
PROVA 3 - SÉRIES TEMPORAIS

Paulo Renato Batista Laeber

7 de dezembro de 2023

1 Introdução

Seja y o vetor de de tamanho $(n \times 1)$, denotando a variável de interesse para estimação; \mathbf{X} uma matriz de tamanho $(n \times k)$, denotando as variáveis explicativas associadas ao y ; e β um vetor de tamanho $(k \times 1)$, denotando os parâmetros de cada covariável. O estimador de mínimo quadrados ordinários (MQO) é definido por:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \epsilon) = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon \quad (1)$$

Para que este estimador seja aplicável e o melhor possível para os dados disponíveis é necessário que as seguintes suposições sejam satisfeitas:

- S1. Linearidade: O processo gerador de y deve ser $y = \mathbf{X}\beta + \epsilon$.
- S2. Posto coluna completo: A matriz \mathbf{X} deve ter posto coluna completo para que seja singular e, assim, invertível.
- S3. Exogeneidade das covariáveis: Não há correlação entre as variáveis explicativas e os erros. $E[\epsilon|\mathbf{X}] = 0$.
- S4. Esfericidade dos erros: Cada ϵ_i , $i = 1, \dots, n$, possui a mesma variância σ^2 e são não-correlacionados entre si, ou seja, $E[\epsilon_i\epsilon_j|\mathbf{X}] = 0$, $i \neq j$, $i, j = 1, \dots, n$.
- S5. Normalidade dos erros: Todos os erros $\epsilon|\mathbf{X}$ são normalmente distribuídos.

Com as suposições S1-S4, pelo teorema de Gauss-Markov, o estimador $\hat{\beta}$ é o estimador não viesado de mínima variância de β . Apesar de o teorema de Gauss-Markov ser resiliente a violação da suposição S5, ela é necessária para construção de intervalos de confiança e estatísticas importantes sobre o estimador.

Em um cenário com séries temporais há problemas de violações de suposição. Já que a séries temporais naturalmente possuem autocorrelação, a suposição S4 é violada e o teorema de Gauss-Markov é perdido.

1.1 Estacionariedade

Teorema 1 (Estacionariedade Forte): Seja $\{X_t\}_{t=-\infty}^{t=\infty}$, um processo de série temporal. $\{X_t\}$ é dito fortemente estacionário se qualquer conjunto de k observações na sequência $[X_t, X_{t+1}, \dots, X_{t+k}]$ possui a mesma distribuição conjunta.

Adicionalmente, X_t é dito fracamente estacionário se $E[X_t]$ é constante por todo tempo t e $Cov[X_t, X_{t-k}] < \infty$ e independente de t .

Estacionariedade forte automaticamente implica estacionariedade fraca, mas o inverso não é verdadeiro. Mas como é difícil encontrar exemplos de processos fracamente estacionários e não estacionariamente fortes, exceto em alguns casos teóricos, é assumido estacionariedade forte para todos os processos estacionários neste trabalho.

1.2 Teorema ergótico

Teorema 2 (Ergodicidade): Seja $\{X_t\}_{t=-\infty}^{t=\infty}$ um processo de série temporal. $\{X_t\}$ é ergótico se qualquer duas funções limitadas que transforma vetores de dimensões a e b em escalares reais, $f: \mathbb{R}^a \rightarrow \mathbb{R}$, e $g: \mathbb{R}^b \rightarrow \mathbb{R}$,

$$\lim_{k \rightarrow \infty} |E[f(X_t, \dots, X_{t+a})g(X_{t+k}, \dots, X_{t+b+k})]| = |E[f(X_t, \dots, X_{t+a})]| |E[g(X_{t+k}, \dots, X_{t+b+k})]| \quad (2)$$

O que significa que se dois eventos estão afastados o suficiente no tempo, então se tornam "asintoticamente independente". Isso implica que em cada observação da série, existe uma informação única, o que nos permite considerar MQO em séries temporais.

