运筹学与控制论

DOI:10.11721/cqnuj20150617

非凸半定规划的鞍点存在性研究*

李永玲, 罗洪林, 向彦宁
(重庆师范大学 数学学院, 重庆 401331)

摘要: 主要利用矩阵分析的谱分解、Frobenius 内积及其相关性质,凸分析的凸集分离定理来研究非凸半定规划问题的鞍点的存在性,通过 3 种不同的方式给出并证明了鞍点存在的一些充分、必要以及充分必要条件。首先,利用一个不等式系统给出了与文献[1]中的对偶定理等价的一个鞍点存在的充分必要条件。然后,给出了广义的 KKT 条件,并在不变凸性的假设下,证明了广义 KKT 条件是鞍点存在的一个充分条件;若 $\bar{x}\in \operatorname{int} C$,则广义 KKT 条件是鞍点存在的一个必要条件。最后,定义了一个扰动函数 ν,并在非凸半定规划问题的最优解存在的假设下,利用此扰动函数给出了鞍点存在的一个充分必要条件:若非凸半定规划问题的最优解存在,则对偶可达且无对偶间隙等价于扰动函数 ν 的上图在点 $(0,\nu(0))$ 处存在支撑超平面。

关键词: 非凸半定规划;鞍点;广义 KKT 条件;不变凸

中图分类号: O224 **文献标志码:** A **文章编号:** 1672-6693(2015)06-0009-06

文中用 $\mathbf{R}^m$, S^n, S^n_+ 分别表示 m 维向量空间,n 阶对称矩阵空间及 n 阶半定矩阵锥。考虑如下形式的非凸半定规划问题:

$$\begin{aligned}&\min \quad f(x)\\&\text{s.t.} \quad g(x)\leqslant 0, G(x)\leqslant 0, x\in C\subset \mathbf{R}^m。\end{aligned}\tag{1}$$

其中函数 $f:\mathbf{R}^m\to\mathbf{R}$, $g:\mathbf{R}^m\to\mathbf{R}^k$, $G:\mathbf{R}^m\to S^n$ 均不要求是凸的,且 $G(x)\leqslant 0$ 当且仅当 $-G(x)\in S^n_+$。

鞍点在优化问题的对偶理论以及最优性条件的讨论中扮演着重要的角色。鞍点可以把原问题与对偶问题的最优解以及其最优值紧密联系起来,因此鞍点存在性的研究对于优化问题是很重要的。

文献[1]中给出了对偶定理,事实上即为鞍点存在的一个充分必要条件。此文献在 Slater 条件成立,以及类凸(预不变凸)性的假设下,证明了鞍点存在的一个充分条件。文献中还研究了问题(1)的一种特殊情况,即把"$g(x)\leqslant 0$"退化为等式约束即"$g(x)=0$",此时则不满足 Slater 条件,文献中也给出了对应的鞍点存在的充分必要条件。文献[2]中证明了鞍点存在的一个充分必要条件,并在凸性和 Slater 条件的假设下给出了鞍点存在的一个充分条件。事实上,文献[2]中的非线性半定规划问题是问题(1)的特殊情况[1,3],因此本文给出的部分结论比文献[2]中的结论更一般,所以文献[2]中的部分结论是本文结论的特殊情况。

在一般的非线性规划问题中,无论是在最优性条件还是鞍点存在性的研究,Karus-Kuhn-Tucker 条件(简称 KKT 条件)都起着重要的作用[4]。在非线性半定规划一个局部最优解处满足 Robinson 约束品性的前提下,文献[5]中证明了以下条件是等价的:强二阶充分条件和约束非退;KKT 系统的 Clarke's Jacobian 非奇异;KKT 点满足强正则性。文献[6]中也给出了在不同的假设条件下 KKT 点是广义方程的强正则解。此外,从计算的角度来看,由于 KKT 系统的求解易在计算机上实现,因此本文给出了所谓的广义 Karus-Kuhn-Tucker 条件(6)式,简称广义 KKT 条件。本文讨论了鞍点存在性与广义 KKT 条件的关系,证明了在不变凸性的假设下,广义 KKT 条件是鞍点存在的充分条件。还给出了与文献[1]中的对偶定理等价的一个定理即鞍点存在的另一个充分必要

* **收稿日期:** 2014-11-24 **修回日期:** 2015-04-07 **网络出版时间:** 2015-9-28 12:02
资助项目: 国家自然科学基金(No. 11431004)
作者简介: 李永玲,女,研究方向为最优化理论与算法,半定规划,E-mail:1152280286@qq.com;**通信作者:** 罗洪林,副教授,E-mail:1071025013@fudan.edu.cn
网络出版地址: http://www.cnki.net/kcms/detail/50.1165.n.20150928.1202.022.html

