第 38 卷第 1 期　西南大学学报(自然科学版)　2016 年 1 月
Vol. 38 No. 1　Journal of Southwest University (Natural Science Edition)　Jan. 2016

DOI: 10.13718/j.cnki.xdzk.2016.01.016

非凸半定规划的最优性条件[①]

李永玲，　罗洪林，　向彦宁

重庆师范大学 数学科学学院，重庆 401331

摘要： 研究了非凸半定规划的一阶和二阶充分性条件. 在不变凸性的假设下，给出并证明了广义 Karush-Kuhn-Tucker 条件是非凸半定规划具有全局最优解的一阶充分性条件. 在没有任何广义凸性的假设下，给出了非凸半定规划具有严格局部最优解的二阶充分条件.

关 键 词： 非凸半定规划；最优性条件；不变凸

中图分类号： O224　　**文献标志码：** A　　**文章编号：** 1673-9868(2016)01-0103-06

在本文中，我们用 $\mathbb{R}^m$，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. 我们考虑如下形式的非凸半定规划问题：

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

其中 $f: \mathbb{R}^m \longrightarrow \mathbb{R}$，$\boldsymbol{g}: \mathbb{R}^m \longrightarrow \mathbb{R}^k$，$\boldsymbol{G}: \mathbb{R}^m \longrightarrow S^n$ 均不要求是凸的，$\boldsymbol{G}(\boldsymbol{x}) \leqslant \boldsymbol{0}$ 当且仅当 $-\boldsymbol{G}(\boldsymbol{x}) \in S^n_+$.

在无约束以及含约束的优化问题中，最优性条件即最优解存在的必要条件、充分条件、充分必要条件的研究一直是优化理论与方法研究中的基础和热点课题.

文献[1]建立了广义 Fakars 引理，并在类凸(预不变凸)性的假设下利用此广义 Fakars 引理证明了非凸半定规划问题的最优性必要条件以及充分必要条件. 文献[2-3]则对一般的非线性半定规划问题的最优性条件进行了研究. 文献[2]中以正则条件作为约束品性来证明了非线性半定规划问题的一阶和二阶最优性条件；结合矩阵形式的 Fakars 引理证明了一阶必要条件；类似于讨论一般的非线性规划问题的二阶最优性条件的方式证明了非线性半定规划问题的二阶必要以及二阶充分条件. 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[3]证明了以下条件是等价的：强二阶充分条件和约束非退化；Karush-Kuhn-Tucker 系统的 Clarke's Jacobian 非奇异；Karush-Kuhn-Tucker 点满足强正则性. 文献[4]中也给出了在不同的假设条件下 K-K-T 点是广义方程的强正则解.

文献[1]在类凸的假设下讨论了非凸半定规划问题的一阶最优性条件，而本文在另一种广义凸性——不变凸的假设下来讨论一阶最优性条件. 类似于文献[2]，本文也利用类似于讨论一般的非线性规划问题的二阶最优性条件的方式来证明了非凸半定规划问题的二阶最优性充分条件. 从文献[3]中可以看出

---

① 收稿日期：2014-11-24
基金项目：国家自然科学基金项目(11431004).
作者简介：李永玲(1989-)，女，重庆梁平人，硕士研究生，主要从事半定规划的研究.
通信作者：罗洪林，博士，副教授.

Karush-Kuhn-Tucker 条件在非线性半定规划问题中扮演着重要的角色. 受此启发，为了给出非凸半定规划问题的最优性充分条件，本文也给出了所谓的广义 Karush-Kuhn-Tucker 条件(式(3))，简称广义 K-K-T 条件. 此外，从计算的角度来看，由于本文引入的广义 K-K-T 系统的求解易在计算机上实现，因此广义 K-K-T 条件的引入是有必要的.

