Modelos GARCH para estimativa da volatilidade em séries temporais financeiras univariadas

Nelio MACHADO

Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

December 2, 2019

Overview

Qual o objetivo da análise de séries temporais?

- Extrair informações de uma série temporal.
  1. Descrever: calcular medidas estatísticas descritivas da série temporal como, por exemplo tendência e sazonalidade.
  2. Explicar: a variação de uma série temporal pode nos ajudar a entender e explicar o processo estocástico gerador da série.
  3. Prever: usamos os valores passados da série temporal para nos ajudar a prever o futuro.
  4. Controlar: podemos usar os valores previstos de uma série temporal para controlar, monitorar e otimizar as configurações de um sistema ou processo.

Principais Fatos estilizados em Séries Temporais Financeiras

- Agrupamento de volatilidade (também conhecido como heterocedasticidade condicional) - Surge nos períodos de calma ou nervosismo do mercado financeiro. Ou seja, a volatilidade do instrumento financeiro não é constante no tempo (MORETTIN; TOLOI, 2004);
- São (em geral) não-correlacionadas. Se há correlação, então elas são significativas apenas para pequenas defasagens no tempo;
- Os quadrados dos retornos são auto-correlacionados;
- Apresentam média amostral próxima de zero;
- Apresentam caudas (esquerda e direita) pesadas;
- Apresentam alguma forma de não-linearidade, reagindo de forma diferente a choques pequenos ou grandes;

Processos Estocásticos Estacionários

- Seja T (geralmente o conjunto dos números inteiros) um conjunto arbitrário. Um Processo Estocástico é uma família

$$Z_t = Z(t), t \in T \tag{1}$$

  tal que para cada $t \in T$, $Z(t)$ é uma variável aleatória. Então, por definição, Série Temporal é uma realização ou trajetória de um Processo Estocástico. De forma simples, Séries Temporais são simples pontos ordenados no tempo.
- O Processo Estocástico Z_t é dito estacionário quando as propriedades estatísticas de qualquer sequência finita $Z_1, Z_2, ..., Z_k$ são semelhantes á sequência $Z_{1+\tau}, Z_{2+\tau}, ..., Z_{k+\tau}$ para qualquer número inteiro τ. Em outras palavras, as propriedades estatísticas (média μ e variância σ^2) são todas constantes ao longo do tempo, não apresentando tendência (variações em torno da média tem amplitude constante).

Exemplos de Processos Estocásticos

A seguir, exemplos de Séries Temporais:

- Série do GDP (Gross Domestic Product);
- População de Portugal;
- Consumo de eletricidade e água;
- Uso do cartão de crédito;
- Aquecimento Global;
- Retornos de ativos financeiros;

Sazonalidade e tendência

- Muitas séries temporais apresentam um comportamento que tende a se repetir a cada s períodos de tempo. Isso é o que chamamos de sazonalidade.
- Nos 3 slides a seguir, apresentamos exemplos de séries temporais sazonais, onde é fácil ver a tendência de crescimento (ou decrescimento) e o padrão sazonal a se repetir ao longo do tempo.

Exemplos de séries temporais sazonais: AirPassengers

A seguir apresentamos o gráfico da série temporal AirPassengers. Esta série traz informações do número de passageiros transportados por linhas aéreas internacionais nos EUA no período entre 1949 e 1960 (NGO, 2013).

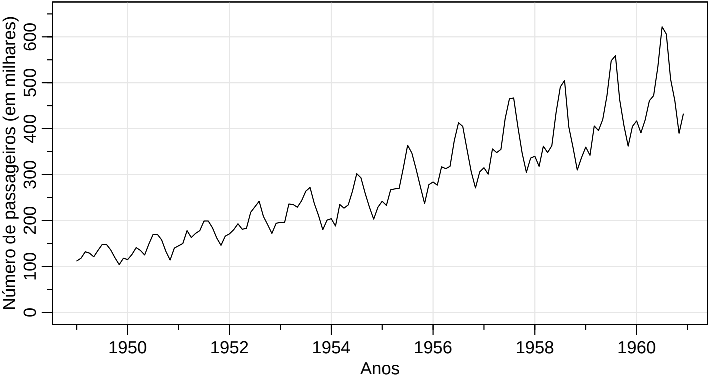


Figure 1: Série temporal sazonal AirPassengers

Exemplos de séries temporais sazonais: Retornos de Ações da Johnson & Johnson

A seguir apresentamos a série temporal dos ganhos trimestrais por ação da empresa norte-americana Johnson & Johnson durante o período de 1960 a 1980 0 (ROBERT H. SHUMWAY; STOFFER, 2010).

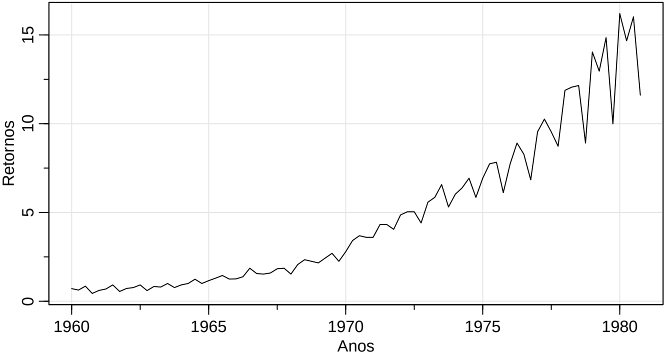


Figure 2: Série temporal sazonal da Johnson & Johnson)

Exemplos de séries temporais sazonais: Aquecimento Global

A figura a seguir apresenta a série do índice médio global da temperatura do oceano de 1880 a 2015. A tendência ascendente desta série durante a última parte do século XX tem sido usado como argumento para a hipótese da mudança climática (ROBERT H. SHUMWAY; STOFFER, 2010).