条件,这一充分必要条件由方程组的形式给出易在计算机上求解。引入扰动函数,并利用扰动函数来刻画鞍点存在的一个充分必要条件。

本文在第1节中给出了一些记号与定义,在第2节中讨论了非凸半定规划问题(1)的鞍点存在的一些充分、必要以及充分必要条件。第3节为总结部分。

1 预备知识与记号

这一节将介绍一些定义、性质与符号。在对称矩阵空间 S^n 中可定义内积为 $\boldsymbol{A}\cdot\boldsymbol{B}=\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}B_{ij}=\mathrm{tr}(\boldsymbol{AB})$, $\forall \boldsymbol{A},\boldsymbol{B}\in S^n$,其中 tr 表示矩阵的迹。基于此可以定义 $\boldsymbol{A}\cdot\boldsymbol{C}=\begin{pmatrix}\boldsymbol{A}\cdot\boldsymbol{C}_{11} & \cdots & \boldsymbol{A}\cdot\boldsymbol{C}_{1q}\\ \boldsymbol{A}\cdot\boldsymbol{C}_{21} & \cdots & \boldsymbol{A}\cdot\boldsymbol{C}_{2q}\\ \vdots & \vdots & \vdots\\ \boldsymbol{A}\cdot\boldsymbol{C}_{p1} & \cdots & \boldsymbol{A}\cdot\boldsymbol{C}_{pq}\end{pmatrix}$,其中 $\boldsymbol{A}\in S^n$, $\boldsymbol{C}=\begin{pmatrix}\boldsymbol{C}_{11} & \cdots & \boldsymbol{C}_{1q}\\ \boldsymbol{C}_{21} & \cdots & \boldsymbol{C}_{2q}\\ \vdots & \vdots & \vdots\\ \boldsymbol{C}_{p1} & \cdots & \boldsymbol{C}_{pq}\end{pmatrix}$, $\boldsymbol{C}_{ij}\in S^n$, $\boldsymbol{A}\cdot\boldsymbol{C}_{ij}$ 为对称矩阵空间 S^n 中定义的内积。

用 $\boldsymbol{Z}$ 表示 $\mathbf{R}^i$(S^j 或 $\mathbf{R}^i\times S^j$)时,其中 i,j 可以为任意的正整数,当且仅当 $\boldsymbol{Z}$ 表示 $\mathbf{R}^i$(S^j 或 $\mathbf{R}^i\times S^j$)时,用 $\boldsymbol{Z}_+$ 表示 $\mathbf{R}^i_+$(S^j_+ 或 $\mathbf{R}^i_+\times S^j_+$)。在此基础上可定义如下的偏序关系:

$$\boldsymbol{A}\geqslant\boldsymbol{B}\Leftrightarrow\boldsymbol{A}-\boldsymbol{B}\in\boldsymbol{Z}_+,\ \mathbf{A}>\mathbf{B}\Leftrightarrow\mathbf{A}-\boldsymbol{B}\in\mathrm{int}\,\boldsymbol{Z}_+,\ \forall\boldsymbol{A},\boldsymbol{B}\in\boldsymbol{Z}。$$

若 $\boldsymbol{A}\geqslant0,\boldsymbol{B}\geqslant0$,则有 $\boldsymbol{A}\cdot\boldsymbol{B}\geqslant0$。事实上,因为 $\boldsymbol{A}\geqslant0,\boldsymbol{B}\geqslant0$ 则存在 n 阶矩阵 $\boldsymbol{P},\boldsymbol{H}$ 使得 $\boldsymbol{A}=\boldsymbol{PP}^{\mathrm{T}},\boldsymbol{B}=\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$,因此有 $\boldsymbol{A}\cdot\boldsymbol{B}=\mathrm{tr}(\boldsymbol{PP}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})=\mathrm{tr}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{HP})=\mathrm{tr}((\boldsymbol{HP})^{\mathrm{T}}\boldsymbol{HP})\geqslant0$。

称 $G:\mathbf{R}^m\to S^n$ 是可微的,若 $G(x)$ 作为 $x\in\mathbf{R}^m$ 的多元函数是可微的,令

$$\frac{\mathrm{d}G(x)}{\mathrm{d}x}=\begin{pmatrix}\dfrac{\partial G(x)}{\partial x_1}\\ \dfrac{\partial G(x)}{\partial x_2}\\ \vdots\\ \dfrac{\partial G(x)}{\partial x_m}\end{pmatrix}=\begin{pmatrix}G_1(x)\\ G_2(x)\\ \vdots\\ G_m(x)\end{pmatrix},$$

