



18 a 21 de novembro de 2014, Caldas Novas - Goiás

ALGUNS MÉTODOS CLÁSSICOS DE OTIMIZAÇÃO

Pedro Henrique Rodrigues da Silva, pedro.henrique-02@hotmail.com¹

Milena Almeida Leite Brandão, milena@pontal.ufu.br¹

¹Universidade Federal de Uberlândia, Faculdade de Ciências Integradas do Pontal, Ituiutaba - MG - Brasil

Resumo. Em linhas gerais, otimizar significa planejar ou desenvolver com maior eficiência. Após a formulação do modelo, um método de otimização é utilizado para obter a solução do problema. Pode-se dividir os métodos de otimização entre os que dependem ou não do conhecimento de derivadas. Os métodos de otimização que não dependem do uso das derivadas geralmente utilizam apenas informações da função a ser otimizada, que pode ser descontínua e multimodal. A maior vantagem destes métodos é o enorme potencial prático devido à crescente necessidade de resolver problemas de otimização definidos por funções para as quais não existem derivadas ou estas são disponíveis a um custo computacional muito elevado. Nos métodos de otimização que dependem do uso das derivadas o modelo é completamente conhecido e são para funções contínuas, convexas e semi modais. Neste trabalho, serão apresentados dois métodos de otimização: o Método das Coordenadas Cíclicas, que não depende de derivadas pois utiliza a direção dos eixos coordenados como direções de busca e o Método de Newton cuja direção de busca utiliza as derivadas da função a ser otimizada. Em seguida, um problema de modelagem matemática será resolvido por ambos os métodos e os resultados obtidos serão analisados.

Palavras-chave: Otimização Determinística, Método das Coordenadas Cíclicas, Método de Newton.

1. INTRODUÇÃO

O termo otimização refere-se ao estudo de problemas em que se busca minimizar ou maximizar uma função através da escolha dos valores de variáveis dentro de um conjunto viável. Segundo Takahashi (2003), no contexto da engenharia, os sistemas de otimização se destinam a empregar técnicas e métodos para determinar a melhor solução de problemas abstratos, para os quais se deseja quantificar o grau de adequação de cada possível solução às necessidades que os causaram.

De acordo com Hartman e Mutmanský (2002), o processo de identificação dos objetivos, variáveis e restrições é denominado modelagem, que significa criar um modelo que explique as características de funcionamento e comportamento de um determinado problema, facilitando seu entendimento e projeto.

Existem diversos métodos de otimização e desta forma vários autores já apresentaram diferentes propostas para classificá-los, não sendo possível escolher uma proposta universalmente aceita. Os métodos dependem do tipo do problema em estudo (programação linear ou não-linear, problemas restritos ou irrestritos), das variáveis de projeto (contínuas ou discretas), do tipo de resposta desejada (otimização local ou global; projeto ótimo ou controle ótimo), das técnicas empregadas (determinísticas, estocástica ou híbrida). Além disso, os problemas podem ser classificados segundo o número de variáveis (pequenos ou grandes) ou segundo a suavidade das funções (diferenciáveis ou não-diferenciáveis).

Ainda, pode-se dividir os métodos de otimização entre os que dependem ou não do conhecimento de derivadas. Os métodos de otimização que não dependem do uso das derivadas geralmente utilizam apenas informações da função a ser otimizada, que pode ser descontínua, não diferenciável e multimodal. A maior desvantagem destes métodos é o alto custo computacional já que necessitam de muitas avaliações da função objetivo. Por outro lado, os métodos de otimização sem derivadas tem tido um enorme potencial prático devido à crescente necessidade de resolver problemas de otimização definidos por funções para as quais não existem derivadas ou estas são disponíveis a um custo computacional muito elevado.

Nos métodos de otimização que dependem do uso das derivadas o modelo é completamente conhecido e são para funções contínuas, convexas e semi modais. Os melhores resultados destes métodos são para funções contínuas, convexas

e semi-modais. A grande vantagem na utilização destes métodos é o baixo número de avaliações da função objetivo, fazendo com que tenham convergência rápida e baixo custo computacional.

Visto que não existe um algoritmo universal, mas um conjunto de métodos entre os quais alguns são mais apropriados para determinadas aplicações específicas, a escolha do método de otimização dependerá do usuário e poderá determinar a eficácia ou falha para a solução do problema, que após a formulação do modelo, será resolvido por meio de um algoritmo de otimização, com a ajuda de códigos computacionais.

Neste trabalho, serão apresentados as principais formulações matemáticas de um problema de otimização, a estratégia de busca linear para o algoritmo se mover de uma iteração para a outra e, de forma resumida, os métodos de otimização das Coordenadas Cíclicas, que não depende do uso de derivadas pois utiliza a direção dos eixos coordenados como direções de busca e o Método de Newton cuja direção de busca utiliza as derivadas da função a ser otimizada. Em seguida, um problema de modelagem matemática será resolvido por ambos os métodos e os resultados obtidos serão analisados.

