Processos de média móveis: parte 2

Prof. Caio Azevedo

Estimação - método dos momentos

- Considere $\{Y_t\}$ um processo MA(q) de média μ (link).
- Uma vez que a média do processo é μ, então o EMM (estimador pelo método dos momentos) para μ é dado por $\widehat{\mu} = \bar{Y}$ (semelhante ao processo AR(p), link).
- Por outro lado, para $\theta_1, \ldots, \theta_q$ os respectivos EMM's são obtidos resolvendo-se as equações definidas pelas auto-correlações amostrais.
- Por exemplo, considere o modelo MA(1) de média μ: $Y_t = \mu + \theta\epsilon_{t-1} + \epsilon_t$, em que $\epsilon_t \sim RB(0, \sigma^2)$. Para este processo temos que $\gamma(0) = \sigma^2(1+\theta^2)$, $\gamma(1) = \sigma^2\theta$ e portanto $\rho(1) = \theta/(1+\theta^2)$.

Estimação - método dos momentos

- Então um estimador para θ é encontrado resolvendo-se:

$$\widehat{\rho}(1) = \frac{\widehat{\theta}}{1 + \widehat{\theta}^2}, \tag{1}$$

a qual é uma equação do 2^o grau.

- Nesse caso, devemos escolher a solução cujo valor de $\widehat{\theta}$ torne o processo invertível (ser possível escrever $\epsilon_t = \sum_{j=0}^{\infty} (-\theta)^j \, Y_{t-j}$).

Estimação - método dos momentos

- Além disso, as soluções serão números reais se e somente se $|\widehat{\rho}(1)| < 1/2$ (para esse caso).
- Se $|\widehat{\rho}(1)| < 1/2$, resolvendo a equação (1) temos que a solução inversível é dada por:

$$\widehat{\theta} = \frac{1 - \sqrt{1 - 4\widehat{\rho}(1)^2}}{2\widehat{\rho}(1)}.$$

Estimação - máxima verossimilhança

- Diferentemente do modelo AR(p), para um modelo MA(q) é útil escrevermos a verossimilhança associada em termos dos erros de previsão a um passo a frente $Y_t - Y_t(1)$ (veremos adiante que pode ser calculada tanto por (3) no slide 14 quanto por (4) no slide 17) uma vez que temos uma verossimilhança mais tratável neste caso.
- Para um modelo MA(q), seja $\boldsymbol{\beta} = (\mu, \theta_1, \ldots, \theta_q)'$ o vetor de dimensão $q + 1$ de parâmetros. Temos que a respectiva verossimilhança é dada por:

$$L(\boldsymbol{\beta}, \sigma^2) = \prod_{t=1}^{n} f(y_t | y_{t-1}, \ldots, y_1).$$

Estimação - máxima verossimilhança

- Com efeito, temos que a distribuição de $Y_t | Y_{t-1} = y_{t-1}, \ldots, Y_1 = y_1$ é normal com média y_{t+1}^t e variância $Var(P_{t+1}^t) = \sigma^2$. Portanto a verossimilhança pode ser escrita como:

$$L(\boldsymbol{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^{n} (y_t - y_{t+1}^t)^2 \right\}, \qquad (2)$$

em que y_{t+1}^t é função somente de β (como veremos adiante).

- Por fim os estimadores são obtidos maximizando-se a equação (2) com respeito a $\boldsymbol{\beta}$ e σ^2.

Estimação - mínimos quadrados condicionais

- Podemos escrever um modelo MA(q) de média μ em termos de seus erros:

$$\epsilon_t = Y_t - \mu - \sum_{k=1}^{q} \theta_k \epsilon_{t-k}.$$

- Neste método de estimação aproximamos a soma dos quadrados dos erros pelo condicionamento em $\epsilon_0 = \epsilon_1 = \cdots = \epsilon_{1-q} = 0$. Portanto, os estimadores de $\boldsymbol{\beta} = (\theta_1, \ldots, \theta_q)'$ são dados minimizando-se:

$$S_c(\boldsymbol{\beta}) = \sum_{t=1}^{n} \epsilon_t^2.$$

- Exercício (expanda o somatório acima).

