



Universidade Federal
de São João del-Rei

Ana Cláudia Machado

MOCHILA BINÁRIA: comparações e análises dos conjuntos de parâmetros

São João Del-Rei - MG

2025

1 . INTRODUÇÃO

O problema da Mochila Binária consiste em um problema de otimização que têm como objetivo maximizar a utilidade do conjunto dos n objetos transportados na mochila, porém, de forma que eles não ultrapassem sua capacidade. Matematicamente, sendo v e p , respectivamente, a utilidade e o peso dos objetos, deseja-se encontrar o subconjunto X de $\{0, 1, 2, \dots, n - 1\}$ que maximize $v(X)$, considerando a restrição $\sum p(X) \leq c$, sendo c a capacidade da mochila.

Assim como nas implementações de algoritmos bioinspirados desenvolvidas anteriormente, a escolha do conjunto de parâmetros é algo essencial para o bom desempenho do algoritmo. Neste trabalho, variou-se os seguintes parâmetros: tamanho da população, número de gerações, taxas de cruzamento e de mutação e, por fim, o método de seleção dos pais, sendo implementado tanto seleção por torneio quanto por roleta. Utilizando duas instâncias diferentes, executa-se o algoritmo 10 vezes com cada combinação de parâmetros, de modo a validar estatisticamente as afirmações sobre o melhor conjunto. Para esse conjunto, realiza-se algumas análises gráficas.

2 . IMPLEMENTAÇÃO E METODOLOGIA

Nesta seção, detalha-se quais parâmetros são fixos e quais seus valores, bem como quais são variados e sua faixa de valores. Como forma de representação, cada mochila é definida por um vetor binário V de n posições, de modo que, para uma posição qualquer i , se $V[i] = 1$, o objeto i faz parte da mochila. Caso $V[i] = 0$, o objeto não faz parte daquela solução.

Já para a escolha dos pais, utiliza-se tanto o método do torneio quanto da roleta (com o uso de um *offset* de valor igual ao módulo menor *fitness* somado a um, de modo que todas os *fitness* se tornem positivos e facilitem a definição da probabilidade de cada solução ser escolhida, proporcionalmente ao ser valor de *fitness*). Para o cruzamento, utiliza-se cruzamento de n pontos, fixando $n = 2$.

Por fim, para as taxas de mutação e cruzamento varia-se, respectivamente, $[0.01, 0.05, 0.1]$ e $[0.6, 0.8, 1]$. Já para o tamanho da população p e o número de gerações g , definiu-se que $p \cdot g = 10000$, de modo que adotou-se as seguintes tuplas de valores (p, g) : $[(25, 400), (50, 200), (100, 100)]$.

Para definir qual o melhor conjunto de parâmetros, executa-se o algoritmo dez vezes com cada conjunto a fim de obter as médias dos seguintes valores: melhor aptidão com o seu desvio padrão e aptidão média com o seu desvio padrão. Por fim, para o melhor conjunto de parâmetros, faz-se duas plotagens gráficas: a do melhor *fitness* em cada execução ao longo das gerações e a média dos valores de todas as execuções de melhor *fitness*, pior *fitness*, média e mediana dos *fitness* ao longo das gerações.

3 . RESULTADOS

Verificou-se o funcionamento do algoritmo com duas [instâncias](#), sendo elas as instâncias 01 e 07. A seguir, são apresentados os resultados para cada uma.

3.1 INSTÂNCIA 01

Nesta instância, têm-se um conjunto de dez objetos, com a capacidade máxima da mochila igual a 165. A solução ótima é dada pela sequência $[1, 1, 1, 1, 0, 1, 0, 0, 0, 0]$, resultando no *fitness* igual a 309 e no peso total 165.

