## CAPÍTULO 6

## IDENTIFICAÇÃO DE MODELOS ARIMA

### INTRODUÇÃO

Dado que especificamos a classe geral de modelo ARIMA, o próximo passo é a identificação do particular modelo a ser ajustado;

| | | | |
|---|---|---|---|
| $p + q \leq 3$ | MA(1) | AR(1) | ARIMA(1,1) |
| | MA (2) | AR(2) | |
| | | AR(3) | |

A escolha baseia-se primariamente nas autocorrelações e autocorrelações parciais ($facv$) estimadas;

Lembrando que a fac é estimada por: $r_j = \frac{c_j}{c_o}$, $\quad$ j = 0, 1, ..., N – 1.

Onde $c_j$ é a estimativa da $facv$ $\gamma_j$:

$$cj = \frac{1}{N}\sum_{t=1}^{N-j}[(Z_t - \bar{Z})\,(Z_{t+j} - \bar{Z})]\,,\ \ j = 0, 1, \dots, N-1.$$

Onde $\bar{Z} = \frac{\sum_{t=1}^{N-j} Z_t}{N}$ é a média amostral e $r_{-j} = r_j$.

A idéia geral é determinar um IC para $r_j$, para saber se $r_j$ é nula além de um certo *lag*. Uma expressão aproximada para a variância de $r_j$, para um processo estacionário normal, é dado por:

$$\text{Var}\ (r_j) \simeq \frac{1}{N}\sum_{v=-\infty}^{\infty}(\rho_v{}^2 + \rho_{v+j}.\rho_{v-j} - 4\rho_j\rho_v\rho_{v-j} + 2\rho_v{}^2\rho_j{}^2)$$

Para um processo em que as autocorrelações são nulas para $v > q$, todos os termos do lado direito se anulam para j > $q$, exceto o primeiro termo $r_v^2$. Temos, então que:

$$\text{Var}\ (r_j) \simeq \frac{1}{N}\left[1 + 2\ \sum_{v=1}^{q} r_v{}^2\right],\ j > q.$$

Como desconhecemos $\rho_v$ vamos substituir por $r_v$, obtendo a seguinte estimativa:

$$\hat{\sigma}^2\ (r_j) \simeq \frac{1}{N}\left[1 + 2\ \sum_{v=1}^{q} r_v{}^2\right],\ j > q.$$

Para N suficientemente grande e sob a hipótese que $\rho_j = 0$, para j > $q$, a distribuição de $r_j$ é aproximadamente normal, com média igual a zero e variância $\sigma^2\ (r_j)$;

$$r_j \xrightarrow{\text{D(convergência ditrib.)}} \text{N}\left[0, \frac{1}{N}(1 + 2\sum r_v{}^2)\right],\ j > q.$$

Dessa forma, podemos construir um intervalo de confiança aproximado para as autocorrelações;

$$r_j \ \pm t_\gamma . \hat{\sigma}\ (r_j)$$

Onde $t_\gamma$ é o valor da estatística t de Student com N-1 graus de liberdade, t. q. P ($-t_\gamma <$ t $< t_\gamma$) = γ. Na prática, $t_\gamma = 2$.

Podemos considerar $\rho_j$ como sendo significativamente diferente de zero se:

$$|r_j| > 2.\hat{\sigma}\ (r_j),\ j > q.$$

Para a facp, sob a hipótese de que o processo é AR(p):

$$\text{Var}\ (\hat{\Phi}_{jj}) \simeq \frac{1}{N},\ \text{j} > \text{p}.$$

De modo que:

$$\hat{\sigma}\ (\hat{\Phi}_{JJ}) \simeq \frac{1}{\sqrt{N}},\ \text{j} > \text{p}.$$

Além disso, para N grande e sob a hipótese de que o processo é AR(p), $\hat{\Phi}_{jj}$ terá distribuição aproximadamente normal, com média zero e variância Var($\hat{\Phi}_{jj}$);

$$\hat{\Phi}_{jj} \xrightarrow{D} \text{N}\ (0, \frac{1}{N}),\ \text{j} > \text{p}.$$

Podemos considerar $\hat{\Phi}_{jj}$ significativamente diferente de zero se:

$$|\hat{\Phi}_{jj}| > \frac{2}{\sqrt{N}},\ \text{j} > \text{p}.$$

**PROCEDIMENTO DE IDENTIFICAÇÃO**

O objetivo da identificação consiste em determinar os valores "p", "d" e "q" do modelo ARIMA (p,d,q).

