

Introdução a Séries Temporais

Pedro Valls*
EESP-FGV e CEQEF-FGV

Sumário

1	Processos Estocásticos	5
1.1	Introdução	5
1.1.1	Exemplos	8
2	Restrições na heterogeneidade temporal - Processos Estacionários	9
2.1	Exercício	10
3	Restrições na memória do processo - assintoticamente independente e m-dependente.	11
3.1	Introdução	11
3.2	Processo Assintoticamente Independente	12
3.2.1	Inflação no Brasil medida pelo IPCA	12
3.2.2	O erro é ruído branco	16
3.2.3	Momentos não Condicionais e Autocorrelação do processo X_t	16
3.2.4	Correlograma para o processo AR(1)	20
3.2.5	Gerando processos AR(1) no R	21
3.2.6	Exercícios	23
3.3	Processo 1-dependente	23
3.3.1	Correlograma para o $MA(1)$	24

*EESP-FGV Rua Doutor Plínio Barreto 365, s.1319, 01313-020, São Paulo, S.P. E-mail:pedro.valls@fgv.br. ©2024-Pedro Valls

3.3.2	Gerando processos MA(1) no R	26
3.3.3	Exercícios	28
4	Processos Autorregressivos de Ordem Superior	28
4.1	Introdução	28
4.2	Processo AR(2) - Condições de Estacionaridade e Função de Autocorrelação	29
4.3	Função de Autocorrelação Parcial para o processo AR(2)	34
4.4	Distribuição da Função de Autocorrelação e de Autocorrelação Parcial	36
4.5	Testando cada elemento da Função de Autocorrelação	37
4.6	Teste Portmanteau	38
4.7	Função de Autocorrelação e de Autocorrelação Parcial no R . .	38
4.8	Exercícios	40
5	Processo Média Móvel de Ordem Superior	40
5.1	Introdução	40
5.2	Função de Autocorrelação	41
5.3	Função de Autocorrelação Parcial	42
5.4	Função de Autocorrelação e de Autocorrelação Parcial no R .	42
5.5	Exercícios	44
6	Processo ARMA	44
6.1	Introdução	44
6.2	Gerando um ARMA(1,1) no R	45
6.3	Função de Autocorrelação e de Autocorrelação Parcial para o ARMA(1,1)	45
6.4	Função de Autocorrelação e Autocorrelação Parcial no R. . . .	47
6.5	Exercícios	47

7	Representações de Variáveis Integradas	48
7.1	Introdução	48
7.2	Especificação de um processo $I(1)$	50
7.2.1	Modelo de Passeio Aleatório com Drift ou Modelo de Nível Local com Taxa de Crescimento Constante	51
7.2.2	Modelo de Tendência Local	52
7.2.3	Tendência Determinística ou Tendência Estacionária	53
7.2.4	Forma Reduzida do Modelo de Tendência Local sem Erro Observacional	55
7.2.5	Forma Reduzida do Modelo de Tendência Local com Erro Observacional	56
7.2.6	Principais diferenças entre séries $I(0)$ e $I(1)$	57
7.3	Gerando um passeio aleatório no R	57
7.4	Gerando um Passeio Aleatório com drift através do Modelo de Nível Local no R	60
8	Apêndice I - Algumas Transformações	62
9	Apêndice II - Solução de uma equação de primeira ordem no operador defasagem e sua inversão.	64
10	Apêndice III - Solução do Sistema (17-19)	65
11	Exercícios	66
12	Referências	68

Lista de Tabelas

Lista de Figuras

1	Três realizações de (1)	6
2	Série Temporal Observada	7
3	Inflação medida pelo IPCA	13
4	Diagrama de Dispersão $\text{IPCA}(-1) \times \text{IPCA}$	14
5	Inflação medida pelo IPCA - 80:01 - 03:04	15

6	Diagrama de Dispersão para o período de 80:02 até 03:04 . . .	15
7	Correlograma para AR(1) com $\phi = 0.5$	20
8	Correlograma para AR(1) com $\phi = -0.5$	20
9	Correlograma para AR(1) com $\phi = 0.99$	21
10	AR(1) com $\phi = 0.5$	21
11	Diagrama de dispersão da série e defasagens	22
12	Correlograma para MA(1) com $\theta = 0.5$	25
13	Correlograma para MA(1) com $\theta = -0.5$	25
14	Correlograma para MA(1) com $\theta = 0.99$	26
15	MA(1) com $\theta = 0.5$	27
16	MA(1) com $\theta = 0.5$	27
17	AR(2) com raízes reais	32
18	Função de Autocorrelação para AR(2) com raízes reais	33
19	AR(2) com raízes complexas	33
20	Função de Autocorrelação para AR(2) com raízes complexas .	34
21	AR(2) com raízes reais	39
22	Função de Autocorrelação e de Autocorrelação Parcial para AR(2) com raízes reais	39
23	MA(2)	43
24	Função de autocorrelação e autocorrelação parcial para MA(2)	43
25	ARMA(1,1)	45
26	FAC e FACP para ARMA(1,1)	47
27	PIB Brasileiro desde 1860 até 2001	49
28	Passeio Aleatório	58
29	FAC e FACP para Passeio Aleatório	58
30	Janela de Valls para Passeio Aleatório	59
31	Passeio Aleatório com drift	60
32	FAC e FACP do Passeio Aleatório com drift	61
33	Janela de Valls para Passeio Aleatório sem drift	61
34	Diferenças Diversas	63

1 Processos Estocásticos

1.1 Introdução

Considere uma função $X : \mathbb{T} \times \Omega \longrightarrow \mathbb{R}$ onde \mathbb{T} representa um conjunto de índices que pode ser discreto ou contínuo e Ω um conjunto de resultados possíveis.

Observe que esta função está indexada por dois conjuntos. Podemos então fixar um deles fazendo que X seja função só do outro conjunto.

Por exemplo ao se fixar \mathbb{T} em t_n temos que $X(t_n, \omega)$ para $\omega \in \Omega$ é uma variável aleatória (v.a.) uma vez que para o tempo t_n ela pode assumir qualquer valor no conjunto de resultados possíveis desta v.a.

Por outro lado se fixarmos Ω em ω_k teremos uma função $X(t, \omega_k)$ para $t \in \mathbb{T}$ que é uma função do tempo para este dado valor de Ω .

Observe que mudando ω teremos outra função do tempo. Esta função é chamada de **processo estocástico** e para cada ω teremos uma **realização** deste processo, também chamada de **série temporal**.

Vamos considerar a seguinte função X

$$X(t, \omega) = Y(\omega) \cos[Z(\omega) + t * U(\omega)] \quad (1)$$

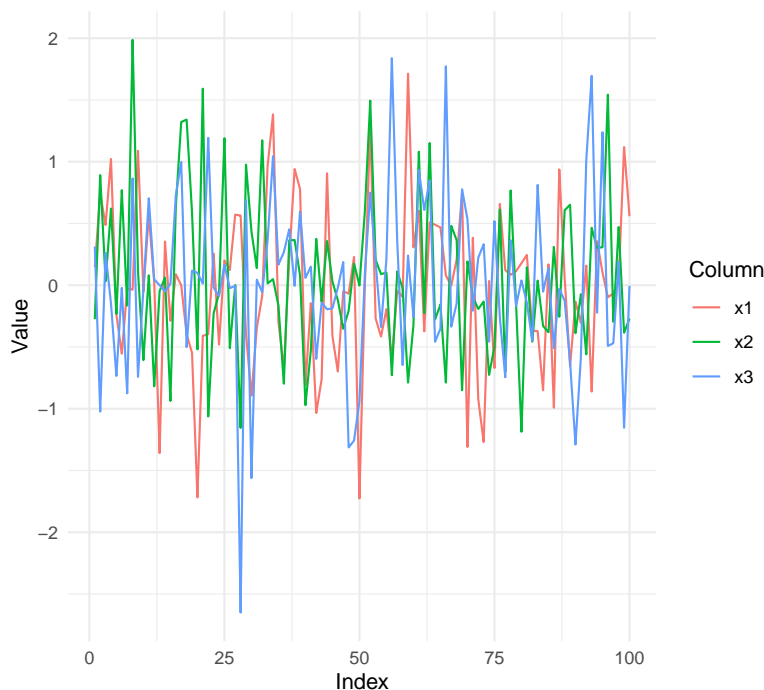
onde Y, Z e U são v.a. independentes entre si, por exemplo - Y tem distribuição normal com média zero e variância σ_y^2 - $Y \sim N(0, \sigma_y^2)$ - , Z tem distribuição normal com média zero e variância σ_z^2 - $Z \sim N(0, \sigma_z^2)$ - e U tem distribuição uniforme no intervalo $[-\pi, \pi]$ - $U \sim U[-\pi, \pi]$.

Para gerar cem possíveis realizações deste processo estocástico com uma amostra de tamanho 100, podemos usar o seguinte programa do R:

Stochastic Process

O gráfico abaixo apresenta as realizações **x1**, **x2** e **x3**, onde o tempo está variando e para **x1** foi fixando o valor de ω em ω_1 , para **x2** foi fixando o valor de ω em ω_2 e para **x3** foi fixando o valor de ω em ω_3 .

Figura 1: Três realizações de (1)

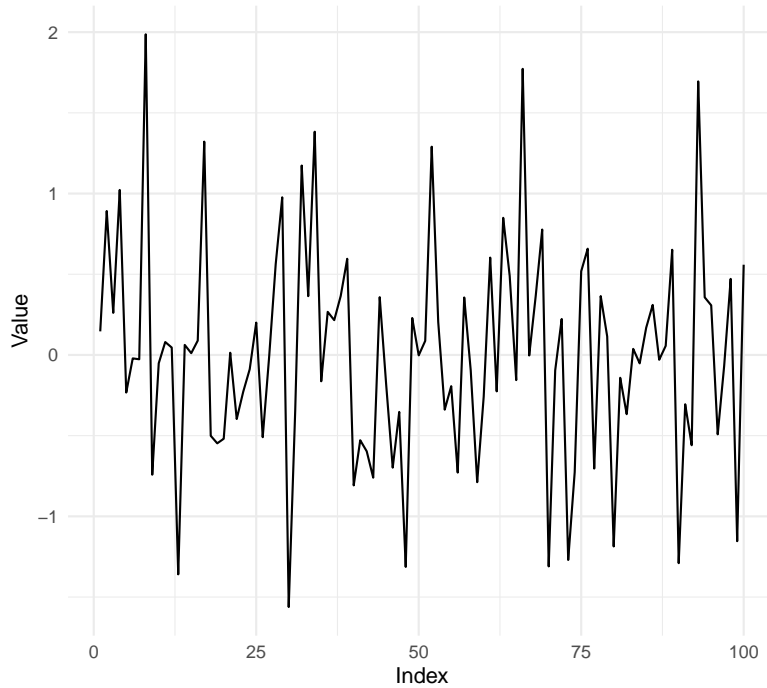


Observe que as três realizações não parecem cosenoides embora o processo gerador dos dados seja.

Na prática ao se observar uma série temporal, não sabemos se a realização deste processo é exatamente x_1 ou x_2 ou x_3 ou uma mistura destas três realizações. Por exemplo suponha que em cada instante de tempo temos um mecanismo aleatório de escolha entre as três realizações. Este mecanismo, embora não aleatório, pode ser escolha a primeira observação em x_1 , a segunda observação em x_2 , a terceira observação em x_3 , a quarta observação em x_1 , a quinta observação em x_2 , a seis observação em x_3 , etc. O programa para fazer esta escolha está abaixo:

Processo Estocástico Observado - Série Temporal Observada

Figura 2: Série Temporal Observada



A série temporal resultante tem a forma acima, e em geral é este tipo de série que é observada na prática e através dela desejamos recuperar todas as características do processo $X(t, \omega)$.

Observe que ao se definir o processo estocástico foi necessário introduzir três características, a saber:

- (i) o espaço onde está definido o processo, que no nosso caso é o dos números reais;
- (ii) o conjunto dos índices \mathbb{T} , que no nosso caso é o dos inteiros e, portanto, será chamado de processo a tempo discreto; e
- (iii) a estrutura de dependência da v.a., que no nosso caso é dada pelo processo gerador dos dados (1), isto é, X_t , para $t \in \mathbb{T}$.

Como o processo estocástico X_t , $t \in \mathbb{T}$, é uma v.a. estaremos interessados em obter a distribuição conjunta de (X_1, \dots, X_T) . Mas na prática só dispo-

mos de uma única realização do processo, isto é, o segundo gráfico, sendo portanto necessário introduzir certas restrições para que seja possível obter esta distribuição conjunta.

Por exemplo no instante $t = 1$ só observamos x_1 , o valor que a v.a. X_1 assume neste instante de tempo, e desejamos recuperar a distribuição da v.a. X_1 . Observe que esta distribuição depende de mais de um parâmetro e, portanto, é impossível recuperar esta distribuição com uma única observação.

No caso de processos estocásticos as restrições que devem ser impostas são de dois tipos:

- (i) na heterogeneidade temporal; e
- (ii) na memória do processo.

1.1.1 Exemplos

1. O programa abaixo escrito no R gerar 1000 replicações do modelo abaixo que é conhecido na literatura por $AR(1)$ e será apresentado na próxima seção:

$$X_t = 0.5 * X_{t-1} + u_t$$

onde $u_t \sim NI(0, 1)$ e a amostra tem tamanho 100.

Ao programa apresentado abaixo acrescente linha de comando de tal forma que para cada instante de tempo seja obtida a distribuição de frequência de X_t , na notação do programa são as séries tar. Acrescente também linhas de comando para construir gráficos para as diversas realizações do processo, que na notação do programa são as séries sar. Quais as diferenças entre as séries tar e sar?

Exemplo AR(1)

2. O programa abaixo escrito no R gerar 1000 replicações do modelo abaixo

$$X_t = u_t + 0.5 * u_{t-1}$$

onde $u_t \sim NI(0, 1)$ e a amostra tem tamanho 100.

Além disto, para cada instante obtenha a distribuição de frequência de X_t , na notação do programa são as séries tma. Construa os gráficos para as diversas realizações do processo, que na notação do programa são as séries sma. Quais as diferenças entre as séries tma e sma?

Exemplo MA(1)

2 Restrições na heterogeneidade temporal - Processos Estacionários

Primeiramente vamos apresentar os tipos de restrições que devem ser introduzidas na heterogeneidade do processo de tal sorte que seja possível recuperar a densidade de X_t para cada instante de tempo através de uma única realização do processo.

Seja dado um subconjunto do conjunto de índices \mathbb{T} , por exemplo, $\{t_1, \dots, t_n\}$. Uma restrição na heterogeneidade temporal é, por exemplo, assumir que a distribuição conjunta de $\{X_{t_1}, \dots, X_{t_n}\}$ seja invariante por translações, isto é, a distribuição conjunta de $\{X_{t_1}, \dots, X_{t_n}\}$ é a mesma que a distribuição conjunta de $\{X_{t_1+\tau}, \dots, X_{t_n+\tau}\}$ para qualquer valor de τ . Um processo que satisfaz esta restrição é chamado de **estritamente estacionário**.

Observe que esta restrição implica que tanto a **média** quanto a **variância** do processo são **invariantes no tempo** dado que se o subconjunto do conjunto de índices for t_1 temos que a distribuição de X_{t_1} tem que ser a mesma que a de $X_{t_1+\tau}$ para qualquer τ , logo em particular os dois primeiros momentos tem que ser iguais implicando que tanto a média quanto a variância do processo são invariantes no tempo.