Teorema 3 (Funções Ergóticas): Se $\{X_t\}_{t=-\infty}^{t=\infty}$ é um processo estacionário e ergótico e se $y_t = f\{X_t\}$ é uma função mensurável com o mesmo espaço probabilístico que define X_t , então y_t também é estacionário e ergótico. Expandindo esse teorema, seja $\{X_t\}_{t=-\infty}^{t=\infty}$ denotar um vetor de k processos estocásticos, onde cada processo é estacionário e ergótico, então as funções de $\{X_t\}_{t=-\infty}^{t=\infty}$ também são estacionárias e ergóticas.

Com isso, se definirmos

$$y_t = X_t \beta + \epsilon_t \quad (3)$$

então temos um modelo de MQO válido para séries temporais, enquanto $\{X_t, \epsilon_t\}_{t=-\infty}^{t=\infty}$ ser juntamente estacionário e ergótico. Caso y_t e X_t não forem estacionários, é possível obter a estacionariedade se ambos y_t e X_t forem cointegráveis, modificando o modelo na seguinte maneira:

$$\Delta^k y_t = \Delta^k X_t + \epsilon_t \quad (4)$$

onde Δ é o operador de diferenciação dos processos e $\{\Delta X_t, \epsilon_t\}_{t=-\infty}^{t=\infty}$ é juntamente estacionário e ergótico.

1.3 Convergência a normalidade

Para finalizar a adaptação do MQO para séries temporais, precisamos encontrar uma convergência a distribuição. Como não podemos utilizar o teorema central do limite de Lindeberg-Levy, é necessário encontrar uma alternativa.

Seja Z_t um vetor de sequência de Martingale, definido, especificamente, como um passeio aleatório:

$$Z_t = Z_{t-1} + u_t \quad (5)$$

onde $Cov(u_t, u_s) = 0$ para $t \neq s$. Assim temos $E[Z_t | Z_{t-1}, Z_{t-2}, \dots] = Z_{t-1}$. Seja Z_t^* um vetor de sequência de diferenças de Martingale se $E[Z_t^* | Z_{t-1}^*, Z_{t-2}^*, \dots] = 0$, com este vetor em específico, temos o seguinte teorema:

Teorema 4 (Teorema Central do Limite em diferenças de Martingale): Se Z_t^* é um vetor de sequência de diferenças de Martingale, onde $E[Z_t^* Z_t^{*'}] = \Sigma$ com Σ sendo uma matriz positiva definida, e $\bar{Z}_T^* = (1/T) \sum_{t=1}^T Z_t^*$, então $\sqrt{T} \bar{Z}_T^* \xrightarrow{d} N(0, \Sigma)$.

O Teorema 4 ainda não é robusto o suficiente para cobrir casos de autocorrelação, mas já é uma evolução do teorema de Lindeberg-Levy. Suponha um processo multivariado $\{X_t\}_{t=-\infty}^{t=\infty}$ estacionário e ergótico e $\{\epsilon_t\}_{t=-\infty}^{t=\infty}$ um processo i.i.d. Com isso temos um processo $\{X_t \epsilon_t\}_{t=-\infty}^{t=\infty}$ que pode ser visto como a sequência de diferenças de Martingale, possibilitando a ligação do modelo obtido na Equação (3) a uma distribuição.

Temos que considerar agora um teorema central do limite que seja robusto o bastante para cobrir casos de dependência estocástica a X_t e autocorrelação em ϵ_t . Seja $\{Z_t\}$ um processo estacionário e ergótico onde $\sqrt{T} \bar{Z}_T$ segue as seguintes condições:

C1. Não correlação assintótica: $E[Z_t | Z_{t-k}, Z_{t-k-1}, \dots] \xrightarrow{MQ} 0$ quando $k \rightarrow \infty$.

