São José dos Campos - SP
LV SBPO
LV Simpósio Brasileiro de Pesquisa Operacional
São José dos Campos, SP - 6 a 9 de novembro de 2023

Sobre o Parâmetro de Penalidade em Métodos de Lagrangiano Aumentado

Mariana da Rosa
Universidade Estadual de Campinas
Rua Sérgio Buarque de Holanda, 651, Campinas, SP, Brasil
marianadarosa13@gmail.com

Roberto Andreani
Universidade Estadual de Campinas
Rua Sérgio Buarque de Holanda, 651, Campinas, SP, Brasil
andreani@unicamp.br

Leonardo Delarmelina Secchin
Universidade Federal do Espirito Santo
Rodovia Governador Mário Covas, Km 60 - Bairro Litorâneo, São Mateus - ES
leonardo.secchin@ufes.br

RESUMO

Métodos de lagrangiano aumentado são ferramentas eficazes para a resolução de problemas de otimização com restrições gerais. Um dos inconvenientes, porém, é o aumento excessivo do parâmetro de penalização, o que pode gerar instabilidades numéricas, sobretudo em problemas não convexos. Uma das estratégias que visam minimizar este problema é o uso de penalidades exatas. No entanto, as funções de penalização exata usuais não são diferenciáveis, o que na prática dificulta a utilização de algoritmos eficientes. Neste trabalho, apresentamos um novo resultado que, com hipóteses mais fracas do que aquelas apresentadas na literatura, permite manter o parâmetro de penalização limitado.

PALAVRAS CHAVE. Otimização Contínua, Métodos de Lagrangiano Aumentado.

Tópicos (Otimização)

ABSTRACT

Augmented Lagrangian methods are effective tools for solving optimization problems with general constraints. One of the drawbacks, however, is the excessive increase of the penalty parameter, which can generate numerical instabilities, especially in problems not convex. One of the strategies that aim to minimize this problem is the use of exact penalties. However, the usual exact penalty functions are not differentiable, which in practice makes it difficult to use efficient algorithms. In this work, we present a new result that, with weaker hypotheses than those presented in the literature, allows keeping the penalty parameter limited.

KEYWORDS. Continuous Optimization, Augmented Lagrangian Methods.

Paper Topics (Optimization)

http://live-proceedings.pantheonsite.io/p/175218

1. Introdução

Neste trabalho, consideramos o problema geral de programação não linear

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

onde $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ 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 a 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.

2. Funções de penalidade exatas diferenciáveis

Uma forma de contornar tal inconveniente é utilizar *funções de penalidade exatas diferenciáveis*, proposto inicialmente por [Di Pillo e Grippo, 1979]. 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 [de André e Silva, 2009], um método para desigualdades variacionais baseado em (1) foi proposto. Assim como no trabalho inicial de Di Pillo e Grippo, em [Andreani et al., 2012; de André e Silva, 2009] o multiplicador λ 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 [Glad e Polak, 1979] 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 $H(x) = \text{diag}\,(h_1(x), \ldots, h_m(x))$ 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 do método. De fato, sob LICQ, $\lambda = \lambda(x)$ é único e dado por [de André e Silva, 2009]: $\lambda(x) = -N^{-1}(x) J_h(x) \nabla f(x)$, onde $N(x) = J_h(x) J_h(x)^T + \xi^2 H(x)^2$ é uma matriz positiva definida (resultado similar em [Andreani et al., 2012] 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, se x é um ponto KKT de (PNL), para $C > 0$ suficientemente grande, qualquer solução desse problema é um multiplicador KKT associado a x, 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), podemos provar o resultado abaixo, considerando $\{x^k\} \subset \mathbb{R}^n$ e $\{c_k\} \subset \mathbb{R}_+$, tais que $x^k \to \bar{x} \in X$, definindo $I_- := \{i \mid g_i(\bar{x}) < 0\} \neq \emptyset$, $c_k \geq C\gamma$, em que $\gamma = \min\{-g_i(\bar{x})^{-1} \mid i \in I_-\}$, e ainda supondo 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.

3. Conclusão

Neste trabalho, mostramos que sob hipóteses razoáveis, a estimativa dada em (3) permite gerar uma sequência de multiplicadores limitados, sem que o parâmetro de penalização 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.

Referências

Andreani, R., Fukuda, E. H., e Silva, P. J. S. (2012). A Gauss-Newton approach for solving constrained optimization problems using differentiable exact penalties. *Journal of Optimization Theory and Applications*, 156(2):417–449.

de André, T. A. e Silva, P. J. S. (2009). Exact penalties for variational inequalities with applications to nonlinear complementarity problems. *Computational Optimization and Applications*, 47(3): 401–429.

Di Pillo, G. e Grippo, L. (1979). A new class of augmented Lagrangians in nonlinear programming. *SIAM Journal on Control and Optimization*, 17(5):618–628.

Glad, T. e Polak, E. (1979). A multiplier method with automatic limitation of penalty growth. *Mathematical Programming*, 17(1):140–155.