Journal of Computational and Applied Mathematics 229 (2009) 129–144

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Log-Sigmoid nonlinear Lagrange method for nonlinear optimization problems over second-order cones

Jian Gu *, Liwei Zhang, Xiantao Xiao

Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China

ARTICLE INFO

Article history:
Received 27 January 2008
Received in revised form 6 October 2008

MSC:
90C26
90C30

Keywords:
Log-Sigmoid function
Nonlinear Lagrangian method
Nonlinear second-order cone programming
Dual algorithm
Rate of convergence

ABSTRACT

This paper analyzes the rate of local convergence of the Log-Sigmoid nonlinear Lagrange method for nonconvex nonlinear second-order cone programming. Under the componentwise strict complementarity condition, the constraint nondegeneracy condition and the second-order sufficient condition, we show that the sequence of iteration points generated by the proposed method locally converges to a local solution when the penalty parameter is less than a threshold and the error bound of solution is proportional to the penalty parameter. Finally, we report numerical results to show the efficiency of the method.

1. Introduction

In this paper, we consider the nonconvex nonlinear second-order cone programming problem of the form

$$\text{(NLSOP)} \quad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & h(x) = 0, \\ & g^j(x) \succeq_{Q_{m_j+1}} 0, \quad j = 1, 2, \ldots, J, \end{array} \tag{1.1}$$

where $f : \Re^n \to \Re$, $h : \Re^n \to \Re^l$ and $g^j : \Re^n \to \Re^{m_j+1}, j = 1, 2, \ldots, J$ are twice continuously differentiable. The second-order cone (or ice-cream cone, or Lorentz cone) of dimension $m + 1$ is defined by

$$Q_{m+1} := \{x = (x_0; \bar{x}) \in \Re^{m+1} : \|\bar{x}\| \le x_0\},$$

and the order relation $\succeq_{Q_{m+1}}$ induced by Q_{m+1} is given by

$$x \succeq_{Q_{m+1}} 0 \quad \text{if and only if} \quad x \in \Re^{m+1}, \|\bar{x}\| \le x_0.$$

The interior of the cone Q_{m+1}, denoted by int Q_{m+1}, is the set of $x \in \Re^{m+1}$ such that $x_0 > \|\bar{x}\|$. In that case, we say that $x \succ_{Q_{m+1}} 0$ for $x \in \text{int}\, Q_{m+1}$. The boundary of Q_{m+1}, denoted by ∂Q_{m+1}, is the set of $x \in \Re^{m+1}$ such that $x_0 = \|\bar{x}\|$.

Let $Q := Q_{m_1+1} \times Q_{m_2+1} \times \cdots \times Q_{m_J+1}$. Denote

$$g(x) := (g^1(x); g^2(x); \ldots; g^J(x)) \in \Re^q, \quad g^j(x) \in \Re^{m_j+1},$$
$$u := (u_1; u_2; \ldots; u_J) \in \Re^q, \quad u_j \in \Re^{m_j+1},$$

* Corresponding author.
E-mail address: gujian82@yahoo.cn (J. Gu).

0377-0427/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2008.10.016

where $q := \sum_{j=1}^{J}(m_j + 1)$. Then the problem (NLSOP) can be expressed as

$$\text{minimize } f(x) \quad \text{subject to } h(x) = 0, \quad g(x) \in Q.$$

The standard (linear) Lagrangian of (NLSOP) is defined by

$$L(x, \zeta, u) = f(x) - \langle \zeta, h(x) \rangle - \sum_{j=1}^{J} \langle u_j, g^j(x) \rangle, \tag{1.2}$$

which plays an important role in describing the optimality conditions for second-order cone optimization problem (1.1) and designing algorithms for solving (1.1). For convex programming, the saddle point theory can be established in terms of the standard Lagrangian and dual algorithms based on solving minimizing $L(x, \zeta_k, u_k)$ can be developed as well, where (ζ_k, u_k) is the estimate multiplier at the kth iteration. For nonconvex nonlinear programming, $L(x, \zeta, u)$ is usually not convex even for (ζ_k, u_k) being near to (ζ^*, u^*) and x in a neighborhood of x^*, where (x^*, ζ^*, u^*) is a Karush–Kuhn–Tucker point to the optimization problem, and this leads to difficulties in numerical implementations. To solve this problem, many scholars pay much attention to studying the variants of the standard Lagrangian. The augmented Lagrangian method was initiated in [10, 19] for solving nonlinear programming with only equality constraints and was generalized in [20–22] to include inequality constrained problems. For more details on proximal augmented Lagrangian method we refer to [1] or [2]. Besides these, Polyak and his collaborators have developed many nonlinear Lagrangians for solving nonlinear programming problems, for instance, see [8,14–18]. Among them, Polyak (2001) [16] constructed a nonlinear Lagrangian based on the Log-Sigmoid function for solving nonconvex NLP problems.

