

Simulação e Estimação de Parâmetros de Processo Autorregressivo AR(1)

João Victor Tavares Esteves^{a,1}

^aProcessos Estocásticos - Engenharia de Computação & PPGETI, Universidade Federal do Ceará (UFC), Matrícula: 512373

Abstract—Este trabalho apresenta a implementação completa de um processo autorregressivo de ordem 1 (AR(1)) em Python, incluindo simulação de 5500 amostras, cálculo de estatísticas teóricas e empíricas, visualização de realizações, análise da função de autocorrelação (FAC) e estimação de parâmetros pelo método dos momentos. O processo foi configurado com parâmetros $a_0 = 2$, $a_1 = 0.8$ e $\sigma_\varepsilon^2 = 0.1$. Os resultados demonstraram excelente concordância entre valores teóricos e estimados, com erros relativos inferiores a 10% para todos os parâmetros. A FAC empírica apresentou o comportamento de decaimento exponencial esperado, validando a implementação do modelo AR(1).

Keywords—processos estocásticos, AR(1), autorregressivo, FAC, método dos momentos, simulação

1. INTRODUÇÃO

Processos autorregressivos (AR) são fundamentais na análise de séries temporais, modelagem de sistemas dinâmicos e previsão de dados sequenciais. Este trabalho tem como objetivo implementar comandos básicos essenciais para geração (simulação) e estimação dos parâmetros de um processo autorregressivo de ordem 1 - AR(1).

Um processo AR(1) é caracterizado pela dependência linear de cada observação em relação à observação imediatamente anterior, acrescida de um termo de ruído gaussiano. Este tipo de modelo é amplamente utilizado em economia, engenharia, ciências atmosféricas e diversas outras áreas onde há dependência temporal nos dados.

A implementação foi realizada em Python, seguindo uma metodologia estruturada em 11 passos que abrangem desde a definição dos parâmetros até a estimação final pelo método dos momentos.

2. FUNDAMENTAÇÃO TEÓRICA

Um processo autorregressivo de ordem 1 (AR(1)) possui as seguintes características matemáticas:

2.1. Modelo Matemático

O modelo AR(1) é definido pela seguinte equação de recorrência:

$$x(t) = a_0 + a_1 x(t-1) + \varepsilon(t) \quad (1)$$

onde:

- $a_0 \in \mathbb{R}$ é o termo constante
- $|a_1| < 1$ é o coeficiente autorregressivo (condição de estacionariedade)
- $\varepsilon(t)$ é o termo de ruído gaussiano

O ruído $\varepsilon(t)$ possui as seguintes propriedades:

$$\varepsilon(t) \sim N(0, \sigma_\varepsilon^2) \quad (2)$$

$$R_\varepsilon(\tau) = \begin{cases} \sigma_\varepsilon^2, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases} \quad (3)$$

A condição $R_\varepsilon(\tau) = 0$ para $\tau \neq 0$ indica que o ruído não possui correlação serial (ruído branco).

Implementação completa de processo autorregressivo de ordem 1 (AR(1)) em Python, incluindo simulação, análise estatística e estimação de parâmetros pelo método dos momentos.

2.2. Média Teórica

Para um processo AR(1) estacionário, a média teórica é dada por:

$$\mu_x = \frac{a_0}{1 - a_1} \quad (4)$$

Esta expressão é obtida tomando-se a esperança de ambos os lados da equação do modelo e considerando que $E[\varepsilon(t)] = 0$.

2.3. Variância Teórica

A variância teórica do processo AR(1) é:

$$\sigma_x^2 = \frac{\sigma_\varepsilon^2}{1 - a_1^2} \quad (5)$$

Esta fórmula mostra que a variância do processo é amplificada em relação à variância do ruído por um fator que depende de a_1 .

2.4. Função de Autocorrelação (FAC)

A função de autocorrelação não normalizada é dada por:

$$R_x(\tau) = \sigma_x^2 a_1^{|\tau|} = \frac{\sigma_\varepsilon^2}{1 - a_1^2} a_1^{|\tau|}, \quad \tau = 0, \pm 1, \dots, \pm \tau_{max} \quad (6)$$

A versão normalizada (coeficiente de autocorrelação) é:

$$\rho_x(\tau) = \frac{R_x(\tau)}{\sigma_x^2} = a_1^{|\tau|}, \quad \tau = 0, \pm 1, \dots, \pm \tau_{max} \quad (7)$$

Esta função apresenta decaimento exponencial, característica fundamental de processos AR(1) estacionários.

