08/06 (2025), **xx**:xx 1–8

Trabalho Computacional 1

ELE077 Otimização Não Linear - Irrestrita

Tasso Augusto Tomaz Pimenta[1]

[1]Departamento de Engenharia de Produção, UFMG. E-mail: TassoossaT842@gmail.com.

Abstract
Neste trabalho apresento os resultados e discuções sobre a resolução de problemas usando algoritmos de otimização não linear irrestrita para quatro problemas

Contents

1. Problema 1

1.1. Definição do problema

Formulação Matemática do Problema de Otimização PID

O problema consiste em encontrar os parâmetros ótimos de um controlador PID para minimizar o erro quadrático médio da temperatura de um sistema ao longo do tempo.

Dinâmica do Sistema e Controle

O sinal de controle $u(t)$ é definido pela equação do controlador PID:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{d}{dt} e(t) \tag{1.1}$$

onde $e(t)$ é o erro entre a temperatura de referência T_{ref} e a temperatura atual $T(t)$:

$$e(t) = T_{\text{ref}} - T(t) \tag{1.2}$$

Os ganhos do controlador, que são nossas variáveis de decisão, são o ganho proporcional (K_p), o integral (K_i) e o derivativo (K_d).

A temperatura do sistema $T(t)$ evolui de acordo com a seguinte Equação Diferencial Ordinária (EDO) de primeira ordem:

$$\frac{dT}{dt} = -\frac{T}{\tau} + \frac{K}{\tau} u(t) \tag{1.3}$$

onde τ é a constante de tempo do sistema e K é uma constante de proporcionalidade.

Função-Objetivo

O objetivo é minimizar o critério de desempenho do erro quadrático integrado (Integral Square Error - ISE) ao longo de um período de simulação T_{sim}. A função de custo J a ser minimizada é definida como:

$$J(K_p, K_i, K_d) = \int_0^{T_{\text{sim}}} [e(t)]^2 dt \tag{1.4}$$

Problema de Otimização

Portanto, o problema de otimização pode ser formulado de maneira completa como:

$$\min_{K_p, K_i, K_d} \int_0^{T_{\text{sim}}} \left(T_{\text{ref}} - T(t, K_p, K_i, K_d)\right)^2 dt \tag{1.5}$$

sujeito à dinâmica do sistema descrita pela equação (1.3) e às condições iniciais do problema.

1.2. Resolução e Escolha do método

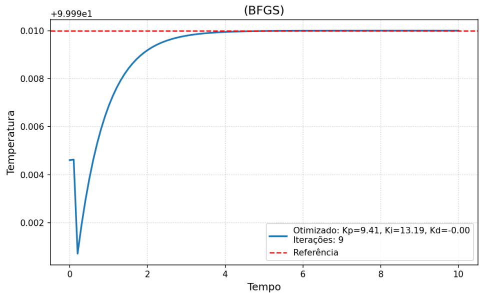
Como é dito "O problema consiste em encontrar os parâmetros ótimos de um controlador PID para minimizar o erro quadrático médio da temperatura de um sistema ao longo do tempo."

Então as aproximações quadraticas seriam as melhores escolhas, temos então como opção:

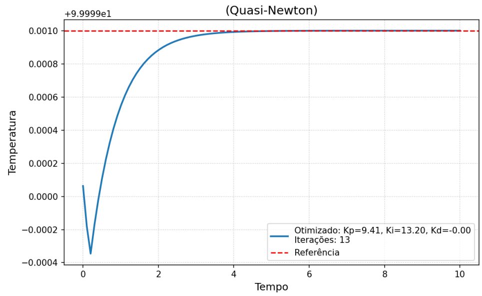
- Newton

- Quase-Newton
- DFP (Não consegui aplicar)
- BFGS

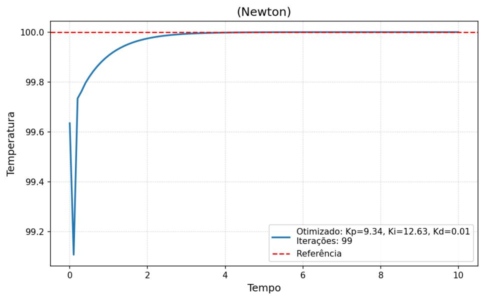
Todos vão ter resultados parecidos, a diferença vai está basicamente no custo, e na quantidade de iterações necessarias para convergir, por isso rodei para todos para determinar o melhor:



(a) BFGS



(b) Quase-Newton



(c) Newton

Figure 1: Comparação da resposta do sistema PID com diferentes métodos de otimização.

