

ESTUDO COMPARATIVO ENTRE BUSCAS LINEARES NÃO MONOTÔNICAS EMPREGADAS NO MÉTODO DO GRADIENTE ESPECTRAL PROJETADO EM CONJUNTOS CONVEXOS

Samara Viriato Vilar Dias¹, Marcio Antônio de Andrade Bortoloti²

Universidade Estadual do Sudoeste da Bahia, BA, Brasil

samaravvilar@gmail.com¹, mbortoloti@uesb.edu.br²

Resumo

Neste trabalho, apresentamos um estudo comparativo entre duas buscas lineares não monotônicas empregadas no Método do Gradiente Espectral Projetado. Elas foram utilizadas na minimização de funções diferenciáveis restritas a um conjunto convexo fechado. A restrição considerada neste trabalho foi a de limitação de variáveis. O conjunto de problemas foi estabelecido por meio da biblioteca de funções teste CUTEst. Os experimentos numéricos evidenciaram que a busca ao longo do arco de projeção apresentou melhor desempenho que a busca ao longo de direções viáveis, embora empregue uma quantidade maior de projeções.

Palavras-Chave: Gradiente Projetado; Passo espectral; Busca linear não monotônica.

Introdução

Desde o seu surgimento, o Método do Gradiente Espectral Projetado (SPG) tem sido amplamente utilizado na resolução de problemas restritos, como os que ocorrem na tomografia de reflexão sísmica, [3]. O método SPG incorpora duas adaptações em relação ao Método do Gradiente Projetado: o passo espectral de Barzilai e Borwein, [1], e uma busca linear não monotônica. Neste trabalho, analisamos o método SPG para a minimização de funções definidas em conjuntos convexos fechados, restritos por limitações nas variáveis, utilizando duas buscas lineares não monotônicas. Investigamos a busca desenvolvida por Grippo, Lampariello e Lucidi, [5], que é caracterizada pela estratégia de determinar comprimentos de passo ao longo do arco de projeção, e a busca proposta por Birgin, Martínez e Raydan, [2], que é realizada ao longo de direções viáveis.

Fundamentação Teórica

Seja $f : \Omega \rightarrow \mathbb{R}$ uma função diferenciável e $\Omega \subset \mathbb{R}^n$ um conjunto convexo e fechado. Considerando o problema de minimizar f em Ω , abordaremos o método SPG equipado com duas buscas lineares não monotônicas. Para isso, denotaremos $x_k(\alpha_k) = P(x_k - \alpha_k \nabla f(x_k))$, onde $P : \mathbb{R}^n \rightarrow \Omega$ é uma projeção ortogonal. Na busca proposta por Grippo, Lampariello e Lucidi, [5], a qual denominaremos GLL, dados $\gamma \in (0, 1)$ e um inteiro $M > 0$, busca-se determinar um comprimento de passo $\alpha_k > 0$ tal que

$$f(x_k(\alpha_k)) \leq \max_{0 \leq j \leq \min\{k, M-1\}} f(x_{k-j}) + \gamma \nabla f(x_k)^\top (x_k(\alpha_k) - x_k). \quad (1)$$

Nota-se que, em (1), o cálculo do produto $\nabla f(x_k)^\top (x_k(\alpha_k) - x_k)$ é realizado para cada ponto teste $x_k(\alpha_k)$. Segundo Birgin, Martínez e Raydan, [2], este emprego sucessivo de projeções pode aumentar o esforço computacional, o que os levou à formulação de uma nova busca linear não monotônica para o método SPG. Nesta busca, que denominaremos BMR, ao rejeitar o primeiro ponto de teste, os próximos são calculados ao longo da mesma direção viável $d_k = x_k(\lambda_k) - x_k$, onde λ_k é o passo espectral. Isso resulta em um único cálculo da direção d_k e uma única operação de projeção por iteração. Assim, a busca BMR consiste, dados $\gamma \in (0, 1)$ e um inteiro $M > 0$, em determinar $\alpha_k > 0$ tal que

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq \min\{k, M-1\}} f(x_{k-j}) + \gamma \alpha_k \nabla f(x_k)^\top d_k. \quad (2)$$

Desenvolvimento e Metodologia

Neste trabalho, quando o método SPG for equipado com a busca GLL, referimo-nos a ele como SPG1; enquanto, quando equipado com a busca BMR, será denominado SPG2. O Algoritmo 1 apresenta formalmente os métodos SPG1 e SPG2.

Algoritmo 1: MÉTODO DO GRADIENTE ESPECTRAL PROJETADO (SPG)

- 1 Considere $x_0 \in \Omega$, $\varepsilon > 0$, $\gamma \in (0, 1)$, $1 > \sigma_2 > \sigma_1 > 0$, um inteiro $M > 0$, $\lambda_{\max} > \lambda_{\min} > 0$, $\lambda_0 \in [\lambda_{\min}, \lambda_{\max}]$ e $k = 0$.
- 2 Se $\|x_k(1) - x_k\|_\infty < \varepsilon$, então pare e declare que x_k é estacionário.
- 3 Comece com $\alpha_k = \lambda_k$, calcule α_k que satisfaça (1) e $x_{k+1} := x_k(\alpha_k)$ para o SPG1; ou comece com $\alpha_k = 1$, calcule α_k que satisfaça (2) e $x_{k+1} := x_k + \alpha_k(x_k(\lambda_k) - x_k)$ para o SPG2.
- 4 Se (1) não for válida ou (2) não for válida, tome $\alpha_{\text{new}} \in [\sigma_1 \alpha_k, \sigma_2 \alpha_k]$, faça $\alpha_k = \alpha_{\text{new}}$ e retorne para o passo 3.
- 5 Calcule $s_k = x_{k+1} - x_k$ e $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$. Se $s_k^\top y_k \leq 0$, então $\lambda_{k+1} := \lambda_{\max}$. Caso contrário, $\lambda_{k+1} := \max\{\lambda_{\min}, \min\{s_k^\top s_k / s_k^\top y_k, \lambda_{\max}\}\}$. Faça $k = k + 1$ e retorne para o passo 2.

