

Métodos de Descida Máxima e de Newton aplicados em minimização irrestrita

Maximum Descent and Newton methods applied in unrestricted minimization

RESUMO

Caroline Alves Batista
caroline.alves.b7@gmail.com
Universidade Tecnológica Federal do Paraná, Campo Mourão, Paraná, Brasil

Leandro Waidemam
waidemam@utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Campo Mourão, Paraná, Brasil

Tatiane Cazarin da Silva
tatianecazarin@utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Campo Mourão, Paraná, Brasil

O objetivo deste trabalho foi realizar um estudo numérico comparativo, em relação ao tempo computacional e número de iterações, entre os métodos de Descida Máxima e de Newton aplicados a alguns problemas irrestritos. A implementação dos algoritmos foi realizada com o auxílio do *software* Matlab, e como tamanhos de passo foram utilizadas a busca linear exata e a condição de Armijo, respectivamente, para os métodos de Descida Máxima e Newton. Verificou-se, como esperado pela literatura, que o método de Newton se sobressai em eficiência para a maioria dos problemas, entretanto em outros pode obter um alto número de iterações, tendo em vista a relação direta com escolha do ponto inicial, que influencia diretamente no seu raio de convergência. O método de Descida Máxima, por sua vez, clássico na resolução de minimização de funções quadráticas, tem a convergência diretamente associada às características do problema, podendo ser extremamente lento na prática.

PALAVRAS-CHAVE: Otimização irrestrita. Método de Descida Máxima. Método de Newton.

Recebido: 19 ago. 2019.

Aprovado: 01 out. 2019.

Direito autorial: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



ABSTRACT

The objective of this work was to perform a comparative numerical study, regarding computational time and number of iterations, between the Maximum Descent and Newton methods applied to some unrestricted problems. The implementation of the algorithms was performed with the aid of Matlab software, and as step sizes were used the exact linear search and the Armijo condition, respectively, for the Maximum Descent and Newton methods. As expected from the literature, it was found that Newton's method excels in efficiency for most problems, but in others it can obtain a high number of iterations, given the direct relationship with choice of the starting point, which directly influences in its radius of convergence. The method of Maximum Descent, in turn, classical in the resolution of minimization of quadratic functions, has the convergence directly associated with the characteristics of the problem and can be extremely slow in practice.

KEYWORDS: Unrestricted optimization. Maximum Descent method. Newton method.

INTRODUÇÃO

A otimização, de acordo com Avriel (1976), é a ciência detentora de ferramentas que direcionam para a escolha da melhor decisão dentre muitas possíveis.

Dentro das vertentes que compõem a otimização, a programação não linear é utilizada nos casos onde a função objetivo e/ou o conjunto de restrições são não-lineares (Bertsekas, 1995). Sua aplicação destaca-se nos problemas das áreas de engenharia, economia, administração de empresas, análise de sistemas, ciência da computação ou em qualquer outra área onde a situação possa ser expressa por um modelo matemático (Avriel, 1976).

Segundo Martínez e Santos (1995), a minimização irrestrita de uma função contínua de n variáveis é um dos problemas clássicos da otimização não linear, que fornece a possibilidade de contornar as manipulações isoladas em casos quando o número de variáveis ou a complexidade da função aumentam, fazendo uso de métodos numéricos quase sempre iterativos.

Existem diversos métodos que se diferem pela escolha da direção de descida ou pelo método de cálculo do tamanho do passo. Assim, considerando as diferenças dos algoritmos de cada método, o objetivo desse estudo é realizar um comparativo voltado para otimização irrestrita entre os métodos de Descida Máxima (método do Gradiente) e de Newton aplicados em alguns problemas.

MATERIAL E MÉTODOS

Os algoritmos aplicados em problemas de programação linear consistem em escolher a partir de cada ponto obtido, uma nova direção para dar o próximo passo afim de que o valor da função objetivo diminua (Brandão, 2010). Tais direções são chamadas direções de descida e variam em cada método aplicado.

Estes algoritmos, dentro dos parâmetros específicos de cada método, funcionam da seguinte maneira: a partir de um ponto x^k , determina-se uma direção d^k , voltada para diminuir o valor da função objetivo, e calcula-se o tamanho do passo que possa permitir uma diminuição razoável da função (Martínez e Santos, 1995).

