes. Camacho (2004) descreve esses testes detalhadamente.

3.2.3. Modelo SBTVAR

O modelo TVAR proposto por Tsay (1998) é um modelo linear local com matrizes auto-regressivas diferentes em cada regime, determinados por uma variável de limiar (uma das variáveis endógenas), um *lag* e um limiar. O modelo SBTVAR (*Structural Break Threshold Vector Autoregressive*) proposto por Galvão (2006) também divide a amostra em dois períodos, determinados por um ponto de quebra, o qual permite diferentes dinâmicas antes e depois desta quebra. O que mostra que este modelo caracteriza em mudanças abruptas de um regime para o outro.

Apesar dos modelos não-lineares capturarem algumas características de modelos de quebras estruturais, pode ser que a quebra também implique em mudanças nos parâmetros que determinam a não-linearidade.

3.2.3.1. Formulação matemática

Defina $x_t = (x_{1t}, \dots, x_{mt})'$ como um vetor ($m \times 1$) de m variáveis endógenas e defina $x_{t-1} = (1, x_{t-1}, \dots, x_{t-p})$ como uma matriz ($m \times (mp+1)$), onde p é a ordem auto-regressiva. O modelo SBTVAR pode ser escrito como:

$$x_t = \{[(x_{t-1}\beta_1)I_{1,t-d_1}(r_1) + (x_{t-1}\beta_2)(1 - I_{1,t-d_1}(r_1))]I_t(\tau)\} + \{[(x_{t-1}\beta_3)I_{2,t-d_2}(r_2) + (x_{t-1}\beta_4)(1 - I_{2,t-d_2}(r_2))](1 - I_t(\tau))\} + u_t \quad (3.12)$$

onde,

$I_{i,t-d_i}(r_i)$ é uma função indicadora, a qual depende de uma variável de transição z_{t-d_i} , do limiar r_i e do lag d_i , e $I_t(\tau)$ é uma função indicadora que depende do ponto de quebra.

O SBTVAR tem um TVAR em cada subconjunto determinado pelo ponto de quebra, ou seja, a quebra também afeta os parâmetros da função indicadora que determina os regimes. Se não houver limiar, o VAR com quebra estrutural (SBVAR) é dado por:

$$x_t = (x_{t-1}\beta_1)I_t(\tau) + (x_{t-1}\beta_2)(1 - I_t(\tau)) + u_t \quad (3.13)$$

Por outro lado, se houver limiar, mas não quebra estrutural, temos o VAR com limiar (TVAR), que pode ser escrito como:

$$x_t = (x_{t-1}\beta_1)I_{t-d}(r) + (x_{t-1}\beta_2)(1 - I_{t-d}(r)) + u_t \quad (3.14)$$

3.2.3.2. Estimação

A estimação do SBTVAR pode seguir duas abordagens, a de mínimos quadrados condicionais, usada em Tsay (1998), apresentada na seção 2.3, ou máxima verossimilhança, sugerida em Hansen and Seo (2002).

Usando os resíduos, a matriz de covariância é computada de forma consistente como

$$\hat{\Sigma}(r_1, r_2, \tau) = (\hat{u}'\hat{u})/T$$

O estimador de mínimos quadrados condicionais (MQC) é obtido fazendo

$$\hat{r}_1, \hat{r}_2, \hat{\tau} = \min_{\substack{r_1 \leq r_1 \leq r_M \\ r_1 \leq r_2 \leq r_M \\ \tau_1 \leq \tau \leq \tau_M}} T * \text{trace}(\hat{\Sigma}(r_1, r_2, \tau))$$

Da mesma forma, o estimador de máxima verossimilhança (ML) é obtido fazendo

$$\hat{r}_1, \hat{r}_2, \hat{\tau} = \min_{\substack{r_1 \leq r_1 \leq r_M \\ r_1 \leq r_2 \leq r_M \\ \tau_1 \leq \tau \leq \tau_M}} \log(\det(\hat{\Sigma}(r_1, r_2, \tau)))$$

O estimador de ML é construído assumindo que as matrizes de covariância são as mesmas em cada regime. Essa hipótese pode não ser válida quando aplicada a dados macroeconômicos com variância não constante no tempo, mas o estimador pode ser modificado para este caso.