其中 $G_i(x)=\dfrac{\partial G(x)}{\partial x_i}$ 为 $n\times n(i=1,\cdots,m)$ 偏导数矩阵,则称 $\dfrac{\mathrm{d}G(x)}{\mathrm{d}x}$ 为 $G(x)$ 对 x 的导数。用 $DG(x)$ 表示 $G(\cdot)$ 在 x 的微分映射,即 $DG(x)$ 为从 $\mathbf{R}^m$ 到 S^n 的线性映射,定义为 $DG(x)y=\sum_{i=1}^{m}y_iG_i(x)$,且记 $\dfrac{\mathrm{d}G(x)}{\mathrm{d}x}:=DG(x)$。

定义1[1] 设集合 $C\subset\mathbf{R}^m$,如果存在一个向量值函数 $\eta:\mathbf{R}^m\times\mathbf{R}^m\to\mathbf{R}^m$ 使得 $\forall x,y\in C,\lambda\in[0,1]$,有 $y+\lambda\eta(x,y)\in C$,则称 C 是关于函数 η 的不变凸集。

定义2[1] 设集合 $C\subset\mathbf{R}^m$ 是函数 $\eta:\mathbf{R}^m\times\mathbf{R}^m\to\mathbf{R}^m$ 的不变凸集,$T:C\to Z$,若对 $\forall x,y\in C$, $\lambda\in[0,1]$,有

$$T(y+\lambda\eta(x,y))\leqslant\lambda T(x)+(1-\lambda)T(y),\tag{2}$$

则称 T 是关于函数 η 的预不变凸函数。

类似于 $\mathbf{R}^m\to\mathbf{R}$ 上的不变凸函数的定义[7],可定义 $\mathbf{R}^m\to S^n$ 上的不变凸函数,即有如下定义。

定义3 设集合 $C\subset\mathbf{R}^m$,$G:C\to S^n$ 可微,如果存在一个向量值函数 $\eta:\mathbf{R}^m\times\mathbf{R}^m\to\mathbf{R}^m$ 使得

$$G(x)-G(y)\geqslant DG(y)\eta(x,y),\ \forall x,y\in C,$$

则称 G 是关于函数 η 的不变凸函数。

事实上,若 $T:C\to Z$ Fréchet 可微[8]且是关于向量值函数 η 的预不变凸函数,则 T 为不变凸函数。这是因为

T Fréchet 可微且对 $\forall x,y\in C,\lambda\in[0,1]$，$T$ 满足(2)式，则(2)式可改写为 $\lambda(T(x)-T(y))\geqslant T(y+\lambda\eta(x,y))-T(y)$，两边同时除以 λ，令 $\lambda\to 0_+$ 有 $T(x)-T(y)\geqslant DT(y)\eta(x,y)$，因此 T 为不变凸函数，反之不然。

2 鞍点存在性的研究

本节讨论非凸半定规划问题(1)的鞍点存在的充分必要条件、充分条件及必要条件。用 Ω 表示问题(1)的可行集，即 $\Omega:=\{x\in C\mid g(x)\leqslant 0,G(x)\leqslant 0\}$。

定义问题(1)的 Lagrange 函数如下：$L(x,\mu,\boldsymbol{U})=f(x)+\mu^{\mathrm{T}}g(x)+\boldsymbol{U}\cdot G(x)$，且称 $\theta(\mu,\boldsymbol{U})=\min\limits_{x\in C}L(x,\mu,\boldsymbol{U})$ 为$L(x,\mu,\boldsymbol{U})$（其中$(\mu,\boldsymbol{U})\geqslant 0$）的对偶目标函数，则问题(1)的对偶问题可以写成

$$\max_{(\mu,\boldsymbol{U})\geqslant 0}\theta(\mu,\boldsymbol{U})。\tag{3}$$

$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$称为 Lagrange 函数 $L(x,\mu,\boldsymbol{U})$的一个鞍点，如果 $\bar{x}\in C$，$(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$，且满足

$$L(\bar{x},\mu,\boldsymbol{U})\leqslant L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})\leqslant L(x,\bar{\mu},\bar{\boldsymbol{U}}),\ \forall x\in C,\ \forall(\mu,\boldsymbol{U})\geqslant 0。\tag{4}$$

(4)式的解与问题(1)和对偶问题(3)的关系由以下定理给出。

在文献[1]的定理 3.3 证明了“若 Ω 非空，则$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,U)$的鞍点且 $\bar{x}\in\Omega$ 当且仅当 $\bar{x}$，$(\bar{\mu},\bar{\boldsymbol{U}})$分别为原始问题(1)与对偶问题(3)的最优解，且无对偶间隙即 $f(\bar{x})=\theta(\bar{\mu},\bar{\boldsymbol{U}})$”。本文将给出鞍点存在性的另一个充分必要条件如下。