3.1.1 MELHOR CONJUNTO DE PARÂMETROS

Duas tabelas com os resultados descritos na seção 2 foram construídas a fim de analisar a melhor combinação de parâmetros possíveis. A primeira tabela, possui os resultados com a seleção por torneio e, a segunda, com a seleção por roleta. As linhas dessas tabelas foram ordenadas de acordo com o valor da média da melhor aptidão em todas as execuções (com um determinado conjunto de parâmetros). Como tais tabelas demandam muito espaço, suas versões na íntegra estão disponibilizadas junto com as demais saídas do algoritmo. Para fim de análise, as cinco primeiras linhas de cada uma são evidenciadas abaixo:

Taxa de Cruzamento	Taxa de Mutação	Tamanho da População	Número de Gerações	Melhor Aptidão (Média)	Melhor Aptidão (Desvio Padrão)	Aptidão Média (Média)	Aptidão Média (Desvio Padrão)
0.8	0.05	50	200	309.0	0.0	-6742.066	3422.3984072258772
0.8	0.05	100	100	309.0	0.0	-9102.819	2341.907390689383
1.0	0.1	50	200	309.0	0.0	-15760.15	4825.0268570375165
1.0	0.1	25	400	309.0	0.0	-15033.545833333334	5568.850963857641
0.8	0.1	100	100	309.0	0.0	-17402.976	2815.2763237546214

Tabela 1. Cinco melhores combinações de parâmetros - Roleta

Taxa de Cruzamento	Taxa de Mutação	Tamanho da População	Número de Gerações	Melhor Aptidão (Média)	Melhor Aptidão (Desvio Padrão)	Aptidão Média (Média)	Aptidão Média (Desvio Padrão)
0.8	0.05	50	200	309.0	0.0	-6286.318	1442.7743377088154
0.8	0.05	100	100	309.0	0.0	-6241.716	1685.8287844934512
1.0	0.1	50	200	309.0	0.0	-9444.568	1718.329322367126
1.0	0.1	25	400	309.0	0.0	-10320.608333333334	3965.658156848029
0.8	0.1	100	100	309.0	0.0	-10125.155	2478.2783746753175

Tabela 2. Cinco melhores combinações de parâmetros - Torneio

Pode-se perceber algumas semelhanças nas Tabelas 1 e 2, como, por exemplo, os valores altos para a taxa de cruzamento, a qual está entre 0.8 e 1.0 nos dois casos. Também, pode-se perceber que a taxa de mutação fica entre 0.05 e 0.1, não aparecendo o menor valor possível, 0.01. Em todos esses casos a solução ótima foi encontrada em todas as soluções, resultando no desvio padrão da melhor aptidão zerado.

Assim, como forma de desempate, pode-se analisar a aptidão média da última população, bem como seu desvio padrão. Como padrão, percebe-se um alto desvio padrão, com as médias sendo muito negativas, o que demonstra a existência de indivíduos com *fitness* muito ruins, os quais compõem mochilas inviáveis. Dentre todos os casos, o que apresenta melhor aptidão média da última população, bem como o segundo menor desvio padrão é composto pelo seguinte conjunto de parâmetros: seleção

por torneio, com 100 indivíduos, 100 gerações, 0.8 de probabilidade de cruzamento e 0.05 de probabilidade de mutação.

3.1.2 ANÁLISE GRÁFICA PARA O MELHOR CONJUNTO DE PARÂMETROS

Definido o melhor conjunto de parâmetros com os resultados apresentados na seção 3.1.1, realizou-se duas análises gráficas: na primeira, o comportamento do melhor *fitness* ao longo das gerações é analisado para cada uma das execuções, como mostrado na Figura 1.