3.2.3.3. Seleção do modelo

Em Galvão (2006) é apresentado um procedimento de seleção entre modelos de limiar. A questão a ser estudada é qual modelo é mais adequado ao conjunto de dados, um VAR, TVAR, SBVAR ou um SBTVAR.

Mesmo se podendo estimar modelos SBTVAR's, não fica claro a necessidade de ter limiares ou transições que variam no tempo para capturar a estrutura dinâmica dos dados. Testes para limiar em um SBVAR ou para quebra estrutural em um TVAR são complicados devido à descontinuidade das mudanças e da presença de parâmetros mal comportados.

A autora propõe um método de especificação do modelo baseado nos limites assintóticos para os testes LM e de Wald, derivados por Altissimo e Corradi (2002). A regra de decisão para a seleção do modelo usa limites assintóticos e os valores máximos das estatísticas de Wald e LM em uma grade de possíveis valores para os parâmetros mal comportados, como proposto por Altissimo e Corradi (2002). As estatísticas de Wald e LM são calculadas usando a soma do quadrado dos resíduos (*SSR*) sob a hipótese nula e alternativa:

$$W(\theta_2) = n \left(\frac{SSR(\hat{\theta}_1) - SSR(\theta_2)}{SSR(\theta_2)} \right); \quad LM(\theta_2) = n \left(\frac{SSR(\hat{\theta}_1) - SSR(\theta_2)}{SSR(\hat{\theta}_1)} \right)$$

O vetor θ_1 contém parâmetros como limiares e quebras do modelo sob a hipótese nula, e o vetor θ_2 contém os mesmos parâmetros sob a hipótese alternativa.

3.2.4. Modelo TVEC

Lo e Zivot (2001) definiram um modelo de co-integração com limiar multivariado, chamado TVEC (*Threshold Vector Error Correction*), que é um caso especial do TVAR do Tsay (1998).

3.2.4.1. Formulação matemática

De acordo com todas as considerações e suposições feitas no modelo TVAR, Tsay (1998), define-se um modelo de limiar bivariado com 3 regimes pela expressão:

$$p_t = \alpha^{(j)} + \Phi_1^{(j)} p_{t-1} + \Phi_2^{(j)} p_{t-2} + \dots + \Phi_k^{(j)} p_{t-k} + \epsilon_t^{(j)}, \quad \text{if } c^{(j-1)} \leq z_{t-d} \quad (3.15)$$

Pode-se reescrever este modelo como:

$$\Delta p_t = \alpha^{(j)} + \prod^{(j)} p_{t-1} + \sum_{i=1}^{k-1} \Psi_i^{(j)} \Delta p_{t-i} + \epsilon_t^{(j)}, \quad \text{if } c^{(j-1)} \leq z_{(t-d)} \leq c^{(j)} \quad (3.16)$$

onde

$$\prod^{(j)} = \sum_{i=1}^k \Phi_i^{(j)} - I_2 \quad \text{and} \quad \Psi_i^{(j)} = - \sum_{l=i+1}^k \Phi_l^{(j)}$$

Se, em cada regime j , p_t é $I(1)$ e co-integrado com o vetor de co-integração comum $\beta' = (1, -\beta_2)$, então o rank $(\prod^{(j)}) = 1$ e

$$\prod^{(j)} = \gamma^{(j)} \beta' = \begin{pmatrix} \gamma_1^{(j)} \\ \gamma_2^{(j)} \end{pmatrix} (1, -\beta_2).$$