1 预备知识与记号

在这一节，我们将介绍一些定义、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. 基于此定义我们可以定义

$$\boldsymbol{A}\cdot\boldsymbol{C}=\begin{pmatrix}\boldsymbol{A}\cdot\boldsymbol{C}_{11} & \cdots & \boldsymbol{A}\cdot\boldsymbol{C}_{1q}\\ \boldsymbol{A}\cdot\boldsymbol{C}_{21} & \cdots & \boldsymbol{A}\cdot\boldsymbol{C}_{2q}\\ \vdots & & \vdots\\ \boldsymbol{A}\cdot\boldsymbol{C}_{p1} & \cdots & \boldsymbol{A}\cdot\boldsymbol{C}_{pq}\end{pmatrix}$$

其中 $\boldsymbol{A}\in S^n,\boldsymbol{C}=\begin{pmatrix}\boldsymbol{C}_{11} & \cdots & \boldsymbol{C}_{1q}\\ \boldsymbol{C}_{21} & \cdots & \boldsymbol{C}_{2q}\\ \vdots & & \vdots\\ \boldsymbol{C}_{p1} & \cdots & \boldsymbol{C}_{pq}\end{pmatrix}$，$\boldsymbol{C}_{ij}\in S^n$，$\boldsymbol{C}\cdot\boldsymbol{C}_{ij}$ 为对称矩阵空间 S^n 中定义的内积.

如果我们用 Z 表示 $\mathbb{R}^i$(S^j 或 $\mathbb{R}^i\times S^j$)，则用 Z_+ 表示 $\mathbb{R}_+$(S_+^j 或 $\mathbb{R}_+\times S_+^j$)，其中 i,j 可以为任意的正整数. 在此基础上可定义如下的偏序关系：

$$\boldsymbol{A}\geqslant\boldsymbol{B}\Leftrightarrow\boldsymbol{A}-\boldsymbol{B}\in Z_+$$

$$\boldsymbol{A}>\boldsymbol{B}\Leftrightarrow\boldsymbol{A}-\boldsymbol{B}\in\mathrm{int}\,Z_+,\ \forall\,\boldsymbol{A},\boldsymbol{B}\in Z$$

若 $\boldsymbol{A},\boldsymbol{B}\in S_+^n$，则有 $\boldsymbol{A}\cdot\boldsymbol{B}\geqslant 0$. 事实上，因为 $\boldsymbol{A}\geqslant\boldsymbol{0}$，$\boldsymbol{B}\geqslant\boldsymbol{0}$，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

$$\boldsymbol{A}\cdot\boldsymbol{B}=\mathrm{Tr}(\boldsymbol{AB})=\mathrm{Tr}(\boldsymbol{PP}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})=\mathrm{Tr}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{P})=\mathrm{Tr}((\boldsymbol{HP})^{\mathrm{T}}\boldsymbol{HP})\geqslant 0$$

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

$$\frac{\mathrm{d}\boldsymbol{G}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}=\begin{pmatrix}\dfrac{\partial\boldsymbol{G}(\boldsymbol{x})}{\partial x_1}\\ \dfrac{\partial\boldsymbol{G}(\boldsymbol{x})}{\partial x_2}\\ \vdots\\ \dfrac{\partial\boldsymbol{G}(\boldsymbol{x})}{\partial x_m}\end{pmatrix}=\begin{pmatrix}\boldsymbol{G}_1(\boldsymbol{x})\\ \boldsymbol{G}_2(\boldsymbol{x})\\ \vdots\\ \boldsymbol{G}_m(\boldsymbol{x})\end{pmatrix}$$

其中

$$\boldsymbol{G}_i(\boldsymbol{x})=\frac{\partial\boldsymbol{G}(\boldsymbol{x})}{\partial x_i}\qquad i=1,\cdots,m$$

为 $n\times n$ 偏导数矩阵，则称 $\dfrac{\mathrm{d}\boldsymbol{G}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}$ 为 $\boldsymbol{G}(\boldsymbol{x})$ 对 $\boldsymbol{x}$ 的导数. 我们用 $D\boldsymbol{G}(\boldsymbol{x})$ 表示 $\boldsymbol{G}(\cdot)$ 在 $\boldsymbol{x}$ 的微分映射，即 $D\boldsymbol{G}(\boldsymbol{x})$ 为从 $\mathbb{R}^m$ 到 S^n 的线性映射，定义为

$$D\boldsymbol{G}(\boldsymbol{x})\boldsymbol{y}=\sum_{i=1}^{m}y_i\boldsymbol{G}_i(\boldsymbol{x})$$