Agora considere um subconjunto do conjunto de índices formado por $\{t_1, t_2\}$ podemos mostrar que além da média e variância serem invariantes no tempo, as **covariâncias dependem dos intervalos entre os tempo**.

Dizemos que o processo é estacionário de primeira ordem para o primeiro caso e **estacionário de segunda ordem** para o segundo caso.

Este tipo de restrição na heterogeneidade temporal é mais operacional dado que a do processo estritamente estacionário dado que é mais fácil verificar se os dois primeiros momentos do processos são independentes do tempo

e se as covariâncias dependem dos intervalos entre os tempos, do que mostrar a invariância por translação da distribuição conjunta. Deste modo este tipo de restrição, processos estacionários de segunda ordem, nos permite, com uma única realização do processo, recuperar alguns parâmetros da densidade de X_t . Observe também que se a distribuição subjacente a X_t for normal, o fato do processo ser estacionário de segunda ordem nos permite recuperar todos os parâmetros da distribuição de X_t dado que média, e a matriz de variância e covariância caracterizam totalmente a distribuição normal.

Em várias situações os momentos da distribuição dependem das condições iniciais do processo. Nestes casos o processo só se torna estacionário assintoticamente.

A seguir vamos apresentar restrições na memória do processo, isto é, na dependência temporal do processo.

2.1 Exercício

1. Mostre que se o processo for:

- (a) estritamente estacionário e
- (b) se o subconjunto de índice só contiver um elemento.

então a média e variância do processo são invariantes no tempo.

2. Mostre que se o processo for:

- (a) estritamente estacionário e
- (b) se o subconjunto de índices contiver dois elementos.

então a média e variância do processo são invariantes no tempo e as covariâncias dependem dos intervalos entre os tempos.

3. Mostre que se o processo é $X_t = u_t$ com $u_t \sim NI(0, \sigma^2)$ ele é estacionário de segunda ordem.

3 Restrições na memória do processo - assintoticamente independente e m-dependente.

3.1 Introdução

A primeira idéia é fazer com que o processo não tenha memória, isto é, seja não correlacionados ou independente. No caso em que tanto a média do processo quanto a variância são finitas temos um processo que é chamado de **ruído branco**. Este processo é definido por:

Definição 1 *O processo u_t que satisfaz:*

- (i) $E(u_t) = m < \infty$
- (ii) $Var(u_t) = \sigma^2 < \infty$
- (iii) $Cov(u_t, u_{t-j}) = 0$ para todo $j \neq 0$

*é chamado de **ruído branco**.*

Observe que nesta definição a condição (iii) implica que o processo é não correlacionado mas não necessariamente independente. No caso de processos gaussianos teremos que a condição (iii) implica em independência.

A restrição de independência é muito forte e pouco plausível dado que séries temporais econômicas apresentam alguns tipo de dependência temporal. Uma forma de abrandar esta hipótese é fazer com que para instante de tempo muito afastados não exista autocorrelação, que no caso gaussiano corresponde a independência, isto é, que o processo seja **assintoticamente independente** (não correlacionado). Deste modo o processo tem memória, mas esta memória vai diminuindo com o aumento dos intervalos entre os instantes de tempo. Um caso particular deste tipo de processo é aquele em que a dependência existe quando se considera, por exemplo, m observações consecutivas, mas que a partir da observação $m+1$ não exista correlação com as m observações anteriores. Este processo é chamado de **m-dependente**.

Vamos apresentar dois modelos que tem memória mas que para intervalos de tempo distantes podem ser considerados independentes.

3.2 Processo Assintoticamente Independente

No primeiro exemplo vamos considerar um processo gaussiano, isto é, normal, em que a média do processo é linear no passado mais recente, isto é, o conjunto de informação que estamos usando é $\mathbf{I}_t = \{X_t\}$, e a variância do processo é constante, isto é:

1. $E(X_t | \mathbf{I}_{t-1}) = E(X_t | X_{t-1}) = \phi X_{t-1}$
2. $Var(X_t | \mathbf{I}_{t-1}) = E([(X_t - \phi X_{t-1})^2 | X_{t-1}]) = \sigma^2$

Este tipo de modelo pode ser racionalizado da seguinte forma, se X_t representa uma série de inflação, o processo gerador da inflação é tal que, dado o passado, é esperado que a inflação hoje seja a inflação passada mais choques não previsível e que a inflação passada está sujeita a um redutor porque o parâmetro ϕ , por hipótese, é menor do que um. Deste modo, temos que a inflação hoje, a menos de choques não previsíveis, deverá ser menor do que a inflação do instante anterior.

Podemos então representar o processo gerador por¹:

$$X_t = \phi X_{t-1} + u_t \quad (2)$$

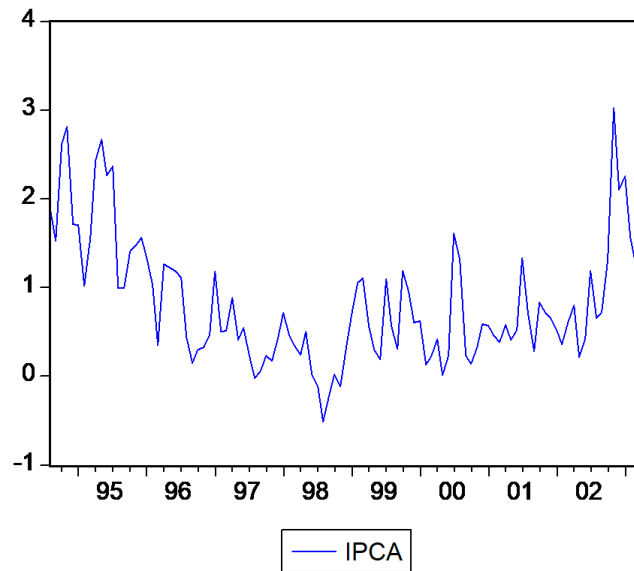
onde u_t representa os desvios de X_t de sua média condicional e porque o processo é gaussiano estes desvios também são gaussianos e vamos mostrar que eles tem média zero, variância constante e igual a σ^2 e covariância igual a zero, logo são independentes.

3.2.1 Inflação no Brasil medida pelo IPCA

O gráfico abaixo apresenta a inflação medida pelo IPCA para o período de 1994:08 até 2003:04.

¹Implicitamente estamos assumindo que o processo X_t tem média não condicional igual a zero. Isto não é restritivo uma vez que se X_t tiver média m podemos definir $X_t^* = X_t - m$ e a especificação(3) é válida para X_t^* . Observe que o processo gerador de dados para X_t^* é dado por $X_t^* = \phi X_{t-1}^* + u_t \Leftrightarrow (X_t - m) = \phi(X_{t-1} - m) + u_t \Rightarrow X_t = m(1 - \phi) + \phi X_{t-1} + u_t$, implicando que $E(X_t) = m$.

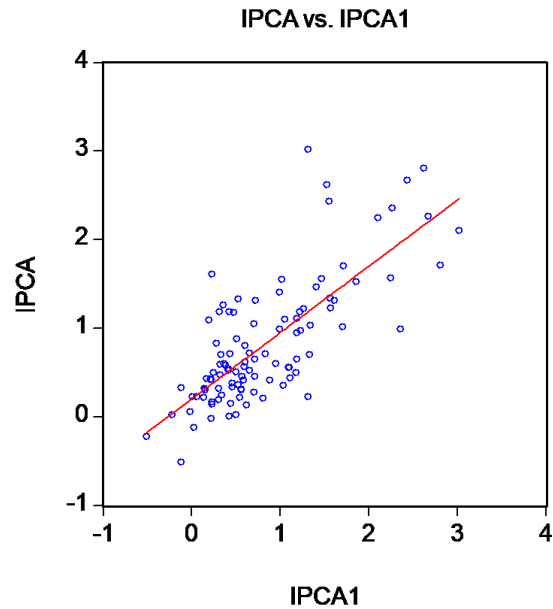
Figura 3: Inflação medida pelo IPCA



Observe que o gráfico contra o tempo nos dá uma idéia da evolução da série de inflação. Podemos observar que a série tem uma tendência declinante até meados de 98, passando a oscilar num patamar perto de 0,5 % e no final de 2002 a série apresenta um crescimento.

Uma forma de verificar se a inflação está relacionada a inflação no período anterior é através do diagrama de dispersão de $IPCA(-1)$ contra $IPCA$. Este gráfico, para o período de 1994:09 até 2003:04 é apresentado abaixo.

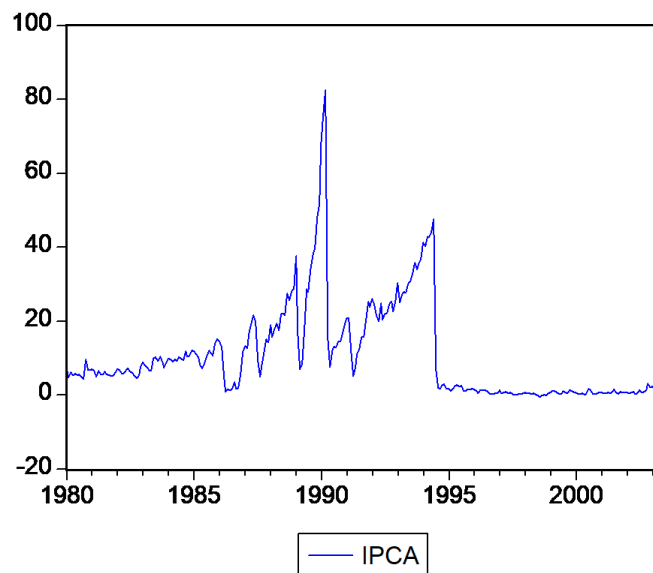
Figura 4: Diagrama de Dispersão $\text{IPCA}(-1) \times \text{IPCA}$



Observe que existe uma associação positiva entre as duas variáveis com correlação de 0,75 que é um estimador para ϕ em (3). Observe que neste período 75% da inflação passada é repassada para a inflação presente.

Agora se estivermos considerando o IPCA para o período de 1980:01 até 2003:04 o gráfico contra o tempo é apresentado abaixo.

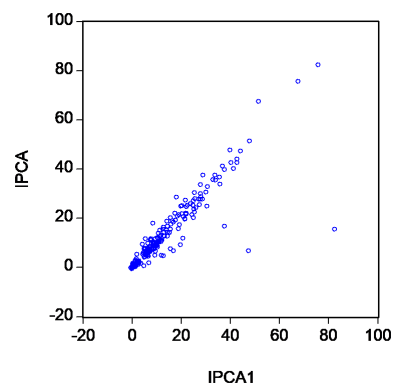
Figura 5: Inflação medida pelo IPCA - 80:01 - 03:04



Observe que em diversos períodos a inflação apresenta tendência, sendo que em alguns períodos este comportamento é explosivo.

O gráfico abaixo apresenta o diagrama de dispersão para este período.

Figura 6: Diagrama de Dispersão para o período de 80:02 até 03:04



e neste caso a correlação é de 0,92 que é próxima de um, implicando que a inflação passada e totalmente repassada a inflação corrente.

3.2.2 O erro é ruído branco

Para mostrar que a média não condicional de u_t é zero, vamos mostrar primeiro que a média condicional é zero.

$$\begin{aligned} E(u_t \mid X_{t-1}) &= E(X_t - \phi X_{t-1} \mid X_{t-1}) = E(X_t \mid X_{t-1}) - \phi E(X_{t-1} \mid X_{t-1}) = \\ &= \phi X_{t-1} - \phi X_{t-1} \\ &= 0 \end{aligned} \tag{3}$$

onde foi usado que $E(X_t \mid X_{t-1}) = \phi X_{t-1}$ e $E(X_{t-1} \mid X_{t-1}) = X_{t-1}$. Mas com $E(E(Y \mid Z)) = E(Y)$, temos que $E(u_t) = E(E(u_t \mid X_{t-1})) = 0$.

Vamos mostrar que $Var(u_t) = \sigma^2$

$$\begin{aligned} Var(u_t) &= E(u_t^2) = E(X_t - \phi X_{t-1})^2 = E(E[(X_t - \phi X_{t-1})^2 \mid X_{t-1}]) \\ &= E(Var(X_t \mid X_{t-1})) = E(\sigma^2) = \sigma^2 \end{aligned}$$

Vamos mostrar que as covariâncias são nulas e vamos assumir que $t > s$:

$$\begin{aligned} Cov(u_t, u_s) &= E(u_t u_s) = E(E[u_t u_s \mid X_{t-1}]) \\ &= E(u_s E[u_t \mid X_{t-1}]) \\ &= E(u_s \cdot 0) \\ &= 0 \end{aligned}$$

3.2.3 Momentos não Condicionais e Autocorrelação do processo

X_t

Agora, vamos mostrar que a média, não condicional, de X_t é assintoticamente igual a zero. Para mostrar isto, vamos assumir que X_0 é fixo e podemos substituir para trás no tempo em (2) e temos a seguinte expressão:

$$X_t = \phi^t X_0 + \sum_{j=0}^{t-1} \phi^j u_{t-j} \tag{4}$$

e tomando-se o operador esperança em (4) temos:

$$\begin{aligned} E(X_t) &= \phi^t E(X_0) + \sum_{j=0}^{t-1} \phi^j E(u_{t-j}) \\ &= \phi^t X_0 + \sum_{j=0}^{t-1} \phi^j \cdot 0 \\ &= \phi^t X_0 \\ \Rightarrow E(X_t) &= \phi^t X_0 \end{aligned}$$

mas a média do processo será invariante no tempo sómente assintoticamente e também se $|\phi| < 1$. Observe, também, que as covariâncias não condicionais podem ser escritas da seguinte forma:

$$\begin{aligned}
Cov(X_t, X_{t-s}) &= \gamma_s = E[(X_t - \phi^t X_0)(X_{t-s} - \phi^{t-s} X_0)] \\
&= E \left[\left(\sum_{j=0}^{t-1} \phi^j u_{t-j} \right) \left(\sum_{k=0}^{t-s-1} \phi^k u_{t-s-k} \right) \right] \\
&= E \left[\left(\sum_{j=0}^{s-1} \phi^j u_{t-j} + \sum_{j=s}^{t-1} \phi^j u_{t-j} \right) \left(\sum_{k=0}^{t-s-1} \phi^k u_{t-s-k} \right) \right] \\
&= E \left[\left(\sum_{j=0}^{s-1} \phi^j u_{t-j} + \sum_{k=0}^{t-s-1} \phi^{s+k} u_{t-s-j} \right) \left(\sum_{k=0}^{t-s-1} \phi^k u_{t-s-k} \right) \right] \\
&= E \left[\left(\sum_{j=0}^{s-1} \phi^j u_{t-j} \right) \left(\sum_{k=0}^{t-s-1} \phi^k u_{t-s-k} \right) \right] \\
&\quad + E \left[\left(\sum_{k=0}^{t-s-1} \phi^{s+k} u_{t-s-j} \right) \left(\sum_{k=0}^{t-s-1} \phi^k u_{t-s-k} \right) \right] \tag{5}
\end{aligned}$$

$$\begin{aligned}
&= E \left[\left(\sum_{j=0}^{s-1} \sum_{k=0}^{t-s-1} \phi^j \phi^k u_{t-j} u_{t-s-k} \right) \right] \\
&\quad + E \left[\left(\phi^s \sum_{k=0}^{t-s-1} \phi^{2k} u_{t-s-j}^2 \right) \right] \tag{6}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j=0}^{s-1} \sum_{k=0}^{t-s-1} \phi^j \phi^k E(u_{t-j} u_{t-s-k}) + \phi^s \sum_{k=0}^{t-s-1} \phi^{2k} E(u_{t-s-j}^2) \\
&= \sum_{j=0}^{s-1} \sum_{k=0}^{t-s-1} \phi^j \phi^k \cdot 0 + \phi^s \sum_{k=0}^{t-s-1} \phi^{2k} \cdot \sigma^2 \\
&\Rightarrow Cov(X_t, X_{t-s}) = \gamma_s = \phi^s \sigma^2 \frac{1 - \phi^{2(t-s)}}{1 - \phi^2} \tag{7}
\end{aligned}$$

e portanto a variância do processo é dada por:

$$Var(X_t) = \gamma_0 = \sigma^2 \frac{1 - \phi^{2t}}{1 - \phi^2} \tag{8}$$

Deste modo as autocorrelações deste processo serão dadas por:

$$\begin{aligned}
Corr(X_t, X_{t-s}) &= \rho_s = \frac{Cov(X_t, X_{t-s})}{\sqrt{Var(X_t)Var(X_{t-s})}} = \frac{\phi^s \sigma^2 \frac{1-\phi^{2(t-s)}}{1-\phi^2}}{\sqrt{\sigma^2 \frac{1-\phi^{2t}}{1-\phi^2} \sigma^2 \frac{1-\phi^{2(t-s)}}{1-\phi^2}}} \\
&= \frac{\phi^s (1 - \phi^{2(t-s)})}{\sqrt{(1 - \phi^{2t}) (1 - \phi^{2(t-s)})}} = \frac{\phi^s}{\sqrt{\frac{(1-\phi^{2t})}{(1-\phi^{2(t-s)})}}} \leq \phi^s
\end{aligned}$$

e observe que se a sequência $\{\phi^s ; s \geq 0\}$ que majora as correlações satisfaz as seguintes condições:

(i) $0 \leq \phi < 1$; e

(ii) $\sum_{s=0}^{\infty} \phi^s < \infty$

então dizemos que o processo é assintoticamente não correlacionado. Como o processo é gaussiano teremos que este também será assintoticamente independente.

Para que (7) dependa somente dos intervalos entre os tempos devemos ter que $\phi^{t-s} \approx 0$ e deste modo teremos que (7) será dado por:

$$Cov(X_t, X_{t-s}) = \gamma_s = \frac{\phi^s \sigma^2}{1 - \phi^2} \quad (9)$$

O modelo (2) é conhecido na literatura por modelo **autorregressivo** de primeira ordem e é denotado por $AR(1)$. Observe que o fato das covariâncias terem a expressão acima, implica que as autocorrelações serão dadas por ϕ^s e, portanto, irão decair exponencialmente.

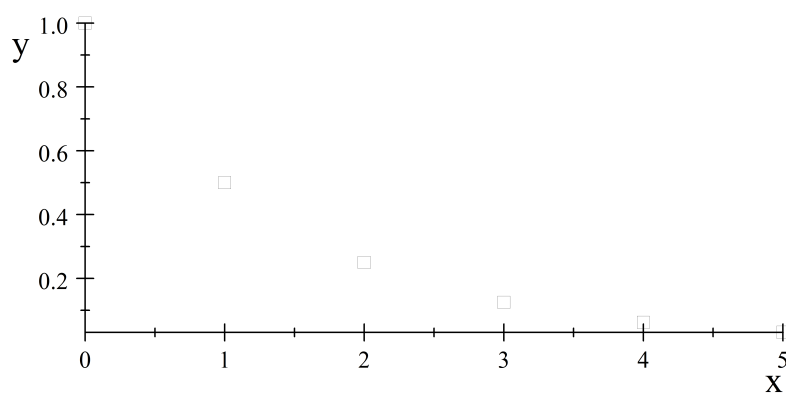
Observe a diferença entre a distribuição não condicional de X_t e a distribuição de X_t condicional na informação passada. No caso da distribuição não condicional se o início dos tempos é no passado distante e se $|\phi| < 1$ temos que a média é zero a variância é constante e igual a $\frac{\sigma^2}{1-\phi^2}$ e as autocorrelações são dadas ϕ^s sendo portanto estacionário em covariância. Mas para a distribuição condicional temos que a média condicional é dada por ϕX_{t-1} a variância é constante e igual a σ^2 e são não correlacionados uma vez que $X_t - E(X_t | \mathbf{I}_{t-1}) = u_t$.

3.2.4 Correlograma para o processo AR(1)

O gráfico das autocorrelações contra as defasagens é chamado de **correlograma** e pode ser usado para identificar este processo. Abaixo são apresentados alguns exemplo.

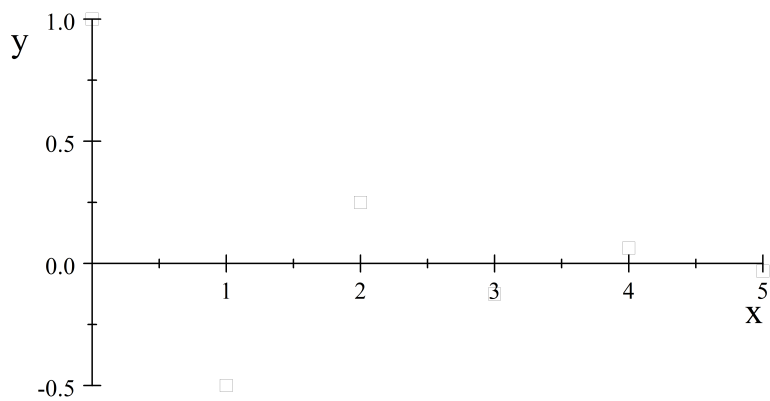
O primeiro caso, quando $\phi = 0.5$, o decaimento é exponencial e rápido

Figura 7: Correlograma para AR(1) com $\phi = 0.5$



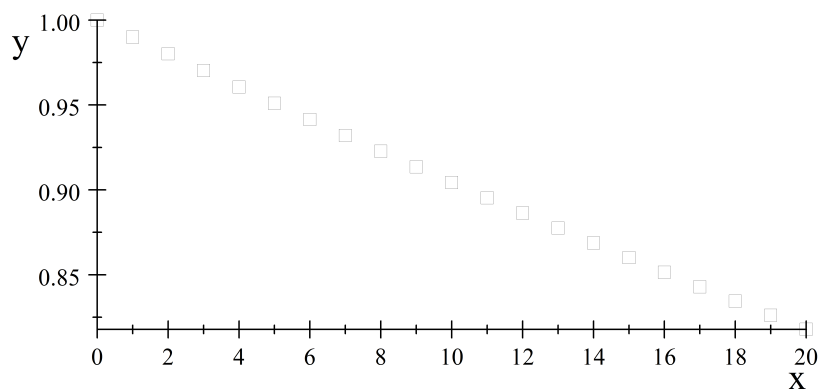
Para o segundo caso, como $\phi = -0.5$, o decaimento é rápido mas alternando entre positivos e negativos

Figura 8: Correlograma para AR(1) com $\phi = -0.5$



Para o terceiro caso como $\phi = 0.99$, o decaimento é muito lento uma vez que depois de 20 defasagens a autocorrelação é ainda superior a 0.8.

Figura 9: Correlograma para AR(1) com $\phi = 0.99$

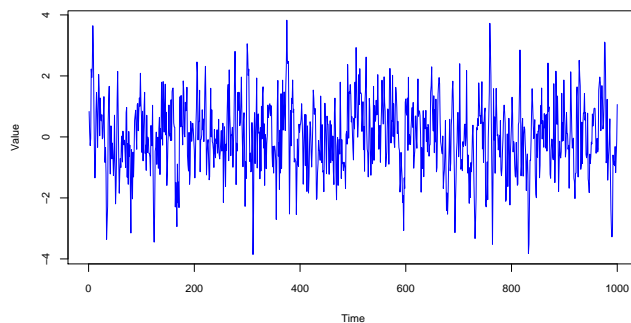


3.2.5 Gerando processos AR(1) no R

Para se gerar um AR(1) com $\phi = 0.5$, e fazer o gráfico da série no R podemos usar o seguinte programa

AR(1) com $\phi = 0.5$

Figura 10: AR(1) com $\phi = 0.5$

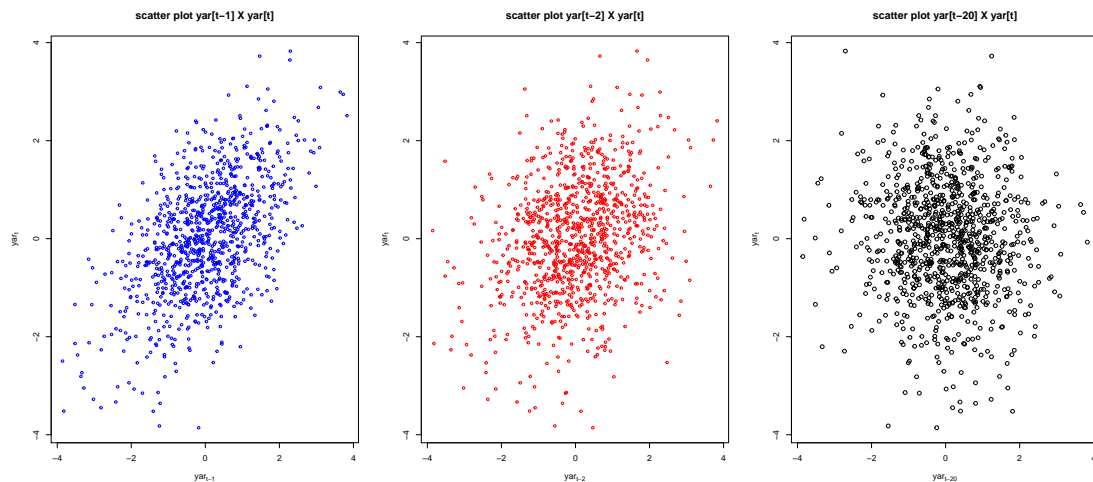


Observe que o processo cruza frequentemente o zero que é a média do processo e que a maioria dos pontos estão entre -2.67 e 2.67 que corresponde a dois desvios padrões.

Uma forma de verificar o grau de dependência desta série é fazer o diagrama de dispersão da série contra a série defasada um, dois,..., n períodos e verificar se a associação vai desaparecendo com o aumento do defasagem usada. Vamos fazer isto para $n = 1, 2, 20$. O comando no R para definir o operador defasagem é `lag(x, k = -n)` que defasa a série n períodos e o comando para definir o diagrama de dispersão é `plot(x1, x)` e queremos multiplos gráficos sendo sempre a primeira variável da lista contra as outras.

Diagrama de dispersão das defasagens da série contra a própria série

Figura 11: Diagrama de dispersão da série e defasagens



Observe que no gráfico da esquerda existe uma associação entre as duas variáveis e esta associação ainda está presente no gráfico do meio mas parece menor e no gráfico da direita esta associação desaparece. Logo para esta série o grau de associação vai diminuindo com o aumento dos intervalos entre os tempos.

3.2.6 Exercícios

1. Escreva um programa no R para gerar um $AR(1)$ com $\phi = .99$ e uma amostra de tamanho 1000. Como voce pode verificar se este processo é assintoticamente independente ou não?
2. E se no ítem exercício anterior $\phi = 1.01$, qual seria a sua conclusão sobre independência assintótica.
3. Para o modelo gerado no exercício 1, construa um programa para gerar a função de autocorrelação para até 20 defasagens. Esta função é parecida com o gráfico 9? Porque? Quais as diferenças que voce observa?

3.3 Processo 1-dependente

A seguir apresentaremos o exemplo de um modelo 1-dependente.

Considere agora, o processo que só depende dos choques de hoje e ontem. Este modelo pode ser racionalizado como sendo o preço de um produto que é afetado somente por eventos externos ao mercado, por exemplo, fatores climáticos. Neste tipo de modelo o processo gerador dos dados pode ser escrito da seguinte forma:

$$X_t = u_t + \theta u_{t-1} \quad (10)$$

onde u_t são os choques externos que serão v.a. gaussianas independentes com média zero e variância constante σ^2 . Na literatura este modelo é chamado de **médias-móveis** de primeira ordem e denotado por $MA(1)$. É fácil mostrar que X_t tem média zero e que a variância é dada por $(1 + \theta^2) \sigma^2$. Para mostrar que este processo é 1-dependente falta mostrar que as covariâncias para intervalos maiores do que um são zero. Considere então a covariância de primeira ordem, isto é, $Cov(X_t, X_{t-1})$ que é dada por:

$$\begin{aligned} Cov(X_t, X_{t-1}) &= E[(u_t + \theta u_{t-1})(u_{t-1} + \theta u_{t-2})] \\ &= E(u_t u_{t-1} + \theta u_t u_{t-2} + \theta u_{t-1} u_{t-1} + \theta^2 u_{t-1} u_{t-2}) \\ &= E(u_t u_{t-1}) + \theta E(u_t u_{t-2}) + \theta E(u_{t-1} u_{t-1}) + \theta^2 E(u_{t-1} u_{t-2}) \end{aligned}$$

e como os u 's são não correlacionados a única contribuição não nula para a expressão acima é a terceira que corresponde a variância do processo. Tem-se então:

$$Cov(X_t, X_{t-1}) = \gamma_1 = \theta \sigma^2 \quad (11)$$

e a primeira autocorrelação é dada por:

$$Corr(X_t, X_{t-1}) = \rho_1 = \frac{\theta}{(1 + \theta^2)}$$

Considere agora a covariância de segunda ordem, isto é, $Cov(X_t, X_{t-2})$, que é dada por:

$$\begin{aligned} Cov(X_t, X_{t-2}) &= \gamma_2 = E[(u_t + \theta u_{t-1})(u_{t-2} + \theta u_{t-3})] \\ &= E(u_t u_{t-2} + \theta u_t u_{t-3} + \theta u_{t-1} u_{t-2} + \theta^2 u_{t-1} u_{t-3}) \\ &= E(u_t u_{t-2}) + \theta E(u_t u_{t-3}) + \theta E(u_{t-1} u_{t-2}) + \theta^2 E(u_{t-1} u_{t-3}) \end{aligned}$$

e como os u 's são não correlacionados todas as contribuições são nula o que implica que a covariância é zero e como o processo é gaussiano, temos que são independentes. É fácil mostrar que todas as covariâncias de todas as ordens superiores, incluindo, a segunda são zero, porque X_t depende de $\{u_t, u_{t-1}\}$ enquanto que X_{t-j} depende de $\{u_{t-j}, u_{t-j-1}\}$ e como $j \geq 2$ todos os produtos terão esperança nula. Tem-se então que o processo MA(1) é sempre estacionário de segunda ordem.