2. FUNDAMENTOS DE OTIMIZAÇÃO IRRESTRITA

Segundo Camponogara (2006), normalmente procura-se encontrar uma solução ótima global. No entanto, é extremamente difícil encontrar uma solução global visto que, geralmente, o conhecimento que se tem de uma função é apenas local.

A maioria dos algoritmos usados para resolver os problemas de otimização são capazes de encontrar uma solução ótima para certa vizinhança de uma função. Eles iniciam com uma estimativa inicial e seguem com sequências de aproximações para obter o ponto ótimo.

A formulação matemática do problema de otimização pode ser escrito como:

$$\begin{array}{ll} \min f(x) & \text{sujeito a} \\ x \in \mathbb{R}^n & c_i(x) = 0, i \in I \\ & c_i(x) \leq 0, i \in D \end{array} \quad (1)$$

onde,

- x é o vetor das variáveis de projeto;
- f é a função objetivo;
- c_i são as funções de restrições;
- I e D são os conjuntos de índices das restrições de igualdade e desigualdade, respectivamente.

Em seguida, são apresentadas algumas definições e teoremas a fim de se estabelecer as condições de existência do ponto ótimo. Um estudo detalhado das demonstrações dos teoremas pode ser visto em Brandão (2010).

Sejam $B_\delta(x^*)$ uma bola aberta de centro x^* e raio δ , $\nabla f(x)$ o gradiente da função f no ponto x e $\nabla^2 f(x)$ a matriz Hessiana de f no ponto x .

Definição 2.1 Um ponto $x^* \in \mathbb{R}^n$ é mínimo global se $f(x^*) \leq f(x)$ para todo $x \in \mathbb{R}^n$. Um ponto $x^* \in \mathbb{R}^n$ é mínimo local se existe uma bola aberta de centro x^* e raio δ tal que $f(x^*) \leq f(x, y)$ para todo $x \in B_\delta(x^*)$. E um ponto $x^* \in \mathbb{R}^n$ é mínimo local estrito se existir uma $B_\delta(x^*)$ tal que $f(x^*) < f(x)$ para todo $x \in B_\delta(x^*)$, com $x \neq x^*$.

Teorema 2.1 (Condição necessária de primeira ordem): Se x^* é um minimizador local de f possui primeira derivada contínua em uma vizinhança aberta de x^* , então $\nabla f(x^*) = 0$. Chama-se x^* de ponto estacionário.

O ponto x^* é chamado de ponto estacionário de f se $\nabla f(x^*) = 0$. Logo, pelo teorema acima um mínimo local é um ponto estacionário.

Teorema 2.2 (Condição necessária de segunda ordem): Se x^* é um minimizador local de f e $\nabla^2 f$ é contínua em uma vizinhança aberta de x^* então $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ é semi definida positiva.

Teorema 2.3 (Condição suficiente de segunda ordem): Suponha $\nabla^2 f(x^*)$ seja contínua em uma vizinhança aberta de x^* , que $\nabla f(x^*) = 0$ e que $\nabla^2 f(x^*)$ seja definida positiva. Então x^* é um mínimo local estrito de f .

O objetivo dos algoritmos de otimização é encontrar uma aproximação x^{k+1} que corresponde a um valor da função inferior a $f(x^k)$, ou seja, $f(x^{k+1}) < f(x^k)$. Nos métodos de busca linear, a cada iteração é calculado a direção de busca p_k e então se decide o quanto se deve mover ao longo desta direção. A iteração é dada por:

$$x^{k+1} = x^k + \alpha_k p_k, \quad (2)$$

sendo que α_k é um escalar positivo chamado de tamanho ou comprimento de passo. O sucesso do método de busca linear depende de uma escolha eficaz tanto da direção p_k quanto do tamanho do passo α_k .

Durante a computação do passo α_k , depara-se com um conflito. Por um lado, deseja-se escolher α_k que produza uma redução substancial no valor de f mas, por outro lado, não se deseja gastar muito esforço computacional durante a escolha. O passo ideal é o que minimiza a função $\varphi(\alpha)$, definida por:

$$\varphi(\alpha) = f(x_k + \alpha p_k), \text{ onde } \alpha > 0. \quad (3)$$

Geralmente é muito caro identificar o passo ideal, pois mesmo para encontrar um mínimo local, usando uma precisão moderada, são requeridas muitas avaliações da função objetivo e de seu gradiente. A estratégia prática é uma busca

inexata, identificando um comprimento de passo que reduz o valor da função. Normalmente, em uma busca típica, obtém-se uma sequência de valores para α , finalizando quando alguma condição de parada é satisfeita. A busca é feita em duas fases: na primeira encontra-se um intervalo que contém o passo desejado e na segunda, calcula-se o melhor passo neste intervalo.