且记

$$\frac{\mathrm{d}\boldsymbol{G}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}=D\boldsymbol{G}(\boldsymbol{x})$$

定义 1[5]　考虑问题$\min\limits_{x\in\Omega} f(\boldsymbol{x})$，若可行点$\overline{\boldsymbol{x}}$存在一个ε(ε > 0)邻域$N_\varepsilon(\overline{\boldsymbol{x}})$，使得对 $\forall \boldsymbol{x}\in\Omega\cap N_\varepsilon(\overline{\boldsymbol{x}})$且$\boldsymbol{x}\neq\overline{\boldsymbol{x}}$，都满足$f(\boldsymbol{x})>f(\overline{\boldsymbol{x}})$，则称$\overline{\boldsymbol{x}}$为严格局部极小点.

定义 2[1]　设集合$C\subset\mathbb{R}^m$，如果存在一个向量值函数$\boldsymbol{\eta}:\mathbb{R}^m\times\mathbb{R}^m\longrightarrow\mathbb{R}^m$使得$\forall\boldsymbol{x},\boldsymbol{y}\in C$，$\lambda\in[0,1]$，有$\boldsymbol{y}+\lambda\boldsymbol{\eta}(\boldsymbol{x},\boldsymbol{y})\in C$，则称$C$是关于函数$\boldsymbol{\eta}$的不变凸集.

定义 3[1]　设集合$C\subset\mathbb{R}^m$是关于向量值函数$\boldsymbol{\eta}:\mathbb{R}^m\times\mathbb{R}^m\longrightarrow\mathbb{R}^m$的不变凸集，$\boldsymbol{T}:C\longrightarrow Z$，若对$\forall\boldsymbol{x},\boldsymbol{y}\in C$，$\lambda\in[0,1]$，有

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

则称$\boldsymbol{T}$是关于函数$\boldsymbol{\eta}$的预不变凸函数.

类似于$\mathbb{R}^m\longrightarrow\mathbb{R}$上的不变凸函数的定义，我们可定义$\mathbb{R}^m\longrightarrow S^n$上的不变凸函数，即：

定义 4　设集合$C\subset\mathbb{R}^m$，$\boldsymbol{G}:C\longrightarrow S^n$可微，如果存在一个向量值函数$\boldsymbol{\eta}:\mathbb{R}^m\times\mathbb{R}^m\longrightarrow\mathbb{R}^m$使得

$$\boldsymbol{G}(\boldsymbol{x})-\boldsymbol{G}(\boldsymbol{y})\geqslant D\boldsymbol{G}(\boldsymbol{y})\boldsymbol{\eta}(\boldsymbol{x},\boldsymbol{y}),\ (\forall\boldsymbol{x},\boldsymbol{y}\in C)$$

则称$\boldsymbol{G}$是关于函数$\boldsymbol{\eta}$的不变凸函数.

事实上，若$\boldsymbol{T}:C\longrightarrow Z$是 Fréchet 可微的且是关于向量值函数$\boldsymbol{\eta}$的预不变凸函数，则$\boldsymbol{T}$为不变凸函数. 这是因为$\boldsymbol{T}$是 Fréchet 可微的且对$\forall\boldsymbol{x},\boldsymbol{y}\in C$，$\lambda\in[0,1]$，$\boldsymbol{T}$满足(2)式，则(2)式可改写为

$$\lambda(\boldsymbol{T}(\boldsymbol{x})-\boldsymbol{T}(\boldsymbol{y}))\geqslant\boldsymbol{T}(\boldsymbol{y}+\lambda\boldsymbol{\eta}(\boldsymbol{x},\boldsymbol{y}))-\boldsymbol{T}(\boldsymbol{y})$$

两边同时除以λ，令$\lambda\to0_+$有

$$\boldsymbol{T}(\boldsymbol{x})-\boldsymbol{T}(\boldsymbol{y})\geqslant D\boldsymbol{T}(\boldsymbol{y})\boldsymbol{\eta}(\boldsymbol{x},\boldsymbol{y})$$

反之不然.

