Identificação da Ordem de um Processo Auto-Regressivo Estacionário

CNPq/UFRGS
Apoio

Felipe Lazzari Vargas
Bolsista

Sílvia Regina Costa Lopes
Orientadora

Introdução

Os modelos auto-regressivos (AR) são frequentemente usados para descrever séries temporais de caudas pesadas, que apresentam variância infinita por serem simples e descrever seus diferentes comportamentos. Dentro das aplicações de modelos AR com caudas pesadas podemos citar: tempo entre redes (Resnick,1997), temperatura da superfície do mar (Gallagher,2001), log-retorno do mercado de ações (Ling, 2005 e Andrews e Davis, 2013).

Este trabalho tem como objetivo estudar a identificação da ordem de um modelo AR(p) estacionário com inovações α-estáveis. Esta classe de distribuições caracteriza-se pela capacidade de modelar várias características. Uma delas é a presença de caudas pesadas.

A identificação da ordem do processo é feita por diversos critérios de seleção de modelos, com diferentes formulações e em diferentes cenários, tais como, diferentes tamanhos amostrais e diferentes burn-in. Os critérios de seleção estudados são: o Critério de Informação de Akaike(AIC), Critério de Informação Bayesiana(BIC) e o Critério de Hannan-Quin(HQC), além do critério de Determinação Eficiente(EDC).

Método

Existem três principais estágios nas técnicas de modelagem "Box-Jenkins" para séries temporais Gaussianas. São elas a identificação do modelo, estimação dos parâmetros e medidas de diagnósticos.

Para a etapa da identificação do modelo, no caso Gaussiano, existem diferentes técnicas que podem ser aplicadas, como por exemplo, observação dos gráficos da função de autocorrelação e da função de autocorrelação parcial, função de autocorrelação estendida ou métodos baseados numa função penalizadora (entre eles, encontram-se o AIC, BIC, HQC e EDC).

A ideia dos métodos baseados numa função penalizadora é que eles sejam uma medida de qualidade relativa de um modelo estatístico para um conjunto de dados. A funcionalidade dos critérios é baseada no fato de que, ao construir modelos, é possível aumentar a informação adicionando parâmetros. Porém, ao fazer isso, pode-se resultar em super ajuste e, para corrigi-lo, é adicionado um termo de penalidade relativo à quantidade de parâmetros do modelo. Ou seja, ele oferece uma estimativa da informação perdida, quando um determinado modelo é usado para representar o processo que gera os dados.

Estamos interessados em verificar se estes conceitos, dados para situações onde as inovações são Gaussianas, também são válidos, em questão de desempenho e propriedades estatísticas, quando as inovações advêm de distribuições α-estáveis.

Uma variável aleatória X é dita α-estável se sua função característica é dada por

$$\varphi_Z(t) = \mathbb{E}(e^{itZ}) = \begin{cases} \exp\left\{ i\delta t - \sigma_a^\alpha |t|^\alpha \left[1 - i\beta \frac{t}{|t|} \tan(\frac{\pi\alpha}{2}) \right] \right\}, & \alpha \neq 1, \\ \exp\left\{ i\delta t - \sigma_a^\alpha |t| \left[1 + i\beta \frac{2t}{|t|\pi} \ln|t| \right] \right\}, & \alpha = 1. \end{cases}$$

- α é o parâmetro de estabilidade ($0 < \alpha \leq 2$)
- β é o parâmetro de simetria ($|\beta| \leq 1$)
- σ_a é o parâmetro de escala ($\sigma_a > 0$)
- δ é o parâmetro de locação ($\delta \in \mathbb{R}$).

A ideia é escolher a ordem p do processo AR que minimiza a quantidade:

$$\text{critério} = -2\ln(L(.)) + k\, c_n,$$

onde $\ln(L(.))$ é o logaritmo da função de verossimilhança, k é o número de parâmetros do modelo e c_n é uma função penalizadora.

- Critério de Informação de Akaike(1976)

$$\text{AIC}(p) = -2\ln(L(.)) + 2k$$

- Critério de Informação Bayesiana(1978)

$$\text{BIC}(p) = -2\ln(L(.)) + 2\frac{1}{2}k\ln(n)$$

- Critério de Hannan-Quinn(1979)

$$\text{HQC}(p) = -2\ln(L(.)) + 2\,c\,k\,\ln(\ln(n)),$$

onde c é uma constante real maior do que 1.

