

2

Modelos STAR-Tree

Este capítulo apresenta os principais conceitos envolvidos na definição do modelo STAR-Tree. Sua origem foi marcada pelos processos autoregressivos com limiar, TAR (Threshold AutoRegressive), propostos por Tong [16], e discutido detalhadamente em Tong e Lin [17], e Tong [18]. Estes modelos dividem o espaço Euclidiano unidimensional em k regimes, sendo cada regime governado por um modelo autoregressivo.

Posteriormente, Chan e Tong [3], propõem a substituição da transição abrupta por uma transição suave, originando modelos que ficaram conhecidos como STAR (Smooth Transition AutoRegressive). Granger e Teräsvirta [6], Teräsvirta [15], e Luukkone, Saikkonen e Teräsvirta [12], fazem apresentações detalhadas desta metodologia.

Modelos de transição suave estavam limitados a apresentar apenas 2 regimes, mas em van Dijk e Franses [21], e Medeiros e Veiga [13], esta aplicação foi estendida para a abordagem que suporta a inclusão de múltiplos regimes, denominada MRSTAR (Multiple Regime STAR).

A modelagem STAR-Tree se baseia no conceito do STAR com múltiplos regimes definidos por uma árvore binária. Este capítulo fornece a teoria na qual o STAR-Tree se baseia. Os algoritmos de crescimento da árvore e estimação de parâmetros são detalhados, assim como a metodologia usada para prosseguir com o processo.

2.1

CART

Proposto por Breiman, Friedman, Olshen e Stone [2], o algoritmo CART (Classification And Regression Trees) é usado para prever variáveis dependentes contínuas (regressão) e categóricas (classificação) através do particionamento recursivo do espaço de variáveis de transição, sendo apresentado na forma de árvores de decisão binárias. Sua grande vantagem em

comparação com os demais métodos não lineares e não-paramétricos é a capacidade de interpretação.

Um conceito importante é a profundidade da árvore, que determina a quantidade de ancestrais que um determinado nó possui. A raiz é adotada como profundidade 0. A profundidade fornece uma idéia do grau de complexidade que a árvore estimada apresenta.

A numeração da árvore segue uma convenção na qual o nó raiz recebe o valor 0 e os nós seguintes assumem valores seqüências a partir da esquerda de cada profundidade. A figura (2.1) apresenta um exemplo de enumeração de árvore.

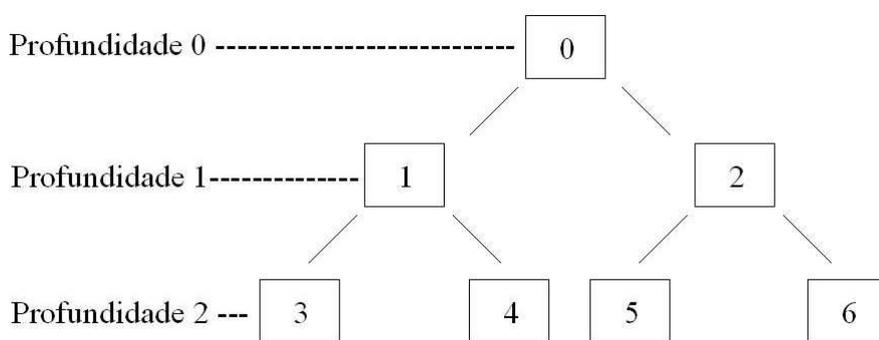


Figura 2.1: Árvore CART

Representação Matemática

A variável de interesse é expressa como a soma de uma função não linear e um termo de erro aleatório.

$$y_t = f(\mathbf{x}_t) + \epsilon_t \tag{2-1}$$

onde $y_t \in \mathbb{R}$ e $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{mt})'$ é um vetor com m variáveis explanatórias. Nenhuma hipótese é feita a respeito da distribuição do termo aleatório ϵ_t .