2　最优性条件

在这一节，我们主要讨论非凸半定规划问题(1)的一阶与二阶最优性充分条件. 用Ω表示问题(1)的可行集，即

$$\Omega=\{\boldsymbol{x}\in C:\boldsymbol{g}(\boldsymbol{x})\leqslant\boldsymbol{0},\ \boldsymbol{G}(\boldsymbol{x})\leqslant\boldsymbol{0}\}$$

为了给出非凸半定规划问题(1)的最优性充分条件，本文给出了所谓的广义 Karush-Kuhn-Tucker 条件(简称广义 K-K-T 条件)如下：

$$\begin{cases}\nabla f(\overline{\boldsymbol{x}})+\nabla\boldsymbol{g}(\overline{\boldsymbol{x}})^{\mathrm{T}}\overline{\boldsymbol{\mu}}+\overline{\boldsymbol{U}}\cdot D\boldsymbol{G}(\overline{\boldsymbol{x}})=\boldsymbol{0}\\ \overline{\boldsymbol{\mu}}^{\mathrm{T}}\boldsymbol{g}(\overline{\boldsymbol{x}})=0\\ \overline{\boldsymbol{U}}\cdot\boldsymbol{G}(\overline{\boldsymbol{x}})=0\end{cases}\tag{3}$$

其中：$\overline{\boldsymbol{x}}\in\Omega$，$(\overline{\boldsymbol{\mu}},\overline{\boldsymbol{U}})\geqslant\boldsymbol{0}$.

一般情况下广义 K-K-T 条件不能保证$\overline{\boldsymbol{x}}\in\Omega$的最优性，现证明在不变凸性的假设下，广义 K-K-T 条件是全局最优性的充分条件.

定理 1(充分条件)　考虑问题(1)，设$\boldsymbol{g}(\boldsymbol{x})=(g_1(\boldsymbol{x}),g_2(\boldsymbol{x}),\cdots,g_k(\boldsymbol{x}))^{\mathrm{T}}$，$I=\{i:g_i(\overline{\boldsymbol{x}})=0,\ i=1,\cdots,k\}$，$\overline{\boldsymbol{x}}\in\Omega$，假设$f$，$g_i(i\in I)$，$\boldsymbol{G}$是关于相同的$\boldsymbol{\eta}:\mathbb{R}^m\times\mathbb{R}^m\longrightarrow\mathbb{R}^m$的不变凸函数，如果存在$\overline{\boldsymbol{\mu}}=(\overline{\boldsymbol{\mu}}_1,\overline{\boldsymbol{\mu}}_2,\cdots,\overline{\boldsymbol{\mu}}_k)^{\mathrm{T}}\geqslant\boldsymbol{0}$，$\overline{\boldsymbol{U}}\geqslant\boldsymbol{0}(\overline{\boldsymbol{U}}\in S^n)$使得(3)式成立，则$\overline{\boldsymbol{x}}$是问题(1)的全局极小点.

证　当$i\notin I$时，$g_i(\overline{\boldsymbol{x}})\neq0$，由

$$\overline{\boldsymbol{\mu}}^{\mathrm{T}}\boldsymbol{g}(\overline{\boldsymbol{x}})=0$$

可知

$$\{\overline{\boldsymbol{\mu}}_i=0(i\notin I)\}$$

从而有

$$\nabla \boldsymbol{g}(\bar{\boldsymbol{x}})^{\mathrm{T}}\bar{\boldsymbol{\mu}} = \sum_{i\in I}\bar{\boldsymbol{\mu}}_i \nabla g_i(\bar{\boldsymbol{x}})$$

