
Capítulo 2

Desenvolvimento

O principal método de otimização desenvolvido a fim de solucionar os problemas de programação quadrática é o método de Newton que busca uma aproximação de segunda ordem da função objetivo por meio da série de Taylor. Este método, considerando funções perfeitamente quadrática, possui convergência rápida. Todavia, devido a necessidade de determinar a derivada de primeira e de segunda ordem da função objetivo, o seu esforço computacional pode ser elevado. Além disso, em muitos casos os problemas são considerados uma caixa preta, ou seja, a função objetivo não está disponível; e em outros casos, mesmo a função objetivo estando disponível, o cálculo de suas derivadas de primeira e de segunda ordem é uma operação extremamente complexa. Isso implica na utilização de métodos numéricos para a determinação das derivadas, acarretando em um custo computacional ainda mais elevado.

Todo esse cenário proporcionou o desenvolvimento de novos métodos de otimização, que buscam manter a qualidade de convergência do método de Newton e, simultaneamente, diminuir o custo computacional por meio de calcular as derivadas da função objetivo de forma aproximada. Como já abordado nesse trabalho, esses métodos são conhecidos como métodos Quase-Newton.

A fim de verificar e comparar o desempenho dos algoritmos da família de métodos Quase-Newton (o BFGS e o DFP nas suas formas tradicionais e com as adaptações de Huang e de Biggs) serão consideradas seis funções de teste, as quais são apresentadas nas próximas seções.

2.1. Primeira Função de Teste

A primeira função de teste se trata de uma função objetivo com variável de decisão e dimensão dada por: n

$$f(x) = \frac{1}{2} \cdot (x - c)^T \cdot A \cdot (x - c) \quad (21)$$

Onde c é um vetor unitário e A é uma matriz diagonal. Ressalta-se que para os casos estudados A é a matriz identidade. Esses parâmetros são apresentados abaixo:

$$x_{1 \times n} = [x_1 \ x_2 \ \cdots \ x_n] \quad c_{1 \times n} = [1 \ 1 \ \cdots \ 1] \quad A_{n \times n} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & & 0 & 1 \end{bmatrix}$$

O gradiente da função apresentada em (21) é dado por:

$$\nabla f(x) = A \quad (x - c) \quad (22)$$

Os gráficos dessa função, do gradiente e das curvas de nível considerando duas variáveis são ilustrados abaixo:

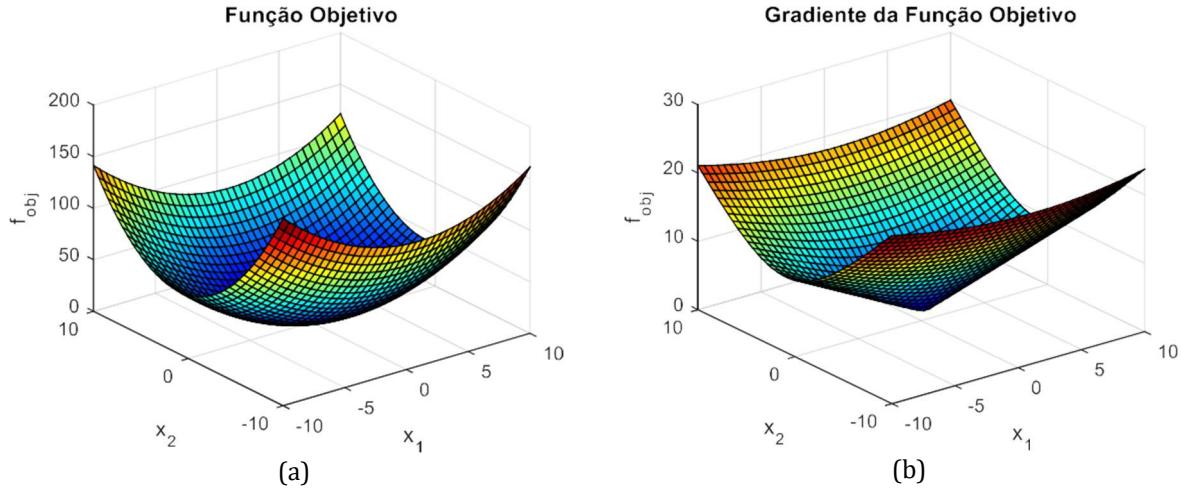


Figura 1 Primeira função de teste (a) Função objetivo; (b) Gradiente da função objetivo.