Seguindo a abordagem usada em Corrêa da Rosa et al [8], o modelo de árvore de regressão com H nós é um modelo de particionamento recursivo que aproxima $f(\mathbf{x}_t)$ por uma função constante por partes definida pelo vetor de parâmetros $\boldsymbol{\psi} \in \mathbb{R}^r$ onde r é o número total de parâmetros.

$$f(\mathbf{x}_t; \boldsymbol{\psi}) \approx \sum_{i=1}^{H+1} \beta_i g_i(\mathbf{x}_t) \tag{2-2}$$

onde

$$g_i(\mathbf{x}_t) = \prod_{k=0}^H I_k(\mathbf{x}_t; c_k)^{j_k(1+j_k)/2} (1 - I_k(\mathbf{x}_t; c_k))^{(1+j_k)(1-j_k)} \quad (2-3)$$

onde $j_{i,k} = -1$ se o caminho até a folha i não inclui o nó k ; $j_{i,k} = 0$ se o caminho até a folha i inclui o ramo direito do nó k ; $j_{i,k} = 1$ se o caminho até a folha i inclui o ramo esquerdo do nó k ; e

$$I(\mathbf{x}_t; c_i) = \begin{cases} 1 & \text{se } \mathbf{x}_t \in A_i \\ 0 & \text{cc} \end{cases} \quad (2-4)$$

onde A_i , $i = 0, 1, \dots, H$ é uma partição induzida por \mathbf{x}_t no espaço \mathbb{R}^m e $\boldsymbol{\psi} = (\beta_1, \dots, \beta_{H+1}, c_0, \dots, c_H)'$, sendo c_i , $i = 0, 1, \dots, H$ o valor limite da partição A_i que determinará a inclusão da observação na região.

Exemplo Ilustrativo

A figura (2.2) apresenta um exemplo de árvore de regressão com duas variáveis explanatórias, x_1 e x_2 , e a resposta univariada prevista \hat{y} . A figura (2.2) apresenta ainda a partição do espaço induzido pela árvore resultante. O modelo final para esta arquitetura tem a forma

$$\begin{aligned} y_t &= [\beta_3 I(x_{2t}; c_1) + \beta_4 (1 - I(x_{2t}; c_1))] I(x_{1t}; c_0) + [(\beta_{11} I(x_{1t}; c_5) + \\ &\quad + \beta_{12} (1 - I(x_{1t}; c_5))) I(x_{2t}; c_2) + \beta_6 (1 - I(x_{2t}; c_2))] (1 - I(x_{1t}; c_0)) = \\ &= [1.5 I(x_{2t}; 6) + 2(1 - I(x_{2t}; 6))] I(x_{1t}; 10) + [(4 I(x_{1t}; 13) + \\ &\quad + 6(1 - I(x_{1t}; 13))) I(x_{2t}; 11) + 7(1 - I(x_{2t}; 11))] (1 - I(x_{1t}; 10)) \end{aligned}$$

A árvore da figura (2.2) pode ser interpretada como um conjunto de regras de predição, como a seguir:

1. Se $(x_1 < 10 \ \& \ x_2 < 6)$, então o valor previsto é $\hat{y} = 1.5$.
2. Se $(x_1 < 10 \ \& \ x_2 > 6)$, então o valor previsto é $\hat{y} = 2$.
3. Se $(x_1 > 10 \ \& \ x_2 > 11)$, então o valor previsto é $\hat{y} = 7$.
4. Se $(x_1 > 10 \ \& \ x_2 < 11 \ \& \ x_1 < 13)$, então o valor previsto é $\hat{y} = 4$.
5. Se $(x_1 > 10 \ \& \ x_2 < 11 \ \& \ x_1 > 13)$, então o valor previsto é $\hat{y} = 6$.

