

UM ESTUDO NUMÉRICO DE DUAS ESTRATÉGIAS PARA A BUSCA DE ARMIJO EM OTIMIZAÇÃO RESTRITA

EMANUEL MENDES QUEIROZ¹, MARCIO ANTÔNIO DE ANDRADE BORTOLOTI²

Universidade Estadual do Sudoeste da Bahia, BA, Brasil

emanuelmqueiroz.emq@gmail.com¹, mbortoloti@uesb.edu.br²

Resumo

Neste trabalho, realizamos um estudo comparativo de um método do Gradiente Projetado na minimização de uma função utilizando a busca de Armijo ao longo das direções viáveis e ao longo do arco de projeção de diferentes conjuntos viáveis. Nossos experimentos evidenciaram que a escolha do conjunto viável influenciou no desempenho das estratégias, mas, em geral, a estratégia ao longo do arco de projeção obteve melhor desempenho considerando o tempo de CPU e o número de iterações. Por outro lado, estratégia ao longo das direções viáveis foi mais eficiente em relação ao número de avaliações da função.

Palavras-Chave: Busca de Armijo; Método do Gradiente Projetado; Otimização restrita.

Introdução

Em geral, os métodos para otimização com restrições buscam resolver o problema

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

onde $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 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 a projeção, no conjunto C , dos termos da sequência obtida no processo iterativo. A convexidade do conjunto viável torna possível a utilização do operador projeção ortogonal $P_C: \mathbb{R}^n \rightarrow C$ para se obter direções viáveis, as quais também são de descida, conforme [1]. No processo iterativo de um método do Gradiente Projetado, a partir de x_k , toma-se um comprimento de passo na direção $-\nabla f(x_k)$ e 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 podem ser empregadas para a determinação do comprimento de passo utilizado em um método de resolução de problemas de otimização restrita, conforme [3]. Neste trabalho, trataremos de duas estratégias baseadas na busca de Armijo.

Fundamentação Teórica

As estratégias analisadas neste trabalho são baseadas em Iusem [2]. A primeira estratégia, a qual denominaremos GPA1, trata-se do emprego da busca de Armijo ao longo das direções viáveis de 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

$$\gamma_k = 2^{-\ell(k)}, \quad (1)$$

com $\ell(k) = \min \{j \in \mathbb{Z}_+ : f(x_k + 2^{-j}(z_k - x_k)) \leq f(x_k) - \sigma 2^{-j} \nabla f(x_k)^\top (x_k - z_k)\}$ e $\sigma \in (0, 1)$. Além disso, tomamos $z_k = P_C(x_k - \beta_k \nabla f(x_k))$ com $\beta_k \in [\hat{\beta}, \bar{\beta}]$ para $0 < \hat{\beta} \leq \bar{\beta}$. A segunda estratégia, a qual denominaremos GPA2, consiste no emprego da busca de Armijo ao longo do arco de projeção do conjunto C . Neste caso, a determinação dos termos da sequência é dada por $x_{k+1} = P_C(x_k - \beta_k \nabla f(x_k))$, onde

$$\beta_k = \bar{\beta} 2^{-\ell(k)}, \quad (2)$$

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

Desenvolvimento e Metodologia

A seguir, apresentamos formalmente, no Algoritmo 1, o método do Gradiente Projetado equipado com a estratégia GPA1 ou com a estratégia GPA2.

Algoritmo 1: MÉTODO DO GRADIENTE PROJETADO

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

Os resultados de convergência para o Método do Gradiente Projetado com a busca de Armijo ao longo das direções viáveis e ao longo do arco de projeção podem ser vistas em [2].