Comentários

- Essencialmente, os comentários feitos aqui, slides 21, 24 e 41.
- São necessários métodos numéricos para obter as estimativas (MQC e MV).
- Erros-padrão, intervalos de confiança e testes de hipótese podem ser obtidos/construídos através de teoria assintótica ou métodos numéricos.

Previsão - Método 1

- Considere $\{Y_t\}$ um processo $MA(q)$ invertível. Sem perda de generalidade assuma que média zero. Caso não o tenha, basta centrá-lo em sua média $Y_t = X_t - \mu$. Logo temos que $Y_t = \theta(B)\epsilon_t$, em que $\epsilon_t \sim RB(0, \sigma^2)$.
- Desejamos obter previsões m passos a frente para esse processo (semelhante ao que fora feiro para o processo $AR(p)$).

Previsão Método 1

- Para isso consideraremos $Y_n(m) = \mathcal{E}(Y_{n+m}|Y_n, \ldots, Y_1)$ o preditor ótimo sob o erro quadrático médio, baseado na amostra finita $(Y_1, \ldots, Y_n)$, e $\widetilde{Y}_n(m) = \mathcal{E}(Y_{n+m}|Y_n, \ldots, Y_1, Y_0, Y_{-1}, \ldots)$ o preditor correspondente, baseado na amostra infinita $(\ldots, Y_{-1}, Y_0, Y_1, \ldots, Y_n)$ ou simplesmente $(Y_n, \ldots)$.
- Utilizaremos $\widetilde{Y}_n(m)$ pois as respectivas contas são mais simples se comparadas àquelas relativas à $Y_n(m)$ (para os processos MA(q) e também para os processos ARMA(p,q), que veremos adiante).
- Uma vez encontrado $\widetilde{Y}_n(m)$ ele é truncado para o caso finito (amostra de tamanho n).

Previsão - Método 1

- Primeiramente, temos que como $\{Y_t\}$ é invertível, logo

$$
\begin{aligned}
Y_{n+m} &= \theta(B)\epsilon_{n+m} \\
\epsilon_{n+m} &= \sum_{j=0}^{\infty} \pi_j Y_{n+m-j}, \ \pi_0 = 1, \\
\mathcal{E}(\epsilon_{n+m}|Y_n, \ldots, Y_{-1}, \ldots) &= \sum_{j=0}^{\infty} \pi_j \mathcal{E}(Y_{n+m-j}|Y_n, \ldots, Y_{-1}, \ldots) \\
0 &= \sum_{j=0}^{\infty} \pi_j \widetilde{Y}_n(m-j)
\end{aligned}
$$

Previsão - Método 1

- Continuando:

$$
\begin{aligned}
0 &= \widetilde{Y}_n(m) + \sum_{j=1}^{\infty} \pi_j \widetilde{Y}_n(m-j) \\
\widetilde{Y}_n(m) &= -\sum_{j=1}^{\infty} \pi_j \widetilde{Y}_n(m-j) = -\sum_{j=1}^{m-1} \pi_j \widetilde{Y}_n(m-j) - \sum_{j=m}^{\infty} \pi_j \widetilde{Y}_n(m-j).
\end{aligned}
$$

Porém note que $\widetilde{Y}_t(0) = \mathcal{E}(Y_t | Y_n, \ldots, Y_{-1}, \ldots) = Y_t$, para $t \leq n$, devido a forma do processo MA (exercício). Logo ($\widetilde{Y}_n(m-j) = Y_{n+m-j}$) e

$$
\widetilde{Y}_n(m) = -\sum_{j=1}^{m-1} \pi_j \widetilde{Y}_n(m-j) - \sum_{j=m}^{\infty} \pi_j Y_{n+m-j},
$$