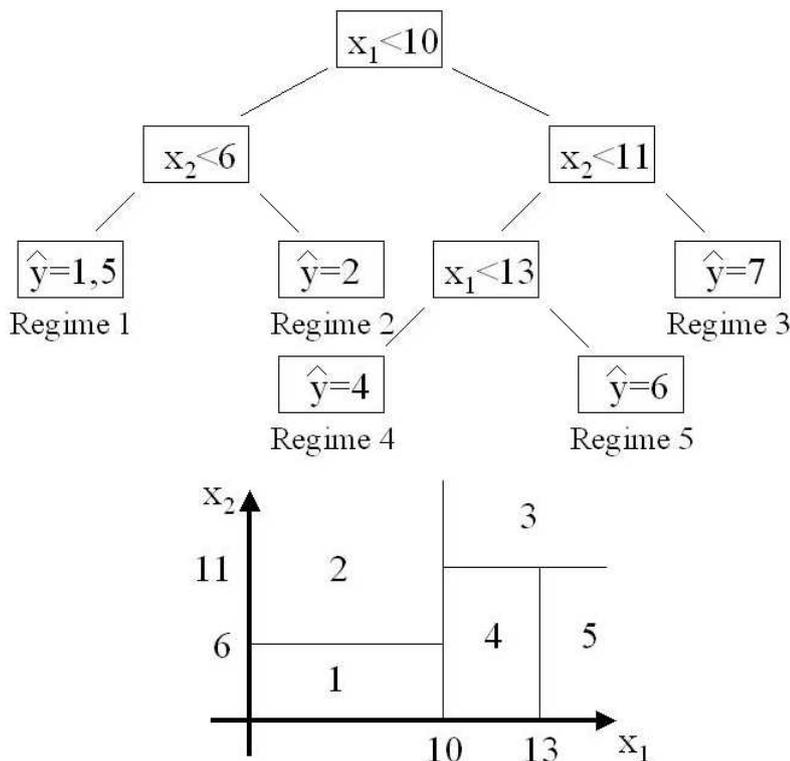


Figura 2.2: Exemplo de Árvore de Regressão

Algoritmo de Crescimento

A determinação da arquitetura da árvore é feita a partir de um algoritmo iterativo em que a cada passo um nó é escolhido para ser subdividido. O nó a ser dividido, a variável e o limiar da divisão são obtidos através da busca da variável explanatória e do limiar da divisão c_i que minimiza a soma dos erros quadráticos de previsão. Para a raiz da árvore, a equação a ser minimizada pode ser vista em (2-5). O critério de parada adotado foi sugerido por Breiman, Friedman, Olshen e Stone [2], e afirma que o nó com 5 ou menos observações é declarado como terminal.

$$SEQ = \sum_{t=1}^T \{y_t - \beta_1 I(x_t; s_0, c_0) - \beta_2 [1 - I(x_t; s_0, c_0)]\}^2 \quad (2-5)$$

Medidas de custo e complexidade e o desempenho preditivo do modelo obtido após o encerramento do algoritmo avaliam se há necessidade de reestimação do mesmo. Uma técnica conhecida como podagem (corte de algumas folhas) é utilizada neste processo. "Suponha uma medida de qualidade de ajuste R_i calculada dentro da i -ésima folha e que a árvore possua N folhas. Em Breiman, Friedman, Olshen e Stone [2], há a sugestão para avaliar o custo complexidade através de uma função", Corrêa da Rosa

et al [8] :

$$R^*(N, \alpha) = \sum_{i=1}^N R_i + |\alpha|N \quad (2-6)$$

onde α é um parâmetro que penaliza a árvore pelo seu tamanho.

2.2 STAR

O modelo STAR pode ser visto como um modelo autoregressivo linear com os coeficientes sendo determinados pela posição de um vetor de variáveis explanatórias no chamado espaço de transição. A representação para dois regimes é dada por

$$y_t = \phi'_1 x_t G(s_t; \gamma, c) + \phi'_2 x_t (1 - G(s_t; \gamma, c)) + \epsilon_t \quad (2-7)$$

onde y_t é a série temporal a ser modelada; $x_t = (1, y_{t-1}, y_{t-2}, \dots, y_{t-p})'$; $\phi_i = (\phi_{i,0}, \phi_{i,1}, \dots, \phi_{i,p})'$, $i = 1, 2$. O erro aleatório ϵ_t tem uma distribuição condicional normal com média $E[\epsilon_t | y_{t-1}, y_{t-2}, \dots, y_{t-p}] = 0$ e variância $\text{var}[\epsilon_t | y_{t-1}, y_{t-2}, \dots, y_{t-p}] = \sigma^2$.