Desta forma, a representação do modelo TVEC é dada por:

$$\Delta p_t = \alpha^{(j)} + \gamma^{(j)} \beta' p_{t-1} + \sum_{i=1}^{k-1} \Psi_i^{(j)} \Delta p_{t-i} + \epsilon_t^{(j)}, \quad \text{if } c^{(j-1)} \leq z_{t-d} \leq c^{(j)} \quad (3.17)$$

3.2.4.2. Teste de co-integração

Lo e Zivot (2001) consideram testes de não co-integração contra co-integração linear e co-integração com limiar, além de testes de linearidade depois de determinado que existe co-integração nos dados.

Balke e Fomby (1997) discutiram alguns problemas associados a testes de co-integração com limiar. Os autores notaram que testar a hipótese nula de não co-integração contra a hipótese alternativa de co-integração com limiar é complicado. Além disso, para construir testes com alto poder para um tipo específico de TVEC, é preciso especificar e estimar a forma do modelo de limiar sob a hipótese alternativa, e isto pode ser difícil, uma vez que existem muitos tipos de modelos de limiar.

Baseados em resultados de simulações de Monte Carlo, Balke e Fomby (1997) sugeriram a seguinte estratégia, que Lo e Zivot (2001) estenderam:

- 1) Testa a hipótese nula de não co-integração contra a alternativa de co-integração linear.
- 2) Se a hipótese de não co-integração for rejeitada, testa a hipótese nula de co-integração linear contra a alternativa de co-integração não-linear (com limiar).
- 3) Se a hipótese de linearidade for rejeitada, é realizada a especificação e estimação do modelo de limiar.

3.2.4.3. Teste de linearidade

Para testar a linearidade, os autores usaram o teste generalizado apresentado em Tsay (1998), que também é válido para processos co-integrados. Para implementar esse teste, foi considerada uma regressão arranjada multivariada para VEC.

3.2.4.4. Especificação do modelo

Após a realização dos testes e de rejeitar a não co-integração e linearidade, é necessário determinar que tipo de modelo de limiar é mais apropriado para o conjunto de dados. Algumas questões a serem respondidas são o número de regimes do modelo, se os valores dos limiares são simétricos, qual modelo é mais apropriado, entre outras.

Duas linhas gerais foram seguidas para determinar a especificação do modelo de limiar apropriada. A primeira, adotada por Tong (1990), Clements e Krolzig (1998), e Tsay (1998), usa um critério de seleção como AIC para determinar a melhor especificação do modelo. A segunda, recentemente revisada por Hansen (1999), usa um procedimento de testes seqüenciais baseados em modelos aninhados. Lo e Zivot (2001) seguiram Hansen (1999) e consideraram testes de hipóteses em ninhos baseadas em estimação irrestrita do modelo TVEC.

3.2.4.5. Estimação

A estimação do modelo é realizada usando mínimos quadrados condicional seqüenciais, como em Hansen (1999).

3.2.5. Modelo STVEC

O modelo STVEC (*Smooth Transition Vector Error Correction*) é a versão não-linear do modelo VEC (*Vector Error Correction*). A julgar pelas aplicações destes modelos multivariados não-lineares que estão atualmente disponíveis, um modelo de particular interesse é aquele em que os componentes do Y_t linear estão ligados por uma relação de equilíbrio de longo prazo, enquanto a adaptação para este equilíbrio é não-linear e pode ser caracterizado como troca de regimes, com os regimes determinados pelo tamanho e/ou o desvio de sinal de equilíbrio. Em modelos lineares de séries temporais, este tipo de comportamento é capturado pelo modelo vetorial de correção de erros, consulte Johansen (1995) para os

tratamentos em profundidade. Recentemente, extensões não-lineares destes conceitos foram consideradas na literatura.

