Processos com Memória Longa

1. Introdução

O processo ARMA(p,q) é considerado um processo de "memória curta", uma vez que a f.a.c. ρ_j decresce rapidamente para zero. Na realidade, pode-se demonstrar que, para tal processo,

$$|\rho_j| \leq Cr^{-j}, \quad j = 1, 2, \ldots \tag{1}$$

onde $C > 0$ e $0 < r < 1$. A expressão (9.1) garante que a função de autocorrelação decai para zero exponencialmente.

Um processo de memória longa é um processo estacionário em que a função de autocorrelação decresce hiperbolicamente para zero, isto é,

$$\rho_j \sim Cj^{-\alpha}, \quad j \to \infty, \tag{2}$$

onde $C > 0$ e $0 < \alpha < 1$. Pode-se usar o *coeficiente de Hurst* $H = 1 - \alpha/2$, de modo que $1/2 < H < 1$. Quanto maior H, maior a memória longa do processo. Pode-se provar que o espectro $f(\lambda)$ do processo, cuja função de autocorrelação é como em (9.2), tende a $C_f \lambda^{\alpha-1}$, para $\lambda \to 0$, onde $C_f > 0$ constante. Ou seja, a função densidade espectral de um processo de memória longa diverge para $+\infty$ na freqüência zero.

Estudos empíricos, principalmente em Climatologia e Hidrologia (década de 50) revelaram a presença de memória longa (ML) em dados de séries temporais e espaciais. Estas séries apresentam persistência nas autocorrelações amostrais, isto é, dependência significativa entre observações separadas por um longo intervalo de tempo. Estas autocorrelações apresentam o comportamento dado por (9.2). Outra característica desse tipo de série é que sua função densidade espectral é não limitada na freqüência zero, o que equivale a dizer que sua função de autocorrelação não é absolutamente somável.

Formalmente, temos a

Definição 9.1. *Suponha que X_t tenha autocorrelação ρ_j. Dizemos que X_t possui memória longa se*

$$\lim_{n\to\infty} \sum_{j=-n}^{n} |\rho_j| \tag{3}$$

é não-finita.

O fenômeno de ML foi notado por Hurst (1951, 1957), Mandelbrot e Wallis (1968) e McLeod e Hipel (1978), em conjunção com problemas na área de Hidrologia. Modelos de ML também são de interesse na análise de estudos climáticos, como no estudo da aparente tendência crescente em temperaturas globais devido ao efeito estufa. Veja Seater (1993), por exemplo.

Recentemente (década de 80), os economistas notaram que há evidências que processos de ML descrevem de modo satisfatório dados econômicos e financeiros, tais como taxas de juros e de inflação. Estudos recentes na modelagem da volatilidade de ativos financeiros mostram que tais processos são de grande utilidade. Uma excelente revisão sobre processos de ML em econometria é feita por Baillie (1996).

Exemplo 9.1. A Figura 9.1 mostra a conhecida série de índices de preços de trigo de Beveridge (1925) e suas autocorrelações amostrais, notando o seu lento decaimento.

Exemplo 9.2. Temos, na Figura 9.2 as auto-correlações amostrais das séries de valores absolutos dos retornos diários do Ibovespa, Dow Jones, Banespa e Petrobrás. Estes valores absolutos representam a volatilidade da série. O lento decaimento das auto-correlações mostra claramente a persistência da volatilidade. As figuras mostram, também, as auto-correlações de modelos auto-regressivos AR(p) ajustados às séries. Os valores de p paras as séries do Ibovespa, Dow, Banespa e Petrobrás são, respectivamente, 12, 12, 6 e 17. A série de volatilidades menos persistente é a do Banespa. Vemos que as auto-correlaçẽs dos modelos auto-regressivos são boas estimativas para “lags” baixos. Notamos, ainda, o número excessivo de parâmetros do modelo auto-regressivo necessários para capturar a dependência nas séries.

Uma outra característica de séries com memória longa é que as autocorrelações da série original indicam não-estacionariedade, ao passo que a série diferençada pode parecer ser “super-diferençada”. Ou seja, processos de ML situam-se entre processos I(0) e I(1).