Previsão - Método 1

- Dessa forma, começando com $m = 1$, temos um procedimento recursivo para obtermos previsões para um modelo MA(q).
- Por definição, esse procedimento também pode ser utilizado para acharmos previsões para um processo ARMA(p, q), o qual estudaremos adiante.
- Note que esse método necessita que o modelo seja invertível e que seja fácil a obtenção de π_j.
- Na prática fazemos:

$$\widetilde{Y}_n(m) = -\sum_{j=1}^{m-1} \pi_j \widetilde{Y}_n(m-j) - \sum_{j=m}^{n} \pi_j Y_{n+m-j}$$

Estrutura - Previsão - Método 1

$$
\begin{aligned}
\widetilde{Y}_n(1) &= -\sum_{j=1}^{n} \pi_j Y_{n+1-j} \qquad (3)\\
\widetilde{Y}_n(2) &= -\pi_1 \widetilde{Y}_n(1) - \sum_{j=2}^{n} \pi_j Y_{n+2-j}\\
\widetilde{Y}_n(3) &= -\pi_1 \widetilde{Y}_n(2) - \pi_2 \widetilde{Y}_n(1) - \sum_{j=3}^{n} \pi_j Y_{n+3-j}\\
&\vdots
\end{aligned}
$$

Exemplo - Previsão - Método 1

- Considere o modelo MA(1) em que $Y_t = -0,3\epsilon_{t-1} + \epsilon_t$, $\epsilon_t \sim RB(0, \sigma^2)$.
- Sabemos que $Y_t = (1 - 0,3B)\epsilon_t$ e que $\theta(z) = 1 - 0,3z = 0 \iff z \approx 3,33 > 1$, logo $\{Y_t\}$ é invertível e portanto:

$$\epsilon_t = \frac{1}{1 - 0,3B} Y_t = \sum_{j=0}^{\infty} 0,3^j Y_{t-j} = \sum_{j=0}^{\infty} \pi_j Y_{t-j},$$

em que $\pi_j = 0,3^j$.

Exemplo - Previsão - Método 1

- Logo, a previsão a m passos é dada recursivamente por:

$$\widetilde{Y}_n(m) = -\sum_{j=1}^{m-1} 0,3^j \widetilde{Y}_n(m-j) - \sum_{j=m}^{n} 0,3^j Y_{n+m-j}.$$

Previsão - Método 2

- Do processo MA(q) temos que:

$$\begin{aligned} Y_{n+m} &= \theta_1 \epsilon_{n+m-1} + \cdots + \theta_q \epsilon_{n+m-q} + \epsilon_{n+m} \\ \mathcal{E}(Y_{n+m}|Y_n, \ldots) &= \theta_1 \mathcal{E}(\epsilon_{n+m-1}|Y_n, \ldots) + \cdots + \theta_q \mathcal{E}(\epsilon_{n+m-q}|Y_n, \ldots) \\ &+ \mathcal{E}(\epsilon_{n+m}|Y_n, \ldots) \\ \widetilde{Y}_n(m) &= \theta_1 \widetilde{\epsilon}_n(m-1) + \cdots + \theta_q \widetilde{\epsilon}_n(m-q) + \widetilde{\epsilon}_n(m), \qquad (4) \end{aligned}$$

notando (devido a causalidade e a invertibilidade) que $\widetilde{\epsilon}_n(t) = \mathcal{E}(\epsilon_t|Y_n, \ldots) = \mathcal{E}(\epsilon_t|\epsilon_n, \ldots) = \epsilon_t$ para $t \leq n$ (exercício).

- Mas como imputar ϵ_t na prática dado que ele é estocástico?