In this paper, we focus on the study of the following nonlinear Lagrangian for (NLSOP):

$$G(x, \zeta, u, t) = f(x) - \langle \zeta, h(x) \rangle + (2t)^{-1} \|h(x)\|^2 + t \sum_{j=1}^{J} \langle \psi_{LS}(-t^{-1} g^j(x)), u_j \rangle, \tag{1.3}$$

where

$$\psi_{LS}(w) = 2(\ln(1 + \mathrm{e}^{\lambda_1^w}) - \ln 2) c_1^w + 2(\ln(1 + \mathrm{e}^{\lambda_2^w}) - \ln 2) c_2^w \tag{1.4}$$

is the Löwner operator associated Log-Sigmoid function $\psi_{LS}(s) = 2(\ln(1 + \mathrm{e}^s) - \ln 2)$ for $s \in \Re$. We will extend the results in [16] to the case of nonlinear optimization problems with second-order cone constraints and equality constraints.

We should point out that, for the proximal augmented Lagrange method, Liu and Zhang [12] discussed the rate of convergence for nonconvex semidefinite programming when the strict complementarity is satisfied and Liu and Zhang [13] studied its rate of convergence without strict complementarity condition.

This paper is organized as follows. In Section 2, we discuss properties of the Log-Sigmoid Löwner operator. In Section 3, we introduce a set of basic assumptions needed for the convergence analysis and discuss properties of the nonlinear Lagrangian $G(x, \zeta, u, t)$. In Section 4, we focus on analyzing the rate of convergence of the Log-Sigmoid nonlinear Lagrange method under the given conditions. Finally, in Section 5, we report numerical results implemented by the Log-Sigmoid nonlinear Lagrange method.

The following notations and terminologies are used throughout the paper. If F is differentiable at $(x, y) \in \Re^n \times \Re^m$, then we use $\mathcal{J}F(x, y)$ (respectively, $\mathcal{J}_x F(x, y)$) to denote the derivative of F at (x, y) (respectively, the partial derivative of F at (x, y) with respect to x) and $\nabla F(x, y) := \mathcal{J}F(x, y)^*$ the adjoint of $\mathcal{J}F(x, y)$ (respectively, $\nabla_x F(x, y) := \mathcal{J}_x F(x, y)^*$ the adjoint of $\mathcal{J}_x F(x, y)$). Moreover, if F is twice differentiable at $(x, y) \in \Re^n \times \Re^m$, we define

$$\nabla^2 F(x, y) := \mathcal{J}(\nabla F)(x, y), \quad \nabla^2_{xx} F(x, y) := \mathcal{J}_x(\nabla_x F)(x, y).$$

For any index set $I \in \{1, 2, \ldots, J\}$, we denote by $\mathrm{diag}(A_j)_{j \in I}$ the block diagonal matrix where its diagonal entries A_j are arranged in the increasing order of $j \in I$. For any $x, y \in \Re^n$, the Euclidean inner product and norm are denoted by $\langle x, y \rangle = x^{\mathrm{T}} y$, $\|x\| = \sqrt{x^{\mathrm{T}} x}$, respectively. For any two matrices C and D in $\Re^{m \times n}$, we write

$$\langle C, D \rangle := \mathrm{Tr}(C^{\mathrm{T}} D), \qquad \|C\| = \sqrt{\mathrm{Tr}(C^{\mathrm{T}} C)}$$

for the Frobenius inner product between C and D and the Frobenius norm, respectively, where "Tr" denotes the trace of a square matrix.

2. The Log-Sigmoid Löwner operator

For any $u = (u_0; \bar{u})$, $v = (v_0; \bar{v})$ in $\Re^{m+1}$, we define their Jordan product as

$$u \circ v = (u^{\mathrm{T}} v;\ v_0 \bar{u} + u_0 \bar{v}).$$

It is easy to check that $\mathbb{e} := (1; 0) \in \Re^{m+1}$ is a unit element satisfying $u \circ \mathbb{e} = \mathbb{e} \circ u = u$ for any $u \in \Re^{m+1}$. Then $(\Re^{m+1}, \circ)$ becomes a Jordan algebra, see [6]. For $u \in \Re^{m+1}$, its spectral decomposition is

$$u = \lambda_1^u c_1^u + \lambda_2^u c_2^u, \tag{2.1}$$