Uma simples condição que pode ser imposta é requerer a redução do valor de f , ou seja,

$$f(x^k + \alpha_k p_k) < f(x^k) \quad (4)$$

Mas esta condição não é suficiente para garantir a convergência para o ótimo, conforme citado por Nocedal e Wright (2000). Além disso, é preciso garantir uma condição de decrescimento suficiente. Neste estudo será utilizado o Método da Seção Áurea para obter o comprimento do passo. Definido o processo iterativo $x^{k+1} = x^k + \alpha_k p_k$, a otimização de uma função de várias variáveis recai, a cada iteração, na busca unidimensional de α_k que minimiza a função $f(x^k + \alpha p_k)$. Este processo é conhecido como Método da Seção Áurea. O problema é definido em um intervalo de incerteza de busca $I = [a_0; b_0]$. Assumindo que $\varphi(\alpha)$ é unimodal, então $\varphi(\alpha)$ é decrescente em $[a_0; a_k]$ e crescente em $[a_k; b_0]$.

Definição 2.2 Seja α_k o ponto mínimo em $[a_0; b_0]$ de uma função unimodal $\varphi(\alpha)$. Sejam $c_0, d_0 \in I_0$ sendo $c_0 < d_0$.

Caso 1 Se $\varphi(c_0) > \varphi(d_0)$, então $\alpha_k \in [c_0; b_0]$

Caso 2 Se $\varphi(c_0) < \varphi(d_0)$, então $\alpha_k \in [a_0; d_0]$

Caso 3 Se $\varphi(c_0) = \varphi(d_0)$, então $\alpha_k \in [c_0; d_0]$

Uma ilustração de cada caso apresentado na Definição 2.2 é dada nas Figs. 1, 2 e 3, respectivamente.

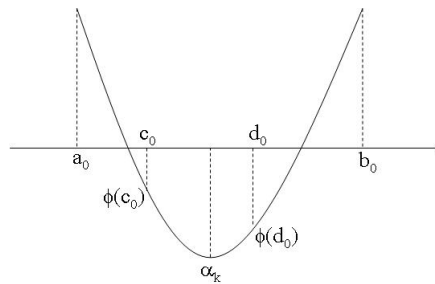


Figura 1. Caso 1 do método da Seção Áurea: $\phi(c_0) > \phi(d_0) \Rightarrow \alpha_k \in [c_0, b_0]$

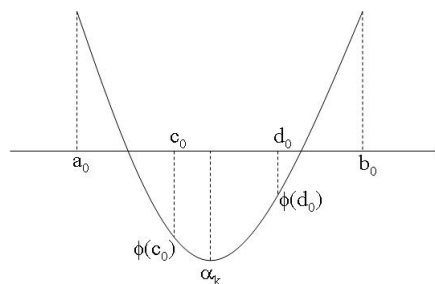


Figura 2. Caso 2 do método da Seção Áurea: $\phi(d_0) > \phi(c_0) \Rightarrow \alpha_k \in [a_0, d_0]$

A aplicação iterativa desta definição conduz a intervalos cada vez menores. Quando os critérios de parada são obedecidos escolhe-se no último intervalo o α que minimiza a função φ . Para uma consideração mais detalhada do Método da Seção Áurea veja Nocedal e Wright (2000) e Brandão (2010).

3. MÉTODO DAS COORDENADAS CÍCLICAS

Segundo Camponogara (2006), os problemas de otimização sem derivadas surgem de modelos para os quais as derivadas das funções e das restrições envolvidas, por alguma razão, não estão disponíveis. Os motivos variam desde usuários que não querem programar as derivadas até funções excessivamente complexas, oriundas de simulações só possíveis devido ao crescimento na capacidade de processamento dos computadores. Por isso, o número de algoritmos para resolver problemas de otimização sem derivadas aumentou nos últimos anos. Porém, poucos são aqueles que conseguem lidar de forma eficiente com problemas cujos domínios são pequenos. Portanto uma solução quase ótima obtida através de um

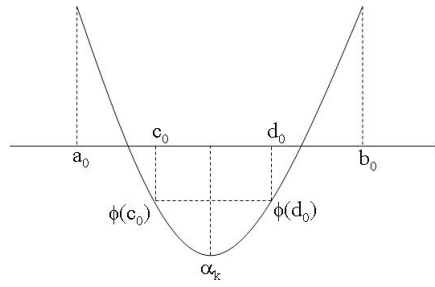


Figura 3. Caso 3 do método da Seção Áurea: $\phi(c_0) = \phi(d_0) \Rightarrow \alpha_k \in [c_0, d_0]$

método sem derivadas é muitas vezes menos precisa do que a obtida por um método com derivadas, assumindo que a informação da derivada está disponível. Mesmo assim, para muitas aplicações estas limitações são aceitáveis.

