Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Elétrica

ENG 1469 - Análise de Séries Temporais

Notas de aula
(versão preliminar)

Prof. Cristiano Fernandes

2013

Introdução e Motivação

- Uma **Série Temporal** (ST) é um conjunto de observações, geralmente equiespaçadas, obtidas a partir da medição/observação de uma variável ao longo do tempo.

- A freqüência de observação/medição de uma ST pode variar dependendo do fenômeno observado: minuto, diária, semanal, mensal, anual etc.

- **Exemplos** de ST's:

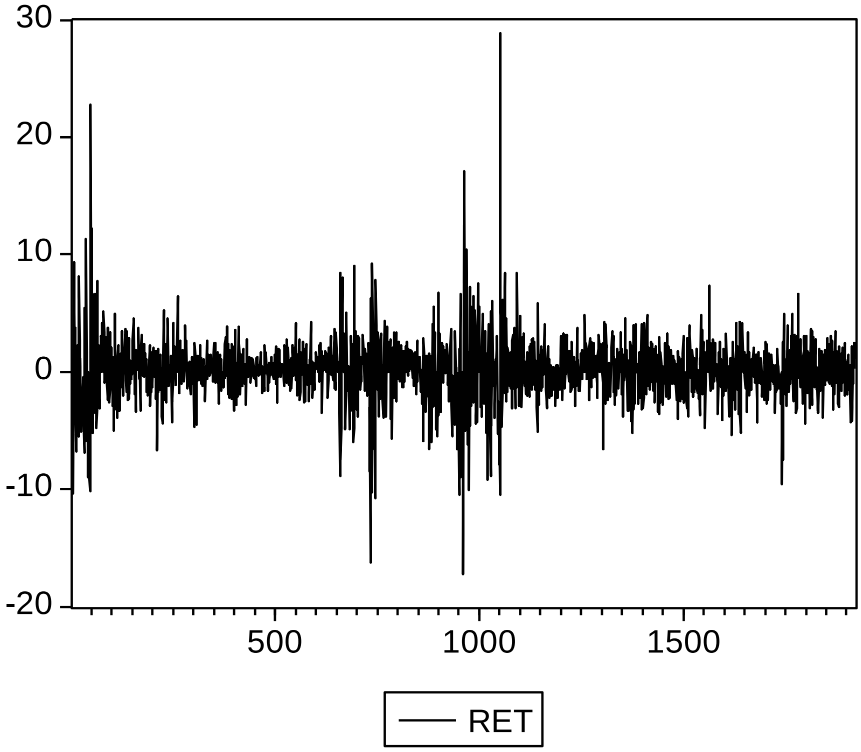
  - Retornos diários das cotações do IBOVESPA;

  - Taxa de desemprego mensal no Brasil;

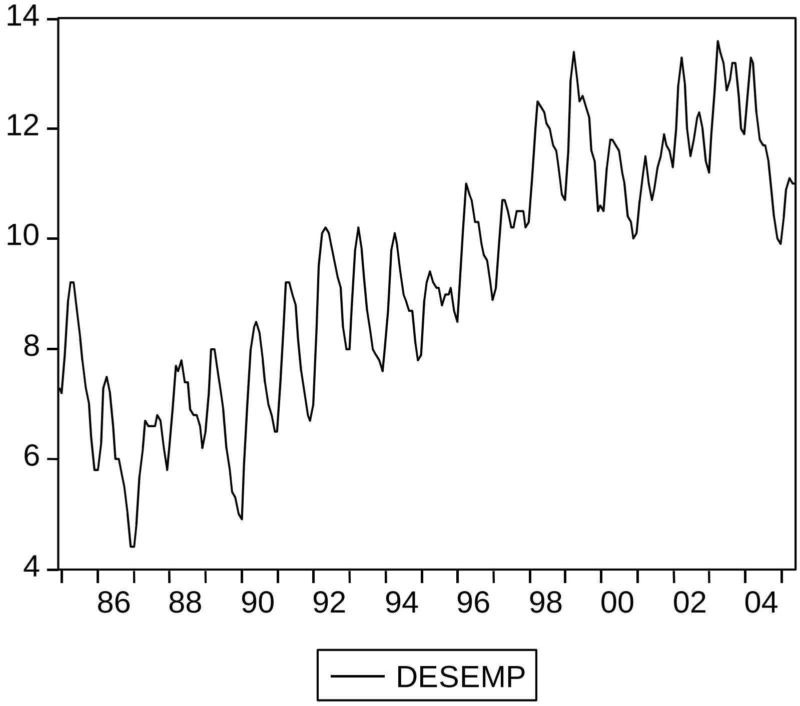
  - Venda mensal de aparelhos de ar condicionado;

  - Média dos índices diários de CO (monóxido de carbono) em São Paulo.

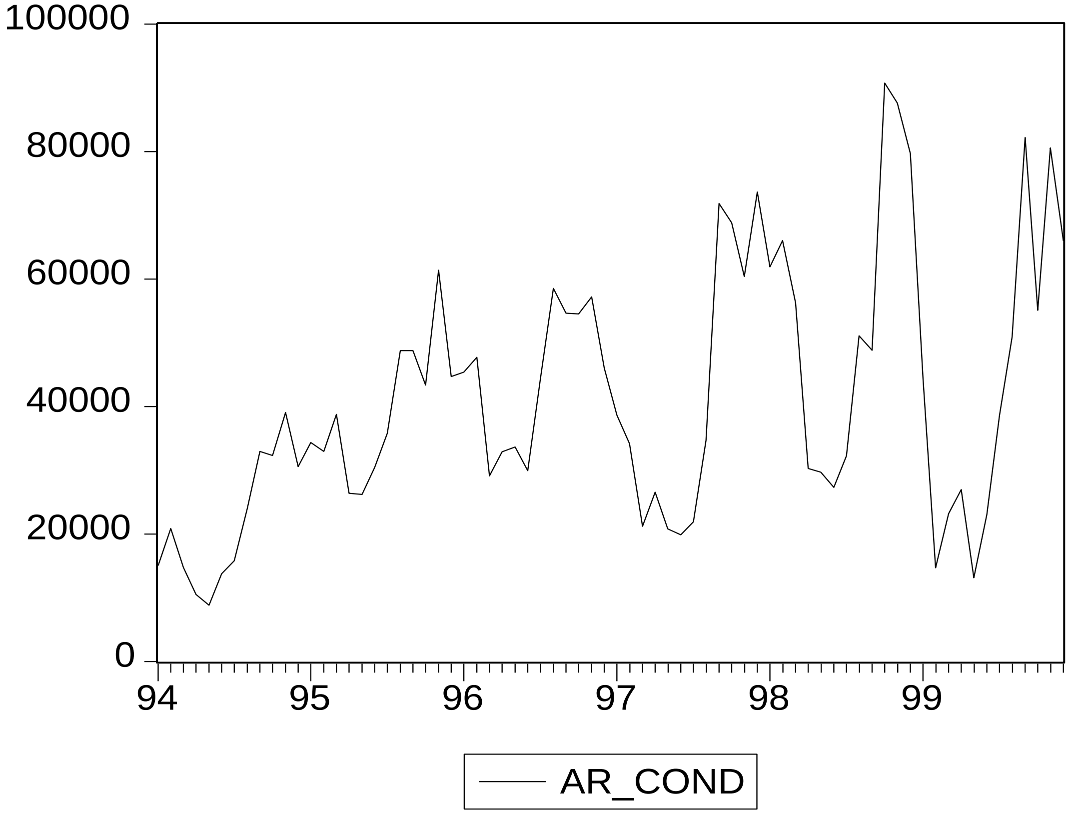
Exemplo1: Retornos diários do IBOVESPA (fechamento) no período de 02/01/1995 a 17/05/2002.



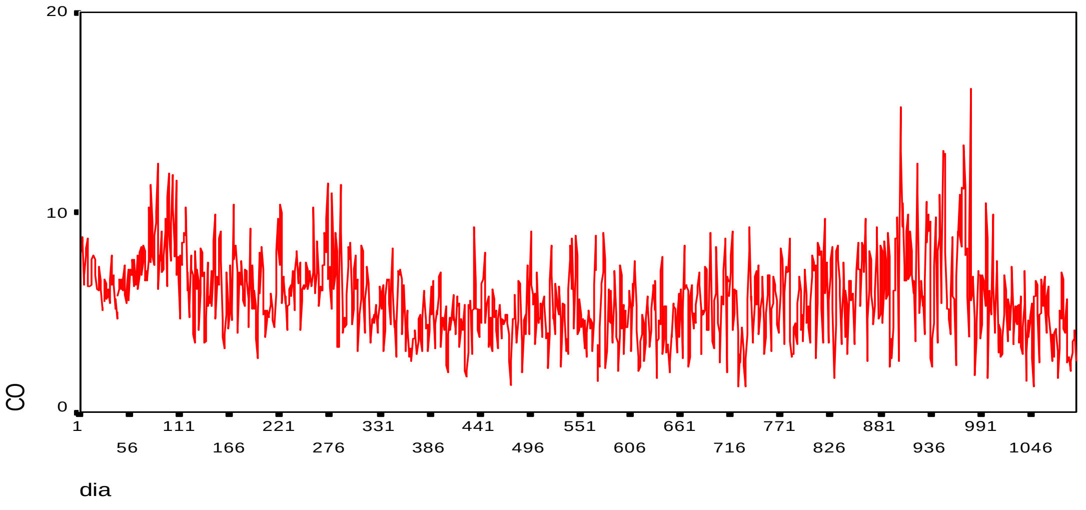
Exemplo 2: Taxa de desemprego aberto na RMSP (Seade/Dieese/PED, dez/1984 a jun/2005)



Exemplo 3: Venda mensal de aparelhos de ar condicionado (indústria, jan/1994 à dez/1999)



Exemplo 4: Média dos índices diários de CO (monóxido de carbono) em São Paulo (01/01/91 a 31/12/93; Tolerável = 9 ppm).



Análise de séries temporais: é a utilização de ferramentas estatísticas (FAC, periodograma, testes de dependência, testes de raiz unitária, etc) para revelar características importantes da série:

- as observações apresentam <u>autocorrelação</u> ?
- a série é <u>estacionária</u> ?
- a série é periódica/possui <u>sazonalidade</u> ?
- existe dependência <u>não-linear</u> ?

- A partir da análise de séries temporais, é possível obter subsídios para a escolha de um modelo adequado para modelar a série, escolhido dentro de uma classe de modelos pré-existentes.

- Neste curso iremos estudar duas classes de modelos univariados para séries temporais, a saber:

  1. **Modelos ARIMA**

     modelos para séries estacionárias/não estacionárias e lineares: AR, MA, ARMA, ARIMA, SARIMA

  2. **Modelos GARCH**

     modelos para séries de retornos financeiros utilizados para cálculo de risco de mercado.

- Uma vez construído, um modelo de ST pode ser utilizado para efetuar **previsões probabilísticas** sobre o futuro da série.

- A capacidade de realizar previsões é fundamental no processo de tomada de decisões em diversos contextos e em diversos lugares como órgãos públicos e empresas:

  - Em quanto % devemos aumentar a capacidade instalada da nossa fábrica de televisores digitais ?

  - Qual será a dimensão da nova estação de tratamento de água de Búzios?

  - Em quanto % devemos baixar a taxa de juros básica para que haja um aumento significativo nos empregos oferecidos pela indústria de bicicletas ?

- A formalização dos modelos de séries temporais é realizada a partir do conceito de **processo estocástico,** a ser visto na próxima seção.

Séries Temporais e Processos Estocásticos

- Um **processo estocástico** é uma função aleatória indexada pelo tempo t, y_t, em que para cada t, o valor de y_t é uma variável aleatória.

$$\{y_t : t \in T\}$$

y_t é o estado do processo no instante t.

- O conjunto T é o conjunto de índices ou espaço paramétrico do processo estocástico.

Processos Discretos e Contínuos

- Um processo estocástico $\{y_t:\ t \in T\}$ é um **processo contínuo** (ou processo com parâmetro contínuo) se o conjunto T é um **intervalo, finito ou não, de números reais.**

  Ex: série da pressão sangüínea para um paciente, monitorada continuamente por um aparelho, no período entre 17 e 18 hs de uma 4ª feira.

- Um processo estocástico $\{y_t:\ t \in T\}$ é um **processo discreto** (ou processo com parâmetro discreto) se o conjunto de índices T é um conjunto **contável**.

Ex: série da pressão sangüínea de uma paciente tomada à cada 15 minutos, num período de 1 semana.

Espaço de Estados

- **Espaço de estados** de um processo estocástico é o conjunto de todos os valores possíveis da variável aleatória y_t.

  pode ser **discreto** ou **contínuo**, se as variáveis aleatórias y_t forem discretas ou contínuas.

Exs:

1. **Discreto:** número de atendimentos diários no Barra d'Or devido ao consumo excessivo de álcool;

2. **Contínuo:** temperatura diária, as 21:00 h, no Alto da Boa Vista.

- Os modelos para processos estocásticos são específicos para a combinação entre o tipo de <u>espaço paramétrico</u> (discreto e contínuo) e o tipo do <u>espaço de estado</u> (discreto e contínuo).

- Neste curso iremos estudar modelos para processos estocásticos com **espaço paramétrico <u>discreto</u>** e **espaço de estado <u>contínuo</u>*.***

Ex: Série mensal de demanda residencial de energia elétrica na cidade de Rio de Janeiro entre 1990 e 2005.

Séries Temporais como realizações de Processos Estocásticos

- Um processo estocástico é um mecanismo gerador de dados (no tempo e/ou no espaço), cujo **comportamento não pode ser descrito por uma função determinística** no tempo**.**

- O comportamento futuro de um processo estocástico só pode ser descrito probabilisticamente.

- A estrutura probabilística de um processo estocástico pode ser completamente definida através da especificação da sua distribuição de probabilidade conjunta:

$$\mathrm{F}\left(y_{t_1}, y_{t_2}, \ldots, y_{t_T}\right) = \mathrm{Pr}\left(y_{t_1} \leq a_1, y_{t_2} \leq a_2, \ldots, y_{t_T} \leq a_T\right)$$

para qualquer subconjunto t_i, i=1,2,...,T.

