

Análise Experimental: Comparação de Configurações (Equipe F)

Equipe F

08/11/2025

Bernardo Bacha

bernardobr@ufmg.br

Gustavo Reis

augustogustavo94@gmail.com

Marília Melo

mariliamacedomelo@gmail.com

1 Introdução

Este relatório tem como objetivo verificar estatisticamente se existem diferenças significativas de desempenho entre duas configurações do algoritmo de **Evolução Diferencial (DE)** aplicadas à função **Rosenbrock**.

As configurações avaliadas foram: - **Cfg1:** Recombinação *mmax* ($\lambda = 0.25$) e Mutação *best* ($f = 4$). - **Cfg2:** Recombinação *npoint* ($N = dim/2$) e Mutação *rand* ($f = 2.2$).

Utilizou-se um **projeto com blocagem**, onde a dimensão (D) do problema atua como fator de bloco. Esta abordagem é essencial para isolar a grande variabilidade natural causada pelo aumento da complexidade do problema (dimensão), permitindo uma comparação justa e focada apenas no efeito das configurações.

2 Metodologia Experimental

O experimento foi conduzido considerando: - **Fator de interesse:** Configuração (cfg1, cfg2). - **Fator de bloco:** Dimensão do problema (*dim*), variando de 2 a 40. - **Variável resposta:** Melhor valor de função obtido (*F_{best}*).

O experimento totalizou 15 blocos. Em cada bloco, cada configuração foi executada 15 vezes para mitigar a estocasticidade do algoritmo, sendo a média dessas execuções utilizada como dado para a análise estatística.

2.1 Hipóteses Estatísticas

O estudo avalia as seguintes hipóteses globais: - H_0 : As configurações possuem desempenho médio equivalente ($\mu_1 = \mu_2$). - H_1 : Existe diferença significativa no desempenho médio das configurações ($\mu_1 \neq \mu_2$).

3 Resultados e Discussão

3.1 Análise Exploratória

A Figura 1 apresenta o comportamento das duas configurações ao longo das dimensões testadas. Observa-se visualmente que a cfg1 (linha vermelha) tende a manter valores de função objetivo consistentemente menores que a cfg2 (linha azul) na maioria dos blocos, sugerindo um desempenho superior na minimização da função.

