运筹学与控制论

DOI:10.11721/cqnuj20160502

一类混合整数约束三次规划问题的全局最优性条件*

陈　露，李国权
(重庆师范大学 数学学院，重庆 401331)

摘要:通过构造目标函数的二次上估计函数和二次下估计函数,给出了一类混合整数三次规划问题的全局最优性条件。首先利用二次上估计函数给出全局最优性必要条件,其次再利用二次下估计函数获得全局最优性充分条件。最后给出一个数值例子来说明如何利用所给出的全局最优性条件来判定一个给定的点是否是全局最优解。

关键词:全局最优性条件;混合整数三次规划;二次上估计;二次下估计

中图分类号:O221.1　　**文献标志码**:A　　**文章编号**:1672-6693(2016)05-0001-06

1 预备知识

考虑如下带有混合整数约束的三次规划问题(CP):

$$\min f(\boldsymbol{x})=\sum_{i=1}^{n}\frac{1}{3}\alpha_i x_i^3+\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x},$$

$$\text{s.t. } \boldsymbol{x}\in D_B:=\{(x_1,\cdots,x_n)^{\mathrm{T}}\mid x_i\in[u_i,v_i],i\in I;x_i\in\{u_i,u_i+1,\cdots,v_i\},i\in J\}。$$

其中 $\alpha_i,u_i,v_i\in\mathbf{R},\boldsymbol{b}\in\mathbf{R}^n,\boldsymbol{A}\in S^n$,$S^n$ 为所有 $n\times n$ 阶实对称矩阵构成的集合,$u_i<v_i$,$i=1,2,\cdots,n$,u_i,v_i 都为整数,且 $I\cap J=\varnothing$,$I\cup J=\{1,2,\cdots,n\}$。

三次规划问题(CP)包含了一大类的最优化问题,包括二次优化问题和组合优化问题,它在三次多项式近似优化、凸优化、工程设计和结构优化等领域有着广泛应用。此外,由于二次优化问题是三次优化问题的特殊情形,所以关于三次优化问题的研究成果可以应用到二次优化问题,但是由于带有非负特征值的二次规划问题是NP-难的,因此三次规划问题(CP)也是 NP-难的。近年来,一些文献研究了特殊形式的三次规划问题的全局最优性条件。例如,文献[1-2]利用 L-次微分的方法,给出了当带箱子约束或双值约束特殊三次规划问题的全局最优性充分条件和全局最优性必要条件,但是这些条件比较难验证;Quan 等人[3]也利用 L-次微分的方法,给出了当 $I=\varnothing$时的一些特殊多项式(包括三次)整数规划问题的全局最优性充分条件和全局最优性必要条件,所给出的条件是很容易验证的;Wu 等人[4]给出了带有混合变量的三次多项式优化问题的一些必要局部最优性条件和必要全局最优性条件,并结合局部优化方法和一些辅助函数给出了求解三次多项式优化问题的一个全局优化方法;Zhou 等人[5]通过构造原目标函数的二次上估计函数和二次下估计函数给出了带箱子约束或双值约束特殊三次规划问题的全局最优性必要条件和全局最优性充分条件;Zhang 等人[6]研究了一类特殊的双值约束非凸三次规划问题,利用问题结构的特殊性给出了该问题的一个全局最优充分必要条件;Li 等人[7]利用 L-次微分和 L-正则锥,给出了混合整数二次规划问题的全局最优充分性条件和全局最优性必要条件。

受上述文献的启发,本文将利用二次上(下)估计函数研究混合整数约束三次规划问题的全局最优性条件,所给出的必要条件和充分条件比较容易验证,并且推广了文献[5-7]中相应的一些结果。