Como observado os metodos quase newton tiveraam resultados superiores em relação ao newton, isso se deve ao fato da função objetivo não é exatamente quadratica pura, erro quadratico é diferente de função quadratica, a newton foi desenvolvida para lidar com funções no formato f(x)=ax²+bx+c, então é normal que ela não perfome tão bem em outros formatos, então as funções quase newtonianas são mais adequadas nesse problema. Tabela comparativa com outros metodos:

Table 1: Resultados da otimização dos parâmetros PID para diferentes algoritmos.

Método	K_p	K_i	K_d	**Iterações**
Gradiente	9.279	14.463	0.004	574
Newton	9.344	12.626	0.007	99
Quasi-Newton	9.409	13.198	-0.000	13
BFGS	9.410	13.185	-0.000	9
Gradiente Conjugado	9.407	13.202	0.000	93
Hooke-Jeeves	9.403	13.218	0.000	266
Nelder-Mead Simplex	9.408	13.200	0.000	98

A tablema 1 serve para exemplificar que os melhores metodos são os quase newton.

2. Problema 2

2.1. Definição do problema

Regressão Logística

O objetivo da regressão logística é encontrar um vetor de pesos $\mathbf{w}$ que melhor classifica um conjunto de dados.

Função de Hipótese (Sigmoide)

A probabilidade de uma amostra $\mathbf{x}_i$ pertencer à classe $y = 1$ é dada pela função de hipótese h, também conhecida como função sigmoide:

$$h(\mathbf{x}_i; \mathbf{w}) = \frac{1}{1 + e^{-\mathbf{w}^T \mathbf{x}_i}} \tag{2.1}$$

Função de Custo (Entropia Cruzada com Regularização L2)

Para encontrar o vetor de pesos $\mathbf{w}$ ideal, minimizamos a função de custo $J(\mathbf{w})$. Ela é composta pela média da entropia cruzada sobre todas as amostras e um termo de regularização L2 para evitar sobreajuste (overfitting):

$$J(\mathbf{w}) = -\frac{1}{m} \sum_{i=1}^{m} \left[y_i \log\left(h(\mathbf{x}_i; \mathbf{w})\right) + (1 - y_i) \log\left(1 - h(\mathbf{x}_i; \mathbf{w})\right) \right] + \frac{\lambda}{2} \|\mathbf{w}\|_2^2 \tag{2.2}$$

Onde:

- m é o número total de amostras.
- λ é o hiperparâmetro de regularização.
- $\|\mathbf{w}\|_2^2$ é a norma L2 ao quadrado do vetor de pesos, que penaliza valores de pesos muito grandes.

Problema de Otimização

O problema de otimização é, portanto, encontrar o vetor de pesos $\mathbf{w}$ que minimiza a função de custo $J(\mathbf{w})$:

$$\min_{\mathbf{w} \in \mathbb{R}^n} J(\mathbf{w}) \tag{2.3}$$

2.2. Resolução e Escolha do Método

O Problema 2 (Regressão Logística) apresenta um desafio de otimização em alta dimensão ($n = 500$) e com um grande número de amostras ($m = 10000$). Embora o gradiente da função de custo seja analiticamente disponível, o cálculo sobre o conjunto de dados completo torna cada avaliação de função e gradiente computacionalmente cara. Além disso, o Método de Newton é inviável devido ao custo proibitivo de calcular e inverter a matriz Hessiana de 500×500 a cada passo.

Diante disso, os candidatos mais promissores são os métodos que usam informações de primeira ordem de forma inteligente: os métodos Quasi-Newton (como o BFGS) e o método do Gradiente Conjugado. A escolha entre eles envolve um importante trade-off:

- **BFGS:** Tende a convergir em menos iterações, pois constrói uma rica aproximação da curvatura da função. No entanto, possui um custo maior por iteração devido às operações matriciais ($O(n^2)$).
- **Gradiente Conjugado:** Possui um custo por iteração muito menor, envolvendo apenas operações com vetores ($O(n)$). Em contrapartida, geralmente requer mais iterações para atingir a mesma precisão que o BFGS.

Para determinar o método mais adequado na prática, ambos foram implementados e os resultados são apresentados na Tabela 2.

Table 2: Resultados da otimização para o Problema 2.

Otimizador	**Custo Final**	**Iterações**	**Tempo (s)**
BFGS	0.0948	19	2228.98
Gradiente Conjugado	0.0937	1	-
Gradiente Conjugado	$\approx 10^{21}$	2 (divergiu)	336.63

Análise dos Resultados e Conclusão

Os resultados experimentais mostram que a implementação do BFGS foi robusta e bem-sucedida, convergindo para uma solução válida com um custo final de 0.0948 em 19 iterações. O tempo de execução elevado é consistente com o custo computacional esperado para este método em um problema de alta dimensão.