定理 1　$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 Lagrange 函数 $L(x,\mu,\boldsymbol{U})=f(x)+\mu^{\mathrm{T}}g(x)+\boldsymbol{U}\cdot G(x)$，$\bar{x}\in C$，$(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$ 的鞍点的充分必要条件是
$$\begin{cases}a:L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})=\min\{L(x,\bar{\mu},\bar{\boldsymbol{U}}):x\in C\},\\ b:g(\bar{x})\leqslant 0,G(\bar{x})\leqslant 0,\\ c:\bar{\mu}^{\mathrm{T}}g(\bar{x})=0,\bar{\boldsymbol{U}}\cdot G(\bar{x})=0。\end{cases}$$

证明　假设$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点，由鞍点的定义可知 a 成立；由(4)式知

$$f(\bar{x})+\mu^{\mathrm{T}}g(\bar{x})+\boldsymbol{U}\cdot G(\bar{x})\leqslant f(\bar{x})+\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x}),\ \forall(\mu,\boldsymbol{U})\geqslant 0,\tag{5}$$

从而 $g(\bar{x})\leqslant 0,G(\bar{x})\leqslant 0$，否则假设 $g(\bar{x})\leqslant 0,G(\bar{x})\leqslant 0$ 不成立，则 $g(\bar{x})\leqslant 0$ 不成立或者 $G(\bar{x})\leqslant 0$ 不成立。对于第 1 种情况，设 $g(x)=(g_1(x),g_2(x),\cdots,g_k(x))^{\mathrm{T}}$，则存在 $i\in\{1,\cdots,k\}$使得 $g_i(\bar{x})>0$，让 μ_i 充分大，$\mu_j=0(j\neq i)$，$\boldsymbol{U}=\bar{\boldsymbol{U}}$，则(5)式中左边充分大，与(5)式矛盾，从而 $g(\bar{x})\leqslant 0$。对于第 2 种情况，设 $G(\bar{x})$的特征值为 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$，由 $G(\bar{x})\leqslant 0$ 不成立有 $\lambda_1>0$，则存在正交矩阵 $\boldsymbol{P}$ 使得

$$G(\bar{x})=\boldsymbol{P}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}\boldsymbol{P}^{\mathrm{T}}，取\ \mu=\bar{\mu},\boldsymbol{U}=\boldsymbol{P}\begin{pmatrix}t&&&\\&0&&\\&&\ddots&\\&&&0\end{pmatrix}\boldsymbol{P}^{\mathrm{T}}\geqslant 0(t>0)，则$$

$$\boldsymbol{U}\cdot\boldsymbol{G}(\bar{x})=\mathrm{tr}(\boldsymbol{P}\begin{pmatrix}t&&&\\&0&&\\&&\ddots&\\&&&0\end{pmatrix}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}\boldsymbol{P}^{\mathrm{T}})=t\lambda_1,$$

让 t 充分大，则(5)式中左边充分大，与(5)式矛盾，从而 $G(\bar{x})\leqslant 0$，b 成立。

(5)式中令 $\mu=0$，$\boldsymbol{U}=\bar{\boldsymbol{U}}$，有 $\bar{\mu}^{\mathrm{T}}g(\bar{x})\geqslant 0$，又由 $\bar{\mu}^{\mathrm{T}}g(\bar{x})\leqslant 0$，得到 $\bar{\mu}^{\mathrm{T}}g(\bar{x})=0$。(5)式中令 $\mu=\bar{\mu}$，$\boldsymbol{U}=0$，有 $\bar{\boldsymbol{U}}\cdot G(\bar{x})\geqslant 0$，又由 $\bar{\boldsymbol{U}}\cdot G(\bar{x})\leqslant 0$，得到 $\bar{\boldsymbol{U}}\cdot G(\bar{x})=0$，$c$ 成立。

假设 a,b,c 成立，由 a 可得 $L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})\leqslant L(x,\bar{\mu},\bar{\boldsymbol{U}})$，$\forall x\in C$。由 b,c 可得

$$L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})=f(\bar{x})+\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x})\geqslant f(\bar{x})+\mu^{\mathrm{T}}g(\bar{x})+\boldsymbol{U}\cdot G(\bar{x})=L(\bar{x},\mu,\boldsymbol{U})(\forall(\mu,\boldsymbol{U})\geqslant 0),$$