A função de transição $G(s_t; \gamma, c)$, descrita na equação (2-8), é não linear, contínua e limitada entre 0 e 1. O conjunto de variáveis de transição pode ser formado por variáveis endógenas defasadas, exógenas e funções de variáveis endógenas.

$$G(s_t; \gamma, c) = \frac{e^{-\gamma(s_t-c)}}{1 + e^{-\gamma(s_t-c)}}, \quad \gamma > 0 \quad (2-8)$$

onde γ^1 é o parâmetro que determina o grau de suavidade da transição e c é o limiar entre dois regimes.

A figura (2.3) apresenta o gráfico da função logística descrita na equação (2-8) com $\gamma = 1$ e $c = 0$.

Com relação ao parâmetro γ , caso $\gamma \rightarrow \infty$, a função $G(s_t; \gamma, c)$ se aproxima de uma função degrau² e o modelo STAR tem a forma de um modelo

¹O parâmetro de suavidade γ , por ser dependente da escala, é padronizado através da divisão pelo desvio padrão da variável de transição.

$$\gamma^* = \frac{\gamma}{\sigma_s}$$

onde σ_s é o desvio padrão da variável de transição s

²A função degrau é definida pela equação

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

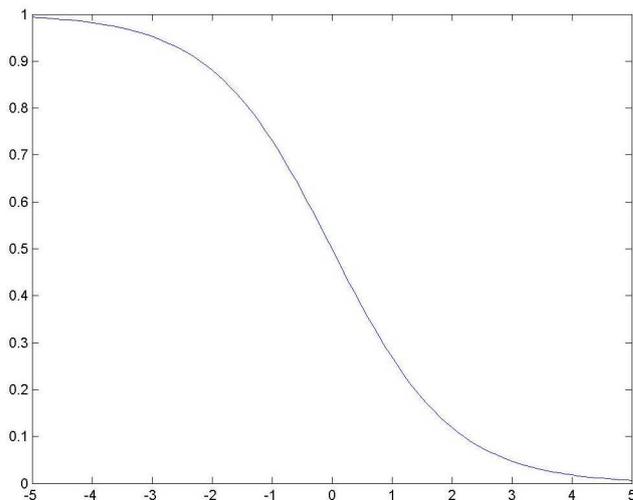


Figura 2.3: Função de Transição

TAR(L) (Threshold Autoregressive) (2-9). Se a variável de transição for uma variável endógena defasada, o modelo é chamado SETAR(L, k_1, \dots, k_L, d) (Self-Exciting Threshold Autoregressive). Caso $\gamma \rightarrow 0$, a função logística é igual a 0.5 e o modelo é linear.

$$y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^{k_1} \phi_i^{(1)} y_{t-i} + \epsilon_t^{(1)} & \text{se } y_{t-d} \in \mathbb{R}_1 \\ \phi_0^{(2)} + \sum_{i=1}^{k_2} \phi_i^{(2)} y_{t-i} + \epsilon_t^{(2)} & \text{se } y_{t-d} \in \mathbb{R}_2 \\ \phi_0^{(L)} + \sum_{i=1}^{k_L} \phi_i^{(L)} y_{t-i} + \epsilon_t^{(L)} & \text{se } y_{t-d} \in \mathbb{R}_L \end{cases} \quad (2-9)$$

O modelo STAR com 2 regimes (2-7) também pode ser visto como uma média ponderada de dois modelos AR, onde os pesos, $G(s_t; \gamma, c)$ e $(1 - G(s_t; \gamma, c))$, são determinados pelo valor assumido pela função de transição. Uma vez fixado γ , a distância entre o valor de s_t e c irá determinar o grau de pertinência da observação aos regimes do modelo. Por exemplo, para a situação na qual $s_t = c$, a observação fará parte de ambos os regimes com o mesmo grau de pertinência.