Procurando respeitar as características de uma série de memória longa, citadas anteriormente, foram definidos dois modelos importantes, nos quais a função de densidade espectral é proporcional a λ^{-r}, $1 < r < 2$, para λ próximo de zero e o decaimento da função de autocorrelação é do tipo (9.2). Primeiro foi introduzido o ruído gaussiano fracionário por Mandelbrot e Van Ness (1968). Mais tarde Granger e Joyeux (1980) e Hosking (1981) introduziram o modelo ARIMA fracionário (ou ARFIMA), que é uma generalização do modelo ARIMA.

Há trabalhos recentes incorporando ML a processos GARCH, como nos processos FIGARCH (“fractionally integrated generalized autoregressive conditional heteroskedasticity”), introduzidos por Baillie at al. (1996).

Também, processos de ML associados a modelos de volatilidade estocástica foram considerados por Harvey (1998) e Breidt et al. (1993).

2. Estimação e Testes para Memória Longa

Nesta seção apresentaremos dois procedimentos para testar se uma série temporal apresenta memória longa e estimar o parâmetro de longa dependência. Um é baseado na estatística R/S e outro no periodograma, que é um estimador do espectro de um proceso estacionário.

O modelo proposto para a série X_t é o *processo integrado fracionário*

$$(1 - B)^d (X_t - \mu) = u_t, \qquad (4)$$

onde u_t é um processo estacionário, com espectro $f_u(\lambda)$, e para qualquer número real $d > -1$, define-se o operador de diferença fracionária

$$\begin{aligned}(1-B)^d &= \sum_{k=0}^{\infty}\binom{d}{k}(-B)^k \\ &= 1 - dB + \frac{1}{2!}d(d-1)B^2 - \frac{1}{3!}d(d-1)(d\end{aligned}$$

ou seja,

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}.$$

A relação existente entre d e H é $d = H - 1/2$. Se $0 < d < 1/2$, então X_t é estacionário com memória longa. Se $-1/2 < d < 0$, dizemos que X_t é estacionário com memória curta, ou anti-persistente.

- Estatística R/S

  A estatística R/S foi introduzida por Hurst (1951) com o nome "rescaled range" (ou "range over standard deviation"), com o propósito de testar a existência de memória longa numa série temporal.

Dadas as observações $X_1, \ldots, X_T$, a estatística R/S é dada por

$$Q_T = \frac{1}{S_T}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(X_j-\overline{X}) - \min_{1\le k\le T}\sum_{j=1}^{k}(X_j-\overline{X})\right], \tag{6}$$

onde $\overline{X}$ é a média amostral e S_T^2 é a variância amostral.

Pode-se demonstrar que se X_t são i.i.d. normais, então $Q_T/\sqrt{T}$ converge fracamente para uma v.a. que está no domínio de atração de uma ponte browniana. Lo (1991) fornece os quantis desta variável limite. Ele nota que a estatística definida por (9.6) não é robusta à dependência de curta memória e propõe substituir Q_T por

$$\tilde{Q}_T = \frac{1}{\hat{\sigma}_T(q)}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(X_j-\overline{X}) - \min_{1\le k\le T}\sum_{j=1}^{k}(X_j-\overline{X})\right], \tag{7}$$

onde $\widehat{\sigma}_T(q)$ é a raiz quadrada do estimador da variância de longo prazo de Newey-West, com largura de faixa q, dado por

$$\widehat{\sigma}_T^2(q) = S_T^2(1 + \frac{2}{T}\sum_{j=1}^{q} w_{qj} r_j),$$

sendo $w_{qj} = 1 - j/(q+1)$, $q < T$ e r_j são as auto-correlações amostrais usuais de X_t. Newey and West sugerem escolher $q = [4(T/100)^{2/9}]$.

Se o processo X_t não tiver ML, a estatística S/L converge para sua distribuição limite à taxa $T^{1/2}$, mas se há ML presente, a taxa de convergência é T^H.

Estes fatos sugerem construir gráficos (na escala log-log) de R/S contra o tamanho amostral. Para uma série com MC os pontos devem estar ao longo de uma reta com inclinação 1/2, ao passo que para uma série com ML, a reta deve ter inclinação $H > 1/2$, para grandes amostras.

Para a construção deste gráfico, considerar os valores de R/S contra k_i, para $k_i = fk_{i-1}$, $i = 2, \ldots, s$, k_1 grande inicialmente e f um fator conveniente. Por exemplo, divida a amostra em $[T/k_i]$ blocos.

