Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 9, n. 1, 2022.

Trabalho apresentado no XLI CNMAC, Unicamp - Campinas - SP, 2022.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics
Preprint

Otimização Irrestrita por métodos computacionais-Método do Gradiente

Elvis D. S. Rodrigues[1] Edfram Rodrigues Pereira[2]
CSTB/UEA, AM

A otimização trata-se de um área da matemática cujo objetivo é encontrar a melhor maneira de realizar algo. Muitos problemas podem ser modelados por um problema de otimização, deste modo saber resolvê-lo tem importância fundamental para a sociedade. Este problema consiste em encontrar os máximos ou mínimos de uma função de várias variáveis considerando uma determinada região de escolha de entradas (valores dentro do espaço multidimensional). Dessa forma, seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, minimizá-la é achar um valor mínimo para esta considerando um conjunto, $D \subset \mathbb{R}^n$, chamado viável. O problema é descrito como: $min \quad f(x) \quad$ sujeito a $\quad x \in D$, esta função f é dita objetivo e os pontos de D são chamados pontos viáveis.

Um ponto $\bar{x} \in D$ será minimizador global se:

$$f(\bar{x}) \leq f(x) \quad \forall x \in D \tag{1}$$

Um ponto é dito mínimo local se

$$f(\bar{x}) \leq f(x) \quad \forall x \in V \cap D \tag{2}$$

sendo V uma vizinhança de $\bar{x}$. Como neste trabalho trataremos de otimização sem restrições, consideramos a partir deste ponto $D = \mathbb{R}^n$.

O método iterativo do gradiente vem para facilitar o processo de busca pela solução do problema. Este método consiste em encontra r um ponto x^{k+1}, dada uma aproximção x^k da solução, de forma que se obtenha $f(x^{k+1})$ um valor menor que $f(x^k)$. Para isso é tomada como direção $-f'(x^k) \in \mathbb{R}^n$ e um comprimento de passo α_k estritamente positivo tal que x^{k+1} seja igual a $x^k - \alpha_k f'(x^k)$ (Izmailov e Solodov [2]A seguir serão apresentados dois teoremas, condição necessária de primeira ordem e condição necessária de segunda ordem, respectivamente, que um ponto deve satisfazer para ser minimizador local de um problema. Os teoremas a seguir e suas demonstrações são encontradas em Izmailov e Solodov [1].

Teorema 1. Seja a função $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável no ponto $\bar{x} \in \mathbb{R}^n$. Se $\bar{x}$ é minimizador local, então o gradiente neste ponto é nulo.

Teorema 2. Suponhamos que $f : \mathbb{R}^n \to \mathbb{R}$ seja duas vezes diferenciável em $\bar{x} \in \mathbb{R}^n$. Se $\bar{x}$ é minimizador local irrestrito de f, então o gradiente é nulo em $\bar{x}$ e a Hessiana de f é semi-definida positiva neste ponto.

O próximo teorema garante que se f é convexa todo minimizador local é minimizador global do problema. A demonstração deste teorema se encontra em Karas e Ribeiro [3].

Teorema 3 Sejam $D \subset \mathbb{R}^n$ e $f : D \to \mathbb{R}^n$ convexa. Se $\bar{x}$ é minimizador local de f, então $\bar{x}$ é minimizador global de f.

[1]edsr.mat19@uea.edu.br
[2]erpereira@uea.edu.br

O teste numérico aqui apresentado, utiliza o método gradiente, e é aplicado na função quadrática

$$f(x_1, x_2) = 12x_1^2 - 12x_1x_2 + 4x_2^2 + 2x_1 \tag{3}$$

Todas as iterações serão feitas utilizando comprimento do passo fixo. Será considerado como válido (sem provas) o fato de que se $\alpha_k < \frac{2}{L}$, sendo L o maior autovalor da matriz Hessiana, então a sequência x^k converge para um ponto estacionário. O passo fixo escolhido dessa maneira garante a convergência do método. O critério de parada utilizado neste método será $\| f'(x^{k+1}) \| < 10^{-8}$. Foi considerado como ponto inicial nos testes $x = (1, 0)$, o passo escolhido está no intervalo $[0, 0.0657]$, considerando que o maior autovalor da matriz Hessiana é $L \approx 30.42$.

Tabela 1: Problema (1)

N° de Iterações.	Comprimento de passo.	Ponto inicial.	Ponto final.
595	0.02	(1,0)	$(-\frac{1}{3},-\frac{1}{2})$
395	0.03	(1,0)	$(-\frac{1}{3},-\frac{1}{2})$
356	0.033	(1,0)	$(-\frac{1}{3},-\frac{1}{2})$
292	0.04	(1,0)	$(-\frac{1}{3},-\frac{1}{2})$
232	0.05	(1,0)	$(-\frac{1}{3},-\frac{1}{2})$
192	0.06	(1,0)	$(-\frac{1}{3},-\frac{1}{2})$

É possível notar neste teste que a velocidade de convergência aumenta a medida que aproximamos o passo de 0.0657 diminuindo o número de iterações e diminui a medida que se distanciamos do mesmo. A função utilizada neste problema é convexa, isso garante que o ponto estacionário encontrado é minimizador global do problema. O método do gradiente pode não ser eficiente quando a função que está sendo tratada possui curvas de nível muito alongadas (como a usada no teste), já que isso aumentaria o tempo de convergência, ou seja, o método para esses casos é lento e por este fato, menos interessante.

Agradecimentos

Agradecemos ao Centro de Estudo Superiores de Tabatinga da Universidade do Estado do Amazonas pelo espaço para estudo em Nível Superior e a Fundação de Amparo a Pesquisa do Etado do Amazonas-FAPEAM pelo apoio financeiro para realização desta pesquisa.

Referências

[1] A. Ismailov e M. Solodov. **Otimização: Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade**. 2a. ed. Rio de Janeiro: Impa, 2014. ISBN: 9786599052804.

[2] A. Ismailov e M. Solodov. **Otimização: Métodos Computacionais**. 2a. ed. Rio de Janeiro: Impa, 2012. ISBN: 9788524404542.

[3] E. W. Karas e A. A. Ribeiro. **Otimização contínua: aspectos teóricos e computacionais**. 2a. ed. São Paulo: Cengage Learning, 2013.