



**Universidade de Brasília**  
**Departamento de Estatística**

**Otimização na Alocação de Docentes**  
**Uma Proposta de Ferramenta Computacional para o Departamento de**  
**Estatística da UnB**

**Karl Peixoto**

Projeto apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília**  
**2025**

**Karl Peixoto**

**Otimização na Alocação de Docentes  
Uma Proposta de Ferramenta Computacional para o Departamento de  
Estatística da UnB**

Orientador(a): Eduardo Monteiro

Relatório apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília  
2025**

---

incluir dedicatória (opcional)

---

# Agradecimentos

- Escrever os agradecimentos...

---

# **Resumo**

Texto...

Palavras-chaves:

## **Lista de Tabelas**

1	Distribuição das preferências por disciplina . . . . .	26
2	Distribuição das preferências por docente . . . . .	28
3	Resultados do estudo de convergência . . . . .	30
5	Hiperparâmetros fixados para o Algoritmo Genético . . . . .	34
6	Resumo das 10 execuções do algoritmo de colônia de formiga. . . . .	37
7	Resumo organizado das execuções do AG (Apêndice). . . . .	42

# Lista de Figuras

1	Gráfico de linhas do estudo de convergência do algoritmo genético . . . . .	30
2	Distribuição do Fitness Final . . . . .	32
3	Gráficos de dispersão dos hiperparâmetros pelo tempo e pelas falhas . . . . .	33
4	Gráficos de dispersão dos hiperparâmetros pelo tempo e pelas falhas . . . . .	34
5	Gráficos de dispersão dos hiperparâmetros pelo tempo e pelas falhas . . . . .	35
6	Histograma do tempo até convergência com os parâmetros otimizados . . . . .	35
7	Boxplot da dispersão dos tempos de convergência . . . . .	35
8	Gráficos de linhas do estudo de convergência do ACO . . . . .	36

# Sumário

<b>1 Introdução . . . . .</b>	<b>8</b>
<b>2 Referencial Teórico . . . . .</b>	<b>9</b>
<b>3 Metodologia . . . . .</b>	<b>10</b>
3.1 Programação Linear Inteira . . . . .	10
3.1.1 Características da PLI . . . . .	10
3.1.2 Vantagens e Desvantagens . . . . .	11
3.1.3 Aplicação ao problema . . . . .	11
3.1.4 Componentes do Modelo . . . . .	11
3.1.5 Função Objetivo . . . . .	12
3.1.6 Restrições do Modelo . . . . .	13
3.2 Algoritmos Genéticos . . . . .	14
3.2.1 Representação e População . . . . .	14
3.2.2 Função de Avaliação . . . . .	15
3.2.3 Operadores Genéticos . . . . .	15
3.2.4 Estrutura Geral do Algoritmo . . . . .	15
3.2.5 Critérios de Parada . . . . .	15
3.2.6 Vantagens e Desvantagens . . . . .	16
3.2.7 Aplicação de Algoritmos Genéticos à Alocação de Docentes . . . . .	16
3.2.8 Representação do Indivíduo (Cromossomo) . . . . .	16
3.2.9 Função de Avaliação (Fitness Function) . . . . .	17
3.2.10 Operadores Genéticos Aplicados . . . . .	18
3.3 Algoritmos de Colônias de Formigas . . . . .	19
3.3.1 Princípios Fundamentais . . . . .	19
3.3.2 Componentes do ACO . . . . .	20
3.3.3 Procedimento Geral do ACO . . . . .	20

3.4 Aplicação de Algoritmos de Colônia de Formigas (ACO) . . . . .	21
3.4.1 Estrutura do Problema e Representação . . . . .	21
3.4.2 Componentes do Modelo ACO . . . . .	21
3.5 Comparação entre Algoritmos de Otimização . . . . .	23
<b>4 Resultados . . . . .</b>	<b>25</b>
4.1 Caracterização do Cenário de Estudo . . . . .	25
4.1.1 Análise da Procura por Disciplinas . . . . .	26
4.1.2 Conclusão . . . . .	29
4.2 Análise e Calibração do Algoritmo Genético . . . . .	29
4.2.1 Estudo de Convergência . . . . .	29
4.2.2 Análise de sensibilidade dos hiperparâmetros . . . . .	31
4.2.3 Conclusão . . . . .	34
4.3 Análise e Calibração do Algoritmo de Colônia de Formigas . . . . .	36
4.3.1 Estudo de Convergência . . . . .	36
4.3.2 Análise de sensibilidade dos hiperparâmetros . . . . .	38
<b>5 Conclusão . . . . .</b>	<b>39</b>
<b>Referências . . . . .</b>	<b>40</b>
<b>Apêndice . . . . .</b>	<b>41</b>
<b>A Tabela Hiperparâmetros Ajuste - AGs . . . . .</b>	<b>41</b>
<b>B nome apêndice 2 . . . . .</b>	<b>43</b>
<b>Anexo . . . . .</b>	<b>44</b>
<b>A nome anexo 1 . . . . .</b>	<b>44</b>

## 1 Introdução

A alocação de professores a disciplinas é uma atividade fundamental para o funcionamento eficiente de instituições de ensino superior. Esse processo, embora rotineiro, apresenta elevada complexidade devido à necessidade de conciliar diversas restrições e preferências, tais como carga horária dos docentes, compatibilidade com as disciplinas, restrições curriculares e disponibilidade de salas e horários. Além disso, é desejável que a satisfação dos professores seja considerada, equilibrando a distribuição das atividades e contribuindo para um ambiente de trabalho mais justo e eficiente.

Esse problema está inserido no contexto mais amplo dos problemas de *timetabling* e *scheduling*, os quais vêm sendo extensivamente estudados desde as décadas de 1960 e 1970, especialmente com o avanço da Pesquisa Operacional e da Inteligência Artificial. Segundo Burke e Petrovic (2002), o problema de alocação de aulas está entre os mais desafiadores na área de otimização combinatória, e a literatura especializada sugere uma ampla gama de abordagens, incluindo Programação Linear Inteira, Algoritmos Genéticos, Busca Tabu, Colônias de Formigas e Algoritmos Híbridos.

No contexto brasileiro, a maioria das alocações é feita manualmente ou com o uso de planilhas, o que limita a eficiência do processo. Este projeto propõe o desenvolvimento de uma ferramenta computacional de apoio à tomada de decisão para o Departamento de Estatística da Universidade de Brasília, que automatize a alocação de professores considerando restrições operacionais e preferências individuais. A solução será baseada em algoritmos de otimização, com comparação entre diferentes abordagens clássicas e metaheurísticas.

## **2 Referencial Teórico**

Aqui você deve escrever a Revisão teórica ou incluir na Metodologia.

### 3 Metodologia

A proposta inicial para o estudo é a comparação da utilização de 3 algoritmos de otimização diferentes, sendo um deles exato (Programação Linear Inteira) e os outros dois sendo heurísticos, ou seja, que selecionam a melhor opção dentro das apresentadas, mas não garantem a convergência para a melhor solução global.

Os métodos propostos seguem com uma breve descrição abaixo:

#### 3.1 Programação Linear Inteira

A Programação Linear Inteira (PLI) é uma técnica de otimização utilizada para resolver problemas nos quais as variáveis de decisão assumem apenas valores inteiros, geralmente binários (0 ou 1). Trata-se de um caso particular da Programação Linear (PL), no qual se impõe restrições de integralidade às variáveis.

A estrutura geral de um problema de PLI pode ser expressa da seguinte forma:

$$\text{Minimizar (ou Maximizar)} \quad c^\top x \quad (3.1.1)$$

$$\text{sujeito a} \quad Ax \leq b \quad (3.1.2)$$

$$x \in \mathbb{Z}^n \quad (3.1.3)$$

onde:

- $x$  é o vetor de variáveis de decisão;
- $c$  é o vetor de coeficientes da função objetivo;
- $A$  é a matriz de coeficientes das restrições;
- $b$  é o vetor de constantes do lado direito das restrições;
- $\mathbb{Z}^n$  denota que as variáveis são inteiras.

##### 3.1.1 Características da PLI

Problemas de PLI são comuns em contextos nos quais decisões discretas precisam ser tomadas, como alocação de recursos, escalonamento, roteamento e planejamento.

Uma classe frequente de variáveis são as binárias, indicadoras de escolha ou ativação de determinada alternativa:

$$x_i \in \{0, 1\}$$

Esses modelos permitem a formulação de restrições complexas, tais como exclusões mútuas, limites mínimos/máximos, prioridades e preferências, por meio de combinações lineares das variáveis.

### 3.1.2 Vantagens e Desvantagens

A principal vantagem da PLI é a obtenção de soluções exatas e interpretáveis. No entanto, a escalabilidade pode ser uma limitação, especialmente quando o número de variáveis e restrições cresce, o que pode demandar grande poder computacional.

### 3.1.3 Aplicação ao problema

A alocação de docentes às disciplinas é um problema clássico de otimização combinatória que pode ser modelado por meio da PLI. A modelagem a seguir tem como propósito encontrar uma distribuição de disciplinas que maximize a satisfação do corpo docente, representada por suas preferências de alocação, ao mesmo tempo que respeita as restrições operacionais de conflito de horário entre as disciplinas e limite de carga horária máxima.