Pode-se citar como exemplo de métodos de otimização independentes de derivadas os métodos determinísticos das Coordenadas Cíclicas, de Hooke e Jeeves e de Rosenbrock e os métodos heurísticos tais como Algoritmo Genético e Evolução Diferencial.

De acordo com Brandão (2010), Coordenadas Cíclicas é um método de otimização que utiliza a direção dos eixos coordenados como direções de busca. Mais especificamente, o método caminha através das direções d_1, \dots, d_n onde $d_j, j = 1, \dots, n$, é um vetor de zeros, exceto na j -ésima posição onde assume o valor 1.

Desta forma, caminhando na direção de busca d_j , a variável x_j é modificada, enquanto todas as outras permanecem fixas. Assume-se que a minimização é feita nas componentes x_1, \dots, x_n , seguindo esta ordem, ou seja, primeiro na variável x_1 , depois x_2 e assim, sucessivamente. Note que a minimização é realizada em todas as direções a cada iteração. Além disso, se a função objetivo for diferenciável o método das coordenadas cíclicas convergirá para um ponto estacionário.

A Fig. 4 mostra um esquema das primeiras duas iterações do Método das Coordenadas Cíclicas de um problema bidimensional.

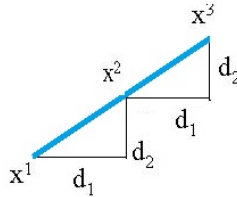


Figura 4. Primeiras duas iterações do Método das Coordenadas Cíclicas.

Em seguida, será apresentado um algoritmo do métodos das Coordenadas Cíclicas conforme proposto por Bazaraa et al. (2006). Observe que a minimização da função objetivo é feita sem usar informações das derivadas.

Dados Iniciais: Escolha um $\epsilon > 0$ para interromper o algoritmo e seja d_1, \dots, d_n as direções coordenadas. Escolha um ponto inicial x^1 faça $y^1 = x^1$ e $k = j = 1$ e inicie o processo iterativo.

Passo 1: Seja α_j uma solução ótima do problema: $\min f(y^j + \alpha d_j)$, com $\alpha \in \mathbb{R}$. Seja $y^{j+1} = y^j + \alpha_j d_j$. Se $j < n$, troque j por $j + 1$ e repita o passo 1. Caso contrário, se $j = n$, vá para o Passo 2.

Passo 2: Seja $x^{k+1} = y^{n+1}$. Se $\|x^{k+1} - x^k\| < \epsilon$ então pare. Caso contrário, seja $y^1 = x^{k+1}$ e $j = 1$, substitua k por $k + 1$ e repita o passo 1.

4. MÉTODO DE NEWTON

A direção de busca dos métodos de Newton, de acordo com Nocedal e Wright (2000), é uma das mais notáveis já que o método converge rapidamente para a solução ótima quando adota um ponto inicial suficientemente próximo da solução. A direção de busca é obtida através da aproximação de $f(x^k + p)$ por uma série de Taylor de 2º ordem, ou seja:

$$f(x^k + p) \approx f^k + p^T \nabla f^k + \frac{1}{2} p^T \nabla^2 f^k p = m_k(p). \quad (5)$$

Por isso, o método de Newton costuma ser mais preciso, já que considera a aproximação da função até 2ª ordem. Assumindo que $\nabla^2 f_k$ é positiva definida, obtem-se a direção de Newton encontrando o vetor p que minimiza $m_k(p)$.

Derivando $m_k(p)$ com relação a p e igualando a zero, obtém-se a seguinte relação:

$$\frac{\partial m_k(p)}{\partial p} = \nabla f^k + \nabla^2 f^k p_k = 0, \quad (6)$$

ou seja,

$$p_k = -(\nabla^2 f^k)^{-1} \nabla f^k. \quad (7)$$

Métodos que utilizam a direção Newton tem uma rápida convergência local, tipicamente quadrática. Além disso, quando a vizinhança da solução é alcançada, a convergência geralmente ocorre em poucas iterações. A principal desvantagem da direção de Newton é o uso da inversa da matriz Hessiana $\nabla^2 f(x)$ já que o cálculo desta matriz de segundas derivadas é, na maioria dos casos, um processo caro e sujeito a erros.

Quando a Hessiana $\nabla^2 f^k$ não é positiva definida, a direção p_k pode não ser uma direção de descida. No entanto, de acordo com o Teorema 4.1, se para todo x vizinho da solução x^* a matriz $\nabla^2 f(x)$ também é positiva definida então, o Método de Newton será bem definido nesta região e convergirá quadraticamente.