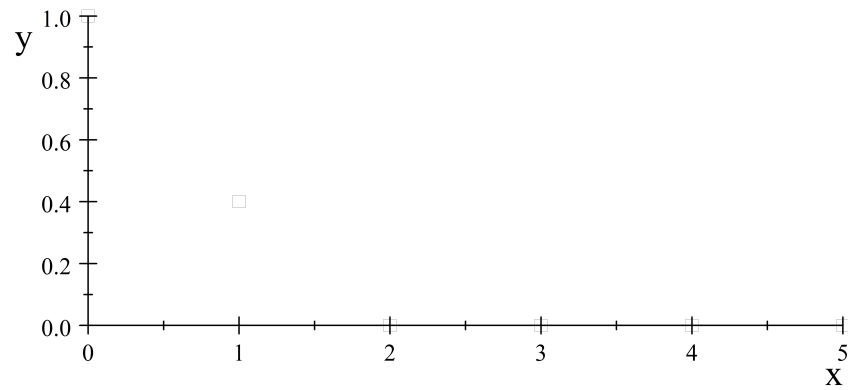
3.3.1 Correlograma para o MA(1)

Observe que o correlograma neste é caso tem sómente a primeira autocorrelação diferente de zero e as de ordem superiores são todas nula e este padrão das autocorrelações pode ser usado para identificar o processo.

Abaixo são apresentados alguns exemplo

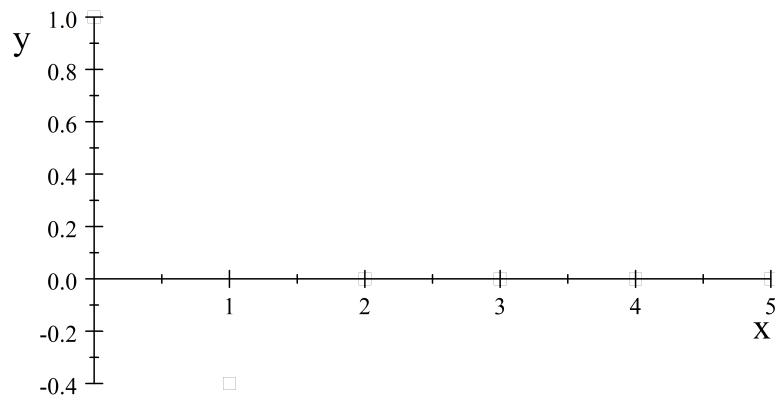
Para o primeiro caso foi considera aquele em que $\theta = 0.5$, temos então o seguinte correlograma

Figura 12: Correlograma para MA(1) com $\theta = 0.5$



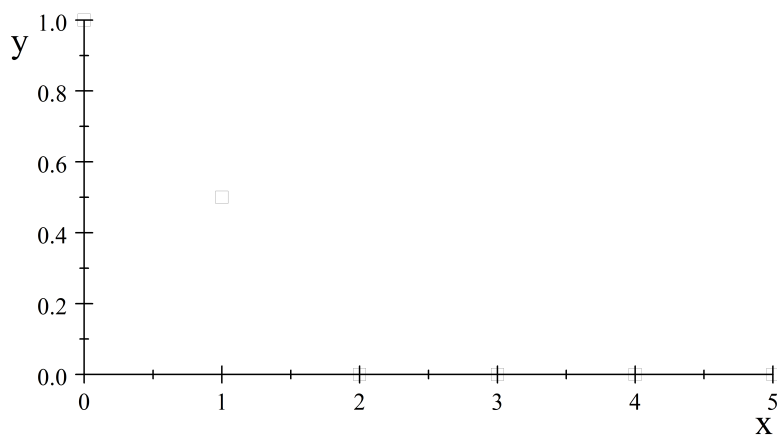
e para o segundo caso foi considerado $\theta = -0.5$, temos então o seguinte correlograma

Figura 13: Correlograma para MA(1) com $\theta = -0.5$



e para o terceiro caso $\theta = 0.99$, temos então o seguinte correlograma

Figura 14: Correlograma para MA(1) com $\theta = 0.99$



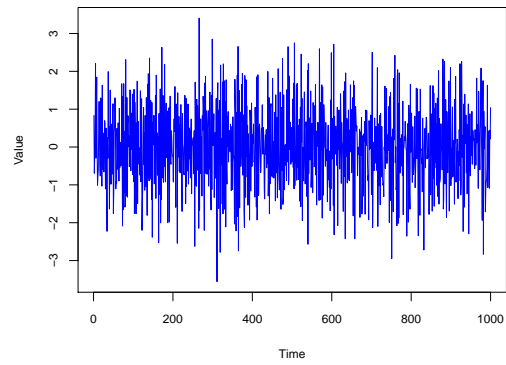
Observe que em todos os três casos temos o mesmo padrão, primeira autocorrelação diferente de zero e todas as outras iguais a zero, mesmo para o caso em que θ é próximo de um.

3.3.2 Gerando processos MA(1) no R

Para se gerar um MA(1) com $\theta = 0.5$, e fazer o gráfico da série no R podemos usar o seguinte programa

```
MA(1)
```

Figura 15: MA(1) com $\theta = 0.5$

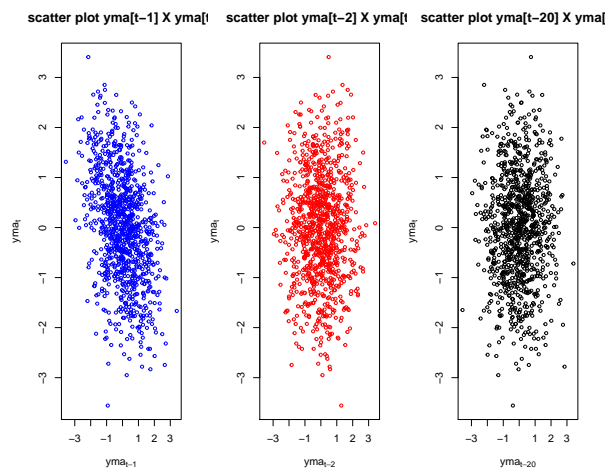


Pelo diagrama de dispersão da série contra o seu passado é possível verificar que o grau de associação desta série é de somente um período. Incluindo os comandos

Diagrama de dispersão

no programa acima temos os seguintes gráficos

Figura 16: MA(1) com $\theta = 0.5$



Observe que neste caso somente o gráfico da esquerda apresenta associação entre as variáveis e os outros dois gráficos as séries parecem não ter associação entre si.

3.3.3 Exercícios

1. Escreva um programa no R para gerar um $MA(1)$ com $\theta = 0,99$ e uma amostra de tamanho 1000. Como voce pode verificar se este processo é 1-dependente ou não?
2. E se no item exercício anterior $\theta = 1.01$, qual seria a sua conclusão sobre dependência temporal do processo.
3. Para o modelo gerado no exercício 1, construa um programa para gerar a função de autocorrelação para até 20 defasagens. Esta função é parecida com o gráfico 14? Porque? Quais as diferenças que voce observa?

4 Processos Autorregressivos de Ordem Superior

4.1 Introdução

O processo autorregressivo pode ser generalizado para ordens superiores a 1 e o caso geral corresponde ao modelo $AR(p)$ onde o processo depende das defasagens até a ordem p , isto é, o processo gerador dos dados é dado por:

$$X_t = E(X_t \mid X_{t-1}, \dots, X_{t-p}) + u_t$$

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + u_t \quad (12)$$

onde os parâmetros ϕ 's terão que satisfazer certas condições para garantir independência assintótica do processo. Esta condição é dada por: se ϕ_j^* para $j = 1, \dots, p$ são as raízes do polinômio $1 - \phi_1 L + \dots + \phi_p L^p$, onde $L^j X_t = X_{t-j}$, então devemos ter $|\phi_j^*| < 1$. Esta condição pode ser mostrada reescrevendo

(12) em forma vetorial, isto é, defina $\mathbf{Z}_t = [X_{t-1}, \dots, X_{t-p}]'$ então (12) pode ser reescrito da seguinte forma:

$$\mathbf{Z}_t = \Phi \mathbf{Z}_{t-1} + \varepsilon_t \quad (13)$$

onde a matriz Φ é definida por:

$$\Phi = \begin{bmatrix} \phi_1 & \phi_2 & \cdot & \cdot & \cdot & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 1 & 0 \end{bmatrix}$$

e $\varepsilon_t = (u_t, 0, \dots, 0)$.

Assumindo que a condição inicial, isto é, \mathbf{Z}_0 , é fixa, para que a média do processo (13) seja independente do tempo é necessário que a matriz Φ seja estável, isto é, quando t cresce ela converge para a matriz zero. A condição para que isto aconteça é que os autovalores da matriz Φ estejam dentro do círculo unitário. É fácil mostrar que os autovalores desta matriz são as raízes do polinômio $\Phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$.

4.2 Processo AR(2) - Condições de Estacionaridade e Função de Autocorrelação

Vamos considerar o caso de $p = 2$, isto é, um AR(2). O processo gerador de dados é dado por:

$$X_t = m^* + \phi_1 X_{t-1} + \phi_2 X_{t-2} + u_t \quad (14)$$

onde u_t tem média zero variância constante e igual a σ^2 e são não correlacionados.

Seja $m = m^*(1 - \phi_1 - \phi_2)$ e fazendo $X_t^* = X_t - m$, (14) pode ser escrito da seguinte forma:

$$X_t^* = \phi_1 X_{t-1}^* + \phi_2 X_{t-2}^* + u_t \quad (15)$$

Para se obter as condições de estacionaridade para este processo, multiplique (15) por X_t^* e aplicando o operador esperança em ambos os lados obtém-se:

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + E(X_t^*u_t) \quad (16)$$

Mas por (15) $E(X_t^*u_t) = \sigma^2$ e portanto (16) pode ser escrito da seguinte forma:

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2 \quad (17)$$

E multiplicando-se (15) por X_{t-1}^* e X_{t-2}^* e aplicando o operador esperança em ambos os lados obtém-se:

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 \quad (18)$$

$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0 \quad (19)$$

A solução do sistema (17-19) é dada por:

$$\gamma_0 = \frac{(1 - \phi_2)\sigma^2}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)} \quad (20)$$

$$\gamma_1 = \frac{\phi_1\sigma^2}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)} \quad (21)$$

$$\gamma_2 = \frac{(\phi_2 + \phi_1^2 - \phi_2^2)\sigma^2}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)} \quad (22)$$

e como γ_0 tem que ser positivo e finito as seguintes condições tem que ser satisfeitas:

$$|\phi_2| < 1 \quad (23)$$

$$\phi_1 + \phi_2 < 1 \quad (24)$$

$$\phi_2 - \phi_1 < 1 \quad (25)$$

Pode-se mostrar que as condições (23-25) são necessárias para que o processo seja estacionário.²

As equações 18 e 19 podem ser reescritas da seguinte forma:

$$\rho_1 = \phi_1 + \phi_2 \rho_1 \quad (26)$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2 \quad (27)$$

que são chamadas de **equações de Yule-Walker** cuja solução dá as duas primeiras autocorrelações e que são dadas por:

$$\rho_1 = \frac{\phi_1}{(1 - \phi_2)} \quad (28)$$

$$\rho_2 = \frac{\phi_1^2}{(1 - \phi_2)} + \phi_2 \quad (29)$$

e a função de autocorrelação para defasagens superiores a dois é dada pela seguinte recursão

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad (30)$$

A função de autocorrelação dada em (30) satisfaz a seguinte equação de diferença de segunda ordem:

$$(1 - \phi_1 L - \phi_2 L^2) \rho_k = 0 \quad (31)$$

A solução de (31), fazendo $L = x$ é dada por:

$$x = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2} \quad (32)$$

Se ω_1 e ω_2 são as duas soluções reais o $AR(2)$ pode ser escrito como $(1 - \omega_1 L)(1 - \omega_2 L)$ e a função de autocorrelação é uma mistura de dois decaimentos exponenciais.

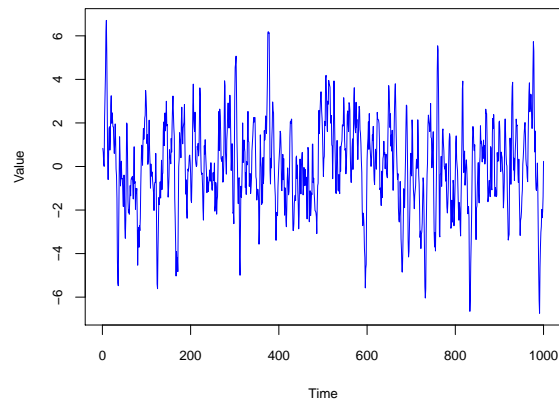
²Estas condições implicam que as raízes do polinômio $1 - \phi_1 L - \phi_2 L^2$ estão fora do círculo unitário, ou se considerarmos o polinômio $z^2 - \phi_1 z - \phi_2$ as raízes estão dentro do círculo unitário. Iremos usar esta "raízes dentro do círculo" mesmo quando estivermos usando a primeira especificação do polinômio.

Se $\phi_1^2 + 4\phi_2 < 0$ as duas raízes são complexas e a função de autocorrelação será um coseno ou seno amortecido.

Simulando um $AR(2)$ com as raízes reais no R temos

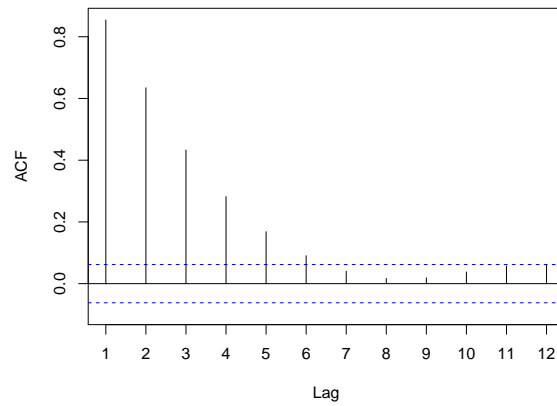
$AR(2)$ com raízes reais

Figura 17: $AR(2)$ com raízes reais



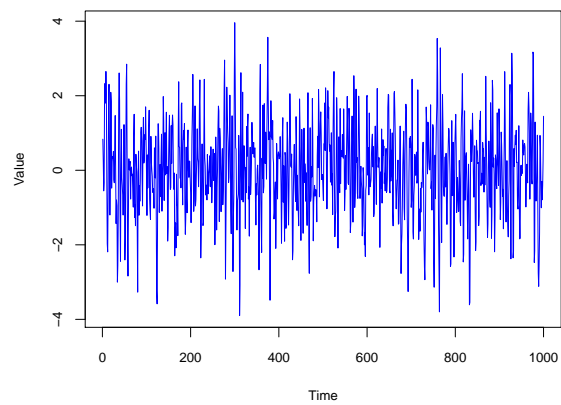
A Função de Autocorrelação para o $AR(2)$ com raízes reais é dada por:

Figura 18: Função de Autocorrelação para $AR(2)$ com raízes reais



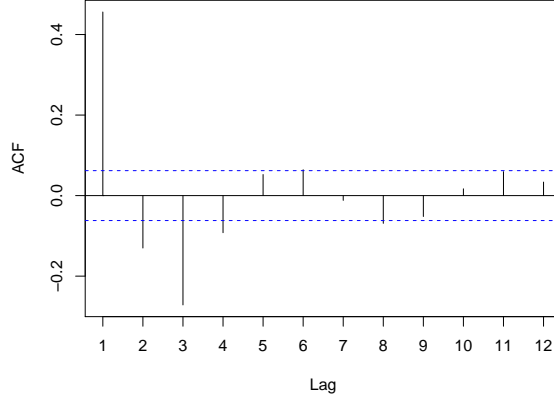
Simulando um $AR(2)$ com raízes complexas no R temos
 $AR(2)$ com raízes complexas

Figura 19: $AR(2)$ com raízes complexas



E a FAC do $AR(2)$ com raízes complexas é dada por

Figura 20: Função de Autocorrelação para AR(2) com raízes complexas



Observe que a FAC neste caso tem um comportamento senoidal. Desta forma dizemos que o $AR(2)$ com raízes complexas é pseudo-periódico.