Uma outra interpretação do modelo leva em conta que os parâmetros lineares variam suavemente, de acordo com os valores das funções de pertinência obtidos. O conceito de função de pertinência é emprestado da Lógica Fuzzy, sendo Zadeh [19], uma boa referência para o assunto.

A abordagem que faz uso de mais de dois regimes recebe o nome de MRSTAR (Multiple Regime Smooth Transition AutoRegression), cuja a descrição detalhada pode ser vista em van Dijk e Franses [21]. Por exemplo,

a representação de um modelo com 4 regimes pode ser vista como

$$y_t = [\phi'_1 x_t G_1(s_{1t}; \gamma_1, c_1) + \phi'_2 x_t (1 - G_1(s_{1t}; \gamma_1, c_1))] G_2(s_{2t}; \gamma_2, c_2) + [\phi'_3 x_t G_1(s_{1t}; \gamma_1, c_1) + \phi'_4 x_t (1 - G_1(s_{1t}; \gamma_1, c_1))] (1 - G_2(s_{2t}; \gamma_2, c_2)) + \epsilon_t \quad (2-10)$$

Ciclo de Modelagem STAR

A modelagem STAR segue a abordagem específico-geral para a construção do modelo, isto é, parte de um modelo simples e aumenta a sua complexidade caso os testes de diagnósticos apontem para esta direção. As etapas do ciclo, proposto em van Dijk, Teräsvirta e Franses [20], estão dispostas abaixo:

1. Escolhe-se a ordem p do modelo AR

A ordem p pode ser obtida através de critérios de informação, como o AIC (Akaike Information Criterion) ou o BIC (Bayesian Information Criterion). É importante que a ordem seja tal que os resíduos do modelo sejam aproximadamente ruído branco.

$$AIC = -2 \frac{l}{T} + 2 \frac{k}{T} \quad BIC = -2 \frac{l}{T} + k \frac{\log(T)}{T} \quad (2-11)$$

onde k é o número de parâmetros, T a quantidade de observações e l é o valor do log da função de verossimilhança.

2. Testa-se a hipótese de linearidade contra não linearidade para cada uma das variáveis explicativas candidatas. Caso a hipótese alternativa seja aceita, escolhe-se a variável s_t com o menor p-valor. O teste utilizado é do tipo Multiplicador de Lagrange.

Este teste apresenta como hipótese nula $H_0 : \phi_1 = \phi_2$ e hipótese alternativa $H_1 : \phi_{1,i} \neq \phi_{2,i}$ com pelo menos um valor de i no conjunto $(0, 1, 2, \dots, p)$. Porém, sob a hipótese nula, os parâmetros γ e c podem assumir diferentes valores sem afetar a função de verossimilhança, impossibilitando a identificação do modelo. Outra forma de testar linearidade seria considerando $H_0 : \gamma = 0$, mas enfrenta-se o mesmo problema observado anteriormente, com ϕ_1 e ϕ_2 podendo assumir qualquer valor desde que a média deles permaneça fixa.

A solução para este problema foi proposta em Luukkonen, Saikkonen e Teräsvirta [12]. Eles sugerem a aproximação da função de transição

por um polinômio de Taylor de terceira ordem ³. Reescrevendo a equação (2-7) como

$$y_t = \phi_1' x_t + (\phi_2 - \phi_1)' x_t G(s_t; \gamma, c) + \epsilon_t \quad (2-12)$$