A fim de desenvolver o estudo numérico com as estratégias apresentadas, consideramos o problema de minimizar a função Dixon-Price, dada por $f: \mathbb{R}^n \rightarrow \mathbb{R}$, onde

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

Nos testes desenvolvidos, consideramos 4 conjuntos viáveis: $C_1 = B(\hat{x}, \delta)$, onde $\hat{x} = (-50, 0, \dots, 0)$ e $\delta = 20$; $C_2 = \{x \in \mathbb{R}^n; \langle a, x \rangle = 0\}$, onde $a = (1, 1, \dots, 1)$; $C_3 = \{x \in \mathbb{R}^n; \langle a, x \rangle \leq 0\}$, onde $a = (1, 1, \dots, 1)$; $C_4 = \{x \in \mathbb{R}^n; Ax = (0, 0)\}$, onde $A \in \mathbb{R}(2, n)$ tem posto 2. Para cada conjunto viável, tomamos 10 dimensões (variando de 5 a 500) e 5 chutes iniciais dados aleatoriamente, totalizando 200 problemas. Além disso, tomamos $\bar{\beta} = 1.0$, $\hat{\beta} = 0.01$, $\hat{\beta} = 0.9$ e $\beta_0 = 0.01$. Nós consideramos convergência na iteração k quando $\|x_k - x_{k+1}\| < 10^{-5}$. O algoritmo foi interrompido quando o número máximo $k = 40000$ de iterações foi atingido ou quando o comprimento de passo atingiu um valor menor que 10^{-5} . Os códigos utilizados, implementados na linguagem de programação Julia, estão livremente disponibilizados em https://github.com/petimatematica/MGP_EA.

Resultados

No estudo numérico, analisamos o desempenho das estratégias GPA1 e GPA2 considerando o tempo de CPU, o número de avaliações de função e a quantidade de iterações. Conforme as figuras 1 e 2, a GPA2 obteve melhor desempenho relativo ao tempo de CPU e ao número de iterações. A Figura 3 mostra que a GPA1 foi mais eficiente em relação à avaliação de função.

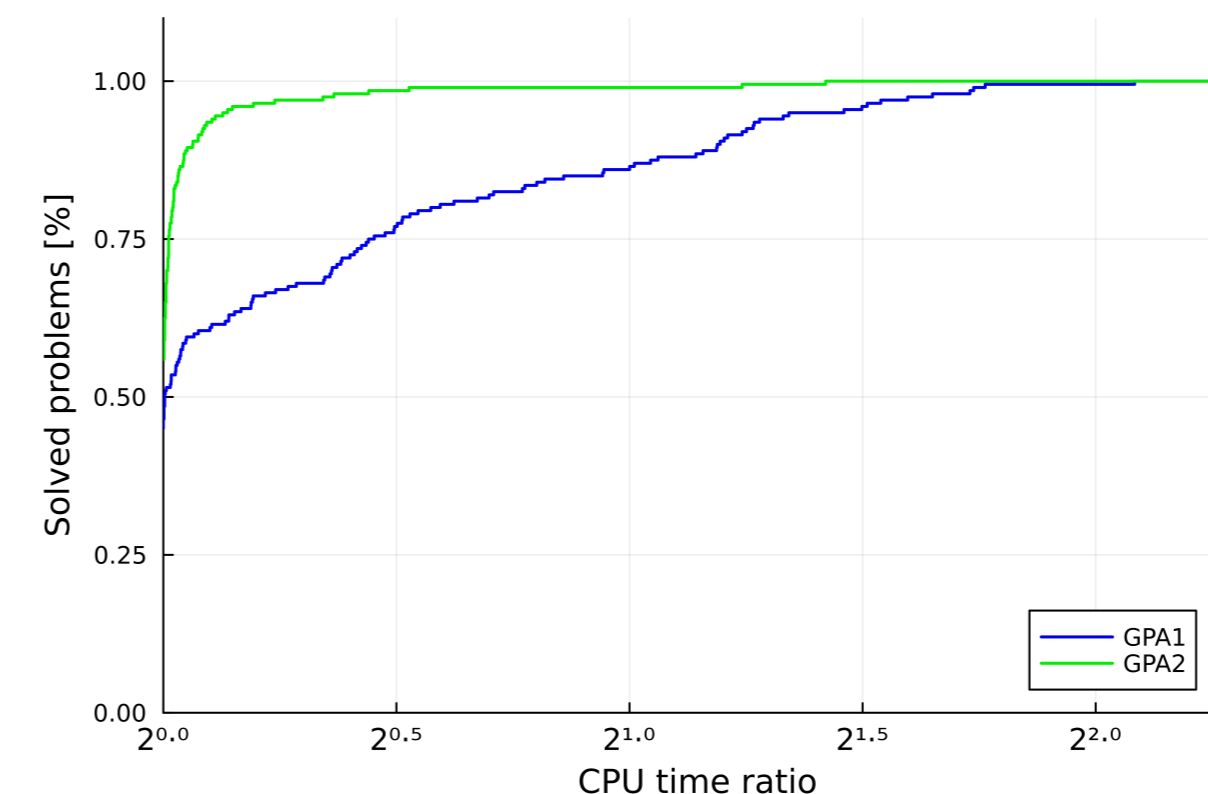


Figura 1: Perfil de desempenho relativo ao tempo de CPU.