A função rosTest do S+FinMetrics calcula (9.7) com $q = [4(T/100)^{1/4}]$. Esta função pode ser usada para testar se há ML na série temporal. A função d.ros estima o valor de H segundo esse procedimento gráfico.

Exemplo 9.3. Considere os retornos diários do Ibovespa de 1995 a 2000 e a série de volatilidades, dada pelos valores absolutos dos retornos. Esta śerie é mostrada na Figura 9.3. O Quadro 9.1 mostra o resultado da aplicação da função rosTest. O valor da estatística $\tilde{Q}_T$ é $2,4619$, significativa com o nível $0,01$, o que confirma que a série apresenta memória longa. A Figura 9.4 apresenta o log-log plot de R/S, com a reta ajustada. O valor estimado de $H = 0,71$, do que resulta $d = 0,21$. O gráfico foi feito com $k_1 = 5$ e $f = 2$. A reta pontilhada no gráfico indica MC $(H = 1/2)$.

- Procedimento GPH

Este método para estimação do parâmetro de longa memória foi proposto por Geweke e Porter-Hudak (1983) e se baseia na equação que exibe relação entre os espectros de X_t e de u_t em (9.4). Tal equação foi reescrita para que se assemelhasse a uma equação de regressão linear, onde o coeficiente de inclinação envolve o parâmetro d.

De (9.4) temos que

$$f_x(\lambda) = |1 - e^{-i\lambda}|^{-2d} f_u(\lambda), \qquad (8)$$

em que $f_u(\lambda)$ é o espectro de u_t.

Multiplicando ambos os lados de (9.8) por $f_u(0)$ e aplicando o logaritmo obtemos

$$\ln f_x(\lambda) = \ln f_u(0) - d \ln |1 - e^{-i\lambda}|^2 + \ln\left(\frac{f_u(\lambda)}{f_u(0)}\right). \qquad (9)$$

Substituindo λ por $\lambda_j = 2\pi j/T$ (freqüência de Fourier) e adicionando $\ln(I_x(\lambda_j))$, a ambos os lados de (9.9), obtemos

$$\begin{aligned} \ln I_x(\lambda_j) &= \ln f_u(0) - d \ln\left(4\text{sen}^2\left(\frac{\lambda_j}{2}\right)\right) \\ &\quad + \ln\left(\frac{f_u(\lambda_j)}{f_u(0)}\right) + \ln\left(\frac{I_z(\lambda_j)}{f_x(\lambda_j)}\right), \end{aligned} \qquad (10)$$

em que

$$I_x(\lambda_j) = (2\pi T)^{-1} \left| \sum_{t=1}^{T} X_t \exp(-i\lambda_j t) \right|^2$$

é o *periodograma*.

O termo $\ln(\frac{f_u(\lambda_j)}{f_u(0)})$ pode ser desprezado quando se considerar apenas as freqüências λ_j próximas de zero. Assim, podemos reescrever (9.10) como um modelo de regressão linear

$$Y_j = a - dX_j + \varepsilon_j, \quad j = 1, \ldots, m, \quad (11)$$

em que

$$\begin{aligned} Y_j &= \ln I_x(\lambda_j), \\ X_j &= \ln(4\mathrm{sen}^2(\frac{\lambda_j}{2})), \\ \varepsilon_j &= \ln\left(\frac{I_x(\lambda_j)}{f_x(\lambda_j)}\right), \\ a &= \ln f_u(0) \text{ e } m = cT^\alpha, \quad 0 < \alpha < 1, \end{aligned}$$

c uma constante. A relação linear (9.11) sugere a utilização de um estimador de mínimos quadrados para d, isto é,

$$\hat{d}_{MQ} = -\frac{\sum_{i=1}^{m}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{m}(X_i - \overline{X})^2}. \quad (12)$$

Geweke e Porter-Hudak (1983) demonstram que

$$\hat{d}_{MQ} \xrightarrow{\mathcal{D}} \mathcal{N}\left(d, \frac{\pi^2}{6\sum_{j=1}^{m}(X_j - \overline{X})^2}\right).$$

Sob $H_0 : d = 0$, isto é, o processo não tem memória longa, a estatística

$$t_{d=0} = \hat{d}\left(\frac{\pi^2}{6\sum_{j=1}^{m}(X_j - \overline{X})^2}\right)^{-1/2} \tag{13}$$

tem distribuição normal padrão.