A aproximação de Taylor em torno de $\gamma = 0$ pode ser reescrita na forma

$$y_t = \beta_0' x_t + \beta_1' x_t s_t + \beta_2' x_t s_t^2 + \beta_3' x_t s_t^3 + e_t \quad (2-13)$$

onde $\beta_i = (\beta_{i,0}, \beta_{i,1}, \dots, \beta_{i,p})'$, $i = 0, 1, 2, 3$, e $e_t = \epsilon_t + (\phi_2 - \phi_1)' x_t R_3(s_t; \gamma, c)$, sendo $R_3(s_t; \gamma, c)$ o termo restante da expansão de Taylor, que, sob H_0 , é nulo e $e_t = \epsilon_t$. Os parâmetros β_i são função de ϕ_1, ϕ_2, γ e c , o que faz com que a hipótese nula possa ser construída como $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$. A estatística de teste é então obtida através das seguintes etapas:

- (a) Estimaco do modelo LSTAR por mnimos quadrados no lineares e obteno da soma dos quadrados dos resduos, $SSR_0 = \sum_{t=1}^T \hat{\epsilon}_t^2$ sob a hiptese nula.
- (b) Estimaco dos resduos e_t e obteno da soma dos quadrados de \hat{e}_t como $SSR_1 = \sum_{t=1}^T \hat{e}_t^2$.
- (c) A estatstica de teste sob a hiptese nula seguir uma distribuo χ^2 com $3(p+1)$ graus de liberdade e ter a forma

$$LM = \frac{T(SSR_0 - SSR_1)}{SSR_0} \quad (2-14)$$

Para pequenas amostras, adota-se uma forma alternativa da estatstica de teste, apresentada na equaco (2-15), e que ter uma distribuo F.

$$LM = \frac{(SSR_0 - SSR_1)/3(p+1)}{SSR_1/(T - 4(p+1))} \quad (2-15)$$

sendo $3(p+1)$ e $(T - 4(p+1))$ os graus de liberdade da distribuo.

3. Estimaco dos parmetros do modelo STAR

³Maiores detalhes sobre a aproximao pelo polinmio de Taylor de terceira ordem pode ser visto no Apndice (A).

A estimação dos parâmetros é feita por mínimos quadrados não lineares (MQNL)

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{t=1}^T (y_t - F(x_t; \theta))^2 \quad (2-16)$$

onde $\theta = (\phi'_1, \phi'_2, \gamma, c)'$ e $F(x_t; \theta) = E[y_t|x_t]$. Vale observar que, sob a normalidade dos erros ϵ_t , MQNL é equivalente à máxima verossimilhança.

A solução por MQNL faz uso de procedimentos de otimização não linear. Neste trabalho, é aplicado o algoritmo de Levenberg-Marquardt⁴ [11] e [14], recomendado para superar o problema de mínimos locais. Em Greene [5], pode ser encontrado mais detalhes sobre algoritmos de otimização não linear.

O problema da estimação por MQNL pode ser simplificado através da fixação dos parâmetros não lineares γ e c e a posterior estimação dos demais elementos por Mínimos Quadrados Ordinários (MQO). Este procedimento recebe o nome de Mínimos Quadrados Concentrados (2-17).

$$\hat{\phi}(\theta) = (Z(\theta)'Z(\theta))^{-1}Z(\theta)'\mathbf{y} \quad (2-17)$$

onde $\theta = (\gamma, c)'$ e $Z(\theta)$ é uma matriz $T \times 2(p+1)$, com a t-ésima linha dada por

$$\begin{bmatrix} G(s_t; \theta) \\ y_{t-1}G(s_t; \theta) \\ \vdots \\ y_{t-p}G(s_t; \theta) \\ 1 - G(s_t; \theta) \\ y_{t-1}(1 - G(s_t; \theta)) \\ \vdots \\ y_{t-p}(1 - G(s_t; \theta)) \end{bmatrix}' \quad (2-18)$$

Então, a estratégia utilizada para estimar os parâmetros segue abaixo:

- (a) Obtém-se os valores iniciais para (γ, c) através da maximização da função de log-verossimilhança concentrada aplicada à uma grade de valores de (γ, c) .

