Operations Research Letters 44 (2016) 336–341

Contents lists available at ScienceDirect

Operations Research Letters

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

Error bounds for rank constrained optimization problems and applications

Shujun Bi [a], Shaohua Pan [b,*]

[a] State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Science, Beijing, PR China

[b] Department of Mathematics, South China University of Technology, Tianhe District, Guangzhou 510641, PR China

ARTICLE INFO

Article history:
Received 31 May 2015
Received in revised form
3 March 2016
Accepted 4 March 2016
Available online 11 March 2016

Keywords:
Rank constrained optimization
Error bounds
Calmness
Exact penalty

ABSTRACT

For the rank constrained optimization problem whose feasible set is the intersection of the rank constraint set $\mathcal{R} = \{X \in \mathbb{X} \mid \mathrm{rank}(X) \le \kappa\}$ and a closed convex set Ω, we establish the local (global) Lipschitzian type error bounds for estimating the distance from any $X \in \Omega$ ($X \in \mathbb{X}$) to the feasible set and the solution set, under the calmness of a multifunction associated to the feasible set at the origin, which is satisfied by three classes of common rank constrained optimization problems.

1. Introduction

Let $\mathbb{X}$ denote the vector space $\mathbb{R}^{n_1\times n_2}$ of all $n_1 \times n_2$ real matrices or the vector space $\mathbb{H}^n$ of all $n \times n$ Hermitian matrices, both endowed with the trace inner product $\langle\cdot,\cdot\rangle$ and its induced norm $\|\cdot\|_F$. Given a positive integer κ and a suitable continuous loss function $f : \mathbb{X} \to \mathbb{R}$, we are concerned with the following rank constrained optimization problem

$$\min_{X\in\mathbb{X}} \left\{ f(X) \mid \mathrm{rank}(X) \le \kappa,\ X \in \Omega \right\}, \tag{1}$$

where Ω is a convex compact subset of $\mathbb{X}$. Such a problem has many applications in a host of fields including statistics, signal and image processing, system identification and control, collaborative filtering, quantum tomography, finance, and so on (see, e.g., [4,6,16,20,22,28]). In the sequel, we denote by $\mathcal{F}$ the feasible set of (1) and assume that $\mathcal{F} \neq \emptyset$, which implies that the solution set of (1), denoted by $\mathcal{F}^*$, is nonempty.

A common way to deal with the NP-hard problem (1) is to adopt convex relaxation technique, which yields a desirable local optimal even feasible solution by solving a single or a sequence of tractable convex optimization problems. The popular nuclear norm convex relaxation method proposed in [4] belongs to the single-stage convex relaxation class, which received active research in the past several years from many fields such as optimization, statistics, information, computer science, and so on (see, e.g., [2,7,10,20,24]). However, when the set Ω is characterizing some structure conflicted with low rank, such as the correlation or density matrix structure, the nuclear norm relaxation method will fail in yielding a low rank solution. In view of this, many researchers recently develop effective solution methods based on the sequential convex relaxation models arising from the penalty problems [6,11], the nonconvex surrogate problems [5,9,17,18], and the rank constrained optimization problem itself [15,23,25]. We notice that to measure the distance from any given point to the feasible set or the solution set plays a key role in the analysis of these methods. Motivated by this, we in this work take the first step towards the study on Lipschitzian type error bounds for (1).

In this paper, we show that the calmness of a multifunction associated to the feasible set $\mathcal{F}$ at the origin is a sufficient and necessary condition for the local Lipschitzian error bounds to estimate the distance from any $X \in \Omega$ to $\mathcal{F}$, which is specially satisfied by three classes of common rank constrained optimization problems (1) where Ω is a ball set, a density matrix set or a correlation matrix set, and under this condition derive the global error bound for estimating the distance from any $X \in \mathbb{X}$ to $\mathcal{F}$. In addition, under an additional mild assumption for the objective function f, we also establish the local (global) Lipschitzian error bounds for estimating the distance from any $X \in \Omega$ ($X \in \mathbb{X}$) to the solution set $\mathcal{F}^*$. To the best of our knowledge, this paper is the first one to study the Lipschitzian type error bounds for low-rank optimization problems, though there are many works on error bounds for the system

