Otimização teórica e computacional aplicada à resolução de problemas restritos

Theoretical and computational optimization applied to the resolution of restricted problems

Helena Sumiko Yakuda (orientada)*, Tatiane Cazarin Da Silva (orientadora)†

RESUMO

Problemas de otimização matemática possuem aplicações nas mais diversas áreas tal como economia, estatística, medicina, biologia, localização, transportes e engenharias. Tais problemas exigem uma solução precisa, o que pode aumentar a sua complexidade e as variáveis de projeto. Sendo assim, esses problemas podem ser classificados como problemas de otimização irrestrita ou restrita. Para a convergência desses problemas geralmente fazem do uso de códigos computacionais, não se tratando apenas de um algoritmo universal, mas sim de um conjunto de métodos que se adaptam a diferentes situações, ao fazer a escolha pode-se então determinar a eficácia ou a falha na solução. No presente estudo estão dispostos métodos de otimização não-linear irrestrito e restrito. Para o método irrestrito, é abordado a respeito da escolha do tamanho do passo e a escolha do método de descida, e qual a sua influência diante do algoritmo; já para o método restrito serão apresentados meios de converter um problema restrito para um irrestrito.

Palavras-chave: otimização, irrestrita, restrita, métodos de otimização

ABSTRACT

Math optimization problems have applications in several areas, such as economics, statistics, medicine, biology, location, transport, and engineering. Such problems require a precise solution what can increase your complexity and design variables. In general, these problems can be classified as unconstrained and constrained optimization problems. For the convergence of these problems they usually use computational codes, not just a universal algorithm, but a set of methods that adapt to different situations, by making the choice one can then determine the effectiveness or failure of the solution. In the present study, unrestricted and constrained nonlinear optimization methods are arranged. For the unrestricted method, it is discussed about the choice of step size and the choice of the descent method and what is its influence on the algorithm; for the restricted method, means of converting a restricted problem to an unrestricted one will be presented.

Keywords: optimization, unrestricted, restricted, optimization methods

1 INTRODUÇÃO

A otimização matemática aplicada à análise de um problema pode variar muito em função das características do sistema. Por exemplo, as variáveis de projeto, ou também denominada de variáveis de decisão, podem estar restritas por condições impostas pelo problema analisado. Assim, a modelagem do problema é a etapa da identificação dos objetivos, variáveis e restrições, e considerada o passo mais importante no processo (BRANDÃO, 2010).

* Engenharia Civil, Universidade Tecnológica Federal do Paraná, Campo Mourão, Paraná, Brasil; helenayakuda@alunos.utfpr.edu.br
† Universidade Tecnológica Federal do Paraná, Campus Campo Mourão; tatianecazarin@utfpr.edu.br

Para a análise do problema, torna-se fundamental estabelecer conceitos teóricos associados às funções. Assim como (FRIEDLANDER, 1994), consideremos a hipótese de que os resultados das funções de uma variável apresentem $f'(x^*) = 0$ e $f''(x^*) \geq 0$, sendo que $f: \mathbb{R} \to \mathbb{R}$, $f \in C^1$ e $f \in C^2$, e x^* um minimizador de f em $\mathbb{R}$. Temos algumas considerações a serem feitas para o uso do método irrestrito.

Preposição 1. (Condições necessária de 1ª ordem): Seja $f: \mathbb{R}^n \to \mathbb{R}$, $f \in C^1$. Se x^* é um minimizador local de f em $\mathbb{R}^n$, então $\nabla f(x^*) = 0$.

Preposição 2. (Condições necessárias de segunda ordem): Seja $f: \mathbb{R}^n \to \mathbb{R}$, $f \in C^2$. Se x^* é um minimizador local de f em $\mathbb{R}^n$, então $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ é semidefinida positiva, ou seja $d^T \nabla^2 f(x^*) d \geq 0$, para qualquer vetor $d \in \mathbb{R}^n$, não nulo.

Preposição 3. (Condições suficientes de segunda ordem): Seja $f: \mathbb{R}^n \to \mathbb{R}$, $f \in C^2$. Se $x^* \in \mathbb{R}^n$, $\nabla f(x^*) = 0$, e $\nabla^2 f(x^2) > 0$, então x^* é um minimizador local estrito de f em $\mathbb{R}^n$.

O seguinte teorema justifica a importância de a função ser convexa.