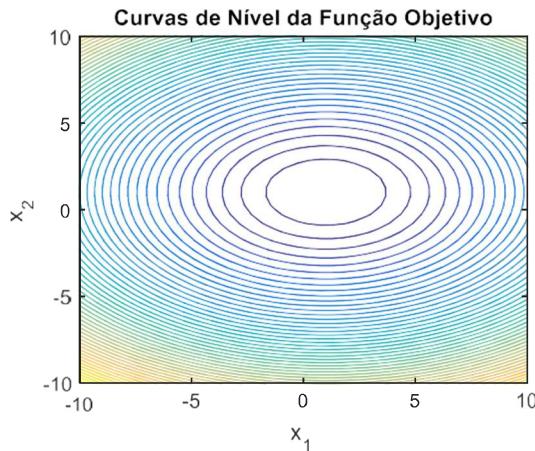


Figura 2-Curvas de nível da primeira função de teste.

Ao observar os gráficos ilustrados nas figuras 1 e 2 é fácil notar que a função apresentada em (21) é uma função quadrática deslocada com o mínimo posicionado em $x_{1 \times n} = [1 \quad 1 \quad \dots \quad 1]$ (em duas dimensões um paraboloide).

2.2. Segunda Função de Teste

A segunda função de teste se trata de uma função objetivo com duas variáveis independentes com termos quadráticos e um cúbico, em que o termo cúbico deve

apresentar menor influência no caso de se utilizar métodos Quase-Newton para determinar o mínimo da função. Essa função é definida como apresentado abaixo:

$$f(x) = 12 \cdot x_1^2 + 4 \cdot x_2^2 - 12 \cdot x_1 \cdot x_2 + 2 \cdot x_1 + a \cdot (x_1^3 + x_2^3) \quad (23)$$

Onde a é um parâmetro que assume valores que ponderam a influência do termo cúbico na função. O gradiente da função apresentada em (23) pode ser escrito como segue:

$$\nabla f(x) = \begin{cases} 24 \cdot x_1 - 12 \cdot x_2 + 2 + 3 \cdot a \cdot x_1^2 \\ 8 \cdot x_2 - 12 \cdot x_1 + 3 \cdot a \cdot x_2^2 \end{cases} \quad (24)$$

A solução ótima desse problema pode ser determinada ao considerar $\nabla f(x) = 0$. Nesse caso, conforme pode ser observado pelo sistema de equações em (24), a solução ótima é dependente da escolha de a , podendo ser obtida por meio da resolução desse sistema.

Os gráficos da função em (23), do gradiente e das curvas de nível considerando valores do parâmetro a igual a zero e igual a um são ilustrados abaixo:

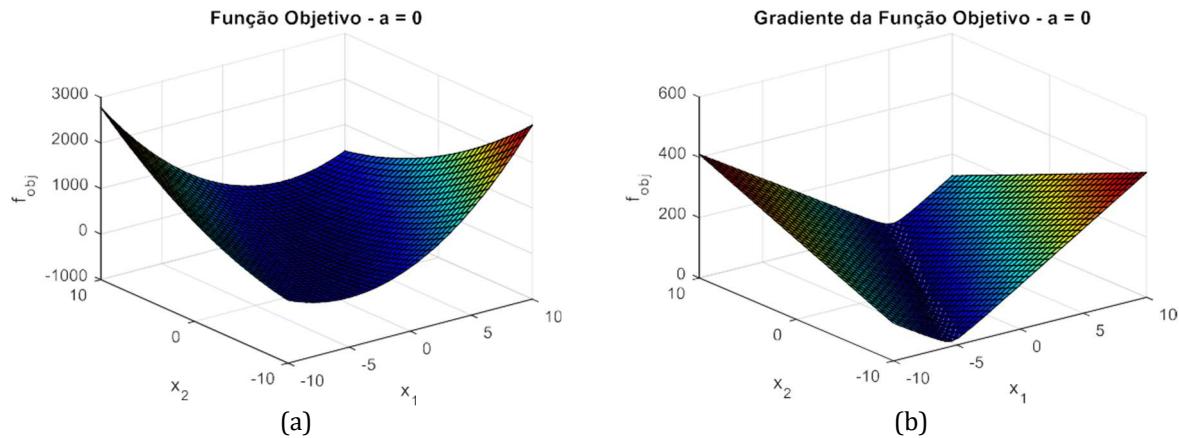


Figura 3 – Segunda função de teste para $a = 0$ (a) Função objetivo; (b) Gradiente da função objetivo.

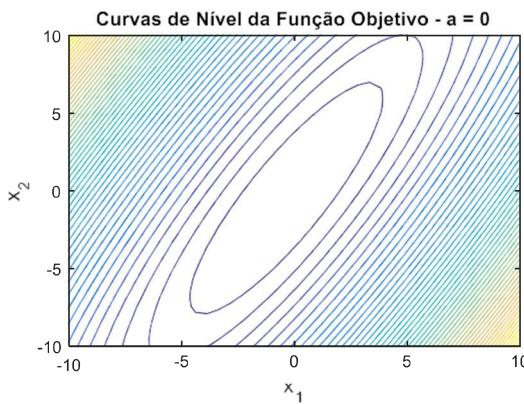


Figura 4 – Curvas de nível da segunda função de teste para $a = 0$.