4.3 Função de Autocorrelação Parcial para o processo AR(2)

Numa regressão simples o parâmetro de inclinação corresponde a correlação entre a variável dependente e a independente multiplicada pela razão entre o desvio padrão da variável dependente e do desvio padrão da variável independente. Numa regressão múltipla podemos interpretar cada um dos coeficientes de forma análogo, a saber, o parâmetro corresponde a correlação parcial entre a variável dependente e a variável independente que corresponde aquele a associação entre a variável dependente e aquela variável independente mantendo todas as outras variáveis independentes constante e também multiplicada pela razão dos desvios padrões mantendo as outras variáveis independentes constante.

Como o processo $AR(2)$ é uma regressão de X_t em X_{t-1} e X_{t-2} podemos, por exemplo, interpretar o coeficiente ϕ_1 como a correlação parcial de X_t e X_{t-1} mantendo X_{t-2} constante.

Sabemos que esta correlação é dada por:

$$r_{01.2} = \frac{r_{01} - r_{02}r_{21}}{1 - r_{12}^2} \quad (33)$$

onde 0, 1, 2 representam X_t , X_{t-1} e X_{t-2} respectivamente.

Mas como $r_{01} = r_{12} = \rho_1$ e $r_{02} = \rho_2$ substituindo-se em (33) temos:

$$r_{01.2} = \phi_{1,1} = \frac{\rho_1 - \rho_1\rho_2}{1 - \rho_1^2} \quad (34)$$

E o coeficiente ϕ_2 pode ser interpretado como a correlação parcial de X_t e X_{t-2} mantendo X_{t-1} constante.

Sabemos que esta correlação é dada por:

$$r_{02.1} = \phi_{2,2} = \frac{r_{02} - r_{01}r_{12}}{1 - r_{12}^2} \quad (35)$$

Mas como $r_{01} = r_{12} = \rho_1$ e $r_{02} = \rho_2$ substituindo-se em (35) temos:

$$r_{02.1} = \phi_{2,2} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \quad (36)$$

Resolvendo as equações de Yule-Walker para $\phi_{1,1}$ e $\phi_{2,2}$ temos:

$$\begin{pmatrix} 1 & -\rho_1 \\ -\rho_1 & 1 \end{pmatrix} \begin{pmatrix} \phi_{1,1} \\ \phi_{2,2} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} \quad \begin{pmatrix} \phi_{1,1} \\ \phi_{2,2} \end{pmatrix} = \begin{pmatrix} \frac{\rho_1 - \rho_1\rho_2}{1 - \rho_1^2} \\ \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \end{pmatrix} \quad (37)$$

que são exatamente as expressões (34) e (36), isto é, para se obter as autocorrelações parciais basta resolver as equações de Yule-Walker para os $\phi's$. Observe que neste caso a função de autocorrelação parcial será zero para defasagem superior a dois.

Observe que se processo é um $AR(1)$ a primeira autocorrelação parcial é $\phi_{1,1} = \rho_1 = \phi_1$ e a segunda é igual a zero uma vez que $\rho_2 = \phi_1^2$ e o numerador de $\phi_{2,2}$ em (37) é igual a zero. Pode se mostrar que todas as outras autocorrelações parciais de ordem superiores a dois, para o processo $AR(1)$, também são iguais a zero.

Para um processo $AR(2)$ temos que a primeira autocorrelação parcial é dada por $\phi_{1,1} = \rho_1 = \phi_1$ e a segunda por $\phi_{2,2}$ em (37) e todas as autocorrelações de ordem superior a três são iguais a zero

Uma outra forma de obter a função de autocorrelação parcial é através da estimação de modelos AR com ordem crescentes, isto é

$$\begin{aligned}
X_t &= \phi_{1,1}X_{t-1} + u_{1t} \\
X_t &= \phi_{2,1}X_{t-1} + \phi_{2,2}X_{t-2} + u_{2t} \\
X_t &= \phi_{3,1}X_{t-1} + \phi_{3,2}X_{t-2} + \phi_{3,3}X_{t-3} + u_{3t} \\
X_t &= \phi_{4,1}X_{t-1} + \phi_{4,2}X_{t-2} + \phi_{4,3}X_{t-3} + \phi_{4,4}X_{t-4} + u_{4t} \\
&\vdots \\
&\vdots \\
&\vdots
\end{aligned}$$

Na primeira equação $\hat{\phi}_{1,1}$ estima a primeira autocorrelação parcial.

Na segunda equação $\hat{\phi}_{2,2}$ estima a segunda autocorrelação parcial.

Na terceira equação $\hat{\phi}_{3,3}$ estima a terceira autocorrelação parcial e assim por diante.

4.4 Distribuição da Função de Autocorrelação e de Autocorrelação Parcial

Considere uma série temporal X_t que seja estacionária fraca. A autocovariância de ordem k foi definida por $\gamma_k = Cov(X_t, X_{t-k})$ e satisfaz:

1. $\gamma_0 = Var(X_t)$
2. $\gamma_k = \gamma_{-k}$

A autocorrelação entre X_t e X_{t-k} , denotada por ρ_k foi definida por:

$$\rho_k = \frac{Cov(X_t, X_{t-k})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t-k})}} = \frac{Cov(X_t, X_{t-k})}{Var(X_t)} \quad (38)$$

Por (38) temos que:

1. $\rho_0 = 1$
2. $\rho_k = \rho_{-k}$
3. $-1 \leq \rho_k \leq 1$

Dada uma amostra da série temporal $\{X_t\}_{t=1}^T$ a primeira autocorrelação amostral é definida por:

$$\hat{\rho}_1 = \frac{\sum_{t=2}^T (X_t - \bar{X})(X_{t-1} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} \quad (39)$$

Se $\{X_t\}$ é i.i.d. com $E(X_t^2) < \infty$ então $\hat{\rho}_1 \stackrel{a}{\sim} N(0, \frac{1}{T})$.

Para testar $H_0 : \rho_1 = 0$ versus $H_a : \rho_1 \neq 0$ a estatística de teste é $\sqrt{T}\hat{\rho}_1$ que tem distribuição normal padrão.

A autocorrelação amostral de defasagem k ($0 < k \leq T - 1$) é definida por:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^T (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} \quad (40)$$

Se $\{X_t\}$ é i.i.d. com $E(X_t^2) < \infty$ então $\hat{\rho}_k \stackrel{a}{\sim} N(0, \frac{1}{T})$ para qualquer inteiro positivo k .

Se X_t é fracamente estacionário satisfazendo $X_t = \mu + \sum_{i=0}^q \psi_i u_{t-i}$ onde $\psi_0 = 1$, q é um inteiro não negativo e $\{u_t\}$ é ruído branco gaussiano então $\hat{\rho}_k \stackrel{a}{\sim} N(0, (1 + 2 \sum_{i=0}^q \rho_i^2)/T)$ para $k > q$.

4.5 Testando cada elemento da Função de Autocorrelação

Para se testar $H_0 : \rho_k = 0$ versus $H_a : \rho_k \neq 0$ a estatística de teste é:

$$t = \frac{\hat{\rho}_k}{\sqrt{(1 + 2 \sum_{i=0}^q \hat{\rho}_i^2)/T}} \quad (41)$$

Se X_t é fracamente estacionário e gaussiano com $\rho_j = 0$ para $j > k$ então a estatística dada em (41) tem distribuição assintótica normal padrão.

4.6 Teste Portmanteau

Para se testar $H_0 : \rho_1 = \dots = \rho_m = 0$ versus $H_a : \rho_i \neq 0$ para algum $i \in \{1, \dots, m\}$ a estatística de teste é:

$$Q^*(m) = T \sum_{l=1}^m \hat{\rho}_l^2 \quad (42)$$

que se $\{X_t\}$ é i.i.d. com certos momentos finitos tem distribuição qui-quadrada com m graus de liberdade, veja Box and Pierce (1970).

Ljung and Box (1978) modificam $Q^*(m)$ para melhorar a potência do teste em pequenas amostras para:

$$Q(m) = T(T+2) \sum_{l=1}^m \frac{\hat{\rho}_l^2}{T-l} \quad (43)$$

que também tem distribuição qui-quadrada com m graus de liberdade.

Para se testar se autocorrelações parciais são iguais a zero Quenouille (1949) mostrou que se o processo é um $AR(p)$ a autocorrelação parcial de ordens $p+1, p+2, \dots$ são independentemente e $Var(\hat{\phi}_{k,k}) \simeq \frac{1}{T}$.

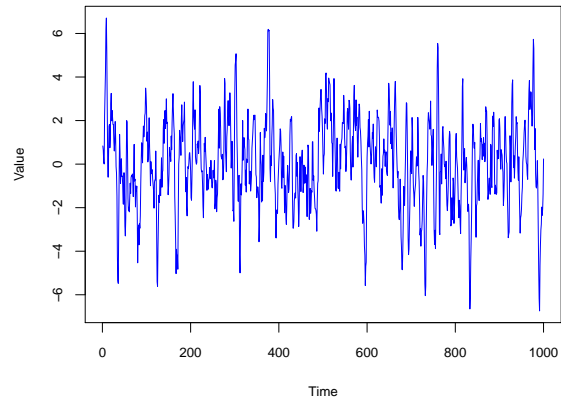
4.7 Função de Autocorrelação e de Autocorrelação Parcial no R

O seguinte programa gera um $AR(2)$ com raízes $\phi_1^* = 0.5$ e $\phi_2^* = 0.6$, assim como salva e apresenta o gráfico da série e também apresenta a função de autocorrelação e de autocorrelação parcial até 12 lags.

$AR(2)$ com raízes reais

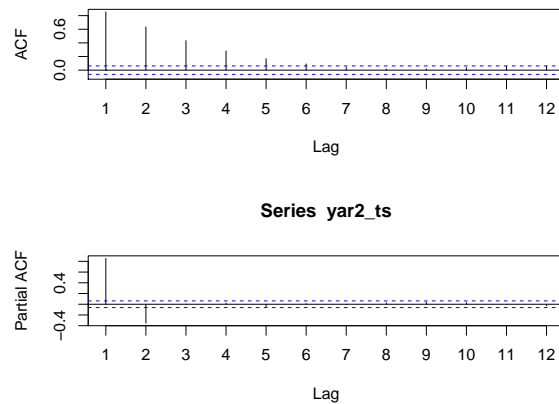
O gráfico abaixo apresenta a série gerada segundo um $AR(2)$

Figura 21: AR(2) com raízes reais



e a função de autocorrelação e autocorrelação parcial é dada abaixo

Figura 22: Função de Autocorrelação e de Autocorrelação Parcial para AR(2) com raízes reais



Observe que com os valores simulados era esperado que a primeira autocorrelação fosse 0,8461 e que a segunda fosse 0,6308, que são próximos dos

dois primeiros valores na coluna AC . Quanto as autocorrelações parciais temos que a primeira é próxima de 0,8461 e a segunda próxima de $-0,3$ - que é o valor teórico para este processo gerado. Observe também que todas as autocorrelações parciais para ordens superiores a dois são iguais a zero, uma vez que $2/\sqrt{T} = 0,06$.

4.8 Exercícios

1. Considere o seguinte processo $AR(2)$

$$X_t = 1.6 * X_{t-1} - .89 * X_{t-2} + u_t$$

com $u_t \sim NI(0, 1)$.

- (i) mostre que o processo em operador defasagem pode ser escrito da seguinte forma: $(1 - 1.6L + .89L^2)X_t = u_t$.
- (ii) obtenha as raízes do polinômio $(1 - 1.6L + .89L^2)$
- (iii) obtenha a função de autocorrelação para este processo. Esta função parece periódica? Porque?
- (iv) obtenha a função de autocorrelação parcial para este processo.

2. Refaça o exercício anterior para o processo: $X_t = 1.6 * X_{t-1} - X_{t-2} + u_t$.

Comente as semelhanças e diferenças entre os dois processos.

5 Processo Média Móvel de Ordem Superior

5.1 Introdução

O modelo (13) pode ser generalizado para ordens superiores a 1 e o caso geral corresponde ao modelo $MA(q)$ onde o processo gerador dos dados é dado por:

$$X_t = u_t + \theta_1 u_{t-1} + \dots + \theta_q u_{t-q} \quad (44)$$

Observe que o modelo AR(1) estacionário dado por (3) pode ser reescrito da forma:

$$\begin{aligned} X_t &= \phi X_{t-1} + u_t \Rightarrow (1 - \phi L)X_t = u_t \Rightarrow X_t = \frac{1}{(1 - \phi L)} u_t \\ \Rightarrow X_t &= \sum_{i=0}^{\infty} \phi^i L^i u_t \Rightarrow X_t = \sum_{i=0}^{\infty} \phi^i u_{t-i} \end{aligned} \quad (45)$$

que corresponde a um modelo $MA(\infty)$, dado que depende de todos os choques passados. Observe também que no modelo (13) quando θ satisfizer a condição $|\theta| < 1$ podemos reescrever (13) da seguinte forma:

$$\sum_{j=0}^{\infty} (-\theta)^j X_{t-j} = u_t \quad (46)$$

que corresponde a um modelo $AR(\infty)$. Observe que a condição $|\theta| < 1$ nos permitiu escrever o processo $MA(1)$ como um $AR(\infty)$ e é conhecida como **condição de invertibilidade**. Esta condição pode ser generalizada para o caso em que estamos considerando um modelo $MA(q)$ e será dada por: o modelo $MA(q)$ dado por (44) é invertível se θ_j^* para $j = 1, \dots, q$; as raízes do polinômio, $1 + \theta_1 L + \dots + \theta_q L^q$, satisfazem $|\theta_j^*| < 1$.

5.2 Função de Autocorrelação

A função de autocorrelação do modelo $MA(q)$ dado por (44) é facilmente obtida uma vez que a variância do processo é dada por:

$$\gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma^2 \quad (47)$$

e as autocovariâncias são dadas por:

$$\gamma_i = \left(\sum_{j=0}^{q-i} \theta_j \theta_{j+i} \right) \sigma^2 \quad i = 1, \dots, q \quad (48)$$

e para defasagens superiores a q as autocovariâncias são todas nulas. Portanto a função de autocorrelação é dada por:

$$\rho_i = \frac{\sum_{j=0}^{q-i} \theta_j \theta_{j+i}}{1 + \theta_1^2 + \dots + \theta_q^2} \quad i = 1, \dots, q \quad (49)$$

e $\rho_i = 0$ para $i > q$.

Deste modo o correlograma tem valores não nulos para defasagens até a ordem do processo.

5.3 Função de Autocorrelação Parcial

Para obter a função de autocorrelação parcial do processo MA devemos escrevê-lo como um AR . Por exemplo para o $MA(1)$ temos que a representação AR é dada por (46) e portanto a função de autocorrelação parcial é dada por:

$$r_{0j.2,3,\dots,j-1,j+1,\dots} = (-1)^{j+1} \theta^j \quad (50)$$

e, portanto, tem um decaimento exponencial como a função de autocorrelação do processo AR .

