EAE1223: Econometria III

Estimação do ARMA(p,q) condicional

Estimação condicional do MA(1)

Considere, de início o MA(1) sem intercepto:

$$Y_t = \epsilon_t + \theta_1 \epsilon_{t-1}$$

Se conhecêssemos ϵ_0, poderíamos estimar o parâmetro θ_1 como:

$$\check{\theta}_1 \in \text{argmin}_{c\in\mathbb{R}} \frac{1}{T}\sum_{t=1}^{T}(Y_t - c\check{\epsilon}_{t-1}(c))^2 \,, \tag{1}$$

onde $\check{\epsilon}_0(c) = \epsilon_0$, e, para $t > 0$, $\check{\epsilon}_t(c) = Y_t - c\check{\epsilon}_{t-1}(c)$.

O estimador $\check{\theta}_1$ não é factível, visto que desconhecemos ϵ_0. Nós notamos, no entanto, que, dado um chute $\tilde{\epsilon}_0 = 0$ inicial para o ruído branco, somos capazes de construir o estimador alternativo

$$\hat{\theta}_1 \in \text{argmin}_{c\in\mathbb{R}} \frac{1}{T}\sum_{t=1}^{T}(Y_t - c\tilde{\epsilon}_{t-1}(c))^2 \,, \tag{2}$$

onde $\tilde{\epsilon}_0(c) = \tilde{\epsilon}_0$, e, recursivamente:

$$\begin{aligned}
\tilde{\epsilon}_1(c) &= Y_1 - c\tilde{\epsilon}_0(c), \\
\tilde{\epsilon}_2(c) &= Y_2 - c\tilde{\epsilon}_1(c) \\
&\vdots \\
\tilde{\epsilon}_T(c) &= Y_t - c\tilde{\epsilon}_{T-1}(c)
\end{aligned}$$

Note que o estimador factível efetivo $\hat{\theta}_1$ na prática não usa a primeira observação, visto que $(Y_1 - c\tilde{\epsilon}_0(c))^2 = (Y_1)^2$.

No Apêndice A, nós mostramos que o chute inicial $\tilde{\epsilon}_0$ torna-se irrelevante, quando T é grande, na região de invertibilidade do MA(1), $|c| < 1$.

Obeservação: note que a função objetivo $\frac{1}{T}\sum_{t=1}^{T}(Y_t - \tilde{\epsilon}_t(c) - c\tilde{\epsilon}_{t-1}(c))^2$ é trivialmente igual a zero, para qualquer valor de c. Logo, ela não pode ser usada na minimização. Devemos remover o ruído branco contemporâneo na estimação.

Estimação condicional do ARMA(p,q)

A estimação do ARMA(p,q) procede analogamente. Minimizamos:

$$\min_{a,b_1,\ldots,b_p,c_1,\ldots,c_q} \sum_{t=p+1}^{T} (Y_t - a - b_1 y_{t-1} \ldots - b_p y_{t-p} - c_1 \tilde{\epsilon}_{t-1}(a;\boldsymbol{b};\boldsymbol{c}) - \ldots - c_q \tilde{\epsilon}_{t-q}(a;\boldsymbol{b};\boldsymbol{c}))^2 ,$$

onde $\tilde{\epsilon}_{p-q+1} = \tilde{\epsilon}_{p-q+2} = \ldots = \tilde{\epsilon}_p = 0$ e, recursivamente, para $t > p$:

$$\tilde{\epsilon}_t(a;\boldsymbol{b};\boldsymbol{c}) = y_t - a - b_1 y_{t-1} \ldots - b_p y_{t-p} - c_1 \tilde{\epsilon}_{t-1}(a;\boldsymbol{b};\boldsymbol{c}) - \ldots - c_q \tilde{\epsilon}_{t-q}(a;\boldsymbol{b};\boldsymbol{c}) .$$

Note que perdemos as p primeiras observações, visto que não observamos os Y_t anteriores a $t = 1$. Além disso, na observação $t = p + 1$, não há informação referente aos coeficientes $c_1, \ldots, c_q$, visto que os $\tilde{\epsilon}$ correspondentes são todos zero. Similarmente, em $t = p + 2$, não há informação para os coeficientes $c_2, \ldots, c_q$. Somente em $t = p + q + 1$ há informação para todos os coeficientes. Assim, perdemos informação nas $p+q$ primeiras observações. Em virtude desse fato, algumas implementações da estimação condicional consideram somente os períodos em que há informação para todos os coeficientes, isto é:

$$\min_{a,b_1,\ldots,b_p,c_1,\ldots,c_q} \sum_{t=p+q+1}^{T} (Y_t - a - b_1 y_{t-1} \ldots - b_p y_{t-p} - c_1 \tilde{\epsilon}_{t-1}(a;\boldsymbol{b};\boldsymbol{c}) - \ldots - c_q \tilde{\epsilon}_{t-q}(a;\boldsymbol{b};\boldsymbol{c}))^2 ,$$

Com T grande, ambas as abordagens (de $p + 1$ a T e de $p + q + 1$ a T) produzem resultados similares.