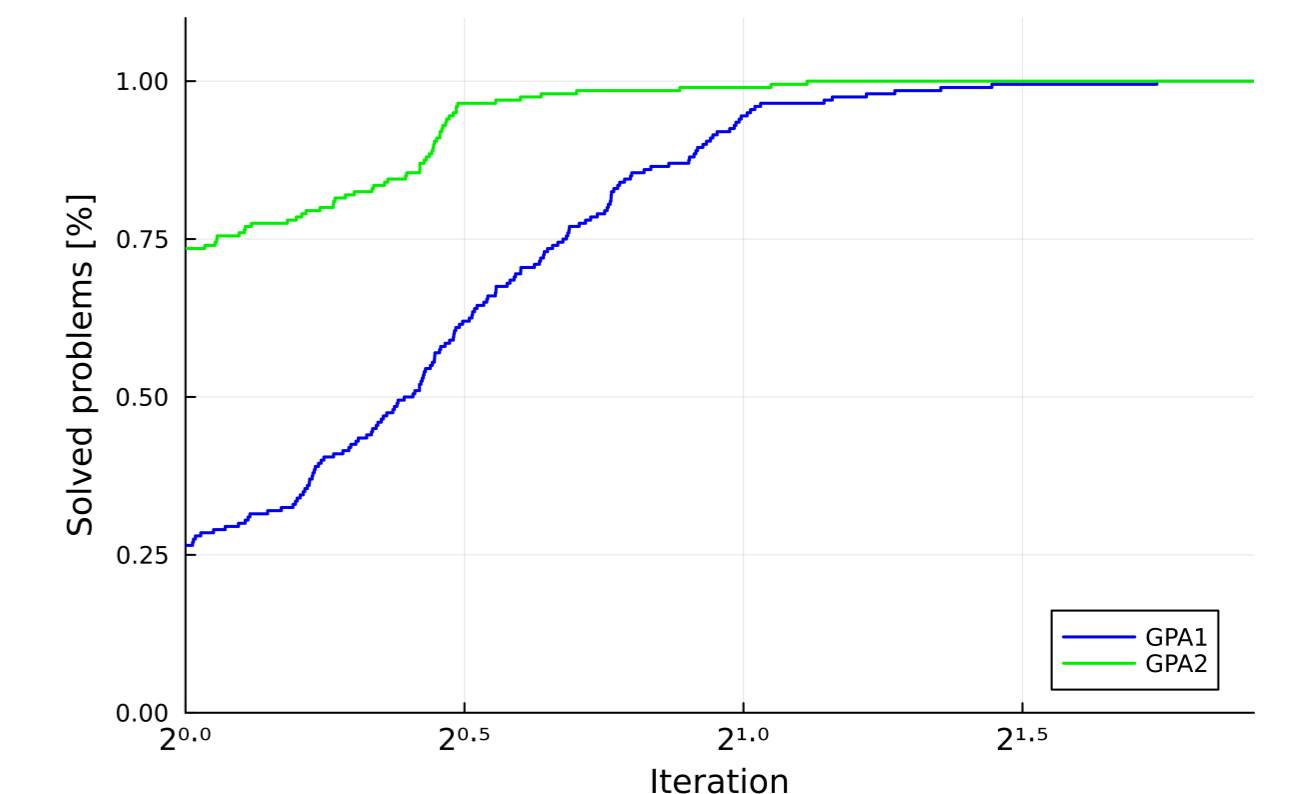


Figura 2: Perfil de desempenho relativo ao número de iterações.