Para a implementação do método de Descida Máxima, utilizou-se a direção dada em (1) e o tamanho do passo definido por (2), aplicados no algoritmo.

$$d^k = -\nabla f(x^k) \quad (1)$$

$$\lambda_k = \frac{\nabla^t f(x^k) \cdot \nabla f(x^k)}{\nabla^t f(x^k) \cdot \nabla^2 f(x^k) \cdot \nabla f(x^k)} \quad (2)$$

Já para o método de Newton, fez-se o uso da direção dada em (3) e o do tamanho do passo definido pelo algoritmo da condição de Armijo, utilizando $\gamma = 0,5$ e $\eta = 0,1$, aplicados no algoritmo.

$$d^k = -\left(\nabla^2 f(x^k)\right)^{-1} \cdot \nabla f(x^k) \quad (3)$$

O algoritmo de cada método foi implementado em linguagem de programação com o auxílio do *software* matemático Matlab, para as cinco funções que constam

na Tabela 1. Para mensurar o tempo computacional, realizou-se a média de três tempos obtidos ao rodar o código do algoritmo, em cada função, para cada método.

Em ambos os modelos matemáticos, o critério de parada foi definido como $\|\nabla f(x^k)\| < 10^{-6}$ com o intuito de aumentar o tempo de obtenção dos resultados, levando em consideração que o critério já garante uma boa proximidade do ponto ótimo.

O algoritmo genérico e os algoritmos específicos de cada método podem ser encontrados em vários textos de literatura, como por exemplo, Friedlander (1994) e Gouvêa (2016).

RESULTADOS E DISCUSSÃO

A Tabela 1, a seguir, traz os resultados obtidos nas compilações e apresenta: a função objetivo utilizada; o método aplicado, sendo as opções método de Descida Máxima (ou método do Gradiente), G, e método de Newton, N; o ponto inicial $x^{(0)}$, escolhido arbitrariamente; o número de iterações realizadas pelo método, k ; e por fim, o tempo computacional gasto para compilar a função no respectivo método, dado em segundos.

Tabela 1 – Resultado do comparativo entre os métodos do Gradiente (G) e de Newton (N)

Função objetivo	Método	$x^{(0)}$	k	t(s)
$f_1: -2 + 2x_1^2 + x_2^2 - (x_1x_2)^2$	G	[1,1]	2	0,2111
	N	[1,1]	3	0,5882
$f_2: 100(x_2 - x_1^2)^2 + (1 - x_1)^2$	G	[1,2]	5	-
	N	[1,2]	1	0,1879
$f_3: 5 + x_1^2 + x_2^2$	G	[1,0]	1	0,3703
	N	[1,0]	1	0,4685
$f_4: 5 + x_1^2 + 4x_2^2$	G	[1,0]	15	3,22533
	N	[1,0]	1	0,6092
$f_5: (2x_1 - 3x_2)^2 + (x_2 - 1)^4$	G	[0,2]	7	-
	N	[0,2]	13	1,6231

Fonte: Autoria própria (2019).

Partindo para a análise dos resultados, foi possível constatar em f_3 ínfima diferença entre os métodos, que convergiram em apenas uma iteração e utilizaram quase o mesmo tempo para fazê-lo, 0,3703 e 0,4685 segundos, sendo uma função bem comportada numericamente, com curvas de nível sendo circunferências.

Na função f_1 (pouco complexa), com a utilização de um ponto inicial próximo do ótimo, a proximidade do número de iterações, 2 e 3, e do tempo computacional, 0,2111 e 0,5882 segundos, demonstrou um comportamento semelhante entre os métodos. Os resultados de Newton para f_1 podem ser explicados pelo fato de que ele converge rapidamente quando o ponto inicial está suficientemente próximo do minimizador (Martínez e Santos, 1995).

Em contrapartida, na função f_4 , quando é modificada a abertura do parabolóide em comparação a f_3 , o método do Gradiente demorou mais para convergir, utilizando 15 iterações e 3,2253 segundos, sendo inferior ao método de

Newton, que demorou 1 iteração e 0,6092 segundos. Segundo Brandão (2010), isto pode ser explicado pela complexidade da função, que torna a convergência lenta para o primeiro método, e pela característica quadrática da função, que garante a convergência para um mínimo local com apenas uma iteração para o segundo método. Vale ainda destacar que pelo fato de f_3 e f_4 serem convexas o mínimo local atingido por ambos os métodos é um mínimo global para tais funções.