A Mostrando que chute inicial é irrelevante para T grande, na região em que $|c| < 1$

Vamos mostrar agora que, para a região de parâmetros em que o MA é invertível, $|c| < 1$, o chute inicial faz pouca diferença, com T grande. De fato, observe que, para todo t:

$$\tilde{\epsilon}_t(c) = \check{\epsilon}_t(c) - c^t(\tilde{\epsilon}_0 - \epsilon_0) .$$

Usando esse fato, notamos que podemos escrever

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T}(Y_t - c\tilde{\epsilon}_{t-1}(c))^2 &= \frac{1}{T}\sum_{t=1}^{T}(Y_t - c\check{\epsilon}_{t-1}(c))^2 \\ -\frac{2}{T}(\tilde{\epsilon}_0 - \epsilon_0)\sum_{t=1}^{T}(Y_t - c\check{\epsilon}_{t-1}(c))c^t &+ \frac{1}{T}(\tilde{\epsilon}_0 - \epsilon_0)^2\sum_{t=1}^{T}c^{2t} \end{aligned} \tag{3}$$

Observe que, como $|c| < 1$ temos que o terceiro termo é limitado por cima por:

$$\left| \frac{1}{T}(\tilde{\epsilon}_0 - \epsilon_0)^2 \sum_{t=1}^{T} c^{2t} \right| \leq \frac{1}{T}\frac{|c|^2}{1 - |c|^2}|\tilde{\epsilon}_0 - \epsilon_0|^2 .$$

Assim, a contribuição desse termo desaparece com T grande.

Quanto ao segundo termo, nós notamos que

$$(Y_t - c\check{\epsilon}_{t-1}(c)) = \sum_{i=1}^{t}(-c)^{t-i}Y_i + (-c^t)\epsilon_0$$

Assim, podemos reescrever a parte relevante do segundo termo como:

$$\frac{1}{T}\sum_{t=1}^{T}(Y_t - c\check{\epsilon}_{t-1}(c))c^t = \frac{1}{T}\sum_{t=1}^{T}c^t((-c)^0+(-c)^1+\ldots+(-c)^{T-t})Y_t+\frac{\epsilon_0}{T}\sum_{t=1}^{T}(-c^{2t})\,.$$

O segundo termo da expressão acima desaparece, com T grande, quando $|c| < 1$. Quanto ao outro termo, observe que ele possui média zero, visto que o MA(1) que estamos analisando possui média zero. Além disso:

$$\begin{aligned}
&\mathbb{V}\left[\frac{1}{T}\sum_{t=1}^{T}c^t((-c)^0 + (-c)^1 + \ldots + (-c)^{T-t})Y_t\right] = \\
&\frac{1}{T^2}\sum_{t=1}^{T}c^{2t}((-c)^0 + (-c)^1 + \ldots + (-c)^{T-t})^2\mathbb{V}[Y_t] \\
&+2\frac{1}{T^2}\sum_{i<j}c^{i+j}((-c)^0 + (-c)^1 + \ldots + (-c)^{T-i})((-c)^0 + (-c)^1 + \ldots + (-c)^{T-j})\operatorname{cov}(Y_i, Y_j) = \\
&\frac{1}{T^2}\mathbb{V}[Y_1]\sum_{t=1}^{T}c^{2t}((-c)^0 + (-c)^1 + \ldots + (-c)^{T-t})^2 \\
&+2\frac{1}{T^2}\operatorname{cov}(Y_1, Y_2)\sum_{i=1}^{T-1}c^{2i+1}((-c)^0 + (-c)^1 + \ldots + (-c)^{T-i})((-c)^0 + (-c)^1 + \ldots + (-c)^{T-i-1})
\end{aligned} \tag{4}$$

onde a última igualdade usa que o MA(1) é estacionário e que somente as autocovariâncias de primeira ordem são distintas de zero. Segue, então, que, quando $T \to \infty$, a variância acima vai a zero, de onde concluímos que:

$$\operatorname{plim}_{T\to\infty}\frac{2}{T}(\tilde{\epsilon}_0 - \epsilon_0)\sum_{t=1}^{T}(Y_t - c\check{\epsilon}_{t-1}(c))c^t = 0$$

Assim, temos que, para $|c| < 1$:

$$\operatorname{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}(Y_t - c\tilde{\epsilon}_{t-1}(c))^2 = \operatorname{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}(Y_t - c\check{\epsilon}_{t-1}(c))^2\,,$$

de modo que a escolha de $\tilde{\epsilon}$ não afeta o comportamento do objetivo, para T grande, na região invertível do MA.