- A distribuição de probabilidade conjunta geralmente é de difícil especificação.

- Assim sendo, caracterizamos um PE por uma equação estocástica **(modelo)**, a partir da qual pode-se obter a evolução de alguns dos seus momentos (como sua média, variância).

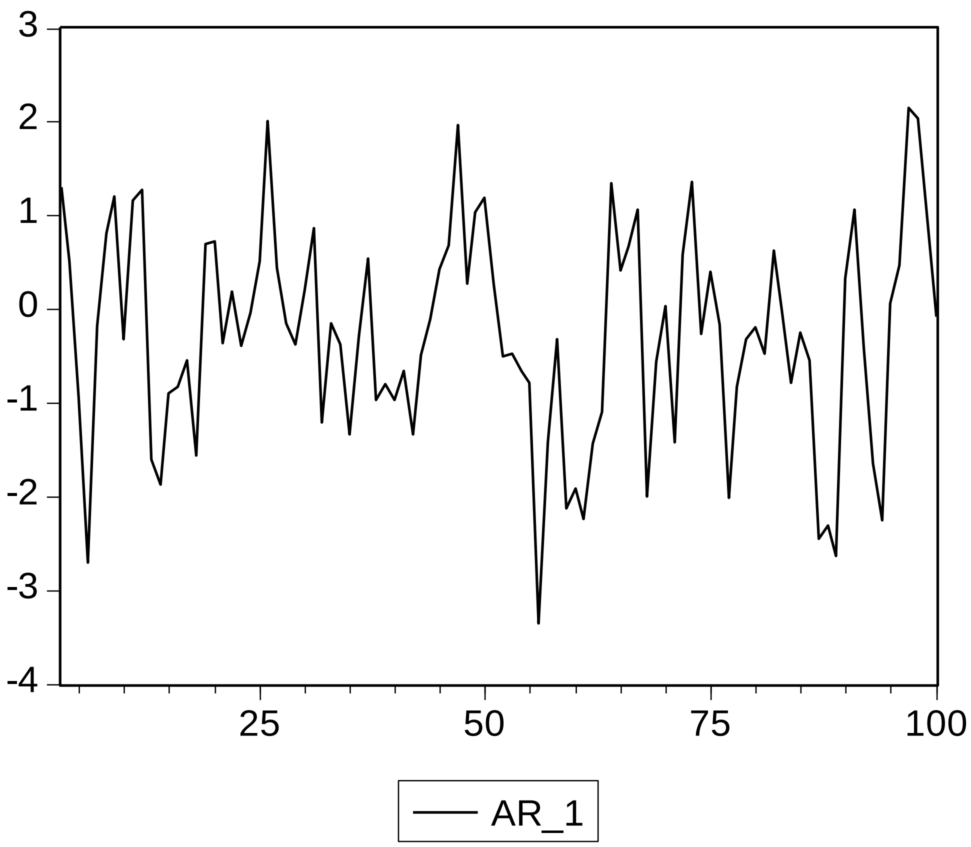
Ex:

Processo AR(1) (ou seja, ***auto-regressivo*** de 1a ordem)

$$y_t = 0.6\, y_{t-1} + e_t \qquad e_t \sim NID(0,1).$$

estrutura **termo aleatório**

Gráfico no tempo de um processo AR (1)



- No gráfico acima, cada valor de y_t é uma variável aleatória, e assim sendo, não pode ser completamente determinado a partir dos seus valores passados.

- Nos processos estocásticos, **o futuro é incerto!**

- Os modelos de séries temporais tentam capturar a estrutura de dependência existente no passado da série para, dentre outros objetivos, realizar previsões.

- Assim, o grau de acerto das previsões depende de dois fatores chaves:

  - Que o modelo seja uma boa aproximação da estrutura de dependência da série;

    $\Rightarrow$ **testes de especificação**

  - Que a estrutura de dependência identificada no passado permaneça razoavelmente estável no futuro.

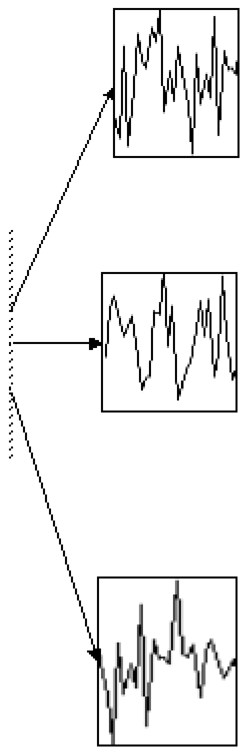
    $\Rightarrow$ **testes de previsão fora da amostra**

- Na prática somente observamos uma única realização de um PE, i.e., uma única ST (inflação, câmbio, etc).

Processo Estocástico **Realizações**

Processo AR(1)

$y_t = 0.6y_{t-1} + e_t \quad e_t \sim NID(0,1),$



- A partir de uma única realização, a análise de séries temporais procura inferir estatisticamente sobre o processo estocástico subjacente que deu origem à série.

- Para que isso seja possível, é necessário impor algumas hipóteses restritivas sobre o processo estocástico: **estacionariedade** e **ergodicidade**.

Estacionariedade

- A noção de estacionariedade está associada à idéia de "equilíbrio estatístico" no tempo: algumas das características probabilísticas do processo permanecem invariantes no tempo.

- Existem diversos dois tipos de estacionariedade:

  - Estacionariedade forte: a característica que permanece invariante no tempo é a **distribuição de probabilidade conjunta** do processo estocástico.

$$\mathrm{F}\left(y_{t_1}, y_{t_2} \ldots, y_{t_T}\right) = \mathrm{F}\left(y_{t_1+k}, y_{t_2+k} \ldots, y_{t_T+k}\right) \quad \forall \ t_1,\ t_2, \ldots, t_T \ \ e \ k$$

Esta condição é muito restritiva para processos reais, e assim, na prática, adota-se um tipo de estacionariedade mais branda.

  - Estacionariedade Fraca (ou estacionariedade de 2ª ordem): alguns dos **momentos (incondicionais)** do processo estocástico permanecem invariantes ao longo do tempo.

$$i. \ \mathrm{E}\left[y_t\right] = \mu \qquad \forall t$$

$$ii. \ \mathrm{E}\left[\left(y_t - \mu\right)\left(y_{t-k} - \mu\right)\right] = \gamma_k \qquad \forall t,$$

$$donde \ segue \ que:$$

$$\mathrm{E}\left[\left(y_t - \mu\right)^2\right] = \gamma_0 = \sigma_y^2 \qquad \forall t$$

- Assim sendo num PE de 2ª ordem:

  i. a média e a variância (incondicionais) do processo são constantes no tempo;

  ii. a estrutura de dependência linear γ_k depende apenas da distância entre as observações e neste caso pode-se mostrar que a intensidade da dependência diminui com esta distância.

- Os processos estocásticos são modelos matemáticos úteis, pois oferecem, muitas vezes, uma <u>descrição parcimoniosa</u> das séries reais (econômicas, físicas, biológicas, financeiras etc), as quais podem ser utilizadas para simulação e previsão de uma ST.

- Deve-se, entretanto, estar atentos às simplificações implicadas por estas descrições matemáticas:

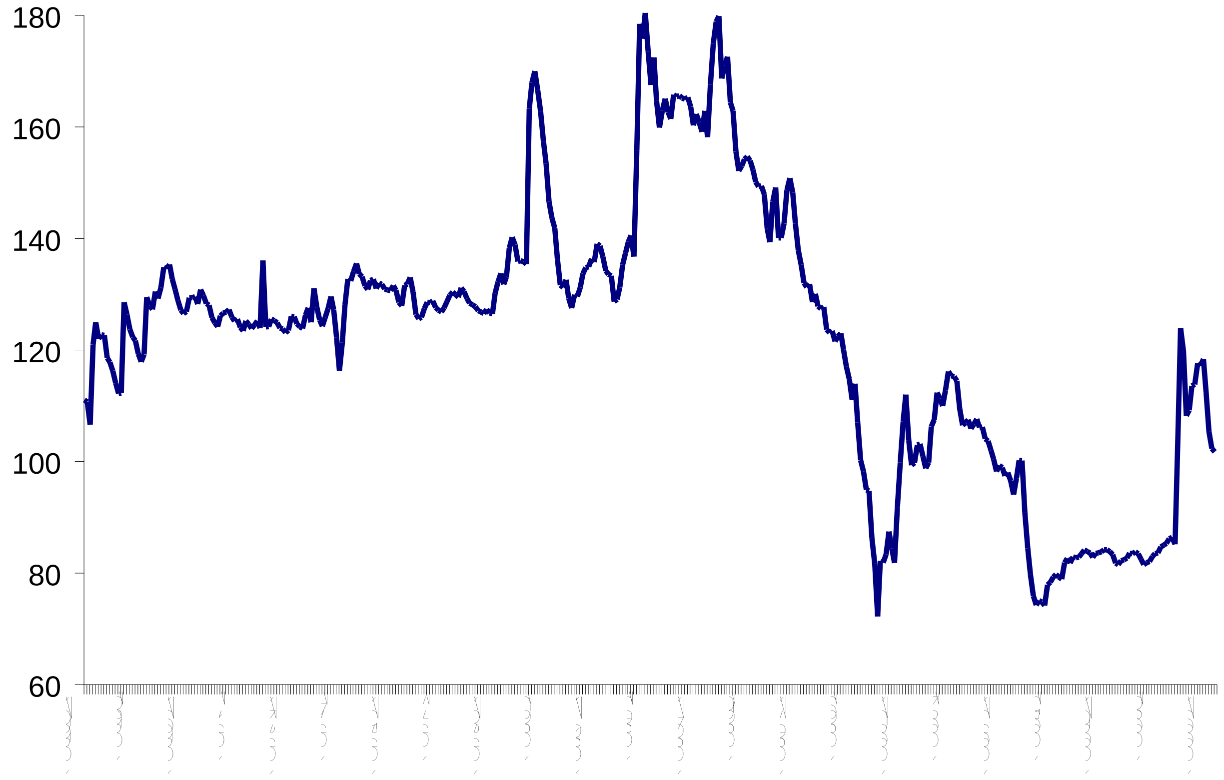
“All models are wrong, but some are useful.”

G.E.P. Box

- As realidades econômicas, físicas, biológicas, que são o verdadeiro "processo" gerador das séries temporais, podem ser <u>extremamente dinâmicas e mutáveis</u>.

- A mudança do processo gerador pressuposto "sob" a série observada é chamada **quebra estrutural.**

Exemplo clássico de quebra estrutural: Taxa de câmbio real R$/US$.



- Muitas séries reais são não-estacionárias, pois apresentam tendência no tempo.

- Como veremos se uma ST for originalmente não-estacionária temos que torná-la estacionária para que possamos utilizar modelos estacionários.

- Existem dois tipos de tendência:

  - Tendência Determinística: o modelo da tendência é uma função determinista do tempo (TS):

Caso particular: considere uma tendência linear sobre um processo a_t ~AR (1):

$$y_t = \beta_0 + \beta_1 t + a_t,$$
$$a_t = \phi a_{t-1} + \varepsilon_t, \quad |\phi| < 1, \quad \varepsilon_t \sim NID(0, \sigma^2)$$

Segue que:

$$E(y_t) = \beta_o + \beta_1 t$$
$$Var(y_t) = \sigma^2 / (1 - \phi^2)$$
$$\rho(y_t \; y_{t-k}) = \phi^k$$

- Portanto o processo não satisfaz as condições de estacionariedade de 2ª ordem, e assim é não estacionário.

- Se uma ST tem tendência deste tipo, devemos ajustar o modelo e trabalhar com o resíduo, o qual será, por construção estacionário.

$$\hat{\varepsilon}_t = y_t - [(\hat{\beta}_0 + \hat{\beta}_1 t) + \hat{\phi}\, a_{t-1}]$$

- Tendência Estocástica: o modelo da tendência é uma função estocástica (DS).

- Exemplo canônico de modelo com tendência estocástica

y_t ~ RW + drift (passeio aleatório + constante)

$y_t = a + y_{t-1} + e_t, \quad e_t \sim NID(0, \sigma^2)$

$y_t = at + y_o + \sum_{i=1}^{t} e_i$

- Segue que:

$E(y_t) = at + y_o$

$Var(y_t) = t\,\sigma^2$

$\rho(k) = [1 - (k / t)]^{1/2}$

- Os momentos variam explicitamente no tempo, e assim sendo o PE **não é** estacionário de 2ª ordem.

- Se uma ST tem tendência estocástica então para torná-la estacionária basta tomar a 1ª diferença da série

$$z_t = y_t - y_{t-1} = a + e_t,$$

Ergodicidade

- Ergodicidade na média: a média amostral das observações de uma realização converge para a média do processo:

$$\frac{1}{T}\sum_{t=1}^{T} y_t \xrightarrow{p} \mu$$

- Ergodicidade de 2a. ordem: a covariância amostral converge para a covariância do processo:

$$\frac{1}{T-k}\sum_{t=k+1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y}) \xrightarrow{p} \gamma_j$$

- Todo processo ergódico é estacionário, mas nem todo processo estacionário é ergódico.

Investigando Estacionariedade

- A hipótese de estacionariedade de uma série tem papel fundamental no desenvolvimento da teoria que será apresentada durante o curso, pois a classe dos modelos aqui estudados foi desenvolvida para PE estacionários.

- Dado um PE, utilizando a definição de PE estacionário, podemos saber se este processo é estacionário ou não.

- O que precisamos é de ferramentas estatísticas para investigar a estacionariedade de uma <u>série temporal</u>.