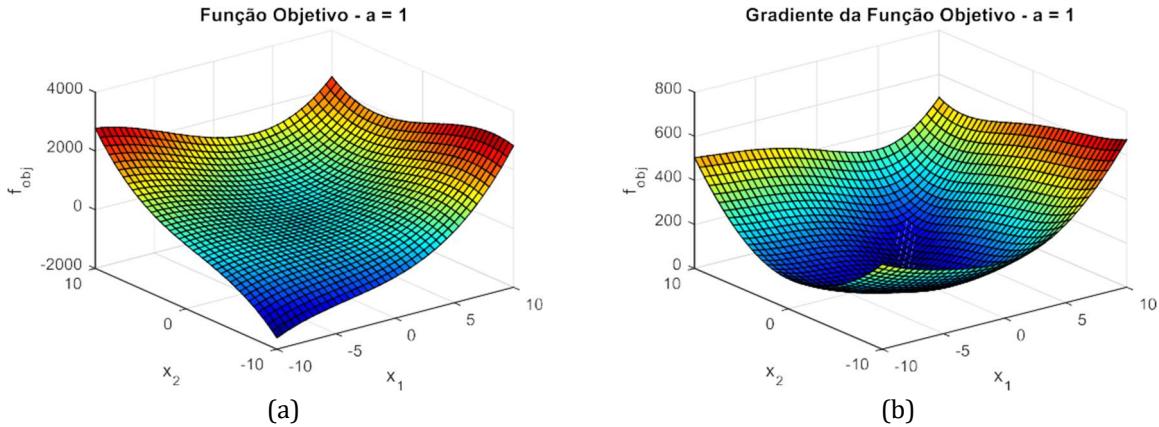


Figura 5 – Segunda função de teste para $a = 1$ (a) Função objetivo; (b) Gradiente da função objetivo.

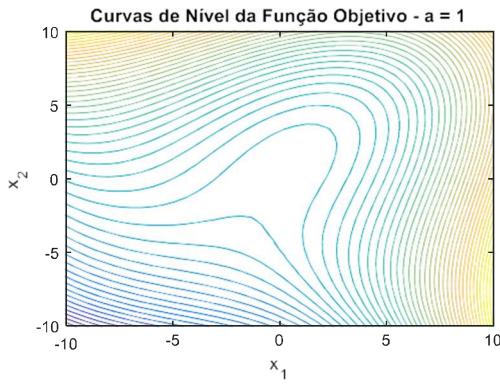


Figura 6 – Curvas de nível da segunda função de teste para $a = 1$.

Ao observar os gráficos ilustrados nas figuras 3 e 4 nota-se que a função apresentada em (23) é uma função quadrática com o mínimo igual a $x = [-1/3 \quad -1/2]$. Já os gráficos ilustrados nas figuras 5 e 6 mostram que esta mesma função é uma função não quadrática perfeita devido ao parâmetro a não nulo (nesse caso, $a = 1$). Além disso, o mínimo dessa função, como já mencionado, depende do parâmetro a .

Percebe-se que um importante aspecto relacionado à análise da função estudada é relativo ao intervalo de variação do parâmetro a . Ressalta-se que somente para pequenos valores de a ($-0,0263 \leq a \leq 0,0263$) a Hessiana da função em (23) será semi-definida positiva, exceto no caso em que a é nulo onde a Hessiana da função é definida positiva. Fora dessa faixa, ou seja, para valores de a fora do intervalo, a função deixa de ser convexa. Nesse caso, o termo cúbico passa a ter maior influência e, consequentemente, no processo de busca de soluções pode ocorrer a convergência para bacias de atração indesejadas. Dessa forma, a faixa de variação do parâmetro a deve ser dada de forma a respeitar a convexidade da função.

A matriz Hessiana da função estudada é apresentada pela equação abaixo:

$$H = \begin{bmatrix} 24 + 6 \cdot a \cdot x_1 & -12 \\ -12 & 8 + 6 \cdot a \cdot x_2 \end{bmatrix} \rightarrow \det(H) = (24 + 6 \cdot a \cdot x_1) \cdot (8 + 6 \cdot a \cdot x_2) - 144 \quad (25)$$

No caso em que a função em (23) apresenta termos não quadráticos insignificativos, o determinante apresentado na equação (25) é diferente de zero. Já no caso contrário, o determinante se aproxima de zero, o que evidencia a singularidade da hessiana e a expressividade do termo não quadrático.