Vamos dividi-la em três passos:

1. Verificar se existe necessidade de uma transformação na série original, com o objetivo de estabilizar a sua variância. Para se ter uma idéia do tipo de transformação, pode-se utilizar a série de dados em K subconjuntos com k observações em cada subconjunto. Calcula-se a média $\bar{Z}$ de cada subconjunto, tal que

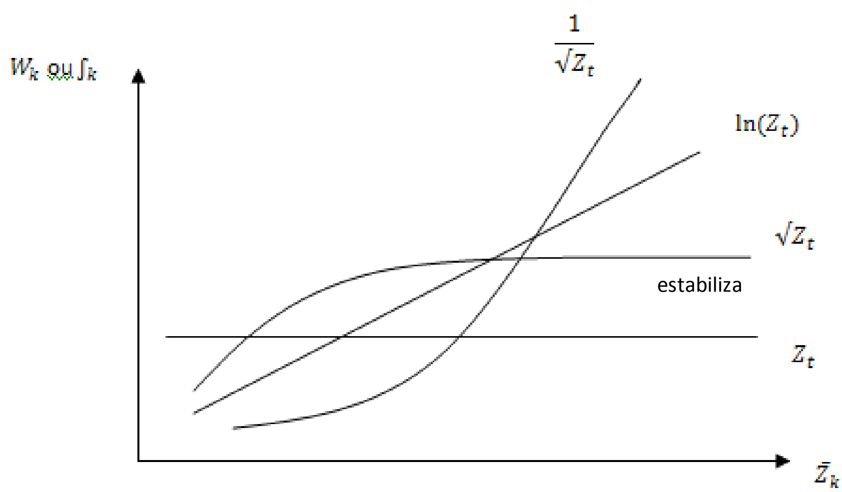
$$\bar{Z}_k = \frac{1}{k}\sum_{i=1}^{k} Z_{ti}$$

i. Calcula-se, também, a amplitude W ou o desvio padrão S de cada subconjunto, tal que:

ii. $W_k = \max(Z_{ti}) - \min(Z_{ti})$, e

iii. $S_k = \sqrt{\frac{\sum_{i=1}^{k}(Z_{ti} - \bar{Z}_k)^2}{k-1}}$

iv. A seguir, construímos um gráfico que traz no eixo das abscissas a variável $\bar{z}_K$ e no eixo das ordenadas, $W_k$ ou $S_k$. O par $(\bar{z}_K, W_k)$ ou $(\bar{z}_K, S_k)$ será um ponto do gráfico. Podemos obter as seguintes variações gráficas:




As transformações logarítmicas e raiz quadrada são as mais comuns.

Nota: a série pode esconder variância não constante;

2. Após a checagem da variabilidade da série, tomar diferenças da série, tantas quantas necessárias, de modo a torná-la estacionária. Ou seja, até um processo $\Delta^d Z_t$ se reduzir a um ARMA (p, q). O número de diferenças necessárias para tornar a série estacionária é alcançado quando a fac amostral de $W_t = \Delta^d Z_t$ decresce rapidamente para zero.

$r_j \longrightarrow 0, \quad j \rightarrow \infty \quad$ (d=1 ≅ 90% casos)

Pode-se testar a existência de raiz unitária no polinômio autorregressivo (teste de raiz unitária ou Dickey-Fuller).

3. Identificar o processo ARMA (p, q) resultante, através das fac e facp estimadas, cujos comportamentos devem ser semelhantes às quantidades teóricas.

| Ordem | (1,d,0) | (0,d,1) |
|---|---|---|
| Comportamento de $\rho_k$ | decai exponencialmente | somente $\rho_1 \neq 0$ |
| Comportamento de $\phi_{kk}$ | somente $\phi_{11} \neq 0$ | decai exponenc. |
| **Ordem** | **(2,d,0)** | **(0,d,2)** |
| Comportamento de $\rho_k$ | mistura de exponenciais ou ondas senóides amortecidas | somente $\rho_1 \neq 0$ e $\rho_2 \neq 0$ |
| Comportamento de $\phi_{kk}$ | somente $\phi_{11} \neq 0$ e $\phi_{22} \neq 0$ | mistura de exponenciais ou ondas senóides amortecidas |
| | **(1,d,1)** | |
| Comportamento de $\rho_k$ | decai exponencialmente após lag 1 | |
| Comportamento de $\phi_{kk}$ | decai exponencialmente após lag 1 | |

Modelos ARMA → fac e facp não são muito úteis. Sugestão: a) Ajustar modelos de baixa ordem e utilizar critérios que permitam escolher o modelo mais adequado. b) utilizar a função de autocorrelação estendida.

