



Universidade Federal
de São João del-Rei

Ana Cláudia Machado

PROBLEMA DO CAIXEIRO VIAJANTE: otimização na permutação de caminhos

São João Del-Rei - MG

2025

1 . INTRODUÇÃO

O problema do Caixeiro Viajante consiste em um problema de otimização relacionada a permutação de n cidades, com o objetivo de encontrar a menor rota possível, isto é, minimizar a distância. Nesse sentido, dada uma cidade qualquer i , o caixeiro deve realizar um circuito que passe por todas as cidades uma única vez e, ao final, retornar para i .

Matematicamente, dado um conjunto $C = \{c_1, c_2, \dots, c_n\}$ de n cidades, deseja-se encontrar a permutação $\pi \in S_n = \{s: \{1, \dots, n\} \rightarrow \{1, \dots, n\}\}$, de modo que a função objetivo $f: S_n \rightarrow \mathbb{R}$, definida por

$$f(\pi) = \sum_{i=1}^{n-1} \rho(\pi(i), \pi(i+1)) + \rho(\pi(n), \pi(1))$$

seja minimizada.

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 caminho possível é definido por um vetor V de n posições, de modo que, cada posição é uma das cidades. Cada caminho possui um *fitness* associado, o qual é a distância percorrida no circuito.

Já para a escolha dos pais, utiliza-se tanto o método do torneio quanto da roleta, no qual utiliza-se a distância invertida, isto é, para a distância d , $\frac{1}{d}$. Isso é necessário porque se trata de um problema de minimização, em que indivíduos com menor valor de *fitness* devem ter maior probabilidade de serem selecionados.

Para o cruzamento, utiliza-se a técnica de *OX-Crossing* (*Order Crossover*), que garante que os filhos sejam permutações válidas dos caminhos. Isso é feito copiando uma subsequência de um dos pais e preenchendo o restante com as cidades do outro pai, respeitando sua ordem relativa e evitando repetições.

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) : $[(100, 100), (200, 50), (500, 20)]$.

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 a [instância LAU15](#). A seguir, são apresentados seus resultados.

3.1 INSTÂNCIA LAU15

Nesta instância, têm-se um conjunto de quinze cidades, com o menor circuito tendo distância igual a 291. A solução ótima é dada pela sequência $[1, 13, 2, 15, 9, 5, 7, 3, 12, 14, 10, 8, 6, 4, 11]$.

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 na pasta do projeto. 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)
0.6	0.01	100	100	333.5	28.11187965572167	481.088	65.26575906757438
1.0	0.01	100	100	341.2	25.121040318160926	579.061	30.943251086105647
0.8	0.01	100	100	348.2	29.264312737530673	540.699	43.42126909706806
0.8	0.05	100	100	363.8	35.29179948814046	615.498	15.652779518879948
1.0	0.01	200	50	364.6	21.229957868802074	620.0575	10.713511005060646

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)
1.0	0.01	200	50	299.2	11.013123484683575	357.5675	47.28526052058929
1.0	0.01	100	100	299.4	12.285492799775406	320.594	12.93772279464633
0.8	0.01	200	50	302.1	14.62456077213185	333.5425	38.69491477140978
0.8	0.01	100	100	302.4	13.599019572503503	319.084	13.023742592323883
0.6	0.01	100	100	303.7	17.83909564224973	319.491	18.51858012915677

Tabela 2. Cinco melhores combinações de parâmetros - Torneio

Pode-se perceber que esse é um problema mais complexo, pelo fato de que nenhuma das combinações de parâmetros encontrou em todas as execuções a solução ótima. Nesse sentido, têm-se que a combinação de parâmetros que chegou mais próximo encontra-se na *Tabela 2*, a qual é referente

ao torneio. Tendo as taxas de cruzamento e de mutação, respectivamente, os valores 1.0 e 0.01 ; uma população com 200 indivíduos e com 50 gerações, essa combinação de parâmetros alcançou um melhor *fitness* médio de 299.2 , com o menor desvio padrão, 11.01 . Entretanto, é interessante observar que, nesse caso, tanto a média quanto o desvio padrão do *fitness* médio da última população não configuram entre os melhores apresentados. Inclusive, esses valores são os maiores da *Tabela 2*, indicando uma maior dispersão nas distâncias dos caminhos encontrados na última iteração.