O estimandor d, calculado por meio de (9.12) é chamado estimador GPH (de Geweke e Porter-Hudak).

O programa S+FinMetrics usa a função gphTest para estimar d e testar H_0 dada acima, usando como "default", m=T$^{\alpha}$, com $\alpha = 0,5$.

Observação: o parâmetro d pode também ser estimado por máxima verossimilhança, juntamente com os parâmetros de um processo ARFIMA ajustado à uma série temporal com memória longa. Veja a seção seguinte.

Exemplo 9.4. Consideremos novamente os retornos diários do Ibovespa, do exemplo 9.3. O Quadro 9.2 apresenta o resultado da aplicação da função gphTest. O valor da estatística é $4,3335$, e a hipótese nula de que não há memória longa é rejeitada com o nível 1%. O valor estimado de d é $\hat{d} = 0,5256$, o que sugere ML e não-estacionariedade. Além disso, baseado no erro padrão assintótico fornecido, $0,12128$, obtemos o intervalo de confiança $[0,283; 0,768]$ para d, com coeficiente de confiança de 95%.

3. Modelos ARFIMA

Nesta seção estudaremos uma classe de modelos que são capazes de descrever, simultaneamente, as dinâmicas de memórias curta e longa de processos estacionários.

Definição 9.2. *Dizemos que* $\{X_t\}$ *é um processo auto-regressivo fracionário integrado de média móveis, ou ARFIMA(*p, d, q*) com* $d \in (-\frac{1}{2}, \frac{1}{2})$*, se* $\{X_t\}$ *for estacionário e satisfizer a equação*

$$\phi(B)(1-B)^d X_t = \theta(B) a_t, \tag{14}$$

onde $a_t \sim$ *RB*$(0, \sigma_a^2)$ *e* $\phi(B)$ *e* $\theta(B)$ *são polinômios em* B *de graus* p *e* q*, respectivamente.*

A razão da escolha dessa família de processos, para fins de modelagem das séries com comportamento de memória longa, é que o efeito do parâmetro d em observações distantes decai hiperbolicamente conforme a distância aumenta, enquanto os efeitos dos parâmetros ϕ e θ decaem exponencialmente. Então, d deve ser escolhido com o objetivo de explicar a estrutura de correlação de ordens altas da série, enquanto os parâmetros ϕ e θ explicam a estrutura de correlação de ordens baixas.

A) Estacionariedade e Invertibilidade

Hosking (1981) demonstra que o processo ARFIMA(p, d, q), dado por (9.14) é:

(i) estacionário se $d < \frac{1}{2}$ e todas as raízes de $\phi(B) = 0$ estiverem fora do círculo unitário;

(ii) invertível se $d > -\frac{1}{2}$ e todas as raízes de $\theta(B) = 0$ estiverem fora do círculo unitário.

B) Funções de autocorrelação e densidade espectral

Hosking (1981) também mostra que se X_t, dado por (9.14), for estacionário e invertível e se $f(\lambda)$ for a função densidade espectral de X_t, então

(i) $\lim_{\lambda \to 0} \lambda^{2d} f(\lambda)$ existe e é finito;

(ii) $\lim_{k \to \infty} k^{1-2d} \rho_k$ existe e é finito.

Exemplo 9.5. O caso mais simples é o *ruído branco fracionário*, ou seja, um ARFIMA$(0, d, 0)$, dado por

$$(1 - B)^d X_t = a_t, \quad a_t \sim \text{RB}(0, \sigma_a^2). \quad (15)$$

Se a_t é gaussiano, teremos o *ruído gaussiano fracionário*.

Quando $d < \frac{1}{2}$, X_t é um processo estacionário e tem representação na forma $X_t =$

$\psi(B)a_t$ com os pesos dados por

$$\psi_k = \frac{d(1+d)\cdots(k-1+d)}{k!} = \frac{(k+d-1)!}{k!(d-1)!}.$$

Como $\Gamma(d+k) = d(d+1)\cdots(d+k-1)/\Gamma(d)$, podemos escrever

$$\psi_k = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)},$$

e temos

$$\psi_k \sim \frac{k^{d-1}}{(d-1)!} = c_1 k^{d-1}, k \to \infty$$

sendo c_1 uma constante.