Teorema 1. Sejam $C \subset \mathbb{R}^n$ convexo e $f: C \to \mathbb{R}$ uma função convexa. Se $x^* \in C$ é minimizador local de f, então x^* é minimizador global de f.

A partir do conhecimento teórico são propostos métodos numéricos que, por meio de um processo iterativo, gera uma sequência de valores que atendem as condições citadas e garantem uma maior precisão nos resultados. O objetivo é: a partir de um ponto inicial x^0, o algoritmo irá gerar uma série de aproximações, até que a função objetivo decresça até atingir um ponto ótimo, isso gerará uma sequência $x^k \subset \mathbb{R}^n$. O algoritmo só encerrará quando atender a condição imposta: $\nabla f(x) = 0$. Para gerar tal sequência os algoritmos escolhem, a partir de cada ponto obtido uma melhor direção d para dar o seguinte passo, uma opção é na qual f decresce. Garantimos que d é de descida quando, para cada x^*, existe $\delta > 0$ tal que $f(\bar{x} + td) < f(\bar{x})$, para todo $t \in (0, \delta)$. Uma condição suficiente para garantir que d seja de descida é que $\nabla f(\bar{x})^T\, d < 0$ (RIBEIRO e KARAS, 2011). A seguir é apresentado um algoritmo básico de descida para minimizar uma função f.

Algoritmo 1. Algoritmo de descida
Dado $x^0 \in \mathbb{R}^n$
$k = 0$
REPITA enquanto $\nabla f(x) \neq 0$
    Calcule d^k tal que $\nabla f(x^k)^T dk < 0$
    Escolha $t_k > 0$ tal que $f(x^k + t_k d^k) < f(x^k)$
    Faça $x^{k+1} = x^k + t_k d^k$
    $k = k + 1$

O Algoritmo 1 encontra um ponto estacionário em um número finito de iterações ou gera uma sequência ao longo da qual f decresce. A cada iteração, os algoritmos de descida buscam minimizar $f(x^k)$ na direção $f(x^k + t_k d^k)$, sujeita a $t \geq 0$. Esse processo é chamado de busca linear exata. Existem diversos métodos para a escolha da direção d^k, são eles: Método do Gradiente, Newton, Quase-Newton e Gradientes Conjugados; e para o tamanho do passo, t_k além da Busca Exata temos a Seção Áurea e a Condição de Armijo (RINCÃO, 2008). Tais métodos, aplicados na minimização de funções irrestritas, serão discutidos a seguir; assim como os métodos aplicados em minimização restrita.

2 MÉTODOS

Para a escolha da direção d^k proposta pelo algoritmo 1, são apresentados a seguir os principais métodos da literatura.

- *Método do Gradiente:* A função objetivo necessita avançar na direção que fornece o menor valor da derivada direcional, seja ela: $d^k = -\nabla f(x^k)$.

- *Método de Newton:* Neste método a cada iteração, calcula-se o gradiente da função e também a matriz hessiana, o que pode consumir muito tempo de computação ou ainda, muitas vezes, ser muito complexa a sua obtenção, a direção é dada por: $d_k = (\nabla^2 f(x^k))^{-1}\, \nabla f\left(x^k\right)$.

- *Método das direções conjugadas:* Este método possui convergência mais rápida que o método de Cauchy e um custo computacional inferior que ao de Newton. A direção é dada por: $d^k = -\nabla f(x^{k+1}) + \beta_k d^k$, onde $\beta_k = \frac{(d^k)^T A \nabla f(x^{k+1})}{(d^k)^T A d^k}$. Há mais duas variantes para o uso de β_k, uma é proposta por Polak and Ribière e outra por Fletcher e Reeves (RINCÃO, 2008).

- *Método Quase-Newton:* Esse método tem uma melhor performance que Cauchy e é mais barato computacionalmente que Newton. Nele é feito uma aproximação para a Hessiana da função objetivo ao longo das iterações, temos a direção dada por: $d^k = -H_k \nabla f(x^k)$, onde $H_k \in \mathbb{R}^{nxn}$ é definida positiva. Há duas variações para o método Quase-Newton, que se diferem no cálculo de H_k, são eles o Método DFP, proposta por Davindon, Fletcher e Powell e o Método BFGS pertencente aos Broyden, Fletcher, Goldfarb e Shanno (BFGS), que possui um desempenho computacional superior ao de DFP.

