Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 11, n. 1, 2025.

Trabalho apresentado no XLIII CNMAC, Centro de Convenções do Armação Resort - Porto de Galinhas - PE, 2024

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

Globalização do Método de Newton: Análise comparativa entre estratégias de Buscas Lineares em um Método Híbrido

Márcio Antônio de Andrade Bortoloti[1], Maria Clara Brito dos Reis[2]
Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista, BA

Neste trabalho, apresentamos um estudo de um método de Newton globalizado para determinar o minimizador de uma função objetivo $f : \mathbb{R}^n \to \mathbb{R}$ duas vezes continuamente diferenciável. O Método de Newton (MN) destaca-se por apresentar taxa de convergência superlinear, podendo chegar a quadrática, sob algumas hipóteses. Contudo, uma dificuldade do MN reside na exigência de um chute inicial próximo à solução. Para contornar essa dificuldade, pode-se empregar estratégias para o fornecimento de chutes iniciais mais adequados para o MN. Uma delas consiste em utilizar a direção de descida do gradiente da função objetivo. O método que utiliza essa direção de descida é conhecido como Método do Gradiente (MG) e é caracterizado pela convergência global da sequência gerada. Apesar disso, essa convergência tende a ser lenta com taxa linear, tornando-o menos eficiente, veja [1]. Buscando reunir a característica global do MG com as boas taxas de convergência do MN, consideramos um método que gera uma sequência com o emprego da direção de descida do gradiente. Este fornece um chute apropriado para as direções de descida dadas pelo MN. Essas direções são obtidas pelas soluções d^k da equação linear

$$H(x^k)d^k = -\nabla f(x^k), \tag{1}$$

onde $H(x^k)$ é a matriz Hessiana de f no ponto x^k. Além da direção de descida, o estabelecimento de uma busca linear é fundamental na construção da sequência x^k.

Neste estudo, consideraremos as buscas monotônicas de Armijo, Goldstein e Wolfe, veja [2]. Estas serão testadas tanto na fase inicial conduzida pelo MG quanto na fase de aceleração promovida pelo MN. O objetivo é investigar o desempenho dessas combinações equipadas em um Método Híbrido (MH) ao determinar o mínimo global da função objetivo.

A Regra de Armijo consiste em encontrar um comprimento de passo que assegure uma redução suficiente em f, isto é, determinar um valor $\alpha^k > 0$ que satisfaça, para $\sigma \in (0,1)$ a condição

$$f(x^k + \alpha^k d^k) \le f(x^k) + \sigma\alpha^k \langle f'(x^k), d^k\rangle. \tag{2}$$

Na Regra de Goldstein é adicionada à condição de Armijo, uma desigualdade com o intuito de descartar comprimentos de passo excessivamente pequenos. Essa regra busca determinar, para $0 < \sigma_1 < \sigma_2 < 1$, um $\alpha^k > 0$ tal que

$$f(x^k) + \sigma_1\alpha^k\langle f'(x^k), d^k\rangle \le f(x^k + \alpha^k d^k) \le f(x^k) + \sigma_2\alpha^k\langle f'(x^k), d^k\rangle. \tag{3}$$

Adicionalmente, a Regra de Wolfe complementa a desigualdade de Armijo com uma condição de curvatura, estipulando que, além da redução suficiente em f, para $0 < \sigma_1 < \sigma_2 < 1$, o comprimento de passo $\alpha^k > 0$ deve cumprir

$$\begin{gathered} f(x^k + \alpha^k d^k) \le f(x^k) + \sigma_1\alpha^k\langle f'(x^k), d^k\rangle \\ \langle f'(x^k + \alpha^k d^k), d^k\rangle \ge \sigma_2\langle f'(x^k), d^k\rangle. \end{gathered} \tag{4}$$

[1]mbortoloti@uesb.edu.br
[2]maria.cbritoreis@gmail.com

A garantia da existência do comprimento de passo para cada uma das buscas comentadas acima está apresentada em [1].

O algoritmo a seguir detalha a globalização do método de Newton realizada por meio do MH.

Algoritmo 1: Método Híbrido (MH)

1 Tome um ponto inicial $x^0 \in \mathbb{R}^n$ e $\varepsilon > 0$.
2 Defina $k = 0$.
3 **Repita** enquanto $\nabla f(x^k) \neq 0$.
4   **Se** $||\nabla f(x^k)|| > \epsilon$
5   Calcule $d^k = -\nabla f(x^k)$
6   **Senão**
7   Calcule d^k por (1).
8 Calcule α^k por (2), (3) ou (4).
9 Defina $x^k = x^k + \alpha^k d^k$
10 Defina $k \leftarrow k + 1$ e retorne para o passo 3

Gostaríamos de destacar que entre as linhas 4 e 7 do algoritmo, o parâmetro ϵ indica quando a troca da direção de descida ocorrerá. Em outras palavras, quando a norma do gradiente se torna suficientemente pequena, a direção do MG é alterada para a direção do MN.

O desempenho do Algoritmo 1 será avaliado utilizando funções do CUTEst, [3]. Esta implementação será realizada em Linguagem de Programação Julia e os códigos-fonte estarão disponíveis no endereço `https://github.com/petimatematica/HGNM`. Neste contexto, *perfomance profiles* serão elaborados com o objetivo de analisar os resultados alcançados ao comparar a eficiência de diferentes combinações das buscas lineares dadas por (2), (3) e (4).

Agradecimentos

A autora Maria Clara Brito dos Reis agradece ao Programa de Educação Tutorial Institucional da Universidade Estadual do Sudoeste da Bahia (PETI/UESB) pelo apoio financeiro.

Referências

[1] Alexey Izmailov e Mikhail Solodov. **Otimização, volume 2: métodos computacionais**. IMPA, 2018.

[2] Jorge Nocedal e Stephen J Wright. **Numerical optimization**. Springer, 1999.

[3] Nicholas IM Gould, Dominique Orban e Philippe L Toint. "CUTEst: a constrained and unconstrained testing environment with safe threads for mathematical optimization". Em: **Computational optimization and applications** 60 (2015), pp. 545–557.