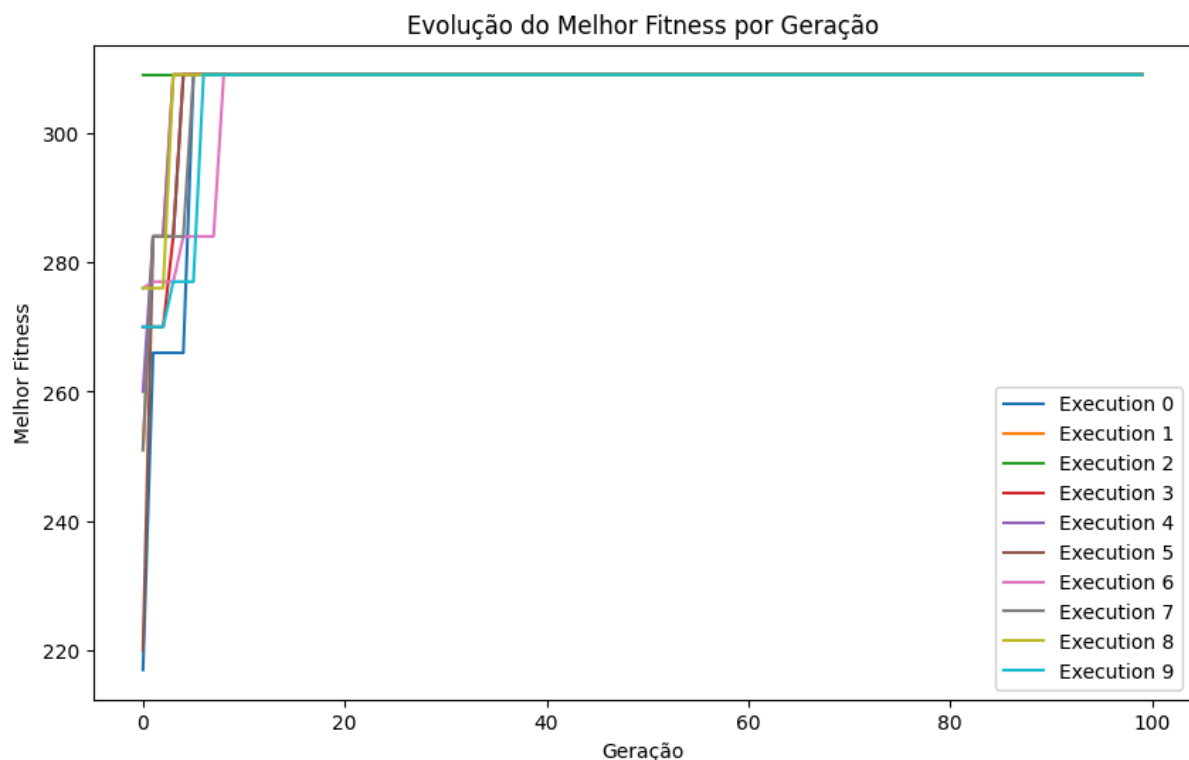


Figura 1. Evolução do melhor fitness por geração

Com isso é possível entender com mais detalhes como o melhor conjunto de parâmetros se comportou em várias execuções. Nesse sentido, em todas as execuções a solução ótima é rapidamente encontrada, necessitando de menos de 20 gerações para isso ocorrer. Nota-se, inclusive, que na execução 2 o melhor resultado foi alcançado logo de início, com a população aleatória definida inicialmente.

A segunda análise gráfica está relacionada com a variação do *fitness* ao longo das gerações. Para isso, calcula-se a média de todas as execuções para os seguintes valores: melhor *fitness*, pior *fitness*, *fitness* médio e mediana do *fitness*. O resultado é apresentado na Figura 2.

