

Lista 5 - Series

Davi Wentrick Feijó

2024-06-03

Considere a serie do consumo mensal de energia eletrica (ConsumoEnergiaEAqua_New.xlsx). Denotando X_t como o valor do consumo registrado no mes t e D_t como o número de dias de leitura, faca o que se pede a seguir.

1. Calcule o consumo médio diário $Y_t = \frac{x_t}{D_t}$ e explique o porquê dessa transformacao.
2. Apresente o gráfico da evolucao temporal de $\{Y_t\}$, e apresente sua descricao, contemplado elementos como o tamanho da serie e periodicidade dos dados.
3. Apresente os gráficos da funcao de autocorrelacao (FAC) e da funcao de autocorrelacao parcial (FACP) de $\{Y_t\}$, considerando um número apropriado de defasagens (lag), incluindo a banda de 95% de confianca sob a hipótese nula de nao haver autocorrelacao serial. Em um parágrafo, descreva as formas da FAC e da FACP, explicando o que se pode diagnosticar/sugerir com base nelas.
4. Aplique o teste aumentado de estacionariedade de Dickey-Fuller do pacote aTSA do R. Para a parte sazonal, faca a avaliacao por meio de um modelo de regressao com funções harmonicas.
5. Calcule a variacao do consumo $Z_t = Y_t - Y_{t-1}$, e explique o papel/significado dessa transformacao para a análise desses dados.
6. Faca o gráfico da evolucao temporal de $\{Z_t\}$, e descreva em um parágrafo o aspecto dessa figura, comparando-a com a forma observada no item 2.
7. Repita os passos 3 e 4, comparando os novos resultados com os anteriores.

Considere que \hat{Y}_{t+h} representa a previsao no instante $t+h$ obtida com base nas informações disponíveis até o tempo t ; ou seja,

$$Y_1, \dots, Y_t \rightarrow \hat{Y}_{t+h}$$

Separe a massa de dados em duas partes, conforme esquema abaixo:

Treinamento (modelagem)	validacao
Y_1, \dots, Y_m	Y_{m+1}, \dots, Y_n

Utilizando os dados de treinamento: 8. Considerando o modelo $\text{SARIMA}(p, d, q) \times (P, D, Q)_s$ para a série Y_t , defina um valor apropriado para a ordem sazonal s e as ordens de diferenciaciones d e D com base nos passos anteriores.

9. Defina uma malha de valores para as ordens autorregressivas p e P e de médias móveis q e Q e obtenha o valor do critério de informacao bayesiano de Schwarz (BIC) para cada combinacao $(p, d, q) \times (P, D, Q)$ por meio da funcao sarima do pacote astsa.

10. Liste os modelos com os menores BIC. Certifique-se que o melhor modelo não possua uma ordem na extremidade da malha definida no item 9. Se houver, retorne para o passo 9 , ampliando a malha.
11. Inicie o diagnóstico com o modelo que apresenta o menor BIC:
 - 11.1 Analise as estimativas dos parâmetros por meio da função `sarima` do pacote `astsa`.
 - 11.2 Faça os gráficos da FAC e FACP residual, e aplique o teste de LjungBox.
 - 11.3 Teste a normalidade residual.
 - 11.4 Caso haja problemas em 11.1 e 11.2 , repita a análise com os próximos modelos candidatos.
 - 11.5 Caso não seja possível encontrar um modelo adequado, será preciso redefinir o modelo no passo 8 . Se as ordens s , d , e D estiverem corretas, então é possível que o modelo SARIMA não seja apropriado. Utilizando os dados de validação:
12. Como o método de estimação é recursivo, a obtenção dos erros de previsão um passo à frente na massa de validação pode ser realizada da seguinte forma:
 - 12.1 Aplique o modelo sobre a base de dados completa, usando a função `sarima` do pacote `astsa`.