Para o processo $MA(q)$ geral uma expressão analítica como (50) é difícil de obter mas o padrão é o mesmo do $MA(1)$, isto é, o decaimento é exponencial.

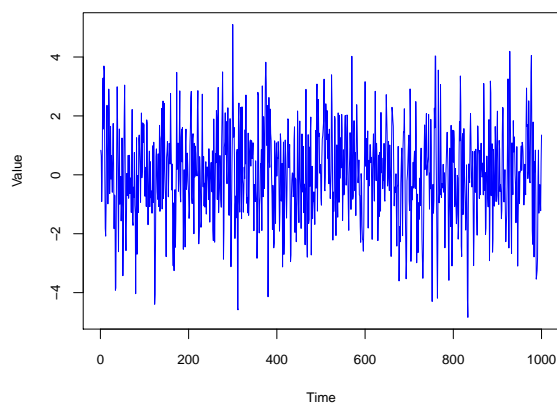
5.4 Função de Autocorrelação e de Autocorrelação Parcial no R

O seguinte programa gera um $MA(2)$ com raízes $\theta_1 = -1.1$ e $\theta_2 = 0.3$, assim como salva e apresenta o gráfico da série e também apresenta a função de autocorrelação e de autocorrelação parcial até 12 lags.

`MA(2)`

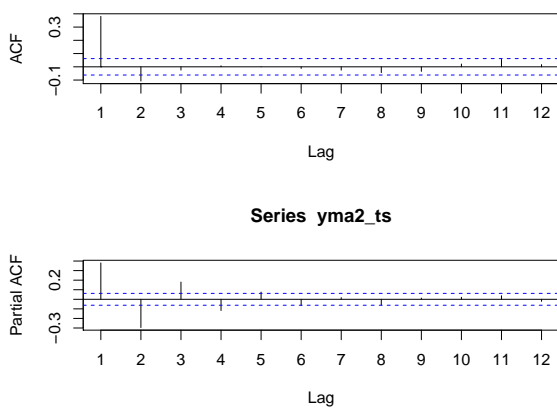
O gráfico abaixo apresenta a série gerada segundo um $MA(2)$

Figura 23: $MA(2)$



e a função de autocorrelação e autocorrelação parcial é apresentada abaixo:

Figura 24: Função de autocorrelação e autocorrelação parcial para $MA(2)$



5.5 Exercícios

1. Mostre que o modelo $MA(1)$ dado por $X_t = u_t + \theta u_{t-1}$ pode ser escrito como um $AR(\infty)$.
2. Seja dado o modelo $MA(2)$ dado por $X_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$, mostre que a representação $AR(\infty)$ deste processo pode ser escrita da seguinte forma: $\sum_{j=0}^{\infty} (-\theta_1 - \theta_2 L)^j X_{t-j} = u_t$
3. Mostre que o modelo $MA(1)$ dado por $X_t = u_t - u_{t-1}$ é estacionário embora não seja invertível.

6 Processo ARMA

6.1 Introdução

Podemos pensar em modelos que tenham as duas características das apresentadas em (3) e (13), isto é, modelos cujo processo gerador dos dados seja dado por:

$$X_t = \phi X_{t-1} + u_t + \theta u_{t-1} \quad (51)$$

Este tipo de modelo é chamado de autorregressivo médias-móveis e denotado por $ARMA(1, 1)$. Observe que a condição para que o modelo seja assintoticamente independente é a mesma do modelo $AR(1)$, isto é, $|\phi| < 1$ e deve ser imposta para garantir a estacionaridade do processo. Como as covariâncias de ordem superior a 1 vão depender sómente da contribuição da parte autorregressiva, tem-se que elas também decairão exponencialmente.

Podemos generalizar este modelo para ordens maiores do que 1, por exemplo, ordem p para a parte autorregressiva e ordem q para a parte de médias-móveis, obtendo-se então o processo $ARMA(p, q)$ cujo processo gerador dos dados toma a forma:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + u_t + \theta_1 u_{t-1} + \dots + \theta_q u_{t-q} \quad (52)$$

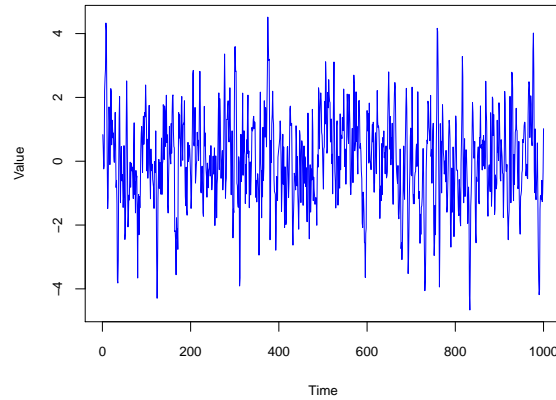
6.2 Gerando um ARMA(1,1) no R

O programa abaixo gera um $ARMA(1,1)$ com parâmetros $\phi = 0.5$, $\theta = 0.3$ e $u \sim N(0,1)$. Apresenta também o gráfico da série contra o tempo.

Modelo $ARMA(1,1)$ com $\phi = 0.5$ e $\theta = 0.3$

O gráfico desta variável é dado por:

Figura 25: $ARMA(1,1)$



Observe que esta série cruza frequentemente o zero que é a média do processo e seus valores estão entre -2.7 e 2.7 .

6.3 Função de Autocorrelação e de Autocorrelação Parcial para o $ARMA(1,1)$

Observe que (51) pode ser escrito como um $MA(\infty)$ da seguinte forma:

$$\begin{aligned} X_t &= \frac{1 + \theta L}{1 - \phi L} u_t \\ &= u_t + \sum_{j=1}^{\infty} (\theta + \phi) \phi^{j-1} u_{t-j} \end{aligned} \quad (53)$$

logo temos que

$$E(X_t) = 0$$

$$\begin{aligned}\gamma_0 &= E(X_t^2) = \sigma^2 + \frac{(\theta + \phi)^2}{1 - \phi^2} \sigma^2 \\ \gamma_0 &= \frac{1 + \theta^2 + 2\theta\phi}{1 - \phi^2} \sigma^2\end{aligned}\tag{54}$$

Para obter a primeira autocovariância, multiplique (51) por X_{t-1} e usando o operador esperança temos:

$$\gamma_1 = \phi\gamma_0 + E(X_{t-1}u_t) + \theta E(X_{t-1}u_{t-1})\tag{55}$$

Agora por (53) temos que:

$$E(X_{t-1}u_t) = 0\tag{56}$$

e

$$E(X_{t-1}u_{t-1}) = \sigma^2\tag{57}$$

logo substituindo-se (54,56,57) em (55) temos:

$$\begin{aligned}\gamma_1 &= \phi \frac{1 + \theta^2 + 2\theta\phi}{1 - \phi^2} \sigma^2 + \theta \sigma^2 \\ \gamma_1 &= \frac{(1 + \phi\theta)(\theta + \phi)}{1 - \phi^2} \sigma^2\end{aligned}\tag{58}$$

Para as autocovariância de ordem superiores a um temos a seguinte relação de recorrência:

$$\gamma_k = \phi\gamma_{k-1}\tag{59}$$

Podemos então definir a função de autocorrelação por:

$$\rho_k = \begin{cases} \frac{(1+\phi\theta)(\theta+\phi)}{1+\theta^2+2\theta\phi} & k = 1 \\ \phi\rho_{k-1} & k > 1 \end{cases}\tag{60}$$

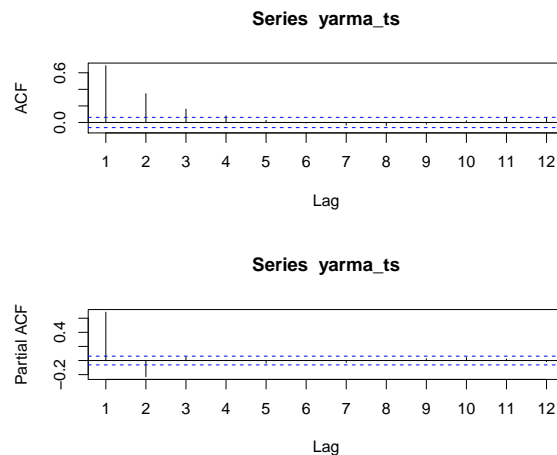
que a partir da segunda defasagem decae exponencialmente.

Como o processo *ARMA* pode ser escrito como uma *AR*(∞) a FACP terá o mesmo padrão de um *MA*, isto é, decairá exponencialmente.

6.4 Função de Autocorrelação e Autocorrelação Parcial no R.

Para o *ARMA* gerado acima temos que a FAC e FACP será dada pelo gráfico abaixo, onde a primeira autocorrelação teórica seria 0,662 e a segunda 0,331 e as próximas seriam dadas por (59)

Figura 26: FAC e FACP para $ARMA(1,1)$



Vamos considerar a seguir outros casos especiais que serão importante tanto no estudo de econometria quanto na caracterização de séries temporais. O caso que consideraremos a seguir é o das variáveis integradas.

6.5 Exercícios

1. Um modelo é dito ser mais parcimonioso do que outro se contém menos parâmetros. Mostre que um modelo $ARMA(1,1)$ é mais parcimonioso do que um modelo $AR(p)$.
2. Escreva um programa no R para gerar o seguinte modelo $X_t = .5X_{t-1} + u_t - 0.5u_{t-1}$. Obtenha a FAC e FACP para este modelo e comente. Esta série parece ruído branco? Porque?

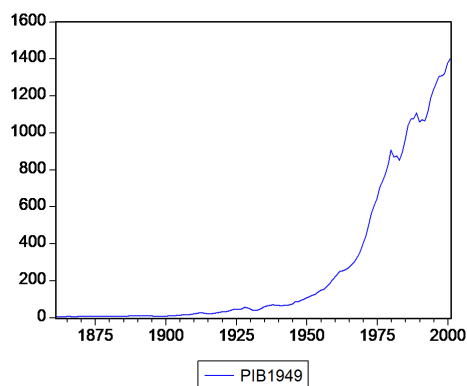
3. No arquivo bolsa.xls apresenta os dados dos índices Bovespa (ibv), IBX (ibx) e do S&P500 (sp500) assim como da taxa de câmbio dolar real para o período de 3/1/2000 até 5/3/2003.
- (i) construa as séries de retornos simples, isto é, $rs_t = \frac{p_t - p_{t-1}}{p_{t-1}}$;
 - (ii) construa as séries de retornos compostos, isto é, $rc_t = \Delta \log(p_t) = \log(p_t) - \log(p_{t-1})$;
 - (iii) usando o R, obtenha as funções de autocorrelação (FAC) e autocorrelação parcial (FACP) para as séries usando as transformações dos itens (i) e (ii) e comente.
 - (iv) construa as séries de retornos ao quadrado, isto é, $rssq_t = rs_t * rs_t$ e $rcsq_t = rc_t * rc_t$ e usando o R, obtenha a FAC e FACP para as séries e comente.

7 Representações de Variáveis Integradas

7.1 Introdução

O conceito de variável integrada pode ser racionalizado utilizando como fato estilizado o produto agregado que tem um crescimento de longo prazo e flutuações em torno desta trajetória de crescimento. Por exemplo, se considerarmos o PIB real para o Brasil de 1860 até 1980, o crescimento médio foi de 6,93 e ao se extrapolar este crescimento para a década de 80 o observado estará bem distante da previsão tendencial dada pela década anterior.

Figura 27: PIB Brasileiro desde 1860 até 2001



Este exercício tenta mostrar que o crescimento de longo prazo do produto, do ponto de vista estatístico, se alterou na década de 80 e do ponto de vista econômico as teorias de crescimento tentam explicar estas mudanças na tendência de longo-prazo da série do produto. Mas, estatisticamente, tratar este crescimento como uma tendência linear determinística é problemático se extrapolações futuras forem feitas. Por outro lado os desvios deste crescimento, tem na teoria dos ciclos econômicos uma fonte de explicação dado que o objetivo é explicar as flutuações de curto-prazo e, também, determinar quando políticas macroeconômicas podem estabilizar ou exacerbar os movimentos entre expansões e recessões. Para uma descrição das teorias estatísticas dos ciclos econômicos, segundo uma visão histórica, veja Morgan (1991). Deste modo a série do produto pode ser pensada com composta de dois componentes: um associado a tendência da série e o outro associado aos ciclos.

Para caracterizar a tendência, ou o crescimento de longo-prazo da série, pode-se adotar, como foi feito acima, uma tendência determinística, mas, também como foi visto acima, a extrapolação desta tendência determinística é problemática.

Uma alternativa é fazer com que esta tendência seja mais adaptativa, permitindo que o nível e/ou a taxa de crescimento evoluam ao longo do tempo. Este tipo de modelo é chamado de processos integrados que podem ser racionalizado como um processo que acumula os erros passados.

7.2 Especificação de um processo $I(1)$

O processo gerador dos dados de um processo integrado de primeira ordem é dado por:

$$\Delta X_t = \alpha + u_t \text{ com } u_t \sim IID(0, \sigma_u^2) \text{ para } t = 1, \dots, T \quad (61)$$

Observe que (61) pode ser reescrito da seguinte forma:

$$\begin{aligned} X_t &= X_{t-1} + \alpha + u_t = X_{t-2} + 2\alpha + u_t + u_{t-1} = \dots = X_0 + \alpha t + \sum_{j=0}^{t-1} u_{t-j} \\ \Rightarrow X_t &= X_0 + \alpha t + \sum_{j=0}^{t-1} u_{t-j} \end{aligned} \quad (62)$$

e desta forma X_t é dado pela acumulação dos erros passados e $X_0 + \alpha t$ que representa uma tendência linear determinística, isto é, uma reta, se X_0 for fixo.

As seguintes características são válidas para (61):

$$E(\Delta X_t) = \alpha \quad (63)$$

$$Var(\Delta X_t) = E(\Delta X_t - \alpha)^2 = E(u_t^2) = \sigma_u^2 \quad (64)$$

$$Cov(\Delta X_t, \Delta X_s) = E[(\Delta X_t - \alpha)(\Delta X_s - \alpha)] = E(u_t u_s) = 0 \quad \forall t \neq s \quad (65)$$

e para (62):

$$E(X_t) = X_0 + \alpha t \quad (66)$$

$$Var(X_t) = E(X_t - X_0 - \alpha t)^2 = E\left(\sum_{j=0}^{t-1} u_{t-j}\right)^2 = \sum_{j=0}^{t-1} E(u_{t-j}^2) = t \sigma_u^2 \quad (67)$$

$$\begin{aligned}
Cov(X_t, X_s) &= E[(X_t - E(X_t))(X_s - E(X_s))] = E \left[\left(\sum_{j=0}^{t-1} u_{t-j} \right) \left(\sum_{i=0}^{s-1} u_{s-i} \right) \right] \\
&= E \left[\sum_{s=0}^{s-1} u_{s-i}^2 + \sum_{j=s}^{t-1} \sum_{i=0}^{s-1} u_{t-j} u_{s-i} \right] = \sum_{s=0}^{s-1} E[u_{s-i}^2] \\
&= s \sigma_u^2
\end{aligned} \tag{68}$$

observe que enquanto o processo ΔX_t tem média e variância independentes do tempo e covariâncias iguais a zero o processo X_t tem média que varia linearmente com o tempo, variância que cresce com o tempo e covariâncias que depende dos instantes de tempo tendo a mesma dimensão que a variância. As correlações decaem muito lentamente. Deste modo as características dos dois processos são bem distintas. Enquanto que o processo ΔX_t é estacionário de segunda ordem, X_t é um processo não estacionário tanto de primeira quanto de segunda ordem. Observe que embora este último processo tenha esta propriedade indesejável uma simples transformação, tomar primeira diferença, o torna estacionário. Esta é a definição de processo **diferença estacionário** ou **integrado**, no nosso caso bastou tomar uma diferença para induzir estacionaridade. Este processo é chamado de integrado de primeira ordem e denotado por $I(1)$ e, portanto, os processos estacionários serão denotados por $I(0)$. Observe que como as correlações decaem muito lentamente esta propriedade pode ser usada para identificar um processo integrado. Este tipo de modelo é bastante importante no estudo de séries econômicas porque estas séries, como foi visto em exemplos anteriores, apresentam algum tipo de tendência e os modelos $I(1)$ como o apresentado acima são modelos que apresentam o que é chamado de **tendência estocástica**, isto é, uma tendência que se modifica ao longo do tempo.