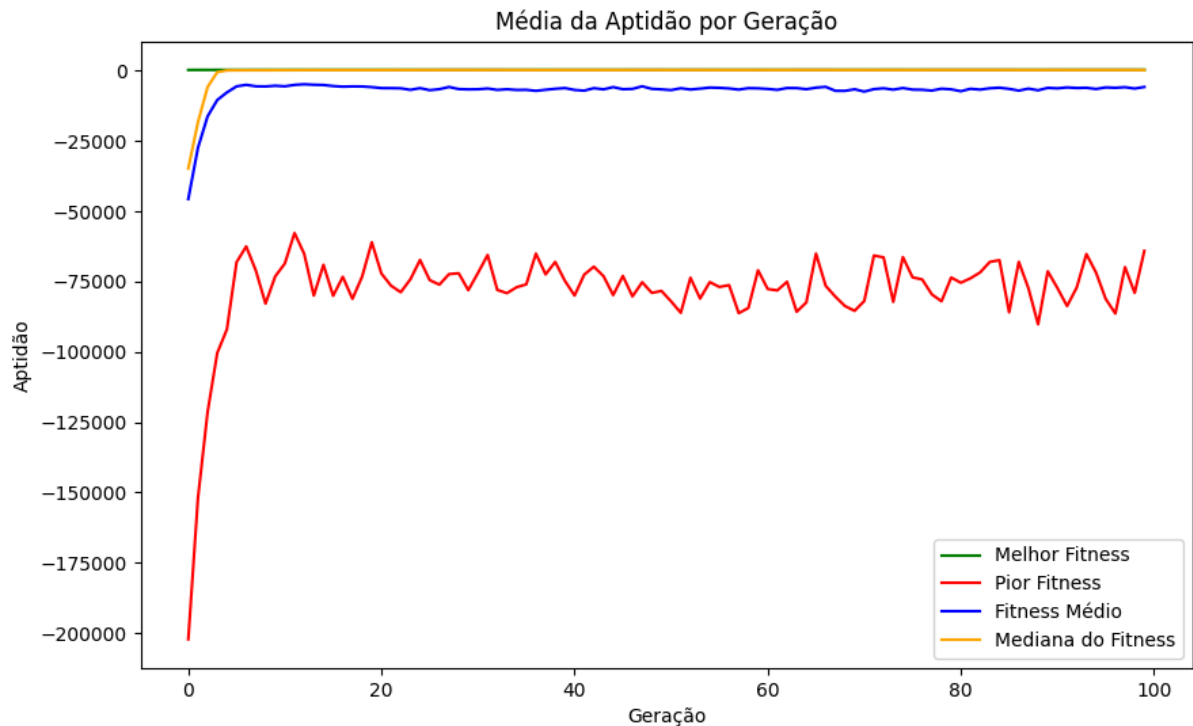


Figura 2. Média de *fitness* por geração

De início, é possível perceber que o pior *fitness* médio se inicia com valores muito baixos, por volta de -200000 , se estabilizando próximo ao valor de -75000 . Isso demonstra que, mesmo alcançando a melhor solução rapidamente, a população ao longo de todas as gerações conta com um conjunto de indivíduos cujas soluções estão muito distantes da adequada. Apesar de isso não representar boas soluções, indica que a variabilidade dos indivíduos gerados é alta, haja vista que soluções não viáveis ainda existem, possibilitando maiores mudanças no espaço de busca.

Já em relação a mediana do *fitness* e ao *fitness* médio, eles possuem valores consideravelmente mais altos, entretanto, ainda no intervalo $[-25000, 0]$.

3.2 INSTÂNCIA 07

A instância 07 possui um conjunto de 15 objetos, com a capacidade máxima da mochila igual a 750. A solução ótima é dada pela sequência $[1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1]$, resultando no *fitness* igual a 1458 e no peso total 749.

3.2.1 MELHOR CONJUNTO DE PARÂMETROS

As tabelas são definidas da forma padronizada na Seção 3.1.1. Assim, abaixo se encontram os resultados para cada conjunto de parâmetros.

Taxa de Cruzamento	Taxa de Mutação	Tamanho da População	Número de Gerações	Melhor Aptidão (Média)	Melhor Aptidão (Desvio Padrão)	Aptidão Média (Média)	Aptidão Média (Desvio Padrão)
0.8	0.1	25	400	1457.2	1.6865480854231356	-44377.708333333336	26363.130664400247
0.8	0.05	25	400	1456.7	1.636391694484477	-28395.204166666666	15493.79890958433
0.8	0.05	50	200	1456.6	2.270584848790187	-16939.638	13663.26137425127
1.0	0.1	50	200	1456.5	2.41522945769824	-45818.966	23074.407144925353
1.0	0.1	25	400	1456.4	2.2211108331943574	-28494.320833333335	18326.0346079293

Tabela 3. Cinco melhores combinações de parâmetros - Roleta

Taxa de Cruzamento	Taxa de Mutação	Tamanho da População	Número de Gerações	Melhor Aptidão (Média)	Melhor Aptidão (Desvio Padrão)	Aptidão Média (Média)	Aptidão Média (Desvio Padrão)
0.6	0.05	50	200	1458.0	0.0	-30267.244000000002	9393.074158136349
0.6	0.05	100	100	1458.0	0.0	-29989.756	8449.893460827643
1.0	0.05	100	100	1457.8	0.6324555320336759	-36392.516	10860.96570494253
0.8	0.1	25	400	1457.4	1.3498971154211057	-49396.254166666666	28955.074553945746
0.6	0.1	50	200	1457.4	1.3498971154211057	-41487.59	14092.614639190748