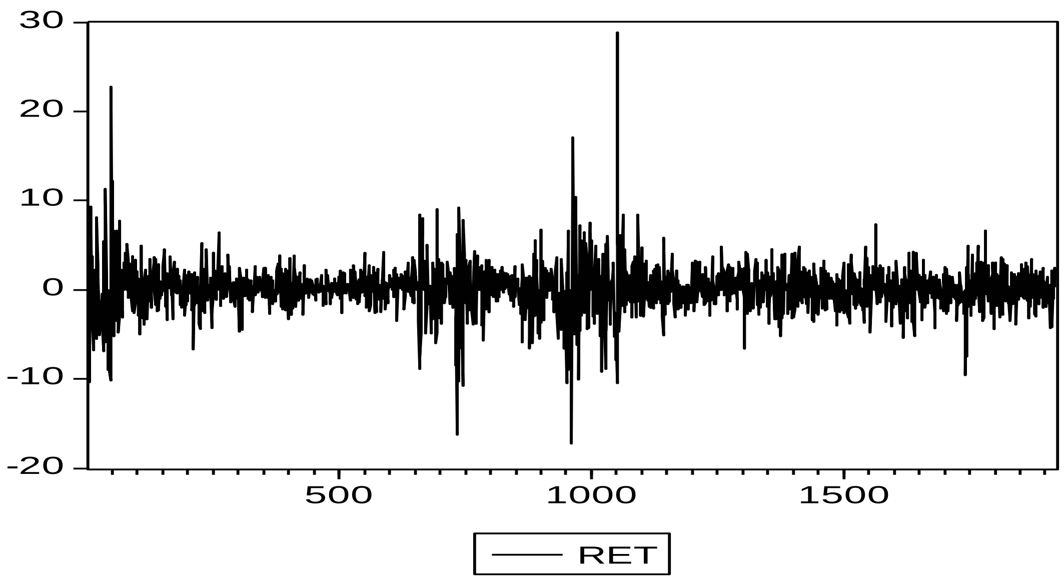
- Esta investigação será realizada através das seguintes ferramentas:

  - gráfico da série no tempo;
  - FAC;
  - Testes de raiz unitária/ não estacionariedade.

Gráfico da série no tempo

- Em uma **série estacionária**, as observações variam em torno de um valor de equilíbrio, a média (incondicional).

Série de retornos do Ibovespa



Fato estilizado: séries de retornos financeiros são em geral estacionárias !

- A maioria das séries econômicas e financeiras que não são estacionárias possuem tendência estocástica.

- Assim sendo para torná-las estacionárias deveremos diferencia-las um número adequado de vezes.

A Função de Autocorrelação (FAC) e a Análise do Correlograma

- **Função de autocorrelação (FAC)** = representa uma medida de associação linear entre o presente do processo e o seu passado, ou seja, uma medida da "memória" do processo.

- A FAC é o procedimento padrão para investigar a dependência linear implicada por um modelo de ST.

- Em geral, autocorrelação nula para todo k em um PE não implica que as observações deste processo sejam independentes entre si, podendo existir algum outro tipo de dependência que não seja linear.

- Entretanto, se um PE for independente, a sua FAC, e qualquer outra medida de auto-associação, será nula.

- A função de autocorrelação (FAC) ou correlograma de um determinado PE é obtida através do cálculo da seguinte expressão:

$$\rho_k = \gamma_k \ / \ \gamma_0 , \qquad k = 1,2,3,...$$

autocovariância variância lags

- Utilizando a fórmula matemática do processo podemos então para cada k calcular o numerador

$$\rho_k = \gamma_k \ / \ \gamma_0$$
$$\gamma_k = E(y_t y_{t\text{-}k}) - E(y_t)E(y_{t\text{-}k}), \ k = 0,1,2,...$$

- **Propriedades da FAC**

i. $\rho_0 = 1;$

ii. $-1 \leq \rho_k \leq 1;$

ii. $\rho_k = \rho_{-k}, \ \ p/ \ \ y_t$ Real.

- **Ex**: y_t= a cos(wt) + b sin(wt), onde

Onde a e b são variáveis aleatórias com E(a)=E(b)=0, E(ab)=0, $E(a^2)= E(b^2)=\sigma^2$.

Segue que:

$E(y_t)=0$, $Var(y_t)= \sigma^2$
$E(y_t\ y_{t-k})= \sigma^2 cos(wk)$, utilizando que

cos(a-b)=cosa cosb + sina sin b

Portanto ρ_k= cos(wk), k=1,2,3,...

Assim sendo, o processo é estacionário de 2ª ordem.

- Grosso modo, pode-se dizer que a FAC de um processo estocástico estacionário é uma "assinatura" deste processo. Processos de uma mesma "família" geram FAC's do mesmo tipo.

- Portanto, a FAC pode ser utilizada como um procedimento para identificar o processo estocástico que gera uma série temporal estacionária.

- Observe que:

- se houver uma relação linear perfeita entre y_t e y_{t-k}, então $| \rho_k | =1$;

- mesmo que $\rho_k = 0$, para todo k, ainda assim pode haver algum outro tipo de relação entre y_t e y_{t-k} (relação não-linear).

Identificando Estacionariedade segundo o correlograma

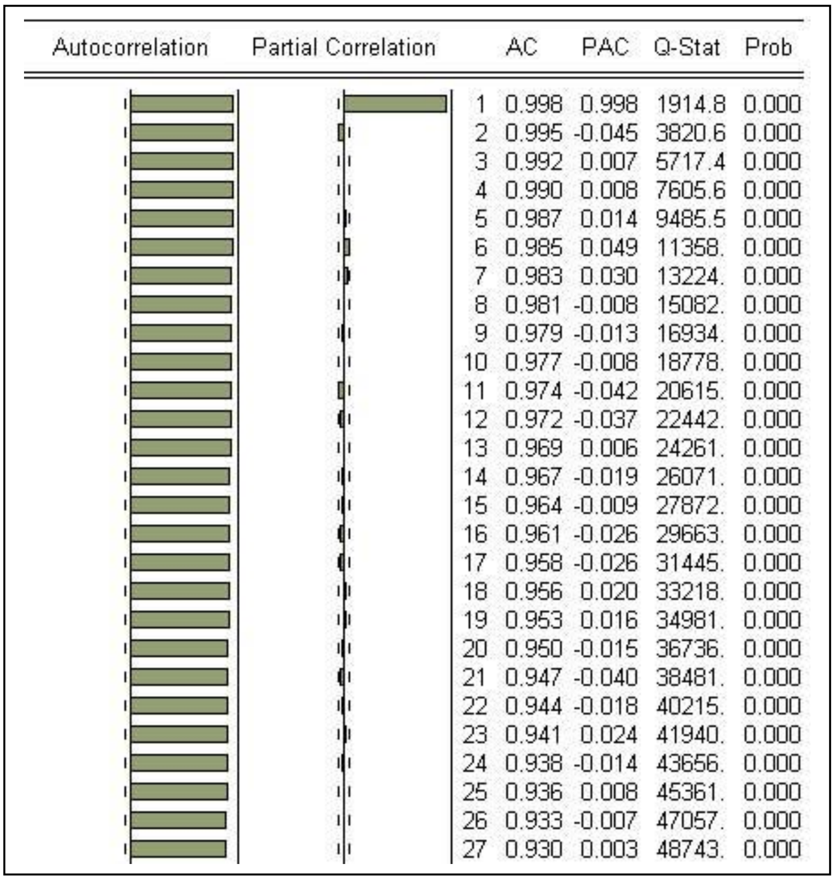
- Para uma **série estacionária**, o correlograma decai rapidamente para zero com o aumento de k, ou é estatisticamente nulo para todo k (ruído branco).

Ex: Correlograma dos **retornos** do Ibovespa.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.045	0.045	3.9799	0.046
		2	-0.006	-0.008	4.0515	0.132
		3	-0.048	-0.047	8.4948	0.037
		4	-0.039	-0.035	11.396	0.022
		5	-0.065	-0.062	19.504	0.002
		6	-0.078	-0.076	31.367	0.000
		7	0.037	0.040	34.049	0.000
		8	0.023	0.012	35.087	0.000
		9	0.035	0.023	37.485	0.000
		10	0.077	0.070	48.865	0.000
		11	0.043	0.033	52.413	0.000
		12	-0.048	-0.048	56.878	0.000
		13	-0.013	0.009	57.190	0.000
		14	-0.021	-0.009	58.013	0.000
		15	0.036	0.047	60.589	0.000
		16	-0.016	-0.010	61.096	0.000
		17	0.000	-0.006	61.096	0.000
		18	-0.003	-0.014	61.111	0.000
		19	0.029	0.029	62.775	0.000
		20	0.033	0.027	64.875	0.000
		21	0.024	0.025	65.989	0.000
		22	-0.013	-0.012	66.300	0.000
		23	0.006	0.015	66.377	0.000
		24	-0.015	-0.013	66.831	0.000
		25	-0.027	-0.023	68.279	0.000
		26	-0.024	-0.020	69.386	0.000
		27	0.009	0.016	69.562	0.000

- Em uma **série não estacionária**, o correlograma decai lentamente para zero com o aumento de k. (Por que ?)

Ex: correlograma da série do **índice** Ibovespa.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.998	0.998	1914.8	0.000
		2	0.995	-0.045	3820.6	0.000
		3	0.992	0.007	5717.4	0.000
		4	0.990	0.008	7605.6	0.000
		5	0.987	0.014	9485.5	0.000
		6	0.985	0.049	11358.	0.000
		7	0.983	0.030	13224.	0.000
		8	0.981	-0.008	15082.	0.000
		9	0.979	-0.013	16934.	0.000
		10	0.977	-0.008	18778.	0.000
		11	0.974	-0.042	20615.	0.000
		12	0.972	-0.037	22442.	0.000
		13	0.969	0.006	24261.	0.000
		14	0.967	-0.019	26071.	0.000
		15	0.964	-0.009	27872.	0.000
		16	0.961	-0.026	29663.	0.000
		17	0.958	-0.026	31445.	0.000
		18	0.956	0.020	33218.	0.000
		19	0.953	0.016	34981.	0.000
		20	0.950	-0.015	36736.	0.000
		21	0.947	-0.040	38481.	0.000
		22	0.944	-0.018	40215.	0.000
		23	0.941	0.024	41940.	0.000
		24	0.938	-0.014	43656.	0.000
		25	0.936	0.008	45361.	0.000
		26	0.933	-0.007	47057.	0.000
		27	0.930	0.003	48743.	0.000

- **Testes de raiz unitária**

- Grande revolução na Econometria na década de 80!

- As seguintes opções de teste de RU estão disponíveis no EViews (a partir da versão 4.0):

  - Augmented Dickey Fuller (ADF)
  - Dickey Fuller GLS (DF-GLS)
  - Phillips-Perron (PP)
  - Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
  - Elliot–Rothenberg-Stock point optimal (ERS)
  - Ng-Perron (NP)

  - Em todos estes testes a hipótese nula é que a série possui uma raiz unitária (a ST é não-estacionária) e a hipótese alternativa a que a série é estacionária, com exceção do teste KPSS, no qual ocorre o contrário.

Ho: a ST possui uma RU « » a série é não estacionária
Ha: a ST não possui RU « » a série é estacionária

- Existem várias modelos compatíveis com um teste de RU. A equação mais simples possui a forma:

$$Y_t = \varphi Y_{t-1} + e_t \quad (I)$$

$$Ho: \varphi = 1 \rightarrow \text{a ST possui uma RU}$$
$$Ha: \varphi < 1 \rightarrow \text{a ST é estacionária}$$

- Como é mais usual testar se o parâmetro é nulo, o teste de RU é efetuado subtraindo-se Y_{t-1} de ambos os lados da eq. I, resultando em:

$$\Delta Y_t = a_1 Y_{t-1} + e_t \quad (II)$$

$$a_1 = \varphi - 1$$

- Portanto nesta nova parametrização o teste será:

$$Ho: a_1 = 0 \rightarrow \text{a ST possui uma RU}$$
$$Ha: a_1 < 0 \rightarrow \text{a ST é estacionária}$$

- Um pouco de reflexão indica que existem outras equações compatíveis com RU. Por exemplo, adicionando-se um intercepto a_0 ou um intercepto mais uma tendência determinística, $a_0 + a_1 t$, resultam em processos adequados para testar RU.

Formas do teste de Raiz Unitária

- Como existem várias especificações consistentes com não-estacionariedade ou RU, vão existir várias formas de testar RU.

- A questão importante na prática é escolher a forma do teste de RU adequada para a ST em questão.

- As seguintes formas para teste de RU se apresentam:

- DF I: $\Delta Y_t = a_1 Y_{t-1} + e_t$, $e_t \sim NID(0,\sigma^2)$

Ho: $a_1=0 \rightarrow Y_t \sim$ não-estacionária (RW)
Ha: $a_1< 0 \rightarrow Y_t \sim$ estacionária

- DF II: $\Delta Y_t = a_o + a_1 Y_{t-1} + e_t$, $e_t \sim NID(0,\sigma^2)$

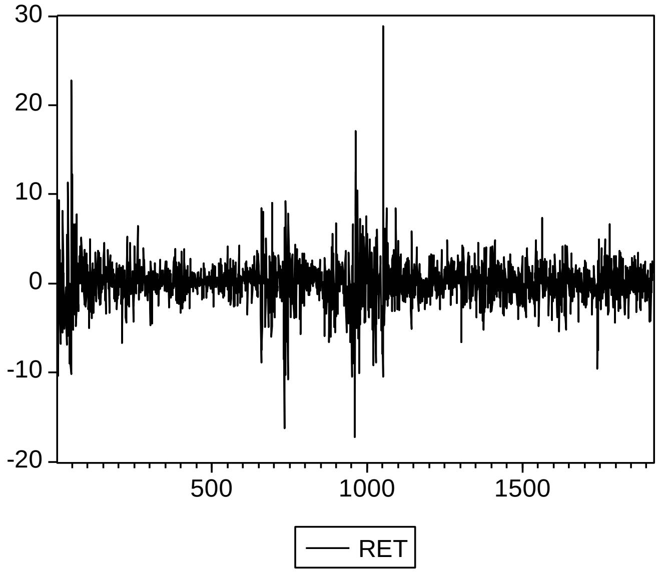
Ho: $a_1=0 \rightarrow Y_t \sim$ RW + drift
Ha: $a_1< 0 \rightarrow Y_t \sim$ estacionária

- DF III: $\Delta Y_t = a_o + a_1 Y_{t-1} + a_2\, t + e_t$, $e_t \sim NID(0,\sigma^2)$

Ho: $a_1=0 \rightarrow Y_t \sim$ RW + drift em torno de uma tendência determinística.
Ha: $a_1< 0 \rightarrow Y_t \sim$ estacionária (supondo $a_2=0$)

- O procedimento prático de testar RU é escolher uma das formas, (i, ii ou iii) que acomode, simultaneamente, as hipóteses nula e alternativa consistente com a série sendo testada.