3.2.5.1. Formulação matemática

Concentrando-se na incorporação do mecanismo de transição suave em um VEC para permitir a não-linearidade ou assimetria dos dados, define-se como um modelo vetorial de correção de erros com transição suave (*Smooth Transition Vector Error Correction* – STVEC) dado por:

$$\begin{aligned} \Delta Y_t = & (\Phi_{1,0} + \alpha_1 z_{t-1} + \sum_{j=1}^{p-1} \Phi_{1,j} \Delta Y_{t-j})(1 - G(s_t; \gamma, c)) \\ & + (\Phi_{2,0} + \alpha_2 z_{t-1} + \sum_{j=1}^{p-1} \Phi_{2,j} \Delta Y_{t-j})G(s_t; \gamma, c) + \end{aligned} \quad (3.18)$$

onde,

os vetores α_i , $i=1,2$ são vetores ($K \times 1$) e $z_t = \beta' Y_t$ para algum vetor ($K \times 1$);

β é o termo de correção de erro, isto é, z_t é o desvio da relação de equilíbrio a qual é dada por $\beta' Y_t = 0$;

a matriz $\Phi_i = (\Phi_{i,1}, \dots, \Phi_{i,p})'$ $i=1,2$ tem dimensão ($K \times K$) e é formada pelos coeficientes dos modelos lineares ligados aos regimes;

o vetor $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})'$ é o vetor k -dimensional de ruído branco com média zero e matriz de variância-covariância positiva definida Σ_ε ;

a matriz $X_t = (1, \dots, 1, Y_{t-1}, Y_{t-2}, \dots, Y_{t-p})'$, é formada por 1 na posição inicial indicando o intercepto do modelo e nas demais posições as defasagens das variáveis endógenas.

Afigura-se que as formas de correção de erros não-lineares freqüentemente afetam diferentes ajustes para desvios positivos e negativos, ou para desvios grandes e pequenos do equilíbrio. Efeitos assimétricos de desvios positivos e negativos do equilíbrio podem ser obtidos definindo a função $G(\cdot)$ como a função logística e $s_t = z_{t-1}$. No modelo resultante, a força de reversão de z_t ao seu atrator muda monotonicamente para valores crescentes de z_t . A constante de locação c pode ser reduzida a zero para tornar a mudança simétrica em torno do valor de equilíbrio zero.

Já os efeitos assimétricos de desvios grandes e pequenos do equilíbrio podem ser obtidos definindo a função $G(\cdot)$ como a função exponencial, dada por:

$$G(s_t; \gamma, c) = 1 - \exp\{-\gamma(s_t - c)^2\}, \quad \gamma > 0.$$

com $s_t = z_{t-1}$ e novamente com a constante de locação c reduzida a zero para centrar a força de equilíbrio em zero.

3.2.5.2. Teste de Linearidade

A seleção da variável de transição é feita testando a linearidade do modelo. A hipótese nula é de que o conjunto de dados segue um modelo VEC e a alternativa é de que seguem um STVEC. Para isso, preparam-se uma seqüência de candidatas a variáveis de transição. Para solucionar o problema de identificação dos parâmetros sob a hipótese nula, segue-se a abordagem de Luukkonen, Saikkonen, Teräsvirta (1988) e substitui-se a função de transição por uma aproximação de Taylor adequada. Desta forma, a metodologia aplicada no teste é a mesma da aplicada em modelos STVAR.

O teste proposto por Luukkonen, Saikkonen, Teräsvirta (1988) vem sendo usado em muitos estudos empíricos. Mas esse teste estatístico é baseado em uma aproximação polinomial, e os erros de aproximação podem afetar a inferência estatística. Além disso, os testes não são diretamente relacionados com o modelo de transição suave, logo não pode apontar o que causa a rejeição da linearidade. Com esta motivação, Seo (2004) considerou testes diretos para ajuste não-linear em um modelo STVEC, baseados na especificação exata da transição suave. Hansen e Seo (2002) consideraram os testes para não-linearidade de limiar em um VEC e Seo (2004) estendeu para VEC de transição suave (STVEC). Os testes são baseados na estatística LM, que pode ser calculada sob a hipótese nula.