2.3. Terceira Função de Teste

A terceira função de teste se trata de uma função objetivo com duas variáveis de decisão como apresentado abaixo:

$$f(x) = -8 \cdot x_1 \cdot x_2 + \frac{4}{a^2} \cdot x_2 \cdot x_1^3 + \frac{4}{a^2} \cdot x_1 \cdot x_2^3 \quad (26)$$

Ao observar a equação (26) percebe-se que a função objetivo $f(x)$ possui duas parcelas cúbicas. Todavia, ao considerar $a = \sqrt{2}$ e os intervalos para as variáveis de decisão igual a $[0,3]$ tem-se que a função $f(x)$ passa a se comportar como uma função quadrática. O gradiente da função apresentada em (26) pode ser escrito como segue:

$$\nabla f(x) = \begin{cases} -8 \cdot x_2 + \frac{12}{a^2} \cdot x_2 \cdot x_1^2 + \frac{4}{a^2} \cdot x_2^3 \\ -8 \cdot x_1 + \frac{12}{a^2} \cdot x_1 \cdot x_2^2 + \frac{4}{a^2} \cdot x_1^3 \end{cases} \quad (27)$$

Os gráficos dessa função, do gradiente e das curvas de nível considerando duas variáveis são ilustrados abaixo:

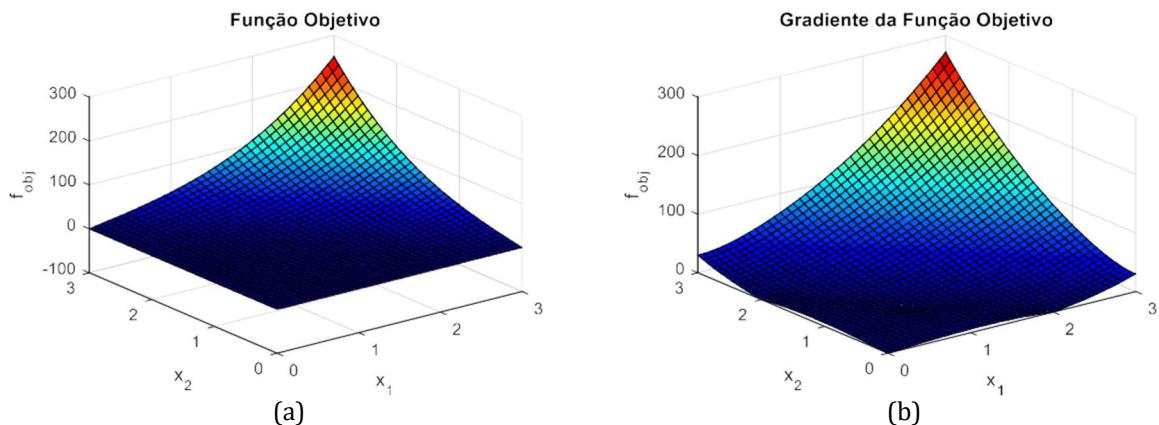


Figura 7-Terceira função de teste (a) Função objetivo; (b) Gradiente da função objetivo.

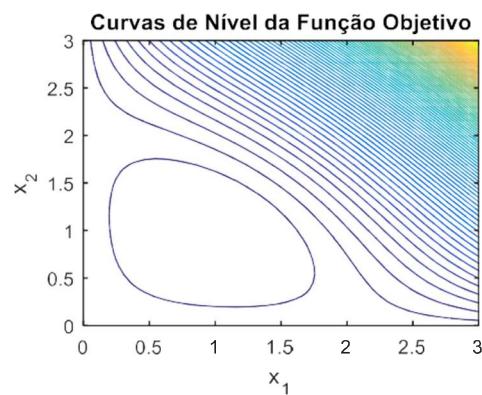


Figura 8-Curvas de nível da terceira função de teste.

Ao observar os gráficos ilustrados nas figuras 7 e 8 observa-se que a função apresentada em (26), no intervalo definido para as variáveis de decisão, pode ser considerada uma função quadrática cujo o mínimo é dado pelo ponto $x = [1 \ 1]$.

2.4. Quarta Função de Teste

A quarta função de teste se trata de uma função objetivo com variável de decisão x e de dimensão n dada por:

$$f(x) = \frac{1}{2} \cdot (x - c)^T \cdot A \cdot (x - c) + a \cdot \sum_{k=1}^n (x_k - 1)^3 \quad (28)$$

Onde a é uma parâmetro que pondera o termo cúbico, c é um vetor unitário e A é uma matriz diagonal. Ressalta-se que para os casos estudados A é a matriz identidade e c é um vetor unitário como apresentado abaixo:

$$x_{1 \times n} = [x_1 \ x_2 \ \dots \ x_n] \quad c_{1 \times n} = [1 \ 1 \ \dots \ 1] \quad A_{n \times n} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

No caso em que o parâmetro a é igual a zero a função $f(x)$ se torna idêntica a primeira função de teste apresentada na seção 2.1. Dessa forma, para o termo cúbico ser considerado, o valor de a deve assumir valores diferentes de zero.

