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**

# Limitação do parâmetro de penalidade em métodos de lagrangiano aumentado

Mariana da Rosa[1] Roberto Andreani[2]
IMECC, Campinas, SP
Leonardo Delarmelina Secchin[3]
UFES, São Mateus, ES

Neste trabalho consideramos o problema geral de programação não linear

$$\min \quad f(x) \quad \text{s.a} \quad x \in X := \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\}, \tag{PNL}$$

em que $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ e $h : \mathbb{R}^n \to \mathbb{R}^p$ são funções de classe $\mathcal{C}^2$.

Uma das ideias mais criativas e eficazes para resolver o problema (PNL) é trocá-lo por uma sequência de subproblemas mais simples (isto é, com restrições facilmente tratáveis ou mesmo sem restrições), que consistem em minimizar uma função que agrega um termo que *penaliza* os pontos inviáveis. Esses termos são controlados por um *parâmetro de penalização*, que aumenta a cada iteração de modo a recuperar viabilidade. O problema prático desta ideia é que, se os subproblemas são diferenciáveis, o parâmetro de penalização deve ser aumentado indefinidamente para que a viabilidade seja atingida. Isso provoca instabilidades numéricas, chegando ao ponto da convergência ocorrer para um ponto inviável.

Uma forma de contornar tal inconveniente é utilizar *funções de penalidade exatas diferenciáveis*, proposto inicialmente por Di Pillo e Grippo [1]. Para o (PNL) em que $g \equiv 0$ e $m \leq n$, os autores trabalham com a função lagrangiano aumentado:

$$f(x) + \lambda^T h(x) + \rho \|h(x)\|_2^2 + \|M(x)(\nabla f(x) + J_h(x)^T \lambda)\|_2^2, \tag{1}$$

onde $M(x)$ é uma matriz $q \times n$, $m \leq q \leq n$, com entradas de classe $\mathcal{C}^2$, e $J_h(x)$ denota a jacobiana de $h$ em $x$. Posteriormente, os próprios autores estenderam (1) para lidar com restrições de desigualdade. Em [2], um método para desigualdades variacionais baseado em (1) foi proposto. Assim como no trabalho inicial de Di Pillo e Grippo, em [2, 3] o multiplicador $\lambda$ em (1) é substituído por uma estimativa $\lambda(x)$ que depende de $x$, tornando (1) uma função apenas de $x$. Em outras palavras, o subproblema do método de lagrangiano aumentado resultante tem somente a variável primal $x$. A estimativa $\lambda^k = \lambda(x^k)$ é inspirada em [4], e consiste na solução de

$$\min_{\lambda} \|\nabla_x L(x^k, \lambda)\|_2^2 + \xi^2 \|H(x^k)\lambda\|_2^2, \tag{2}$$

onde $H(x) = \operatorname{diag}(h_1(x), \ldots, h_m(x))$ e $L(x, \lambda) = f(x) + \lambda^T h(x)$ é a função lagrangiano usual.

A vantagem da função lagrangiano aumentado (1) e sua extensão para desigualdades, além de ser uma penalização exata, é que possui derivadas contínuas. As estratégias de penalidade exata necessitam da hipótese de independência linear dos gradientes das restrições ativas (do inglês, LICQ). Porém, nos trabalhos citados essa hipótese é requerida para boa definição e implementação

[1]marianadarosa13@gmail.com
[2]andreani@unicamp.br
[3]leonardo.secchin@ufes.br

do método. De fato, sob LICQ, $\lambda = \lambda(x)$ é único e dado por [2]: $\lambda(x) = -N^{-1}(x)J_h(x)\nabla f(x)$, onde $N(x) = J_h(x)J_h(x)^T + \xi^2 H(x)^2$ é uma matriz positiva definida (resultado similar em [3] inclui restrições de desigualdade).