Teorema 4.1 *Suponha que a função f seja duas vezes diferenciável e que a Hessiana $\nabla^2 f(x)$ seja Lipschitz contínua. Então na vizinhança da solução x^* as condições suficientes de 2ª ordem definidas no Teorema 2.3 são satisfeitas. Considere a iteração $x^{k+1} = x^k + p_k$ sendo p_k dado pela equação (7), então:*

1. *se o ponto de partida x^0 é suficientemente próximo de x^* , a sequência de iterações x^k converge para x^* ;*
2. *a razão de convergência de x^k é quadrática;*
3. *a sequência de normas do gradiente $\|\nabla f^k\|$ converge quadraticamente para zero.*

A demonstração do Teorema 4.1 pode ser vista em Brandão (2010).

5. MÉTODO DA PENALIDADE QUADRÁTICA

Muitos algoritmos de otimização foram desenvolvidos para resolver problemas irrestritos nos quais as variáveis de otimização podem assumir quaisquer valores na busca pelo ponto onde a função objetivo é minimizada. Este é o caso dos métodos das Coordenadas Cíclicas e de Newton. Assim, torna-se necessário utilizar algum artifício para que os problemas com restrições também se apliquem aos métodos de otimização irrestrita.

Uma das abordagens fundamentais para a otimização restrita é substituir o problema original por uma função de penalidade que é composta da função objetivo original do problema de otimização com restrições somada a um termo adicional para cada restrição, o qual é positivo quando o ponto atual x viola essa restrição e zero caso contrário.

Segundo Dasgupta et al. (1996), os métodos de Penalidade foram criados para resolver um problema P , solucionando uma sequência de problemas sem restrições especialmente escolhidos. Assim é contruída uma nova função objetivo, Φ , que contém informação relativa à função objetivo inicial, f , e, simultaneamente, às restrições, que dependem de um parâmetro positivo, r_p , cujas correspondentes soluções, $x_{r_p}^*$, convergirão para a solução do problema inicial, x^* .

Considere o seguinte problema:

$$\begin{aligned} \min f(x), \quad & X \in \mathbb{R}^n \text{ sujeito a:} \\ g_j(x) \leq 0, \quad & \text{restrições de desigualdade} \\ h_l(x) = 0, \quad & \text{restrições de igualdade.} \end{aligned} \quad (8)$$

Como o objetivo da otimização é minimizar a função objetivo $f(x)$, que está sujeita às restrições, a ideia é penalizar os pontos que não pertencem à região viável por aumentar o valor da função nestes pontos. Desta forma, os pontos não viáveis serão descartados durante a execução de um algoritmo de otimização.

A nova função objetivo, chamada *pseudo objetivo*, é penalizada de acordo com o fator de penalidade r_p , toda vez que encontrar uma restrição ativa. O modelo matemático da função *pseudo objetivo* $\Phi(x)$ utilizada para minimizar a função $f(x)$ é dada por:

$$\Phi(x) = f(x) + r_p P(x) \quad (9)$$

e a função de penalidade $P(x)$ é dada por:

$$P(x) = \sum_{j=1}^m (\max[0, g_j(x)]^2) + \sum_{k=1}^l h_k(x)^2 \quad (10)$$

sendo $g_j(x)$ as restrições de desigualdade, $h_k(x)$ as restrições de igualdade e m e l o número de restrições de desigualdade e igualdade respectivamente.

6. SIMULAÇÕES NUMÉRICAS

Para testar e avaliar os algoritmos das Coordenadas Cíclicas e de Newton, é proposto, a seguir, um problema que se conhece a solução analítica pois, desta forma, pode-se comparar a solução encontrada com os métodos com a analítica. O código foi escrito na forma de um toolbox para *MATLAB*®.

Considere o seguinte problema: Deseja-se projetar uma caixa retangular sem tampa, conforme ilustrado na Fig. 5, com a restrição de que a área total deve ser de 12m^2 . Obtenha o volume máximo dessa caixa.

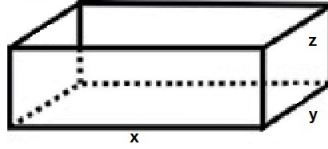


Figura 5. Caixa retangular sem tampa.