A justificativa de (2) é a seguinte:

Para um modelo ARMA estacionário, as fac's são dadas por:

$$\rho_j = A_1 G_1^j + A_2 G_2^j + \ldots + A_p G_p^j, \ j > q - p.$$

Supondo raízes distintas. Como $\Phi(B) = \prod_{i=1}^{p}(1 - G_i B)$ e as raízes de $\Phi(B) = 0$ devem estar fora do círculo unitário, devemos ter $|G_i| < 1$.

Pela equação nota-se que se nenhuma raiz está muito próxima do círculo unitário, as autocorrelações $\rho_j$ decrescerão para zero, conforme j aumenta;

Por outro lado, supondo que uma raiz real $G_i$, esteja próxima de um, ou seja:

$$G_1 = 1 - \varepsilon, \varepsilon > 0 \text{ (pequeno)}$$

Como $j = (1 - \varepsilon)^j \simeq 1 - j\varepsilon$ , ver que para $\varepsilon = 0{,}05$: $0{,}95^2 = 0{,}9025 \simeq 1 - j\varepsilon = 0{,}90$; $0{,}95^3 = 0{,}8574 \simeq 1 - j\varepsilon = 0{,}85$.

$\rho_j \simeq A_1(1 - j\varepsilon)$, o que mostra que a fac decairá lentamente para zero e de forma aproximadamente linear.

Deve-se evitar um excesso de diferenças (**<u>superdiferenciação</u>**).

Notas:

i. Um número excessivo de diferenças resulta em um valor negativo de autocorrelação de ordem 1 da série diferenciada. Neste caso, $\rho_1 = -0{,}5$;
ii. Um número correto de diferenças diminui a variância da série transformada, por outro lado, excesso de diferenças aumentará a variância;

   - O movimento da variância pode ser utilizado para a escolha de <u>d</u>;
iii. Autocorrelações parciais tornam-se negativas;

   - Na prática d = 0, 1 ou 2 e é suficiente verificar as primeiras 20 autocorrelações da série e de suas diferenças.

Convém testar se $\mathrm{E}(W_t) = \mu_W$ é zero, comparando $\bar{W}$ com seu desvio padrão estimado. Lembrando que se d = 0, $\bar{W} = \bar{Z}$ (quadro 6.1);

Quadro 6.1. Variâncias aproximadas, para $\bar{W}$, onde $W_t = \Delta^d Z_t$, n = N – d;

| AR(1) | MA(1) | ARMA(1,1) |
|---|---|---|
| $\frac{C_o(1+r_1)}{n(1-r_1)}$ | $\frac{C_o(1+2r_1)}{n}$ | $\frac{C_o}{n}\left[1+\frac{2{r_1}^2}{r_1-r_2}\right]$ |

| AR(2) | MA(2) |
|---|---|
| $\frac{C_o(1+r_1)(1-2r_1^2+r_2)}{n(1-r_1)(1-r_2)}$ | $\frac{C_o(1+2r_1+2r_2)}{n}$ |

Exercício: Qual o modelo melhor se ajusta as informações abaixo, em que $N = 50$; $\bar{Z} = 0{,}5327$ e $S^2 = 6{,}0579$.

| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $r_j$ | 0,81 | 0,69 | 0,58 | 0,44 | 0,3 | 0,26 | 0,19 | 0,15 |
| $\hat{\Phi}_{jj}$ | 0,81 | 0,11 | -0,03 | -0,12 | -0,13 | 0,17 | -0,01 | 0,02 |

Exemplos R

## FORMAS ALTERNATIVAS DE IDENTIFICAÇÃO

## MODELOS BASEADOS EM UMA FUNÇÃO PENALIZADORA

A idéia é escolher as ordens k e $\ell$ que minimizem a quantidade:

$$P\ (k,\ \ell) = \ln \hat{\sigma}^2_{k,\ell} + (k + \ell)\ \frac{C(N)}{N}$$

Em que $\hat{\sigma}^2_{k,\ell}$ é uma estimativa de variância residual obtida ajustando um modelo ARMA (k, $\ell$) às N observações da série, e C(N) é uma função do tamanho da série;

$(k + \ell)\ \frac{C(N)}{N}$ é o termo penalizador e $\uparrow$ nº parâmetro, $\uparrow$ $(k+ \ell)\ \frac{C(N)}{N}$; $\uparrow$ nº parâmetro, $\downarrow$ $\hat{\sigma}^2_{k,\ell}$ (variância residual);

Assim, minimizar P(k, $\ell$) corresponde a identificar as ordens k e $\ell$ que equilibrem esse comportamento.