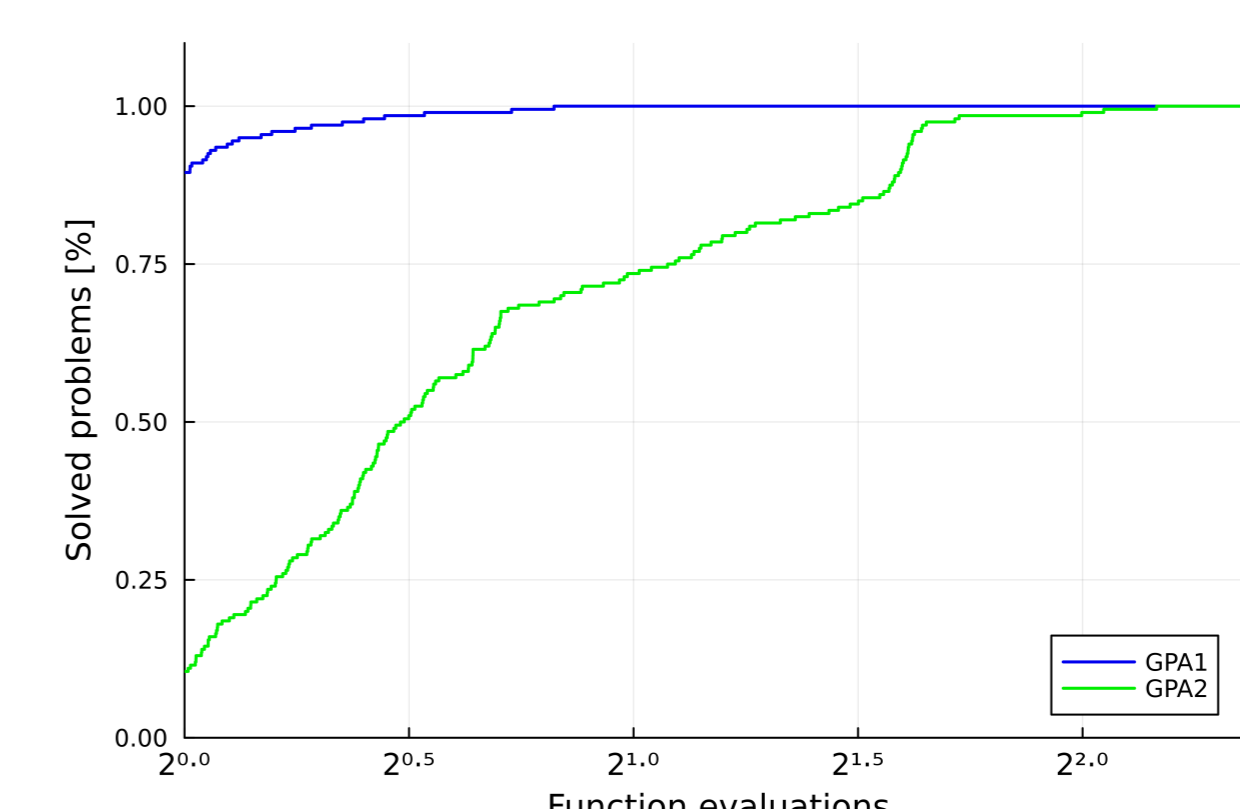


Figura 3: Perfil de desempenho relativo à avaliação de função.

Nossos experimentos mostraram que as restrições dos conjuntos viáveis influenciaram no desempenho das estratégias. Além disso, ressaltamos que a GPA1, a cada iteração, fixa uma direção viável e realiza a busca de um comprimento de passo. Por outro lado, a GPA2 realiza a busca de um comprimento de passo ao longo de diferentes direções viáveis, a cada iteração, o que pode acarretar mais avaliações de função para se obter a direção e o comprimento de passo.

Referências

- [1] A. Izmailov e M. Solodov. *Otimização, volume 2: métodos computacionais*. IMPA, 2018.
- [2] A. N. Iusem. *On the convergence properties of the projected gradient method for convex optimization*. Computational & Applied Mathematics, volume 22, pages 37–52. SciELO Brasil, 2003.
- [3] Nocedal, J., & Wright, S. J. (Eds.). (1999). *Numerical optimization*. New York, NY: Springer New York.

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, à UESB e ao International Council for Industrial and Applied Mathematics (ICIAM) pelo apoio financeiro.