### 3.1.4 Componentes do Modelo

Para a construção do modelo matemático, definem-se os seguintes componentes, baseados nos dados de entrada e nas decisões a serem tomadas.

#### Conjuntos e Índices.

- $p \in P$ : Representa cada professor do departamento disponível para lecionar disciplinas.
- $d \in D$ : Representa cada disciplina ofertada pelo departamento em um dado semestre.

## Parâmetros.

Os parâmetros constituem os dados de entrada do modelo, extraídos de fontes como o questionário de preferências e os sistemas acadêmicos da universidade.

- $pref_{pd}$ : Valor numérico da preferência do professor  $p$  em lecionar a disciplina  $d$ .
- $ch_d$ : Carga horária, em horas por semestre, da disciplina  $d$ .
- $CH_{max,p}$ : Carga horária máxima, em horas por semestre, que pode ser atribuída ao professor  $p$ .
- $W$ : Coeficiente de penalidade, um valor escalar utilizado para realizar penalizações sobre alocações indesejadas.
- $Cf_{d_i d_j} \in \{0, 1\}$ : Parâmetro binário que assume valor 1 se as disciplinas  $d_i$  e  $d_j$  são ofertadas em horários conflitantes, e 0 caso contrário.

## Variáveis de Decisão.

As variáveis de decisão representam as escolhas que o modelo deve realizar.

- $x_{pd} \in \{0, 1\}$ : Variável binária que assume valor 1 se o professor  $p$  for alocado para a disciplina  $d$ , e 0 caso contrário.
- $y_p \in \mathbb{Z}^+$ : Variável inteira auxiliar que quantifica o número de alocações de baixo interesse para o professor  $p$ .

### 3.1.5 Função Objetivo

A função objetivo, apresentada na Equação 3.1.4, é formulada para maximizar o escore total de satisfação dos docentes, ao mesmo tempo que minimiza a ocorrência de alocações consideradas indesejáveis. O termo à esquerda maximiza a soma das preferências das alocações efetivadas. O termo à direita introduz uma penalidade, ponderada pelo coeficiente  $W$ , para cada alocação de baixo interesse realizada.

$$\text{Maximizar } Z = \left( \sum_{p \in P} \sum_{d \in D} pref_{pd} \cdot x_{pd} \right) - W \cdot \left( \sum_{p \in P} y_p \right) \quad (3.1.4)$$

A magnitude do coeficiente  $W$  é crucial para o comportamento do modelo. Um valor adequadamente alto garante que o otimizador evite violações das restrições brandas, a menos que seja estritamente necessário para se obter uma solução viável.

A abordagem proposta utiliza  $W$  como um único valor que penaliza apenas o número de violações de restrições brandas de forma igual. No entanto, é possível propor a criação de  $W_1, W_2, \dots, W_n$ , que apresentem pesos diferentes para violação de cada conjunto de restrições brandas. E também é possível propor  $W$  como um vetor, que tenha um valor diferente para cada professor, possibilitando a criação de uma hierarquia entre as preferências que favoreça professores que tenham outros encargos no departamento além das disciplinas, ou até mesmo utilizar isso para fazer correções ao longo do tempo e garantir que professores que foram menos favorecidos em semestres anteriores tenham suas preferências atendidas.

### 3.1.6 Restrições do Modelo

As restrições definem o espaço de soluções viáveis, garantindo que apenas alocações válidas sejam consideradas. Elas são divididas em restrições duras (*hard constraints*), que devem ser obrigatoriamente satisfeitas, e brandas (*soft constraints*), que são desejáveis, mas podem ser violadas a um custo.

#### Restrições Duras (Hard Constraints).

- Garantia de Oferta de Disciplinas:** Cada disciplina do conjunto de oferta  $D$  deve ser atribuída a exatamente um professor. Esta restrição, definida na Equação 3.1.5, assegura que toda a oferta de disciplinas do semestre seja coberta, e que nenhuma disciplina seja alocada para dois professores diferentes.

$$\sum_{p \in P} x_{pd} = 1, \quad \forall d \in D \quad (3.1.5)$$

- Limite de Carga Horária Docente:** A soma das cargas horárias das disciplinas alocadas a um professor não pode exceder sua capacidade máxima de trabalho, conforme a Equação 3.1.6.

$$\sum_{d \in D} ch_d \cdot x_{pd} \leq CH_{max,p}, \quad \forall p \in P \quad (3.1.6)$$

- Restrição de Conflito de Horário:** Para garantir que um docente não seja

alocado a duas disciplinas que ocorrem simultaneamente, a seguinte restrição é imposta para cada professor  $p$  e para cada par de disciplinas  $(d_i, d_j)$  com conflito de horário.

$$x_{pd_i} + x_{pd_j} \leq 1, \quad \forall p \in P, \forall (d_i, d_j) \text{ tal que } Cf_{d_id_j} = 1 \quad (3.1.7)$$

### **Restrições Brandas (Soft Constraints).**

3. **Minimização de Alocações de Baixo Interesse:** O modelo deve evitar, sempre que possível, alocar um professor a uma disciplina pela qual ele demonstrou baixo interesse (e.g.,  $pref_{pd} = 1$ ). A Equação 3.1.8 conecta a variável de alocação  $x_{pd}$  à variável de penalidade  $y_p$ .

$$\sum_{d \in D | pref_{pd}=1} x_{pd} \leq y_p, \quad \forall p \in P \quad (3.1.8)$$

Se nenhuma disciplina de baixo interesse for alocada ao professor  $p$ , o somatório à esquerda será zero, permitindo que  $y_p$  seja zero e nenhuma penalidade seja incorrida. Contudo, se uma ou mais dessas disciplinas forem alocadas,  $y_p$  será forçado a ser no mínimo 1, ativando a penalidade  $W$  na função objetivo.

## **3.2 Algoritmos Genéticos**

Algoritmos Genéticos (AGs) são técnicas de otimização e busca inspiradas nos princípios da seleção natural e da genética biológica, propostos inicialmente em 1975 e adaptado em Holland (1992). São amplamente utilizados para resolver problemas complexos, especialmente aqueles com espaços de solução grandes, não lineares e com múltiplos ótimos locais.

### **3.2.1 Representação e População**

A primeira etapa na construção de um AG é a definição da representação dos indivíduos, que codificam possíveis soluções do problema. Essa codificação pode ser feita por vetores binários, inteiros, reais ou estruturas mais complexas, dependendo do problema.

Uma *população inicial* é então gerada aleatoriamente, composta por  $N$  indivíduos (soluções candidatas).

### 3.2.2 Função de Avaliação

Cada indivíduo é avaliado por meio de uma *função de aptidão* (*fitness function*), que quantifica a qualidade da solução em relação ao objetivo do problema. Essa função guia o processo evolutivo, favorecendo soluções melhores ao longo das gerações.

### 3.2.3 Operadores Genéticos

A evolução da população ocorre por meio de operadores genéticos:

- **Seleção:** seleciona os indivíduos mais aptos para reprodução. Métodos comuns incluem roleta, torneio e seleção por ranking.
- **Crossover (recombinação):** combina pares de indivíduos selecionados para gerar novos indivíduos, simulando reprodução biológica. O cruzamento pode ocorrer em um ou mais pontos da estrutura do cromossomo.
- **Mutação:** introduz variações aleatórias em indivíduos da nova geração, alterando aleatoriamente partes da solução com baixa probabilidade, promovendo diversidade.

### 3.2.4 Estrutura Geral do Algoritmo

1. Inicializar a população com  $N$  indivíduos aleatórios.
2. Avaliar a aptidão de cada indivíduo.
3. Enquanto o critério de parada não for satisfeito:
  - (a) Selecionar indivíduos com base na aptidão.
  - (b) Aplicar operadores de crossover e mutação para gerar nova população.
  - (c) Avaliar a nova população.
  - (d) Atualizar a população (com ou sem elitismo).

### 3.2.5 Critérios de Parada

Alguns critérios comuns para interromper o algoritmo incluem:

- Número fixo de gerações.

- Convergência da população.
- Tempo de execução máximo.
- Ausência de melhoria após várias gerações.

### **3.2.6 Vantagens e Desvantagens**

AGs são particularmente eficazes em problemas com espaços de busca complexos e funções objetivo não deriváveis ou descontínuas. Sua principal vantagem é a capacidade de explorar amplamente o espaço de soluções e escapar de ótimos locais.

Por outro lado, os AGs não garantem a obtenção da solução ótima global, e sua performance pode ser sensível à configuração dos parâmetros, como taxa de crossover, taxa de mutação e tamanho da população.

### **3.2.7 Aplicação de Algoritmos Genéticos à Alocação de Docentes**

Como alternativa à abordagem exata da PLI, propõe-se a implementação de um Algoritmo Genético (AG), uma meta-heurística inspirada nos princípios da evolução natural. O AG é particularmente eficaz na exploração de espaços de solução vastos e complexos, como o da alocação de docentes, sendo capaz de encontrar soluções de alta qualidade em tempo computacional razoável. A seguir, detalha-se a adaptação dos componentes do AG para este problema.