⁴No Apêndice (B) pode ser encontrado maiores detalhes sobre o algoritmo de Levenberg-Marquardt

- (b) Aplicam-se os valores obtidos para (γ, c) e estimam-se os parâmetros lineares $(\phi'_1, \phi'_2)'$ por Mínimos Quadrados Concentrados.
- (c) Os valores de $(\phi'_1, \phi'_2)'$ encontrados são usados no cálculo de (γ, c) por MQNL.
- (d) Estimam-se por MQO os parâmetros lineares fazendo uso dos valores dos parâmetros não lineares obtidos no passo anterior.
- (e) O algoritmo alternará entre as etapas (c) e (d) até que a diferença absoluta entre as somas dos quadrados dos resíduos convirja.

Neste trabalho, a grade de valores iniciais de γ e c foi formada por:

$$\gamma^{(0)} = \{0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.5, 5, 7.5, 10, 25, 50, 75, 100, 250, 500, 750, 1000\}$$

e

$$c^{(0)} = \{y_{(5)}, y_{(10)}, y_{(15)}, y_{(20)}, y_{(25)}, y_{(30)}, y_{(35)}, y_{(40)}, y_{(45)}, y_{(50)}, y_{(55)}, y_{(60)}, y_{(65)}, y_{(70)}, y_{(75)}, y_{(80)}, y_{(85)}, y_{(90)}, y_{(95)}\}$$

onde $y_{(i)}$ é o i -ésimo percentil empírico do processo simulado.

4. Avalia-se o modelo obtido através de testes de diagnósticos

Os testes realizados para validação do modelo são:

- Resíduos descorrelatados
- Não linearidade não remanescente

Se alguma das hipóteses nulas dos testes de diagnósticos for rejeitada, o modelo deve ser reestimado usando uma nova estrutura.

No apêndice (C) é apresentada a descrição teórica de cada teste.

2.3

STAR-Tree

O modelo STAR-Tree é uma combinação de duas metodologias: STAR (Smooth Transition AutoRegression) e CART (Classification And Regression Tree). O resultado é um modelo estruturado em árvore, onde a transição entre os regimes é suave, e com cada folha sendo formada por modelos AutoRegressivos. No presente trabalho, são adotados modelos AR(1) nas folhas.

Ciclo de Modelagem STAR-Tree

A modelagem é baseada em um algoritmo que contém uma série de testes do tipo Multiplicador de Lagrange (ML), aplicados para determinar a inclusão de regimes adicionais e a variável de transição s_t .

Visando minimizar a possibilidade de superestimação da árvore, os testes ML são aplicados com um nível de significância variável que diminui à medida que os testes são aplicados e as profundidades são criadas. O nível de significância é calculado pela equação

$$\alpha(d, n) = \frac{\alpha}{n^d} \quad (2-19)$$

onde n é o n -ésimo teste aplicado na seqüência e d a profundidade

Assim, para o caso do teste na raiz, tem-se ($n = 1$) e ($d = 0$), logo, o nível de significância é α . Com a criação da primeira profundidade, n passa a valer 2 e d , 2. À medida que a árvore cresce, a significância assume os valores $\alpha/3, \alpha/4^2, \alpha/5^2, \alpha/6^3, \alpha/7^3$, etc.

A variável de transição s_t a ser escolhida faz parte do conjunto de variáveis explicativas propostas, formado por defasagens da série analisada e de variáveis exógenas.

O procedimento de crescimento da árvore é feito a partir da raiz e se baseia na inclusão de nós que constituirão as profundidades da árvore. Seguindo a abordagem usada por Corrêa da Rosa et al [8], os passos são descritos a seguir:

1. Criação da raiz da árvore – A criação da raiz da árvore resulta na formulação de um modelo LSTAR(1) apresentado sob a forma

$$y_t = (\alpha_1 + \beta_1 y_{t-1})G(s_{0t}; \gamma_0, c_0) + (\alpha_2 + \beta_2 y_{t-1})(1 - G(s_{0t}; \gamma_0, c_0)) \quad (2-20)$$

Aplica-se o teste ML descrito no item 2 da seção (2.2) para escolher a variável de transição que apresentar o menor p-valor. Se uma das hipóteses nulas for aceita, deve-se escolher uma nova variável de transição e reestimar os parâmetros do modelo.