下面首先给出要用的一些基本符号与结果:$\mathbf{R}$ 表示实线性空间,$\mathbf{R}^n$ 表示 n 维欧氏空间。对于向量 $\boldsymbol{x},\boldsymbol{y}\in\mathbf{R}^n$,$\boldsymbol{x}\geqslant\boldsymbol{y}$ 是指对任意的 $i=1,\cdots,n$,$x_i\geqslant y_i$,记号 $\boldsymbol{A}\geqslant\boldsymbol{B}$ 指 $\boldsymbol{A}-\boldsymbol{B}$ 是半正定矩阵。单位矩阵用 $\boldsymbol{I}$ 表示,所有 $n\times n$ 阶实

* **收稿日期**:2015-06-12　**修回日期**:2016-05-25　**网络出版时间**:2016-07-13 14:06
资助项目:国家自然科学基金(No. 11471062;No. 11401064);重庆市自然科学基金(No. cstc2013jcyjA-00021);重庆市教委科学技术研究项目(No. KJ1500302)
作者简介:陈露,女,研究方向为最优化理论与算法,E-mail:1203702404@qq.com;**通信作者**:李国权,副教授,E-mail:ligq@cqnu.edu.cn
网络出版地址:http://www.cnki.net/kcms/detail/50.1165.N.20160713.1406.062.html

对称矩阵空间用 S^n 表示,用 $\mathrm{diag}(\alpha_1,\cdots,\alpha_n)$ 表示对角元素为 $\alpha_1,\cdots,\alpha_n$ 的对角矩阵,用 $\boldsymbol{c}(\boldsymbol{x})=(\alpha_1x_1^2,\cdots,\alpha_nx_n^2)^{\mathrm{T}}$ 表示 $\sum_{i=1}^{n}\frac{1}{3}\alpha_ix_i^3$ 在 $\boldsymbol{x}$ 点处的梯度,用 $\boldsymbol{C}(\boldsymbol{x})=\mathrm{diag}(2\alpha_1x_1,\cdots,2\alpha_nx_n)$ 表示 $\sum_{i=1}^{n}\frac{1}{3}\alpha_ix_i^3$ 在 $\boldsymbol{x}$ 点处的 Hessian 矩阵。

定义 1[5] 设 $D\subseteq\mathbf{R}^n$,$\bar{x}\in D$,令 $h:\mathbf{R}^n\to\mathbf{R}$,$f:\mathbf{R}^n\to\mathbf{R}$,如果 h 是一个二次函数,且对 $\forall\,\boldsymbol{x}\in D$,$f(\boldsymbol{x})\leqslant h(\boldsymbol{x})$,$f(\bar{\boldsymbol{x}})=h(\bar{\boldsymbol{x}})$,则函数 h 是函数 f 在 D 上$\bar{\boldsymbol{x}}$处的二次上估计函数;如果 h 是一个二次函数,且对 $\forall\,\boldsymbol{x}\in D$,$f(\boldsymbol{x})\geqslant h(\boldsymbol{x})$,$f(\bar{\boldsymbol{x}})=h(\bar{\boldsymbol{x}})$,则函数 h 是函数 f 在 D 上$\bar{\boldsymbol{x}}$处的二次下估计函数。

本文特别选择如下形式的二次函数:$\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{x}$,其中 $\boldsymbol{Q}=\mathrm{diag}(q_1,\cdots,q_n)$,$q_i\in\mathbf{R}$,$\boldsymbol{d}\in\mathbf{R}^n$。

引理 1[5] 令 $\bar{x}\in D:=\prod_{i=1}^{n}[u_i,v_i]$,$\forall\,i=1,\cdots,n$,$u_i,v_i\in\mathbf{R}$,假设存在一个对角矩阵 $\boldsymbol{Q}$ 使得对任意的 $\boldsymbol{x}\in D$,$(\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{Q})$是半负定矩阵。令 $l(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+(\boldsymbol{A}-\boldsymbol{Q})\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}$,$\boldsymbol{x}\in\mathbf{R}^n$,则函数 $h:\mathbf{R}^n\to\mathbf{R}$,$h(\boldsymbol{x})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})+f(\bar{\boldsymbol{x}})$是 f 在 D 上$\bar{\boldsymbol{x}}$处的二次上估计函数。