### **3.2.8 Representação do Indivíduo (Cromossomo)**

No contexto do AG, uma solução candidata é denominada *indivíduo* ou *cromossomo*. Para o problema de alocação, cada indivíduo representará uma grade de alocação completa e viável. A representação escolhida é a de um **vetor de alocação direta**, denotado por  $C$ .

- A estrutura do cromossomo  $C$  é um vetor de inteiros cujo comprimento é igual ao número total de disciplinas ofertadas,  $|D|$ .
- O índice do vetor,  $d$ , corresponde a uma disciplina específica do conjunto  $D$  (para  $d \in \{0, 1, \dots, |D| - 1\}$ ).
- O valor contido na posição  $d$  do vetor,  $C[d]$ , é o identificador do professor  $p \in P$  alocado para lecionar a disciplina  $d$ .

Um exemplo de cromossomo para um conjunto hipotético de 4 disciplinas seria:  $C = [p_3, p_7, p_2, p_3]$ , o que significa que o Professor 3 foi alocado para as disciplinas 0 e 3, o Professor 7 para a disciplina 1, e o Professor 2 para a disciplina 2. Esta representação garante, por construção, que cada disciplina seja alocada a exatamente um professor, satisfazendo nativamente a restrição da alocação única.

### 3.2.9 Função de Avaliação (Fitness Function)

A função de avaliação, ou *fitness function*, é o componente que mede a qualidade de um indivíduo (solução), guiando o processo de seleção. Uma pontuação de fitness mais alta indica uma solução melhor. A função proposta,  $f(C)$ , equilibra a maximização da preferência docente com a penalização de violações de restrições de carga horária.

A função de fitness para um cromossomo  $C$  é definida pela Equação 3.2.1:

$$f(C) = \left( \sum_{d=0}^{|D|-1} pref_{C[d],d} \right) - \left( W_{ch} \cdot \sum_{p \in P} \text{Excesso}_p(C) \right) \quad (3.2.1)$$

Os componentes da função são:

- **Escore de Preferência:** O primeiro termo,  $\sum pref_{C[d],d}$ , é a soma direta das preferências de cada alocação presente no cromossomo  $C$ , análogo ao objetivo principal do modelo de PLI.
- **Penalidade por Carga Horária:** O segundo termo penaliza o excesso de carga horária.  $W_{ch}$  é um coeficiente de penalidade escalar, e  $\text{Excesso}_p(C)$  é o total de créditos que excedem a carga horária máxima do professor  $p$  na solução  $C$ , calculado conforme a Equação 3.2.2.

$$\text{Excesso}_p(C) = \max \left( 0, \left( \sum_{d|C[d]=p} ch_d \right) - CH_{max,p} \right) \quad (3.2.2)$$

- **Penalidade por Conflito de Horário:** O terceiro termo, de forma análoga, penaliza conflitos de horário.  $W_{hor}$  é um coeficiente de penalidade específico para esta violação, e  $\text{Cf}_p(C)$  calcula o número de conflitos de horário para o professor  $p$  na solução  $C$ .

### 3.2.10 Operadores Genéticos Aplicados

Os operadores genéticos são os mecanismos que promovem a evolução da população de soluções a cada geração.

#### Seleção.

Existem diversas formas de selecionar os indivíduos que seguirão para a próxima etapa, sendo os 3 principais: Seleção por Roleta, Seleção Por Ranking e Seleção por Torneio. A seleção por roleta possui a implementação mais simples, em que cada indivíduo é avaliado e a sua probabilidade de seleção é proporcional ao seu escore, no entanto essa forma de seleção apresenta um alto risco de convergência prematura, pois soluções razoáveis podem acabar dominando o processo de seleção nas primeiras iterações, acabando com a diversidade da solução. A seleção por Ranking apresenta uma correção da seleção por Roleta, utilizando a classificação de cada solução para determinar a sua probabilidade de seleção, o que reduz a dominância das soluções inicialmente melhores, mas introduz um grande aumento na complexidade computacional do problema, que requer que a cada rodada todos os indivíduos sejam avaliados e ordenados. A terceira alternativa apresenta uma convergência mais cuidadosa, o que a torna mais lenta, mas que resolve o problema de forma satisfatória, e portanto foi a forma de seleção adotada para o problema. O método de seleção por Torneio consiste em:

1. Sortear aleatoriamente  $k$  indivíduos da população (um "torneio" de tamanho  $k$ ,  $k=3$  por exemplo).
2. Avaliar o fitness dos  $k$  indivíduos selecionados.
3. O indivíduo com o maior valor de fitness é selecionado para reprodução, e eles retornam para a população.
4. Repetir o processo até que o número desejado de pais seja selecionado.

Este método é eficiente e tende a evitar a convergência prematura que pode ocorrer em outros métodos, como o da roleta.

#### Crossover (Reprodução).

O operador de crossover combina o material genético de dois pais ( $P_1$  e  $P_2$ ) para gerar um ou mais descendentes ( $O_1, O_2$ ). Será utilizado o **crossover de um ponto**.

1. Um ponto de corte aleatório,  $k$ , é sorteado ao longo do comprimento do cromossomo (onde  $1 \leq k < |D|$ ).
2. O primeiro descendente ( $O_1$ ) é formado pela combinação dos primeiros  $k$  genes de  $P_1$  com os genes restantes de  $P_2$ .
3. O segundo descendente ( $O_2$ ) é formado pela combinação dos primeiros  $k$  genes de  $P_2$  com os genes restantes de  $P_1$ .

*Exemplo:* Se  $P_1 = [p_A, p_B | p_C, p_D]$  e  $P_2 = [p_E, p_F | p_G, p_H]$ , com corte em  $k = 2$ , os descendentes seriam  $O_1 = [p_A, p_B, p_G, p_H]$  e  $O_2 = [p_E, p_F, p_C, p_D]$ .

### Mutação.

A mutação introduz diversidade na população, alterando aleatoriamente um ou mais genes de um indivíduo com uma baixa probabilidade,  $\mu$  (taxa de mutação). Será aplicada a **mutação de gene único**:

1. Para cada gene (posição  $d$ ) em um cromossomo recém-gerado, um número aleatório é sorteado.
2. Se este número for menor que a taxa de mutação  $\mu$ , o gene sofre uma mutação.
3. A mutação consiste em substituir o professor atualmente alocado,  $C[d] = p$ , por um novo professor,  $p'$ , sorteado aleatoriamente do conjunto de todos os professores  $P$ .

## 3.3 Algoritmos de Colônias de Formigas

Os Algoritmos de Colônia de Formigas (Ant Colony Optimization - ACO) são meta-heurísticas inspiradas no comportamento de formigas reais na busca por caminhos entre a colônia e fontes de alimento. Propostos inicialmente por Dorigo e Gambardella (1997), os ACOs têm sido aplicados com sucesso em diversos problemas de otimização combinatória.

### 3.3.1 Princípios Fundamentais

A ideia central do ACO é simular o processo de construção de soluções por agentes artificiais (*formigas*), que exploram o espaço de soluções de um problema, depositando e seguindo trilhas de feromônio que indicam a qualidade das soluções encontradas.

### 3.3.2 Componentes do ACO

- **Representação do Problema:** O problema é modelado como um grafo onde os nós representam componentes da solução e as arestas representam possíveis transições entre esses componentes.
- **Trilha de Feromônio ( $\tau$ ):** Cada aresta do grafo possui um valor de feromônio que influencia a probabilidade de ser escolhida pelas formigas.
- **Heurística ( $\eta$ ):** Informação adicional sobre o problema que pode guiar as formigas na construção de soluções mais promissoras.
- **Função de Transição:** A probabilidade  $p_{ij}^k$  de uma formiga  $k$  se mover do nó  $i$  para o nó  $j$  é dada por:

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha \cdot [\eta_{ij}]^\beta}{\sum_{l \in \mathcal{N}_i^k} [\tau_{il}]^\alpha \cdot [\eta_{il}]^\beta}$$

onde  $\alpha$  e  $\beta$  são parâmetros que controlam a influência do feromônio e da heurística, respectivamente, e  $\mathcal{N}_i^k$  é o conjunto de nós vizinhos de  $i$  ainda não visitados pela formiga  $k$ .

- **Atualização de Feromônio:** Após todas as formigas construírem suas soluções, os valores de feromônio são atualizados para reforçar boas soluções e evaporar feromônios antigos:

$$\tau_{ij} \leftarrow (1 - \rho) \cdot \tau_{ij} + \sum_{k=1}^m \Delta \tau_{ij}^k$$

onde  $\rho$  é a taxa de evaporação e  $\Delta \tau_{ij}^k$  é a quantidade de feromônio depositada pela formiga  $k$  na aresta  $(i, j)$ .

### 3.3.3 Procedimento Geral do ACO

1. Inicializar os parâmetros do algoritmo e os valores iniciais de feromônio.
2. Para cada iteração:
  - (a) Cada formiga constrói uma solução completa baseada nas probabilidades de transição.
  - (b) Avaliar a qualidade das soluções construídas.
  - (c) Atualizar os valores de feromônio com base nas soluções encontradas.