Para a escolha do tamanho do passo t_k proposta pelo algoritmo 1, são apresentados a seguir os principais métodos da literatura.

- *Busca Linear exata:* Busca-se minimizar a função $\varphi(t) = f(x^k + t_k d^k)$, fazendo $\nabla\varphi(t_k) = 0$. Para o caso em que f é uma função quadrática, definida por $f(x) = \frac{1}{2}x^T Ax + b^T x + c$, com A simétrica definida positiva, a fórmula do passo é dada por: $t_k = \frac{\nabla f(x^k)^T \nabla f(x^k)}{\nabla f(x^k)^T A \nabla f(x^k)}$.

- *Seção Áurea:* Neste método utiliza-se a estratégia de dividir o segmento $[a, b]$ na razão área, conhecida como número de ouro, e descartar uma das partes. A vantagem é que ao descartar uma dessas áreas é descartado mais de 38% ao invés de 33,33% no caso se a divisão for feita em três partes iguais (RIBEIRO e KARAS, 2011).

- *Condição de Armijo:* Trata-se de um método inexato que procura uma boa redução da função ao longo da direção, porém sem tentativa de minimização, sendo medida pela seguinte inequação: $f\left(x^k + td^k\right) \leq f\left(x^k\right) + \eta t(\nabla f^k)^T d^k$. São necessários os seguintes dados: $x_0 \in \mathbb{R}^n$, $d \in \mathbb{R}^n$ e γ, $\eta \in (0,1)$ arbitrários. É importante que o tamanho do passo t não seja muito pequeno. E uma forma de evitar é iniciar com $t = 1$, e caso necessário, reduzi-lo até que a inequação esteja satisfeita. É feito o teste $t = \gamma t$, até que $f(x^k + t_k d^k) > f\left(x^k\right) + \eta t(\nabla f^k)^T d^k$, seja atingida.

Ao impor alguma restrição à função objetivo fará com que o nível de dificuldade do problema aumente à medida que cresce o grau de não linearidade (FRITZSCHE, 1978). A resolução para um problema não linear restrito parte de dois aspectos: (a) transformar um problema restrito em um não irrestrito, sendo possível resolvê-lo através dos métodos já apresentados; e (b) estender os conceitos de programação linear a não linear, onde por meio de aproximações lineares sucessivas substitui a função do problema não linear original por um linear análogo (FORMIGA, 1999). Há inúmeros métodos para a busca da solução de problemas restritos, a seguir apresentaremos alguns deles.

- *Método de Penalidades:* Trata-se de um método indireto para otimização restrita, onde é composto pela transformação do problema restrito original num problema irrestrito equivalente. Os métodos de penalidade podem ser divididos em:

(a) Método de penalidade exterior ou métodos de Penalidade: neste caso o método "pune" qualquer violação as restrições, gerando uma sequência de pontos podendo estar ou não presentes na região factível (FORMIGA, 1999).

(b) Método de penalidade interior ou métodos de barreira: geram os seus pontos dentro da região factível do problema. Dessa forma adiciona a função objetivo uma penalidade que favorece os pontos interiores próximos da fronteira da região factível (FORMIGA, 1999).

- *Método do Lagrangeano aumentado:* Proposto por Hestenes (1969) e Powell (1969), descende do método de penalidade quadrática, mas para não ocorrer o mau condicionamento comum desse método, consideram as estimativas dos multiplicadores de Lagrange. O método do Lagrangeano aumentado preserva a suavidade da função (OLIVEIRA, 2012).

3 RESULTADOS

Os métodos teóricos apresentados foram implementados em Matlab, e aplicados a análise de funções escolhidas. A Tabela 1 traz os resultados para uma série de testes com diferentes funções objetivos, assim como o ponto inicial, x^0; o número de iterações, k; e o tempo médio $t_{méd}$, em segundos, de três execuções dos algoritmos. Algumas informações importantes: Para a Condição de Armijo considerou-se $\gamma = 0{,}95$, $\eta = 0{,}5$ e $t = 1$, onde há testes utilizando a busca exata; No método da seção áurea utilizou os parâmetros $\varepsilon = 0{,}1$ e $\rho = 0{,}5$; Como critério da parada $\|\nabla f(x^k)\| < 10^{-4}$; Pontos iniciais (1,1) para a função $f_1(x_1, x_2) = 3x_1^2 + 7x_2^2 + 4x_1x_2 + 8x_1 - 2x_2 + 11$ e (2,1) para a função $f_2(x_1, x_2) = x_1^2 + 8x_2^2 + 3$.