6.1. Solução analítica

Sejam x, y, z o comprimento, a largura e a altura, respectivamente, da caixa em metros. Tem-se como objetivo maximizar $V = xyz$ sujeito a restrição $g(x, y, z) = 2xz + 2yz + xy = 12$. De acordo com Stewart (2005), utilizando o Método dos Multiplicadores de Lagrange, obtem-se o seguinte sistema linear:

$$\begin{cases} \nabla V = \lambda \nabla g \\ g(x, y, z) = 12 \end{cases} \quad (11)$$

Ou seja,

$$yz = \lambda(2z + y) \quad (12)$$

$$xz = \lambda(2z + x) \quad (13)$$

$$xy = \lambda(2x + 2y) \quad (14)$$

$$2xz + 2yz + xy = 12 \quad (15)$$

Para resolver este sistema primeiro multiplica-se as equações (12), (13) e (14) por x, y e z , respectivamente, o que resulta no seguinte sistema:

$$xyz = \lambda(2xz + xy) \quad (16)$$

$$xyz = \lambda(2yz + 2xy) \quad (17)$$

$$xyz = \lambda(2xz + 2yz) \quad (18)$$

Note que $\lambda \neq 0$ pois, caso contrário, teria-se $yz = xz = xy = 0$ o que contradiz a equação (15). Assim, das equações (16) e (18), tem-se:

$$2xz + xy = 2yz + xy$$

donde $xz = yz$. Mas $z \neq 0$ (uma vez que se $z = 0$ resultaria em $V = 0$), portanto $x = y$. Das equações (17) e (18), obtem-se:

$$2yz + xy = 2xz + 2yz$$

logo $2xz = xy$ e, assim, (já que $x \neq 0$), $y = 2z$. Substituindo $x = y = 2z$ na equação (15), vem que:

$$4z^2 + 4z^2 + 4z^2 = 12$$

Como x, y, z são todos positivos, a solução ótima do problema é $x = 2, y = 2, z = 1$. Portanto, segue que $V = 4\text{m}^3$.

6.2. Solução numérica

Visto que o problema de obter o volume máximo da caixa é restrito foi utilizado o Método da Penalidade para transformá-lo em irrestrito. Dessa forma, a nova função objetivo pode ser expressa pela seguinte equação:

$$F(x) = xyz - r_p(2xz + 2yz + xy - 12)^2 \quad (19)$$

onde $r_p = 10^3$.

Solução numérica por meio do método das Coordenadas Cíclicas: No método das Coordenadas Cíclicas a direção de busca é dada pelos eixos coordenados $d_1 = (1; 0; 0)$, $d_2 = (0; 1; 0)$ e $d_3 = (0; 0; 1)$. O comprimento do passo α^* em cada direção foi obtido pelo Método da Seção Áurea com intervalo de busca $I = [-1; 1]$ e critério de parada que o intervalo final seja reduzido a 1% do intervalo inicial. As primeiras simulações foram realizadas com ponto inicial (3; 3; 3).

No entanto, os resultados obtidos com o Método das Coordenadas Cíclicas com este ponto inicial não foram bons, nas primeiras iterações o algoritmo estagnou encontrando $(x, y, z)^* = (0,9969; 1,6030; 1,9999)$ que resulta no volume $V = 3,1959\text{m}^3$, que está longe do volume ótimo.

Como o método não apresentou o resultado esperado, uma nova simulação foi realizada partindo do ponto (1; 1; 1). Os valores encontrados para este caso são muito satisfatórios visto que $(x, y, z)^* = (1,9962; 2,0062; 0,9969)$ donde o volume é $V = 3,9922\text{m}^3$. Os resultados obtidos para este caso são apresentados na Tab. 1.

Tabela 1. Solução Numérica por meio do Método das Coordenadas Cíclicas.

Iteração k	$(x^k)^T$ $F(x^k)$ $V(x^k)$	j	d_j^T	$(y^j)^T$	α_j	$(y^{j+1})^T$
1	(1; 1; 1) 3,3951 1	1	(1; 0; 0)	(1; 1; 1)	1	(2; 1; 1)
		2	(0; 1; 0)	(2; 1; 1)	1	(2; 2; 1)
		3	(0; 0; 1)	(2; 2; 1)	0,003105	(2; 2; 1,0031)
2	(2; 2; 1,0031) 3,7555 4,0124	1	(1; 0; 0)	(2; 2; 1,0031)	-0,0069	(1,9930; 2; 1,0031)
		2	(0; 1; 0)	(1,9930; 2; 1,0031)	0,0031	(1,9930; 2,0031; 1,0031)
		3	(0; 0; 1)	(1,9930; 2,0031; 1,0031)	-0,0031	(1,9930; 2,0031; 0,9999)
3	(1,9930; 2,0031; 0,9999) 3,7547 3,9918	1	(1; 0; 0)	(1,9930; 2,0031; 0,9999)	0,0031	(1,9961; 2,0031; 0,9999)
		2	(0; 1; 0)	(1,9962; 2,0031; 0,9999)	0,0031	(1,9962; 2,0062; 0,9999)
		3	(0; 0; 1)	(1,9961; 2,0062; 0,9999)	-0,0031	(1,9961; 2,0062; 0,9969)
4	(1,9962; 2,0062; 0,9969) 3,7521 3,9923	1	(1; 0; 0)	(1,9962; 2,0062; 0,9969)	0,0031	(1,9993; 2,0062; 0,9969)
		2	(0; 1; 0)	(1,9993; 2,0062; 0,9969)	0,0031	(1,9993; 2,0093; 0,9969)
		3	(0; 0; 1)	(1,9993; 2,0093; 0,9969)	-0,0031	(1,9993; 2,0093; 0,9938)

Observe que no início do processo iterativo os valores de F e V estão distantes. No entanto, depois de algumas iteração estes valores se tornam próximos o que mostra que as soluções ótimas estão perto de satisfazer a restrição $g(x, y, z) = 2xz + 2yz + xy = 12$. Por exemplo, $g(x^4) = g(1,9962; 2,0062; 0,9969) = 11,98$.