Previsão - Método 2

- Note que do processo MA(q): $\epsilon_t = Y_t - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q}$, portanto assumindo que $y_i = 0$ e $\epsilon_i = 0$ para $i \leq 0$, podemos proceder recursivamente, ou seja:

$$
\begin{aligned}
\epsilon_1 &= y_1 \\
\epsilon_2 &= y_2 - \theta_1\epsilon_1 \\
\epsilon_3 &= y_3 - \theta_1\epsilon_2 - \theta_2\epsilon_1 \\
\epsilon_4 &= y_4 - \theta_1\epsilon_3 - \theta_2\epsilon_2 - \theta_3\epsilon_1 \\
&\vdots \\
\epsilon_n &= y_n - \theta_1\epsilon_{n-1} - \cdots - \theta_q\epsilon_{n-q}.
\end{aligned}
$$

Exemplo - Previsão - Método 2

- Considere novamente o modelo MA(1) em que $Y_t = -0,3\epsilon_{t-1} + \epsilon_t$, $\epsilon_t \sim RB(0, \sigma^2)$.
- Dito isso, temos que $\widetilde{Y}_n(m) = -0,3\widetilde{\epsilon}_n(m-1) + \widetilde{\epsilon}_n(m)$, e portanto:

$$\begin{aligned} \widetilde{Y}_n(1) &= -0,3\widetilde{\epsilon}_n(0) = -0,3\epsilon_n \\ \widetilde{Y}_n(k) &= 0, k \geq 2, \end{aligned}$$

em que (próximo slide)

Exemplo - Previsão - Método 2

$$\begin{aligned}
\epsilon_1 &= y_1 \\
\epsilon_2 &= y_2 + 0,3\epsilon_1 = y_2 + 0,3y_1 \\
\epsilon_3 &= y_3 + 0,3\epsilon_2 = y_3 + 0,3(y_2 + 0,3y_1) \\
&= y_3 + 0,3y_2 + 0,3^2 y_1 \\
&\vdots \\
\epsilon_n &= \sum_{j=1}^{n} 0,3^{n-j} y_j.
\end{aligned}$$

Erro de previsão

- O erro de previsão para m passos a frente de $1 \leq m \leq q$ é dado por (veja slides anteriores):

$$\begin{aligned}
P_n(m) = Y_{n+m} - \widetilde{Y}_n(m) &= (\theta_1\epsilon_{n+m-1} + \cdots + \theta_q\epsilon_{n+m-q} + \epsilon_{n+m}) \\
&- (\theta_1\widetilde{\epsilon}_n(m-1) + \cdots + \theta_q\widetilde{\epsilon}_n(m-q) + \widetilde{\epsilon}_n(m)) \\
&= \sum_{j=0}^{m-1} \theta_j\epsilon_{n+m-j}, \ \theta_0 = 1.
\end{aligned}$$

- A respectiva variância então é dada por:

$$\sigma_n^2(m) = Var(P_n(m)) = \sigma^2 \sum_{j=0}^{m-1} \theta_j^2,$$

Erro de previsão

- E a covariância é dada por:

$$Cov(P_n(m), P_n(m+k)) = \sigma^2 \sum_{j=0}^{m-1} \theta_j \theta_{j+k},$$

logo, temos que os erros de previsão são correlacionados.

Propriedades

- Considere $\{X_t\}$ um processo MA(q) de média μ e que desejamos fazer previsões para X_{n+m}. Dessa forma podemos trabalhar com $Y_t = X_t - \mu$, realizar previsões para Y_{n+m} de interesse e recuperar as previsões para X_{n+m} tal que $\widetilde{X}_n(m) = \widetilde{Y}_n(m) + \mu$.
- Além disso, sabemos que $\widetilde{X}_n(m) = \mu + \theta_1 \widetilde{\epsilon}_n(m-1) + \cdots + \theta_q \widetilde{\epsilon}_n(m-q) + \widetilde{\epsilon}_n(m)$. Logo quando $m \to \infty$, $\widetilde{X}_n(m) \to \mu$ e consequentemente $Var(P_n(m)) \to \gamma_X(0)$.
- Portanto, temos que a previsão para um processo MA(q) a longo prazo torna-se a própria média do processo.