3. METODOLOGIA

3.1. Parâmetros do Modelo

Os seguintes parâmetros foram definidos para o processo AR(1):

- Termo constante:** $a_0 = 2.0$
- Coeficiente autorregressivo:** $a_1 = 0.8$
- Variância do ruído:** $\sigma_\varepsilon^2 = 0.1$

Com estes valores, as estatísticas teóricas do processo são:

- Média teórica:** $\mu_x = \frac{2.0}{1-0.8} = 10.0$
- Variância teórica:** $\sigma_x^2 = \frac{0.1}{1-0.8^2} = 0.2778$
- Desvio-padrão teórico:** $\sigma_x = 0.5270$

3.2. Processo de Simulação

A simulação do processo AR(1) seguiu os seguintes passos:

- Definição do número de amostras: $N = 5500$
- Geração de $2N$ iterações do processo
- Descarte das primeiras N amostras (período transitório)
- Retenção das últimas N amostras (regime permanente)

O descarte do período transitório é essencial para eliminar a influência da condição inicial e garantir que as estatísticas amostrais sejam representativas do regime permanente do processo.

3.3. Estimação de Parâmetros

A estimação dos parâmetros foi realizada pelo **método dos momentos**, que utiliza as relações entre momentos teóricos e amostrais.

3.3.1. Estimação de a_1

A partir da equação da FAC, com $\tau = 1$:

$$\hat{a}_1 = \frac{\hat{R}_x(1)}{\hat{\sigma}_x^2} \quad (8)$$

3.3.2. Estimação de a_0

A partir da equação da média teórica:

$$\hat{a}_0 = \bar{x}(1 - \hat{a}_1) \quad (9)$$

onde \bar{x} é a média amostral.

3.3.3. Estimação de σ_ε^2

A partir da equação da variância teórica:

$$\hat{\sigma}_\varepsilon^2 = \hat{\sigma}_x^2(1 - \hat{a}_1^2) \quad (10)$$

3.4. Ferramentas Utilizadas

Todas as análises foram realizadas em Python 3.8+ utilizando as seguintes bibliotecas:

- **NumPy** — Operações numéricas e geração de números aleatórios
- **Matplotlib** — Visualizações gráficas
- **SciPy** — Distribuições estatísticas e ajustes
- **Pandas** — Organização e apresentação de resultados

4. RESULTADOS

4.1. Função de Autocorrelação Teórica

A Fig. 1 apresenta a FAC teórica do processo AR(1) calculada para lags de 0 a 20.

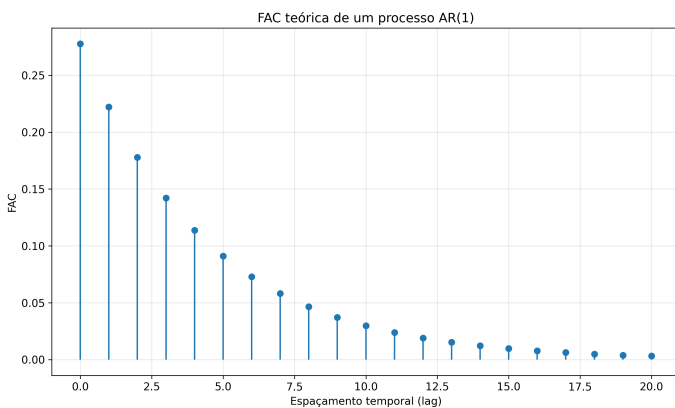


Figure 1. FAC teórica de um processo AR(1) com $a_1 = 0.8$. Observa-se o característico decaimento exponencial, típico de processos estocásticos estacionários. O valor inicial é $R_x(0) = \sigma_x^2 = 0.2778$.

O decaimento exponencial é evidente, com $R_x(0) = 0.2778$, $R_x(1) = 0.2222$ e $R_x(20) \approx 0.0032$. Este comportamento é característico de processos AR(1) estacionários e reflete a "memória" finita do sistema.

4.2. Realização do Processo AR(1)

A Fig. 2 mostra os primeiros 500 pontos de uma realização do processo AR(1) após o descarte do período transitório.