Pode-se citar como vantagens do método das Coordenadas Cíclicas o baixo esforço computacional e o fato de não usar derivadas.

Solução numérica por meio do método de Newton: A direção de busca de Newton é dada pela equação (7). As equações (20) e (21) apresentam, respectivamente, o vetor gradiente e a matriz Hessiana deste problema que são necessários para o cálculo da direção. A fim de simplificar as expressões do gradiente e da Hessiana nas equações (20) e (21), considere $A = 2xz + 2yz + xy - 12$.

$$\nabla F(x) = \begin{pmatrix} yz - 2r_p(2z + y)A \\ xz - 2r_p(2z + x)A \\ xy - 2r_p(2x + 2y)A \end{pmatrix} \quad (20)$$

e

$$\nabla^2 F(x) = \begin{pmatrix} -r_p(8z^2 + 8zy + 2y^2) & z - r_p(8z^2 + 4zy + 4zx + 2xy + 2A) & y - r_p(8xz + 4xy + 8yz + 4y^2 + 4A) \\ -r_p(8z^2 + 4zy + 4zx + 2xy) & -r_p(8z^2 + 8zx + 2x^2) & x - r_p(8zx + 4x^2 + 8yz + 4xy + 4A) \\ y - r_p(8xz + 4xy + 8yz + 4y^2 + 4A) & x - r_p(8zx + 4x^2 + 8yz + 4xy + 4A) & -r_p(8x^2 + 16xy + 8y^2) \end{pmatrix} \quad (21)$$

Novamente, o comprimento do passo α^* foi obtido pelo Método da Seção Áurea. Nas primeiras simulações com o método de Newton adotou-se o ponto inicial (1; 1; 1), o mesmo adotado no método das Coordenadas Cíclicas, obtendo após 500 iterações a solução $(x, y, z)^* = (1,633369; 1,633369; 1,427940)$ cujo volume da caixa é aproximadamente 3,8095. A fim de encontrar uma solução melhor, foram feitas novas simulações com outros pontos iniciais. A Tab. 2 apresenta os valores encontrados partindo do ponto inicial (3; 3; 3).

De acordo com a Tab. 2 é possível observar que o Método de Newton foi eficiente na busca pelo ótimo, principalmente quando o processo teve início no ponto (3; 3; 3). Embora o algoritmo tenha mostrado-se lento, devido ao comprimento do passo α que é muito pequeno, na 7399ª iteração o progresso foi interrompido com o critério de parada

Tabela 2. Solução Numérica por meio do Método de Newton.

Iteração k	$(x^k)^T$	$F(x^k)$	α_k	p_k^T	$\ \nabla F(x^k)\ $
1	(3; 3; 3)	1.088.973	0,0027	(-0,8048; -0,8048; -0,8048)	$1,1545 \times 10^6$
2	(2,9977; 2,9977; 2,9977)	$1,0846 \times 10^6$	0,0027	(-0,8039; -0,8039; -0,8039)	$1,1513 \times 10^6$
3	(2,9955; 2,9955; 2,9955)	$1,0802 \times 10^6$	0,0027	(-0,8030; -0,8030; -0,8030)	$1,1482 \times 10^6$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
10	(2,9802; 2,9802; 2,9802)	$1,0503 \times 10^6$	0,002739	(-0,7967; -0,7967; -0,7967)	$1,1263 \times 10^6$
100	(2,7934; 2,7934; 2,7934)	$7,2986 \times 10^5$	0,002739	(-0,7184; -0,7184; -0,7184)	$8,8009 \times 10^5$
500	(2,1726; 2,1726; 2,1726)	$1,3460 \times 10^5$	0,002739	(-0,4286; -0,4286; -0,4286)	$2,9397 \times 10^5$
1000	(1,7675; 1,7675; 1,7676)	$1,3115 \times 10^4$	0,002737	(-0,1836; -0,1836; -0,1836)	$7,4664 \times 10^4$
2500	(1,5526; 1,5526; 1,5549)	0,8627	0,002735	(-0,0052; -0,0052; -0,0030)	$1,2260 \times 10^3$
5000	(1,7663; 1,7663; 1,2557)	3,8588	0,002735	(0,0058; 0,0058; -0,0060)	$1,4733 \times 10^2$
7399	(1,9629; 1,9629; 1,0375)	3,9965	0,002739	(0,0062; 0,0062; -0,0062)	23,5679

$\|F(x^k) - F(x^{k-1})\| < 10^{-5}$, obtendo solução ótima $(x, y, z)^* = (1, 9629; 1, 9629; 1, 0375)$ no qual o volume da caixa é, aproximadamente, igual a $3,9974\text{m}^3$. Note que devido a penalização dos pontos que não pertencem à região viável, nas primeiras iterações o valor da função objetivo F é muito alto mas após alguns passos este valor começa diminuir e o algoritmo converge para a solução ótima.