Outro ponto a ser considerado é o fato de que, para este problema, os resultados apresentados para o método de seleção da roleta, *Tabela 1*, são consideravelmente piores do que o que foi alcançado para o torneio, na *Tabela 2*. Assim, conclui-se que o melhor método de seleção é o torneio, com o conjunto de parâmetros: 200 indivíduos, 50 gerações, 1.0 de probabilidade de cruzamento e 0.01 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.

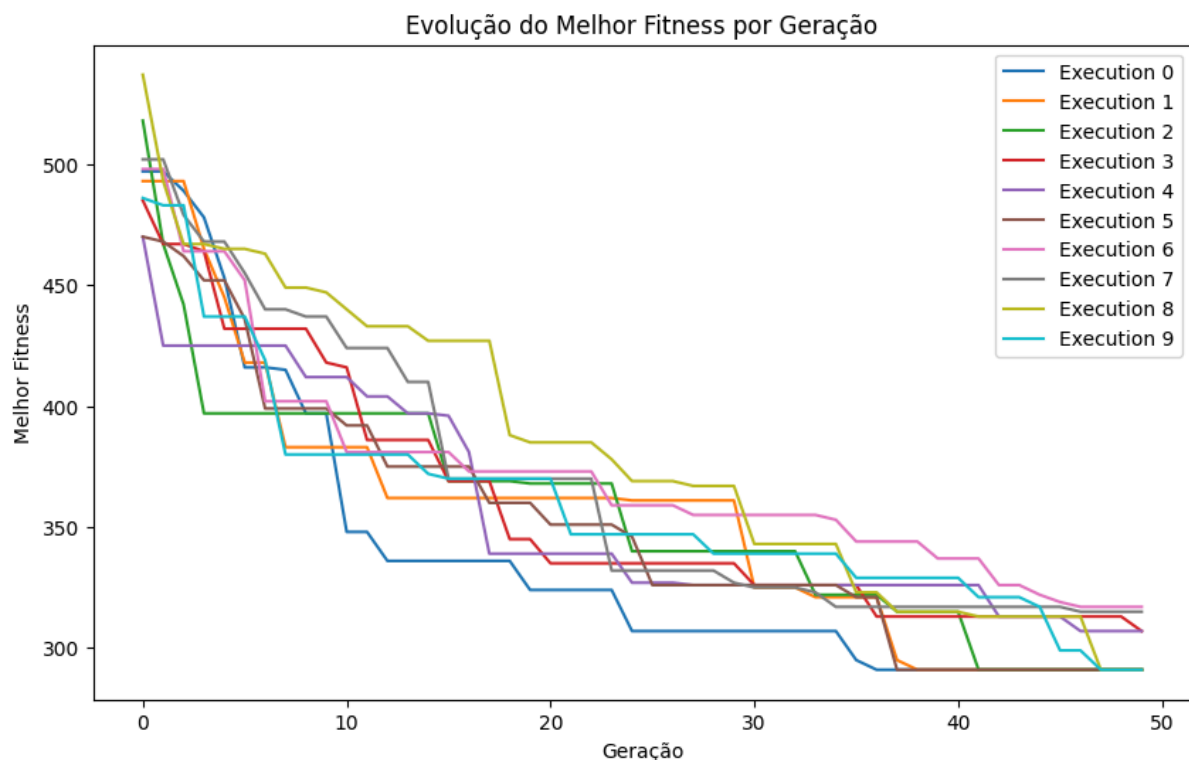


Figura 1. Evolução do melhor fitness por geração

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 demora a ser

encontrada, sendo que, em alguns casos, isso não ocorre. Entretanto é possível notar que para todas as execuções são encontradas distâncias próximas a 300.

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.