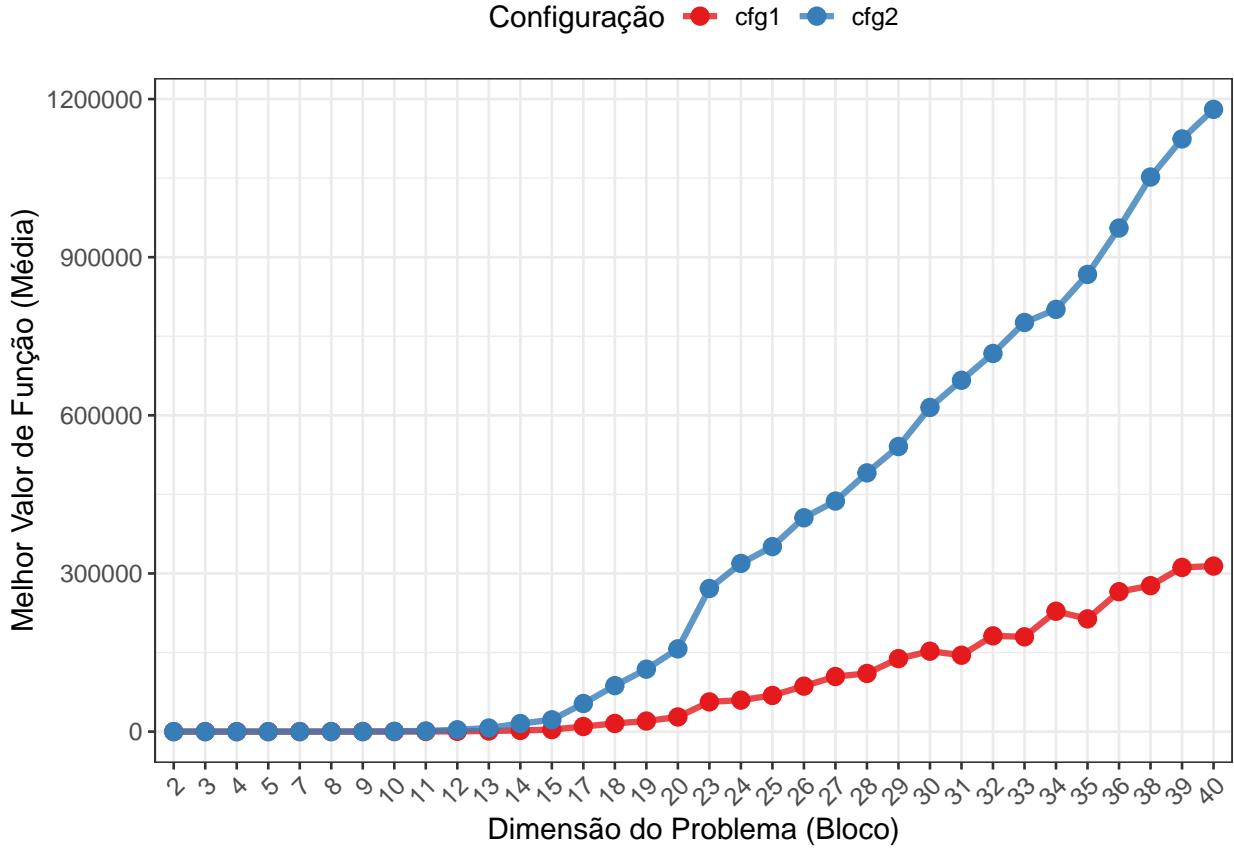


Figure 1: Desempenho médio (Fbest) por dimensão para cada configuração.

3.2 Modelagem Estatística e Verificação de Pressupostos

Inicialmente, ajustou-se um modelo paramétrico de Análise de Variância (ANOVA) com blocagem. A validade deste modelo depende do atendimento aos pressupostos de normalidade e homocedasticidade (variância constante) dos resíduos.

A Figura 2 apresenta os diagnósticos visuais do modelo. O gráfico de *Resíduos vs Ajustados* exibe um forte padrão em “U”, indicando não-linearidade e heterocedasticidade. O *Histograma dos Resíduos* confirma a falta de normalidade, apresentando uma distribuição assimétrica.