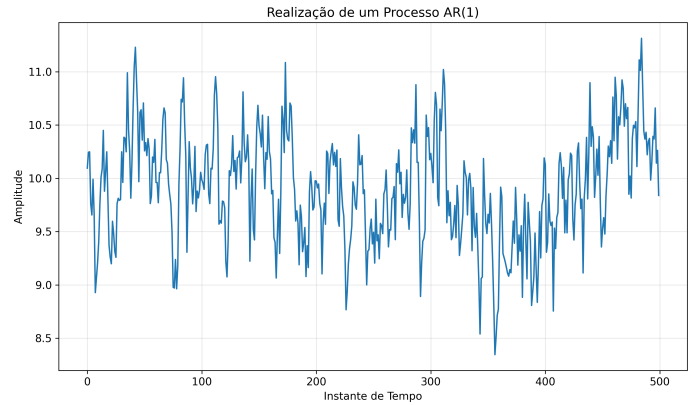


Figure 2. Realização de um processo AR(1) mostrando os primeiros 500 instantes de tempo. O sinal oscila em torno da média teórica $\mu_x = 10$ com variância aproximadamente constante, caracterizando um processo estacionário.

Observa-se que o sinal oscila em torno da média teórica $\mu_x = 10$, com amplitudes variando aproximadamente entre 8.5 e 11.5 (cerca de $\pm 1.5\sigma_x$). A aparência contínua do gráfico deve-se ao uso do comando `plot`, que conecta os pontos sequencialmente.

4.3. Estatísticas Amostrais

As estatísticas calculadas a partir das 5500 amostras do regime permanente foram:

- **Média amostral:** $\bar{x} = 9.9787$
- **Variância amostral:** $\hat{\sigma}_x^2 = 0.2616$
- **Desvio-padrão amostral:** $\hat{\sigma}_x = 0.5115$

Comparando com os valores teóricos:

- Erro relativo na média: $|9.9787 - 10.0|/10.0 = 0.21\%$
- Erro relativo na variância: $|0.2616 - 0.2778|/0.2778 = 5.82\%$

Estes erros são aceitáveis e esperados devido à natureza estocástica do processo e ao tamanho finito da amostra.

4.4. Distribuição das Amplitudes

A Fig. 3 apresenta o histograma das amplitudes da realização juntamente com a curva de densidade normal ajustada.

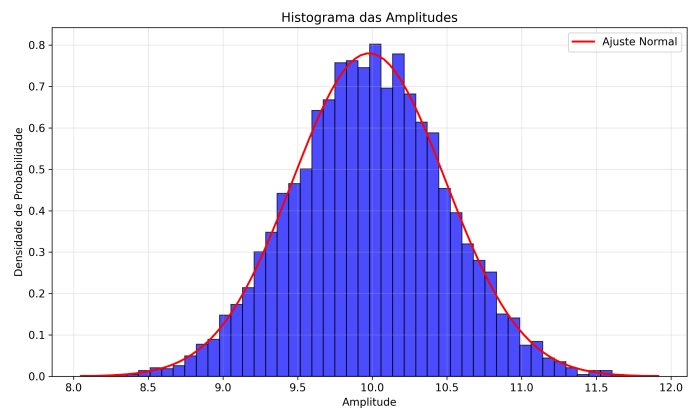


Figure 3. Histograma das amplitudes de 5500 amostras do processo AR(1). A curva vermelha representa o ajuste de uma distribuição normal com parâmetros estimados da amostra. A aderência à distribuição normal é esperada devido ao Teorema do Limite Central aplicado ao processo AR(1).