O gradiente da função apresentada em (28) é dado por:

$$\nabla f(x) = A \cdot (x - c) + 3 \cdot a \cdot \sum_{k=1}^n (x_k - 1)^2 \quad (29)$$

A Hessiana dessa função é descrita como:

$$H_{n \times n} = \begin{bmatrix} 1 + 6 \cdot a \cdot (x_1 - 1) & 0 & \dots & 0 & 0 \\ 0 & 1 + 6 \cdot a \cdot (x_2 - 1) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 + 6 \cdot a \cdot (x_{n-1} - 1) & 0 \\ 0 & 0 & \dots & 0 & 1 + 6 \cdot a \cdot (x_n - 1) \end{bmatrix} \quad (30)$$

$$\det(H) = \underbrace{(1 + 6 \cdot a \cdot (x_1 - 1)) \cdot (1 + 6 \cdot a \cdot (x_2 - 1)) \cdot \dots \cdot (1 + 6 \cdot a \cdot (x_{n-1} - 1)) \cdot (1 + 6 \cdot a \cdot (x_n - 1))}_{\downarrow}$$

No intuito de obter uma Hessiana semi-definida positiva, ou seja, no caso onde não há singularidade, deve-se obter valores de a que respeitem a seguinte condição: $\det(H) \geq 0$

. O valor do parâmetro para função [eq. 28] foi $a = 0,0263$.

Os gráficos dessa função, do gradiente e das curvas de nível considerando duas variáveis são ilustrados abaixo:

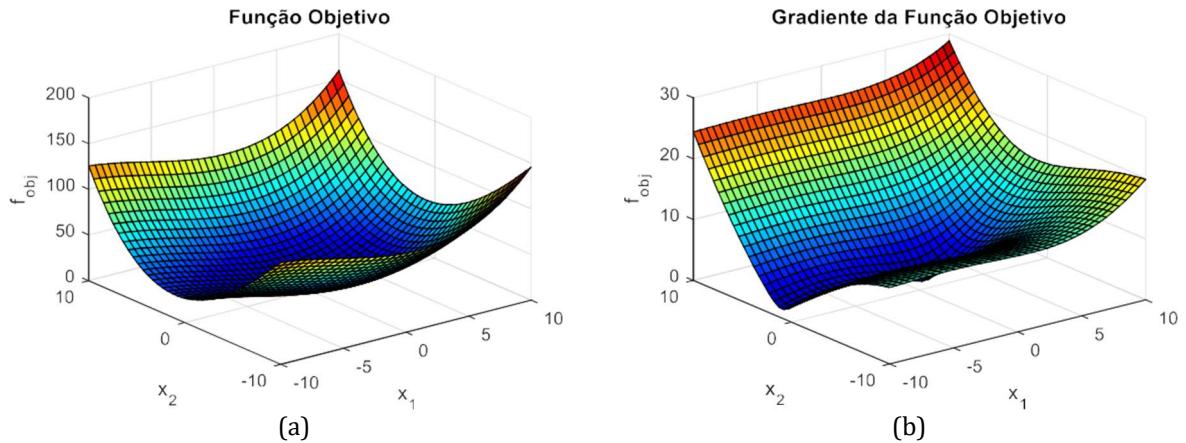


Figura 9-Quarta função de teste (a) Função objetivo; (b) Gradiente da função objetivo.

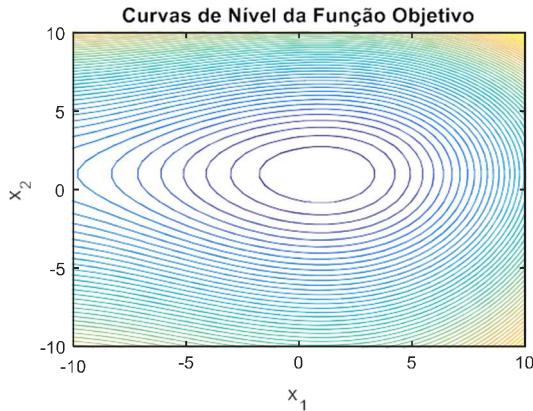


Figura 10-Curvas de nível da quarta função de teste.

Ao observar os gráficos ilustrados nas figuras 9 e 10 é fácil notar que a função apresentada em (28) é uma função não quadrática perfeita devido ao parâmetro $a = 0,0263$, não nulo (nesse caso). Além disso, o mínimo dessa função depende do valor do parâmetro a .

2.5. Quinta Função de Teste

A quinta função de teste é uma função com duas variáveis proveniente do problema de Rosen-Suzuki. Esse problema é considerado um problema de otimização restrito cuja função pseudo-objetivo pode ser considerada uma função quadrática. Vale ressaltar que a função pseudo-objetivo é dada pela soma da função objetivo com a função de penalidade que traduz as restrições não lineares, fazendo com que as soluções fornecidas por algum processo de otimização sejam factíveis.