于是由条件(3)及不变凸性的定义可得

$$\begin{aligned} f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}) \geqslant{} & \boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{x},\bar{\boldsymbol{x}})\nabla f(\bar{\boldsymbol{x}}) = \\ & -\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{x},\bar{\boldsymbol{x}})(\nabla \boldsymbol{g}(\bar{\boldsymbol{x}})^{\mathrm{T}}\bar{\boldsymbol{\mu}}) - \bar{\boldsymbol{U}}\cdot D\boldsymbol{G}(\bar{\boldsymbol{x}})\boldsymbol{\eta}(\boldsymbol{x},\bar{\boldsymbol{x}}) = \\ & -\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{x},\bar{\boldsymbol{x}})\sum_{i\in I}\bar{\boldsymbol{\mu}}_i\nabla g_i(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{U}}\cdot D\boldsymbol{G}(\bar{\boldsymbol{x}})\boldsymbol{\eta}(\boldsymbol{x},\bar{\boldsymbol{x}}) \geqslant \\ & -\sum_{i\in I}\bar{\boldsymbol{\mu}}_i(g_i(\boldsymbol{x}) - g_i(\bar{\boldsymbol{x}})) - \bar{\boldsymbol{U}}\cdot(\boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{G}(\bar{\boldsymbol{x}})) = \\ & -\sum_{i\in I}\bar{\boldsymbol{\mu}}_i g_i(\boldsymbol{x}) - \bar{\boldsymbol{U}}\cdot\boldsymbol{G}(\boldsymbol{x}) = \\ & -\bar{\boldsymbol{\mu}}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}) - \bar{\boldsymbol{U}}\cdot\boldsymbol{G}(\boldsymbol{x}) \geqslant \\ & 0 \qquad \forall\, \boldsymbol{x}\in\Omega \end{aligned}$$

因此 $\bar{\boldsymbol{x}}$ 是问题(1) 的全局极小点.

注 1 在定理 1 中把条件"$f, g_i(i\in I), \boldsymbol{G}$ 是关于相同的 $\boldsymbol{\eta}: \mathbb{R}^m\times\mathbb{R}^m \longrightarrow \mathbb{R}^m$ 的不变凸函数"改为"C 是关于 $\boldsymbol{\eta}: \mathbb{R}^m\times\mathbb{R}^m \longrightarrow \mathbb{R}^m$ 的不变凸集，$f, g_i(i\in I), \boldsymbol{G}$ 是 Fréchet 可微的且是关于向量值函数 $\boldsymbol{\eta}$ 的预不变凸函数"则定理仍成立.

定理 2(二阶充分条件) 考虑问题(1)，假设 C 为非空开集，$f, \boldsymbol{g}, \boldsymbol{G}$ 都是二阶可微的，$\bar{\boldsymbol{x}}\in\Omega$，存在$(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{U}})\geqslant \boldsymbol{0}$，且$(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{U}})\neq\boldsymbol{0}$，使得 $\bar{\boldsymbol{x}}$ 满足(3) 式. 令

$$I = \{i: g_i(\bar{\boldsymbol{x}}) = 0,\ i = 1, \cdots, k\} \qquad I^{+} = \{i\in I: \bar{\boldsymbol{\mu}}_i > 0\} \qquad I^{0} = \{i\in I: \bar{\boldsymbol{\mu}}_i = 0\}$$

分别称为积极约束指标集、强积极约束指标集、弱积极约束指标集. 定义限制的 Lagrange 函数 $\bar{L}(\boldsymbol{x}) = f(\boldsymbol{x}) + \sum_{i\in I}\bar{\boldsymbol{\mu}}_i g_i(\boldsymbol{x}) + \bar{\boldsymbol{U}}\cdot\boldsymbol{G}(\boldsymbol{x})$，$\bar{L}(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 的 Hessian 矩阵可表示为

$$\nabla^2\bar{L}(\bar{\boldsymbol{x}}) = \nabla^2 f(\bar{\boldsymbol{x}}) + \sum_{i\in I}\bar{\boldsymbol{\mu}}_i\nabla^2 g_i(\bar{\boldsymbol{x}}) + \bar{\boldsymbol{U}}\cdot D^2\boldsymbol{G}(\bar{\boldsymbol{x}})$$