- Critério de Determinação Eficiente (Dorea e Zhao, 2006)

$$\text{EDC}(p) = -2\ln(L(.)) + 2\,k\,\eta\ln(\ln(n)),$$

onde $\eta > 0$.

Na literatura estes critérios também apresentam uma formulação, que ao invés de depender da função de máxima verossimilhança, dependem da estimativa de máxima verossimilhança da variância. Como modelos AR(p), com inovações advindas de distribuições α-estáveis, não apresentam variância finita (exceto no caso $\alpha = 2$), usamos no lugar a estimativa de máxima verossimilhança do parâmetro de escala de uma α-estável.

A fórmula geral, neste caso, é dada por:

$$\text{critério} = -2\ln\sigma_a^2 + k\,c_n,$$

E queremos a ordem p que minimize está quantidade.

Desta forma, as fórmulas para os critérios de seleção ficam:

- Critério de Informação de Akaike

$$\text{AIC}(p) = n\ln(\hat{\sigma_a^2}) + 2k$$

- Critério de Informação Bayesiana

$$\text{BIC}(p) = n\ln(\hat{\sigma_a^2}) + 2\frac{1}{2}k\ln(n)$$

- Critério de Hannan-Quinn

$$\text{HQC}(p) = n\ln(\hat{\sigma_a^2}) + 2\,c\,k\,\ln(\ln(n)),$$

onde c é uma constante real maior do que 1.

- Critério de Determinação Eficiente

$$\text{EDC}(p) = n\ln(\hat{\sigma_a^2}) + +2\,k\,\eta\ln(\ln(n)),$$

onde $\eta > 0$.

O valor n indica o tamanho da amostra, $\ln(\hat{\sigma_a^2})$ é o logaritmo da estimativa do parâmetro de escala feito através do método de máxima verossimilhança ao quadrado, k é número de parâmetros do modelo e c_n é uma função penalizadora.

Como $-2\ln(L(.)) + 2k$ é comum aos quatro critérios na primeira formulação e $n\ln(\hat{\sigma^2}) + 2k$ é igual nos quatro critérios na segunda formulação as diferenças das funções penalizadoras podem ser vistas na Tabela 1.

Tabela 1: Critério de Seleção de Modelos e sua respectiva Função Penalizadora.

Critério	Função de Penalização
AIC	1
BIC	$\frac{1}{2}\ln(n)$
HQC	$c\ln(\ln(n))$
EDC	$\eta\ln(\ln(n))$

Para dar igual chance a todos os critérios estudados, consideramos que o valor de η está variando no conjunto $\{\frac{1}{2}, 1, 2\}$. De fato, quando $\eta = \frac{1}{2}$, será comparado o critério EDC com o BIC; quando $\eta = 1$, será comparado o EDC com o critério AIC. Quando $\eta = 2$, a formulação do critéro EDC coincide com a do critério HQC, onde o valor da constante c será 2. Além destes valores para η, variou-se η no conjunto $\{1/4, 3, 7\}$ observando o que ocorre ao penalizar mais levemente e mais duramente o critério EDC.

Simulações

Foram geradas séries temporais a partir de um processo AR(1), no software R, com inovações advindas de uma distribuição α-estável.Os parâmetros de escala e simetria do processo de inovação foram considerados iguais à zero enquanto que o índice de estabilidade (α) variou no conjunto de $\{0.5, 0.8, 1, 1.2, 1.5, 1.8, 2\}$ e o parâmetro de escala (σ_a) no conjunto $\{1, 2, 3\}$. Foram utilizados diferentes tamanhos de amostras, burn-in, e replicações. (Ver Tabela 2)

Define-se um processo AR(p) com inovação advinda de uma α-estável como:

$$X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t, \quad \text{para todo} \quad t \in \mathbb{Z},$$

onde $\varepsilon_t \sim \mathcal{S}_\alpha(\sigma_a, 0, 0)$ são variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) e $\boldsymbol{\eta} = (c, \phi_1, \phi_2, \cdots, \phi_p, \alpha, \sigma^2)$ é o vetor de parâmetros.