3. Repetir até que um critério de parada seja satisfeito (número máximo de iterações, tempo limite, convergência, etc.).

## 3.4 Aplicação de Algoritmos de Colônia de Formigas (ACO)

No problema de alocação de docentes, as "formigas" serão agentes computacionais que construirão soluções completas. Cada passo está associado à uma disciplina, e o caminho formado representa a seleção de professores que ela fez. E o "feromônio" representará a qualidade aprendida das decisões de alocação.

### 3.4.1 Estrutura do Problema e Representação

Conforme a análise de abordagens, o problema será modelado segundo a estrutura de **Construção Sequencial**. Neste modelo, a construção de uma solução (uma grade de alocação completa) é decomposta em uma sequência de  $|D|$  decisões, uma para cada disciplina a ser alocada. Cada formiga artificial do algoritmo executará este processo construtivo para gerar uma solução candidata a cada iteração.

### 3.4.2 Componentes do Modelo ACO

A seguir, são detalhados os componentes fundamentais do ACO e sua adaptação para o problema em questão.

#### Construção da Solução pelas Formigas.

Cada formiga  $m$  constrói uma solução completa  $S_m$ , que é um conjunto de pares (disciplina, professor). O processo é iterativo:

1. Para cada disciplina  $d_k \in D$ , onde  $k = 1, \dots, |D|$ , a formiga deve escolher um professor  $p$  do conjunto de professores  $P$ .
2. A escolha é probabilística e guiada pela trilha de feromônio e pela informação heurística.
3. **Construção Consciente das Restrições:** Antes de uma formiga escolher um professor para a disciplina  $d_k$ , ela avalia a viabilidade daquela alocação. Um professor  $p$  só é considerado um candidato elegível se a adição da carga horária  $ch_{d_k}$

não exceder sua carga horária máxima ( $CH_{max,p}$ ), considerando as alocações que a mesma formiga já realizou nos passos anteriores. Isso garante que toda solução gerada seja, por construção, viável no que tange à carga horária.

### **Informação Heurística ( $\eta$ ).**

A informação heurística, ou atratividade *a priori*, de se alocar o professor  $p$  para a disciplina  $d$ , é uma informação local que auxilia na decisão da formiga. Para este problema, a preferência do docente pela disciplina é uma excelente medida heurística. Assim, define-se:

$$\eta_{dp} = pref_{pd} + \epsilon \quad (3.4.1)$$

Onde  $pref_{pd}$  é a preferência do professor  $p$  pela disciplina  $d$ . Adiciona-se uma pequena constante  $\epsilon > 0$  (e.g.,  $\epsilon = 0.1$ ) para evitar que uma preferência nula anule a probabilidade de escolha, permitindo uma exploração mínima do espaço de soluções.

### **Trilha de Feromônio ( $\tau$ ).**

A trilha de feromônio  $\tau_{dp}$  está associada a cada possível alocação do par  $(d, p)$ . Este valor representa a "memória" coletiva da colônia, indicando o quanto desejável foi, em iterações passadas, alocar o professor  $p$  para a disciplina  $d$  como parte de uma solução global de alta qualidade. No início do algoritmo, todas as trilhas de feromônio são iniciadas com um valor pequeno e constante.

### **Regra de Transição Probabilística.**

No passo  $k$  da construção, ao decidir qual professor alocar para a disciplina  $d_k$ , uma formiga  $m$  escolhe o professor  $p$  dentre um conjunto de professores viáveis  $\mathcal{N}_k^m$  com base na probabilidade definida pela Equação 3.4.2.

$$p(p|d_k) = \begin{cases} \frac{[\tau_{d_k p}]^\alpha \cdot [\eta_{d_k p}]^\beta}{\sum_{j \in \mathcal{N}_k^m} [\tau_{d_k j}]^\alpha \cdot [\eta_{d_k j}]^\beta} & \text{se } p \in \mathcal{N}_k^m \\ 0 & \text{caso contrário} \end{cases} \quad (3.4.2)$$

Onde:

- $\mathcal{N}_k^m$  é o conjunto de professores elegíveis para a disciplina  $d_k$  na visão da formiga  $m$  (i.e., aqueles cuja alocação não viola as restrições de carga horária nem as de conflito de horário).

- $\alpha$  é um parâmetro que controla a influência da trilha de feromônio (a experiência passada da colônia).
- $\beta$  é um parâmetro que controla a influência da informação heurística (a preferência do professor).

### Atualização de Feromônio.

Após todas as formigas terem construído suas soluções, as trilhas de feromônio são atualizadas. Este processo tem duas etapas: evaporação e depósito.

1. **Evaporação:** Uma fração  $\rho$  (taxa de evaporação,  $0 < \rho \leq 1$ ) do feromônio é evaporada de todas as trilhas. Isso permite o ”esquecimento” de soluções ruins e incentiva a exploração de novas partes do espaço de busca.

$$\tau_{dp} \leftarrow (1 - \rho) \cdot \tau_{dp}, \quad \forall (d, p) \in D \times P \quad (3.4.3)$$

2. **Depósito:** As formigas que construíram as melhores soluções depositam uma quantidade de feromônio nas alocações que utilizaram. O valor depositado,  $\Delta\tau_{dp}$ , é proporcional à qualidade (fitness) da solução.

$$\Delta\tau_{dp}^m = \begin{cases} Q/f(S_m) & \text{se a formiga } m \text{ alocou } (d, p) \\ 0 & \text{caso contrário} \end{cases} \quad (3.4.4)$$

Onde  $f(S_m)$  é a função de fitness da solução  $S_m$  da formiga  $m$  (pode-se utilizar a mesma função da Equação 3.2.1), e  $Q$  é uma constante de desempenho. A atualização completa é dada por:  $\tau_{dp} \leftarrow \tau_{dp} + \sum_m \Delta\tau_{dp}^m$ .

## 3.5 Comparação entre Algoritmos de Otimização

A comparação entre os algoritmos utilizados, Programação Linear Inteira (PLI), Algoritmos Genéticos (AG) e Algoritmos de Colônia de Formigas (ACO), será realizada com base em métricas quantitativas e qualitativas. Este tipo de análise é essencial para compreender o desempenho relativo das abordagens em problemas reais de alocação.

## Critérios de Comparação

- **Qualidade da Solução:** Valor da função objetivo (nível de preferência dos docentes e penalidades);
- **Viabilidade:** Proporção de soluções que satisfazem as restrições duras;
- **Robustez:** Variação dos resultados em múltiplas execuções;
- **Tempo de Execução:** Tempo necessário para obter soluções de boa qualidade;
- **Escalabilidade:** Desempenho com aumento no número de professores e disciplinas.

## Procedimentos Experimentais

A avaliação e comparação do desempenho dos algoritmos serão conduzidas por meio de um conjunto de experimentos computacionais controlados. O objetivo é analisar não apenas a performance em um cenário real, mas também a robustez e a adaptabilidade dos métodos sob diferentes condições.

O ponto de partida será a utilização dos dados reais coletados para o semestre 2025.1, o que estabelecerá uma linha de base para a performance dos algoritmos. Adicionalmente, para uma análise mais profunda, serão gerados cenários sintéticos que simulam diferentes padrões de resposta dos docentes.

Como exercício de validação qualitativa, as soluções ótimas ou sub-ótimas geradas pela ferramenta serão comparadas com as alocações realizadas manualmente em semestres anteriores. Esta análise não visa apenas uma comparação direta de métricas, mas também entender as novas configurações de alocação que a otimização pode revelar, oferecendo um entendimento prático sobre a qualidade e a "razoabilidade" das soluções propostas pelo modelo.

Para cada cenário, os algoritmos serão executados múltiplas vezes para garantir a robustez estatística dos resultados. As métricas de desempenho coletadas serão comparadas utilizando testes não apropriados, permitindo conclusões fundamentadas sobre a superioridade relativa de cada abordagem.

## 4 Resultados

### 4.1 Caracterização do Cenário de Estudo

Para a definição do problema e avaliação dos resultados, iremos considerar os dados da consulta de preferência realizada com os professores para o semestre 2025.1 e a oferta disponível no SIGAA para o mesmo semestre.

O questionário de consulta da oferta é organizado da seguinte forma:

	1	2	3	Em princípio, não tenho interesse.
Análise de Regressão Linear	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Computação em Estatística 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Delineamento e Análise de Experimentos 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Estatística Exploratória	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inferência Estatística	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Introdução à Probabilidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Métodos Estatísticos 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Métodos Estatísticos 2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Técnicas de Amostragem	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
TCC1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Em que cada professor avalia todas as matérias ofertadas pelo departamento de acordo com a sua preferência. Para este trabalho, a divisão dos pontos da preferência foi reorganizada, de forma que: 0 representa o não interesse, e o 3 foi considerado como preferência máxima.

E comparando os dados da oferta com a pesquisa por interesse, apenas 23 dos 28 professores haviam respondido o questionário. Dos 5 que não responderam, 1 ocupava a posição de coordenador, então foi adicionado ao arquivo de preferências com preferência

3 pela disciplina TCC1 e 0 nas demais, e os demais professores foram considerados como professores substitutos, contratados para preencher as disciplinas de serviço do departamento, portando ficando com preferência 3 pelas turmas de serviço e 0 pelas demais.