Quando $d > -\frac{1}{2}$ o processo é invertível e tem representação na forma $\pi(B)X_t = a_t$ com os pesos dados por

$$\pi_k = \frac{-d(1-d)\cdots(k-1-d)}{k!} = \frac{(k-d-1)!}{k!(-d-1)!},$$

e como $\Gamma(k-d) = (k-d-1)\cdots(1-d)(-d)\Gamma(-d)$, podemos também escrever

$$\pi_k = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)}$$

e

$$\pi_k \sim \frac{k^{-d-1}}{(-d-1)!} = c_2 k^{-d-1}, k \to \infty,$$

c_2 constante. A seguir, assumiremos $-\frac{1}{2} < d < \frac{1}{2}$.

As funções de densidade espectral, autocorrelação, autocorrelação parcial e a variância são dadas, respectivamente, por

$$f(\lambda) = \begin{cases} \frac{\sigma_a^2}{2\pi}\left(2\text{sen}\left(\frac{\lambda}{2}\right)\right)^{-2d}, & 0 < \lambda \le \pi, \\ \lambda^{-2d}, & \lambda \to 0 \ , \end{cases} \qquad (16)$$

$$\rho_h = \frac{(-d)!(h+d-1)!}{(d-1)!(h-d)!} = \prod_{0<k\le h} \frac{k-1+d}{k-d}, \ h=1,2,\ldots \qquad (17)$$

$$\phi_{hh} = \frac{d}{h-d}, \quad h = 1, 2, \ldots$$

$$\gamma_0 = \frac{(-2d)!}{(-d)!^2}.$$

Em particular, temos que

$$\rho_1 = \frac{d}{1-d}, \tag{18}$$

$$\rho_h \sim \frac{(-d)!h^{2d-1}}{(d-1)!} = c_3 h^{2d-1}, \quad h \to \infty,$$

sendo c_3 constante e

$$f(\lambda) \sim \lambda^{-2d}. \tag{19}$$

A Figura 9.5(a) apresenta $N = 100$ observações simuladas de um modelo ARFIMA$(0, d, 0$ com $d = 0,45$ e a Figura 9.6 (a) apresenta o gráfico das autocorrelações.

Exemplo 9.6. Consideremos, agora, o processo ARFIMA$(1, d, 0)$, dado por

$$(1-B)^d(1-\phi B)X_t = a_t,$$

que é um processo estacionário e invertível se $|d| < \frac{1}{2}$ e $|\phi| < 1$.

Além disso, temos que

(a) os pesos ψ_j e π_j das representações $X_t = \psi(B)a_t$ e $\pi(B)X_t = a_t$ são dados por

$$\psi_j = \frac{(j+d-1)!}{j!(d-1)!}F(1,-j;1-d-j,\phi) \sim \frac{j^{d-1}}{(1-\phi)(d-1)!}$$

e

$$\pi_j = \frac{(j-d-2)!}{(j-1)!(-d-1)!}\{1-\phi-(1+d)/j\} \sim \frac{(1-\phi)}{(-d-1)!}j^{-d-1},$$

respectivamente, em que $F(a,b;c;z) = 1 + \frac{ab}{c}z + \frac{a(a+1)b(b+1)}{c(c+1)}z^2 + \cdots$ é a denominada *função hipergeométrica* e a aproximação vale para $j \to \infty$;

(b) a função densidade espectral é

$$f(\lambda) = \begin{cases} \frac{\sigma_a^2}{2\pi}\frac{\left(2\mathrm{sen}\left(\frac{\lambda}{2}\right)\right)^{-2d}}{1+\phi^2-2\phi\cos\lambda}, & 0 < \lambda \leq \pi, \\ \frac{\lambda^{-2d}}{(1-\phi)^2}, & \lambda \to 0; \end{cases}$$

(c) a expressão para a f.a.c. é bastante complicada mas, em particular, temos que

$$\rho_1 = \frac{(1+\phi^2)F(1,d;1-d;\phi)-1}{\phi[2F(1,d;1-d;\phi)-1]}$$

e

$$\rho_j = \frac{(-d)!(1+\phi)j^{2d-1}}{(d-1)!(1-\phi)^2F(1,1+d;1-d;\phi)}, \ \ j \to \infty.$$

Além disso,

$$\gamma_0 = \frac{(-2d)!F(1,1+d;1-d;\phi)}{(1+\phi)[(-d)!]^2}.$$

A Figura 9.5(b) apresenta $N = 100$ observações simuladas de um processo ARFIMA$(1,d,$ com $\phi = 0,8$ e $d = 0,45$ e a Figura 9.6 (b) apresenta o gráfico das f.a.c.

