Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 10, n. 1, 2023.

Trabalho apresentado no XLII CNMAC, Universidade Federal de Mato Grosso do Sul - Bonito - MS, 2023

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Um Estudo Numérico do Método dos Gradientes Conjugados para as buscas lineares de Armijo e Goldstein

Giselle Lopes da Cruz[1] Márcio Antônio de Andrade Bortoloti[2] Wéllington Moutinho Dias [3]
Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista, BA

Um dos métodos mais utilizados para a minimização de uma função $f : \mathbb{R}^n \to \mathbb{R}$, continuamente diferenciável, é o **Método do Gradiente** (MG). Esse método gera sequências que convergem linearmente quando são empregadas buscas lineares para a determinação do comprimento do passo na geração da sequência. Essas buscas procuram fornecer um comprimento de passo que assegure um decréscimo da função objetivo. O **Método dos Gradientes Conjugados** (MGC) é uma alternativa que melhora a eficiência na determinação de minimizadores de funções, quando comparado com o MG. Cabe ressaltar que para o caso de funções quadráticas, o MGC admite uma fórmula fechada para a determinação do comprimento de passo. Contudo, para funções que não se enquadram nesse quesito, é necessário uma regra de busca linear para a determinação do comprimento de passo.

Neste trabalho, apresentaremos um estudo computacional comparando o MG com o MGC, quando equipados com as buscas de Armijo, [1, Cap. 4], e Goldstein, [2, Cap. 3]. Para um $x_k \in \mathbb{R}^n$, $\rho_1 \in (0, 1)$ e uma direção de descida $d_k \in \mathbb{R}^n \setminus \{0\}$, a **Busca de Armijo** consiste em determinar $\alpha_k \in \mathbb{R}^+$ tal que

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \rho_1 \alpha_k \nabla f(x_k)^T d_k. \tag{1}$$

Uma característica dessa busca é a possibilidade de fornecimento de passos muito pequenos. Isto pode gerar um maior esforço computacional nos problemas de minimização.

Para prevenir essa situação é possível adicionar uma segunda condição sobre (1), visando obter um parâmetro de comprimento de passo satisfatório, veja [2, Cap. 3]. Isso assegura que o passo dado seja consideravelmente significativo, isto é, evita passos muito pequenos. Matematicamente, consiste em encontrar um $\alpha_k \in \mathbb{R}^+$ que obedece as desigualdades

$$f(x_k) + \rho_2 \alpha_k \nabla f(x_k)^T d_k \leq f(x_k + \alpha_k d_k) \leq f(x_k) + \rho_1 \alpha_k \nabla f(x_k)^T d_k, \tag{2}$$

onde $d_k \in \mathbb{R}^n \setminus \{0\}$ é uma direção de descida a partir de $x_k \in \mathbb{R}^n$ e $0 < \rho_1 < \rho_2 < 1$ são os parâmetros dados. A condição dada por (2) é chamada de **Regra de Goldstein**.

O MGC está apresentado no algoritmo abaixo.

[1]giselle.lcz@gmail.com
[2]mbortoloti@uesb.edu.br
[3]wellingtonmoutinhodias@gmail.com

Algoritmo 1: Método dos Gradientes Conjugados

1 Tome um ponto inicial $x_0 \in \mathbb{R}^n$, $\varepsilon > 0$, faça $d_0 := -\nabla f(x_0)$
2 $k := 0$
3 **Repita** enquanto $\nabla f(x_k) \neq 0$
4   Se $\|d_k\| > \varepsilon$, calcule α_k por (1) ou (2)
5     $x_{k+1} := x_k + \alpha_k d_k$
6   **Se** $(k+1) \bmod n \neq 0$
7         $\beta_k = \beta_k^Q$ ou $\beta_k = \beta_k^{PR}$
8   **Senão**
9         $\beta_k = 0$
10     $d_{k+1} := -\nabla f(x_{k+1}) + \beta_k d_k$
11     $k \leftarrow k+1$ e retorne para o passo 3

Observamos que se tomarmos $\beta_k = 0$ para todo $k \in \mathbb{N}$, o Algoritmo 1 se torna o método MG. Para o caso de minimização de funções quadráticas, o Algoritmo 1 pode tomar

$$\beta_k^Q = \frac{(d^k)^T A \nabla f(x^{k+1})}{(d^k)^T A d^k}. \tag{3}$$

Já para casos de funções não quadráticas consideramos o parâmetro β_k dado por

$$\beta_k^{FR} = \frac{\nabla f(x^{k+1})^T \nabla f(x^{k+1})}{\nabla f(x^k)^T \nabla f(x^k)}. \tag{4}$$

Para uma discussão sobre a influência desses parâmetros no desempenho do algoritmo, sugerimos o livro [1]. Observamos que a definição da busca a ser utilizada é feita na linha 4 do algortimo.

Com o objetivo de analisar o comportamento desse algoritmo, vamos apresentar um estudo numérico comparativo entre os métodos MGC e MG equipados com cada uma das buscas lineares apresentadas em (1) e (2). Para esse estudo vamos utilizar algumas das funções teste apresentadas em [3]. Analisaremos os *performance profiles* de cada um dos métodos com cada uma das buscas lineares para determinar o comportamento desses algoritmos na minimização dessas funções.

Agradecimentos

Os autores agradecem ao **Programa de Educação Tutorial Institucional da Universidade Estadual do Sudoeste da Bahia (PETI/UESB)** pelas bolsas de estudo e à **UESB** pelo apoio financeiro na hospedagem e na alimentação.

Referências

[1] A. A. Ribeiro e E. W. Karas. **Otimização Contínua: Aspectos teoricos e computacionais**. São Paulo: Cengage Learning, 2013. ISBN: 978-85-221-2002-4.

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

[3] S. Surjanovic e D. Bingham. **Virtual Library of Simulation Experiments: Test Functions and Datasets**. Retrieved March 9, 2023, from `http://www.sfu.ca/~ssurjano`.