Por outro lado, a implementação do Gradiente Conjugado demonstrou instabilidade numérica e divergiu após a primeira iteração, resultando em um valor de custo astronomicamente alto. Porém essa falha decorreu de um erro de código e não de algoritmo

O método do Gradiente Conjugado ainda é o mais recomendado para o caso. Ele consegue chegar em resultads mais rapidos do que o BFGS, considerando que o erro está no algoritmo e a primeira interação me retornou um valor menor que o BFGS.

3. Problema 3

3.1. Definição do problema

O objetivo é encontrar a combinação de parâmetros de uma campanha – desconto (d), tempo (t) e orçamento de marketing (m) – que maximiza o lucro $L(d, t, m)$.

Funções de Receita e Custo

O lucro é definido como a receita R menos o custo C:

$$L(d, t, m) = R(d, t, m) - C(d, t, m) \tag{3.1}$$

A receita é modelada por:

$$R(d, t, m) = V_B \cdot (1 + f_1(d) + f_2(t)) \cdot \log(1 + m) \tag{3.2}$$

onde $f_1(d)$ e $f_2(t)$ são os incrementos nas vendas:

$$f_1(d) = -0.005d^2 + 0.2d \tag{3.3}$$

$$f_2(t) = 0.05t \tag{3.4}$$

O custo total é composto por um custo fixo C_B, o custo de marketing m, e uma função de penalidade $P(d, t)$:

$$C(d, t, m) = C_B + m + P(d, t) \tag{3.5}$$

Função de Penalidade (Não-Contínua)

A função de penalidade $P(d, t)$ introduz descontinuidades na função de custo, sendo o ponto chave para a escolha do método de otimização. Ela é definida da seguinte forma:

$$P(d, t) = \begin{cases} 0 & \text{se } d \leq 30 \text{ e } t \leq 15 \\ 5000 & \text{se } d > 30 \text{ e } t \leq 15 \\ 2000 & \text{se } d \leq 30 \text{ e } t > 15 \\ 7000 & \text{se } d > 30 \text{ e } t > 15 \\ 100.000 & \text{se } d \leq 0 \text{ ou } 50 \leq d \\ 100.000 & \text{se } t \leq 1 \text{ ou } 30 \leq t \\ 100.000 & \text{se } m \leq 1000 \text{ ou } 50000 \leq m \end{cases} \tag{3.6}$$

Devido a estes "saltos", a função-objetivo não é diferenciável em todo o seu domínio.

Problema de Otimização

O problema completo consiste em maximizar o lucro, o que é equivalente a minimizar o negativo do lucro, sujeito às restrições de cada variável.

$$\min_{d,t,m} -L(d, t, m) \tag{3.7}$$

sujeito a:

$$\begin{aligned} 0 \leq d &\leq 50 \\ 1 \leq t &\leq 30 \\ 1000 \leq m &\leq 50000 \end{aligned}$$

As restrições de caixa (box constraints) podem ser implementadas adicionando uma penalidade muito grande à função de custo caso sejam violadas.

3.2. Resolução e Escolha do método

O Algoritmo Escolhidos foram os Métodos de Busca Direta (sem Derivada), como Hooke-Jeeves ou Nelder-Mead Simplex. Dados que o problema apresenta função Não-Diferenciável por causa das penalidades baseadas em condições (if...then...) torna a função-objetivo não-contínua e não-diferenciável. Existem "saltos" e "penhascos" onde o gradiente não está definido.

A natureza não-diferenciável da função invalida completamente a base matemática dos métodos baseados em gradiente. Eles não funcionam de forma confiável em funções com "saltos". Os Métodos de gradientes até podem funcionar dependendo do salto, porém é instável e não garantido. Os métodos de busca direta, por outro lado, não se importam com a suavidade da função. Eles lidam naturalmente com os "penhascos" das penalidades, tratando-os apenas como regiões de custo muito alto a serem evitadas. Esta é a escolha correta e robusta para este tipo de problema. A tabela 3 demonstra o resultado para os algortitmos não diferenciaveis:

Table 3: Resultados da otimização para o Problema 3 (Maximização de Lucro).