引理 2[5] 令 $\bar{x}\in D:=\prod_{i=1}^{n}[u_i,v_i]$,$\forall\,i=1,\cdots,n$,$u_i,v_i\in\mathbf{R}$,假设存在一个对角矩阵 $\boldsymbol{Q}$ 使得对任意的 $\boldsymbol{x}\in D$,$(\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{Q})$是半正定矩阵。令 $l(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+(\boldsymbol{A}-\boldsymbol{Q})\bar{\boldsymbol{x}})^{\mathrm{T}}x$,$\boldsymbol{x}\in\mathbf{R}^n$,则函数 $h:\mathbf{R}^n\to\mathbf{R}$,$h(\boldsymbol{x})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})+f(\bar{\boldsymbol{x}})$是 f 在 D 上$\bar{\boldsymbol{x}}$处的二次下估计函数。

注 1 对$\bar{\boldsymbol{x}}\in D_B$,由于 $D_B\subseteq D$,则 f 在 D 上$\bar{\boldsymbol{x}}$处的二次上估计函数也是 f 在 D_B 上$\bar{\boldsymbol{x}}$处的二次上估计函数,f 在 D 上$\bar{\boldsymbol{x}}$处的二次下估计函数也是 f 在 D_B 上$\bar{\boldsymbol{x}}$处的二次下估计函数。

2 必要条件与二次上估计

本节利用原目标函数的二次上估计来推导问题(CP)的全局最优性必要条件。

引进下面的一些记号:

$$\tilde{\tilde{\boldsymbol{x}}}:=\begin{cases}-1,\bar{x}_i=u_i\\1,\bar{x}_i=v_i\\\mathrm{sign}(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i),\bar{x}_i\in(u_i,v_i)\end{cases},\tilde{\tilde{\boldsymbol{X}}}:=\mathrm{diag}(\tilde{\tilde{x}}_1,\cdots,\tilde{\tilde{x}}_n),\boldsymbol{p}_{\bar{\boldsymbol{x}}}:=(p_{\bar{x}_1},\cdots,p_{\bar{x}_n})^{\mathrm{T}},$$

$$\boldsymbol{p}_{\bar{x}_i}:=\begin{cases}\tilde{\tilde{x}}_i\dfrac{\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i}{u_i-v_i},i\in I\\\max\left\{\tilde{\tilde{x}}_i(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i),\tilde{\tilde{x}}_i\dfrac{\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i}{u_i-v_i}\right\},i\in J\end{cases},$$

其中 $\mathrm{sign}(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i)=\begin{cases}-1,(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})<0\\0,(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})=0\\1,(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})>0\end{cases}$,$(\boldsymbol{A}\bar{\boldsymbol{x}})_i$ 表示 $\boldsymbol{A}\bar{\boldsymbol{x}}$的第 i 个分量。

对 $\boldsymbol{Q}=\mathrm{diag}(q_1,\cdots,q_n)$,$q_i\in\mathbf{R}$,$i=1,\cdots,n$,令$\tilde{q}_i=\begin{cases}\min\{0,q_i\},i\in I\\q_i,i\in J\end{cases}$,$\tilde{\boldsymbol{Q}}=\mathrm{diag}(\tilde{q}_1,\cdots,\tilde{q}_n)$。

定理 1 对于问题(CP),令$\bar{\boldsymbol{x}}\in D_B$,假设存在一个对角矩阵 $\boldsymbol{Q}$,使得 $\boldsymbol{Q}=\mathrm{diag}(q_1,\cdots,q_n)$,$q_i\in\mathbf{R}$,$i=1,\cdots,n$,且对任意 $\boldsymbol{x}\in D_B$,$(\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{Q})$是半负定矩阵,其中 $\boldsymbol{C}(\boldsymbol{x})=\mathrm{diag}(2\alpha_1x_1,\cdots,2\alpha_nx_n)$。如果$\bar{\boldsymbol{x}}$是问题(CP)的全局极小点,则[NC1] $\mathrm{diag}(\boldsymbol{p}_{\bar{\boldsymbol{x}}})\leqslant\frac{1}{2}\tilde{\boldsymbol{Q}}$成立。