Exemplo 9.7. Considere, agora, um processo ARFIMA$(0,d,1)$, dado por

$$(1-B)^d X_t = (1-\theta B)a_t,$$

que pode ser visto como uma média móvel de primeira ordem de um ruído branco fracionário; X_t é estacionário e invertível se $|\theta| < 1$ e $|d| < \frac{1}{2}$. Além disso, temos que:

(a) os pesos ψ_j e π_j das representações auto-regressiva e de médias móveis infinitas são dadas por

$$\psi_j = \frac{(j-d-1)!}{j!(-d-1)!}F(1,-j;1+d-j,\theta) \sim \frac{j^{-d-1}}{(1-\theta)(-d-1)!}$$

e

$$\pi_j = \frac{(j+d-2)!}{(j-1)(d-1)!}\left[1-\theta-\frac{(1+d)}{j}\right] \sim \frac{(1-\theta)}{(d-1)!}j^{d-1},$$

respectivamente, em que $F(\cdot)$ é a função hipergeométrica dada no Exemplo 9.6 e a aproximação vale para $j \to \infty$;

(b) $f(\lambda) = \frac{\sigma_a^2}{2\pi}[1+\theta^2-2\theta\cos\lambda]\left[2\text{sen}\left(\frac{\lambda}{2}\right)\right]^{-2d} \sim [(1-\theta)^2\lambda^{-2d}]$ quando $\lambda \to 0$;

(c) a expressão para a f.a.c. é bastante complicada mas, em particular, temos que

$$\rho_1 = \frac{(1+\theta^2)d(2-d) - 2\theta(1-d+d^2)}{(1-d)(2-d)\{1+\theta^2 - 2\theta d/(1-d)\}}$$

e

$$\rho_j = \frac{(-d)!}{(d-1)!} a j^{2d-1}, \quad j \to \infty,$$

em que $a = \frac{(1-\theta^2)}{(1+\theta^2-2\theta d/(1-d))}$.

A Figura 9.5 (c) apresenta $N = 100$ observações de um processo ARFIMA

$(0, d, 1)$ com $d = 0,45$ e $\theta = 0,3$ e a Figura 9.6 (c) apresenta o respectivo gráfico das f.a.c.

Exemplo 9.8. Finalmente, a Figura 9.5 (d) apresenta $N = 100$ observações simuladas de um processo ARFIMA$(1, d, 1)$ com $\phi = 0,8$, $\theta = 0,3$ e a Figura 9.6 (d) apresenta o gráficos das autocorrelações.

Em todos os exemplos citados do processo ARFIMA(p, d, q) podemos notar o comportamento da função de autocorrelação, que tem decaimento hiperbólico.

Para mais detalhes, veja Hosking (1981) e Granger e Joyeux (1980).

4. Estimação de modelos ARFIMA

Nesta seção vamos estudar dois métodos de estimação do modelo (9.14): máxima verossimilhança e estimação semi-paramétrica no domínio da freqüência.

- Estimação de máxima verossimilhança
  A função de verossimilhança de $\mathbf{X} = (X_1, \ldots, X_T)$ proveniente de um processo ARFIMA(p, d, q) pode ser expressa na forma

$$L(\boldsymbol{\eta}, \sigma_a^2) = (2\pi\sigma_a^2)^{-T/2}(r_0 \cdots r_{T-1})^{-1/2}$$

$$\exp\left[-\frac{1}{2\sigma_a^2}\sum_{j=1}^{T}(X_j-\hat{X}_j)^2/r_{j-1}\right],$$

em que $\boldsymbol{\eta} = (d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$, $\hat{X}_j$, $j = 1, \ldots, T$, são as previsões um passo à frente e $r_{j-1} = (\sigma_a^2)^{-1}E(X_j - \hat{X}_j)^2$.