Isto faz que esta estratégia dependa de LICQ, uma condição de regularidade muito restritiva e que, em geral, não se cumpre em problemas reais ou em problemas teste, os quais verificam regularidades menos exigentes. Ainda mais, pode-se mostrar que mesmo $x \in X$ seja uma solução de (PNL), satisfazendo uma condição de qualificação mais fraca do que a exigida pelos autores, por exemplo a condição de posto constante (do inglês, CRCQ), o multiplicador obtido para as restrições de desigualdade pode ter entradas negativas.

Em nosso estudo, definimos uma nova estimativa baseada na resolução do problema

$$\min \quad \|\nabla_x L(x,\lambda,\mu)\|^2 + \zeta\left(\|G(x)\mu\|^2 + \|H(x)\lambda\|^2\right) \quad \text{s.a} \quad -\mu \leq 0, \quad \mu_i, \lambda_j \in [-C, C], \tag{3}$$

onde $i = 1,\ldots,m$, $j = 1,\ldots,p$, $G(x) = \text{diag}\,(g_1(x),\ldots,g_m(x))$, $H(x) = \text{diag}\,(h_1(x),\ldots,h_p(x))$, o qual pode ser reescrito de forma equivalente como em um problema de quadrados mínimos com restrições de caixa, que sempre admite um minimizador global, não necessariamente único.

A vantagem da estimativa obtida via (3) é que, para $C > 0$ suficientemente grande, qualquer solução desse problema é um multiplicador KKT associado a $x$ para (PNL), sem que nenhuma restrição de qualificação seja exigida. Além disso, sob a hipótese de Mangasarian-Fromovitz (MFCQ) em um ponto de acumulação viável para (PNL) temos o seguinte resultado:

**Teorema 0.1.** *Sejam $\{x^k\} \subset \mathbb{R}^n$ e $\{c_k\} \subset \mathbb{R}_+$ tais que $x^k \to \bar{x} \in X$ e, se $I_- := \{i \mid g_i(\bar{x}) < 0\} \neq \emptyset$, $c_k \geq C\gamma$, em que $\gamma = \min\{-g_i(\bar{x})^{-1} \mid i \in I_-\}$. Suponha que as soluções obtidas de (3) satisfazem*

$$\nabla_x L\left(x^k, \mu^k + c_k(g(x^k) + y_{c_k}(x^k)), \lambda^k + c_k h(x^k)\right) = 0, \quad \forall k \in \mathbb{N}$$

*onde $y_{c_k}(x^k) = \max\left\{0, -g(x^k) - \mu^k/c_k\right\}$. Se $\bar{x}$ satisfaz MFCQ, então $\bar{x}$ é um ponto KKT.*

Nesse trabalho, mostramos que sob hipóteses razoáveis, a estimativa dada em (3) permite gerar uma sequência de multiplicadores limitados, sem que $c$ seja acrescido indefinidamente. Em especial, isso sugere que, a partir de uma certa cota inferior para o parâmetro de penalização, o valor de $c$ pode ser mantido fixo de modo a evitar instabilidades numéricas.

## Agradecimentos

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo suporte financeiro (Projeto Temático - processo 2022/06745-7).

## Referências

[1] G. Di Pillo e L. Grippo. “A New Class of Augmented Lagrangians in Nonlinear Programming”. Em: **SIAM Journal on Control and Optimization** 17.5 (1979), pp. 618–628. DOI: 10.1137/0317044.

[2] A. André e P.J.S Silva. “Exact penalties for variational inequalities with applications to nonlinear complementarity problems”. Em: **Computational Optimization and Applications** 47.3 (2009), pp. 401–429. DOI: 10.1007/s10589-008-9232-3.

[3] R. Andreani, E.H. Fukuda e P.J.S Silva. “A Gauss-Newton Approach for Solving Constrained Optimization Problems Using Differentiable Exact Penalties”. Em: **Journal of Optimization Theory and Applications** 156.2 (2012), pp. 417–449. DOI: 10.1007/s10957-012-0114-6.

[4] T. Glad e E. Polak. “A multiplier method with automatic limitation of penalty growth”. Em: **Mathematical Programming** 17.1 (1979), pp. 140–155. DOI: 10.1007/bf01588240.

 