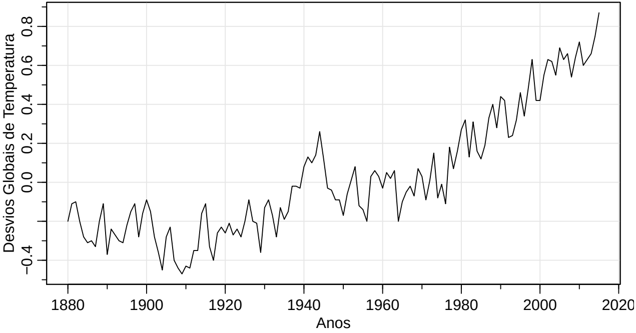


Figure 3: Aquecimento Global)

Ruído Branco

Uma Série Temporal X_t é chamada de Ruído Branco se $X_t \sim N(0, \sigma^2)$ iid. Ou seja, para $t = 1, 2, \ldots, n$, temos:

- X_t são independentes e identicamente distribuídos com $\mathbb{E}[X_t] = 0$ e $\mathbb{V}[X_t] = \sigma^2$;
- $Cov(X_i, X_j) = 0 \quad \forall i \neq j$;
- O Ruído Branco é estacionário - Observe que as variações em torno da média são constantes.

Exemplo de Ruído Branco

O gráfico abaixo apresenta a série temporal de um Ruído Branco Gaussiano composto de 500 pontos.

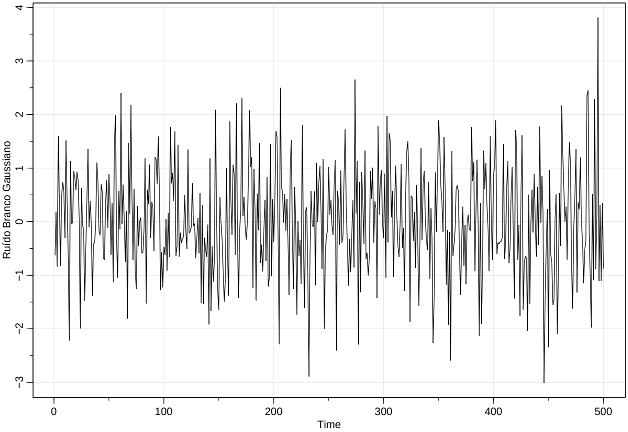


Figure 4: Ruído Branco Gaussiano

Log-retornos

- Foco na análise da série de log-retornos ao invés da análise da série dos preços dos ativos propriamente ditos, pois os log-retornos possuem propriedades estatísticas mais interessantes do que as séries de preços como, por exemplo, estacionariedade (MORETTIN; TOLOI, 2004);
- Seja p_t o preço de um ativo no momento t e suponha que não sejam pagos dividendos no período. Definimos $\Delta p_t = p_t - p_{t-1}$ a variação dos preços do ativo entre os instantes $t-1$ e t e r_t sendo o retorno do ativo dado por (MORETTIN; TOLOI, 2004):

$$r_t = \frac{\Delta p_t}{p_{t-1}} = \frac{p_t - p_{t-1}}{p_{t-1}}.$$

Se fizermos $P_t = \log p_t$, então o retorno composto continuamente (ou simplesmente log-retorno, como é mais conhecido) será dado:

$$R_t = \log\left(\frac{p_t}{p_{t-1}}\right) = \log(1 + r_t) = P_t - P_{t-1}. \tag{2}$$

Retornos da Oracle (ORCL)

Figure 5: Retornos da Oracle (ORCL)

Log-retornos da Oracle (ORCL)

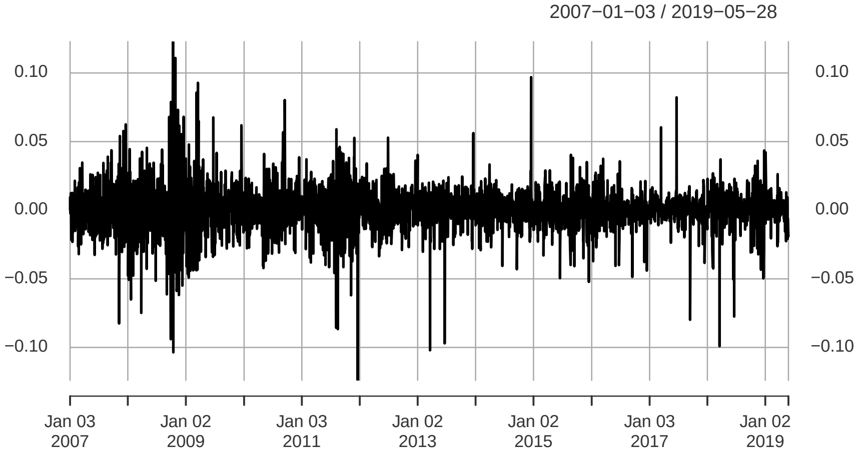


Figure 6: Log-retornos da Oracle (ORCL)

Correlograma: O que é? — Para que serve?