O método do Gradiente não convergiu em f_2 e f_5 , demonstrando a inviabilidade da sua aplicação em funções complexas como essas, devido, segundo Gouvêa (2016), a sua demora de convergência na maioria dos problemas que apresentam funções complexas.

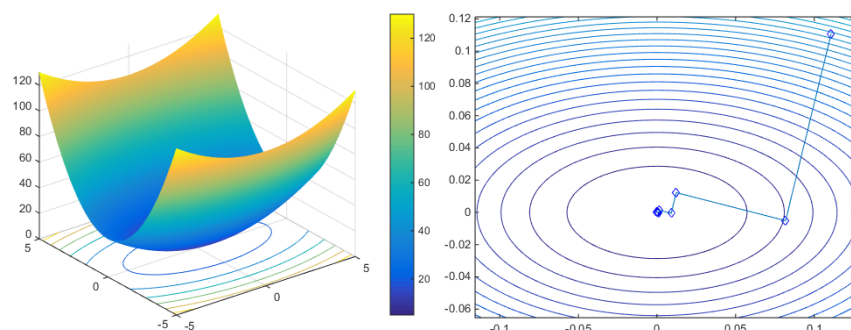
Já o método de Newton apresentou convergência em f_2 , com apenas uma iteração, e em f_5 demorou 13 iterações, elucidando a rapidez de convergência em uma função quadrática, no primeiro caso, e a demora ao utilizar um ponto inicial $x^{(0)}$ longe do ponto ótimo x^* , no segundo caso. Ainda, a necessidade de calcular a inversa da matriz hessiana da função, a cada iteração, torna o método de Newton custoso do ponto de vista computacional, de acordo com Gouvêa (2016), o que pode explicar o grande número de iterações em f_5 .

Com o intuito de realizar uma comparação ilustrativa, foram escolhidas as funções f_3 , f_4 e f_5 para apresentar uma discussão gráfica, exposta a seguir, entre as diferenças dos métodos. As funções, suas respectivas curvas de nível e o caminho percorrido pelas iterações foram plotados com o auxílio do software Matlab.

A primeira função (f_3) não apresentou distinção no número de iterações entre os métodos, onde o caminho percorrido pelas iterações de ambos é igual e linear, com apenas uma iteração.

Por outro lado, pelo fato da função f_4 ser quadrática convexa, o caminho formado pelo método do Gradiente, ilustrado por algumas iterações na Figura 1, formou direções ortogonais entre si. O fato de gerar direções ortogonais justifica a convergência lenta para alguns exemplos, já que o caminho percorrido pelo algoritmo tende a ser maior. Como o método de Newton realizou apenas uma iteração nesta função, o seu caminho é semelhante ao da função f_3 .

Figura 1 – Gráfico da função f_4 e o caminho iterativo gerado pelo método G

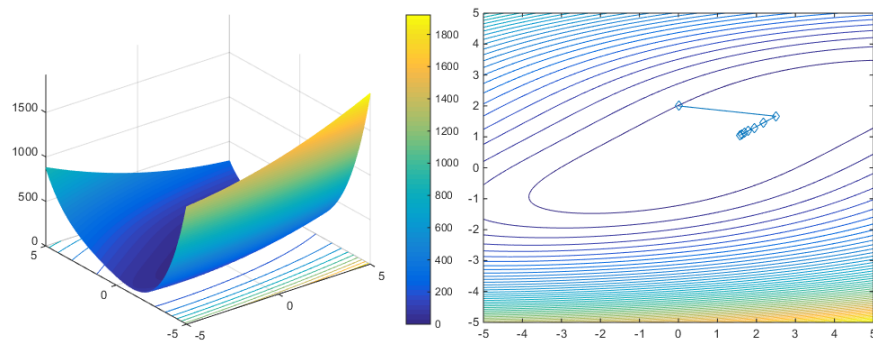


Fonte: Autoria própria (2019).