证明 对$\bar{\boldsymbol{x}}\in D_B$,令 $l(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+(\boldsymbol{A}-\boldsymbol{Q})\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}$,$h(\boldsymbol{x})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})+f(\bar{\boldsymbol{x}})$,由引理 1 知 $h(\boldsymbol{x})$是 f 在 D 上$\bar{\boldsymbol{x}}$处的一个二次上估计函数且 $h(\boldsymbol{x})$是 f 在 D_B 上$\bar{\boldsymbol{x}}$处的一个二次上估计函数。则有

$$f(\boldsymbol{x})-f(\bar{\boldsymbol{x}})\leqslant h(\boldsymbol{x})-h(\bar{\boldsymbol{x}})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})=$$

$$\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+(\boldsymbol{A}-\boldsymbol{Q})\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}-\frac{1}{2}\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{Q}\bar{\boldsymbol{x}}-(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+(\boldsymbol{A}-\boldsymbol{Q})\bar{\boldsymbol{x}})^{\mathrm{T}}\bar{\boldsymbol{x}}=$$

$$\frac{1}{2}(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{x}-\bar{\boldsymbol{x}})+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+\boldsymbol{A}\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})=\sum_{i=1}^{n}\left(\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)。$$

因为$\bar{x}$是问题(CP)的全局极小点，则对任意$\bar{x}\in D_B$，有

$$\sum_{i=1}^{n}\left(\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)\geqslant 0。\tag{1}$$

则对$\forall i=1,\cdots,n$，(1)式成立当且仅当：

1) $\forall i\in I,\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\geqslant 0,\forall x_i\in[u_i,v_i]$；

2) $\forall i\in J,\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\geqslant 0,\forall x_i\in[u_i,u_i+1,\cdots,v_i]$。

事实上，如果条件 1)和 2)成立，明显有(1)式成立。相反，如果(1)式成立，若存在$i_0\in I$且$y_{i_0}\in[u_i,v_i]$或存在$i_0\in I$且$y_{i_0}\in\{u_i,\cdots,v_i\}$使得$\frac{1}{2}q_{i_0}(x_{i_0}-\bar{x}_{i_0})^2+[\alpha_{i_0}\bar{x}_{i_0}^2+b_{i_0}+(\boldsymbol{A}\bar{\boldsymbol{x}})_{i_0}](x_{i_0}-\bar{x}_{i_0})\leqslant 0$，则令$x_{i_0}=y_{i_0}$且$x_i=\bar{x}_i$，$i=1,\cdots,n,i\neq i_0$，则$\boldsymbol{x}=(x_1,\cdots,x_n)^{\mathrm{T}}\in D_B$，且

$$\sum_{i=1}^{n}\left(\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)=$$
$$\frac{1}{2}q_{i_0}(x_{i_0}-\bar{x}_{i_0})^2+[\alpha_{i_0}\bar{x}_{i_0}^2+b_{i_0}+(\boldsymbol{A}\bar{\boldsymbol{x}})_{i_0}](x_{i_0}-\bar{x}_{i_0})\leqslant 0,$$

这与(1)式矛盾。

通过讨论$\bar{x}_i$的 3 种情况：$\bar{x}_i=u_i,\bar{x}_i=v_i,\bar{x}_i\in(u_i,v_i)$，且讨论$q_i$的两种情况：$q_i\geqslant 0$和$q_i<0$，容易验证条件 1)等价于$\tilde{\bar{x}}_i(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i)\leqslant\frac{v_i-u_i}{2}\min\{0,q_i\}$。

