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 Método dos Gradientes Conjugados Projetados

Márcio Antônio de Andrade Bortoloti[1], Wéllington Moutinho Dias[2]
UESB, Vitória da Conquista, BA

Neste trabalho apresentamos um estudo numérico do método dos Gradientes Conjugados Projetados, proposto por [1], para resolver problemas do tipo

$$\mathcal{F}(x) = 0, x \in \Omega, \tag{1}$$

onde Ω é um conjunto não-vazio, fechado e convexo e $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^n$ é continuamente diferenciável e monótona, isto é, satisfaz a desigualdade $(\mathcal{F}(x) - \mathcal{F}(y))^\top (x - y) \geq 0$ para todo $x, y \in \mathbb{R}^n$. Tipicamente, problemas dessa natureza podem ser resolvidos por meio de algoritmos que produzem sequências que convergem para as suas soluções. Uma estratégia importante, utilizada em [1], para criar uma sequência $\{x_k\}_{k\in\mathbb{N}}$ que satisfaça a condição $x_k \in \Omega$ para todo $k \in \mathbb{N}$, é definir um operador projeção $P_\Omega[\cdot] : \mathbb{R}^n \to \Omega$, tal que $P_\Omega[x] = \arg\min\{\|y - x\|; y \in \Omega\}$, para todo $x \in \mathbb{R}^n$. Gostaríamos de observar que a determinação de projeções não é uma tarefa simples e pode aumentar consideravelmente o esforço computacional. Contudo, alguns conjuntos possuem uma fórmula fechada para a sua projeção, veja [2, Capítulo 4, Seção 1].

A sequência $\{x_k\}_{k\in\mathbb{N}}$ é construída por meio do emprego de direções d_k com a propriedade

$$\mathcal{F}(x_k)^\top d_k \leq -c\|\mathcal{F}(x_k)\|^2,$$

onde $c \in (0, 1)$ é uma constante. Uma boa definição para d_k, que foi apresentada em [1], é considerar

$$d_k = \begin{cases} -\mathcal{F}(x_0), & \text{se } k = 0 \\ -\mathcal{F}(x_k) + \beta_k^+(\tau_{k-1})d_{k-1}, & \text{se } k \geq 1 \end{cases} \tag{2}$$

onde o parâmetro $\beta_k^+(\tau_{k-1})$ é dado por

$$\beta_k^+(\tau_{k-1}) = \max\left\{\beta_k(\tau_{k-1}),\ \eta\frac{\mathcal{F}(x_k)^\top d_{k-1}}{\|d_{k-1}\|^2}\right\}, \tag{3}$$

com $\eta \in [0, 1)$,

$$\beta_k(\tau_{k-1}) = \frac{\mathcal{F}_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2}\right)\frac{\mathcal{F}_k^\top\ s_{k-1}}{d_{k-1}^\top y_{k-1}}, \tag{4}$$

e τ_{k-1} em (4) é definida por

$$\tau_{k-1} = \theta\frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} + (1-\theta)\frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2},\ \theta \in [0, 1]. \tag{5}$$

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

Ademais, $y_k = \gamma_k + \lambda_k t_k \|\mathcal{F}(x_k)\| d_k$, com $\gamma_k = \mathcal{F}(x_{k+1}) - \mathcal{F}(x_k)$, $s_k = t_k d_k$, com $t_k > 0$ e

$$\lambda_k = 1 + \|\mathcal{F}(x_k)\|^{-1} \max\left\{0,\ -\frac{\gamma_k^\top d_k}{t_k \|d_k\|^2}\right\}. \tag{6}$$

Desse modo, o método estudado pode ser representado pelo Algoritmo 1, que pode ser visto abaixo.

Algoritmo 1: MGCP

Passo 0 Escolha $x_0 \in \Omega$; $\xi, \sigma, \rho \in (0,1)$; $\eta \in [0,1)$ e $\theta \in [0,1]$. Faça $k := 0$;
Passo 1 Pare se $\mathcal{F}(x_k) = 0$. Caso contrário, determine d_k por (2);
Passo 2 Encontre $t_k = \max\{\xi\rho^i : i = 0,1,2,\dots\}$, que satisfaça

$$-\mathcal{F}(x_k + t_k d_k)^\top d_k \geq \sigma t_k ||d_k||^2. \tag{7}$$

Faça $z_k := x_k + t_k d_k$;
Passo 3 Se $\mathcal{F}(z_k) = 0$, pare. Caso contrário, calcule $x_{k+1} = P_\Omega[x_k - \alpha_k \mathcal{F}(z_k)]$, em que $\alpha_k = F(z_k)^T (x_k - z_k)/||F(z_k)||^2$;
Passo 4 Faça $k := k + 1$. Retorne ao *Passo 1.*

Uma análise de convergência é apresentada em [1], onde é mostrado que o método apresentado converge globalmente com taxa linear. Os testes numéricos apresentados em [1] apontam para uma dependência do desempenho do método com relação à escolhas do parâmetro η, uma vez que problemas que necessitaram de um número maior de iterações apresentaram-se mais sensíveis à escolha desse parâmetro. Buscando uma maior compreensão desse comportamento, apresentaremos um estudo desse método analisando o tempo de CPU, o número de avaliação de função e o número de iterações com relação ao parâmetro η.

Os experimentos numéricos utilizaram algumas funções disponibilizadas no endereço eletrônico `https://www.sfu.ca/~ssurjano/optimization.html`. Os códigos utilizados neste trabalho foram implementados na Linguagem de Programação Julia e estão livremente disponibilizados em `https://github.com/petimatematica/CMECG`.

Agradecimentos

O autor Wéllington Moutinho Dias agradece ao Programa de Educação Tutorial Institucional da Universidade Estadual do Sudoeste da Bahia PETI/UESB pela bolsa de estudos.

Referências

[1] Yanyun Ding, Yunhai Xiao e Jianwei Li. "A class of conjugate gradient methods for convex constrained monotone equations". Em: **Optimization** 66.12 (2017), pp. 2309–2328.

[2] Alexey Izmailov e Mikhail Solodov. **Otimização - volume 2: métodos computacionais**. Terceira Edição. IMPA, 2018. ISBN: 978-85-244-0454-2.