7. CONCLUSÃO

Os resultados encontrados com os algoritmos dos métodos das Coordenadas Cíclicas e de Newton são muito satisfatórios visto que o valor da solução ótima encontrada é muito próximo à solução analítica.

No problema proposto, a dificuldade de convergência dos métodos das Coordenadas Cíclicas e principalmente, de Newton, é devido a restrição de igualdade que é imposta ao problema de maximizar o volume da caixa sem tampa. Além da região viável ficar muito limitada, o problema original precisa ser transformado em um problema irrestrito para então ser resolvido pelos métodos de otimização irrestrita. Além disso, no caso do método de Newton os valores encontrados para o comprimento de passo α são da ordem de 10^{-2} , o que torna o algoritmo muito lento.

Visto que não existe um algoritmo universal, a escolha do método de otimização e do ponto inicial podem determinar a eficácia ou falha para a solução do problema.

Portanto, pretende-se para trabalhos futuros, utilizar outros algoritmos de otimização irrestrita, como por exemplo o de Hooke e Jeeves e Quase Newton, a fim de resolver o mesmo problema e comparar as soluções obtidas.

8. AGRADECIMENTOS

Ao Programa de Bolsas de Graduação da Universidade Federal de Uberlândia pelo auxílio concedidos durante a realização deste trabalho.

REFERÊNCIAS

- Bazaraa, S.M., Sherali, H.D., Shetty, C.M., 2006. *Nonlinear Programming Theory and Algorithms*, 3rd ed., John Wiley & Sons, Atlanta.
- Brandão, M. A. L. 2010. *Estudo de alguns métodos determinísticos de otimização irrestrita* Dissertação de Mestrado, Universidade Federal de Uberlândia, (pp 101). Uberlândia, Brasil.
- Camponogara, E. 2006. *Métodos de otimização: Teoria e Prática* Universidade Federal de Santa Catarina, Departamento de Automação e Sistemas, Florianópolis, Brasil.
- Dasgupta, D., Le Riche, R., Michalewicz, Z. e Schönhauer, M. 1996. *Evolutionary Algorithms for Constrained Engineering Problems* Computers e Industrial Engineering Journal
- Hartman, H. L. ; Mutmanský, J. M. 2002. *Introductory Mining Engineering* New Jersey: John Wiley e Sons, second edition.
- Nocedal, J., Wright, S.J. 2000. *Numerical Optimization*, Springer Series in Operations Research
- Stewart, J. 2007. *Cálculo*, Vol. 2. 5ed. São Paulo: Thomson Learning
- Takahashi, R. 2003. *Otimização Escalar e Vetorial* Universidade Federal de Minas Gerais, Departamento de Matemática, Curso de Verão.

RESPONSABILIDADE AUTURAL

“Os autores são os únicos responsáveis pelo conteúdo deste trabalho”.

SOME CLASSICAL OPTIMIZATION METHODS

Pedro Henrique Rodrigues da Silva, pedro.henrique-02@hotmail.com¹

Milena Almeida Leite Brandão, milena@pontal.ufu.br¹

¹Universidade Federal de Uberlândia, Faculdade de Ciências Integradas do Pontal, Ituiutaba - MG - Brasil

Abstract. *In general terms, optimizing means to plan or develop more efficiently. After the formulation of the model, an optimization method is used to obtain the solution of the problem. We can divide the optimization methods among those dependent or not the knowledge derived. The optimization methods that do not depend on the use of derivatives generally using only information of the function to be optimized, which may be discontinuous and multimodal. The major advantage of these methods is the great practical potential due to the increasing need to solve optimization problems defined by functions for which there are no derived or these are available at a very high computational cost. In optimization methods that dependent upon use of derivatives the model is completely known and are for continuous, convex and semi-modal functions. In this study, two optimization methods are presented: the Method of Cyclic Coordinates, which does not depend on derivatives it uses the direction of the coordinate axes as search directions and the Newton Method whose search direction using the derivatives of the function to be optimized. Then a mathematical modeling problem will be solved by both methods and the results will be analyzed.*

Keywords: *Deterministic Optimization, Method of Cyclic Coordinates, Newton's Method.*