所以$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点。　　证毕

注 1　事实上“$\bar{x}$，$(\bar{\mu},\bar{\boldsymbol{U}})$分别为原始问题(1)与对偶问题(3)的最优解，且无对偶间隙即 $f(\bar{x})=\theta(\bar{\mu},\bar{\boldsymbol{U}})$”与定理 1 中的 a,b,c 等价，由定理 1 中的 b 可以看出$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为鞍点蕴含着 $\bar{x}$ 可行，因此文献[1]中的定理 3.3 可以改写为“若 Ω 非空，则$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点当且仅当 $\bar{x}$，$(\bar{\mu},\bar{\boldsymbol{U}})$分别为原始问题(1)与对偶问题(3)的最优

解,且无对偶间隙即 $f(\bar{x})=\theta(\bar{\mu},\bar{\boldsymbol{U}})$”,所以定理 1 与文献[1]中的定理 3.3 等价。相对于文献[1]中的定理 3.3,本文的定理 1 中的 a,b,c 所构成的系统的解容易在计算机上实现。

为了讨论鞍点存在性的必要、充分条件,本文仿照一般的非线性规划问题中由 Lagrange 函数产生的 Karush-Kuhn-Tucker 条件,本文也给出了所谓的广义 Karush-Kuhn-Tucker 条件,简称广义 KKT 条件:

$$\begin{cases}\nabla f(\bar{x})+\bar{\mu}^{\mathrm{T}}\nabla g(\bar{x})+\bar{\boldsymbol{U}}\cdot DG(\bar{x})=0,\\ \bar{\mu}^{\mathrm{T}}g(\bar{x})=0,\\ \bar{\boldsymbol{U}}\cdot G(\bar{x})=0。\end{cases}\tag{6}$$

其中 $\bar{x}\in\Omega$,$(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$。

如下定理讨论了鞍点存在性与广义 KKT 条件的关系。

定理 2 考虑问题(1)$\bar{x}\in\Omega$,假设存在$(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$ 满足(6)式,设 $g(x)=(g_1(x),g_2(x),\cdots,g_k(x))^{\mathrm{T}}$,$I=\{i:g_i(\bar{x})=0,i=1,\cdots,k\}$,假设 $f,g_i(i\in I),G$ 是关于相同的 $\eta:\mathbf{R}^m\times\mathbf{R}^m\to\mathbf{R}^m$ 的不变凸函数,则$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点;反之,假设$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点,$\bar{x}\in\mathrm{int}C$,$(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$,则 $\bar{x}$ 为问题(1)的可行点,且$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$满足(6)式。

证明 假设$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$,$\bar{x}\in\Omega$,$(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$,使得(6)式成立,$\forall x\in C$,由不变凸性有

$$f(x)-f(\bar{x})\geqslant\nabla f(\bar{x})^{\mathrm{T}}\eta(x,\bar{x}),\tag{7}$$

$$g_i(x)-g_i(\bar{x})\geqslant\nabla g_i(\bar{x})^{\mathrm{T}}\eta(x,\bar{x}),\ \forall i\in I,\tag{8}$$

$$G(x)-G(\bar{x})\geqslant DG(\bar{x})\eta(x,\bar{x}),\tag{9}$$

(8)式乘 $\bar{\mu}_i$,(9)式与 $\bar{\boldsymbol{U}}$ 作内积,与(7)式相加可得

$$\begin{aligned}&f(x)+\sum_{i\in I}\bar{\mu}_i g_i(x)+\bar{\boldsymbol{U}}\cdot G(x)-(f(\bar{x})+\sum_{i\in I}\bar{\mu}_i g_i(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x}))\geqslant\\&\quad(\nabla f(\bar{x})+\sum_{i\in I}\bar{\mu}_i\nabla g_i(\bar{x})+\bar{\boldsymbol{U}}\cdot DG(\bar{x}))^{\mathrm{T}}\eta(x,\bar{x}),\end{aligned}\tag{10}$$

由(6)式中 $\bar{\mu}^{\mathrm{T}}g(\bar{x})=0$,可得 $\bar{\mu}_i=0\ (i\notin I)$,(10)式可转化为

$$f(x)+\bar{\mu}^{\mathrm{T}}g(x)+\bar{\boldsymbol{U}}\cdot G(x)-(f(\bar{x})+\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x}))\geqslant(\nabla f(\bar{x})+\bar{\mu}^{\mathrm{T}}\nabla g(\bar{x})+\bar{\boldsymbol{U}}\cdot DG(\bar{x}))^{\mathrm{T}}\eta(x,\bar{x})=0,$$

即对 $\forall x\in C$ 都有 $L(x,\bar{\mu},\bar{\boldsymbol{U}})\geqslant L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$,由 $g(\bar{x})\leqslant 0$,$G(\bar{x})\leqslant 0$,$\bar{\mu}^{\mathrm{T}}g(\bar{x})=\bar{\boldsymbol{U}}\cdot G(\bar{x})=0$,可得

$$f(\bar{x})+\mu^{\mathrm{T}}g(\bar{x})+\boldsymbol{U}\cdot G(\bar{x})\leqslant f(\bar{x})=f(\bar{x})+\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x})(\forall(\mu,\boldsymbol{U})\geqslant 0),$$