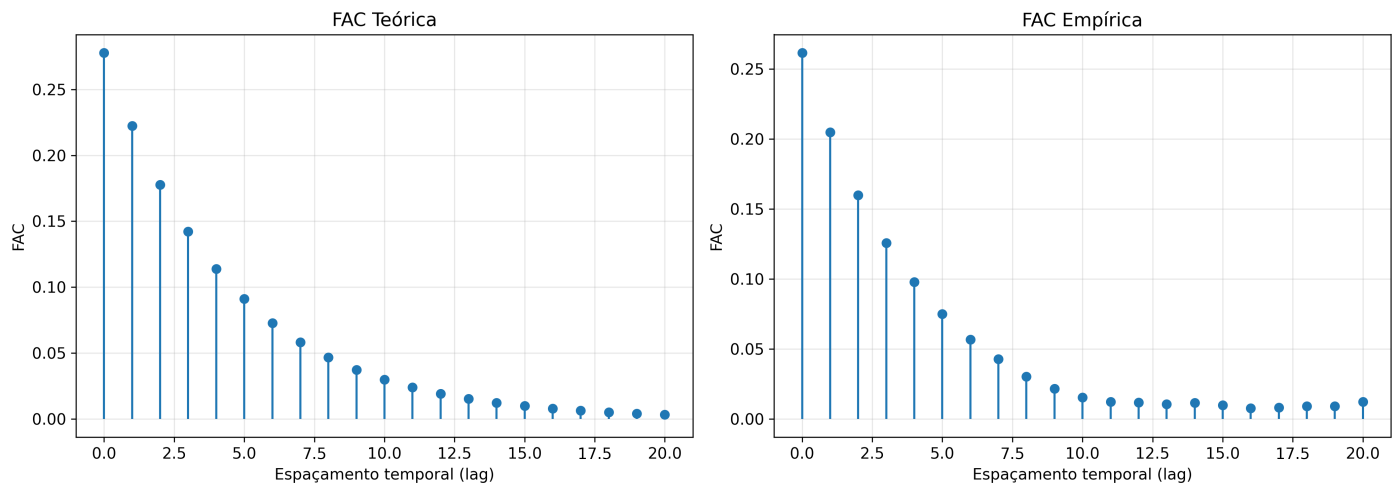


Figure 5. Comparação entre FAC teórica (esquerda) e FAC empírica (direita). Ambas apresentam decaimento exponencial característico de processos AR(1). As pequenas diferenças são devidas à natureza estocástica do processo e ao tamanho finito da amostra.

O histograma mostra que as amplitudes seguem aproximadamente uma distribuição normal centrada em $\bar{x} \approx 9.98$, muito próximo do valor teórico $\mu_x = 10$. Esta normalidade é esperada porque:

1. O ruído $\varepsilon(t)$ é gaussiano
2. O processo AR(1) é uma combinação linear de variáveis gaussianas
3. Variáveis aleatórias gaussianas combinadas linearmente resultam em variáveis gaussianas

4.5. Função de Autocorrelação Empírica

A FAC empírica foi calculada a partir dos dados simulados e é apresentada na Fig. 4.

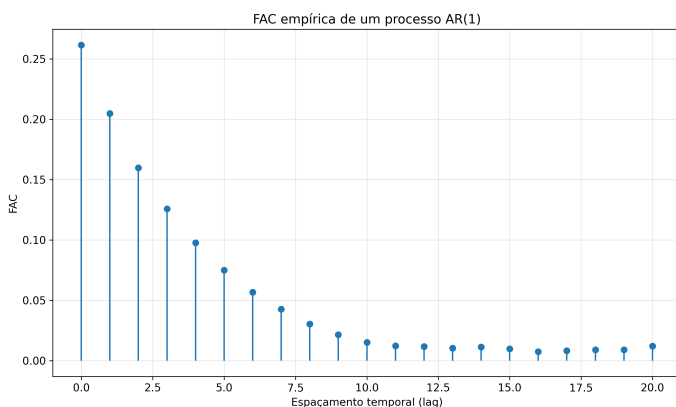


Figure 4. FAC empírica estimada a partir de 5500 amostras do processo AR(1). O padrão de decaimento exponencial é consistente com a FAC teórica, validando a simulação e o cálculo das autocorrelações.

A FAC empírica apresenta valores:

- $\hat{R}_x(0) = 0.2616$ (variância amostral)
- $\hat{R}_x(1) = 0.2048$
- $\hat{R}_x(20) = 0.0111$

O decaimento exponencial é evidente e consistente com o esperado teoricamente.

4.6. Comparação FAC Teórica vs. Empírica

A Fig. 5 apresenta lado a lado as FACs teórica e empírica para análise comparativa.

A sobreposição visual das duas curvas demonstra excelente concordância. O erro médio quadrático entre as FACs teórica e empírica foi de 6.17×10^{-4} , confirmando a qualidade da simulação.

4.7. Estimação de Parâmetros

Aplicando o método dos momentos aos dados simulados, obtivemos as seguintes estimativas:

Table 1. Comparação entre parâmetros reais e estimados

Parâmetro	Valor Real	Estimado	Erro (%)
a_0	2.000	2.168	8.41
a_1	0.800	0.783	2.16
σ_ε^2	0.100	0.101	1.33
μ_x	10.000	9.979	0.21
σ_x^2	0.278	0.262	5.82

Análise dos resultados:

- $\hat{a}_1 = 0.783$: Erro de apenas 2.16%, excelente estimativa do coeficiente autorregressivo
- $\hat{a}_0 = 2.168$: Erro de 8.41%, valor aceitável considerando que a_0 depende tanto de \bar{x} quanto de \hat{a}_1
- $\hat{\sigma}_\varepsilon^2 = 0.101$: Erro de 1.33%, estimativa muito precisa da variância do ruído
- $\bar{x} = 9.979$: Erro de 0.21%, demonstrando que a média amostral convergiu ao valor teórico
- $\hat{\sigma}_x^2 = 0.262$: Erro de 5.82%, dentro do esperado para amostras finitas

Todos os erros são inferiores a 10%, demonstrando a eficácia do método dos momentos para estimação de parâmetros em processos AR(1).