7.2.1 Modelo de Passeio Aleatório com Drift ou Modelo de Nível Local com Taxa de Crescimento Constante

Uma outra forma de escrever (61) que faz com que a denominação de tendência estocástica fique mais clara é a seguinte:

$$X_t = \mu_t \tag{69}$$

$$\mu_t = \mu_{t-1} + \alpha_t + \zeta_t \quad (70)$$

$$\alpha_t = \alpha_{t-1} \quad (71)$$

observe que μ_t representa o nível da tendência, α_t é a taxa de crescimento e, por (71) esta taxa de crescimento é constante, enquanto que o nível, (70), varia ao longo do tempo e $\zeta_t \sim IID(0, \sigma_\zeta^2)$. Este modelo é chamado, na literatura de séries temporais, de modelo de **nível local com taxa de crescimento constante**, porque o nível da série varia localmente. Para mostrar que (69-71) é equivalente a (61), observe que (70) pode ser reescrito da seguinte forma:

$$\Delta\mu_t = \alpha_t + \zeta_t \quad (72)$$

e como por (71) $\alpha_t = \alpha_{t-1}$, isto implica que α_t é constante e, portanto (72) pode ser reescrito da seguinte forma:

$$\Delta\mu_t = \alpha + \zeta_t \quad (73)$$

e tomando-se a primeira diferença em (69) e usando (73) chega-se na equivalência entre os dois modelos se $\sigma_u^2 = \sigma_\zeta^2$.

Na literatura de séries temporais o modelo(69-71) é chamado de representação estrutural enquanto que (73) é a forma reduzida desta representação. Observe que no primeiro caso o processo gerador dos dados é especificado para o nível da série enquanto que no segundo é para a primeira diferença, perdendo-se, portanto a informação do nível.

Esta última representação, isto é, (73), está associada a metodologia tradicional, ou de Box and Jenkins (1976) enquanto que a primeira, isto é (69-71), está associada a metodologia estrutural de séries temporais desenvolvida por Harvey (1990) num contexto clássico e por Harrison and Stevens (1976) num contexto Bayesiano (para uma comparação ver Migon and Valls Pereira (1984)).

7.2.2 Modelo de Tendência Local

Uma extensão natural é permitir que a taxa de crescimento também varie ao longo do tempo e isto pode ser feito acrescentando-se um outro erro a equação (71), que neste caso fica:

$$\alpha_t = \alpha_{t-1} + \eta_t \quad (74)$$

onde η_t são $IID(0, \sigma_\eta^2)$ e também são independentes dos $\zeta's$.

O modelo composto por (69,70,74) é chamado de **tendência local sem erro observacional** ou de medida, porque tanto o nível quanto a taxa de crescimento variam localmente e a extração do componente de tendência da série é feita sem erro, que é um hipótese pouco plausível. O modelo mais geral contemplaria este erro de medida e (69) seria escrito da seguinte forma:

$$X_t = \mu_t + u_t \quad (75)$$

e o modelo composto por (75, 70,74) é chamado de **tendência local** porque tanto o nível quanto a taxa de crescimento variam localmente.

Observe que este modelo engloba os modelos de passeio aleatório (modelo de nível local sem erro de medida e com taxa de crescimento constante e igual a zero); passeio aleatorio com "drift" (modelo de nível local sem erro de medida e com taxa de crescimento constante e diferente de zero); passeio aleatório mais ruído (modelo de nível local com erro de medida e com taxa de crescimento constante e igual a zero)

7.2.3 Tendência Determinística ou Tendência Estacionária

Um outro sub-modelo que será apresentado abaixo, **tendência linear determinística**, será obtido quando não se tem choques em (70) e (74) . Tem-se então a seguinte representação estrutural para o modelo de tendência linear determinística:

$$X_t = \mu_t + u_t \quad (76)$$

$$\mu_t = \mu_{t-1} + \alpha_t \quad (77)$$

$$\alpha_t = \alpha_{t-1} \quad (78)$$

e se substituirmos para trás, e assumindo que μ_0 e α_0 são fixos, obteremos uma tendência linear determinística, isto é:

$$X_t = \mu_0 + \alpha_0 t + u_t \quad (79)$$

Observe que este processo tem as seguintes características:

$$E(X_t) = \mu_0 + \alpha_0 t \quad (80)$$

$$Var(X_t) = E(y_t - \mu_0 - \alpha_0 t)^2 = E(u_t)^2 = \sigma_u^2 \quad (81)$$

$$Cov(X_t, X_s) = E[(X_t - \mu_0 - \alpha_0 t)(X_s - \mu_0 - \alpha_0 s)] = E(u_t u_s) = 0 \quad \forall t \neq s \quad (82)$$

e como a média é uma função linear do tempo, este processo é não estacionário na média.

Para o modelo diferença estacionário a transformação que induzia estacionaridade era diferenciar a série. Agora para o modelo de tendência global esta transformação é inadequada porque ao se diferenciar (79) obtém-se:

$$\Delta X_t = \alpha_0 + \Delta u_t \quad (83)$$

Este processo tem as seguintes características:

$$E(\Delta X_t) = \alpha_0 \quad (84)$$

$$Var(\Delta X_t) = E(\Delta X_t - \alpha_0)^2 = E(\Delta u_t)^2 = \sigma_u^2 \quad (85)$$

$$Cov(\Delta X_t, \Delta X_{t-1}) = E(\Delta X_t - \alpha_0)(\Delta X_{t-1} - \alpha_0) = E(\Delta u_t \Delta u_{t-1}) = -\sigma_u^2 \quad (86)$$

$$Cov(\Delta X_t, \Delta X_{t-j}) = E(\Delta X_t - \alpha_0)(\Delta X_{t-j} - \alpha_0) = 0 \quad \forall j > 1 \quad (87)$$

sendo portanto estacionário. Observe que a não estacionaridade na média foi eliminada mas induziu-se uma maior variância e também autocorrelação negativa no processo, fazendo com que este processo seja muito errático. Pela estrutura de autocovariância pode-se notar que este processo pode ser caracterizado por um $MA(1)$ mas com $\theta = -1$, isto é, embora estacionário este processo é não invertível, que é uma propriedade indesejável.

Por outro lado pode-se induzir estacionaridade na série retirando-se a tendência determinística, isto é, definindo a nova variável, $X_t^d = X_t - \mu_0 -$

α_0 $t = u_t^d$ que terá média zero, variância constante e será não serialmente correlacionada, portanto estacionária. Podemos então chamar o modelo (79) de tendência estacionário, dado que para induzir estacionaridade devemos retirar a tendência determinística (*destrendar* a série).

7.2.4 Forma Reduzida do Modelo de Tendência Local sem Erro Observacional

O modelo composto por (69, 70, 71) tem por forma reduzida:

$$(1 - L)^2 X_t = \eta_t + (1 - L) \zeta_t$$

Observe também que a estrutura dos erros em (7.2.4) pode ser rescrita da seguinte forma:

$$\varepsilon_t + \theta \varepsilon_{t-1} \quad (88)$$

onde ε_t são $IID(0, \sigma_\varepsilon^2)$ e os parâmetros θ e σ_ε^2 estão relacionados aos parâmetros σ_η^2 e σ_ζ^2 pelas seguintes equações:

$$(1 + \theta^2) \sigma_\varepsilon^2 = \sigma_\eta^2 + 2\sigma_\zeta^2 \quad (89)$$

$$\theta \sigma_\varepsilon^2 = - \sigma_\zeta^2 \quad (90)$$

onde a primeira equação foi obtida igualando-se a variância de (7.2.4) com a de (88) e a segunda igualando-se a covariância de ordem 1. Como todas as covariâncias de ordem superior a um de (7.2.4) são zero, os erros podem ser representados por um processo $MA(1)$, como o dado por (88).

Desta forma por (7.2.4) e (88) pode-se reescrever o modelo da seguinte forma:

$$(1 - L)^2 X_t = \varepsilon_t + \theta \varepsilon_{t-1} \quad (91)$$

Observe que o modelo dado por (91) é estacionário de segunda ordem, mas isto foi obtido após tomar duas diferenças na série X_t . Deste modo a série X_t é uma série integrada de segunda ordem e será denotada por $I(2)$. Observe também que a segunda diferença de X_t pode ser descrita como um modelo $MA(1)$ e será denotada por $IMA(2, 1)$ isto é, um processo integrado de segunda ordem com parte MA de ordem 1, que é um caso particular

dos modelos autorregressivos integrados médias-móveis que são denotados por $ARIMA(p, d, q)$ onde p corresponde a ordem da parte autorregressiva, d corresponde a ordem de integração do processo e q é a ordem da parte média móvel.

7.2.5 Forma Reduzida do Modelo de Tendência Local com Erro Observacional

No caso em que existe erro de observação, isto é, o processo gerador dos dados é dado por (76,70,74) a forma reduzida é dada por

$$(1 - L)^2 X_t = \eta_t + (1 - L) \zeta_t + (1 - L)^2 u_t \quad (92)$$

onde a estrutura dos erros de (92) pode ser reescrita da seguinte forma:

$$\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \quad (93)$$

onde ε_t são $IID(0, \sigma_\varepsilon^2)$ e os parâmetros θ_1 , θ_2 e σ_ε^2 estão relacionados as parâmetros σ_η^2 , σ_ζ^2 e σ_u^2 pelas seguintes equações:

$$(1 + \theta_1^2 + \theta_2^2) \sigma_\varepsilon^2 = \sigma_\eta^2 + 2 \sigma_\zeta^2 + 6 \sigma_u^2 \quad (94)$$

$$\theta_1(1 + \theta_2) \sigma_\varepsilon^2 = - \sigma_\zeta^2 - 4 \sigma_u^2 \quad (95)$$

$$\theta_2 \sigma_\varepsilon^2 = \sigma_u^2 \quad (96)$$

onde a primeira equação foi obtida igualando-se a variância de (92) com a de (93) e a segunda e terceira igualando-se a covariância de ordem 1 e 2. Como todas as covariâncias de ordem superior a dois de (92) são zero aqueles erros podem ser representados por um processo $MA(2)$, como o dado por (93).

Desta forma por (92) e (93) pode-se reescrever o modelo da seguinte forma:

$$(1 - L)^2 X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \quad (97)$$

Observe que o modelo dado por (97) é estacionário de segunda ordem, mas isto foi obtido após tomar duas diferenças na série X_t . Deste modo a série X_t é uma série integrada de segunda ordem e será denotada por $I(2)$. Observe também que a segunda diferença de X_t pode ser escrita como um modelo $MA(2)$ e será denotada por $IMA(2, 2)$ ou $ARIMA(0, 2, 2)$.

7.2.6 Principais diferenças entre séries $I(0)$ e $I(1)$

Podemos então ilustrar as diferenças entre $\mathbf{I}(0)$ e $\mathbf{I}(1)$ pelas seguintes características:

(i) séries $\mathbf{I}(0)$ são menos suaves e com flutuações mais óbvias do que séries $\mathbf{I}(1)$;

(ii) uma série $\mathbf{I}(0)$ retornará ao seu valor médio frequentemente, enquanto que uma série $\mathbf{I}(1)$ raramente retornará a qualquer valor, inclusive seu valor inicial.

(iii) As previsões com uma série $\mathbf{I}(0)$ vão decair ao longo do tempo, isto é seus efeitos são transitórios, mas para uma série $\mathbf{I}(1)$ as previsões serão dadas pelo nível da série na origem do horizonte de previsão indicando uma certa persistência na série.

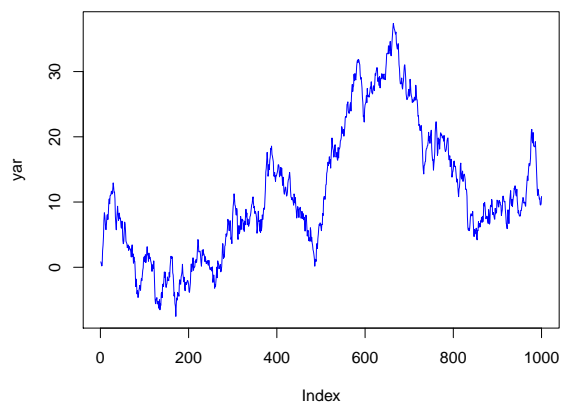
7.3 Gerando um passeio aleatório no R

O programa abaixo gera um passeio aleatório no R com a condição inicial de que $X_0 = 0$, apresenta o gráfico da série gerada contra o tempo e também a função de autocorrelação e de autocorrelação parcial para 12 defasagens.

Passeio Aleatório

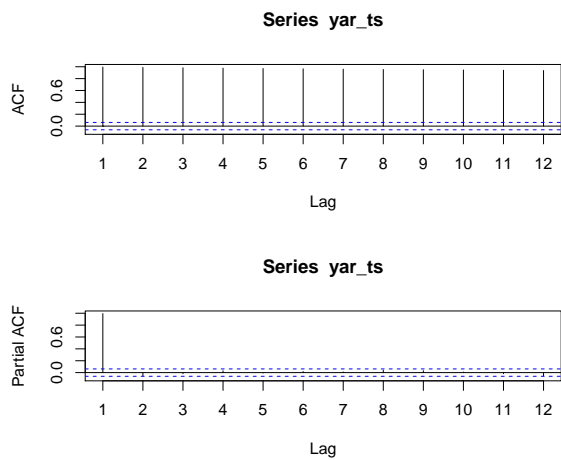
O gráfico deste passeio aleatório é apresentado abaixo

Figura 28: Passeio Aleatório



e o gráfico da FAC e FACP é apresentado abaixo

Figura 29: FAC e FACP para Passeio Aleatório



Observe que a FAC decaie lentamente neste caso e que a FACP tem sómente a primeira autocorrelação parcial diferente de zero, caracterizando

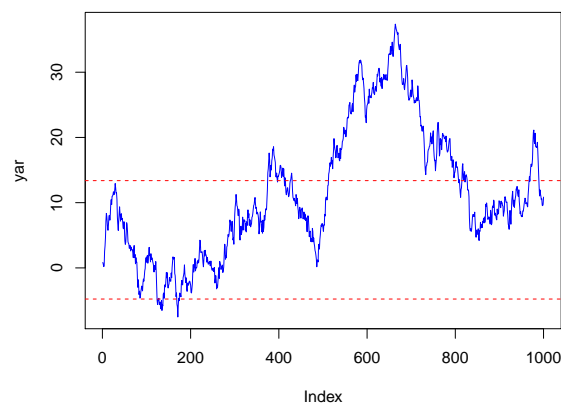
um processo do tipo $AR(1)$. Observe que o coeficiente do $AR(1)$ deve ser igual a um, caracterizando um processo não estacionário.