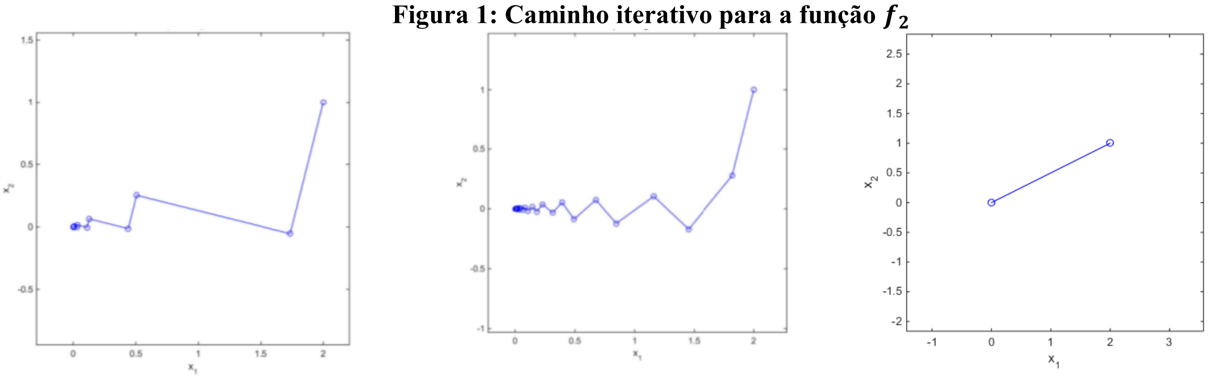
Tabela 1. Resultados obtidos da execução das funções nos diferentes métodos

Função	Passo / Método	Busca Linear		Seção Áurea		Condição de Armijo	
		k	**Tempo**	**k**	**Tempo**	**k**	**Tempo**
f_1	Gradiente	14	1,232	43	9,952	13	28,811
	Newton	132	33,135	3	2,842	1	0,343
	Gradiente Conjugado - β_k *original*	2	0,398	19	11,191	6	17,167
	Quase Newton (DFP)	2	0,438	4	2,857	4	6,622
	Quase Newton (BFGS)	2	0,699	5	5,321	4	6,522
f_2	Gradiente	13	1,311	35	16,629	8	19,878
	Newton	143	40,023	3	2,216	1	0,353
	Gradiente Conjugado - β_k *original*	2	0,424	15	10,130	6	15,248
	Quase Newton (DFP)	2	0,391	4	2,755	5	5,778
	Quase Newton (BFGS)	2	0,640	4	2,351	5	6,364

Fonte: Autoria Própria (2021)

É possível notar que o número de iterações não é o fator que garante a rapidez da convergência, por isso é importante levar em consideração fatores como que algumas funções garantem a convergência quadrática em n passos, como no caso do método de direções conjugadas. Para melhor facilitar o entendimento apresentaremos um comparativo gráfico com a função f_2.

Figura 1: Caminho iterativo para a função f_2



(a) Gradiente - Busca linear exata **(b) Gradiente - Seção Áurea** **(c) Newton - Condição de Armijo**

Fonte: Autoria Própria (2021)

Podemos notar que as 13 iterações, para busca linear exata, formam ângulos ortogonais entre si (Figura 1(a)) e apresenta a sua convergência com $t_{méd} = 1{,}311921$s, diferente do caso com o uso do passo de Seção Áurea, $t_{méd} = 16{,}62952367$s (Figura 1(b)). Neste caso foram utilizadas 35 iterações, e ao usar esse tipo de passo, já não ocorre a ortogonalidade entre eles.

O método de Cauchy, por se tratar de um método exato, a sua convergência pode ser lento, porém os seus resultados são precisos, e também por ser uma função quadrática podemos garantir a sua convergência. Mas podemos destacar também que o melhor método para a resolução de funções quadráticas é com a utilização do método de Newton com passo de Armijo, que apesar do alto custo computacional alcança a convergência em apenas 1 iteração (Figura 1(c)).