C2. Somabilidade das autocovariâncias: Com observações dependentes temos

$$\lim_{T \rightarrow \infty} Var[\sqrt{T} \bar{Z}_T] = \sum_{t=0}^{\infty} \sum_{s=0}^{\infty} Cov[Z_t Z_s'] = \sum_{k=-\infty}^{\infty} \Gamma_k = \Gamma \quad (6)$$

onde Γ é uma matriz finita

C3. Negligibilidade assintótica de inovações: Seja

$$r_{tk} = E[Z_t | Z_{t-k}, Z_{t-k-1}, \dots] - E[Z_t | Z_{t-k-1}, Z_{t-k-2}, \dots] \quad (7)$$

uma observação Z_t pode ser vista como o acúmulo de informações que entraram desde o processo começou até o tempo t

$$Z_t = \sum_{s=0}^t r_{ts} \quad (8)$$

Com essas condições temos o seguinte teorema:

Teorema 5 (Teorema Central do Limite de Gordin): Se $\{Z_t\}$ é um processo fortemente estacionário e ergótico cumprindo as condições C1-C3, então $\sqrt{T}\bar{Z}_T \xrightarrow{d} N(0, \Gamma)$

2 Mínimos Quadrados Ordinários em séries temporais

De agora em diante considere $\{X_t\}_{t=1}^{t=T}$ um vetor de processos de séries temporais servindo como covariáveis a $\{y_t\}_{t=1}^{t=T}$. É necessário o uso da forma assintótica do estimador MQO, definido na seguinte maneira:

$$\hat{\beta} = (X_t' X_t)^{-1} X_t' y_t = \beta + \left(\frac{X_t' X_t}{T} \right)^{-1} \left(\frac{X_t' \epsilon_t}{T} \right) \quad (9)$$

onde β . Se o processo gerador de X_t for estacionário e ergótico, então $(X_t' X_t)/T$ converge a uma matriz positiva \mathbf{Q} , na qual podemos aplicar o teorema de Slutsky para obter a inversa. Se ϵ_t não possuir autocorrelação, então $Z_t^* = X_t \epsilon_t$ é uma sequência de diferenças de Martingale, e assim $(X_t' \epsilon_t)/T$ converge a zero:

$$\lim_{T \rightarrow \infty} \hat{\beta} = \beta + \lim_{T \rightarrow \infty} \left(\frac{X_t' X_t}{T} \right)^{-1} \left(\frac{X_t' \epsilon_t}{T} \right) \quad (10)$$

$$= \beta + \mathbf{Q}^{-1} \cdot 0 = 0 \quad (11)$$

As Equações (10) e (11) mostram que o estimador é assintoticamente não-viesado e consistente, mas, como não temos amostras infinitas, o desempenho do MQO poderá ser fraco. Já para a normalidade assintótica podemos invocar o Teorema 4 para encontrar a distribuição do estimador.

Caso haver autocovariância em ϵ_t , se $\{X_t, \epsilon_t\}_{t=1}^{t=T}$ serem juntamente estacionários e ergóticos, então podemos utilizar o Teorema 3 para encontrar as suas matrizes de momentos e estabelecer consistência. Para encontrar a normalidade, basta seguir o Teorema 5, desde que suas condições sejam cumpridas.

2.1 Estimando a variância do estimador

O problema de estimar a variância é que a autocovariância presente no modelo torna o estimador da variância viesado. A variância do modelo é dada por $\sigma^2 (X_t' X_t)^{-1} (X_t' \Omega X_t) (X_t' X_t)^{-1}$, enquanto o estimador da variância é dado por $s^2 (X_t' X_t)^{-1}$.

Para fazer qualquer uso de MQO, é necessário estimar a matriz de autocovariâncias Ω , ou usar um estimador para Heterocedasticidade. A única outra alternativa é usar mínimos quadrados generalizados, que já leva em conta a autocovariância e heterocedasticidade.