Otimizador	**d (%)**	**t (dias)**	**m (R$)**	**Lucro (R$)**	**Iterações**
Hooke-Jeeves	20.000	30.000	10020.000	4122576.33	57
Nelder-Mead Simplex	23.342	29.999	10005.504	4070463.84	31

4. Problema 4

4.1. Definição do problema

Formulação Matemática do Problema 4: Aleta Unidimensional

O objetivo é encontrar as temperaturas de regime permanente, x_1 e x_2, que minimizam uma função de energia potencial.

Função-Objetivo

A função a ser minimizada é uma função quadrática de duas variáveis:

$$f(x_1, x_2) = 0.6382x_1^2 + 0.3191x_2^2 - 0.2809x_1x_2 - 67.906x_1 - 14.29x_2 \tag{4.1}$$

Forma Quadrática Matricial

Para analisar a natureza do problema, podemos reescrever a função $f(\mathbf{x})$ na forma matricial padrão para uma função quadrática, onde $\mathbf{x} = [x_1, x_2]^T$:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x} + \mathbf{g}^T\mathbf{x} \tag{4.2}$$

Neste caso, a matriz Hessiana $\mathbf{H}$ e o vetor do gradiente no ponto zero $\mathbf{g}$ são constantes, confirmando a natureza puramente quadrática do problema:

$$\mathbf{H} = \begin{pmatrix} 1.2764 & -0.2809 \\ -0.2809 & 0.6382 \end{pmatrix} \quad \text{e} \quad \mathbf{g} = \begin{pmatrix} -67.906 \\ -14.29 \end{pmatrix} \tag{4.3}$$

Problema de Otimização

O problema de otimização consiste em encontrar o vetor $\mathbf{x}^* = [x_1^*, x_2^*]^T$ que resolve:

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) \tag{4.4}$$

4.2. Resolução e Escolha do método

A natureza quadrática da Equação (4.1) é a principal justificativa para a escolha do Método de Newton como o algoritmo mais eficiente. Não preciso elaborar muito, a propria definção do algoritmo de Newton me diz: "Se a função for exatamente quadrática, basta se conhecer o gradiente e a Hessiana em um ponto qualquer x 0 para se determinar, em uma única iteração, o ponto de mínimo x, através da equação" A otimização da função-objetivo através do Método de Newton resultou nos seguintes valores ótimos:

- **Temperaturas de Regime Permanente (x_1^*, x_2^*):** $(64.3641, 50.7203)$ °C
- **Valor Mínimo da Função ($f(\mathbf{x}^*)$):** -2547.7231
- **Número de Iterações:** 1

A convergência em uma única iteração confirma a teoria de que o Método de Newton é o mais eficiente para uma função puramente quadrática, encontrando o mínimo exato em um único passo.

5. Conclusão

Este trabalho explorou a aplicação de um diversificado conjunto de algoritmos de otimização não linear irrestrita a quatro problemas com características matemáticas fundamentalmente distintas. Através da análise teórica e da implementação prática, foi possível extrair conclusões importantes sobre a eficácia e adequação de cada método.

A análise dos resultados revelou uma lição central na otimização: não existe um algoritmo universalmente superior. A escolha do método mais eficiente está intrinsecamente ligada às propriedades da função-objetivo. Para o **Problema 4**, de natureza puramente quadrática, o **Método de Newton** demonstrou sua superioridade teórica, convergindo para a solução exata em uma única iteração. Em contraste, no **Problema 1 (PID)**, cuja função-objetivo é resultado de uma simulação complexa e não-quadrática, a robustez e a capacidade adaptativa dos métodos **Quasi-Newton (BFGS)** se provaram mais eficientes que a abordagem potencialmente instável do Método de Newton.

Os desafios de problemas em alta dimensão foram evidenciados no **Problema 2 (Regressão Logística)**. Neste cenário, o custo computacional por iteração se tornou um fator crítico, destacando o importante trade-off entre a rápida convergência em iterações do **BFGS** e o menor custo por passo do **Gradiente Conjugado**. Por fim, o **Problema 3 (Maximização de Lucro)** foi categórico ao demonstrar que, para funções não-diferenciáveis devido a penalidades, a escolha correta recai obrigatoriamente sobre os **métodos de busca direta**, como Hooke-Jeeves e Nelder-Mead, que não dependem da existência do gradiente para operar.

Conclui-se, portanto, que o sucesso de um processo de otimização não reside apenas na implementação do algoritmo, mas fundamentalmente na análise criteriosa das características do problema — como dimensionalidade, diferenciabilidade, convexidade e custo de avaliação. A seleção do método mais compatível com a estrutura matemática da função-objetivo é o passo mais crítico para se obter uma solução de forma eficiente e confiável.