Ressaltamos que, quando α_k não satisfaz a desigualdade (1) ou a desigualdade (2), o ajuste para obtenção do novo α_k , no passo 4, é realizado por meio de interpolação quadrática. Além disso, as propriedades de convergência do Algoritmo 1 são provadas nos Teoremas 2.3 e 2.4, que podem ser consultados em [2].

Os experimentos numéricos foram desenvolvidos a partir de problemas com restrições de limitação das variáveis da coleção CUTEst, [4], com dimensões menores ou iguais a 14884. A escolha dos problemas foi feita a partir daqueles apresentados em [2], onde foram selecionados 38 problemas de interesse acadêmico, de aplicações reais e da modelagem. Consideramos os parâmetros $\varepsilon = 10^{-5}$, $\gamma = 10^{-4}$, $\lambda_{\min} = 10^{-30}$, $\lambda_{\max} = 10^{30}$, $\sigma_1 = 0.1$, $\sigma_2 = 0.9$, $M = 10$ e o comprimento de passo inicial igual a $1/\|x_0(1) - x_0\|_\infty$. Ademais, o Algoritmo 1 é interrompido quando o número de iterações atinge 15000. Os códigos utilizados neste trabalho, implementados na linguagem de programação Julia, estão livremente disponíveis em <https://github.com/petimatematica/SPG>.

Resultados Numéricos

Nas Figuras 1, 2 e 3, podemos observar que o método SPG1 obteve melhor desempenho com relação ao número de iterações, tempo de CPU e avaliação de função. Contudo, como previsto na teoria abordada, o SPG2 empregou uma quantidade menor de projeções, conforme mostrado na Figura 4. Neste caso, apesar do emprego sucessivo de projeções pelo SPG1, esse fator não interferiu no seu desempenho, uma vez que a projeção utilizada requereu poucos recursos computacionais.

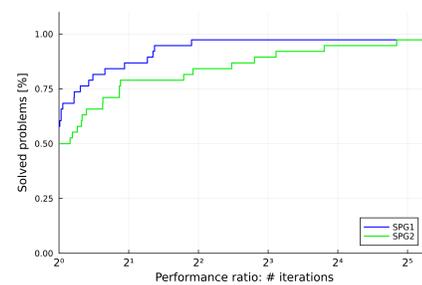


Figura 1: Performance profile relativo ao número de iterações.

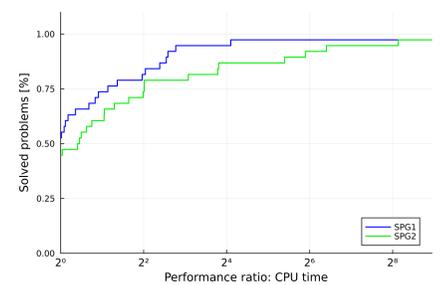


Figura 2: Performance profile relativo ao tempo de CPU.

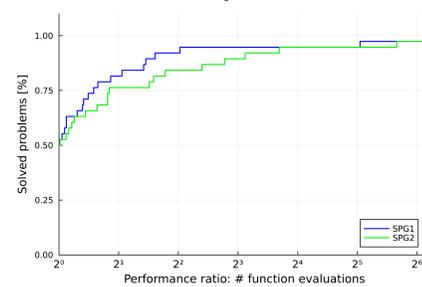


Figura 3: Performance profile relativo à avaliação de função.

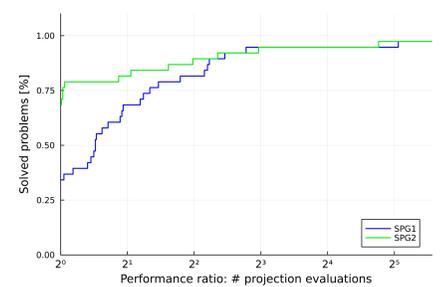


Figura 4: Performance profile relativo à avaliação de projeção.

Referências

- [1] Barzilai, J., & Borwein, J. M. (1988). Two-point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1), 141-148.
- [2] Birgin, E. G., Martínez, J. M., & Raydan, M. (2000). Nonmonotone spectral projected gradient methods on convex sets. *SIAM Journal on Optimization*, 10(4), 1196-1211.
- [3] Bello, L., & Raydan, M. (2007). Convex constrained optimization for the seismic reflection tomography problem. *Journal of Applied Geophysics*, 62(2), 158-166.
- [4] Gould, N. I., Orban, D., & Toint, P. L. (2015). *CUTEst: a constrained and unconstrained testing environment with safe threads for mathematical optimization*. Computational optimization and applications, 60, 545-557.
- [5] Grippo, L., Lampariello, F., & Lucidi, S. (1986). A nonmonotone line search technique for Newton's method. *SIAM Journal on Numerical Analysis*, 23(4), 707-716.

Agradecimentos

A primeira autora agradece ao Programa de Educação Tutorial Institucional da Universidade Estadual do Sudoeste da Bahia (PETI/UESB) pela bolsa de estudo.