#### **4.1.1 Análise da Procura por Disciplinas**

Para analisar a preferência dos professores pelas disciplinas, foi feita uma contagem de quantas vezes cada disciplina obteve cada pontuação de preferência pelos professores, e o seguinte resultado foi obtido:

Tabela 1: Distribuição das preferências por disciplina

Disciplina	Nº de Turmas	Pref 0	Pref 1	Pref 2	Pref 3
Analise de dados categorizados	1	25	2	0	1
Analise de regressao linear	1	21	2	2	3
Analise de series temporais	1	24	2	0	2
Analise de sobrevivencia	1	25	1	0	2
Bioestatistica	2	1	13	5	9
Computacao em estatistica 1	2	21	3	2	2
Computacao em estatistica 2: python	1	25	0	1	2
Computacao em estatistica 2: r	1	24	2	2	0
Delineamento e analise de experimentos	1	25	2	0	1
Demografia	1	27	1	0	0
Estatistica aplicada	10	1	12	4	11
Estatistica computacional	1	23	2	1	2
Estatistica exploratoria	1	21	5	0	2
Inferencia bayesiana	1	26	1	0	1
Inferencia estatistica	1	21	2	1	4
Introducao a probabilidade	2	13	5	2	8
Laboratorio de estatistica 1	1	26	0	2	0
Metodos estatisticos 1	1	19	3	4	2
Metodos estatisticos 2	1	20	4	2	2
Modelos lineares generalizados	1	25	1	0	2
Praticas de extensao em estatistica 2	1	27	0	1	0
Probabilidade e estatistica	11	1	8	2	17
Processos estocasticos	1	20	1	3	4
Tecnicas de amostragem	1	26	1	0	1
Trabalho de conclusao de curso 1	1	27	0	0	1

Podemos observar pela tabela 1 que, de forma geral, as disciplinas possuem poucos professores com interesse. Com exceção das disciplinas de serviço do departamento (Probabilidade e Estatística, Bioestatística e Estatística Aplicada) que praticamente não possuem nenhuma indicação de não interesse em lecionar. Além disso, outra tendência que pode ser observada é que as disciplinas iniciais do curso também possuem um interesse um pouco mais elevado do que as demais, e conforme a especificidade da disciplina aumenta, o número de professores com interesse diminui. Mostrando que no departamento há uma alta especialização nas áreas de cada disciplina e existem poucas opções com interesse em cada disciplina. Além disso, se destacam algumas disciplinas obrigatórias do currículo com pouco ou nenhum interesse pelos professores, como: Demografia, Análise de Dados Categorizados, Inferência Bayesiana e Técnicas de Amostragem.

Quase todas as disciplinas possuem um número de professores com alto interesse maior ou igual ao número de turmas ofertadas, criando a possibilidade da construção de alocações que cada turma seja lecionada por um professor que possui alto interesse, gerando pontuações mais altas na avaliação. O valor máximo teórico de pontuação seria de 141 ( $47 \times 3$ ) que seria obtido se todas as disciplinas fossem alocadas para professores com preferência nível 3. Mas como há 4 disciplinas que não possuem preferência 3, e uma delas possui apenas preferência 1, o valor máximo que pode ser obtido é de 136 pontos. No entanto, esse cenário pode não ocorrer na prática por conflitos de horários entre disciplinas com pouco interesse, selecionadas pelos mesmos docentes.

Podemos também analisar quantas vezes cada docente marcou alguma disciplina com uma preferência específica:

Tabela 2: Distribuição das preferências por docente

<b>Docente</b>	<b>Pref 0</b>	<b>Pref 1</b>	<b>Pref 2</b>	<b>Pref 3</b>
docente_01	23	10	11	3
docente_02	20	2	15	10
docente_03	24	12	0	11
docente_04	19	2	14	12
docente_05	21	12	1	13
docente_06	22	11	1	13
docente_07	11	5	4	27
docente_08	14	27	5	1
docente_09	21	12	0	14
docente_10	24	0	0	23
docente_11	18	0	14	15
docente_12	19	15	10	3
docente_13	24	0	0	23
docente_14	23	23	0	1
docente_15	18	12	0	17
docente_16	24	0	0	23
docente_17	24	0	0	23
docente_18	13	20	2	12
docente_19	15	18	10	4
docente_20	46	0	0	1
docente_21	21	12	0	14
docente_22	19	14	3	11
docente_23	17	11	3	16
docente_24	24	0	0	23
docente_25	5	25	4	13
docente_26	16	14	2	15
docente_27	21	12	0	14
docente_28	22	13	0	12

As inserções feitas na tabela de preferências foram os docentes: 20 (coordenador), 13, 16, 17 e 24 (os professores substitutos).

A Tabela 2 conduz a análise para o perfil dos docentes. Observa-se uma variabilidade significativa no comportamento. Por exemplo, o Docente 12 manifestou desinteresse por 41 das 46 disciplinas, não indicando preferência máxima por nenhuma, o que o caracteriza como um perfil de alocação potencialmente mais restrito. Em contrapartida, os

Docentes 10 e 13 demonstram grande engajamento ou um leque de interesses mais amplo, marcando 27 e 32 disciplinas com o maior nível de preferência, respectivamente. Essa heterogeneidade no perfil docente é um fator chave de complexidade que a ferramenta de otimização deverá gerenciar.

#### **4.1.2 Conclusão**

No cenário considerado temos 28 docentes e 47 disciplinas (considerando turmas de uma mesma disciplina como disciplinas diferentes), e trabalhamos com a restrição de que cada docente deve lecionar no máximo 2 disciplinas. Portanto temos uma capacidade total de 56 alocações para uma demanda de 47 turmas, resultando em uma folga teórica de apenas 9 posições.

### **4.2 Análise e Calibração do Algoritmo Genético**

O desempenho de meta-heurísticas é extremamente dependente da escolha adequada de seus hiperparâmetros. A configuração dos hiperparâmetros define o comportamento do algoritmo, controlando o equilíbrio entre a exploração (capacidade de investigar novas regiões do espaço de busca) e a intensificação (capacidade de refinar e aprimorar as soluções já encontradas). Uma calibração inadequada pode levar à uma convergência para ótimos locais ou a um esforço computacional excessivo sem ganho de qualidade.

#### **4.2.1 Estudo de Convergência**

Para iniciar o estudo da calibração dos hiperparâmetros, foi feito um estudo preliminar sobre o tempo de convergência do algoritmo. Com o objetivo de fixar o número de gerações em cada simulação de forma a evitar que alguma execução seja interrompida antes da convergência ou que seja gasto mais tempo do que o necessário na simulação.

O objetivo é determinar o menor número possível de gerações que permita que o algoritmo chegue à sua melhor resposta para qualquer configuração proposta dos hiperparâmetros, para isso foram realizadas 10 simulações com configurações aleatórias dos hiperparâmetros por 500 gerações. Segue abaixo o gráfico com os resultados obtidos:

Figura 1: Gráfico de linhas do estudo de convergência do algoritmo genético

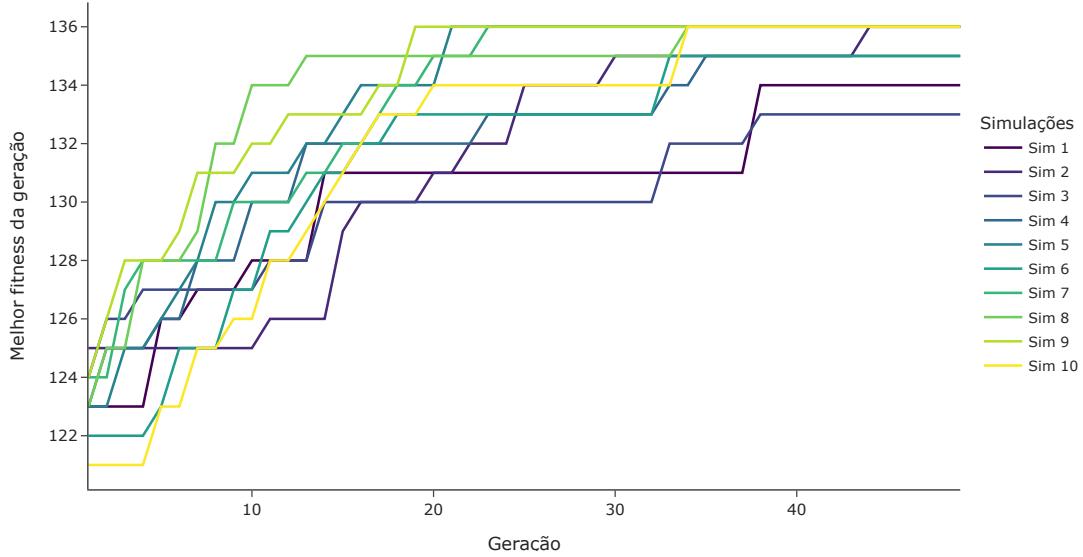


Tabela 3: Resultados do estudo de convergência

Simulação	Fitness Final	Geração Melhor	Tempo (s)	População	Crossover	Mutação	Torneio	Elite
1	136	106	24.77	100	0.92	0.16	2	3
2	136	44	8.32	200	0.84	0.02	2	3
3	136	97	74.36	200	0.94	0.12	2	2
4	136	65	20.38	80	0.77	0.06	2	1
5	136	21	19.87	150	0.69	0.03	4	4
6	136	55	3.54	50	0.69	0.01	4	1
7	136	23	16.96	180	0.89	0.02	3	1
8	136	34	37.03	100	0.94	0.1	5	4
9	136	19	77.4	150	0.84	0.18	5	3
10	136	34	13.38	50	0.71	0.07	5	1