A fim de analisar numericamente também um problema de minimização restrita, consideramos o problema teórico:

$$\min f(x_1, x_2) = 3x_1 + 4x_2^2$$
$$s.a. \begin{cases} 2x_1 + 2x_2 = 3 \\ 5x_1 = 4 \\ 2x_1 \geq 3x_2 \\ 4x_1 - x_2 \leq 10 \\ x_1^2 + 2x_2 = 10 \\ x_1 x_2 + x_2^2 \leq 20 \\ 0 \leq x_1 \leq 10 \\ 1 \leq x_2 \leq 4 \end{cases}$$

O qual foi implementado no software Matlab e resolvido com o auxílio da rotina *fmincon*. Como ponto inicial utilizado foi $x = (0{,}0)$. A solução ótima gerada pelo algoritmo foi $f(1{,}2743;\ 1{,}3920) = 11{,}5739$, em um tempo de $t = 0{,}159340$ segundos. Podemos notar que apesar de possuir condições lineares e não lineares, um intervalo definido a rotina executou em um curto tempo.

Em muitos problemas da Engenharia podem ser utilizadas a rotina *fmincon* para encontrar a solução ótimo. Como exemplo o dimensionamento de estruturas que podem ter em conta como restrição o custo, o peso, a área da seção transversal, ou qualquer outro (SIAS e ALVES, 2015).

4 CONCLUSÃO

O presente estudo propôs importantes tópicos relacionados aos diferentes métodos para a busca de minimizadores. Em Engenharia Civil, por exemplo, tais métodos podem ser aplicados em diversas análises, tais como o dimensionamento de estruturas, trazendo benefício comprovado na busca de excelentes resultados.

A otimização aplica-se em inúmeras situações ou problemas em que se deseja melhorar e obter um desempenho máximo. A eficiência dos resultados pode ser alcançada de forma mais rápida por meio de um estudo prévio, o que permite a resolução dos problemas de forma numérica e de forma otimizada. Sendo assim, a análise teórica e numérica de tais problemas são necessárias para garantir um melhor desempenho numérico alcançado.

Para isso, o estudo dos métodos teóricos aliados à implementação numérica, permitem estabelecer parâmetros que identificam o melhor método para cada problema. Ampliando ao cenário de testes, podemos identificar classes que os tornam mais aplicáveis.

REFERÊNCIAS

FORMIGA, K. T. M. **METODOLOGIA DE OTIMIZAÇÃO DE REDES MALHADAS ATRAVÉS DA PROGRAMAÇÃO NÃO LINEAR**. 1999. 158 f. Dissertação (Mestrado em Engenharia Civil) – Centro de Ciências e Tecnologia, Universidade Federal da Paraíba, Campina Grande, 1999.

FRIEDLANDER, A. **Elementos de Programação Não Linear**. UNICAMP. 1994. Disponível em: < http://www.ime.unicamp.br/~friedlan/livro.pdf>. Acesso em 10 de junho 2021.

FRITZSCHE, H. **Programação Não Linear; Análise e Métodos**. São Paulo: Universidade de São Paulo, 1978.

OLIVEIRA, F. R. **ESTUDO DE ALGUNS MÉTODOS CLÁSSICOS DE OTIMIZAÇÃO RESTRITA NÃO LINEAR**. 2012. 120 F. Dissertação (Mestrado em Matemática) - Universidade Federal de Uberlândia, Uberlândia, 2012. Disponível em: https://repositorio.ufu.br/bitstream/123456789/16795/1/d.pdf. Acesso em: 10 de agosto de 2021.

RIBEIRO, A. A.; KARAS, E. W. **UM CURSO DE OTIMIZAÇÃO**. Curitiba, 2011.

RINCÃO, T. **Otimização Irrestrita sem Derivadas Baseada em Interpolação Polinomial**. 2008. 63f. Dissertação (Mestrado em Matemática Aplicada e Computacional) – Universidade Estadual de Campinas, Campinas, 2008. Disponível em: http://repositorio.unicamp.br/jspui/bitstream/ REPOSIP/ 306044/ 1 /Rincao_ Thiago_M.pdf. Acesso em: 10 de Agosto de 2021.

SIAS, F. M.; ALVES, E. C. Dimensionamento Ótimo de Pilares de Concreto Armado Segundo a NBR 611:2014. **ABPE - Associação Brasileira de Pontes e Estruturas**, v. 14, n. 2, p. 46-57. jul/dez. 2015. Disponível em: http://www.abperevista.com.br/imagens/volume15_02/cap06.pdf. Acesso em 30 de Agosto de 2021.