Os gráficos dessa função, do gradiente e das curvas de nível considerando duas variáveis são ilustrados abaixo:

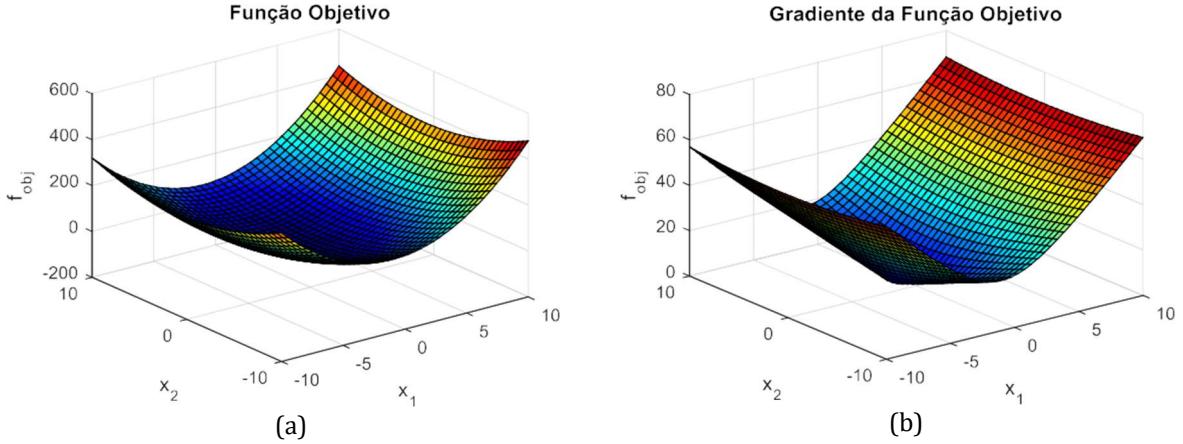


Figura 11 Quinta função de teste (a) Função objetivo; (b) Gradiente da função objetivo.

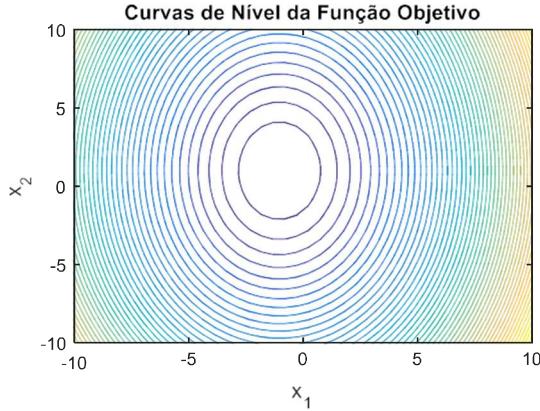


Figura 12-Curvas de nível da primeira função de teste.

Ao observar os gráficos ilustrados nas figuras 11 e 12 nota-se que a função pseudo-objetivo do problema Rosen-Suzuki é uma função quadrática com o mínimo posicionado em $x = [-1 \quad 1]$.

2.6. Sexta Função de Teste

A sexta e última função de teste a ser estudada nesse trabalho se trata de uma função objetivo com n variáveis de decisão conhecida como função de norma condicionada e inclinada (*tilted and conditioned norm*). Essa função na sua forma geral é apresentada abaixo:

$$f(x) = w \cdot \|A \cdot x\|_p + (w - 1) \cdot (A \cdot x)_1 \quad (31)$$

Sendo w e p parâmetros da função $f(x)$ e o índice 1 do termo $(A \cdot x)_1$ indica que apenas o primeiro elemento desse produto escalar é considerado. Além disso, a matriz A

é uma matriz quadrada conhecida como matriz de Hilbert em que os seus elementos são calculados de acordo com a equação abaixo:

$$A(i,j) = \frac{1}{1+i-j} \begin{cases} i \rightarrow N^{\text{o}} \text{ da linha} \\ j \rightarrow N^{\text{o}} \text{ da coluna} \end{cases}$$

Considerando $p = 2$, $w = 8$ e $x = [x_1 \ x_2]$ (problema com duas dimensões), a equação (31) pode ser escrita conforme apresentado abaixo:

$$f(x) = 8 \cdot \left\| x_1 + \frac{x_2}{2}, \frac{x_1}{2} + \frac{x_2}{3} \right\|_2 + 7 \cdot \left(x_1 + \frac{x_2}{2} \right) \quad (32)$$

No caso da equação apresentada em (32), o operador $\|\cdot\|_2$ equivale à norma Euclidiana (\mathbb{L}_2).