A partir da análise do gráfico 1 e da tabela 3, podemos observar que todas as simulações possuem um valor inicial alto para o score da melhor solução, isso se deve à forma com que o algoritmo foi implementado em que a cada alocação que é feita pelo algoritmo, seja na população inicial ou em alguma mutação, o professor selecionado para a matéria é selecionado de forma a não gerar nenhum conflito de horário e com probabilidade de seleção proporcional à preferência pela disciplina. O que faz com que o algoritmo não precise "aprender" as restrições a partir das penalizações e já crie sempre soluções possíveis, e de forma já direcionada para o aumento do score da solução. Obtendo dessa forma, boas soluções iniciais e uma rápida convergência, à um custo de um aumento no tempo computacional da formação de novas gerações, que se compensa na diminuição drástica do número de gerações necessárias.

Além disso, é possível observar que todas as simulações convergiram para o mesmo valor, indicando que a implementação da criação das soluções já garante a convergência e que a calibração dos hiperparâmetros pode ter um foco maior em minimizar o tempo de execução do algoritmo. E, podemos observar que com 50 gerações, 6 das 10 simulações já convergiram para o valor máximo e as demais estão a menos de 4 pontos de distância no escore, tornando-o um valor suficientemente grande para realizar o estudo de calibração dos hiperparâmetros.

#### 4.2.2 Análise de sensibilidade dos hiperparâmetros

Para garantir uma varredura eficiente e representativa do espaço de configurações, foi adotado o método de amostragem por hipercubo latino (LHS - *Latin Hypercube Sampling*). O LHS é uma técnica de amostragem estratificada que divide o intervalo de cada parâmetro em N subintervalos de probabilidade uniforme, garantindo que exatamente uma amostra seja retirada de cada intervalo.

Segundo McKay, Beckman e Conover (2000), que introduziram o método, o LHS oferece uma estimativa da média e da variância da saída do modelo com significativamente menos variabilidade do que a amostragem aleatória simples, sendo particularmente superior em simulações computacionais onde o custo de avaliação é alto. Além disso, Santner et al. (2003) destacam que o LHS é uma das classes de delineamentos mais eficazes para experimentos computacionais (Space-Filling Designs), pois projeta amostras que preenchem o espaço de forma uniforme, maximizando a informação obtida com um número limitado de simulações.

Neste estudo, após a fixação do horizonte de convergência em 100 gerações, aplicou-se o LHS para gerar 100 combinações de parâmetros distintas para o Algoritmo Genético. As faixas de valores exploradas para cada parâmetro foram definidas com base em testes preliminares, conforme detalhado na Tabela ??.

A análise dos resultados do LHS confirma a suposição anterior de que a robusteza na implementação do algoritmo faz com que a configuração dos hiperparâmetros não influencie na convergência dos algoritmos genéticos, pudemos observar que quase 80% das configurações de parâmetros construídas atingiram o Escore máximo, e as demais configurações chegaram em valores muito próximos com apenas 50 iterações. Isso evidencia que a heurística construtiva adotada na inicialização é eficaz em posicionar a busca em regiões promissoras do espaço de soluções.

Portanto, com a complexidade do problema reduzida pela dimensão dos dados uti-

Figura 2: Distribuição do Fitness Final

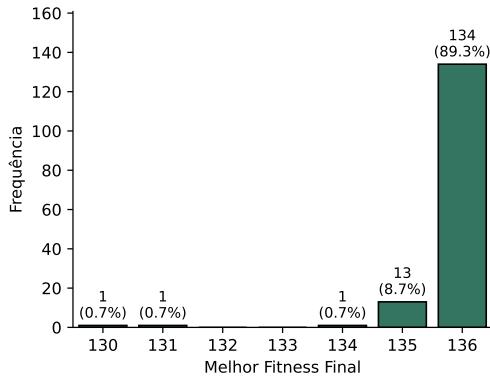


Tabela 4: Frequência de Scores das Simulações

Fitness Final	Frequência Absoluta	Frequência Relativa
128	1	0,01
130	1	0,01
131	1	0,01
132	1	0,01
133	2	0,02
134	6	0,06
135	10	0,1
136	78	0,78

lizados pelo departamento, a qualidade da solução é dominada pela eficácia da inicialização heurística. E como os resultados obtidos são ótimos independente dos hiperparâmetros, eles devem ser otimizados para priorizar a eficiência computacional.

Para maximizar a eficiência computacional, observaremos como cada um dos hiperparâmetros influencia no tempo de geração da solução, destacando as soluções que não obtiveram a pontuação máxima. Os gráficos que mostram o comportamento dos resultados foram agrupados em pares de hiperparâmetros e seguem apresentados abaixo:

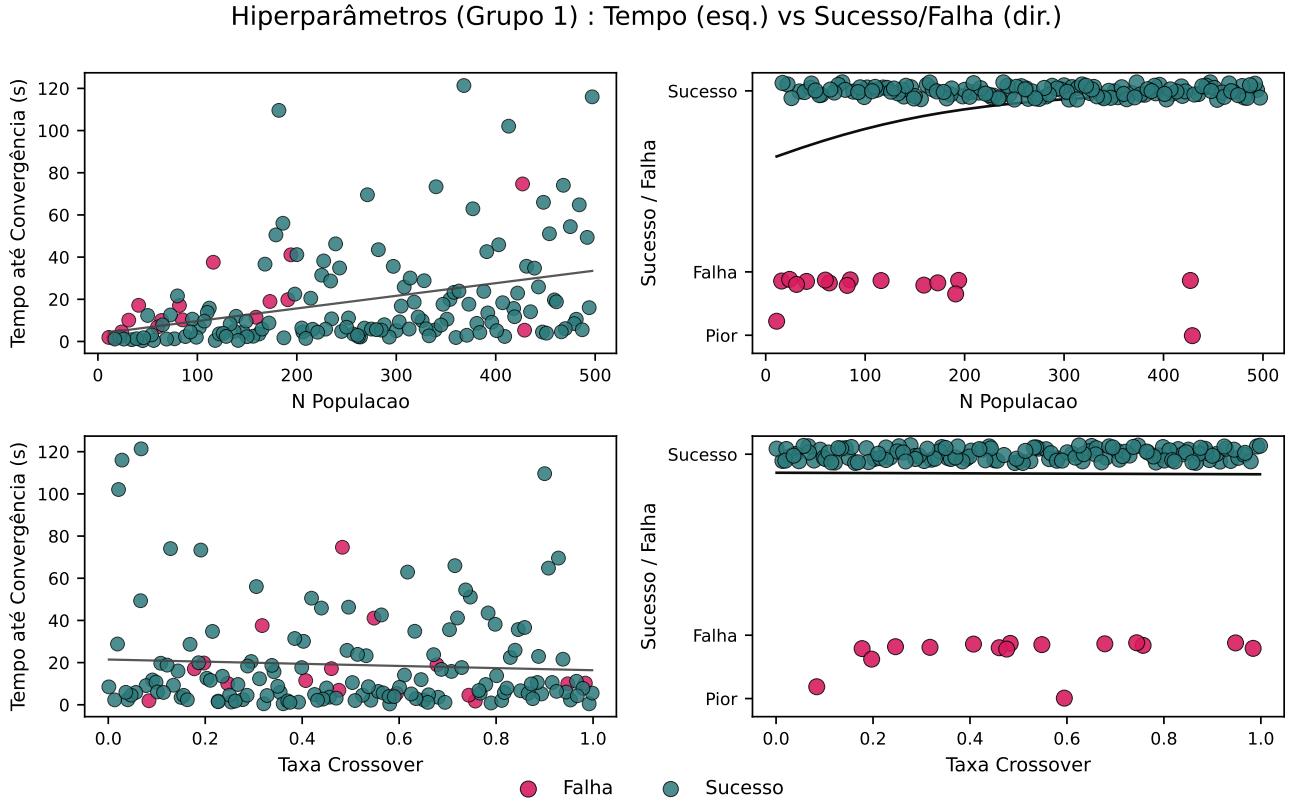
Conforme esperado, o tempo até a convergência aumenta com a quantidade de dados utilizados e quantidade de operações realizadas à cada iteração, então em resumo, devemos buscar pelas soluções que diminuam a complexidade do algoritmo e mantenham um bom resultado de escore.

Podemos observar que o tempo de convergência começa a apresentar valores mais elevados e uma variabilidade muito mais inconstante a partir das populações com mais de 200 indivíduos, e que 9 das 16 falhas vieram das populações com menos de 100 indivíduos. Dessa forma, podemos selecionar 250 indivíduos, que atenderão os dois critérios propostos.