- Serve para nos ajudar a avaliar a presença de autocorrelação (grau de dependência linear) entre X_t e X_{t-l}, com $l = 1, 2, ...$;
- Utilizado no processo de identificação do modelo mais adequado para a Série Temporal. Se o modelo foi bem ajustado à série temporal, então não há autocorrelação visível na série sob estudo.
- Correlograma é um gráfico da FAC (Função de Autocorrelação) para uma sequência de valores dos lags $k = 0, 1, 2, \ldots n$, permitindo-nos testar/avaliar a hipótese de que todos os coeficientes de correlação $\rho_k, \quad k = 1, 2, ...$ de log-retornos são iguais a zero, ou seja, não há correlação entre os valores atuais X_t e os valores passados $X_{t-1}, X_{t-2}, ...$
- Seriamente impactado por *outliers* (GREENE, 2012). Recomenda-se analisar o boxplot;

Correlograma: Fórmulas e propriedades

- As linhas azuis pontilhadas no correlograma representam limites superiores e inferiores sobre as quais, se os valores estão fora destes, temos evidências contra a hipótese nula de que a correlação $\rho(\tau)$ é igual a zero no nível de 5%. Tomando o Intervalo de Confiança de 95%, então espera-se que 5% desses lags excedam esses valores.
- A FAC (Função de Autocorrelação) é dada por:

$$\rho(t, t+\tau) = \rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)} = \frac{\gamma(\tau)}{\sigma^2}, \tau \in \mathbb{Z}$$

- FAC deve satisfazer as seguintes propriedades:
  1. $\rho(0) = 1$;
  2. A correlação entre X_t e $X_{t+\tau}$ é a mesma que entre X_t e $X_{t-\tau}$, ou seja, $\rho(-\tau) = \rho(\tau)$;
  3. $-1 \leq \rho(\tau) \leq 1 \quad \forall \tau$;

Correlograma do Ruído Branco Gaussiano apresentado na figura 4

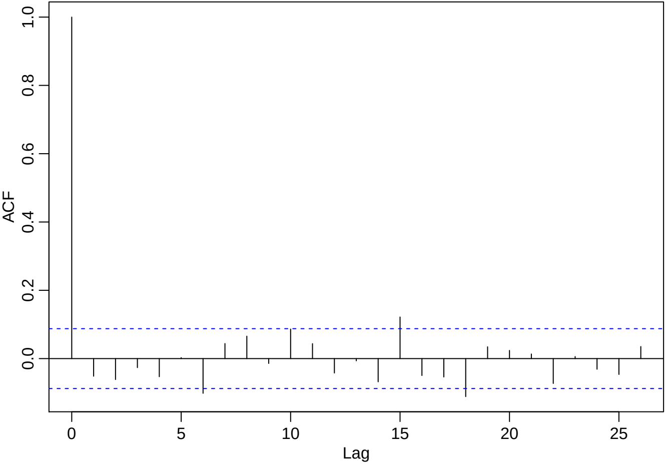


Figure 7: Correlograma do Ruído Branco Gaussiano apresentado na figura 4

Correlograma dos log-retornos da Oracle (ORCL)

Vimos na figura 6 o gráfico da Série Temporal dos log-retornos das ações da Oracle (ORCL). Abaixo, apresentamos o correlograma da série:

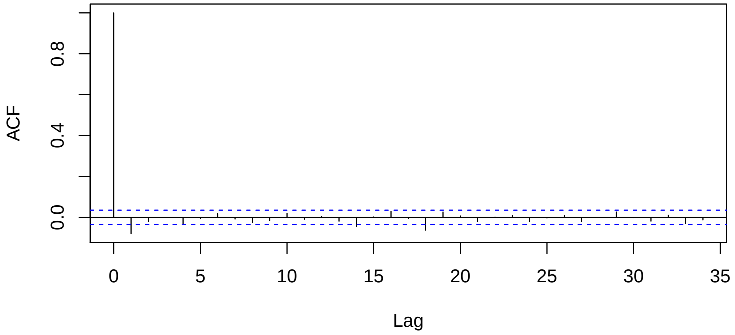


Figure 8: Correlograma dos log-retornos da Oracle (ORCL)

Correlograma do retorno de ações da Johnson & Johnson

Observe na figura 2 que a série é não-estacionária. Em seguida, compare com os gráficos da figuras 6 e 8.

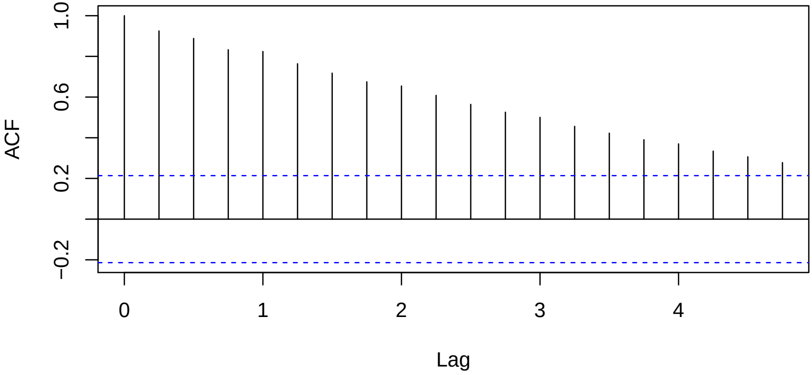


Figure 9: Retorno de Ações da Johnson &Johnson

Correlograma da Série Temporal de Aquecimento Global

Observe na figura 3 que a série é não-estacionária. Em seguida, compare os gráficos das figuras 6 e 8.