其中 $\nabla^2 f(\bar{\boldsymbol{x}}), \nabla^2 g_i(\bar{\boldsymbol{x}})(i\in I), D^2\boldsymbol{G}(\bar{\boldsymbol{x}})$ 分别表示 $f, g_i(i\in I), \boldsymbol{G}$ 在点 $\bar{\boldsymbol{x}}$ 的 Hessian 矩阵，现定义锥

$$\begin{aligned} S = \{\boldsymbol{d}\neq\boldsymbol{0}: & \nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} = 0,\ i\in I^{+} \\ & \nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d}\leqslant 0,\ i\in I^{0} \\ & \bar{\boldsymbol{U}}\cdot D\boldsymbol{G}(\bar{\boldsymbol{x}})\boldsymbol{d} = 0,\ \boldsymbol{d}\in\mathbb{R}^m\} \end{aligned}$$

若

$$\boldsymbol{d}^{\mathrm{T}}\nabla^2\bar{L}(\bar{\boldsymbol{x}})\boldsymbol{d} > 0 \qquad \forall\,\boldsymbol{d}\in S$$

则 $\bar{\boldsymbol{x}}$ 为问题(1) 的严格局部极小点.

证 假设 $\bar{\boldsymbol{x}}$ 不是问题(1) 的严格局部极小点，则存在序列$\{\boldsymbol{x}_k\}\subset\Omega$ 使得 $\boldsymbol{x}_k\to\bar{\boldsymbol{x}}(\boldsymbol{x}_k\neq\bar{\boldsymbol{x}})$ 并且

$$f(\boldsymbol{x}_k)\leqslant f(\bar{\boldsymbol{x}}) \qquad \forall\, k\in\mathbb{N}^*$$

令

$$\boldsymbol{d}_k = \frac{(\boldsymbol{x}_k - \bar{\boldsymbol{x}})}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|} \qquad \lambda_k = \|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|$$

则

$$\boldsymbol{x}_k = \bar{\boldsymbol{x}} + \lambda_k\boldsymbol{d}_k \qquad \|\boldsymbol{d}_k\| = 1 \qquad \lambda_k\to 0^{+}$$

因为 $\|\boldsymbol{d}_k\| = 1$，则存在收敛子列，不失一般性，不妨假设$\{\boldsymbol{d}_k\}$ 收敛，则 $\boldsymbol{d}_k\to\boldsymbol{d}$ 且 $\|\boldsymbol{d}\| = 1$.

进一步，因为 $f, \boldsymbol{g}, \boldsymbol{G}$ 二阶可微，则有

$$0 \geqslant f(\boldsymbol{x}_k) - f(\overline{\boldsymbol{x}}) = \nabla f(\overline{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}_k - \overline{\boldsymbol{x}}) + \frac{1}{2}(\boldsymbol{x}_k - \overline{\boldsymbol{x}})^{\mathrm{T}} \nabla^2 f(\overline{\boldsymbol{x}})(\boldsymbol{x}_k - \overline{\boldsymbol{x}}) + \|\boldsymbol{x}_k - \overline{\boldsymbol{x}}\|^2 \alpha_f(\overline{\boldsymbol{x}};\boldsymbol{x}_k - \overline{\boldsymbol{x}}) \qquad \forall k \in \mathbb{N}^*$$