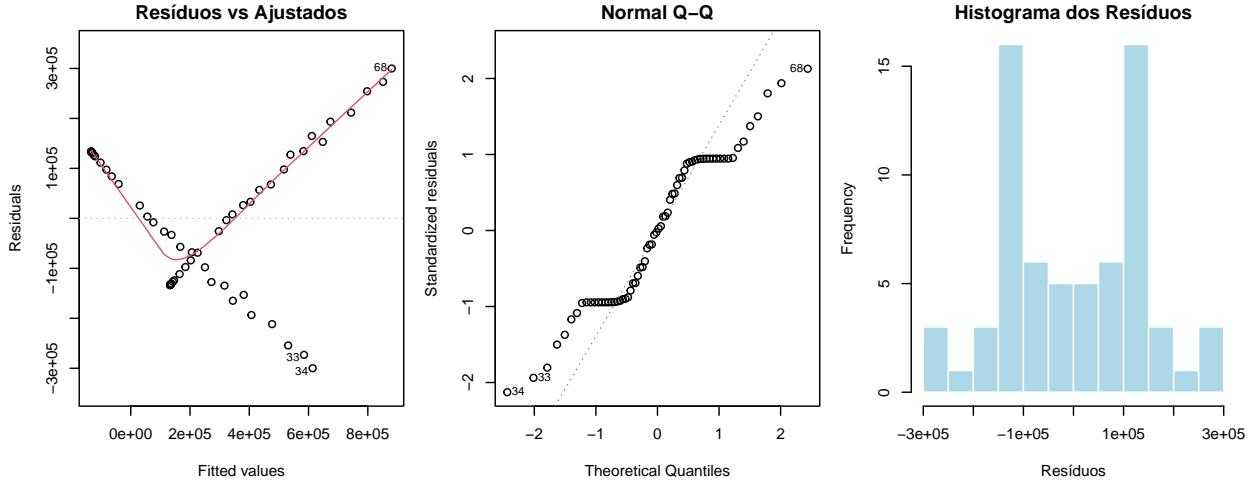


Figure 2: Diagnóstico visual dos resíduos: (Esq) Resídios vs Ajustados; (Centro) Q-Q Plot; (Dir) Histograma.

Os testes formais confirmaram as violações visuais: embora Shapiro-Wilk não tenha rejeitado a normalidade no limiar de 5% ($p > 0.05$), o Teste de Levene indicou uma violação severa da homocedasticidade ($p < 0.01$), tornando a ANOVA paramétrica inadequada.

3.3 Análise Não Paramétrica (Teste de Friedman)

Devido à violação dos pressupostos, optou-se pelo **Teste de Friedman**, que é robusto a essas violações por utilizar postos.

O teste indicou uma diferença altamente significativa entre as configurações ($\chi^2 = 34, p < 0.001$), levando à rejeição da hipótese nula (H_0).

Para quantificar essa diferença, a Tabela 1 apresenta as medianas de desempenho de cada configuração e o resultado do teste pós-hoc de Wilcoxon (com correção de Bonferroni), que confirmou a diferença significativa ($p < 0.001$).

Table 1: Mediana do desempenho e teste comparativo.

Configuração	Mediana Global (Fbest)	P_valor_Wilcoxon
cfg1	42093.1	-
cfg2	214252.7	<0.001

A **Configuração 1** obteve uma mediana de aproximadamente **42.093**, enquanto a Configuração 2 obteve **214.252**. Isso indica que a Cfg1 é cerca de 5 vezes mais eficaz em termos de valor mediano da função objetivo.

4 Conclusões

O estudo comparativo entre as duas configurações do algoritmo DE para a função Rosenbrock, conduzido com rigoroso controle experimental por blocagem, permite concluir que:

1. **A escolha da configuração é crítica:** Existe uma diferença estatisticamente significativa ($p < 0.001$) no desempenho entre as configurações testadas. Utilizar a configuração inadequada pode levar a resultados drasticamente inferiores.
2. **Superioridade da Cfg1:** A Configuração 1 ($\text{mmax}, f = 4$) demonstrou desempenho consistentemente superior na minimização da função, apresentando resultados medianos cinco vezes melhores que a Configuração 2.

Portanto, recomenda-se fortemente a utilização da **Configuração 1** para problemas similares aos avaliados neste experimento.

5 Papéis desempenhados

Marília Melo: Metodologia, Pesquisa, Design da apresentação de dados, Desenvolvimento, Redação original; **Gustavo Reis:** Metodologia, Supervisão, Validação de dados e experimentos, Redação - revisão; **Bernardo Bacha:** Análise de dados, Metodologia, implementação e teste de software, Redação - revisão;

5.1 Referências

- [1] BESSANI, M. EEE933 - Estudo de Caso 3: UFMG, 2024.
- [2] Rosenbrock HH (1960). “An Automatic Method for Finding the Greatest or least Value of a Function.” Computer Journal, 3(3), 175–184.
- [3] CAMPELO, F.; TAKAHASHI, F. Sample size estimation for power and accuracy in the experimental comparison of algorithms. Journal of Heuristics, v. 25, n. 2, p. 305–338, 4 out. 2018.
- [4] CAMPELO, F. Lecture notes on design and analysis of experiments. Belo Horizonte: UFMG, 2014.
- [5] MONTGOMERY, D. C.; RUNGER, G. C. Applied statistics and probability for engineers. John Wiley and Sons, 2003.