- Por exemplo, considere a série de retornos do Ibovespa:



- É óbvio que a série é estacionária com média praticamente nula, e assim sendo a forma adequada para o teste será a I.

O Teste ADF

- No teste ADF, para corrigir a possível presença de autocorrelação dos resíduos nas equações I, II ou III, é adicionado o seguinte somatório em cada uma destas equações:

$$\sum_{i=1}^{P} \beta_i \Delta Y_{t-i}$$

transformando-as em ADF-I, ADF-II e ADF-III, respectivamente:

- ADF-I: $\Delta Y_t = a_1 Y_{t-1} + \sum_{i=1}^{P} \beta_i \Delta Y_{t-i} + e_t$

- ADF-II: $\Delta Y_t = a_o + a_1 Y_{t-1} + \sum_{i=1}^{P} \beta_i \Delta Y_{t-i} + e_t$

- ADF-III: $\Delta Y_t = a_o + a_1 Y_{t-1} + \sum_{i=1}^{P} \beta_i \Delta Y_{t-i} + a_2 t + e_t$

- A ordem p, da autoregressão, é geralmente determinada automaticamente por um dos critérios de informação (AIC, BIC etc).

- A estatística de teste para o teste ADF é:

$$t = (\hat{a}_1 - 0) / se(\hat{a}_1)$$

- Problema: a distribuição desta estatística não é mais a t de Student !

- Possui uma forma não-padrão, obtida por simulação Monte Carlo: Dickey & Fuller (79) e MacKinnon (91).

- Regra de decisão: aceite a Hipótese Nula ao nível de α% se:

$$| t | =< \tau_i^{\alpha\%} \quad i=1,2,3 \quad \text{(valores tabelados)}$$

- Ou seja, o valor crítico da estatística de teste muda com a forma do teste de RU: I, II ou III.

Teste PP

- Para o teste PP, utiliza-se uma correção na estatística de teste baseado num ajuste "não paramétrico" na forma desta estatística, o qual corrige, simultaneamente, a presença de heterocedasticidade e autocorrelação nos resíduos.

- Este ajuste necessita da especificação de um parâmetro *(truncation lag)* o qual aparece na determinação da estimativa da variância pela estimativa da densidade espectral na freqüência zero.

- As formas das equações permanecem inalteradas, ou seja, não introduzimos o "somatório", como em ADF. Por exemplo, para o modelo II a estatística de teste torna-se:

$$t_{PP} = t_2\, \gamma_0^{1/2} / \omega - [\, (\omega^2 - \gamma_0)\, T\, se(\hat{a}_1)] \;/\; (2\omega\sigma')$$

$$\omega^2 = \gamma_0 + 2\sum_{j=1}^{q} (1 - j/(q+1))\gamma_j \;;$$

$$\gamma_j = (1/T) \sum_{t=j+1}^{T} \hat{e}_t\, \hat{e}_{t-j} \;;$$

- T é o tamanho da série;
- σ' é a estimativa de MQO de σ;
- q é o *truncation lag.*

- Os valores críticos permanecem o mesmo dos testes ADF-I, ADF-II e ADF-III (assintoticamente).

- **Exemplos dos testes de Raiz Unitária**

  - IBOVESPA: testado nas estrutura III e II.
  - Retornos do IBOVESPA: testado na estrutura I.

	Teste		valores críticos (*)		
	ADF	PP	**1%**	**5%**	**10%**
Ibov	-2.908 (iii) -1.643 (ii)	-2.17 -1.66	**-3.97** **-3.44**	**-3.41** **-2.86**	**-3.13** **-2.57**
Ret Ibov	-41.88 (i)	-41.84	**-2.57**	**-1.94**	**-1.62**

*MacKinnon critical values for rejection of hypothesis of a unit root.

⇒ Conclusão: aceita não-estacionariedade p/ série IBOVESPA e rejeita não-estacionariedade p/ série de retornos do IBOVESPA.

OBS: sabendo-se de antemão que as séries financeiras apresentam heterocedasticidade condicional, o teste de RU mais adequado seria o PP, o qual corrige por heterocedasticidade.

Modelos ARMA para Séries Temporais Estacionárias

- ***Teorema da Decomposição de Wold =*** qualquer processo estocástico estacionário y_t pode ser escrito na forma

$$y_t = m_t + v_t$$

onde:

- m_t e v_t são processos não-correlacionados;
- m_t é um processo determinístico;
- v_t é definido por

$$v_t = \sum_{i=0}^{\infty} \psi_i \, \varepsilon_{t-i}, \quad \psi_0 = 1, \quad \sum_{i=0}^{\infty} \psi_i^2 < \infty,$$

onde ε_t é um ruído branco.

- <u>Ruído branco</u> é um PE estacionário onde as observações são descorrelatadas, ou seja, o processo não possui “memória” (linear).

- É representado usualmente pelo símbolo ε_t, e tem uma importância fundamental na construção dos modelos para processos com dependência.

  Seus dois primeiros momentos são dados por:

$$E[\varepsilon_t] = 0$$

$$E[\varepsilon_t \, \varepsilon_{t-k}] = \begin{cases} 0, & k \neq 0 \\ \sigma^2, & k = 0 \end{cases}$$

- Note que se o processo segue a distribuição normal, além de ser descorrelatado ele também será independente (pq ?).

- Particularizando o termo determinístico como uma constante (*drift*), temos que:

$$y_t = m + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$$

- Esta é conhecida como a <u>representação linear</u> de um processo estacionário de 2ª ordem.

- Utilizando operador de atraso **L** ($L^k y_t = y_{t-k}$), podemos reescrever a representação geral de um PE estacionário como:

$$y_t = m + \Psi(L)\,\varepsilon_t$$

$$\Psi(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \ldots,\ \psi_0 = 1$$

- polinômio $\Psi(L)$, de grau infinito, poder ser obtido pela razão entre dois polinômios de grau finito, isto é:

$$\Psi(L) = \frac{\Theta_q(L)}{\Phi_p(L)}$$

$$\Theta_q(L) = 1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q$$

$$\Phi_p(L) = 1 - \varphi_1 L - \varphi_2 L^2 - ... - \varphi_p L^p$$

- Substituindo esta expressão na expressão para y_t, obtemos:

$$y_t = m + \Psi(L)\,\varepsilon_t$$

$$= m + \frac{\Theta_q(L)}{\Phi_p(L)}\,\varepsilon_t, \; ou$$

$$\Phi_p(L)\, y_t = c + \Theta_q(L)\varepsilon_t$$

em que:

- $c = (1 - \varphi_1 - \varphi_2 - ... - \varphi_p)m$
- $\Theta_q(L) = 1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q$
- $\Phi_p(L) = 1 - \varphi_1 L - \varphi_2 L^2 - ... - \varphi_p L^p$

- Utilizando a definição dos operadores de atraso, chegamos à seguinte expressão:

$$y_t = c + \varphi_1 y_{t-1} + \ldots + \varphi_p y_{t-p} +$$
$$\varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$$

- Os termos c, ϕ_1, ..., ϕ_p,θ_1, ..., θ_q são <u>parâmetros desconhecidos</u> que devem ser estimados a partir dos dados de uma série temporal real.

- Este é denominado um modelo **ARMA(p,q)**, onde:

  - a componente AR(*auto regressivo*): refere-se aos valores defasados de y_t.

  - a componente MA (*moving average*) refere-se aos valores presente e defasados do ruído branco.

- O nosso estudo será iniciado pela investigação das <u>propriedades estatísticas</u> dos modelos AR puro, em seguida MA puro e por fim os modelos mistos ARMA.

Modos condicional e Incondicional de uma série temporal estacionária

- Inicialmente, é importante distinguir entre os modos condicional e incondicional de um modelo para ST.

- Um mesmo modelo de ST apresenta um modo condicional e um modo incondicional.

- O **modo condicional** explora a dependência existente na série, sendo utilizado para:

  - previsão de curto prazo;

  - estimação dos parâmetros do modelo;

  - diagnósticos do modelo;

- Por sua vez, o **modo incondicional** é utilizado para:

  - checar se a distribuição implicada pelo modelo é adequada para os dados (a série temporal);

  - checar se a FAC implicada pelo modelo tem a mesma forma do correlograma obtido a partir dos dados;

  - estabelecer as condições de estacionariedade do modelo, em termos de restrições nos seus parâmetros;

  - obter a previsão de longo prazo para a série, a partir do modelo.

Processos AR (autoregresivos)

Processo AR(1)

- Considere o processo AR (1)

$$y_t = c + \varphi_1 y_{t-1} + \varepsilon_t$$

em que ε_t é um ruído branco Gaussiano $N(0,\sigma^2)$.

- Obtendo os seus momentos de acordo com o modo ***condicional***, tem-se:

  - $E(y_t \mid \mathbf{Y_{t-1}}) = c + \phi_1 y_{t-1}$
  - $Var(y_t \mid \mathbf{Y_{t-1}}) = \sigma^2$
  - $f(y_t \mid \mathbf{Y_{t-1}}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$

- Um processo AR(1), na sua forma incondicional, é, do ponto de vista matemático, uma equação de diferenças de 1ª ordem, não homogênea.

- A solução desta equação pode ser obtida através de dois modos distintos:

  i. por iteração da equação;

  ii. pela solução formal da equação de diferença

i. **Por iteração da equação**:

$$y_1 = c + \varphi_1 y_0 + \varepsilon_1$$

$$y_2 = c + \varphi_1 y_1 + \varepsilon_2 = c + \varphi_1 (c + \varphi_1 y_0 + \varepsilon_1) + \varepsilon_2$$

$$y_t = c \sum_{i=0}^{t-1} \varphi_1^i + \varphi_1^t \, y_0 + \sum_{i=0}^{t-1} \varphi_1^i \varepsilon_{t-i}$$

$$E(y_t) = c \sum_{i=0}^{t-1} \varphi_1^i + \varphi_1^t \, E(y_0) + \sum_{i=0}^{t-1} \varphi_1^i \, E(\varepsilon_{t-i})$$

$$= c \, (1 - \varphi_1^t) / (1 - \varphi_1) + \varphi_1^t \, y_0$$

$$E(y_t) = c \, / (1 - \varphi_1), \text{ se } |\varphi_1| < 1, \, t \to \infty.$$

$$Var(y_t) = E(y_t - E(y_t))^2 = E(\sum_{i=0}^{t-1} \varphi_1^i \varepsilon_{t-i})^2$$

$$= E(\sum_{i=0}^{t-1} \varphi_1^{2i} \varepsilon_{t-i}^2 + 2 \sum_{i<j}^{t-1} \varphi_1^i \varphi_1^j \, \varepsilon_{t-i} \varepsilon_{t-j})$$

$$= \sigma^2 \sum_{i=0}^{t-1} \varphi_1^{2i} = \sigma^2 (1 - \varphi_1^{2t}) / (1 - \varphi_1^2)$$

$$Var(y_t) = \sigma^2 / (1 - \varphi_1^2), \text{ se } |\varphi_1| < 1, \, t \to \infty$$

- Tem-se, portanto, que a **distribuição incondicional de y_t** é **combinação linear** de v.a's **normais**, e assim sendo também será normal com:

$$f(y_t) \sim N\left(c / (1 - \phi), \frac{\sigma^2}{1 - \varphi_1^2}\right)$$

ii. Por solução formal da equação: a solução desta equação é dada pela soma da solução homogênea (y_t^h) mais a solução particular (y_t^p):

$$y_t = y_t^h + y_t^p$$

- a solução homogênea y_t^h é obtida fazendo todos os termos que não envolvem y_t igual a zero e achando a raiz do polinômio em z, z=1/L.

$$y_t = \varphi_1 y_{t-1}$$
$$(1 - \varphi_1 L)\, y_t = 0$$
$$(1 - \varphi_1 L) = 0 \therefore z = 1/L = \varphi_1$$

Observar que $y_t^h = A\, \varphi_1^t$ é solução

- a solução particular é obtida a equação em termos do operador de atraso (L):

$$y_t = c + \varphi_1 y_{t-1} + \varepsilon_t$$
$$(1 - \varphi_1 L)\, y_t = c + \varepsilon_t$$
$$y_t^p = (1 - \varphi_1 L)^{-1} c + (1 - \varphi_1 L)^{-1} \varepsilon_t$$

- Portanto a solução geral da equação será dada por:

$$y_t = y_t^h + y_t^p$$

$$y_t = A\varphi_1^t + (1 - \varphi_1 L)^{-1} c + (1 - \varphi_1 L)^{-1} \varepsilon_t$$

- Assumindo que $|\varphi_1| < 1$ e que $t \to \infty$ (o que deduziremos como condições suficientes para que um processo AR(1) seja assintoticamente estacionário), tem-se que:

$$y_t = \sum_{j=0}^{\infty} \varphi_1^j L^j c + \sum_{j=0}^{\infty} \varphi_1^j L^j \varepsilon_t$$

$$y_t = c \,/\, (1-\varphi_1) + \sum_{j=0}^{\infty} \varphi_1^j \varepsilon_{t-j}$$

- Nestas condições obtemos os mesmos resultados da solução por iteração para a média, variância e autocovariância.

- Observe que a partir da última expressão derivada para y_t ,**um processo AR (1) pode ser reescrito como um processo MA(∞).** Este é um resultado geral para processos auto-regressivos e que será explorado mais adiante.