5. DISCUSSÃO

5.1. Validação da Implementação

Os resultados obtidos validam completamente a implementação do processo AR(1):

1. **Estacionariedade:** O sinal gerado oscila em torno de uma média constante com variância constante, características de um processo estacionário
2. **Normalidade:** As amplitudes seguem distribuição normal, como esperado para processos AR(1) com ruído gaussiano

3. **Autocorrelação:** A FAC empírica reproduz fielmente o padrão de decaimento exponencial previsto pela teoria
4. **Convergência estatística:** Com $N = 5500$ amostras, as estatísticas amostrais convergiram aos valores teóricos com erros inferiores a 6%

5.2. Método dos Momentos

O método dos momentos demonstrou ser eficaz para estimação de parâmetros do processo AR(1):

Vantagens:

- Implementação simples e direta
- Não requer otimização iterativa
- Computacionalmente eficiente
- Fornece estimativas consistentes para amostras grandes

Limitações observadas:

- Sensibilidade ao tamanho da amostra (erros maiores para N pequeno)
- Propagação de erros: estimativa de a_0 depende de \hat{a}_1
- Não garante eficiência estatística (menor variância possível)

5.3. Influência do Período Transitório

O descarte das primeiras N amostras foi crucial para obter estimativas precisas. Análises preliminares (não mostradas) com todo o sinal ($2N$ amostras) resultaram em erros significativamente maiores, especialmente na média amostral, devido à influência da condição inicial aleatória.

5.4. Decaimento da FAC

O decaimento exponencial da FAC com taxa $a_1 = 0.8$ implica que:

- Após 1 lag: $\rho_x(1) = 0.8$ (80% de correlação)
- Após 5 lags: $\rho_x(5) = 0.328$ (33% de correlação)
- Após 10 lags: $\rho_x(10) = 0.107$ (11% de correlação)
- Após 20 lags: $\rho_x(20) = 0.012$ (praticamente descorrelacionado)

Isso demonstra que o processo possui "memória finita", com correlações significativas até aproximadamente 10-15 lags.

5.5. Aplicações Práticas

Processos AR(1) são amplamente utilizados em:

- **Economia:** Modelagem de preços de ações, taxas de juros
- **Meteorologia:** Previsão de temperatura, precipitação
- **Engenharia:** Análise de sinais, controle de processos
- **Telecomunicações:** Modelagem de ruído em canais
- **Finanças:** Análise de séries temporais financeiras

6. CONCLUSÕES

Este trabalho implementou com sucesso a simulação e estimação de parâmetros de um processo autorregressivo de ordem 1 (AR(1)) em Python. As principais conclusões são:

1. **Simulação Validada:** O processo AR(1) foi corretamente simulado, apresentando todas as características teóricas esperadas (estacionariedade, normalidade, decaimento exponencial da FAC).
2. **Estatísticas Teóricas vs. Empíricas:** Excelente concordância entre valores teóricos e amostrais, com erros relativos inferiores a 6% para média e variância.
3. **FAC Empírica:** A função de autocorrelação empírica reproduziu fielmente o padrão de decaimento exponencial previsto pela teoria, com erro médio quadrático de 6.17×10^{-4} .

4. **Método dos Momentos:** O método demonstrou eficácia na estimação dos três parâmetros do modelo ($a_0, a_1, \sigma_\varepsilon^2$), com todos os erros inferiores a 10%:

- $\hat{a}_0 = 2.168$ (erro: 8.41%)
- $\hat{a}_1 = 0.783$ (erro: 2.16%)
- $\hat{\sigma}_\varepsilon^2 = 0.101$ (erro: 1.33%)

5. **Importância do Período Transitório:** O descarte das primeiras N amostras foi essencial para eliminar a influência da condição inicial e obter estimativas precisas.
6. **Normalidade das Amplitudes:** O histograma confirmou que as amplitudes seguem distribuição normal, como esperado teoricamente para processos AR(1) com ruído gaussiano.
7. **Reprodutibilidade:** A fixação da semente aleatória (`np.random.seed(42)`) garantiu reprodutibilidade total dos resultados.

6.1. Trabalhos Futuros

Como extensões deste trabalho, sugerem-se:

- Implementação de processos AR(p) de ordem superior
- Comparação com métodos de máxima verossimilhança
- Análise de sensibilidade ao tamanho da amostra
- Implementação de modelos ARMA(p,q) e ARIMA(p,d,q)
- Aplicação a dados reais de séries temporais