O gradiente da função $f(x)$ na sua forma geral é dado por:

$$\nabla f(x) = \left(\frac{w \cdot x^T \cdot A^T \cdot A}{\|A \cdot x\|_p} + (w - 1) \cdot (A \cdot x) \right)^T \quad (33)$$

Novamente, considerando $p = 2$, $w = 8$ e $x = [x_1 \ x_2]$ (problema com duas dimensões), a equação (33) pode ser escrita conforme apresentado abaixo:

$$\nabla f(x) = \left(\frac{8 \cdot x^T \cdot A^T \cdot A}{\left\| x_1 + \frac{x_2}{2}, \frac{x_1}{2} + \frac{x_2}{3} \right\|_2} + 7 \cdot (A \cdot x) \right)^T \quad (34)$$

Sendo:

$$A = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix} \quad x = [x_1 \ x_2]$$

Essa função, além de não ser uma função quadrática, apresenta um comportamento cujo valor do determinante da Hessiana é muito próximo de zero. Dessa forma, é constatado um problema de otimização com singularidade, o qual os métodos Quase-Newton possuem dificuldades de resolver, tornando relevante o estudo de como esses métodos conseguem contornar esse problema. Vale ressaltar que a solução ótima, no caso da função objetivo escrita em (32) é igual a $x = [0 \ 0]$.

Os gráficos da função apresentada nessa seção, do gradiente e das curvas de nível considerando duas variáveis são ilustrados abaixo:

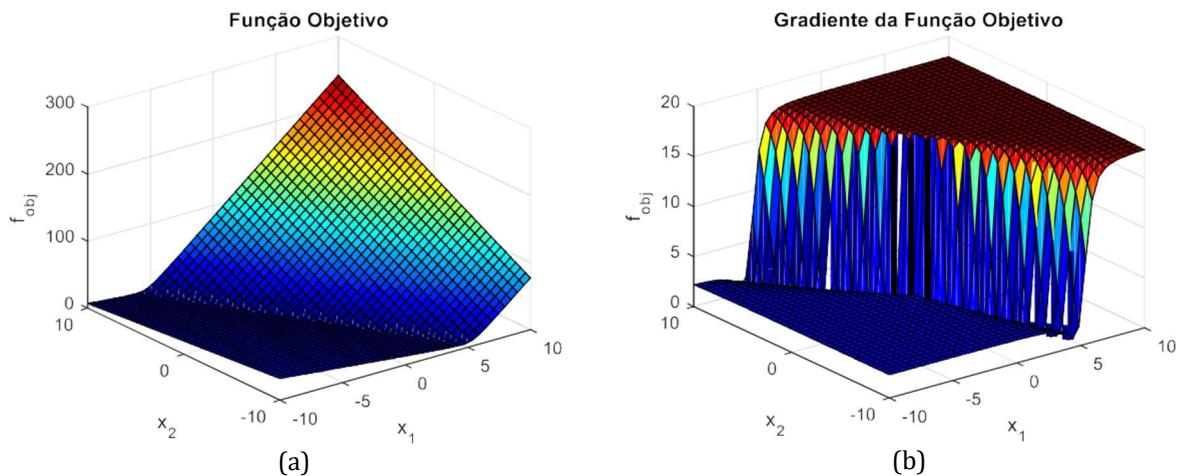


Figura 13 Sexta função de teste (a) Função objetivo; (b) Gradiente da função objetivo.

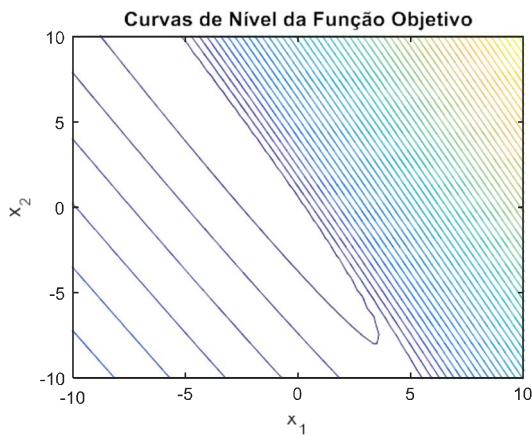


Figura 14-Curvas de nível da sexta função de teste.

Ao observar os gráficos ilustrados nas figuras 13 e 14 percebe-se a dificuldade que a singularidade traz na resolução dos problemas de otimização por meio dos métodos Quase-Newton.

2.7. Considerações dos Experimentos

A Tabela 1 apresenta de forma sumarizada como os experimentos devem ser realizados e analisados.

Tabela 1 – Detalhamento dos experimentos realizados para as análises dos Métodos Quase-Newton.