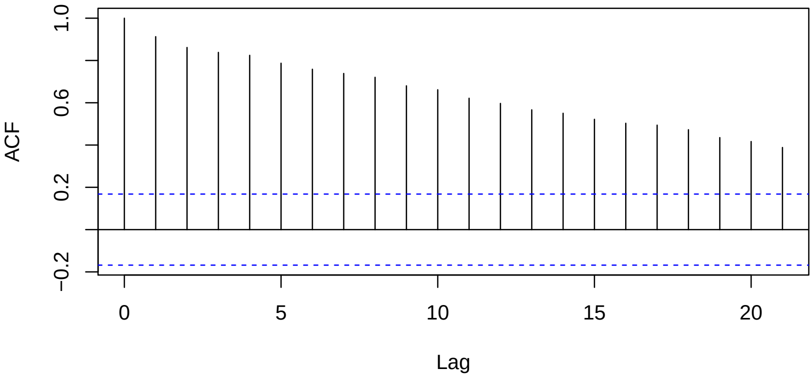


Figure 10: FAC Aquecimento Global

Modelos ARIMA

Os modelos $ARIMA(p, d, q)$ foram propostos por Box e Jenkins (NGO, 2013) e tem a seguinte expressão:

$$\hat{X}_t = \mu + \sum_{i=1}^{p} \phi_i X_{t-i} - \sum_{j=1}^{q} \beta_j \varepsilon_{t-j}, \tag{3}$$

sendo que os parâmetros p, d e q são números inteiros não-negativos que possuem o seguinte significado:

- p: representa o número de termos autoregressivos (defasagens ou lags na parte autoregressiva do modelo);
- d: representa o número de diferenças não-sazonais necessárias para tornar a série estacionária. Diferenciar a série significa simplesmente substrair X_t por X_{t-1} d vezes;
- q: representa o número de médias móveis.

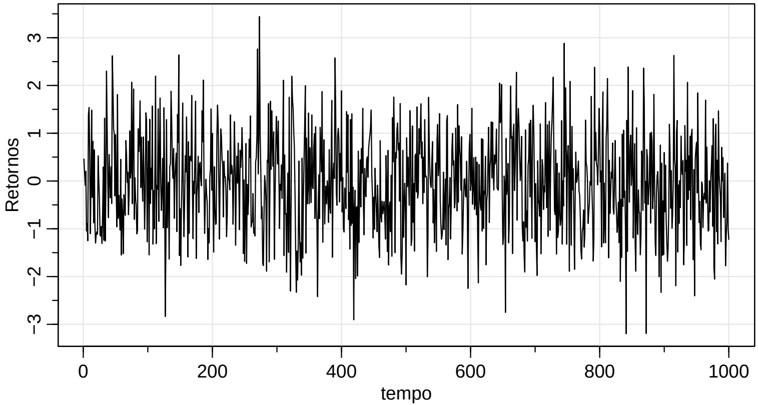
Modelos ARIMA: Considerações importantes

- Na prática (geralmente), os valores de p e q são menores que 2 para séries temporais estacionárias (NGO, 2013);
- Valores passados $X_{t-1}, X_{t-2}, \ldots$ são utilizados para explicar os valores atuais e futuros. Isso é fácil ver na equação 3;
- A série temporal precisa ser estacionária (na verdade, verificar se a série temporal é estacionária deve ser o primeiro passo em análise de séries temporais). Se a série for não-estacionária, deve-se aplicar procedimentos para torná-la estacionária. Geralmente uma diferença (d$=$ 1) torna a série estacionária (GUJARATI, 2011);
- A diferença básica entre os modelos de regressão clássicos / tradicionais e os modelos ARIMA é que nos modelos ARIMA não se pode assumir independência entre observações. Ao contrário, modelos ARIMA tem o objetivo de modelar o grau de autocorrelação entre as observações defasadas da série $(X_{t-1}, X_{t-2}, \ldots)$.

Modelos ARIMA: Exemplo simulado

A seguir, simulamos um processo $ARIMA(1,0,1)$ com $\phi_1 = 0.6$ e $\theta_1 = -0.5$. Ou seja, simulamos o processo abaixo:

$$\hat{X}_t = \phi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} = 0.6X_{t-1} - 0.5\varepsilon_{t-1}$$



Modelos ARIMA: Correlograma do processo simulado

Abaixo apresentamos o correlograma do processo simulado $ARIMA(1, 0, 1)$ com $\phi_1 = 0.6$ e $\theta_1 = -0.5$:

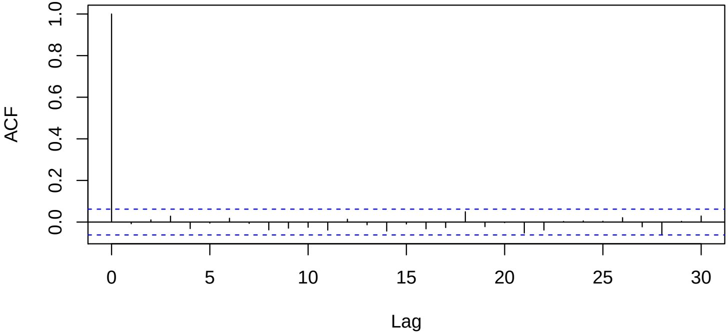


Figure 11: Correlograma

Modelos ARIMA: Observações finais

- Observe que o correlograma dos resíduos na figura 11 se comportam como Ruído Branco, o que é desejável;
- O teste de Ljung-Box avalia se a série temporal é Ruído Branco, o que significa dizer que a correlação entre os valores passados $X_{t-1}, X_{t-2}, ...$ da série são iguais a zero, ou seja, não há autocorrelação serial. Isto significa que os valores defasados da série são independentes e identicamente distribuídas. O teste de Ljung-Box tem $p - value = 0.6252 > 0.05$, permitindo-nos concluir que há fortes evidências de que os resíduos da séria analisada são Ruído Branco, o que é corroborado pelo correlograma apresentado na figura 11.

Modelos $AR(p) = ARMA(p,0) = ARIMA(p,0,0)$

Os modelos autoregressivos de ordem p, ou $AR(p)$, podem ser escritos como:

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \varepsilon_t \\ &= \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t, \end{aligned} \tag{4}$$

onde ε_t é Ruído Branco e $\phi_i \in \mathbb{R}$, com $\phi_i \neq 0 \quad \forall i = 1, 2, \ldots, p$. Ou seja, X_t é combinação linear de seus valores passados $X_{t-1}, X_{t-1}, ..., X_{t-p}$.