A função de autocorrelação para um processo AR (1)

- A FAC é a "assinatura" de um PE, e deve ser sempre calculada, pois será utilizada para identificar o melhor modelo para uma dada série temporal.

$$\rho(k) = E\left[\left(y_t - \mu\right)\left(y_{t-k} - \mu\right)\right] / \mathrm{Var}(y_t),\ k = 1,2,3,...$$
$$= \gamma(k) / \gamma(0), \quad \gamma(0) = \mathrm{Var}(y_t).$$

- Para calcularmos esta expressão, tomamos a equação do processo AR (1), utilizando que $c = \mu\,(1-\varphi_1)$,chegando à:

$$y_t - \mu = \varphi_1(y_{t-1} - \mu) + \varepsilon_t$$

- Multiplicando ambos os lados da equação por $(y_{t-k} - \mu)$, tem-se:

$$(y_t - \mu)(y_{t-k} - \mu) = \varphi_1(y_{t-1} - \mu)(y_{t-k} - \mu) + \varepsilon_t\,(y_{t-k} - \mu)$$

- Tomando a esperança dos dois lados da equação e utilizando que $E\,[\varepsilon_t\,(y_{t-k} - \mu)] = 0$, chega-se à:

$$E[(y_t - \mu)(y_{t-k} - \mu)] = \varphi_1 E[(y_{t-1} - \mu)(y_{t-k} - \mu)]$$

- A hipótese de estacionariedade para o processo implica que:

$$E[(y_{t-1} - \mu)(y_{t-k} - \mu)] = E[(y_t - \mu)(y_{t-k-1} - \mu)] = \gamma_{k-1}$$

- Assim, chegamos a seguinte expressão para $E[(y_t - \mu)(y_{t-k} - \mu)] = \gamma_k$:

$$\gamma_k = \varphi_1 \gamma_{k-1}$$

- Por substituição recursiva, deduzimos a expressão desejada:

$$\gamma_k = \varphi_1^k \gamma_0$$

- A função de autocorrelação é dada por:

$$\rho(k) = \gamma_k / \gamma_0 = \varphi_1^k \text{, k = 1, 2, 3...}$$

- Assim sendo a forma da FAC p/ o AR (1) será o de:

  - senóide amortecida se $-1 < \phi_1 < 0$;
  - exponencial amortecida se $0 < \phi_1 < 1$.

Previsão para o processo AR (1)

- Previsões para valores da série k-passos à frente, a partir do instante t, podem ser obtidas pela projeção do modelo no futuro. A variável **k** é chamada **horizonte de previsão**.

- A função de previsão k-passos à frente, é por definição, a média do processo projetada k passos à frente, condicional nas observações até o tempo t. Formalmente:

$$\hat{y}_{t+k|t} = E(y_{t+k} \mid \mathbf{Y_t})$$

- O erro quadrático médio desta previsão (MSE), também pode ser obtido utilizando que:

$$MSE(\hat{y}_{t+k|t}) = E[(y_{t+k} - \hat{y}_{t+k|t})^2 \mid \mathbf{Y_t})$$

- Finalmente, intervalos de confiança de (1-α)% para a previsão podem ser obtidos, utilizando que

$$\hat{y}_{t+k|t} \pm z^{\alpha/2} \sqrt{MSE(\hat{y}_{t+k|t})}$$

- O cálculo de $MSE(\hat{y}_{t+k|t})$ pode ser feita através de substituição recursiva da expressão do processo:

$$y_t = c + \varphi_1 y_{t-1} + \varepsilon_1 \Rightarrow y_{t+1} = c + \varphi_1 y_t + \varepsilon_{t+1}$$

$$y_{t+2} = c + \varphi_1 y_{t+1} + \varepsilon_{t+2} = c + \varphi_1(c + \varphi_1 y_t + \varepsilon_{t+1}) + \varepsilon_{t+2}$$
$$= c\,(1 + \varphi_1) + \varphi_1^2 y_t + (\varphi_1 \varepsilon_{t+1} + \varepsilon_{t+2})$$

$$y_{t+k} = c \sum_{i=1}^{k} \varphi_1^{k-i} + \varphi_1^k\, y_t + \sum_{i=1}^{k} \varphi_1^{k-i}\, \varepsilon_{t+i}$$

$$\hat{y}_{t+k|t} = E(y_{t+k} \mid \mathbf{Y_t}) = c\,\frac{(1-\varphi_1^k)}{1-\varphi_1} + \varphi_1^k\, y_t$$

- Observe que $\lim_{k\to\infty} \hat{y}_{t+k|t} = c/(1-\varphi_1)$, a média incondicional.

- Finalmente:

$$MSE(\hat{y}_{t+k|t}) = E[(y_{t+k} - \hat{y}_{t+k|t})^2 \mid \mathbf{Y_t}) = \sum_{i=1}^{k} \varphi_1^{2(k-i)}\ E(\varepsilon^2_{t+i} \mid \mathbf{Y_t})$$

$$= \sigma^2 \sum_{i=1}^{k} \varphi_1^{2(k-i)}$$

$$MSE(\hat{y}_{t+k|t}) = \sigma^2 \,(1 + \varphi_1^2 + \varphi_1^4 + \varphi_1^6 + \ldots + \varphi_1^{2(k-1)})$$

Processo AR(2)

- Considere o processo AR(2)

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t$$

- No **modo condicional**:

  - $E(y_t \mid \mathbf{Y_{t-1}}) = c + \phi_1 y_{t-1} + \phi_2 y_{t-2}$
  - $Var(y_t \mid \mathbf{Y_{t-1}}) = \sigma^2$
  - $f(y_t \mid \mathbf{Y_{t-1}}) \sim N(c + \phi_1 y_{t-1} + \phi_2 y_{t-2}, \sigma^2)$

- Quanto ao **modo incondicional**, viu-se que ele estabelece as condições de estacionariedade de um modelo.

- Uma das maneiras de se obter a forma incondicional é através da substituição recursiva. Entretanto, apenas p/ modelos AR(1) este procedimento é recomendável, a não ser que re-escrevamos o modelo AR(p) em forma vetorial.

- Para modelos AR(p), p > 1, utiliza-se a **decomposição do polinômio** $\Phi_p(L)$.

- A “analogia direta” seria dizer que as condições de estacionariedade de um processo AR(2) seriam $|\phi_i| < 1$, i=1,2. Mas isto não está correto!

- Como será visto adiante, as verdadeiras condições de estacionariedade de um processo AR(2) são dadas por:

$$\phi_2 + \phi_1 < 1$$
$$\phi_2 - \phi_1 < 1$$
$$-1 < \phi_2 < 1 \quad (\text{ou} \quad |\phi_2| < 1)$$

- A investigação das condições de estacionariedade de um processo AR(2) é efetuada a partir da solução da equação de diferenças implicada pelo processo.

- Com vimos, a solução desta equação é dada pela soma da solução homogênea (y_t^h) mais a solução particular (y_t^p):

$$y_t = y_t^h + y_t^p$$

- Para construirmos estas soluções devemos, primeiro, obter a decomposição do polinômio $\Phi_2(L)=\left(1 - \varphi_1 L - \varphi_2 L^2\right)$.

- A partir desta decomposição podemos construir as soluções da equação do AR(2), cujo forma irá depender da natureza das raízes λ’s:

1) reais e diferentes
2) reais e iguais
3) complexas conjugadas

1) **Raízes Reais com** $\lambda_1 \neq \lambda_2$:

$$y_t = c + \varphi_1 \, y_{t-1} + \varphi_1 \, y_{t-2} + \varepsilon_1$$

$$\left(1 - \varphi_1 L - \varphi_2 L^2\right) y_t = c + \varepsilon_t$$

Mas

$$\left(1 - \varphi_1 L - \varphi_2 L^2\right) = \left(1 - \lambda_1 L\right)\left(1 - \lambda_2 L\right) \quad \text{(I)}$$

onde os λ's são raízes do polinômio

$$f(z) = z^2 - \varphi_1 z - \varphi_2, \quad z = 1/L$$

De (I), segue que:

a) $\lambda_1 \lambda_2 = -\varphi_2$

b) $\lambda_1 + \lambda_2 = \varphi_1$

Solução da equação homogênea:

$$y_t^h = A_1 \lambda_1^t + A_2 \lambda_2^t$$

em que A_1 e A_2 são constantes determinadas pelas condições iniciais e λ_1 e λ_2 são as raízes do polinômio em z, z =1/L.

Solução da equação particular:

$$y_t^p = \frac{1}{(1-\lambda_1 L)(1-\lambda_2 L)}(c+\varepsilon_t)$$

$$= \frac{1}{(\lambda_1-\lambda_2)}\left[\frac{\lambda_1}{(1-\lambda_1 L)} - \frac{\lambda_2}{(1-\lambda_2 L)}\right](c+\varepsilon_t)$$

Solução geral:

$$y_t = y_t^h + y_t^p$$

$$y_t = A_1\lambda_1^t + A_2\lambda_2^t + \frac{1}{(\lambda_1-\lambda_2)}\left[\frac{\lambda_1}{(1-\lambda_1 L)} - \frac{\lambda_2}{(1-\lambda_2 L)}\right](c+\varepsilon_t)$$

- Note que, <u>sem</u> impormos restrições ao processo, este é <u>não estacionário</u>, pois $E(y_t)$ depende explicitamente do tempo.

- Entretanto, usando que:

  - $|\lambda_i| < 1,\ i = 1,2$ (supomos esta condição);

- o tempo **t** está suficientemente afastado da origem;
- $\dfrac{1}{(1-\lambda_i L)} = \sum_{j=0}^{\infty} (\lambda_i L)^j$, se $|\lambda_i| < 1$, $i = 1, 2$.

segue que:

$$y_t = \frac{1}{(\lambda_1 - \lambda_2)} \{ [\lambda_1 \sum_{j=0}^{\infty} \lambda_1^j L^j - \lambda_2 \sum_{j=0}^{\infty} \lambda_2^j L^j] (c + \varepsilon_t) \}$$

$$= \frac{1}{(\lambda_1 - \lambda_2)} [\sum_{j=0}^{\infty} \lambda_1^{j+1} \varepsilon_{t-j} - \sum_{j=0}^{\infty} \lambda_2^{j+1} \varepsilon_{t-j}] + c \sum_{j=0}^{\infty} \lambda_1^{j+1} - c \sum_{j=0}^{\infty} \lambda_2^{j+1}]$$

$$= \frac{1}{(\lambda_1 - \lambda_2)} \left\{ \left[\frac{\lambda_1}{(1-\lambda_1)} - \frac{\lambda_2}{(1-\lambda_2)} \right] c + \sum_{j=0}^{\infty} (\lambda_1^{j+1} - \lambda_2^{j+1}) \varepsilon_{t-j} \right\}$$

$$= \left[\frac{1}{(1-\lambda_1)(1-\lambda_2)} \right] c + \sum_{j=0}^{\infty} \frac{(\lambda_1^{j+1} - \lambda_2^{j+1})}{(\lambda_1 - \lambda_2)} \varepsilon_{t-j}$$

- Assim, o processo pode ser escrito como:

$$y_t = m + \sum_{j=0}^{\infty} k_j \varepsilon_{t-j}$$

em que

$$m = \frac{c}{(1-\lambda_1)(1-\lambda_2)} = \frac{c}{(1 - \varphi_1 - \varphi_2)}$$

onde utilizamos a relação entre os λ's e os φ's dada pelas eqs (a) e (b).

$$k_j = \frac{(\lambda_1^{j+1} - \lambda_2^{j+1})}{\left(\lambda_1 - \lambda_2\right)}$$

- A partir desta expressão final, podemos, em princípio calcular a média e variância incondicionais e observar que ambas independem do tempo.

- Utilizando que $\varepsilon_{t-j}\,, j = 0 \ldots n$ é um ruído branco, tem-se que:

$$E(y_t) = m = \frac{c}{\left(1 - \varphi_1 - \varphi_2\right)}$$

- A variância incondicional pode ser calculada como a seguir (o termo cruzado desaparece usando que $E(\varepsilon_t\, \varepsilon_s) = 0,\ t \neq s$).

$$Var(y_t) = E(y_t - m)^2$$
$$= \sum_{j=0}^{\infty} k_j^2\ E(\varepsilon_{t-j}^2)$$
$$= \sigma^2 \sum_{j=0}^{\infty} k_j^2$$

- A avaliação deste somatório nos levará a uma expressão complicada em termo dos λ's, a qual poderia ser reparametrizada em termo dos φ's, utilizando as eqs (a) e (b).

- Entretanto uma melhor estratégia para se chegar a expressão da variância será utilizar a eq. original:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t$$

Usando que a média $m = \dfrac{c}{(1 - \varphi_1 - \varphi_2)} \rightarrow c = m(1 - \varphi_1 - \varphi_2)$.

Assim sendo:

$$(y_t - m) = \varphi_1 (y_{t-1} - m) + \varphi_2 (y_{t-2} - m) + \varepsilon_t \quad \text{(I)}$$

$$\mathrm{Var}(y_t) = E[(y_t - m)^2] = \varphi_1 E[(y_{t-1} - m)(y_t - m)] + $$
$$+ \varphi_2 E[(y_{t-2} - m)(y_t - m)] + E[(y_t - m)\, \varepsilon_t]$$

$$\gamma(0) = \varphi_1 \gamma(1) + \varphi_2 \gamma(2) + \sigma^2 \quad \text{(II)}$$

Precisamos das expressões para $\gamma(1)$ e $\gamma(2)$, as quais são obtidas multiplicando a eq.(I) por $E(y_{t-1} - m)$ e $E(y_{t-2} - m)$, respectivamente:

$$\gamma(1) = \varphi_1 \gamma(0) + \varphi_2 \gamma(1) \rightarrow \gamma(1) = \frac{\varphi_1 \gamma(0)}{(1 - \varphi_2)}$$

$$\gamma(2) = \varphi_1 \gamma(1) + \varphi_2 \gamma(0) \rightarrow \gamma(2) = \frac{\gamma(0)(\varphi_1^2 - \varphi_2^2 + \varphi_2)}{(1 - \varphi_2)}$$

Finalmente, substituindo estas expressões na expressão (II), após alguma álgebra obtemos:

$$\mathrm{Var}(y_t) = \frac{(1 - \varphi_2)\ \sigma^2}{(1 + \varphi_2)((1 - \varphi_2)^2 - \varphi_1^2)}$$

- É fácil de ver então que para que a variância incondicional do processo AR(2) esteja bem definida, i.e., seja positiva e não nula, é necessário impormos as seguintes restrições nos parâmetros do processo:

$$\varphi_2 + \varphi_1 < 1, \quad \varphi_2 - \varphi_1 < 1, \quad |\varphi_2| < 1$$

que são as condições de estacionariedade de um processo AR(2).