	Experimentos					
	1º	2º	3º	4º	5º	6º
Descrição	Avaliação das funções quadráticas	Avaliação das funções não quadráticas	Avaliação do custo computacional em função da variação da dimensão	Técnica da Seção Áurea_Avaliação direta da função e aproximações quadráticas da função a cada iteração	Avaliação dos métodos Quase-Newton em problemas com Hessiana singular	Avaliação estatística dos métodos Quase Newton com a variação do parâmetro a
Funções Avaliadas	1ª Função de Teste 2ª Função de Teste 3ª Função de Teste	2ª Função de Teste	4ª Função de Teste	4ª Função de Teste 5ª Função de Teste	6ª Função de Teste	2ª Função de Teste
Código das Funções Avaliadas no Matlab®	<i>fex1</i> <i>fexlivro</i> <i>fex3</i>	<i>Fexlivro</i>	<i>fex1</i>	<i>fexlivro</i> <i>fex1</i> <i>fun_rosensuzuki_irr</i>	<i>f_tiltednormcond</i>	<i>fexlivro</i>

Nos experimentos realizados nesse trabalho alguns parâmetros devem ser fixados a fim executar os algoritmos desenvolvidos em ambiente Matlab e possibilitar uma análise adequada. A descrição de tais parâmetros de uma forma geral e mais específica, considerando os experimentos, é apresentada abaixo:

- **Parâmetro:** *Imetqn* → Chaveia o algoritmo Quase-Newton a ser executado
- **Parâmetro:** *icaso* → Chaveia a função objetivo a ser estudada
- **Parâmetro:** *a* → Pondera os termos não quadráticos
- **Parâmetro:** *dim* → Determina a dimensão do problema a ser resolvido
- **Parâmetro:** *xstar* → Estabelece a solução ótima do problema
- **Parâmetro:** *x0* → Estabelece o ponto inicial
- **Parâmetro:** *isa_FV* → Chaveia a técnica da Seção Áurea

1º Experimento:

Função: fex1	Função: fexlivro	Função: fex3
$Dimensão \rightarrow dim = 2$	$Dimensão \rightarrow dim = 2$	$Dimensão \rightarrow dim = 2$
$Parâmetro \rightarrow a = 0$	$Parâmetro \rightarrow a = 0$	$Parâmetro \rightarrow a = \sqrt{2}$
Função Quadrática	Função Quadrática	Função Quadrática

2º Experimento:

Função: fexlivro
$Dimensão \rightarrow dim = 2$
$Parâmetro \rightarrow a = \begin{cases} -0,0263 \\ 0,0263 \\ 1 \end{cases}$
Função Não Quadrática

3º Experimento:

Função: fex1
$Dimensão \rightarrow dim = \begin{cases} 5 \\ 10 \\ 15 \\ 20 \\ 25 \end{cases}$
$Parâmetro \rightarrow a = 0,0263$

Função Não Quadrática

4º Experimento:

Função: fun_rosensuzuki_irr	Função: fex1
$Dimensão \rightarrow dim = 2$	$Dimensão \rightarrow dim = 30$

Função Quadrática

Função Não Quadrática

5º Experimento:

Função: f_tiltednormcond
$Dimensão \rightarrow dim = 2$
$Parâmetros \rightarrow p = 2$
$w = 8$

Função Não Quadrática

6º Experimento:

Função: fexlivro
$Dimensão \rightarrow dim = 2$
$Parâmetro \rightarrow a = \begin{cases} -0,0263 \\ 0 \\ 0,0263 \end{cases}$

Função Quadrática e Função Não Quadrática

A função desenvolvida no Matlab que soluciona os problemas propostos nesse trabalho é codificada como *otqnmat_a77*. Os parâmetros dessa função são pontuados a seguir:

- **Parâmetro de entrada:** *funcao* → Função a ser aproximada
- **Parâmetro de entrada:** *isa_FV* → Seção Áurea
- **Parâmetro de entrada:** *Imetqn* → Método Quase-Newton escolhido
- **Parâmetro de entrada:** *MAXITER* → Número máximo de iterações
- **Parâmetro de entrada:** *x0* → Ponto inicial
- **Parâmetro de entrada:** *xMin* → Limite inferior
- **Parâmetro de entrada:** *xMax* → Limite superior
- **Parâmetro de entrada:** *epsilon* → Precisão
- **Parâmetro de saída:** *XK1* → Armazena a solução das variáveis de decisão
- **Parâmetro de saída:** *Hfobj* → Armazena os valores da função objetivo
- **Parâmetro de saída:** *icfunc* → Número de avaliações da função objetivo

Finalmente, a precisão que define um dos critérios de parada dos algoritmos Quase-

Newton executados nesse trabalho será 10^{-6} . Sendo ***dim*** a dimensão do problema de otimização a ser resolvido. O número máximo de iterações define um segundo critério de parada e será:

$$N^{\circ} \text{ Máximo de Iteração} = 2 \cdot \dim \cdot 100 \quad (35)$$