- Caso especial: Passeio Aleatório $= AR(1)$ com $\phi_1 = 1$. Neste caso, o processo é não-estacionário. Veja no próximo slide o correlograma de um Passeio Aleatório simulado.

Correlograma de um Passeio Aleatório

Abaixo apresentamos o correlograma de um Passeio Aleatório simulado:

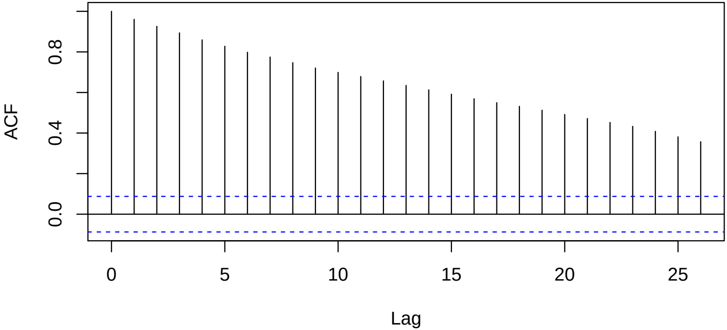


Figure 12: Correlograma

Modelos $MA(q) = ARMA(0, q) = ARIMA(0, 0, q)$

O modelo de Médias Móveis é similar ao modelo auto-regressivo. A diferença reside no fato de que ao invés de termos uma regressão linear dos valores passados $X_{t-1}, X_{t-2}, ...$ da série, teremos uma regressão linear (ou combinação linear) de Ruídos Brancos $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-q}$. Matematicamente, escrevemos da seguinte forma:

$$
\begin{aligned}
X_t &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \\
&= \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j},
\end{aligned} \tag{5}
$$

onde ε_t é Ruído Branco com $\mathbb{E}(\varepsilon_t) = 0$ e $\mathbb{V}(\varepsilon_t) = \sigma^2$. É fácil ver que $\mathbb{E}(\varepsilon_t) = 0$, pois $\mathbb{E}(X_t) = \sum_{i=1}^{q} \mathbb{E}(\varepsilon_i) = 0$.

Modelos $ARMA(p, q) = ARIMA(p, 0, q)$

Se o processo estacionário apresenta características simultaneamente de um processo $AR(p)$ e de um processo $MA(q)$, então temos um processo misto $ARMA(p, q)$, descrito por seus p valores passados $X_{t-1}, X_{t-2}, ..., X_{t-p}$ e pelos q Ruídos Brancos $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-q}$, dados por:

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \\ &= \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \end{aligned}$$

onde ε_t é Ruído Branco com $\mathbb{B}(\varepsilon_t) = 0$ e $\mathbb{V}(\varepsilon_t) = \sigma^2$.

Exemplo simulado: Como escolher o melhor modelo $ARMA(p, q) = ARIMA(p, d, q)$?

Para explicar o método de seleção de modelos, simulamos 1.000 pontos de um processo $ARMA(3, 2)$ com $\phi_1 = 0.5$, $\phi_2 = -0.25$, $\phi_3 = 0.4$, $\theta_1 = 0.5$ e $\theta_2 = -0.3$. A expressão matemática do modelo é dada por:

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \\ &= 0.5X_{t-1} - 0.25X_{t-2} + 0.4X_{t-3} + \varepsilon_t + 0.5\varepsilon_{t-1} - 0.3\varepsilon_{t-2} \end{aligned} \quad (7)$$

O gráfico do modelo simulado pode ser visto no slide a seguir:

Gráfico do processo $ARMA(3,2)$ simulado com $\phi_1 = 0.5$, $\phi_2 = -0.25$, $\phi_3 = 0.4$, $\theta_1 = 0.5$ e $\theta_2 = -0.3$

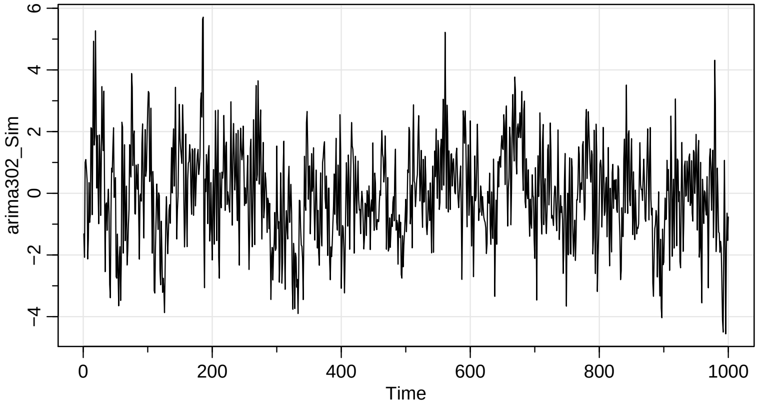


Figure 13: $ARMA(3,2)$ simulado

Encontrando o melhor modelo para a série simulada

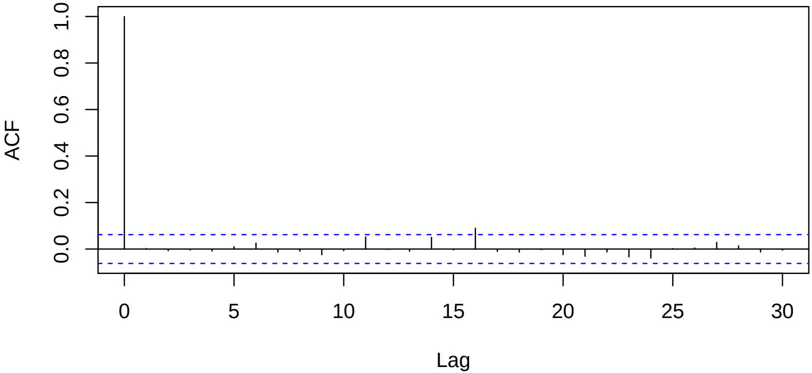
Para encontrarmos o melhor modelo que se ajusta à esta série (simulada), seguimos os seguintes passos:

1. Ajusta o modelo $ARMA(p, q)$ para todas as combinações possíveis de $p = 0, 1, 2, 3, 4$ e $q = 0, 1, 2, 3, 4$;
2. Calcular Log Verossimilhança, AIC e BIC;
3. Escolhe o melhor modelo de acordo com a métrica do passo anterior.

Conclusões sobre o exemplo

- Para todas as combinações possíveis de $p = 0, 1, 2, 3, 4$ e $q = 0, 1, 2, 3, 4$, o melhor modelo que se ajusta à série simulada foi o modelo $ARMA(3, 2)$ com parâmetros $\hat{\phi}_1 = 0.4470, \hat{\phi}_2 = -0.2822, \hat{\phi}_3 = 0.4079, \hat{\theta}_1 = 0.5519$ e $\hat{\theta}_2 = -0.2367$, o que são valores de parâmetros bem próximos dos valores que utilizamos para simular a série $ARMA(3, 2)$.
- O modelo encontrado no passo anterior apresenta menor Log-Verossimilhança, AIC e BIC.
- O correlograma dos resíduos (no próximo slide) se comporta como Ruído Branco, ou seja, não haja nenhuma correlação significativa, indicando que o modelo se ajusta bem aos dados;
- O teste de Ljung-Box aponta $p - value = 0.9963 > 0.05$, o que significa que os resíduos são independentes ao nível de 95% e que, por fim, o modelo $ARMA(3, 2)$ se ajusta bem aos dados simulados.

Correlograma dos resíduos do processo $ARMA(3,2)$ simulado com $\phi_1 = 0.5$, $\phi_2 = -0.25$, $\phi_3 = 0.4$, $\theta_1 = 0.5$ e $\theta_2 = -0.3$



Modelos ARCH/GARCH

Por apresentarem melhores resultados que os modelos $ARMA(p, q)$, esses modelos são muito utilizados nos seguintes casos (WORTHINGTON; HIGGS, 2004):

- Gestão de risco financeiro como, por exemplo, cálculo do VaR (*Value at Risk*);
- Gestão de portfólios (que objetivam uma eficiente alocação de capital/ativos);
- Administração de carteira de ações;
- Precificação de opções, taxas de câmbio, taxa de juros, inflação e etc.

Porque os modelos ARCH/GARCH apresentam melhores resultados?

- Modelos ARCH/GARCH são modelos heterocedásticos e muitos artigos tem demonstrado que esse tipo de modelo lida melhor com as volatilidades, que é o principal parâmetro na gestão de risco de instrumentos/ativos financeiros (MORETTIN; TOLOI, 2004);
- São modelos não-lineares (FRANCQ; ZAKOÏAN, 2010);
- São modelos que assumem que a variância condicional (volatilidade) às informações passadas $X_{t-1}, X_{t-2}, ...$, não é constante e se comporta como um processo $ARMA(p, q)$, ou seja, a variância condicional depende não somente do quadrado das perturbações ε_t passadas, mas também das variâncias condicionais de períodos anteriores (P. WANG, 2014);
- O princípio básico dos modelos GARCH é que, em geral, grandes mudanças tendem a ser seguidas por grandes mudanças e pequenas mudanças tendem a ser seguidas por pequenas mudanças (ENGLE, 1982).

O modelo ARCH(1)

O processo X_t abaixo é um modelo $ARCH(1)$:

$$X_t = \varepsilon_t \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} = \varepsilon_t \sqrt{h_t}. \tag{8}$$

Se elevarmos ao quadrado ambos os lados da expressão 8, então temos:

$$X_t^2 = (\alpha_0 + \alpha_1 X_{t-1}^2)\varepsilon_t^2, \tag{9}$$

o que é muito similar a um processo $AR(1)$ em X_t^2 com ruído multiplicativo de média 1 em vez de ruído aditivo de média de 0.

O modelo ARCH(1) - Continuação

A variância condicional de X_t (volatilidade) dado o passado $X_{t-1}, X_{t-2}, \ldots$ é dado por:

$$\sigma_t^2 = \mathbb{V}[X_t | X_{t-1}, \ldots]$$

e, como ε_t são independentes de X_{t-1} e $\mathbb{E}[\varepsilon_t^2] = \mathbb{V}[\varepsilon_t] = 1$, então

$$\mathbb{E}[X_t | X_{t-1}, \ldots] = 0$$

e

$$\begin{aligned} h_t = \sigma_t^2 &= \mathbb{E}[X_t^2 | X_{t-1}, X_{t-2}, \ldots] \\ &= \mathbb{E}[(\alpha_0 + \alpha_1 X_{t-1}^2)\varepsilon_t^2 | X_{t-1}, X_{t-2}, \ldots] \\ &= (\alpha_0 + \alpha_1 X_{t-1}^2)\mathbb{E}[\varepsilon_t^2 | X_{t-1}, X_{t-2}, \ldots] \\ &= \alpha_0 + \alpha_1 X_{t-1}^2 \Longrightarrow h_t = \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2. \qquad (10) \end{aligned}$$

O modelo ARCH(q) - Forma Geral

A partir da equação 10, podemos escrever a forma geral do modelo $ARCH(q)$, dado por:

$$\begin{aligned} X_t &= \varepsilon_t \sqrt{h_t} = \varepsilon_t \sigma_t, \\ h_t &= \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2 + \ldots + \alpha_q X_{t-q}^2 \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 \end{aligned} \qquad (11)$$

onde $\alpha_0 > 0$ e $\alpha_i \geq 0 \quad \forall i > 0$, o que faz com que $h_t = \sigma_t^2 > 0$. A expressão 11 é a variância condicional (volatilidade) de X_t dado os valores passados $X_{t-1}, X_{t-2}, \ldots$.