通过讨论$\bar{x}_i$的 3 种情况：$\bar{x}_i=u_i,\bar{x}_i=v_i,\bar{x}_i\in\{u_i+1,\cdots,v_i-1\}$，且讨论$q_i$的两种情况：$q_i\geqslant 0$和$q_i<0$，容易验证条件 2)等价于$\max\{(v_i-u_i)\tilde{\bar{x}}_i(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i),\tilde{\bar{x}}_i(\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i)\}\leqslant\frac{v_i-u_i}{2}q_i$。

因此(1)式等价于[NC1]。　　证毕

由定理 1 可以得到以下两个推论，这与文献[5]中的定理 3.1 和定理 5.1 一致。

推论 1[5]　对于问题(CP)，令$\bar{\boldsymbol{x}}\in D_B$，若$J=\varnothing$且存在一个对角矩阵$\boldsymbol{Q}$，使得$\boldsymbol{Q}=\mathrm{diag}(q_1,\cdots,q_n)$，$q_i\in\mathbf{R}$，$i=1,\cdots,n$，且对任意$\boldsymbol{x}\in D_B$，$(\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{Q})$是半负定矩阵，其中$\boldsymbol{C}(\boldsymbol{x})=\mathrm{diag}(2\alpha_1x_1,\cdots,2\alpha_nx_n)$。如果$\bar{\boldsymbol{x}}$是问题(CP)的全局极小点，则对任意的$i=1,\cdots,n$有[NC2] $\tilde{\bar{x}}_i[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i]\leqslant\frac{1}{2}\tilde{q}_i(v_i-u_i)$。

推论 2[5]　对于问题(CP)，令$\bar{\boldsymbol{x}}\in D_B$，若$I=\varnothing$且存在一个对角矩阵$\boldsymbol{Q}$，使得$\boldsymbol{Q}=\mathrm{diag}(q_1,\cdots,q_n)$，$q_i\in\mathbf{R}$，$i=1,\cdots,n$，且对任意$\boldsymbol{x}\in D_B$，$(\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{Q})$是半负定矩阵，其中$\boldsymbol{C}(\boldsymbol{x})=\mathrm{diag}(2\alpha_1x_1,\cdots,2\alpha_nx_n)$。如果$\bar{\boldsymbol{x}}$是问题(CP)的全局极小点，则有条件[NC3] $\mathrm{diag}(\boldsymbol{p}_{\bar{x}})\leqslant\frac{1}{2}\boldsymbol{Q}$成立。

定义 1[11]　矩阵$\boldsymbol{A}=(a_{ij})\in S^n$称为对角占优矩阵，如果对任意的$i=1,\cdots,n$，$|a_{ii}|\geqslant\sum_{j=1,j\neq i,}^{n}|a_{ij}|$。

注 2[11]　每一个有非负对角元的对角占优矩阵是半正定矩阵。

命题 1[5]　对任意的$\boldsymbol{x}\in D_B$，若$\boldsymbol{Q}=\mathrm{diag}(q_1,\cdots,q_n)$满足如下条件：(i) $\forall i=1,\cdots,n,q_i\geqslant 2\alpha_ix_i+a_{ii}$；(ii) $\forall i=1,\cdots,n,|2\alpha_ix_i+a_{ii}-q_i|\geqslant\sum_{j=1,j\neq i}^{n}|a_{ij}|$，则对任意$\boldsymbol{x}\in D_B$，$\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{Q}$是负半定矩阵。

对任意$i=1,\cdots,n$，定义$\beta_i=\max\{2\alpha_ix_i:\boldsymbol{x}\in D_B\}+a_{ii}+\sum_{j=1,j\neq i}^{n}|a_{ij}|$，$\boldsymbol{H}=\mathrm{diag}(\beta_1,\cdots,\beta_n)$，$\tilde{\beta}_i=\begin{cases}\min\{0,\beta_i\},i\in I\\ \beta_i,i\in J\end{cases}$，$\tilde{\boldsymbol{H}}=\mathrm{diag}(\tilde{\beta}_1,\cdots,\tilde{\beta}_n)$。