2) **Raízes Reais com** $\lambda_1 = \lambda_2 = \lambda$:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_1 y_{t-2} + \varepsilon_1$$

$$\left(1 - \varphi_1 L - \varphi_2 L^2\right) y_t = c + \varepsilon_t$$

$$\left(1 - \lambda L\right)^2 y_t = c + \varepsilon_t$$

Solução da equação homogênea

$$y_t^h = A_1 \lambda^t + A_2 \, t \, \lambda^t$$

em que A_1 e A_2 são constantes determinadas pelas condições iniciais (prove que esta solução satisfaz a eq. homogênea).

Solução da equação particular:

$$y_t^p = \frac{1}{(1-\lambda L)^2}(c + \varepsilon_t)$$

Solução geral:

$$y_t = A_1 \lambda^t + A_2 t \lambda^t + \frac{1}{(1-\lambda L)^2}(c + \varepsilon_t)$$

- Novamente, note que o processo não é estacionário, pois $E(y_t)$ depende explicitamente do tempo.

- De forma análoga à condição anterior, utilizando que:

  - $|\lambda_i| < 1, \; i = 1, 2$

$$\frac{1}{(1-\lambda_i L)} = \sum_{j=0}^{\infty} (\lambda_i L)^j, \; se \; |\lambda_i| < 1, \; i = 1, 2, \; t \to \infty$$

(para provar que $\lim_{t\to\infty} t\,\lambda^t = 0$ utilize que $\lim_{t\to\infty} t\,\lambda^t = \lim_{t\to\infty}(\lambda^t + \lambda^t + \ldots + \lambda^t) = \lim_{t\to\infty}(0 + 0 + 0\ldots) = 0$)

Sob essas condições segue que:

$$y_t = \frac{1}{(1\text{-}\lambda L)}\left[\frac{1}{(1\text{-}\lambda L)}(c + \varepsilon_t)\right]$$

$$= \frac{1}{(1\text{-}\lambda L)}\left[\frac{c}{(1\text{-}\lambda L)} + \sum_{j=0}^{\infty}\lambda^j \varepsilon_{t-j}\right]$$

$$= \frac{c}{(1\text{-}\lambda L)^2} + \sum_{j=0}^{\infty}\lambda^j \frac{1}{(1\text{-}\lambda L)}\varepsilon_{t-j},$$

$$= \frac{c}{(1\text{-}\lambda)^2} + \sum_{j=0}^{\infty}(j+1)\lambda^j \varepsilon_{t-j}$$

- De modo semelhante ao caso das raízes distintas, o processo pode ser escrito de outra forma:

$$y_t = m + \sum_{j=0}^{\infty} k_j \varepsilon_{t-j}$$

em que

$$m = \frac{c}{(1\text{-}\ \lambda)^2}$$

$$k_j = (j+1)\lambda^j$$

- A partir desta expressão final, poderíamos calcular a média e variância incondicionais, mas como no caso anterior, é mais indicado utilizar a eq. original (em termo dos φ's) para este propósito.

- É claro que as expressões da média e variância serão dadas pelas mesmas expressões previamente calculadas, i.e.:

$$E(y_t) = \frac{c}{\left(1 - \varphi_1 - \varphi_2\right)}$$

$$Var(y_t) = \frac{(1-\varphi_2)\ \sigma^2}{(1+\varphi_2)(1-\varphi_1-\varphi_2)(1-\varphi_1+\varphi_2)}$$

3) **Raízes Complexas Conjugadas**:

$$y_t = c + \varphi_1\, y_{t-1} + \varphi_1\, y_{t-2} + \varepsilon_1$$
$$\left(1 - \varphi_1 L - \varphi_2 L^2\right) y_t = c + \varepsilon_t$$

$$\left(1-\lambda_1 L\right)\left(1-\lambda_2 L\right) y_t = c + \varepsilon_t$$

- **Solução da equação homogênea:**

- Apesar de complexas, as raízes da equação homogênea são distintas, e a solução segue a forma do primeiro caso:

$$y_t^h = A_1\,\lambda_1^t + A_2\,\lambda_2^t$$

em que A_1 e A_2 também são constantes (complexas) determinadas pelas condições iniciais.

- Se tivermos $\Delta = \varphi_1^2 + 4\varphi_2 < 0$, os λ's serão complexos conjugados com $\lambda_j = a \pm bi,\ j=1,2$ em que:

$$\lambda_1 + \lambda_2 = \varphi_1 \quad \rightarrow a = \varphi_1/2$$
$$\lambda_1 - \lambda_2 = i\sqrt{-\Delta} \quad \rightarrow b = \sqrt{-\Delta}/2$$

- Podemos reescrever as raízes usando coordenadas polares de modo que $\lambda_j = Re^{\pm i\theta}$ em que $R^2 = a^2 + b^2 = -\varphi_2$ e assim segue que $|\lambda_j| = R = \sqrt{-\varphi_2},\ j=1,2$

- Reescrevendo a solução homogênea da equação de diferenças do processo em termos da representação em coordenadas polares das raízes, tem-se:

$$y_t^h = A_1 R^t e^{i\theta t} + A_2 R^t e^{-i\theta t}$$

- Como y_t^h é um número real, A_1 e A_2 devem ser números complexos. Apesar de serem números complexos arbitrários, demonstra-se que A_1 e A_2 devem possuir a seguinte forma:

$$A_1 = B_1\, e^{iB_2} \quad \text{e} \quad A_2 = B_1\, e^{-iB_2}$$

- Após as devidas substituições, podemos finalmente escrever y_t^h como:

$$y_t^h = R^t\, 2B_1 \cos(\theta t + B_2) = C\, R^t \cos(\theta t + B_2), \quad C = 2B_1$$

- Assim para que $\lim_{t\to\infty} y_t^h = 0$, de forma que o processo seja estacionário, é necessário e suficiente que imponhamos a condição $|\lambda_j| = R = \sqrt{-\varphi_2} < 1$.

- **Solução da equação particular:**

$$y_t^p = \frac{1}{(1-\lambda_1 L)(1-\lambda_2 L)}(c + \varepsilon_t)$$

$$= \frac{1}{(\lambda_1-\lambda_2)}\left[\frac{\lambda_1}{(1-\lambda_1 L)} - \frac{\lambda_2}{(1-\lambda_2 L)}\right](c + \varepsilon_t)$$

- **Solução geral:**

$$y_t = R^t\, B_1 \cos(\theta t + B_2) + \frac{1}{(\lambda_1 - \lambda_2)}\left[\frac{\lambda_1}{(1-\lambda_1 L)} - \frac{\lambda_2}{(1-\lambda_2 L)}\right](c + \varepsilon_t)$$

- Note que o processo não é estacionário, pois $E(y_t)$ depende explicitamente do tempo.

- Nota-se também que a condição de estacionariedade está diretamente ligada ao valor de R. De fato, o processo será assintoticamente estacionário se:

$$R = \sqrt{-\varphi_2} < 1$$

- Utilizando ainda que:

  - $$\frac{1}{(1-\lambda_i L)} = \sum_{j=0}^{\infty} (\lambda_i L)^j, \; se \; |\lambda_i| < 1, \; i = 1, 2$$
  - $t \to \infty$

segue que:

$$y_t = \frac{1}{(\lambda_1 - \lambda_2)} \{ [\lambda_1 \sum_{j=0}^{\infty} \lambda_1^j L^j - \lambda_2 \sum_{j=0}^{\infty} \lambda_2^j L^j] (c + \varepsilon_t) \}$$

$$= \frac{1}{(\lambda_1 - \lambda_2)} [\sum_{j=0}^{\infty} \lambda_1^{j+1} \varepsilon_{t-j} - \sum_{j=0}^{\infty} \lambda_2^{j+1} \varepsilon_{t-j}] + c \sum_{j=0}^{\infty} \lambda_1^{j+1} - c \sum_{j=0}^{\infty} \lambda_2^{j+1}]$$

$$= \frac{1}{(\lambda_1 - \lambda_2)} \left\{ \left[\frac{\lambda_1}{(1-\lambda_1)} - \frac{\lambda_2}{(1-\lambda_2)} \right] c + \sum_{j=0}^{\infty} (\lambda_1^{j+1} - \lambda_2^{j+1}) \varepsilon_{t-j} \right\}$$

$$= \left[\frac{1}{(1-\lambda_1)(1-\lambda_2)} \right] c + \sum_{j=0}^{\infty} \frac{(\lambda_1^{j+1} - \lambda_2^{j+1})}{(\lambda_1 - \lambda_2)} \varepsilon_{t-j}$$

- Assim, o processo pode ser escrito como:

$$y_t = m + \sum_{j=0}^{\infty} k_j \varepsilon_{t-j}$$

em que:

$$m = \frac{c}{(1-\lambda_1)(1-\lambda_2)} = \frac{c}{(1 - \varphi_1 - \varphi_2)}$$

$$k_j = \frac{(\lambda_1^{j+1} - \lambda_2^{j+1})}{(\lambda_1 - \lambda_2)}$$

- Observe que mesmo os λ´s sendo números complexos, tanto m quanto k_j serão reais. Pode-se facilmente provar que:

$$(1-\lambda_1)(1-\lambda_2) = (1-Re^{i\theta})(1-Re^{-i\theta}) = 1-2R\cos(\theta)+R^2$$

$$k_j = \frac{(\lambda_1^{j+1} - \lambda_2^{j+1})}{(\lambda_1-\lambda_2)} = \frac{R^{j+1}e^{i\theta(j+1)} - R^{j+1}e^{-i\theta(j+1)}}{Re^{i\theta}-Re^{-i\theta}} = \frac{R^j \sin(\theta(j+1))}{\sin(\theta)}$$

- A partir desta expressão final, poderíamos, em princípio, calcular a média e variância incondicionais, mas como no caso anterior, é mais indicado utilizar a eq. original (em termo dos φ’s) para este propósito.

- É claro que as expressões da média e variância serão dadas pelas mesmas expressões previamente calculadas, i.e.:

$$E(y_t) = \frac{c}{(1 - \varphi_1 - \varphi_2)}$$

$$Var(y_t) = \frac{(1-\varphi_2)\ \sigma^2}{(1+\varphi_2)(1-\varphi_1-\varphi_2)(1-\varphi_1+\varphi_2)}$$

Forma da FAC para um processo AR (2)

- De forma análoga à derivação da expressão para a função de autocorrelação para um processo AR(1) (pg. 41), chega-se a uma expressão para a auto-correlação de ordem **k** de um processo AR(2):

$$(y_t - m) = \varphi_1\ (y_{t-1} - m) + \varphi_2\ (y_{t-2} - m) + \varepsilon_t$$

$$E[(y_t - m)(y_{t-k} - m)] = \varphi_1\ E[(y_{t-1} - m)(y_{t-k} - m)] +$$

$$+ \varphi_2\ E[(y_{t-2} - m)(y_{t-k} - m)] + E[(y_{t-k} - m)\ \varepsilon_t]$$

$$\gamma(k) = \varphi_1\, \gamma(k-1) + \varphi_2\, \gamma(k-2)$$

$$\rho(k) = \varphi_1\, \rho(k-1) + \varphi_2\, \rho(k-2), \quad k = 1,2,...$$

- Nota-se que a FAC para um processo AR(2) segue uma equação de diferenças de 2ª ordem, homogênea.

- Observe a solução da equação acima, que descreve a FAC para o processo AR(2), é **idêntica** a solução da equação homogênea do processo AR(2) a qual nos dedicamos na seção anterior.

$$\left(1 - \varphi_1 L - \varphi_2 L^2\right)\rho(k) = 0$$

- Analisando as 3 possibilidades para as raízes do polinômio $\Phi_2(L)=\left(1 - \varphi_1 L - \varphi_2 L^2\right)$ detalhadas na seção anterior, é interessante observar que:

  - nos casos de raízes reais (diferentes ou idênticas), a FAC do AR(2) terá uma padrão semelhante à de um AR(1), como amortecimento exponencial.

  - no caso de raízes complexas o AR(2) apresenta um comportamento de senóide amortecida.

Condições de estacionariedade para modelos auto-regressivos

- Para o processo AR (2), observa-se que o estabelecimento das condições de estacionariedade são verificadas nas raízes do polinômio $f(z) = \Phi_2(z),\ z = 1/L$, e não diretamente nos parâmetros originais do modelo.

- A questão que se coloca agora é como expressar as restrições de estacionariedade em termo dos <u>parâmetros originais do modelo</u>, φ_1 e φ_2?

- A chave para esta resposta está nas relações entre os φ's e os λ's, as quais são dadas pela solução das raízes do polinômio $\Phi_2(L)$.

- Como já visto, a partir de $\left(1 - \varphi_1 L - \varphi_2 L^2\right) = 0$, fazendo z = 1/L, chegamos à $\left(z^2 - \varphi_1 z - \varphi_2\right) = 0$, que possui as seguintes raízes:

$$\lambda_1 = \frac{\varphi_1 + \sqrt{\varphi_1^{\,2} + 4\varphi_2}}{2} \quad \text{e} \quad \lambda_2 = \frac{\varphi_1 - \sqrt{\varphi_1^{\,2} + 4\varphi_2}}{2}$$

- Estudam-se as condições de estacionariedade em dois casos: se na decomposição da equação característica

$\left(1 - \varphi_1 L - \varphi_2 L^2\right) = \left(1 - \lambda_1 L\right)\left(1 - \lambda_2 L\right)$ as raízes λ's são **reais** ou **complexas**.

- No caso de **raízes reais**, o processo será explosivo se $\lambda_1 > 1$ e/ou $\lambda_2 < -1$.

▪ **Caso 1** : $\lambda_1 > 1 \rightarrow$

$$\frac{\varphi_1 + \sqrt{\varphi_1^2 + 4\varphi_2}}{2} > 1, \; ou$$

$$\sqrt{\varphi_1^2 + 4\varphi_2} > (2 - \varphi_1)$$

Como o lado esquerdo será sempre positivo, a desigualdade será sempre satisfeita para $\varphi_1 > 2$.