Estimador consistente a heterocedasticidade de White: Um estimador que pode ser usado em casos de heterocedasticidade de MQO é dado por:

$$\frac{1}{T} \left(\frac{1}{T} X_t' X_t \right)^{-1} \left(\frac{1}{T} \sum_{i=1}^T e_i^2 X_t' X_t \right) \left(\frac{1}{T} X_t' X_t \right)^{-1} = T (X_t' X_t)^{-1} S_0 (X_t' X_t)^{-1} \quad (12)$$

onde e_i são os resíduos do estimador e $\frac{1}{T} \sum_{i=1}^T e_i^2 X_t' X_t \xrightarrow{P} \frac{1}{T} \sum_{i=1}^T \sigma^2 X_t' X_t$. Vale notar que este estimador é assintótico, dependendo de uma convergência em probabilidade e que não há necessidade de especificar o tipo de heterocedasticidade presente.

3 Aplicação

Como exemplos vamos explorar dois exemplos de aplicação de MQO em séries temporais. Para isso, será feito o uso da função `'tslm'` do pacote `'forecast'` no R.

Exemplo 1: Como primeiro exemplo, será usado os dados `'uschange'` do pacote `'fpp2'`, que é um conjunto de cinco séries temporais. Cada série descreve as mudanças quarternárias em (%) de consumo, renda, produção, popança e desemprego nos Estados Unidos de 1970 a 2016.

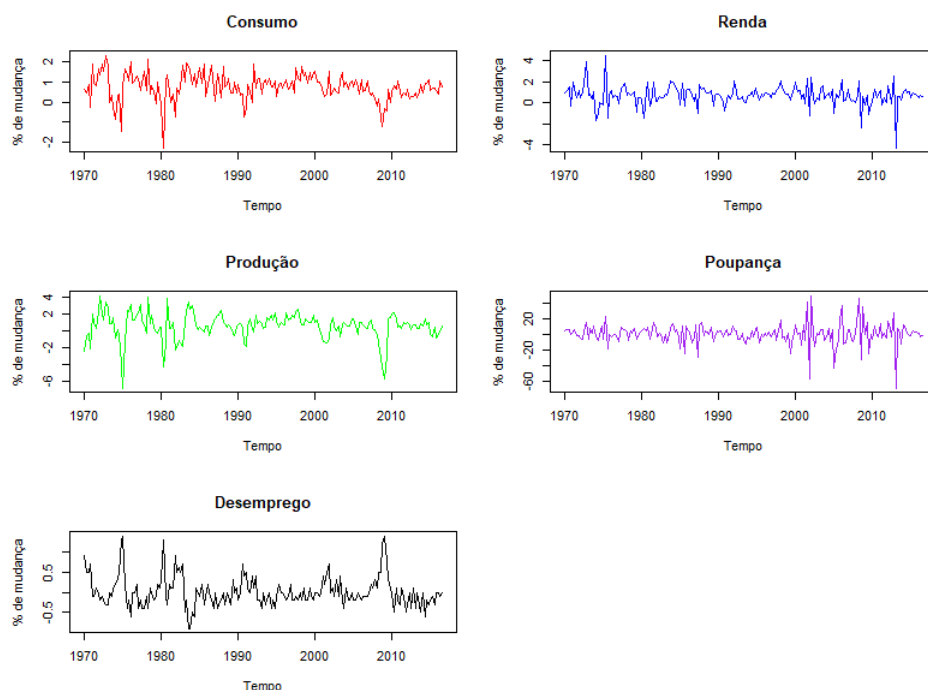


Figura 1: Séries presentes no conjunto `'uschange'`

A série de interesse é a de Consumo, no canto superior esquerdo da Figura 1. Após fazer o tratamento de outliers, a aplicação do teste Dickey-Fuller aumentado mostra que todas as séries são estacionárias. Aplicando a modelagem nas séries, obtemos os seguintes resultados:

Coeficientes	Estimativa	Erro padrão	t-valor	p-valor
(Intercepto)	0,4548	0,0631	7,204	≈ 0
Renda	0,4327	0,0749	5,775	≈ 0
Produção	0,0842	0,0302	2,782	0,0059
Poupança	-0,0267	0,0041	-6,514	≈ 0
Desemprego	-0,4143	0,1375	-3,013	0,0029

Tabela 1: Resultados para `'uschange'`

Pela Tabela 1, todas as estimativas dos coeficientes são significativas, e, em um ponto de vista financeiro, fazem sentido. Quanto maior a renda, maior o consumo; e quanto maior o desemprego, menor o consumo. Mas devemos lembrar que o estimador MQO é assintótico em séries temporais, e portanto o t-valor é apenas uma aproximação.

Aplicando o teste Breusch-Pagan, obtemos uma estatística $BP_4 = 4,8934$ com p-valor 0,2984, o que significa que os resíduos não são heterocedásticos; porém, pela Figura 2, os resíduos possuem autocorrelação.

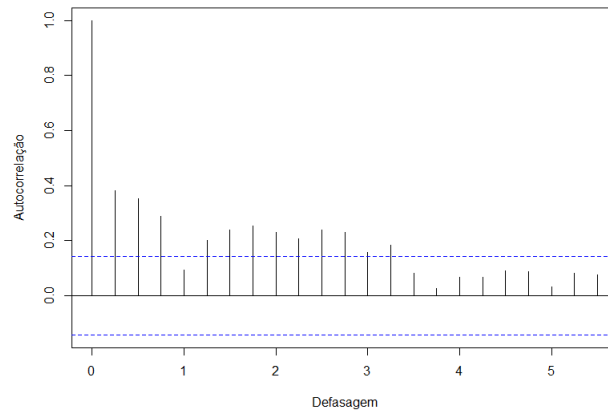


Figura 2: Autocorrelação dos resíduos de 'uschange'

Com isso, temos que invocar o Teorema 5 para conseguir uma convergência a normalidade. Infelizmente podemos apenas assumir ergodicidade, já que não há testes estatísticos para ergodicidade. Finalizando este exemplo com a seguinte figura:



Figura 3: Gráfico QQ dos resíduos

Exemplo 2: Aqui veremos o que fazer se temos apenas uma série. Brockwell e Davis (2016, p. 20) mostra que é possível decompor uma série em diferentes componentes:

$$y_t = m_t + s_t + R_t \quad (13)$$

onde m_t é o componente de tendência, s_t é o componente de sazonalidade e R_t é o ruído branco aleatório e estacionário. Vimos na Equação (4) que caso haver falta de estacionariedade em y_t e X_t , podemos usar as suas diferenciações como variáveis do modelo enquanto forem cointegráveis. Como é óbvio que uma série é cointegrável com sua própria tendência, podemos usar a Equação (13) como modelo.

Neste exemplo vamos usar os dados 'AirPassengers' do pacote 'datasets' como demonstração.

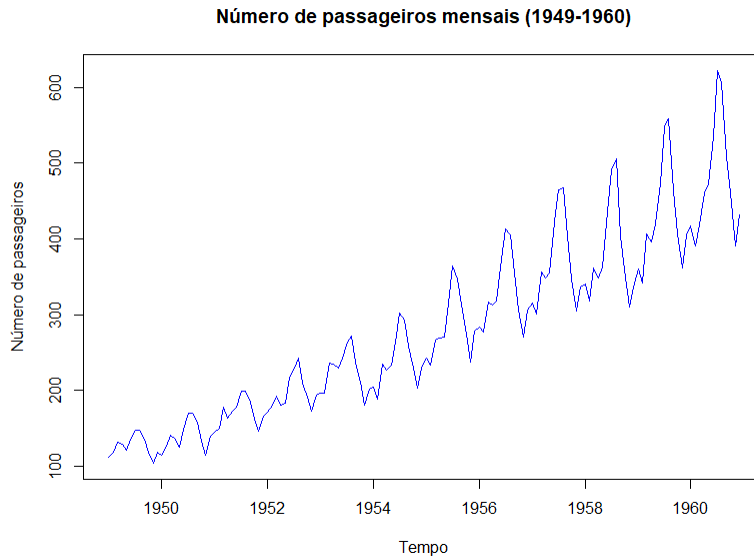


Figura 4: Série 'AirPassengers'