定理 2　对于问题(CP)，令$\bar{\boldsymbol{x}}\in D_B$，如果$\bar{\boldsymbol{x}}$是问题(CP)的全局极小点，则条件[NC4] $\mathrm{diag}(\boldsymbol{p}_{\bar{x}})\leqslant\frac{1}{2}\tilde{\boldsymbol{H}}$成立，其中$\boldsymbol{H}$与$\tilde{\boldsymbol{H}}$分别如前面所定义。

证明　对$\bar{\boldsymbol{x}}\in D_B$，令$\varphi(\boldsymbol{x})=f(\boldsymbol{x})-l(\boldsymbol{x})$，$l(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{x}+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+(\boldsymbol{A}-\boldsymbol{H})\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}$，$h(\boldsymbol{x})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})+$

$f(\bar{\boldsymbol{x}})$,注意到 $\varphi(\boldsymbol{x})$ 是 D 上二次连续可微函数。则对任意 $\boldsymbol{x}\in D$,有

$$\nabla^2\varphi(\boldsymbol{x})=(v_{ij}(\boldsymbol{x}))_{n\times n}=\begin{pmatrix}2\alpha_1x_1+a_{11}-\beta_1 & a_{12} & \cdots & a_{1n}\\ a_{21} & 2\alpha_2x_2+a_{22}-\beta_2 & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & 2\alpha_nx_n+a_{nn}-\beta_n\end{pmatrix},$$

由 β_i 的定义可得 $\forall\,\boldsymbol{x}\in D$, $2\alpha_ix_i+a_{ii}-\beta_i\leqslant-\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|\leqslant 0$, $i=1,\cdots,n$,则 $\nabla^2\varphi(\boldsymbol{x})$ 是一个对角元非正的对角占优矩阵。因此,对 $\boldsymbol{x}\in D$,$(\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{H})\leqslant\boldsymbol{0}$。由引理1知 $h(\boldsymbol{x})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})+f(\bar{\boldsymbol{x}})$ 是 f 在 D 上 $\bar{\boldsymbol{x}}$ 处的一个二次上估计函数,且是 f 在 D_B 上 $\bar{\boldsymbol{x}}$ 处的一个二次上估计函数。则有

$$0\leqslant f(\boldsymbol{x})-f(\bar{\boldsymbol{x}})\leqslant h(\boldsymbol{x})-h(\bar{\boldsymbol{x}})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})=\frac{1}{2}(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{H}(\boldsymbol{x}-\bar{\boldsymbol{x}})+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+\boldsymbol{A}\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})=$$

$$\sum_{i=1}^{n}\left(\frac{1}{2}\beta_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)。$$

因为 $\bar{\boldsymbol{x}}$ 是问题(CP)的全局极小点,则对任意 $\boldsymbol{x}\in D_B$,有

$$\sum_{i=1}^{n}\left(\frac{1}{2}\beta_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)\geqslant 0,\tag{2}$$

则对 $\forall\,i=1,\cdots,n$,(2)式成立当且仅当:

3) $\forall\,i\in I$, $\frac{1}{2}\beta_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\geqslant 0$, $\forall\,x_i\in[u_i,v_i]$;

4) $\forall\,i\in J$, $\frac{1}{2}\beta_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\geqslant 0$, $\forall\,x_i\in[u_i,u_i+1,\cdots,v_i]$。