Uma forma intuitiva de verificar a não estacionridade é Através da **Janela de Valls** que é definida por:

1. Seja dada a série temporal X_t com $t = 1, \dots, T$;
2. Para as primeiras n observações, com por exemplo $n = \frac{T}{10}$, calcule a média e desvio padrão para X_t para $t = 1, \dots, n$, e denote por $media(X_n)$ e $dp(X_n)$;
3. Construa o Intervalo de Confiança em torno da média usando, $[ls, li]$, onde;
 - (i) $li = media(X_n) - 1.96 * dp(X_n)$ e
 - (ii) $ls = media(X_n) + 1.96 * dp(X_n)$.
4. Construa o gráfico contra o tempo para li, X_t, ls .
5. A série temporal será não estacionária se ela não cruzar frequentemente $media(X_n)$.

Temos então o seguinte gráfico

Figura 30: Janela de Valls para Passeio Aleatório



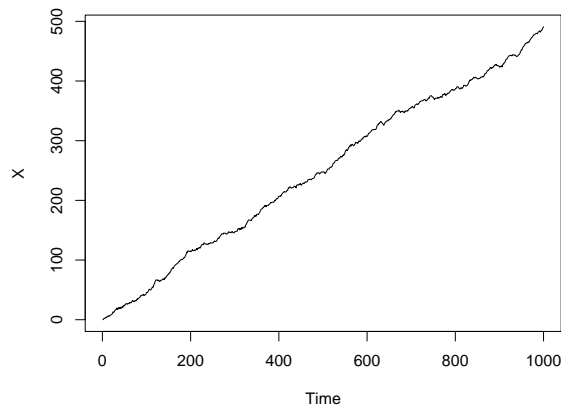
7.4 Gerando um Passeio Aleatório com drift através do Modelo de Nível Local no R

O programa abaixo gera um Modelo de Nível Local no R com a condição inicial de que $\alpha_0 = 0.5$, apresenta o gráfico da série gerada contra o tempo e também a função de autocorrelação e de autocorrelação parcial para 12 defasagens.

Passeio Aleatório com drift

O gráfico deste modelo é apresentado abaixo:

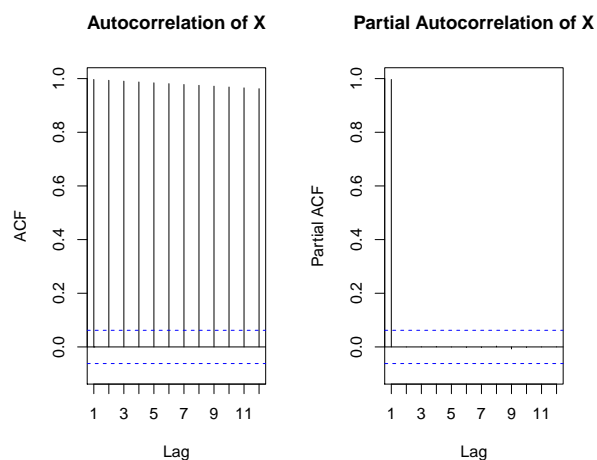
Figura 31: Passeio Aleatório com drift



Observe que os dados tem tendência linear no nível, sendo uma decorrência do drift ser diferente de zero.

Abaixo é apresentada a FAC e FACP para este processo.

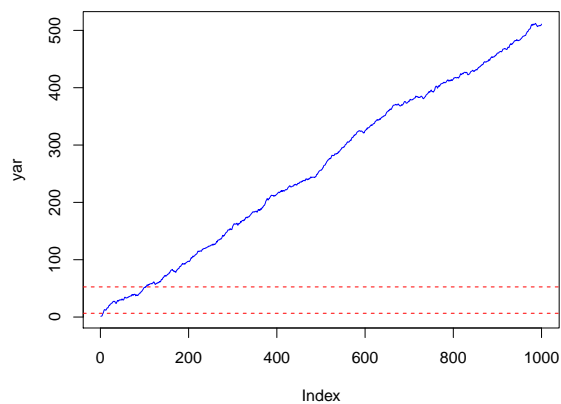
Figura 32: FAC e FACP do Passeio Aleatório com drift



Observe que o padrão da FAC e FACP é semelhante ao do passeio aleatório sem drift indicando que a série é integrada de primeira ordem.

A janela de Valls neste caso é dada pelo seguinte gráfico:

Figura 33: Janela de Valls para Passeio Aleatório sem drift



8 Apêndice I - Algumas Transformações

O **operador defasagem de primeira ordem**, denotado por L , é definido por

$$L X_t = X_{t-1} \quad (98)$$

Observe que podemos definir também **defasagens de ordem superior** a um simplesmente usando o operador diversas vezes. Por exemplo a j -ésima defasagem da série X_t é definida por:

$$L^j X_t = X_{t-j} \quad (99)$$

Podemos definir o **operador diferença de primeira ordem**, denotado por Δ , por:

$$\Delta X_t = X_t - X_{t-1} = (1 - L)X_t \quad (100)$$

e o operador diferença de ordem k , denotado por Δ^k , será definido por:

$$\Delta^k X_t = (1 - L)^k X_t \quad (101)$$

e o **operador diferença sazonal de ordem s** , denotado por Δ_s , será definido por:

$$\Delta_s X_t = (1 - L^s)X_t = X_t - X_{t-s}$$

Vamos supor que a série observada foi transformada usando a transformação logarítmica e denotaremos por letras minúsculas esta transformação, i.e. $x_t = \ln(X_t)$. Observe que neste caso podemos interpretar a primeira diferença como a taxa de crescimento em períodos consecutivos e a diferença sazonal de ordem 12 com a taxa de crescimento entre meses de anos consecutivos.

O operador defasagem de primeira ordem no R é definido por:

```
x1 <- dplyr::lag(x, n=1)
```

e a defasagem de ordem l é definida por:

```
xl <- dplyr::lag(x, n=l)
```

O operador primeira diferença é definido por:

```
dx <- diff(x)
```

e para diferença de ordem n por:

```
dxn <- diff(x, differences = n)
```

e para a diferença sazonal de ordem s por:

```
 $dx_{seasonal} \leftarrow \text{diff}(x, \text{lag} = s)$ 
```

É possível também definir a taxa de crescimento diretamente como sendo a primeira diferença do logaritmo da série, isto é;

```
 $dlog_x \leftarrow \text{diff}(\log(x))$ 
```

e todas as outras transformações definidas acima também podem ser usada para o logaritmo da série em vez da própria série.

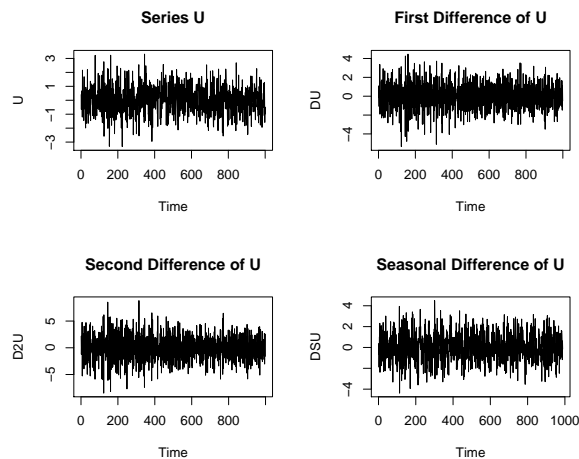
Vamos gerar um série com 1000 observações de uma distribuição normal padrão e definir a primeira diferença, a diferença de segunda ordem e a diferença sazonal de segunda ordem.

Isto pode ser feito através do seguinte programa no R:

Diferenças

E o gráfico das quatro séries é apresentado abaixo

Figura 34: Diferenças Diversas



9 Apêndice II - Solução de uma equação de primeira ordem no operador defasagem e sua inversão.

Observe que o modelo autorregressivo pode ser escrito em operador defasagem da seguinte forma:

$$\begin{aligned}X_t &= \phi X_{t-1} + u_t \\X_t - \phi X_{t-1} &= u_t \\X_t - \phi L X_t &= u_t \\(1 - \phi L)X_t &= u_t\end{aligned}\tag{102}$$

Observe que o lado direito da expressão (102) é uma equação de primeira ordem no operador defasagem cuja solução é dada por:

$$1 - \phi L = 0 \Rightarrow L = 1/\phi\tag{103}$$

ou fazendo $L = 1/z$ temos:

$$1 - \phi/z = 0 \Rightarrow z = \phi\tag{104}$$

y

Para que o processo seja estacionário precisamos que ϕ seja, em módulo, menor do que um. Então se usarmos a solução dada em (103) diremos que ϕ tem que estar fora do círculo unitário e se usarmos (104) diremos que ϕ tem que estar dentro do círculo unitário.

Dado que o processo é estacionário a inversão do polinômio $(1 - \phi L)$ é possível e será dada por:

$$(1 - \phi L)^{-1} = \frac{1}{1 - \phi L} = \sum_{i=0}^{\infty} (\phi L)^i = 1 + \phi L + \phi^2 L^2 + \dots\tag{105}$$

e este resultado é facilmente obtido fazendo-se $a = \phi L$ e usando a propriedade de que a soma de uma Progressão Geométrica infinita é dada pelo primeiro termo dividido por um menos a razão.

10 Apêndice III - Solução do Sistema (17-19)

Considere o sistema:

$$\begin{aligned}\gamma_0 &= \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2 \\ \gamma_1 &= \phi_1\gamma_0 + \phi_2\gamma_1 \\ \gamma_2 &= \phi_1\gamma_1 + \phi_2\gamma_0\end{aligned}$$

Este sistema escrito em forma matricial é dado por:

$$\begin{bmatrix} 1 & -\phi_1 & -\phi_2 \\ -\phi_1 & 1-\phi_2 & 0 \\ -\phi_2 & -\phi_1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ 0 \end{bmatrix} \quad (106)$$

cuja solução é dada por

$$\begin{aligned} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} &= \begin{bmatrix} 1 & -\phi_1 & -\phi_2 \\ -\phi_1 & 1-\phi_2 & 0 \\ -\phi_2 & -\phi_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{-\phi_2+1}{-\phi_2-\phi_1^2-\phi_2^2+\phi_2^3-\phi_1^2\phi_2+1} & -\frac{\phi_1}{2\phi_2+\phi_1^2-\phi_2^2-1} & \frac{-\phi_2+\phi_2^2}{\phi_2+\phi_1^2+\phi_2^2-\phi_2^3+\phi_1^2\phi_2-1} \\ -\frac{\phi_1}{\phi_2+\phi_1^2+\phi_2^2-\phi_2^3+\phi_1^2\phi_2-1} & \frac{\phi_2-1}{2\phi_2+\phi_1^2-\phi_2^2-1} & \phi_1 \frac{\phi_2}{-\phi_2-\phi_1^2-\phi_2^2+\phi_2^3-\phi_1^2\phi_2+1} \\ \frac{\phi_2+\phi_1^2-\phi_2^2}{-\phi_2-\phi_1^2-\phi_2^2+\phi_2^3-\phi_1^2\phi_2+1} & -\frac{\phi_1}{-2\phi_2-\phi_1^2+\phi_2^2+1} & \frac{\phi_2+\phi_1^2-1}{\phi_2+\phi_1^2+\phi_2^2-\phi_2^3+\phi_1^2\phi_2-1} \end{bmatrix} \begin{bmatrix} \sigma^2 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} &= \begin{bmatrix} \sigma^2 \frac{1-\phi_2}{(1+\phi_2)(1-\phi_1-\phi_2)(1+\phi_1-\phi_2)} \\ \sigma^2 \frac{\phi_1}{(1+\phi_2)(1-\phi_1-\phi_2)(1+\phi_1-\phi_2)} \\ \sigma^2 \frac{\phi_2+\phi_1^2-\phi_2^2}{(1+\phi_2)(1-\phi_1-\phi_2)(1+\phi_1-\phi_2)} \end{bmatrix} \end{aligned}$$

11 Exercícios

I. Questões do tipo falso, verdadeiro ou ambígua

1. O modelo $MA(1)$ é sempre estacionário.
2. A função de autocorrelação mede o grau de associação da série com o seu passado.
3. A função de autocorrelação parcial permite identificar se o modelo é do tipo AR ou MA.
4. O modelo $X_t = u_t + 0.5 * u_{t-1} * u_{t-2}$ com $u_t \sim NI(0, 1)$ é um ruído branco fraco, isto é, tem média e variância constante e é não correlacionado.
5. O modelo $AR(1)$ pode representar sómente séries estacionárias
6. O gráfico contra o tempo de uma série de tempo pode ser usado para determinar se o processo é estacionário ou não estacionário.
7. O modelo $X_t = 1.1X_{t-1} - .3X_{t-2} + u_t$ é não estacionário porque o coeficiente de X_{t-1} é maior do que um.
8. O modelo de nível local com erro observacional é adequado para séries $I(1)$.
9. O modelo de tendência local com erro observacional é adequado para séries $I(2)$.
10. O modelo de tendência local sem erro observacional é adequado para séries $I(1)$.

II. As primeira dez autocorrelações de uma série de tamanho 400 são: $r_1 = 0,02$, $r_2 = 0,05$, $r_3 = -0,09$, $r_4 = 0,08$, $r_5 = -0,02$, $r_6 = 0,00$, $r_7 = 0,12$, $r_8 = 0,06$, $r_9 = 0,02$, $r_{10} = -0,08$. Você acha que existe alguma evidência de correlação serial? Porque?

III. Os primeiros dez valores da FAC e FAC de uma série de tamanho 1000 são apresentados abaixo. Que modelo da classe *ARMA* voce acha adequado. Porque?

Defasagem	1	2	3	4	5	6	7	8	9	10
ρ_k	0.626	0.285	0.137	0.060	0.027	0.000	0.002	0.011	0.006	0.006
$\phi_{k,k}$	0.626	-0.176	0.063	-0.031	0.014	-0.028	0.030	-0.001	-0.006	0.007

12 Referências

Referências

- G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, 1976.
- G. E. P. Box and D. A. Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332):1509–1526, dec 1970. doi: 10.1080/01621459.1970.10481180.
- P. J. Harrison and C. F. Stevens. Bayesian forecasting. *Journal of the Royal Statistical Society. Series B (Methodological)*, 38(3):205–247, 1976.
- A. C. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, 1990. doi: 10.1017/cbo9781107049994.
- G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, aug 1978. doi: 10.1093/biomet/65.2.297.
- H. Migon and P. L. Valls Pereira. *Modelagem Estrutural: Abordagens Bayesiana e Clássica*. 1984.
- M. S. Morgan. *The history of econometric ideas*. CUP, 1991. ISBN 0521424658.
- M. H. Quenouille. Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society. Series B (Methodological)*, 11(1):68–84, 1949. ISSN 00359246. URL M.H.<http://www.jstor.org/stable/2983696>.