* Corresponding author.
E-mail addresses: bishujun@lsec.cc.ac.cn (S. Bi), shhpan@scut.edu.cn (S. Pan).

http://dx.doi.org/10.1016/j.orl.2016.03.002
0167-6377/

of linear inequalities and (nondifferentiable) convex inequalities (see, e.g., [12–14,19,27] and references therein). To illustrate the potential applications of the derived error bounds, we show that the penalty problem

$$\min_{X\in\Omega}\Big\{f(X)+\rho\sum_{i=\kappa+1}^{n}\sigma_i(X)\Big\} \tag{2}$$

is exact in the sense that its global optimal solution set coincides with that of (1) when the penalty parameter ρ is over a certain threshold. This does not only affirmatively answer the open question proposed in [6] about whether the penalty problem (32) there is exact or not for the rank constrained correlation matrix problem, but also provides a platform for designing convex relaxation algorithms for (1).

To close this section, we introduce some notations used in this paper. We use $\mathbb{H}_+^n$ to denote the cone of Hermitian positive semidefinite matrices. For any $X\in\mathbb{H}^n$, we assume that X has the eigenvalue decomposition as $X=\sum_{i=1}^n\lambda_i(X)u_iu_i^{\mathbb{T}}$ where $\lambda_1(X)\geq\cdots\geq\lambda_n(X)$ and all u_i are complex orthonormal column vectors. For any $X\in\mathbb{R}^{n_1\times n_2}$, we assume that X has the singular value decomposition (SVD) as $X=\sum_{i=1}^n\sigma_i(X)u_iv_i^{\mathbb{T}}$, where $\sigma_1(X)\geq\cdots\geq\sigma_n(X)$ with $n=\min(n_1,n_2)$, and all $u_i\in\mathbb{R}^{n_1}$ and $v_i\in\mathbb{R}^{n_2}$ are orthonormal column vectors. We denote by $\|X\|_*$ the nuclear norm of $X\in\mathbb{X}$. For a closed subset $\mathcal{S}\subseteq\mathbb{X}$, $\Pi_{\mathcal{S}}(X)$ means the projection of X onto the set $\mathcal{S}$.

2. Lipschitzian type error bounds

Let $\mathbb{Y}$ and $\mathbb{Z}$ be two finite dimensional vector spaces equipped with the inner product $\langle\cdot,\cdot\rangle$ and its induced norm $\|\cdot\|$. Recall that a multifunction $\Upsilon:\mathbb{Y}\rightrightarrows\mathbb{Z}$ is calm at $\bar y$ for $\bar z\in\Upsilon(\bar y)$ if there exist a constant $\alpha\geq 0$ and neighborhoods $\mathcal{U}$ of $\bar y$ and $\mathcal{V}$ of $\bar z$ such that

$$\Upsilon(y)\cap\mathcal{V}\subseteq\Upsilon(\bar y)+\alpha\|y-\bar y\|\mathbb{B}_{\mathbb{Z}}\quad\text{for all } y\in\mathcal{U},$$

where $\mathbb{B}_{\mathbb{X}}$ denotes a closed unit ball of the space $\mathbb{Z}$ centered at the origin. By [3, Exercise 3H.4], we know that there is no need at all to mention a neighborhood $\mathcal{U}$ of $\bar y$ in the description of calmness, i.e., the following equivalent description holds.