Tabela 4. Cinco melhores combinações de parâmetros - Torneio

Para essa instância, a definição do melhor conjunto de parâmetros é mais clara, haja vista que em apenas dois a melhor aptidão alcançada é igual a solução ótima. Em relação a esses dois casos, ambos utilizando o torneio como método de seleção, o desvio padrão em ambos é zero, o que indica que todas as execuções realizadas chegaram ao resultado ótimo. Já para a aptidão média da última geração, ambos apresentam valores próximos a 30000, com um alto desvio padrão. Utilizando essa aptidão média como critério de desempate, tem-se que o melhor conjunto de parâmetros é: seleção por torneio, com 100 indivíduos, 100 gerações. 0.05 probabilidade de mutação e 0.6 probabilidade de cruzamento.

3.2.2 ANÁLISE GRÁFICA PARA O MELHOR CONJUNTO DE PARÂMETROS

No gráfico 3, é possível identificar que as execuções necessitam de um número maior de gerações para alcançar o valor ótimo, sendo em, na execução 2, isso foi alcançado apenas após a 80ª geração. Isso, aliado ao fato de que menos conjuntos de parâmetros alcançaram a solução ótima, indica que essa instância é significativamente mais complexa que a anterior, seja pelo maior número de objetos ou pela maior capacidade da mochila.

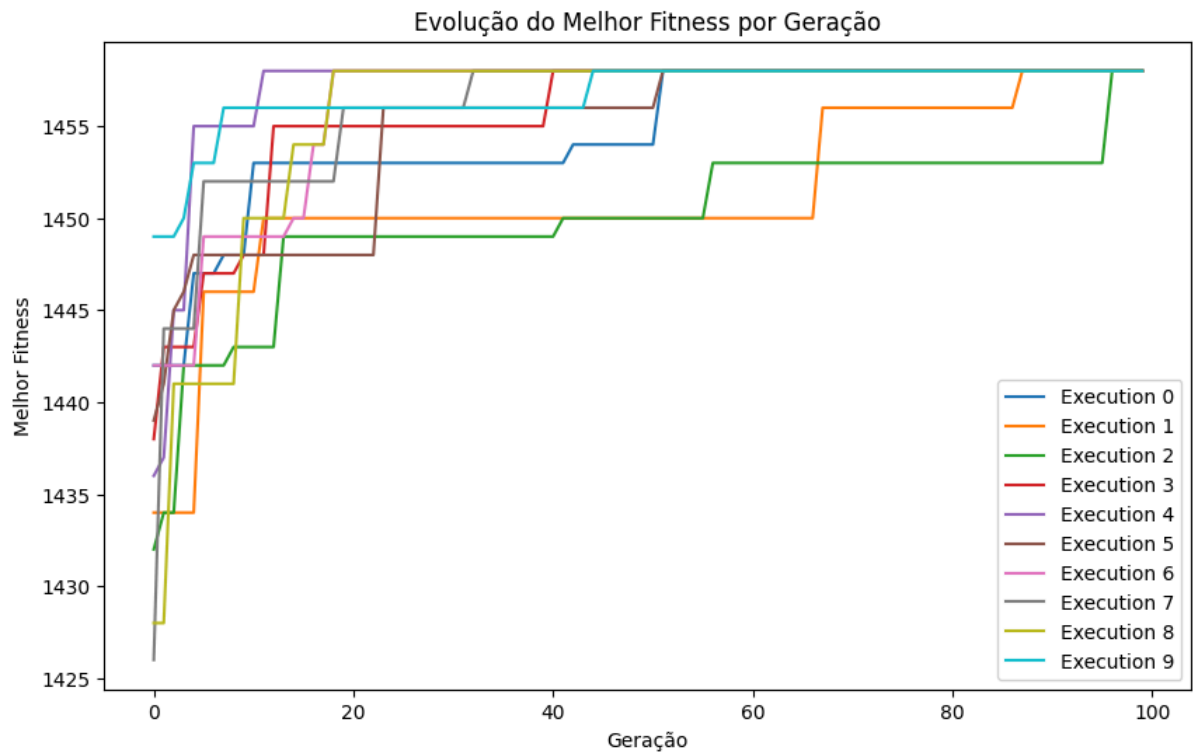


Figura 3. Evolução do melhor fitness por geração

Enfim, na figura 4, para o pior *fitness* observamos valores ainda mais extremos, sendo, no início, próximos a -800000. A estabilização do pior *fitness* ocorre por volta de -500000, o qual ainda é um valor muito aquém do esperado para a solução. Tanto a mediana quanto o melhor *fitness* se sobrepõem em zero, com o *fitness* médio estando muito próximo.