即对 $\forall k \in \mathbb{N}^*$ 有

$$0 \geqslant f(\overline{\boldsymbol{x}} + \lambda_k \boldsymbol{d}_k) - f(\overline{\boldsymbol{x}}) = \lambda_k \nabla f(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d}_k + \frac{1}{2}\lambda_k^2 \boldsymbol{d}_k{}^{\mathrm{T}} \nabla^2 f(\overline{\boldsymbol{x}}) \boldsymbol{d}_k + \lambda_k^2 \alpha_f(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k) \tag{4}$$

$$0 \geqslant g_i(\overline{\boldsymbol{x}} + \lambda_k \boldsymbol{d}_k) - g_i(\overline{\boldsymbol{x}}) = \lambda_k \nabla g_i(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d}_k + \frac{1}{2}\lambda_k^2 \boldsymbol{d}_k{}^{\mathrm{T}} \nabla^2 g_i(\overline{\boldsymbol{x}}) \boldsymbol{d}_k + \lambda_k^2 \alpha_{g_i}(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k)(i \in I) \tag{5}$$

$$\boldsymbol{G}(\overline{\boldsymbol{x}} + \lambda_k \boldsymbol{d}_k) - \boldsymbol{G}(\overline{\boldsymbol{x}}) = \lambda_k D\boldsymbol{G}(\overline{\boldsymbol{x}})\boldsymbol{d}_k + \frac{1}{2}\lambda_k^2 D^2 \boldsymbol{G}(\overline{\boldsymbol{x}})(\boldsymbol{d}_k, \boldsymbol{d}_k) + \lambda_k^2 \boldsymbol{\alpha}_{\boldsymbol{G}}(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k) \tag{6}$$

式(6) 与 $\overline{\boldsymbol{U}}$ 作内积有

$$0 \geqslant \overline{\boldsymbol{U}} \cdot (\boldsymbol{G}(\overline{\boldsymbol{x}} + \lambda_k \boldsymbol{d}_k) - \boldsymbol{G}(\overline{\boldsymbol{x}})) = \overline{\boldsymbol{U}} \cdot (\lambda_k D\boldsymbol{G}(\overline{\boldsymbol{x}})\boldsymbol{d}_k + \frac{1}{2}\lambda_k^2 D^2 \boldsymbol{G}(\overline{\boldsymbol{x}})(\boldsymbol{d}_k, \boldsymbol{d}_k) + \lambda_k^2 \boldsymbol{\alpha}_{\boldsymbol{G}}(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k)) \tag{7}$$

其中 $\alpha_f: \mathbb{R}^m \longrightarrow \mathbb{R}$，$\alpha_{g_i}: \mathbb{R}^m \longrightarrow \mathbb{R}\ (i \in I)$，$\boldsymbol{\alpha}_{\boldsymbol{G}}: \mathbb{R}^m \longrightarrow S^n$，$\lim\limits_{\lambda_k \longrightarrow 0+} \alpha_f(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k) = 0$，$\lim\limits_{\lambda_k \longrightarrow 0+} \alpha_{g_i}(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k) = 0(i \in I)$，$\lim\limits_{\lambda_k \longrightarrow 0+} \boldsymbol{\alpha}_{\boldsymbol{G}}(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k) = 0$. 对式(4),(5),(7) 除 λ_k，令 $k \longrightarrow \infty$ 有

$$\nabla f(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} \leqslant 0,\ \nabla g_i(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} \leqslant 0\ (i \in I),\ \overline{\boldsymbol{U}} \cdot D\boldsymbol{G}(\overline{\boldsymbol{x}})\boldsymbol{d} \leqslant 0 \tag{8}$$

由(3) 式有

$$\nabla f(\overline{\boldsymbol{x}}) + \sum_{i \in I} \overline{\boldsymbol{\mu}}_i \nabla g_i(\overline{\boldsymbol{x}}) + \overline{\boldsymbol{U}} \cdot D\boldsymbol{G}(\overline{\boldsymbol{x}}) = 0$$

从而得到

$$\nabla f(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} + \sum_{i \in I} \overline{\boldsymbol{\mu}}_i \nabla g_i(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} + \overline{\boldsymbol{U}} \cdot D\boldsymbol{G}(\overline{\boldsymbol{x}})\boldsymbol{d} = 0$$