3.3. Metodologia CART

3.3.1. Introdução

A metodologia *Classification and Regression Tree* (CART), proposta por Breiman, Friedman, Olshen e Stone (1984), é um método de particionamento

recursivo, o qual estrutura os modelos definidos para sub-amostra dos dados, dividindo de forma conveniente o problema em partes. Isto define a estruturação por árvores de decisão, servindo de alternativa aos métodos tradicionais de classificação (variável dependente binária) e regressão (variável dependente contínua).

O modelo CART é não-paramétrico, sendo, portanto, não-probabilístico, pois não assume uma distribuição de probabilidade e não seguem suposições sobre componentes aleatórios e a forma do modelo. A principal vantagem do CART vem da facilidade de interpretação da estrutura de árvore de decisão. As variáveis envolvidas na definição da árvore formam um conjunto de sentenças lógicas do modelo final.

O ciclo da modelagem envolve o crescimento da árvore a partir da raiz (nó inicial), que contém todas as observações do conjunto de dados, até as folhas (nós terminais), cada qual contendo parte das observações.

Primeiramente, realiza-se um teste no nó inicial, o qual só admite resposta do tipo binário $\{0,1\}$, sendo então um teste lógico. Este teste é realizado em cada observação de cada uma das variáveis preditoras e, de acordo com as respostas lógicas obtidas, a raiz dará origem a dois filhos (novos nós), contendo parte das observações originais. Por convenção, se a resposta for 1, aloca-se a observação no nó esquerdo, caso contrário, no nó direito. Para cada nó gerado, este procedimento de teste lógico deve ser repetido até que não seja mais possível dividir a árvore. Desta forma, cada um dos nós que não geraram novos nós são os chamados nós terminais. E cada nó que gerou dois filhos são denominados nós ancestrais, ou nós de divisão, ou ainda nós intermediários.

O modelo final estimado é representado por um gráfico com o formato de uma árvore binária de decisão, com os nós ancestrais (ou nós de divisão) e os nós terminais (ou folhas). Um procedimento importante de numeração dos nós deve ser adotado. A raiz é sempre o nó 0. E cada nó gerado a partir do nó 0 segue uma seqüência numérica crescente da esquerda para a direita. Quando os nós não forem gerados, deve-se saltar os seus números correspondentes e prosseguir a numeração com o nó à direita mais próximo. A Figura 3.1 ilustra um exemplo de árvore com ausência de alguns nós.

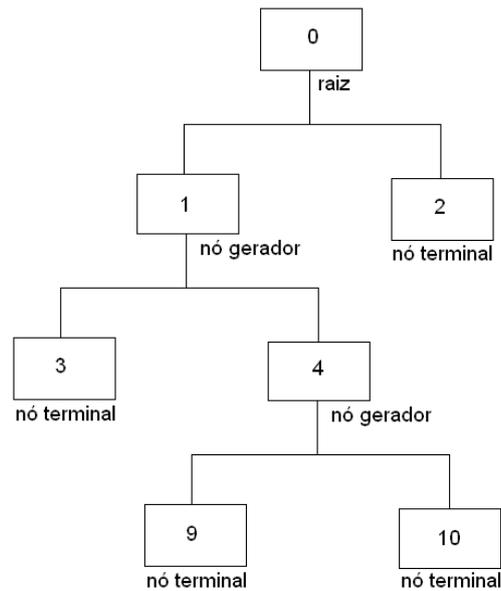


Figura 3.1: Exemplo de árvore com ausência de alguns nós

3.3.2. Formulação matemática

Seja $x_t = x_{1t}, x_{2t}, \dots, x_{pt}' \in \mathbb{X} \subseteq \mathbb{R}^p$ um vetor com p variáveis preditoras de uma determinada resposta univariada e contínua $y_t' \in \mathbb{R}$. Defina $f(\cdot)$ como uma função desconhecida através da expressão:

$$y_t = f(x_t) + \varepsilon_t \quad (3.19)$$

tal que, não há suposições sobre o termo aleatório ε_t .