Para a taxa de Crossover podemos observar que há uma leve diminuição no tempo de convergência conforme ela aumenta, pois apesar de aumentar o número de operações realizadas, são operações simples que ajudam a perpetuar as melhores características dos indivíduos, e não há indícios de relação entre a taxa de crossover e número de falhas obtidas, podendo portanto arbitrar como valor final a taxa de 0.8.

A análise da taxa de mutação mostra de forma clara que o aumento na proporção

Figura 3: Gráficos de dispersão dos hiperparâmetros pelo tempo e pelas falhas

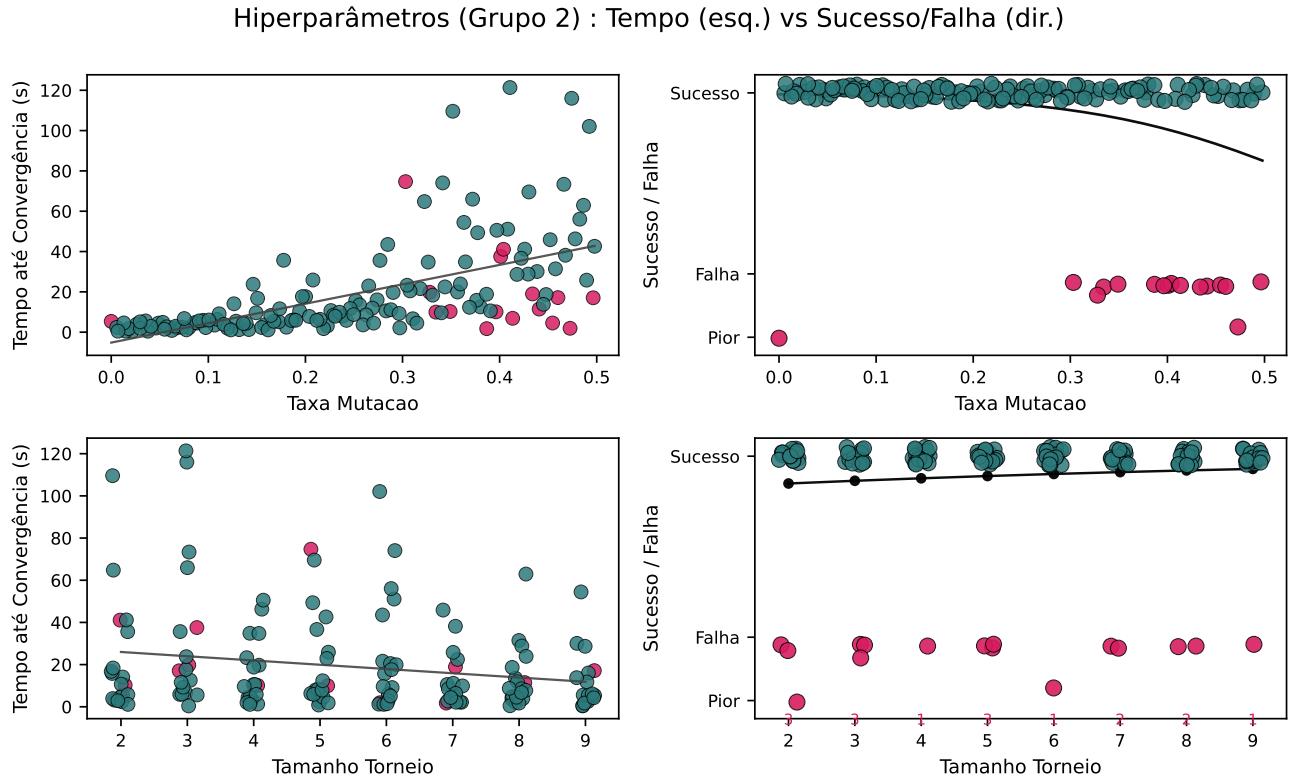


de mutações realizadas faz com que o tempo até a convergência e a sua dispersão começem a aumentar drasticamente, assim como o número de falhas. Permitindo que possamos arbitrar 0.1 como valor para taxa de mutação em razão do seu baixo tempo computacional e ausência de falhas na região.

O tamanho do torneio apresenta pouca influência no tempo de convergência, mas sugere que aumentar o número de indivíduos selecionados à cada rodada de eliminação acelera a convergência, enquanto não indica nenhuma relação aparente com o número de falhas encontradas. Permitindo que possamos arbitrar a nossa escolha no maior tamanho amostrado de 9 indivíduos por seleção.

O gráfico de falhas do número de indivíduos selecionado pelo elitismo pode ser um pouco ilusório, porque como o elitismo é definido à partir de uma % de indivíduos da população, os resultados de altas falhas das simulações com populações baixas se confundem um pouco com os resultados das simulações que foram feitas com um baixo percentual de elitismo, inflando o resultado negativo do segundo grupo. Mas ainda assim, é possível perceber que uma parte considerável das falhas está na região de valores de taxa de elitismo inferiores a 0.1. E podemos ver pelo gráfico de tempo até a convergência,

Figura 4: Gráficos de dispersão dos hiperparâmetros pelo tempo e pelas falhas



que aumentar o percentual de elitismo reduz o tempo de convergência, e novamente temos um resultado no primeiro gráfico que sugere o contrário, mas que é causado pelo mesmo efeito das associações com populações com grandes números de indivíduos, permitindo arbitrar o resultado final em 30% do tamanho da população.

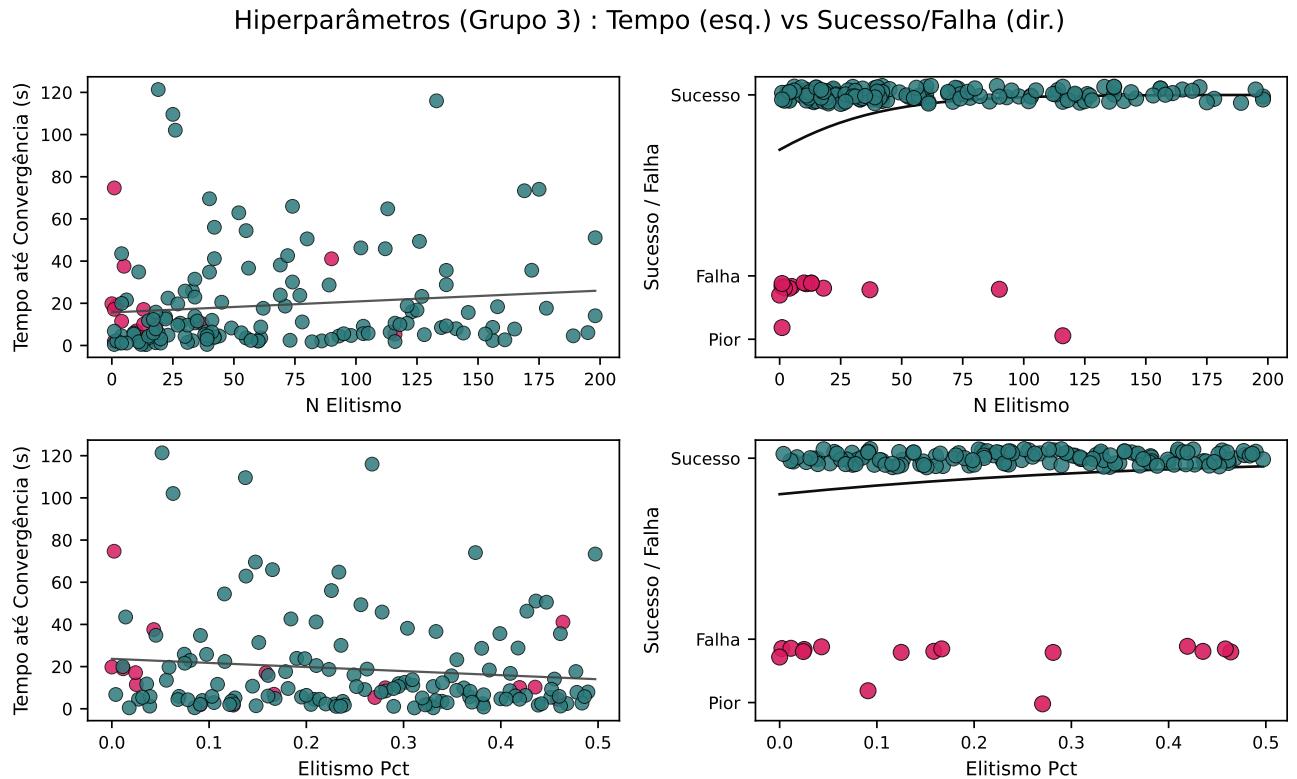
#### 4.2.3 Conclusão

Através da análise da qualidade dos resultados e eficiência computacional, os seguintes valores foram fixados para os hiperparâmetros:

Tabela 5: Hiperparâmetros fixados para o Algoritmo Genético

Parâmetros	Valores
Nº de Gerações	100
Tamanho da População	200
Taxa Crossover	0.8
Taxa Mutação	0.1
Tamanho do Torneio	9
Tamanho da Elite	60

Figura 5: Gráficos de dispersão dos hiperparâmetros pelo tempo e pelas falhas



Foram realizadas 100 repetições com a configuração da tabela 5, que confirmaram a robustez do método proposto. Em 100% das execuções, o algoritmo convergiu para a solução ótima conhecida, demonstrando que a heurística de inicialização e adaptação de soluções garante a eficácia da solução independentemente da aleatoriedade do processo.