再由(8) 式可以得到

$$\nabla f(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} = 0,\ \nabla g_i(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} = 0(i \in I^+)$$

$$\nabla g_i(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} \leqslant 0(i \in I^0),\ \overline{\boldsymbol{U}} \cdot D\boldsymbol{G}(\overline{\boldsymbol{x}})\boldsymbol{d} = 0$$

所以 $\boldsymbol{d} \in S$. (5) 式乘 $\overline{\boldsymbol{\mu}}_i(i \in I)$，再与(4),(7) 式相加，由

$$\nabla f(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d}_k + \sum_{i \in I} \overline{\boldsymbol{\mu}}_i \nabla g_i(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d}_k + \overline{\boldsymbol{U}} \cdot D\boldsymbol{G}(\overline{\boldsymbol{x}})\boldsymbol{d}_k = 0$$

可得

$$0 \geqslant \frac{\lambda_k^2}{2} \boldsymbol{d}_k{}^{\mathrm{T}} \nabla^2 \overline{L}(\overline{\boldsymbol{x}})\boldsymbol{d}_k + \lambda_k^2 (\alpha_f(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k) + \sum_{i \in I} \overline{\boldsymbol{\mu}}_i \alpha_{g_i}(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k) + \overline{\boldsymbol{U}} \cdot \boldsymbol{\alpha}_{\boldsymbol{G}}(\overline{\boldsymbol{x}};\lambda_k \boldsymbol{d}_k)) \tag{9}$$

(9) 式除 λ_k^2，令 $k \to \infty$，得到

$$\boldsymbol{d}^{\mathrm{T}} \nabla^2 \overline{L}(\overline{\boldsymbol{x}})\boldsymbol{d} \leqslant 0$$

其中

$$\|\boldsymbol{d}\| = 1,\ \boldsymbol{d} \in S$$

与条件矛盾，因此 $\overline{\boldsymbol{x}}$ 为问题(1) 的严格局部极小点.

注 2　从定理 2 的证明中知可以进一步限制锥 S 即在锥 S 中添加约束 $\nabla f(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} = 0$，尽管这是有效的，但是在定理 2 中并没有对锥 S 进行进一步的限制，原因在于若广义 K-K-T 条件成立且 $\boldsymbol{d} \in S$，则 $\nabla f(\overline{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} = 0$ 自动成立.

注 3　如果考虑的是无约束优化问题，则定理 2 可简化为：若 $\nabla f(\overline{\boldsymbol{x}}) = \boldsymbol{0}$ 且 $\nabla^2 f(\overline{\boldsymbol{x}})$ 正定，则可断定 $\overline{\boldsymbol{x}}$ 是严格局部极小点.

3 结 论

在本文中对非凸半定规划问题(1)的最优性充分条件进行了讨论.定理1在不变凸性的假设下证明了广义K-K-T条件是全局最优性的充分条件;定理2则给出了问题(1)的最优性二阶充分条件.

参考文献:

[1] 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.

[2] FORSGREN A. Optimality Conditions for Nonconvex Semidefinite Programming [J]. Mathematical Programming,2000,88(1):105-128.

[3] 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.

[4] 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.

[5] BAZARAA M S,SHERALI H D,SHETTY C M. Nonlinear Programming:Theory and Algorithms [M]. 3rd ed. New Jersey:John Wiley & Sons,2006.

The Optimality Conditions for Nonconvex Semidefinite Programming

LI Yong-ling, LUO Hong-lin, XIANG Yan-ning

Department of Mathematics, Chongqing Normal University, Chongqing 401331, China

Abstract: This paper is devoted to study first and second order sufficient condition for nonconvex semidefinite programming problems. Under invex convexity assumption, we prove that the generalized Karush-Kuhn-Tucker condition is first order sufficient condition for the existence of the global optimal solution for nonconvex semidefinite programming problems. Under assumptions of no generalized convexity, the second order sufficient condition for the existence of the strict local optimal solution is derived.

Key words: nonconvex semidefinite programming; optimality conditions; invex

责任编辑 张 枸