Nesta função (f_4), os resultados condizem com os estudos de Junior, Santos e Peričaro (2013), que encontraram um percurso mais curto até o minimizador do problema no método de Newton.

O caminho iterativo do método de Newton na função f_5 , demonstrado na Figura 2, não apresenta ângulos iguais, mas sim discrepantes, quando comparados ao método do Gradiente na Figura 1, pois a sua direção de descida utiliza a aproximação quadrática da função objetivo.

Figura 2 – Gráfico da função f_5 e o caminho iterativo gerado pelo método N



Fonte: Autoria própria (2019).

Os resultados possibilitaram verificar que para as funções f_2 e f_4 o método de Newton sobressaiu-se em eficiência, entretanto na função f_5 obteve alto número de iterações, evidenciando o devido cuidado que deve ser tomado com o seu ponto inicial, pois, de acordo com estudos de Nocedal e Wright (2006), seu raio de convergência não é tão grande.

No âmbito do método do Gradiente, em funções um pouco mais complexas, como f_4 , a convergência é lenta, e em funções com alto grau de complexidade, como f_2 e f_5 , a convergência pode não ser visualizada, tendo em vista os critérios de parada e lentidão do método.

Por fim, em relação aos caminhos em cada iteração, o método de Newton apresentou ângulos discrepantes, enquanto que o método do Gradiente formou uma “escada” com direções ortogonais entre si.

CONCLUSÃO

Através do estudo proposto neste artigo, foi possível identificar a aplicabilidade de cada método em diferentes problemas de minimização não linear irrestrita.

O método de Máxima Descida constitui um algoritmo essencialmente globalmente convergente, mas é muito lento na prática, onde existem problemas com funções complexas.

Por outro lado, conforme os resultados e a literatura, constatou-se que o método de Newton converge rapidamente quando parte de um ponto inicial suficientemente próximo da solução, mas seus passos, quando estão distantes de algum ponto de ótimo, podem levar a uma direção que demora para convergir ou resulta em pontos distantes da solução.

Portanto, concluiu-se que quando a função é simples, convém utilizar o método de Descida Máxima, entretanto em funções mais complexas, é mais indicado o uso do método de Newton para atingir uma possível convergência, tendo em vista uma escolha apropriada do ponto inicial do método.

REFERÊNCIAS

AVRIEL, M. **Nonlinear programming: Analysis and Methods**. New Jersey: Prentice-Hall, 1976.

BERTSEKAS, D. P. **Nonlinear programming**. Belmont: Atena Scientific, 1995.

BRANDÃO, M. A. L. **Estudo de alguns métodos determinísticos de otimização irrestrita**. 2010. Dissertação (Mestrado em Matemática) - Programa de Pós-Graduação em Matemática, Universidade Federal de Uberlândia, Uberlândia, 2010. Disponível em: <https://repositorio.ufu.br/handle/123456789/16782> Acesso em: 20 jun. 2019.

FRIEDLANDER, A. **Elementos de Programação Não Linear**. Campinas: Editora Unicamp, 1994.

GOUVÊA, É. J. C. **Métodos convergentes de otimização global baseados no vetor q-Gradiente**. 2016. Tese (Doutorado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2016. Disponível em: http://bdtd.ibict.br/vufind/Record/INPE_bec800981371e124632a1411e2314112 Acesso em: 18 jun. 2019.

JUNIOR, M. L. D. P.; SANTOS, S. R. dos.; PERIÇARO, G. A. Métodos de Newton e Quase-Newton para otimização irrestrita. In: VIII Encontro de Produção Científica e Tecnológica - EPCT, 2013, Campo Mourão. **O método científico**. Campo Mourão: FECILCAM, 2013. Disponível em: http://www.fecilcam.br/nupem/anais_viii_epct/trabalhos-01.html Acesso em: 23 jun. 2019.

MARTINEZ, J. M.; SANTOS, S. A. **Métodos computacionais de otimização**. Departamento de Matemática Aplicada, IMECC - UNICAMP, 1998.

NOCEDAL, J.; WRIGHT, S. J. **Numerical optimization**. New York: Springer-Verlag New York, Inc, 2006.

RIBEIRO, A. A.; KARAS, E. W. **Otimização Contínua: Aspectos teóricos e computacionais**. Curitiba: Cengage Learning, 2014.