Se $\varphi_1 < 2$, então teremos que:

$$\sqrt{\varphi_1^2 + 4\varphi_2} > (2 - \varphi_1)$$

$$\varphi_1^2 + 4\varphi_2 > (2 - \varphi_1)^2$$

$$\varphi_2 > 1 - \varphi_1$$

▪ **Caso 2** : $\lambda_2 < -1 \rightarrow$

$$\frac{\varphi_1 - \sqrt{\varphi_1^2 + 4\varphi_2}}{2} < -1, \; ou$$

$$\sqrt{\varphi_1^2 + 4\varphi_2} < (2 + \varphi_1)$$

Como o lado esquerdo será sempre positivo, a desigualdade será sempre satisfeita para 2 + $\varphi_1 < 0$, ou $\varphi_1 < -2$.

Se $\varphi_1 > -2$, então teremos que:

$$\sqrt{\varphi_1^2 + 4\varphi_2} > (2 + \varphi_1)$$

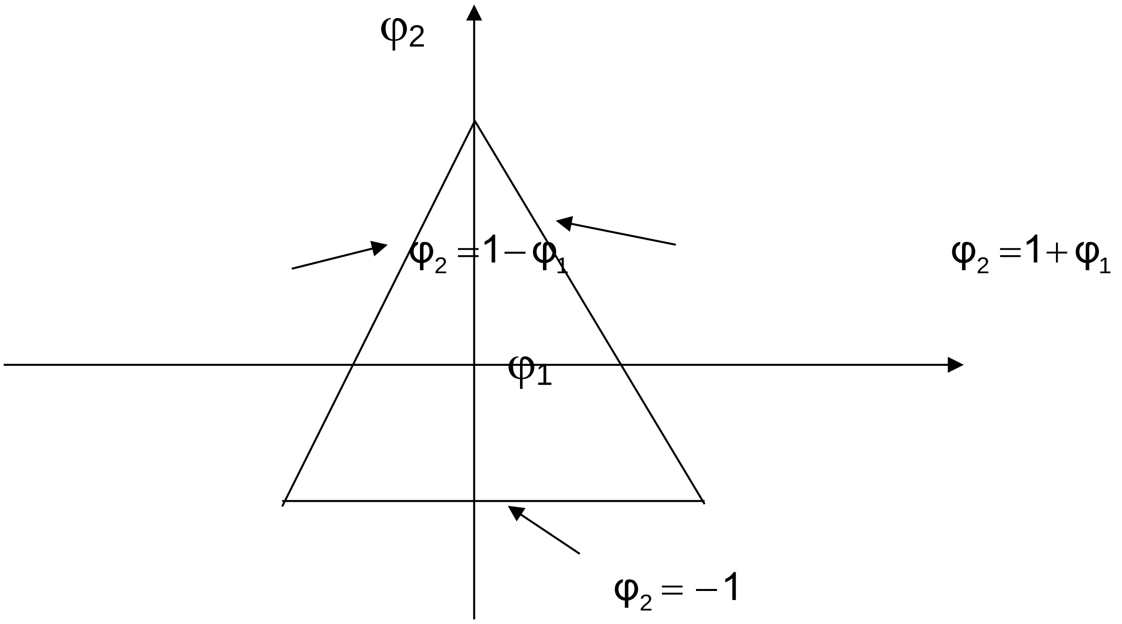
$$\varphi_1^2 + 4\varphi_2 > (2 + \varphi_1)^2$$

$$\varphi_2 > 1 + \varphi_1$$

- No caso de **raízes complexas**, sabemos que o modelo será explosivo se $|\lambda_i| = |R| > 1$, i =1,2 , isto é, se $R^2 = -\varphi_2 > 1$, ou, $\varphi_2 < -1$.

- Unindo as condições para os casos real e complexo, é fácil ver que a região de estacionariedade do modelo AR (2) será formada pelos pares (φ_1,φ_2), dentro do triângulo mostrado a seguir:

Região de estacionariedade de um AR (2)



- Portanto, para um processo AR (2), a condição de estacionariedade é generalizada pela investigação das raízes do polinômio característico, checando se elas localizadas "fora do círculo unitário[1]".

- Relembrando: para processos AR (1), exigimos como condição de estacionariedade $|\phi_1| < 1$, que também pode ser obtida exigindo que as raízes (zeros) do polinômio característico estejam fora do círculo unitário.

[1] Como já se sabe, as raízes podem ser complexas. O círculo unitário nesse caso se refere o conjunto bidimensional dos números no plano complexo que define o círculo centrado na origem de raio unitário.

- A questão que se coloca é como estender a condição de estacionariedade para um processo AR(p) genérico.

- Considere o processo AR(*p*) definido por:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ \dots + \varphi_p\ y_{t-p} + \varepsilon_t$$

- Reescrevendo em termos do operador de atraso L e do conseqüente polinômio $\Phi_p(L)$.

$$\left(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p\right) y_t = c + \varepsilon_t$$
$$\left(1 - \lambda_1 L\right)\left(1 - \lambda_2 L\right)\cdots\left(1 - \lambda_p L\right) y_t = c + \varepsilon_t$$

em que os λ's são as raízes do polinômio de grau **p.**

$$f(z) = \left(z^p - \phi_1 z^{p-1} - \phi_2 z^{p-2} - \cdots - \phi_p\right)$$

- Portanto, investigar as condições de estacionariedade para um processo AR(p) é equivalente a obter as raízes de um polinômio de grau p/ para investigar estacionariedade.

- Definindo as seguintes matrizes:

$$\mathbf{y}_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad \boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- Podemos reescrever a equação de diferenças **escalar** de **ordem p**, como uma equação de diferenças **vetorial** de **ordem 1**, dada por:

$$y_t = F\, y_{t-1} + \varepsilon_t$$

- Por substituição recursiva, chegamos à:

$$\mathbf{y}_t = \mathbf{F}^t \mathbf{y}_0 + \sum_{i=0}^{t-1} \mathbf{F}^i \boldsymbol{\varepsilon}_{t-i}$$

- O valor esperado do vetor y_t será dado por:

$$E(\mathbf{y}_t) = \mathbf{F}^t \mathbf{y}_0$$

- Supondo que a matriz F possui autovalores reais e distintos, é possível fatora-la na seguinte forma

$$F = T\Lambda T^{-1}$$

- A matriz Λ é dada por:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \ldots 0 \\ 0 & \lambda_2 & 0 \ldots 0 \\ \vdots & \vdots & \ddots \quad \vdots \\ 0 & 0 & 0 \cdots \lambda_p \end{bmatrix}$$

em que os elementos da diagonal são obtidos como soluções da equação determinantal (supondo autovalores reais e distintos):

$$\det(F - \lambda I) = \left(\lambda^p - \phi_1\lambda^{p-1} - \phi_2\lambda^{p-2} - \cdots - \phi_p\right) = 0$$

ou seja, os autovalores da matriz F.

- Portanto, investigar a condição de estacionariedade de um processo AR (p) é equivalente a **calcular os autovalores da matriz F**, checando a seguinte condição:

$$|\lambda_i| < 1,\ i = 1,2,\ \ldots\ ,\ p.$$

A Função de Autocorrelação Parcial (FACP)

- Como vimos os processos AR (1) e AR (2) apresentam formas similares para suas FAC's.

- Este comportamento faz com que, dada uma série temporal real, não seja possível identificar precisamente se o modelo gerador da série foi um AR (1) ou AR(2) (ou um AR(p) genérico) a partir da sua FAC estimada (correlograma).

- O que necessitamos é um tipo de função de autocorrelação que seja única para processos AR(p).

- Esta "função" existe e é denominada **Função de Autocorrelação Parcial**, abreviada como **FACP.**

- Para construirmos a FACP devemos, inicialmente, entender por que a FAC falha em identificar a ordem de processos AR(p).

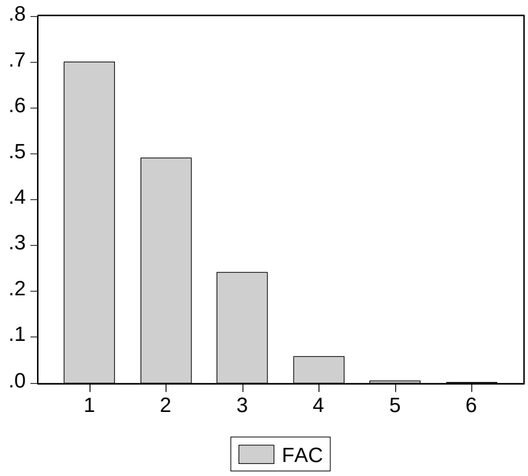
- Por exemplo, considere um processo AR (1):

$$y_t = c + \varphi_1 y_{t-1} + \varepsilon_t$$

- Como vimos, a FAC desse processo é dada por:

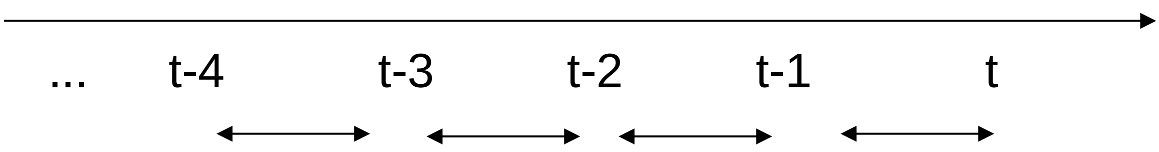
$$\rho(k) = \gamma_k \big/ \gamma_0 = \varphi_1^{\,k}, \; k = 1, 2, 3...$$

- O gráfico da FAC ($0 < \varphi_1 < 1$) é dada por:



- Observe que, embora o processo tenha apenas dependência explícita entre observações distantes uma unidade de tempo, o gráfico da FAC mostra que a dependência é não nula para observações distantes de 2, 3, 4,... unidades de tempo, embora a magnitude da dependência obviamente decresça.

- Não é difícil entender porque isso ocorre. A dependência entre observações distantes de uma unidade de tempo faz com que, por transitividade, observações distantes **k** unidades de tempo sejam também dependentes linearmente.



- Se desejarmos construir uma medida de "dependência líquida" das observações distantes de k unidades de tempo, devemos eliminar da dependência total a contribuição dada pelas observações intermediárias, localizadas em t-1, t-2, t-3,..., t-k+1.

- O coeficiente de correlação parcial faz exatamente isso: calcula a correlação entre y_t e y_{t-k} eliminando as influências das observações intermediárias y_{t-1} y_{t-2} ... $y_{t-k+1.}$

- Se agruparmos num gráfico estes coeficientes para vários valores de **k** obtemos a FACP. Formalmente:

$$\varphi_{kk} = corr(y_t, y_{t-k} \mid y_{t-1} y_{t-2} \ldots y_{t-k+1}),\ k = 1,2,3,\ldots$$

- É possível mostrar que cada coeficiente de correlação parcial de ordem k coincide com o último parâmetro de um modelo AR(k).

- A forma mais direta de se obter os coeficientes de autocorrelação parcial é, primeiramente, construir a série $y_t^* = y_t - \mu$ e então construir autoregressões dessa nova variável nos seus *lags*. Por exemplo:

$$y_t^* = \varphi_{11} y_{t-1}^* + \varepsilon_t$$

φ_{11} é portanto o coeficiente de autocorrelação parcial de ordem 1 para uma série, já que não há valores entre as duas variáveis (já que elas estão separadas por apenas um *lag*).

- Avançando mais um *lag* e formando um polinômio auto-regressivo de segunda ordem, temos:

$$y_t^* = \varphi_{21} y_{t-1}^* + \varphi_{22} y_{t-2}^* + \varepsilon_t$$

- Nesse caso, φ_{22} é o coeficiente de correlação parcial entre y_t e y_{t-2} , isto é, é o efeito líquido da variável y_{t-2} em y_t **controlando** pela variável y_{t-1} Pode-se mostrar que ele será dado por:

$$\varphi_{22} = \text{corr}(y_t, y_{t-2} | y_{t-1}) = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

onde $\rho_k = \text{corr}(y_t, y_{t-k}),\ \ k = 1,2,...$

Prova: a FAC de um processo AR(2) é dada por:

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} \qquad k = 1,2,...$$

Fazendo k =1 e k=2 na expressão acima, chegamos à:

$\rho_1 = \varphi_1 + \varphi_2 \rho_1$

$\rho_2 = \varphi_1 \rho_1 + \varphi_2$ ou em notação matricial:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} \Leftrightarrow p = P\varphi, \text{ ou } \varphi = P^{-1} p.$$

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} = \frac{1}{1 - \rho_1^2} \begin{bmatrix} 1 & -\rho_1 \\ -\rho_1 & 1 \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \Rightarrow \varphi_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \varphi_{22}.$$

- Por extensão, o último coeficiente em um processo AR(3) será o coeficiente de correlação parcial de ordem 3, i.e.,

$$\varphi_{33} = corr(y_t, y_{t-3} \mid y_{t-1} y_{t-2}).$$

- Esta equivalência permite-nos, facilmente observar que a FACP possui um comportamento único para cada processo AR(p).

- Por exemplo, se o processo “verdadeiro” que está gerando a nossa ST é um AR(1), então:

$$y_t = c + \varphi_1 y_{t-1} + \varepsilon_t$$

- Assim sendo para este processo φ_{kk} =0, k >1, pois todos os últimos coeficientes em modelos AR(k), K > 1,serão nulos, pois o nosso processo é AR(1).

- Por exemplo, no processo AR(2), $\varphi_{22} = \varphi_2 = 0$

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t$$

e assim por diante.

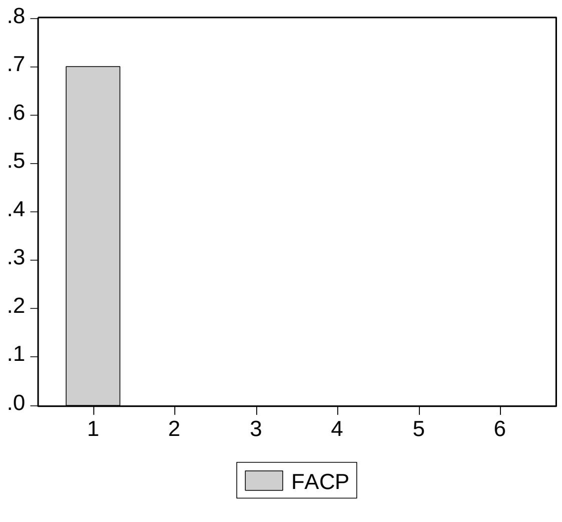
Prova: para um processo AR(2) temos que:

$$\rho_1 = \varphi_1 + \varphi_2\ \rho_1$$
$$\rho_2 = \varphi_1\ \rho_1 + \varphi_2$$

Como o processo verdadeiro é AR(1), então
$\rho_k = \varphi_1^k,\ k = 1,2.$ Fazendo k=1 e 2, e substituindo na 2a expressão acima:

$$\varphi_1^2 = \varphi_1^2 + \varphi_2 \Rightarrow \varphi_2 = \varphi_{22} = 0.$$

- Procedimento similar demonstra que $\varphi_{kk} = 0$, k = 3, 4, ...
- Portanto, a forma da FACP para um processo AR(1) é dada por ($\varphi_1 > 0$) :



- Como seria o gráfico da FACP para um processo AR(3) ?