Os estimadores de máxima verossimilhança dos parâmetros são dados por

$$\hat{\sigma}^2_{MV} = T^{-1}S(\hat{\boldsymbol{\eta}}_{MV}), \qquad (20)$$

onde

$$S(\hat{\boldsymbol{\eta}}_{MV}) = \sum_{j=1}^{T} X_j - \hat{X}_j)^2/r_{j-1}$$

e $\hat{\boldsymbol{\eta}}_{MV}$ é o valor de $\boldsymbol{\eta}$ que minimiza

$$\ell(\boldsymbol{\eta}) = \ln(S(\boldsymbol{\eta})) + T^{-1}\sum_{j=1}^{T}\ln r_{j-1}.$$

Entretanto, o cálculo de $\ell(\boldsymbol{\eta})$ é bastante lento. Um procedimento alternativo é considerar uma aproximação para $\ell(\boldsymbol{\eta})$ dada por

$$\ell(\boldsymbol{\eta}) \simeq \ell_*(\boldsymbol{\eta}) = \ln \frac{1}{T} \sum_j \frac{I_T(w_j)}{2\pi f(w_j;\boldsymbol{\eta})}, \quad (21)$$

em que

$$I_T(w_j) = \frac{1}{T}\left|\sum_{t=1}^{T} X_t e^{-itw_j}\right|^2$$

é o periodograma dos dados,

$$f(w_j;\boldsymbol{\eta}) = \frac{\sigma_a^2}{2\pi}\frac{|1-\theta_1 e^{-iw_j} - \cdots - \theta_q e^{-qiw_j}|^2}{|1-\phi_1 e^{-iw_j} - \cdots - \phi_p e^{-piw_j}|^2}.|1-e^{-iw_j}|^{-2}$$

é a função densidade espectral do processo X_t e $\sum_j$ é a soma sobre todas as freqûências de Fourier, $w_j = 2\pi j/T \in (-\pi,\pi]$.

Hannan (1973) e Fox e Taqqu (1986) mostram que:

(i) o estimador $\widehat{\eta}_{MV}$ que minimiza (9.22) é consistente;

(ii) se $d > 0$,

$$\hat{\boldsymbol{\eta}}_{MV} \xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{\eta}, T^{-1}A^{-1}(\boldsymbol{\eta})), \qquad (22)$$

em que $A(\boldsymbol{\eta})$ é uma matriz de ordem $(p+q+1)\times(p+q+1)$ com (j,k)-ésimo elemento dado por

$$A_{jk}(\boldsymbol{\eta}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\partial \ln f(\lambda;\boldsymbol{\eta})}{\partial \eta_j}\frac{\partial \ln f(\lambda;\boldsymbol{\eta})}{\partial \eta_k} \mathrm{d}\lambda;$$

(iii) a variância σ_a^2 é estimada por

$$\hat{\sigma}^2_{MV} = \frac{1}{T}\sum_j \frac{I_T(w_j)}{2\pi f(w_j;\hat{\boldsymbol{\eta}}_{MV})}.$$

O estimador de d obtido desta maneira é chamado *estimador de Whittle* ou *estimador de Fox-Taqqu*. O programa Splus utiliza a função arima.fracdiff para estimar modelos ARFIMA, incluindo modelos da forma $(0, d, 0)$. Contudo, é necessário especificar os valores de p e q da parte ARIMA. Como vimos, é difícil especificar estes valores e uma possibilidade é encontrar valores $p \leq p$max e $q \leq q$max que minimizam o AIC ou BIC.

O programa S+FinMetrics utiliza a função FARIMA, baseada numa extensão de modelos ARFIMA proposta por Beran (1995), que supõe $d > -1/2$. Veja Zivot e Wang (2003) para detalhes. A função d.whittle também pode ser usada para o caso ARFIMA (0,d,0).

Exemplo 9.9.