即对 $\forall(\mu,\boldsymbol{U})\geqslant 0$ 都有 $L(\bar{x},\mu,\boldsymbol{U})\leqslant L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$,故$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点。

假设$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点,则由定理 1 可得

$$g(\bar{x})\leqslant 0,G(\bar{x})\leqslant 0,\bar{\mu}^{\mathrm{T}}g(\bar{x})=\bar{\boldsymbol{U}}\cdot G(\bar{x})=0,L(\bar{x},\mu,\boldsymbol{U})\leqslant L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}}),\ \forall(\mu,\boldsymbol{U})\geqslant 0,$$

又由 $\bar{x}\in\mathrm{int}\,C$,由费马定理可得$\nabla_x L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})=0$ 即$\nabla f(\bar{x})+\bar{\mu}^{\mathrm{T}}\nabla g(\bar{x})+\bar{\boldsymbol{U}}\cdot DG(\bar{x})=0$。故$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$满足(6)式。

证毕

注 2 定理 2 说明了在不变凸性的假设下,广义 KKT 条件是鞍点存在的一个充分条件。

注 3 在定理 2 中把条件“$f,g_i(i\in I),G$ 是关于相同的 $\eta:\mathbf{R}^m\times\mathbf{R}^m\to\mathbf{R}^m$ 的不变凸函数”改为“C 是关于 $\eta:\mathbf{R}^m\times\mathbf{R}^m\to\mathbf{R}^m$ 的不变凸集,$f,g_i(i\in I),G$ 是 Fréchet 可微的且是关于向量值函数 η 的预不变凸函数”则定理仍成立。

考虑问题(1),定义扰动函数 $\nu:\mathbf{R}^k\times S^n\to\mathbf{R}$,

$$\nu(y)=\min\{f(x):g(x)\leqslant y_1(y_1=(y_1^{(1)},\cdots,y_1^{(k)})^{\mathrm{T}}),G(x)\leqslant y_2,y_2\in S^n,y=(y_1,y_2),x\in C\}。\tag{11}$$

定理 3 假设问题(1)的最优解 $\bar{x}$ 存在,则$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点的充分必要条件是

$$\nu(y)\geqslant\nu(0)-\bar{\mu}^{\mathrm{T}}y_1-\bar{\boldsymbol{U}}\cdot y_2,\ \forall y\in\mathbf{R}^k\times S^n。\tag{12}$$

证明 假设$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$为 $L(x,\mu,\boldsymbol{U})$的鞍点,由文献[1]中的定理 3.3 知 $f(\bar{x})=\theta(\bar{\mu},\bar{\boldsymbol{U}})$。因为 $\nu(0)=\inf P=f(\bar{x})$,则对 $\forall y\in\mathbf{R}^k\times S^n$ 有

$$\begin{aligned}&\nu(0)=\theta(\bar{\mu},\bar{\boldsymbol{U}})=\min\{f(x)+\bar{\mu}^{\mathrm{T}}g(x)+\bar{\boldsymbol{U}}\cdot G(x):x\in C\}=\\&\bar{\mu}^{\mathrm{T}}y_1+\bar{\boldsymbol{U}}\cdot y_2+\min\{f(x)+\bar{\mu}^{\mathrm{T}}(g(x)-y_1)+\bar{\boldsymbol{U}}\cdot(G(x)-y_2):x\in C\}。\end{aligned}$$

对扰动问题(11)应用弱对偶定理有 $\nu(0)\leqslant\bar{\mu}^{\mathrm{T}}y_1+\bar{\boldsymbol{U}}\cdot y_2+\nu(y)(\forall y\in\mathbf{R}^k\times S^n)$，故(12)式成立。

反之，若(12)式成立，$\bar{x}$ 为问题(1)的最优解，有 $\bar{x}\in C,g(\bar{x})\leqslant 0,G(\bar{x})\leqslant 0$，下证 $(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$。