- Um pouco de reflexão permite-nos concluir que a forma geral da FACP para um processo AR(p) será:

$$\varphi_{kk} \neq 0, \;\; k \leq p$$
$$= 0, \;\; k > p.$$

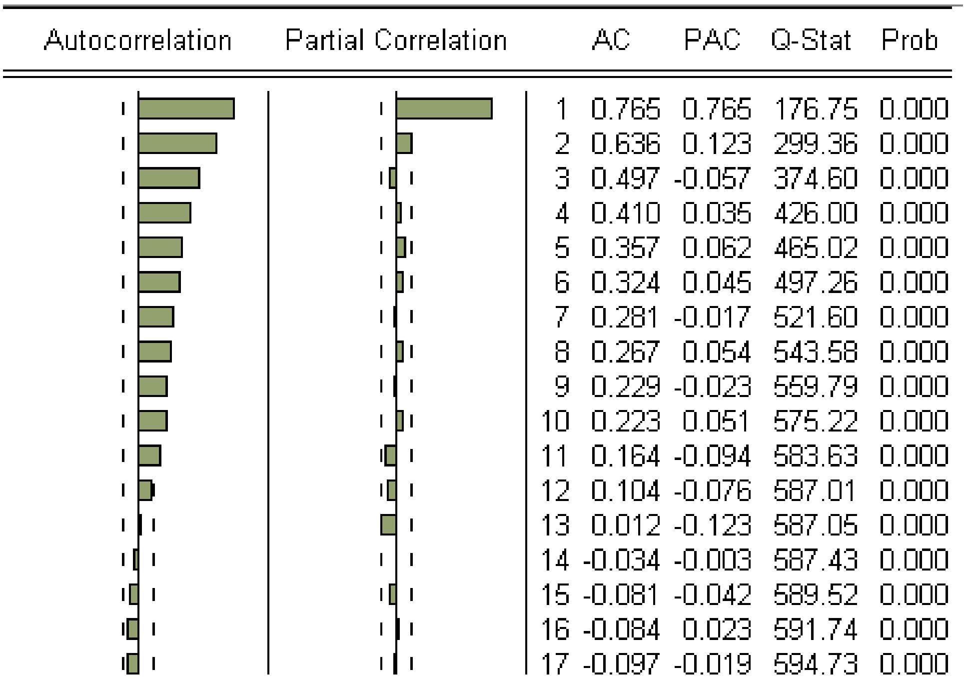
- Observe entretanto que esta é a FACP teórica, obtida a partir do modelo.

- A FACP de ordem k estimada a partir de uma ST é obtida estimando-se o último coeficiente de um processo AR(k).

- Se o valor deste coeficiente estiver dentro do intervalo de confiança de 95%, então, ao nível de 5%, não podemos rejeitar a hipótese de que $\varphi_{kk} = 0$.

- Por exemplo, para um processo AR(1) gerado artificialmente obtemos as seguintes FAC e FACP estimadas:

Fig. – FAC e FACP estimadas para uma realização do processo AR(1).



	AC	PAC	Q-Stat	Prob
1	0.765	0.765	176.75	0.000
2	0.636	0.123	299.36	0.000
3	0.497	-0.057	374.60	0.000
4	0.410	0.035	426.00	0.000
5	0.357	0.062	465.02	0.000
6	0.324	0.045	497.26	0.000
7	0.281	-0.017	521.60	0.000
8	0.267	0.054	543.58	0.000
9	0.229	-0.023	559.79	0.000
10	0.223	0.051	575.22	0.000
11	0.164	-0.094	583.63	0.000
12	0.104	-0.076	587.01	0.000
13	0.012	-0.123	587.05	0.000
14	-0.034	-0.003	587.43	0.000
15	-0.081	-0.042	589.52	0.000
16	-0.084	0.023	591.74	0.000
17	-0.097	-0.019	594.73	0.000

- Observe que na FACP teórica a correlação parcial para k >1 é exatamente zero, enquanto que na FACP estimada não é exatamente zero, mas é estatísticamente não discernivel de zero.

Estimação dos parâmetros

- A estimação dos parâmetros de processos AR pode ser efetuada por **Minimos Quadrados Ordinários** (MQO) ou por **Máxima Verossimilhança** (ML).

- Seja θ = [c, ϕ_1,..., ϕ_p, ,σ^2]' o vetor de parâmetros desconhecidos do modelo.

- A idéia básica da estimação via verossimilhança é encontrar um conjunto de valores para os parâmetros desconhecidos do modelo tal que, dada a observação uma amostra com T observações (y_1, y_2, ..., y_T), a probabilidade de termos observado aquela determinada amostra com os valores escolhidos para os parâmetros é a maior possível. Isto é equivalente a **maximizar a densidade de probabilidade conjunta das observações**.

- O problema a ser resolvido é da maximização da função de verossimilhança:

$$\max L(\boldsymbol{\theta})$$

- Por definição:

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta} \,|\, y_1, y_2, \ldots, y_T) = f\left(y_1, y_2, \ldots, y_T; \boldsymbol{\theta}\right)$$

em que $f\left(y_1, y_2, \ldots, y_T; \boldsymbol{\theta}\right)$ é a densidade conjunta de probabilidade das observações.

- Para observações dependentes, a densidade conjunta das observações de um processo AR(p) pode ser fatorada como o produto de **densidades condicionais marginais**. A forma geral dessa decomposição é dada por:

$$f\left(y_T, y_{T-1}, \ldots, y_1; \boldsymbol{\theta}\right) = \left(\prod_{t=1}^{T} p(y_t \mid \mathbf{Y_{t-1}}, \theta)\right) . f\left(y_p, y_{p-1}, \ldots, y_1; \boldsymbol{\theta}\right)$$

em que $y_p, y_{p-1}, \ldots, y_1$ é o conjunto dos **p** valores iniciais e $p(y_t \mid \mathbf{Y_{t-1}})$ é a densidade condicional marginal para a observação realizada em y_t.

- Se assumirmos ainda que o termo aleatório do processo AR(p) cujos parâmetros estamos estimando segue um processo ruído branco (ou seja, gaussiano com média nula e variância constante e igual a σ^2), $p(y_t \mid \mathbf{Y_{t-1}})$ é dado por:

$$p(y_t \mid \mathbf{Y_{t-1}}) \sim N\,(c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_p y_{t-p}, \sigma^2)$$

- Como já dito, a idéia é escolher como estimador de θ o valor $\hat{\boldsymbol{\theta}}$ que maximiza a verossimilhança.

- Por razões práticas, maximiza-se o logaritmo da verossimilhança (a verossimilhança e seu log possuem o mesmo máximo).

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\arg\max}\ l(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}}{\arg\max}\ \ln\left(L(\boldsymbol{\theta})\right)$$

- A função de log-verossimilhança pode então ser escrita como:

$$\ln\Big(f\left(y_T, y_{T-1}, \dots, y_1; \boldsymbol{\theta}\right)\Big) = \sum_{t=p+1}^{T} \ln\big(p(y_t \mid \mathbf{Y_{t-1}}, \theta)\big) + \ln\Big(f\left(y_p, y_{p-1}, \dots, y_1; \boldsymbol{\theta}\right)\Big)$$

- O primeiro termo do lado direito da equação é chamado de **log-verossimilhança condicional**. O segundo termo é chamado de **log-verossimilhança marginal para os valores iniciais.**

- Nos modelos de séries temporais, dois tipos de MLE podem ser estimados:

  1) <u>Máxima verossimilhança condicional</u>

Baseia-se na maximização da **log-verossimilhança condicional,** isto é**:**

$$\hat{\boldsymbol{\theta}} = \arg\max_{\theta} \sum_{t=p+1}^{T} \ln\big(p(y_t \mid \mathbf{Y_{t-1}}, \theta)\big)$$

2. <u>Máxima verossimilhança exata</u>

- Baseia-se na maximização da função de log-verossimilhança na sua forma completa (condicional + marginal para valores iniciais) **,** isto é**:**

$$\hat{\boldsymbol{\theta}} = \arg\max_{\theta} \sum_{t=p+1}^{T} \ln\left(p(y_t \mid \mathbf{Y_{t-1}}, \theta)\right) + \ln\left(f\left(y_p, y_{p-1}, \ldots, y_1; \boldsymbol{\theta}\right)\right)$$

- Para um processo AR (1), a máxima verossimilhança ***condicional*** é dada por:

$$l(\boldsymbol{\theta}) = \log\left(\prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2\sigma^2}(y_t - (c - \phi_1 y_{t-1}))^2]\right) =$$

$$= \log\left[\left(\frac{1}{2\pi\sigma^2}\right)^{(T-1)/2} e\text{xp}[-\frac{1}{2\sigma^2}\sum_{t=2}^{T}(y_t - (c - \phi_1 y_{t-1}))^2\right] =$$

$$= -\frac{(T-1)}{2}\log(2\pi) - \frac{(T-1)}{2}\log(\sigma^2) - \sum_{t=2}^{T}\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}$$

- Os estimadores de MV são obtidos verificando as seguintes condições de máximo de uma função:

$-$ condição de 1ª ordem: $\dfrac{\partial \mathrm{l}(\boldsymbol{\theta})}{\partial \theta_i} = 0$,i=1,2,...,p (*necessária*)

$-$ condição de 2ª ordem: $H = \dfrac{\partial^2 \mathrm{l}(\theta)}{\partial \theta \partial \theta'}$ *é negativa definida* (*suficiente*)

- É fácil de ver que derivando l(**θ**) em termos de c, ϕ_1 e σ^2 e igualando a zero, encontramos:

$$\hat{c} = \bar{y}_{t,T-1} - \hat{\phi}_1 \bar{y}_{t-1,T-1}, \; onde \, \bar{y}_{s,T-1} = \frac{\sum_{t=2}^{T} y_s}{(T-1)}$$

$$\hat{\phi}_1 = \frac{\sum_{t=2}^{T} y_t \sum_{t=2}^{T} y_{t-1} - (T-1)\sum_{t=2}^{T} y_t y_{t-1}}{\left(\sum_{t=2}^{T} y_{t-1}\right)^2 - (T-1)\sum_{t=2}^{T} y_{t-1}^2}$$

$$\hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=2}^{T}[y_t - (\hat{c} + \hat{\phi}_1 y_{t-1})]^2$$

- Não é difícil observar que estes estimadores seriam também obtidos ao minimizar a soma dos quadrados dos erros de previsão, isto é:

$$\varepsilon_t = y_t - \hat{y}_{t|t-1} = y_t - (c + \phi_1 y_{t-1})$$

$$Q(c, \phi_1) = \sum_{t=2}^{T} \varepsilon_t^{\,2} = \sum_{t=2}^{T} [y_t - (c + \phi_1 y_{t-1})]^2 \; .$$

- Este critério de otimização é chamado de Mínimos Quadrados Ordinários e será equivalente à Máxima Verossimilhança Condicional se $\varepsilon_t \sim N(0, \sigma^2)$.

- Ainda para o processo AR(1), podemos escrever a máxima verossimilhança ***exata***, dada pela da log-verossimilhança condicional deduzida acima somada a densidade marginal para y_1. Dado que:

$$p(y_1|y_0) \sim N\left(\frac{c}{(1-\varphi_1)}, \frac{\sigma^2}{1-\varphi_1^2}\right)$$

tem-se que:

$$f(y_1, \boldsymbol{\theta}) = \left(2\pi \frac{\sigma^2}{1-\phi_1^2}\right)^{1/2} \exp\left[-\left(\frac{1-\phi_1^2}{2\sigma^2}\right)\left(y_1 - \frac{c}{1-\phi_1}\right)^2\right]$$

- Note que a função de log-verossimilhança exata é não linear nos parâmetros. Portanto, não mais será possível obter soluções analíticas para os estimadores de MV, e neste caso não haverá equivalência entre os estimadores de MQO e os de MV. Neste caso, a solução para o problema de maximização é usualmente encontrada pelo uso de métodos numéricos.

- Para *T* suficientemente grande, a distribuição dos estimadores de MV condicionais pode ser aproximada pela seguinte distribuição:

$$\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, T^{-1}\mathbf{I}^{-1})$$

em que **I** é a matriz de informação de Fisher definida por:

$$\mathbf{I} = -\mathrm{E}\left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)$$

- A partir deste resultado é possível então realizar testes de hipótese para investigar a significância dos coeficientes de um modelo AR

Included observations: 299
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.002485	0.326664	-0.007607	0.9939
AR(1)	0.629202	0.045294	13.89160	0.0000

R-squared	0.393849	Mean dependent var	0.016307
Adjusted R-squared	0.391808	S.D. dependent var	2.685514
S.E. of regression	2.094343	Akaike info criterion	4.323024
Sum squared resid	1302.723	Schwarz criterion	4.347776

p-valor para o teste de H_o: $\varphi_1 = 0$

- Para um processo AR(p) genérico, condicional em y_1, y_2, ..., y_p , o log da verossimilhança será dada por:

$$\boxed{\begin{aligned} l(\boldsymbol{\theta}) &= -\frac{(T-p)}{2}\log(2\pi) - \frac{(T-p)}{2}\log(\sigma^2) \\ &\quad - \sum_{t=p+1}^{T} \frac{[y_t - (c - \sum_{i=1}^{p} \phi_i y_{t-p})]^2}{2\sigma^2} \end{aligned}}$$

- Como esperado, nesta situação os estimadores de MV para c e os φ´s são equivalentes aos de MQO.