Lemma 2.1. *For a multifunction $\Upsilon:\mathbb{Y}\rightrightarrows\mathbb{Z}$, the calmness of Υ at $\bar y$ for $\bar z\in\Upsilon(\bar y)$ is equivalent to the existence of a constant $\alpha\geq 0$ and a neighborhood $\mathcal{V}$ of $\bar z$ such that*

$$\Upsilon(y)\cap\mathcal{V}\subseteq\Upsilon(\bar y)+\alpha\|y-\bar y\|\mathbb{B}_{\mathbb{Z}}\quad\textit{for all } y\in\mathbb{Y},$$

or the existence of a constant $\alpha\geq 0$ and a neighborhood $\mathcal{V}$ of $\bar z$ such that

$$\mathrm{dist}(z,\Upsilon(\bar y))\leq\alpha\,\mathrm{dist}(\bar y,\Upsilon^{-1}(z))\quad\textit{for all } z\in\mathcal{V}.$$

Let $\Gamma:\mathbb{R}\rightrightarrows\mathbb{X}$ be a multifunction associated to the feasible set of problem (1) defined by

$$\Gamma(\omega):=\Big\{X\in\Omega\mid\sum_{i=\kappa+1}^{n}\sigma_i(X)=\omega\Big\}\quad\text{for }\omega\in\mathbb{R}. \tag{3}$$

In this section, we study the Lipschitzian error bounds for estimating the distance to the feasible set $\mathcal{F}$ and the solution set $\mathcal{F}^*$, respectively, under the calmness of Γ at 0.

2.1. Error bounds for the feasible set $\mathcal{F}$

First of all, we show that the distance from any $Z\in\Omega$ to the feasible set $\mathcal{F}$ can be bounded above by $\sum_{i=\kappa+1}^n\sigma_i(Z)$ iff the multifunction Γ is calm at 0 for each $X\in\Gamma(0)$.

Theorem 2.1. *The multifunction Γ defined by (3) is calm at 0 for each $X\in\Gamma(0)$ if and only if there exists a constant $c>0$ such that*

$$\mathrm{dist}(Z,\Gamma(0))\leq c\,\mathrm{dist}(0,\Gamma^{-1}(Z))=c\sum_{i=\kappa+1}^{n}\sigma_i(Z)\quad\textit{for all } Z\in\Omega. \tag{4}$$

Proof. "$\Longrightarrow$". By the calmness of Γ at 0 for each $X\in\Gamma(0)=\Omega\cap\mathcal{R}$ and Lemma 2.1, it follows that for each $X\in\Omega\cap\mathcal{R}$, there exist constants $\alpha(X)\geq 0$ and $\epsilon(X)>0$ such that

$$\mathrm{dist}(Y,\Gamma(0))\leq\alpha(X)\,\mathrm{dist}(0,\Gamma^{-1}(Y))\quad\forall\,Y\in\mathbb{B}(X,\epsilon(X)), \tag{5}$$

where $\mathbb{B}(X,\epsilon(X))$ is a closed ball of radius $\epsilon(X)$ centered at X. Notice that the compact set $\Omega\cap\mathcal{R}$ is covered by the set $\bigcup_{X\in\Omega\cap\mathcal{R}}(X+\frac{\epsilon(X)}{2}\mathbb{B}_{\mathbb{X}}^\circ)$, where $\mathbb{B}_{\mathbb{X}}^\circ$ denotes the open unit ball around the origin in $\mathbb{X}$. By the Heine–Borel theorem, there exist a finite number of points $X^1,X^2,\ldots,X^m\in\Omega\cap\mathcal{R}$ such that $\Omega\cap\mathcal{R}\subseteq\bigcup_{i=1}^m(X^i+\frac{\epsilon(X^i)}{2}\mathbb{B}_{\mathbb{X}}^\circ)$. Write

$$\bar\epsilon:=\min\{\epsilon(X^1),\ldots,\epsilon(X^m)\}\quad\text{and}\quad \bar\alpha:=\max\{\alpha(X^1),\ldots,\alpha(X^m)\}.$$

Let Z be an arbitrary point from Ω. We proceed the arguments by two cases as below.