假设 $(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$ 不成立，则 $\bar{\mu}\geqslant 0$ 不成立或者 $\bar{\boldsymbol{U}}\geqslant 0$ 不成立。对于第 1 种情况即 $\bar{\mu}\geqslant 0$ 不成立，则存在 $\bar{\mu}_i<0$ $(i\in\{1,\cdots,k\})$，(12)式中令 $y_1^{(i)}>0,y_1^{(j)}=0,j\neq i,y_2=0$，则有 $\nu(0)\geqslant\nu(y)\geqslant\nu(0)-\bar{\mu}_i y_1^{(i)}\Rightarrow\bar{\mu}_i y_1^{(i)}\geqslant 0$，与 $\bar{\mu}_i y_1^{(i)}<0$ 矛盾，因此有 $\bar{\mu}\geqslant 0$。对于第 2 种情况，设 $\bar{\boldsymbol{U}}$ 的特征值为 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$，由 $\bar{\boldsymbol{U}}\geqslant 0$ 不成立，必有 $\lambda_n<0$，则存在正交矩阵 $\boldsymbol{P}$ 使得 $\bar{\boldsymbol{U}}=\boldsymbol{P}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\boldsymbol{P}^{\mathrm{T}}$，(12)式中令 $y_1=0,y_2=\boldsymbol{P}\begin{pmatrix}0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & 1\end{pmatrix}\boldsymbol{P}^{\mathrm{T}}\geqslant 0$，则 $\nu(0)\geqslant\nu(y)\geqslant\nu(0)-\bar{\boldsymbol{U}}\cdot y_2\Rightarrow\bar{\boldsymbol{U}}\cdot y_2\geqslant 0$，与

$$\bar{\boldsymbol{U}}\cdot y_2=\mathrm{tr}\left(\boldsymbol{P}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\begin{pmatrix}0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & 1\end{pmatrix}\boldsymbol{P}^{\mathrm{T}}\right)=\lambda_n<0,$$

矛盾，因此有 $\bar{\boldsymbol{U}}\geqslant 0$，故 $(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$。

(12)式中令 $y=\bar{y}=(g(\bar{x}),G(\bar{x}))\leqslant 0$，则 $\nu(\bar{y})\geqslant\nu(0)=f(\bar{x})$。因为 $\bar{x}$ 为 $y=\bar{y}$ 时扰动问题(11)的可行点，所以 $\nu(0)=f(\bar{x})\geqslant\nu(\bar{y})$，从而可得 $\nu(0)=\nu(\bar{y})$。由(12)式有 $\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x})\geqslant 0$。又由 $(\bar{\mu},\bar{\boldsymbol{U}})\geqslant 0$，$(g(\bar{x}),G(\bar{x}))\leqslant 0$ 有 $\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x})\leqslant 0$，可得 $\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x})=0$，从而有 $\bar{\mu}^{\mathrm{T}}g(\bar{x})=\bar{\boldsymbol{U}}\cdot G(\bar{x})=0$。

最后由(12)式有

$$L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})=f(\bar{x})+\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x})=f(\bar{x})=\nu(0)\leqslant\nu(y)+\bar{\mu}^{\mathrm{T}}y_1+\bar{\boldsymbol{U}}\cdot y_2,\ \forall y\in\mathbf{R}^k\times S^n。$$

对 $\forall\hat{x}\in C$，令 $\hat{y}=(g(\hat{x}),G(\hat{x}))$，则 $\hat{x}$ 为 $y=\hat{y}$ 时扰动问题(11)的可行点，可得 $f(\hat{x})\geqslant\nu(\hat{y})$。由(12)式有

$$f(\bar{x})=\nu(0)\leqslant\nu(\hat{y})+\bar{\mu}^{\mathrm{T}}g(\hat{x})+\bar{\boldsymbol{U}}\cdot G(\hat{x})\leqslant f(\hat{x})+\bar{\mu}^{\mathrm{T}}g(\hat{x})+\bar{\boldsymbol{U}}\cdot G(\hat{x})=L(\hat{x},\bar{\mu},\bar{\boldsymbol{U}}),\ \forall\hat{x}\in C。$$

由前面的证明知 $\bar{\mu}^{\mathrm{T}}g(\bar{x})=\bar{\boldsymbol{U}}\cdot G(\bar{x})=0$，则

$$L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})=f(\bar{x})+\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x})=f(\bar{x})\leqslant L(\hat{x},\bar{\mu},\bar{\boldsymbol{U}}),\ \forall\hat{x}\in C。$$

即 $L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})\leqslant L(x,\bar{\mu},\bar{\boldsymbol{U}}),\ \forall x\in C$。由 $(g(\bar{x}),G(\bar{x}))\leqslant 0$，可得

$$L(\bar{x},\mu,U)=f(\bar{x})+\mu^{\mathrm{T}}g(\bar{x})+\boldsymbol{U}\cdot G(\bar{x})\leqslant f(\bar{x})=f(\bar{x})+\bar{\mu}^{\mathrm{T}}g(\bar{x})+\bar{\boldsymbol{U}}\cdot G(\bar{x})=L(\bar{x},\bar{\mu},\bar{\boldsymbol{U}}),\ \forall(\mu,U)\geqslant 0。$$