Para analisar o ganho de eficiência computacional, os seguintes gráficos foram elaborados:

Figura 6: Histograma do tempo até convergência com os parâmetros otimizados

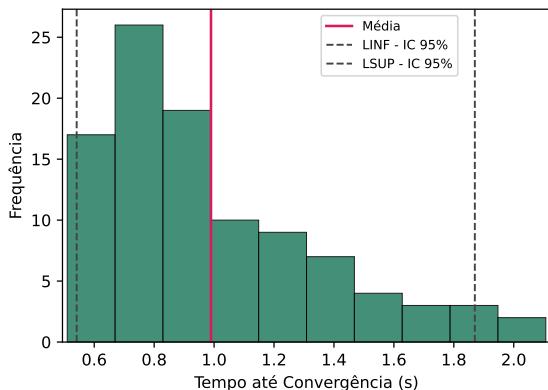
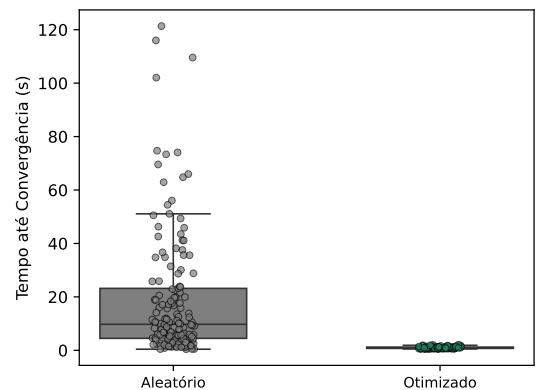


Figura 7: Boxplot da dispersão dos tempos de convergência



Podemos observar que além de garantir resultados mais seguros e estáveis, a otimização também obteve como resultado um grande ganho na eficiência computacional e estabilidade de execução do algoritmo. Diminuindo a média de 18.92s para 0.99s, a mediana de 9.75 para 0.88, e o desvio padrão de 23.55s para 0.37. Deixando o algoritmo com performances confiáveis e constantes.

## 4.3 Análise e Calibração do Algoritmo de Colônia de Formigas

### 4.3.1 Estudo de Convergência

Figura 8: Gráficos de linhas do estudo de convergência do ACO

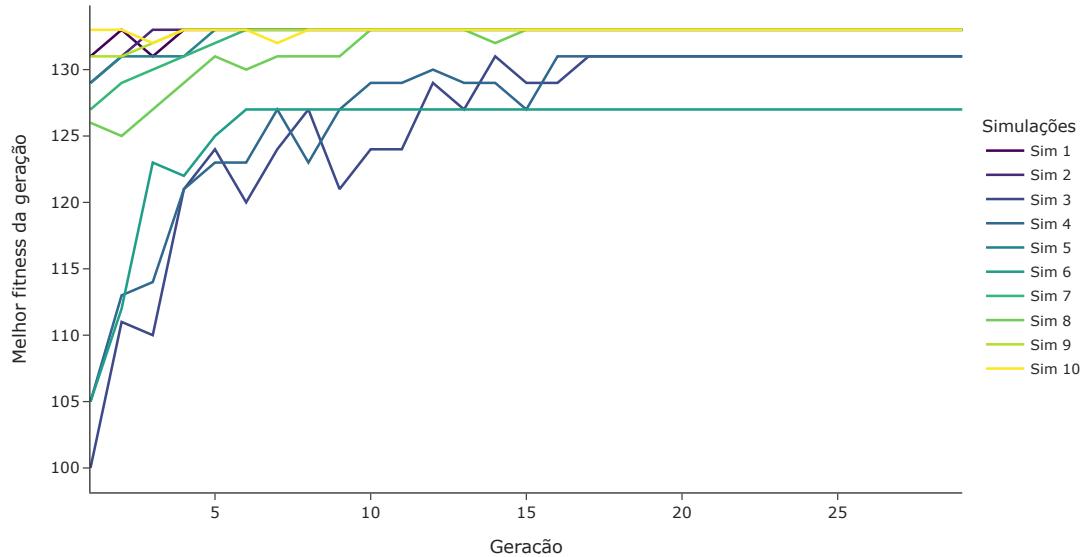


Tabela 6: Resumo das 10 execuções do algoritmo de colônia de formiga.

Simulação	Geração Melhor	Tempo (s)	Nº de formigas	Alfa	Beta	Taxa de Evaporação
1	2	54.27	50	0.76	4.71	0.16
2	3	16.97	15	1.63	4.38	0.45
3	14	11.46	10	1.43	1.16	0.09
4	16	17.13	15	1.31	1.13	0.14
5	5	56.42	50	1.58	3.1	0.25
6	6	22.15	20	1.73	1.03	0.41
7	6	58.02	50	1.19	2.39	0.15
8	10	23.01	20	0.74	2.9	0.21
9	4	23.23	20	1.45	5.04	0.38
10	1	47.77	40	0.77	5.61	0.09

**4.3.2 Análise de sensibilidade dos hiperparâmetros**

## 5 Conclusão

Aqui você deve escrever a conclusão ou considerações finais.

## Referências

- BURKE, E. K.; PETROVIC, S. Recent research directions in automated timetabling. *European journal of operational research*, Elsevier, v. 140, n. 2, p. 266–280, 2002.
- DORIGO, M.; GAMBARDELLA, L. Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, v. 1, n. 1, p. 53–66, 1997.
- HOLLAND, J. H. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. [S.l.]: MIT press, 1992.
- MCKAY, M. D.; BECKMAN, R. J.; CONOVER, W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, Taylor & Francis, v. 42, n. 1, p. 55–61, 2000.
- SANTNER, T. J. et al. *The design and analysis of computer experiments*. [S.l.]: Springer, 2003. v. 1.

## Apêndice

### A Tabela Hiperparâmetros Ajuste - AGs

Tabela 7: Resumo organizado das execuções do AG (Apêndice).

Simulacao	Populacao	Torneio	Elite	Taxa_Crossover	Taxa_Mutacao	Fitness_Final	Tempo_Seg
1	412	5	0	0.499	0.085	136	6.486
2	402	3	0	0.839	0.238	132	14.868
3	311	5	4	0.778	0.292	136	13.405
4	199	5	5	0.883	0.210	136	6.345
5	118	2	0	0.962	0.088	130	1.842
6	219	6	4	0.415	0.111	136	4.134
7	462	9	4	0.598	0.099	136	8.149
8	258	4	1	0.903	0.018	136	1.504
9	453	3	4	0.803	0.283	136	19.541
10	476	7	5	0.992	0.150	136	11.658
11	396	2	1	0.742	0.091	134	6.505
12	270	6	0	0.913	0.286	135	13.024
13	300	9	2	0.865	0.008	136	1.354
14	433	7	3	0.409	0.242	136	25.612
15	87	3	2	0.830	0.154	136	5.419
16	23	5	5	0.554	0.143	136	1.121
17	26	4	4	0.920	0.277	136	2.463
18	499	7	2	0.480	0.003	136	4.773
19	195	5	1	0.816	0.045	136	4.867
20	423	9	5	0.530	0.078	136	15.735
21	206	9	0	0.606	0.060	136	6.217
22	246	8	5	0.475	0.129	136	13.462
23	95	5	5	0.489	0.257	136	9.251
24	108	6	0	0.617	0.056	136	3.088
25	399	2	3	0.810	0.141	135	23.911
26	340	5	2	0.734	0.225	136	30.086
27	131	5	4	0.634	0.274	136	13.477
28	323	9	0	0.964	0.122	136	17.191
29	389	7	1	0.847	0.034	136	8.182
30	238	4	3	0.674	0.161	136	15.705
31	481	6	1	0.929	0.228	136	43.943
32	241	3	0	0.614	0.062	136	7.642
33	136	6	1	0.719	0.290	135	15.626
34	336	7	2	0.449	0.244	136	32.235
35	263	8	2	0.589	0.027	136	4.910
36	333	5	1	0.625	0.031	136	6.685
37	123	7	1	0.533	0.298	135	14.352
38	486	6	2	0.746	0.270	136	51.379
39	377	4	2	0.567	0.135	136	21.628
40	386	5	1	0.630	0.136	136	22.205
41	113	2	4	0.568	0.215	135	9.513
42	82	8	1	0.951	0.235	136	7.686
43	325	7	4	0.649	0.132	136	18.245
44	427	9	3	0.447	0.117	136	21.669
45	371	4	0	0.403	0.083	136	14.240
46	142	4	1	0.754	0.253	135	13.995
47	53	4	4	0.979	0.004	136	0.523
48	437	9	2	0.517	0.049	136	11.572
49	290	6	1	0.713	0.021	136	4.728

## B nome apêndice 2

texto...

## **Anexo**

### **A nome anexo 1**

texto...