Modelo $GARCH(p, q)$

O modelo $GARCH(p, q)$ é uma extensão dos modelos $ARCH(q)$, incluindo termos de variação σ_t^2:

$$h_t = \sigma_t^2 = \underbrace{\alpha_0 + \sum_{i=1}^{q} \alpha_i X_{t-i}^2}_{(1)} + \underbrace{\sum_{j=1}^{p} \beta_j \sigma_{t-j}^2}_{(2)}, \tag{12}$$

onde o componente (1) da equação 12 é o termo ARCH e (2) é o termo GARCH, $\varepsilon_t \sim \mathbb{N}(0, 1)$, ou $\varepsilon_t \sim t_\nu$, $\alpha_0 > 0$, $\alpha_i \geq 0 \quad \forall i > 0$, $\beta_j \geq 0 \quad \forall j > 0$, $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ (esta restrição garante que o modelo tenha covariância estacionária, ou seja, garante que o modelo seja estacionário).

Modelo $GARCH(p, q)$ - Comentários

Na expressão 12, vimos que os parâmetros $\alpha_i, i = 1, 2, ..., q$ estão associados com os termos ao quadrado $X_t^2, X_{t-1}^2, ..., X_{t-q}^2$ e, dessa forma, altos valores de α_i causa alta correlação com X_i^2. Por sua vez, os parâmetros $\beta_j, j = 1, 2, ..., p$ estão associados com os termos de volatilidade $\sigma_{t-1}^2, \sigma_{t-2}^2, ..., \sigma_{t-p}^2$ e, dessa forma, altos valores de β_j causa alta correlação com σ_j^2, dando à variância condicional maior persistência a longo prazo quando a comparamos com os modelos $ARCH(q)$ ou $ARMA(p, q)$.

Modelo $GARCH(1,1)$

O modelo $GARCH(1,1)$ é um dos modelos mais populares e altamente utilizados para estimativa da volatilidade, sendo um modelo dificilmente superado por outros modelos mais sofisticados (HANSEN; LUNDE, 2005). A seguir, o modelo $GARCH(1,1)$:

$$h_t = \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{13}$$

Para se identificar a ordem p e q de um modelo $GARCH(p,q)$, recomenda-se o ajuste de vários modelos de ordens baixas como, por exemplo, $GARCH(1,1)$, $GARCH(1,2)$, $GARCH(2,1)$ e depois escolher o modelo com base em critérios como AIC ou BIC (MORETTIN; TOLOI, 2004).

Simulação $GARCH(1,1)$

A seguir, simulamos um processo $GARCH(1,1)$ com $\alpha_0 = 0.2, \alpha_1 = 0.5, \beta_1 = 0.3$, ou seja, o modelo

$$\begin{aligned} h_t = \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\ &= 0.2 + 0.5 X_{t-1}^2 + 0.2 \sigma_{t-1}^2. \end{aligned} \tag{14}$$

Simulação $GARCH(1,1)$ - Correlograma de h_t

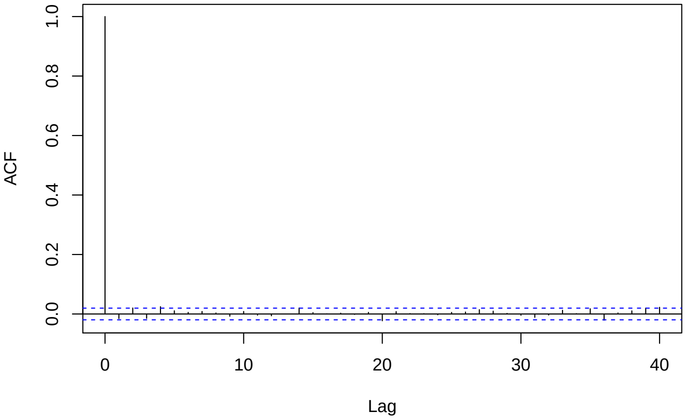
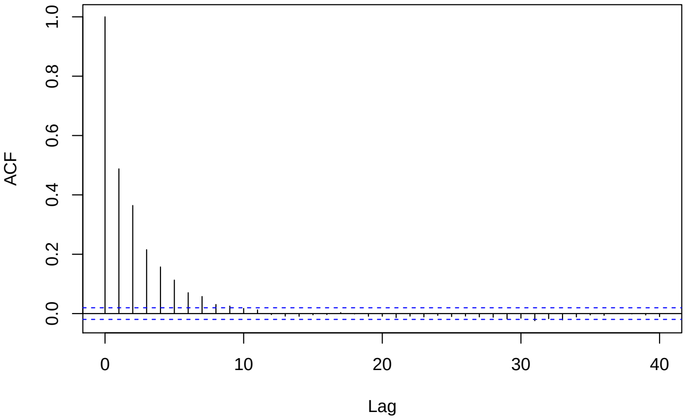


Figure 14: Correlograma de h_t do processo $GARCH(1,1)$

Simulação $GARCH(1,1)$ - Correlograma de h_t^2

Observe que o correlograma abaixo apresenta evidência substancial de heterocedasticidade condicional.



