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

Um Estudo Numérico do Emprego de duas Estratégias para a Busca de Armijo em Otimização Restrita

Emanuel M. Queiroz[1] Marcio A. A. Bortoloti[2]
UESB, Vitória da Conquista, BA

Neste trabalho, objetivamos desenvolver um estudo numérico para analisar o desempenho de um método do Gradiente Projetado na minimização de uma função utilizando duas estratégias para a busca de Armijo apresentadas por Iusem em [2]. Em geral, os métodos para otimização com restrições se preocupam com o problema do tipo

$$\min\ f(x) \text{ sujeito a } x \in C, \tag{1}$$

onde $f : \mathbb{R}^n \to \mathbb{R}$ é uma função diferenciável e $C \subset \mathbb{R}^n$ é um conjunto convexo e fechado, denominado conjunto viável.

Métodos do Gradiente Projetado podem ser vistos como uma combinação de métodos do Gradiente para otimização irrestrita com projeções no conjunto viável do problema. A convexidade do conjunto viável torna possível a utilização do operador projeção ortogonal $P_C : \mathbb{R}^n \to C$ para se obter direções viáveis, as quais também são de descida, conforme [3]. O processo ocorre da seguinte forma: a partir de x_k, toma-se um comprimento de passo na direção oposta à do vetor gradiente, isto é, $-\nabla f(x_k)$. A seguir, aplica-se o operador projeção ortogonal P_C no vetor resultante, obtendo-se assim, uma direção viável para o cálculo do próximo termo da sequência.

Diversas estratégias são utilizadas para se determinar o comprimento de passo utilizado em um método computacional para a resolução de (1). Neste trabalho, trataremos de duas estratégias baseadas na busca de Armijo, veja [2]. A primeira estratégia utilizada, que vamos chamar de GPA1, trata-se do emprego da busca de Armijo ao longo de direções viáveis do conjunto C. Neste caso, a construção da sequência $\{x_k\}$ é feita considerando-se $x_{k+1} = x_k + \gamma_k(z_k - x_k)$, onde o parâmetro γ_k é determinado por

$$\gamma_k = 2^{-\ell(k)}, \tag{2}$$

com $\ell(k) = \min\left\{j \in \mathbb{Z}_+ : f(x_k + 2^{-j}(z_k - x_k)) \leq f(x_k) - \sigma 2^{-j}\nabla f(x_k)^T(x_k - z_k)\right\}$ e $\sigma \in (0,1)$. Além disso, tomamos $z_k = P_C(x_k - \beta_k \nabla f(x_k))$ e a sequência $\{\beta_k\} \subset [\tilde{\beta}, \hat{\beta}]$ para $0 < \tilde{\beta} \leq \hat{\beta}$. A segunda estratégia, denominada GPA2, consiste em utilizar a busca de Armijo para determinar termos da sequência ao longo da fronteira do conjunto C. Neste caso, a determinação dos termos é dada por $x_{k+1} = P_C(x_k - \beta_k \nabla f(x_k))$, onde tomamos

$$\beta_k = \overline{\beta} 2^{-\ell(k)}, \tag{3}$$

com $\ell(k) = \min\left\{j \in \mathbb{Z}_+ : f(z_{k,j}) \leq f(x_k) - \sigma \nabla f(x_k)^T(x_k - z_{k,j})\right\}$ e $z_{k,j} = P_C(x_k - \overline{\beta} 2^{-j}\nabla f(x_k))$, para algum $\overline{\beta} > 0$ e $\sigma \in (0,1)$.

A seguir, apresentamos formalmente, por meio do Algoritmo 1, o método do Gradiente Projetado escolhido equipado com a estratégia GPA1 ou com a estratégia GPA2.

[1]emanuelmqueiroz.emq@gmail.com
[2]mbortoloti@uesb.edu.br

Algoritmo 1: Método do Gradiente Projetado

1 Tome um ponto inicial $x_0 \in C \subset \mathbb{R}^n$, $\varepsilon > 0$, $\sigma \in (0,1)$, $0 < \tilde{\beta} \leq \widehat{\beta}$, $\beta_0 \in [\tilde{\beta}, \widehat{\beta}]$ e faça $k = 0$.
2 Se $\|x_0 - P_C(x_0 - \nabla f(x_0))\| < \varepsilon$, então pare e declare que x_0 é um ponto estacionário.
3 Determine $\beta_k \in [\tilde{\beta}, \hat{\beta}]$ e γ_k por (2) para a GPA1; ou calcule β_k por meio de (3) e tome $\gamma_k = 1$ no passo 5 para a GPA2.
4 Calcule $z_k = P_C(x_k - \beta_k \nabla f(x_k))$.
5 Calcule $x_{k+1} = x_k + \gamma_k(z_k - x_k)$.
6 Se $\|x_k - x_{k+1}\| < \varepsilon$, então pare e declare que x_k é um ponto estacionário.
7 Faça $k = k + 1$ e retorne para o passo 3.

Gostaríamos de observar que, no passo 3, o comprimento de passo β_k para a GPA1 é estabelecido por meio de interpolação quadrática, como pode ser visto em [1].

A fim de desenvolver um estudo numérico do método do Gradiente Projetado com as estratégias apresentadas, vamos considerar o problema de minimizar a função Dixon-Price $f : \mathbb{R}^n \to \mathbb{R}$ dada por

$$f(x) = (x_1 - 1)^2 + \sum_{i=2}^{n} i(2x_i^2 - x_{i-1})^2.$$

Serão considerados diferentes conjuntos viáveis, cujas fórmulas de projeções são apresentadas em [3]. A implementação do algoritmo será feita utilizando a linguagem de programação Julia. O estudo numérico analisará o desempenho das estratégias GPA1 e GPA2 considerando o tempo de CPU, o número de avaliações de função e o número de iterações.

Nota-se que a estratégia GPA2 requer uma projeção no conjunto C para cada comprimento de passo das iteradas internas resultantes da busca Armijo, o que pode gerar muitas projeções para cada iteração k, enquanto a estratégia GPA1 requer apenas uma projeção para cada k. Como consequência, a GPA2 é viável apenas quando o cálculo do operador projeção ortogonal P_C tem custo computacional relativamente baixo.

Agradecimentos

O autor Emanuel Mendes Queiroz agradece ao Programa de Educação Tutorial Institucional da Universidade Estadual do Sudoeste da Bahia - PETI/UESB pela bolsa de estudos.

Referências

[1] E. G. Birgin, J. M. Martínez e M. Raydan. "Spectral projected gradient methods: review and perspectives". Em: **Journal of Statistical Software** 60 (2014), pp. 1–21.

[2] A. N. Iusem. "On the convergence properties of the projected gradient method for convex optimization". Em: **Computational & Applied Mathematics** 22 (2003), pp. 37–52.

[3] A. Izmailov e M. Solodov. **Otimização, volume 2: métodos computacionais**. IMPA, 2018.