故 $(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$ 为 $L(x,\mu,U)$ 的鞍点。 证毕

注 4 定理 3 说明了若 $\bar{x}$ 为问题(1)的最优解，$(\bar{x},\bar{\mu},\bar{\boldsymbol{U}})$ 为 $L(x,\mu,U)$ 的鞍点的充分必要条件是超平面 $H=\{(y,z):\langle(\bar{\mu},\bar{\boldsymbol{U}},1),(y,z)\rangle=\nu(0),(y,z)\in\mathbf{R}^k\times S^n\times\mathbf{R}\}$ 是扰动函数 ν 的上图 $\mathrm{eip}\ \nu=\{(y,z):z\geqslant\nu(y),\ y\in\mathbf{R}^k\times S^n\}$ 在点 $(y,z)=(0,\nu(0))$ 处的支撑超平面，其中 $y=(y_1,y_2)$，$\langle(\bar{\mu},\bar{\boldsymbol{U}},1),(y,z)\rangle=\bar{\mu}^{\mathrm{T}}y_1+\bar{\boldsymbol{U}}\cdot y_2+z$。换言之，若问题(1)的最优解存在，则对偶可达且无对偶间隙等价于扰动函数 ν 的上图在点 $(0,\nu(0))$ 处的支撑超平面存在。

3 结论

文中对非凸半定规划问题(1)的鞍点存在性进行了讨论。定理 1 给出了一个鞍点存在的充分必要条件，且与文献[1]中的定理 3.3 是等价的。定理 2 在不变凸性的假设下证明了广义 KKT 条件是鞍点存在和全局最优性的充分条件。定理 3 则是引入扰动函数来刻画鞍点存在的一个充分必要条件。

参考文献：

[1] Sun W, Li C, Sampaio R J B. On duality theory for nonconvex semidefinite programming[J]. Annals of Operations Research, 2011, 186(1): 331-343.

[2] Fan J. Duality theories in nonlinear semidefinite programming[J]. Applied mathematics letters, 2005, 18(9): 1068-1073.

[3] 李成进，孙文瑜. 非凸半定规划的广义 Fakars 引理及最优性条件[J]. 高等学校计算数学学报, 2008, 30(2): 184-192.

Li C J, Sun W Y. Generalized Fakars lemma and optimal conditions for nonconvex semidefinite programming problems[J]. Numerical Mathematics a Journal of Chinese Universities, 2008, 30(2): 184-192.

[4] Bazaraa M S, Sherali H D, Shetty C M. Nonlinear programming: theory and algorithms[M]. New York: John Wiley & Sons, 2013.

[5] Sun D. The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications[J]. Mathematics of Operations Research, 2006, 31(4): 761-776.

[6] 张立卫. 非线性半定规划若干进展[J]. 运筹学学报, 2014, 18(1): 93-112.
Zhang L W. Nonlinear semidefinite programming some development[J]. Journal of Operational Research, 2014, 18(1): 93-112.

[7] Weir T, Jeyakumar V. A class of nonconvex functions and mathematical programming[J]. Bulletin of the Australian Mathematical Society, 1988, 38(2): 177-189.

[8] 张立卫. 锥约束优化[M]. 北京: 科学出版社, 2010.
Zhang L W. Cone constraints optimization[M]. Beijjing: Science Press, 2010.

Operations Research and Cybernetics

The Study of the Existence of Saddle Point for Nonconvex Semidefinite Programming Problems

LI Yongling, LUO Honglin, XIANG Yanning

(College of Mathematics Sciences, Chongqing Normal University, Chongqing 401331, China)

Abstract: In this paper, we devote to study the existence of the saddle point of nonconvex semidefinite programming problems by means of Spectral decomposition, Inner product and correlative properties of Matrix Analysis and Separation theorem of convex set of Convex Analysis. In this case, some necessary and/or sufficient conditions for existence for the saddle point are derived and proved in three different ways. First, we present a sufficient and necessary condition which is equivalent to dual theorem in ref. [1] by utilizing an inequality system. Then, we give a generalized Karush-Kuhn-Tucker condition and prove this condition is sufficient conditions for existence of the saddle point under invex convexity assumption. In addition, if $\overline{x} \in \mathrm{int}\ C$, this sufficient condition is also necessary condition. Finally, we define a perturbation function which is used to deduce a sufficient and necessary condition for existence of the saddle point: dual attainment and the absence of a duality gap is equivalent to the existence of a supporting hyperplane for the epigraph of ν at the point $(0, \nu(0))$.

Key words: nonconvex semidefinite programming; saddle point; generalized KKT condition; invex

(责任编辑 黄 颖)