Modelo híbrido $ARMA(4, 3) - GARCH(1, 1)$

Nesta seção, vamos utilizar a série de log-retornos da Oracle para demonstrar que modelos $ARMA(p, q)$ não são adequados para modelar séries temporais financeiras.

- Na primeira fase do processo, encontramos o modelo $ARMA(4, 3)$ como sendo o melhor modelo para a série de log-retornos da Oracle. Ou seja, este é o modelo capaz (ou deveria ser capaz) de capturar ou explicar a linearidade presente na série temporal X_t;
- Na segunda fase do processo, encontramos um modelo $GARCH(1, 1)$ aos resíduos gerados pelo modelo $ARMA(4, 3)$, dando origem ao modelo híbrido $ARMA(4, 3) - GARCH(1, 1)$ que esperamos que seja capaz de capturar o padrão não-linear presente nos resíduos.

Primeira fase: Encontrar o melhor $ARIMA(p, d, q)$ para os dados

Abaixo apresentamos o Correlograma de h_t para o modelo $ARMA(4, 3)$:

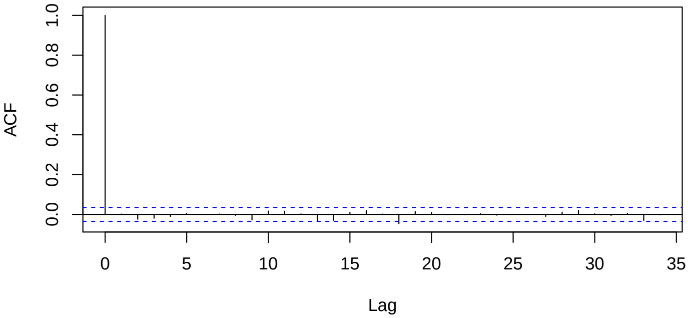


Figure 16: Correlograma de h_t do processo $GARCH(1, 1)$

Primeira fase: Encontrar o melhor $ARIMA(p, d, q)$ para os dados - Continuação

Abaixo apresentamos o Correlograma de h_t^2 para o modelo $ARMA(4, 3)$, onde podemos constatar dependência temporal. Isso é um indicativo de ajuste pobre/insatisfatório aos dados.

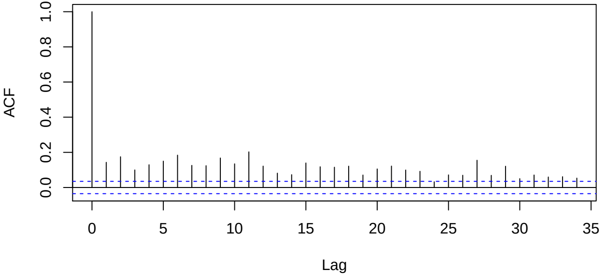


Figure 17: Correlograma de h_t^2 do processo $GARCH(1, 1)$

Segunda fase: Ajustar um modelo $GARCH(1,1)$ aos resíduos do modelo $ARMA(4,3)$

Abaixo apresentamos o Correlograma de h_t. Observe o padrão desejável: Ruído Branco.

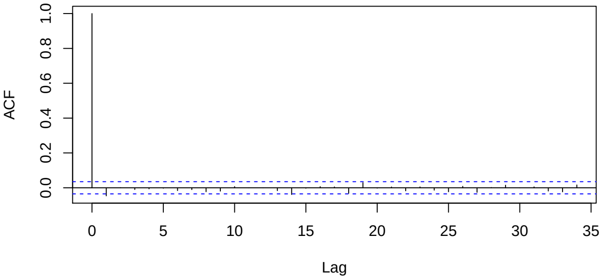


Figure 18: Correlograma de h_t do processo $GARCH(1,1)$

Segunda fase: Ajustar um modelo $GARCH(1,1)$ aos resíduos do modelo $ARMA(4,3)$ - Continuação

Abaixo apresentamos o Correlograma de h_t^2. Observe o padrão desejável: Ruído Branco.

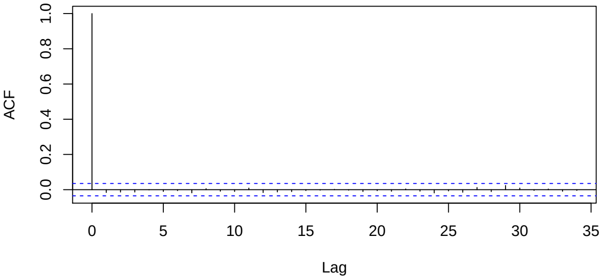


Figure 19: Correlograma de h_t^2 do processo $GARCH(1,1)$

Conclusão: O que eu aprendi com esse seminário?

- Modelos $GARCH(p, q)$ se ajustam e produzem melhores resultados quando estamos analisando séries temporais financeiras, como é o caso da série de log-retornos da Oracle;
- Modelos $ARIMA(p, d, q)$ são inapropriados para modelar séries temporais financeiras devido às características a que chamamos de fatos estilizados (MORETTIN; TOLOI, 2004);
- Vários artigos apresentam comparações com os modelos $GARCH(1, 1)$ e estes, apesar de simples, apresentam resultados superiores aos modelos da família ARIMA (HANSEN; LUNDE, 2005). Inclusive, McNeil et al.(MCNEIL; FREY, 2000) utiliza o modelo $GARCH(1, 1)$ para estimativa do VaR (*Value at Risk*);
- Neste trabalho, $h_t = \sigma^2$ são determinísticas, isto é, sem uma componente de erro aleatório. Uma alternativa aos modelos ARCH/GARCH consiste em assumir que σ_t^2 segue um processo estocástico (Modelos de Volatilidade Estocástica);

Obrigado