Case 1: $\mathrm{dist}(Z,\Omega\cap\mathcal{R})\leq\bar\epsilon/2$. Since the set $\Omega\cap\mathcal{R}$ is closed, there must exist $\bar Z\in\Omega\cap\mathcal{R}$ such that $\|Z-\bar Z\|_F\leq\bar\epsilon/2$. Since $\bar Z\in\Omega\cap\mathcal{R}$, there exists a $k\in\{1,2,\ldots,m\}$ such that $\|\bar Z-X^k\|_F<\epsilon(X^k)/2$. Consequently, $\|Z-X^k\|_F\leq\|Z-\bar Z\|_F+\|\bar Z-X^k\|_F\leq\epsilon(X^k)$. Together with (5), $\mathrm{dist}(Z,\Gamma(0))\leq\bar\alpha\,\mathrm{dist}(0,\Gamma^{-1}(Z))$. This shows that (4) holds with $c=\bar\alpha$.

Case 2: $\mathrm{dist}(Z,\Omega\cap\mathcal{R})>\bar\epsilon/2$. Now there must exist an $\eta>0$ such that $\sum_{i=\kappa+1}^n\sigma_i(Y)\geq\eta$ for all $Y\in\Omega$ with $\mathrm{dist}(Y,\Omega\cap\mathcal{R})>\bar\epsilon/2$. If not, one may select a sequence $\{Z^k\}\subseteq\Omega$ with $\mathrm{dist}(Z^k,\Omega\cap\mathcal{R})>\bar\epsilon/2$ such that $\sum_{i=\kappa+1}^n\sigma_i(Z^k)\leq\eta^k$ for all k, where $\{\eta^k\}$ is a sequence of positive numbers with $\lim_{k\to+\infty}\eta^k=0$. Since Ω is compact, we without loss of generality assume that $\{Z^k\}$ converges to $Z^*\in\Omega$. Then, from the locally Lipschitz continuity of $\sigma_i(\cdot)$, it follows that $\sum_{i=\kappa+1}^n\sigma_i(Z^*)\leq 0$, and then $Z^*\in\Omega\cap\mathcal{R}$. On the other hand, from $\mathrm{dist}(Z^k,\Omega\cap\mathcal{R})>\bar\epsilon/2$ for all k, we have $\mathrm{dist}(Z^*,\Omega\cap\mathcal{R})>\bar\epsilon/2$. Thus, we obtain a contradiction, and the above statement holds. Since Ω is bounded, it follows that $\mathrm{dist}(\cdot,\Omega\cap\mathcal{R})$ is bounded above on Ω, say, by some $M>0$. Thus, for all $Z\in\Omega$ with $\mathrm{dist}(Z,\Omega\cap\mathcal{R})>\bar\epsilon/2$, one has that $\mathrm{dist}(Z,\Omega\cap\mathcal{R})\leq M\leq(M/\eta)\sum_{i=\kappa+1}^n\sigma_i(Z)$. By taking $c=M/\eta$, the desired inequality (4) then follows.

"$\Longleftarrow$. Let X be an arbitrary point from $\Gamma(0)$ and $\epsilon\in(0,1)$ be an arbitrary constant. By Lemma 2.1, we only need to argue that there must exist a constant $c'>0$ such that

$$\mathrm{dist}(Z,\Gamma(0))\leq c'\,\mathrm{dist}(0,\Gamma^{-1}(Z))\quad\forall Z\in\mathbb{B}(X,\epsilon). \tag{6}$$

Indeed, for any $Z\in\mathbb{B}(X,\epsilon)$, if $Z\notin\Omega$, then $\Gamma^{-1}(Z)=\emptyset$ by noting that $\mathrm{dom}\,\Gamma^{-1}\subseteq\Omega$, and inequality (6) holds for any $c'>0$; if $Z\in\Omega$, then by taking $c'=c$, inequality (6) follows directly from (4). Until now, the proof is completed. □

We next illustrate that the sufficient and necessary condition in Theorem 2.1 is especially satisfied by three classes of common rank constrained optimization problems.

(1) **Rank constrained optimization problems over a ball**. The feasible set of this class of rank constrained optimization problems takes the following form

$$\mathcal{F}:=\big\{X\in\mathbb{R}^{n_1\times n_2}\mid\mathrm{rank}(X)\leq\kappa,\ \|X\|\leq\gamma\big\}, \tag{7}$$