No momento de realizar o "goodness of fit", ou seja, de ajustar o modelo para as séries temporais geradas, ajustou-se um AR(p), com p variando no conjunto $\{1, 2, 3, 4, 5\}$. Para cada um dos valores de p, foram solicitados os valores dos critérios de seleção estudados para o modelo, com a finalidade de verificar qual critério selecionava mais vezes corretamente o processo AR gerador dos dados.

Tabela 2: Cenários para o Processo de Geração de Dados.

Cenário	n	burn-in	re
(i)	100	20	500
(ii)	300	50	500
(iii)	500	100	500
(iv)	1000	100	500

Tabela 3: Ordem estimada pelos critérios de seleção de modelo com as formulações estudadas para o processo AR(1)- $S_\alpha(\sigma_a, 0, 0)$ quando $n = 300$, *burn-in* $= 50$, *re* $= 500$.

	AIC					BIC					HQC					EDC_7				
$p_{\alpha,\sigma_a,m}$	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
$\hat{p}_{0.5,1,L}$	388	70	24	15	3	391	70	24	13	2	392	69	24	13	2	396	71	22	10	1
$\hat{p}_{0.5,1,\hat{\sigma}_a}$	82	45	33	132	208	85	49	35	130	201	85	49	35	131	200	172	131	69	86	42
$\hat{p}_{1.5,2,L}$	349	79	45	19	8	469	27	4	0	0	484	15	1	0	0	500	0	0	0	0
$\hat{p}_{1.5,2,\hat{\sigma}_a}$	213	95	79	70	43	347	57	38	36	22	383	49	25	30	13	500	0	0	0	0
$\hat{p}_{2,1,L}$	401	46	32	13	8	495	5	0	0	0	498	2	0	0	0	500	0	0	0	0
$\hat{p}_{2,1,\hat{\sigma}_a}$	188	145	73	58	36	440	50	9	1	0	466	30	4	0	0	500	0	0	0	0

Com base nos resultados obtidos até o momento e sob as condições estudadas, nota-se que os critérios de informação que possuem funções de penalização mais duras obtiveram melhores resultados. Em especial, podemos citar o EDC com $\eta = 7$ que apresentou um desempenho muito bom em quase todos os cenários selecionando corretamente o processo AR(1). Os critérios que apresentaram os piores desempenhos foram o AIC e o EDC com $\eta = 0.25$, justamente aqueles que possuem funções penalizadoras mais brandas.

Além disso, observamos que os critérios de seleção de modelo quando utilizados na formulação que envolve o logaritmo da função de máxima verossimilhança obtiveram ótimos resultados selecionando corretamente o modelo AR(1). Enquanto que a formulação que utiliza a estimativa do parâmetro de escala não obteve, em geral, resultados semelhantes (exceto em alguns casos, o desempenho foi similar ao obtido pelo logaritmo da função de máxima verossimilhança).

Referências

Adler, R. J., R. E. Feldman e C. Gallagher (1998). "Analysing stable time series". *A Practical Guide to Heavy Tails.* (Adler, R. J., R. E. Feldman e M. S. Taqqu, eds.) 133-158. Boston: Birkhäuser.

Andrews, B. e R.A. Davis (2013)."Model Identification for infinite variance autoregressive processes. *Journal of Econometrics*, Vol.172, 222-234.

Dorea, C.C.Y. e L.C. Zhao (2006). "Exponential bounds for the probability of wrong determination of the order of a Markov chain by using the EDC criterion". *Journal of Statistical Planning and Inference*, Vol.136, 3689–3697.

Gallagher, C. (2001). "A method for fitting stable autoregressive models using the autocovariation function". *Statistics and Probability Letters*, Vol.53, 381-390.

Knight, K (1989). "Consistency of Akaike's information criterion for infinite variance autoregressive processes". *Annals of Statistics*, Vol.17, 824-840.

Ling, S. (2005). "Self-weighted least absolute deviation estimation for infinite variance autoregressive models". *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, Vol.67, 381-393.

Morettin, P.A. e C.M.C. Toloi (2004). *Análise de Séries Temporais.* ABE - Projeto Fisher. São Paulo: Edgard Blücher.

Nolan, J.P. (2015). *Stable Distributions: Models for Heavy Tailed Data.* Boston: Birkhäuser. Nota: Em progresso, Capítulo 1 online em academic2.american.edu/∼jpnolan.

Resnick, S.I. (1997). "Heavy tail modeling and teletraffic data". *Annals of Statistics*, Vol.25, 1805-1869.