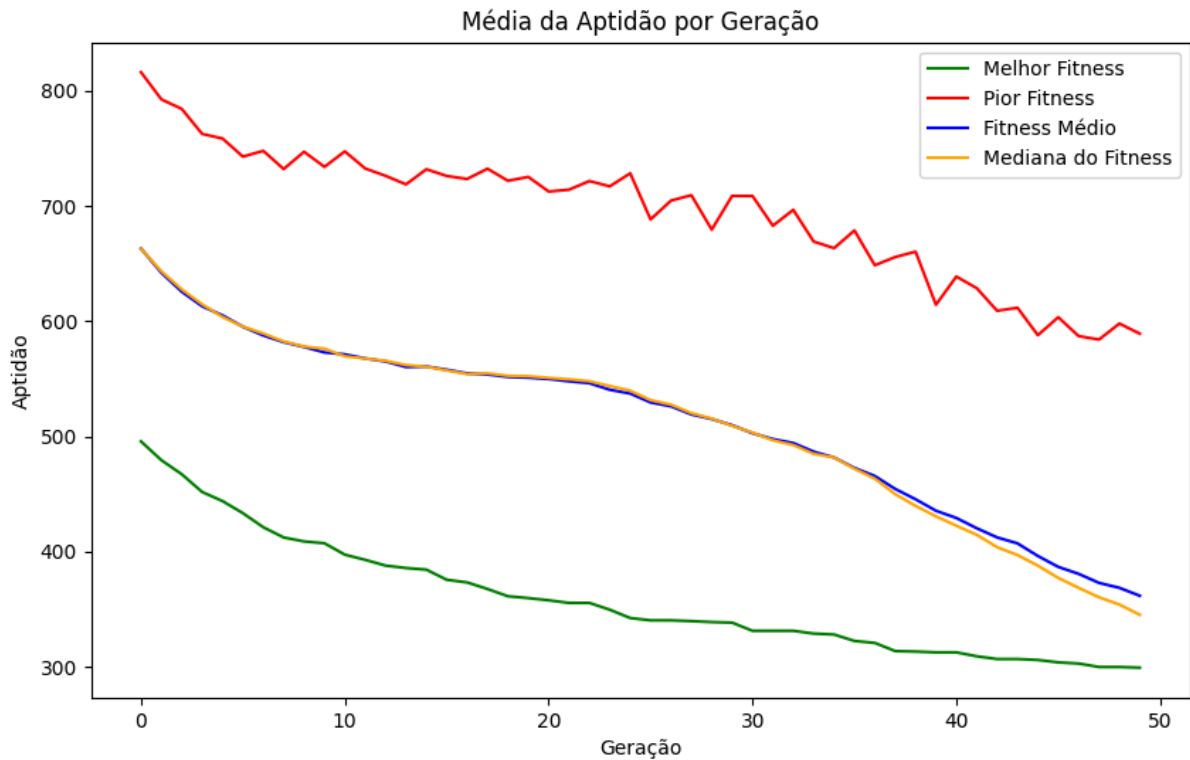


Figura 2. Média de *fitness* por geração

De início, é possível perceber que o pior *fitness* médio se inicia com valores muito altos, por volta de 800, decaindo continuamente ao longo das gerações. Isso demonstra que, mesmo alcançando soluções próximas à ótima, 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 muito próximos, decaindo de forma muito semelhantes. Por fim, para o melhor *fitness*, este rapidamente decai para o intervalo [300, 400], apresentando, após a trigésima geração, valores muito próximos de 300.

4. CONCLUSÃO

O problema do Caixeiro Viajante é um conhecido problema de otimização, o qual trabalha com a permutação de cidades, buscando a distância mínima entre elas. Para isso, utiliza diversos parâmetros, tais como, taxas de mutação e de cruzamento, tamanho da população e número de gerações.

Para avaliar essas combinações de parâmetros, um rigoroso processo metodológico foi aplicado para definir o melhor dentre esses conjuntos. Com base nos resultados tabulares e gráficos apresentados, conclui-se que este problema apresenta maior complexidade para a obtenção de soluções ótimas. Ainda assim, é possível identificar um conjunto de parâmetros que se destacam em desempenho.