Com isso, estimam-se os parâmetros $(\alpha_1, \beta_1, \alpha_2, \beta_2, \gamma_0, c_0)'$ do modelo LSTAR(1) por MQNL e verifica-se, através dos testes das hipóteses de (2-21), se o modelo é adequado. Caso nenhuma das variáveis gere uma divisão estatisticamente significativa, a hipótese de linearidade é

aceita e o modelo AR(1) é estimado.

$$\begin{aligned} H_0 : \beta_i = 0, i = 1, 2; \beta_1 - \beta_2 = 0 | \beta_1, \beta_2 \neq 0 \\ H_1 : \beta_1 \text{ ou } \beta_2 \neq 0 \text{ ou } \beta_1 - \beta_2 \neq 0 | \beta_1, \beta_2 \neq 0 \end{aligned} \quad (2-21)$$

2. Inclusão de nós a partir da primeira profundidade - O teste ML continua sendo aplicado com o objetivo de verificar a necessidade de criação de um regime adicional. Aqui são escolhidos a variável de transição s_t e o nó a ser dividido. Inicialmente, o teste é aplicado a todos os pares $\{n_i, s_{jt}\}$, $i = 1, 2$ $j = 1, 2, \dots, k$, sendo k o índice da última variável de transição utilizada, e o nível de significância α' estipulado segue a equação (2-19) com $n = 2$ e $d = 1$. O par que apresentar o menor p-valor é o escolhido, sendo que é necessário que ele seja inferior a α' .

Os parâmetros da função de transição $G_0(\cdot)$ e do nó que não será alterado permanecem os mesmos. Em seguida, estimam-se por MQNL os parâmetros dos dois regimes criados e da função de transição ($G_1(\cdot)$ ou $G_2(\cdot)$) e aplica-se os testes (2-22) para validação do novo modelo.

$$\begin{aligned} H_0 : \beta_{2m+i} = 0, i = 1, 2; \beta_{2m+1} - \beta_{2m+2} = 0 | \beta_{2m+1}, \beta_{2m+2} \neq 0 \\ H_1 : \beta_{2m+1} \text{ ou } \beta_{2m+2} \neq 0 \text{ ou } \beta_{2m+1} - \beta_{2m+2} \neq 0 | \beta_{2m+1}, \beta_{2m+2} \neq 0 \end{aligned} \quad (2-22)$$

onde o valor de m é tal que $m \in \{1, 2\}$.

O teste ML é aplicado uma última vez em todos os pares formados pelo nó que não foi dividido no teste anterior e as variáveis de transição para verificar a necessidade de criação de um regime adicional neste nó. Para este teste, o nível de significância deve ser $\alpha'' = \alpha/3$.

Caso nenhum par gere divisões significantes, o algoritmo é encerrado.

3. Inclusão de nós na z -ésima profundidade da árvore - O procedimento é semelhante à etapa anterior. Considerando que já foram criados N regimes, o teste ML verifica a hipótese de extensão da árvore à próxima profundidade e a possibilidade de criação de regimes adicionais.

Previsão

A flexibilidade da metodologia STAR-tree permite que sejam criados três tipos de previsão diferentes:

- Combinação de Regimes (RC): esta forma de previsão é a aplicação direta do modelo estimado. Utiliza-se a equação resultante da estimação para obter a previsão um passo à frente.
- Máxima Pertinência (MM): observa-se em qual regime a observação apresentou a maior pertinência e aplica-se o modelo AR(1) correspondente.
- Combinação Adaptativa de Regimes (ARC): esta metodologia é similar à RC, porém os parâmetros lineares são re-estimados usando as últimas 252 observações.