Ajustando a série com os seus componentes de tendência e sazonalidade, obtemos o seguinte resultado:

Coeficientes	Estimativa	Erro padrão	t-valor	p-valor
(Intercepto)	63,5079	8,3885	7,571	≈ 0
Tendência	2,6603	0,0529	50,225	≈ 0
Estação 2	-9,4103	10,749	-0,875	0,3829
Estação 3	23,0960	10,7498	2,149	0,0335
Estação 4	17,3523	10,7504	1,614	0,1089
Estação 5	19,4420	10,7513	1,808	0,0728
Estação 6	56,6150	10,7525	5,265	≈ 0
Estação 7	93,6213	10,7539	8,706	≈ 0
Estação 8	90,7110	10,7556	8,434	≈ 0
Estação 9	39,3840	10,7576	3,661	≈ 0
Estação 10	0,8903	10,7598	0,083	0,9341
Estação 11	-35,5199	10,7623	-3,300	0,0012
Estação 12	-9,1802	10,7650	-0,853	0,3953

Tabela 2: Resultados para 'AirPassengers'

Pela Tabela 2, os coeficientes estimados de sazonalidade variam bastante em valor e significância, não sendo exatamente um valor crescente como esperado. Isso pode ser causado pela definição das estações do algoritmo, que não está sincronizado de forma apropriada com a definição real de estações. Isso resultará em problemas sérios no modelo.

Aplicando o teste Breusch-Pagan foi obtido uma estatística $BP_{12} = 38,889$ com um p-valor de 0,0001, concluindo que os resíduos são heterocedásticos. Graficando a autocorrelação dos resíduos:

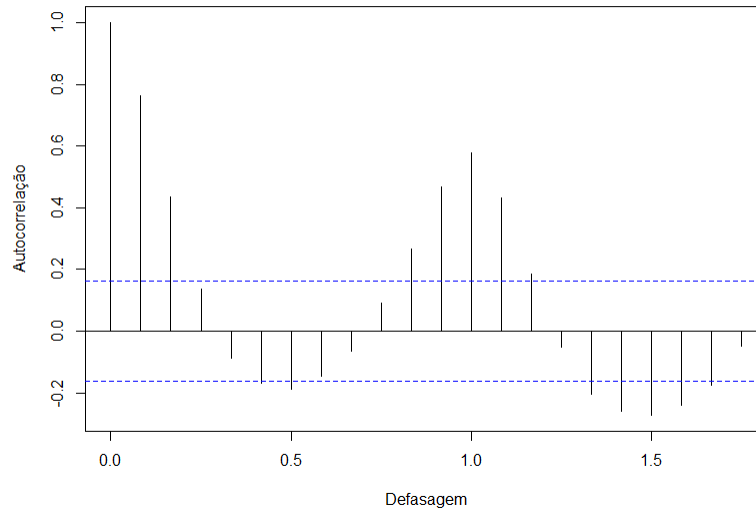


Figura 5: Autocorrelação dos resíduos de '*AirPassengers*'

Está claro que o modelo não conseguiu se encaixar na série de forma correta, resultando em valores inesperados nas estimativas dos coeficientes, resíduos heterocedásticos e autocorrelatados.

Referências

- [1] William E. Greene. *Econometric Analysis* 7th Edition. Capítulos 4, 9 e 20. Pearson, 2012.
- [2] Peter J. Brockwell, Richard A. Davis. *Introduction to time series and forecasting* 3rd Edition. Página 20. Springer, 2016.