Algumas funções penalizadoras particulares:

i. Critério de Informação de Akaike (AIC)

$$AIC(k, d, \ell) = N \ln \hat{\sigma}^2_a + \frac{N}{N-d}\, 2\ (k+\ell+1+ \delta_{do}) + N \ln 2\pi + N$$

$$\delta_{do} \begin{cases} 1, & d = 0 \\ 0, & d \neq 0 \end{cases}$$

$\hat{\sigma}^2_a$ é o estimador de máxima verossimilhança de $\sigma^2_a$.

Com N fixo e com a série apropriadamente diferenciada, obtemos:

$$\text{AIC}(k, \ell) = N\,(\ln \hat{\sigma}^2_a) + 2\,(k+\ell+2)$$

Para o caso de um AR(p), o critério reduz-se a:

$$\text{AIC}\,(k) = N \ln \hat{\sigma}^2_k + 2k,\ k \leq K$$

Onde K é o limite superior para k. Em geral, K = $\ln N$.

ii. Critério de Informação Bayesiano (BIC)

A idéia é minimizar a quantidade BIC:

$$\text{BIC}\,(k, \ell) = \ln \hat{\sigma}^2_{k,\ell} + (k+\ell)\,\frac{\ln N}{N}$$

Onde $\hat{\sigma}^2{}_{k,\ell}$ é a estimativa da variância residual do modelo ARMA(k, $\ell$)

iii. Critério de HANNAN e QUINN (HQC)

Minimizar a quantidade HQC:

$$\text{HQC}\,(k, \ell) = \ln \hat{\sigma}^2_{k,\ell} + 2(k+\ell)c\,\frac{\ln \ln N}{N},\ c > 1.$$

Para o modelo AR(p), o critério HQC reduz-se a

$$\text{HQC}\,(k, \ell) = \ln \hat{\sigma}^2_k + 2ck\frac{\ln \ln N}{N},\ c > 1.$$

**RESUMO**

- fac e facp teóricas:

AR(p) $\rho_k \longrightarrow 0 \quad , \quad k \longrightarrow \infty$

$$\begin{cases} \Phi_{kk} \neq 0 & , \quad k \leq p \\ \Phi_{kk} = 0 & , \quad k > p \end{cases}$$

"corte na facp"
p = 1, 2 (nunca maior)

MA(q) $\Phi_{kk} \longrightarrow 0 \quad , \quad k \longrightarrow \infty$

$$\begin{cases} \rho_k \neq 0 & , \quad k \leq q \\ \rho_k = 0 & , \quad k > q \end{cases}$$

"corte na fac"
p = 1, 2 (nunca menor)

ARIMA (p,q)

$\rho_k \longrightarrow 0$ , k $\longrightarrow \infty$ a partir de lag $p$

Decai exponencialmente a partir de lag $q$

$\Phi_{kk} \longrightarrow 0$ , k $\longrightarrow \infty$

- fac e facp amostrais:
- Determinar as variâncias das fac e facp e construir IC;
- Lembrando que $V\hat{a}r(r_j) \simeq \frac{1}{N}\left[1 + 2\ \sum_{v=1}^{q} {r_v}^2\right]$, j > q.

<u>1º Caso.</u>

- Assim, devemos testar:

$$\begin{cases} H_0 : \text{O modelo é MA(q)} \\ H_1 : \text{O modelo não é MA(q)} \end{cases}$$

Sob $H_0$, então:

$$\hat{\sigma}^2\ (r_j) \simeq \frac{1}{N}\left[1 + 2\ \sum_{v=1}^{q} {r_v}^2\right], \text{ j > q}$$

Ou n = N – d (número de observações com a série diferenciada)

- Desconfiando-se que o processo é MA(2), verifica-se se há outras fac's amostrais $r_j$, j > 2, fora do IC.
- Caso particular: MA(0)$\Rightarrow Z_t = a_t$. Neste caso, q = 0 e $\hat{\sigma}^2\ (r_j) = \frac{1}{N}$

2º Caso.

- Testar:

$$\begin{cases} H_0 : \text{O modelo é AR(p)} \\ H_1 : \text{O modelo não é AR(p)} \end{cases}$$

Lembrando que $\hat{\sigma}^2\ (\Phi_{jj}) \simeq \frac{1}{N}$, j > p sob $H_0$.

- O IC fica $\left[-1{,}96\frac{1}{\sqrt{N}}, 1{,}96\ \frac{1}{\sqrt{N}}\right]$