- Método de Regressão Utilizando o Periodograma

 Aqui utilizaremos o método GPH da seção 9.2. Na expressão (9.8), o espectro $f_u(\lambda)$ de $u_t = (1 - B)^d(X_t - \mu)$ (um processo ARMA(p, q)) é dado por

$$f_u(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}.$$

Estimando d por meio de (9.12) podemos, agora, identificar e estimar os parâmetros do processo livre de componente de longa memória, $u_t = (1-B)^d X_t$. Para isso utilizamos o seguinte procedimento:

(i) Calcule a transformada discreta de Fourier da série original X_t,

$$d_X(\lambda_i) = \sum_{t=1}^{T} X_t e^{-\lambda_i t}, \ \ \lambda_i = \frac{2\pi_i}{T}, \ i = 0, \ldots, T-1.$$

(ii) Calcule

$$d_u(\lambda_i) = (1 - e^{-i\lambda_i})^{\widehat{d}_{MQ}} d_X(\lambda_i), \ \ i = 0, \ldots, T-1.$$

Demonstra-se (veja Brockwell e Davis, 1991) que $d_u(\lambda_i)$ é, aproximadamente, a transformada de Fourier da série filtrada $u_t = (1-B)^{\widehat{d}_{MQ}} X_t$.

(iii) Calcule a transformada inversa de Fourier

$$\tilde{u}_t = \frac{1}{T} \sum_{j=0}^{T-1} e^{iw_j t} d_u(\lambda_j),$$

em que $\tilde{u}_t$ é uma estimativa da série livre da componente de longa memória, u_t.

(iv) Utilize as f.a.c. e f.a.c.p de $\tilde{u}_t$ para identificar os parâmetros p e q.

(v) Estime os parâmetros $\phi_1, \ldots, \phi_p$, $\theta_1, \ldots, \theta_q$, σ_a^2 e d, conjuntamente, utilizando o método de máxima verossimilhança, mencionado na seção anterior.

A vantagem da utilização desse procedimento é que podemos estimar d sem conhecer os valores de p e q. À série $\tilde{u}_t$ podemos aplicar as ferramentas de identificação adequadas aos processos ARMA(p, q) e, finalmente, utilizar o método de máxima verossimilhança para estimar todos os parâmetros do modelo ARFIMA(p, d, q)

O programa S+FinMetrics usa a função d.pgram para implementar um procedimento para o método usando o periodograma.

Exemplo 9.10.

5. Previsão de modelos ARFIMA

Considere o processo ARFIMA(p, d, q) estacionário e invertível,

$$\phi(B)(1-B)^d X_t = \theta_0 + \theta(B)a_t, \quad -0,5 < d < 0,5. \tag{23}$$

Podemos reescrever o processo na forma de choques aleatórios,

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j a_{t-j} \tag{24}$$

e na forma invertida

$$\sum_{j=0}^{\infty} \pi_j Z_{t-j} = \theta_0 + a_t, \tag{25}$$

onde

$$\sum_{j=0}^{\infty} \psi_j B^j = \theta(B)\phi^{-1}(B)(1-B)^{-d}$$

e

$$\sum_{j=0}^{\infty} \pi_j B^j = \phi(B)\theta^{-1}(B)(1-B)^{d}.$$

Assim, podemos fazer previsões de valores futuros do processo X_t, utilizando as equações (9.27) ou (9.28). A variância do erro de previsão, também pode ser calculada de modo usual.

Uma outra forma é usar a da equação de diferenças

$$\varphi(B)X_t = \theta_0 + \theta(B)a_t \qquad (26)$$

em que $\varphi(B) = \phi(B)(1-B)^d = \phi(B)D(B)$ e $D(B) = 1 - d_1B - d_2B^2 - \cdots$ é um polinômio em B, com coeficientes dados por

$$d_j = \frac{-\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} = \prod_{0<k\leq j} \frac{k-1-d}{k}, \;\; j = 0, 1, 2, \ldots \qquad (27)$$

e $\Gamma(\cdot)$ é a função gama, dada por

$$\Gamma(x) = \begin{cases} \int_0^\infty t^{x-1} e^{-t} \mathrm{d}t, & x > 0, \\ \infty, & x = 0, \\ x^{-1}\Gamma(1+x), & x < 0. \end{cases}$$

Utilizando (9.29) e as expressões (9.27) e (9.28), podemos fazer previsões para a série de memória longa X_t.

Note que $D(B)$ é um polinômio de ordem infinita. Na prática, quando temos uma série com T observações, utilizamos somente os L primeiros termos desse polinômio, $L < T$.

Para mais detalhes, ver Brockwell e Davis (1991).

Exemplo 9.11.