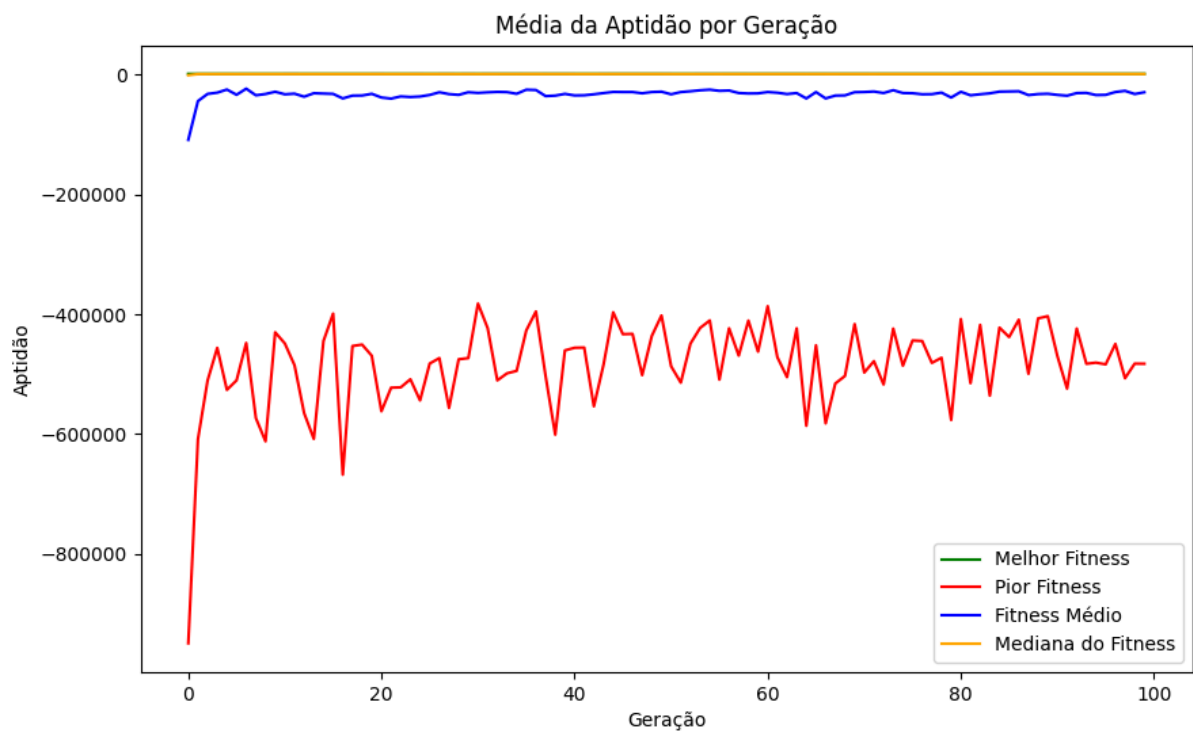


Figura 4. Média de *fitness* por geração

Em conclusão, essa instância apresenta resultados que indicam sua maior complexidade, entretanto, mesmo nesse caso, a implementação realizada obteve sucesso ao encontrar a solução ótima.

4 . CONCLUSÃO

Um rigoroso processo metodológico foi aplicado para definir que o melhor conjunto de parâmetros para cada uma das instâncias analisadas. Conclui-se, com base nos resultados tabulares e gráficos apresentados, que as instâncias possuem diferentes níveis de complexidade, bem como distintos conjuntos de parâmetros que atuam com maior êxito.

Destaca-se, ainda, que a implementação adotada foi eficaz em ambos os casos, permitindo constatar a efetividade da otimização da mochila binária para problemas de maximização.