Define-se, conforme Lewis, Stevens (1991), um modelo estruturado por árvore com K folhas por uma função geral não-linear $H(x_t; \psi)$ de x_t e definida pelo vetor de parâmetros $\psi \in \mathbb{R}^r$, onde r é o número total de parâmetros, através da expressão:

$$f(x_t) \approx H(x_t; \psi) = \sum_{i=1}^K \beta_i I_i(x_t; \theta_i) \quad (3.20)$$

onde, $I(\cdot)$ é uma função indicadora, dada por:

$$I_i(x_t; \theta_i) = \begin{cases} 1, & \text{se } x_t \in k_i(\theta_i) \\ 0, & \text{cc.} \end{cases}$$

e $\psi = (\beta_1, \dots, \beta_K, \theta'_1, \dots, \theta'_K)'$ é o vetor de parâmetros envolvidos na árvore.

Usualmente, $H(\cdot)$ é uma função constante definida por K sub-regiões $k_i(\theta)$, $i = 1, \dots, K$, de algum domínio $K \subset \mathbb{R}^p$.

A Figura 3.2 é um exemplo de um modelo gerado por uma árvore de regressão que explica a relação entre a variável resposta y e um conjunto de duas variáveis preditoras x_1 e x_2 ($q = 2$). Define-se c_j , $j = 0, 1, \dots, N$, como o valor limite da partição k_i que determinará a inclusão da observação na região.

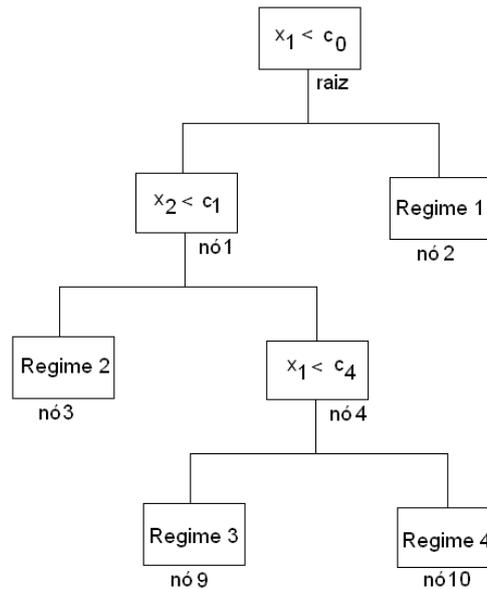


Figura 3.2: Exemplo de um modelo gerado por uma árvore de regressão

3.3.3. Algoritmo de crescimento

A arquitetura da árvore é definida a partir de um ciclo iterativo que escolhe um nó a cada passo para ser subdividido e gerar mais 2 nós. A cada iteração, além do nó a ser dividido, também é especificada uma variável de transição e o limiar desta divisão (c_j). A escolha desta especificação visa minimizar a soma dos erros quadráticos de previsão. Para a raiz da árvore (primeira divisão), a equação a ser minimizada é dada por:

$$SQE(s_0) = \sum_{t=1}^T \{y_t - [\beta_1 I(x_t; s_0, c_0) + \beta_2 (1 - I(x_t; s_0, c_0))]\}^2 \quad (3.21)$$

Após a especificação, estimam-se os parâmetros dos modelos locais para as observações alocadas dentro dos nós gerados pela divisão. Esse ciclo se repete até que não haja mais ganho em efetuar subdivisões na árvore.

Com o modelo final estimado, é possível realizar cortes de algumas folhas, técnica conhecida como podagem (*prunning*), a partir de medidas de custo e complexidade, ou da capacidade preditiva do modelo.