类似于定理1的讨论方法,可得(NC4)式等价于(2)式。 证毕

3 充分条件与二次下估计

本节利用原目标函数的二次下估计来推导问题(CP)的全局最优性充分条件。

定理3 对于问题(CP),令 $\bar{\boldsymbol{x}}\in D_B$,假设存在一个对角矩阵 $\boldsymbol{Q}$,使得 $\boldsymbol{Q}=\mathrm{diag}(q_1,\cdots,q_n)$,$q_i\in\mathbf{R}$,$i=1,\cdots,n$,且对任意 $\boldsymbol{x}\in D_B$,$(\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{Q})$ 是半正定矩阵,其中 $\boldsymbol{C}(\boldsymbol{x})=\mathrm{diag}(2\alpha_1x_1,\cdots,2\alpha_nx_n)$。如果[SC1] $\mathrm{diag}(\boldsymbol{p}_{\bar{x}})\leqslant\frac{1}{2}\tilde{\boldsymbol{Q}}$ 成立,则 $\bar{\boldsymbol{x}}$ 是问题(CP)的全局极小点。

证明 对 $\bar{\boldsymbol{x}}\in D_B$,令 $l(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+(\boldsymbol{A}-\boldsymbol{Q})\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}$,$h(\boldsymbol{x})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})+f(\bar{\boldsymbol{x}})$,由引理2知 $h(\boldsymbol{x})$ 是 f 在 D 上 $\bar{\boldsymbol{x}}$ 处的一个二次下估计函数且 $h(\boldsymbol{x})$ 是 f 在 D_B 上 $\bar{\boldsymbol{x}}$ 处的一个二次上估计函数。则有

$$f(\boldsymbol{x})-f(\bar{\boldsymbol{x}})\geqslant h(\boldsymbol{x})-h(\bar{\boldsymbol{x}})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})=\frac{1}{2}(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{x}-\bar{\boldsymbol{x}})+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+\boldsymbol{A}\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})=$$

$$\sum_{i=1}^{n}\left(\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)。$$

对任意 $\bar{\boldsymbol{x}}\in D_B$,若有

$$\sum_{i=1}^{n}\left(\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)\geqslant 0,\tag{3}$$

则 $\bar{\boldsymbol{x}}$ 是问题(CP)的全局极小点。

对 $\forall\,i=1,2,\cdots,n$,(3)式成立当且仅当:

1) $\forall\,i\in I$, $\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\geqslant 0$, $\forall\,x_i\in[u_i,v_i]$;

2) $\forall\,i\in J$, $\frac{1}{2}q_i(x_i-\bar{x}_i)^2+[\alpha_i\bar{x}_i^2+b_i+(\boldsymbol{A}\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\geqslant 0$, $\forall\,x_i\in[u_i,u_i+1,\cdots,v_i]$。

类似定理1的证明,(3)式与[SC1]等价,因此如果[SC1]成立,$\bar{\boldsymbol{x}}$ 则是问题(CP)的全局极小点。 证毕

推论3 对于问题(CP),令 $\bar{\boldsymbol{x}}\in D_B$,若对所有的 $i=1,\cdots,n$,都有 $\alpha_i=0$,且存在一个对角矩阵 $\boldsymbol{Q}$,其中 $\boldsymbol{Q}=$

$\mathrm{diag}(q_1,\cdots,q_n)$，$q_i\in\mathbf{R}$，$i=1,\cdots,n$，使得 $\boldsymbol{A}-\boldsymbol{Q}$ 是半正定矩阵，如果[SC2] $\mathrm{diag}(\boldsymbol{p}_{\bar{x}})\leqslant\frac{1}{2}\widetilde{\boldsymbol{Q}}$成立，则$\bar{\boldsymbol{x}}$是问题(CP)的全局极小点。

注 3　充分条件[SC2]与文献[7]给出的命题 2.2 等价。

命题 2[5]　对任意的 $\boldsymbol{x}\in D_B$，若 $\boldsymbol{Q}=\mathrm{diag}(q_1,\cdots,q_n)$满足如下条件：(i) $\forall\, i=1,\cdots,n$，$q_i\leqslant 2\alpha_i x_i+a_{ii}$；(ii) $\forall\, i=1,\cdots,n$，$|2\alpha_i x_i+a_{ii}-q_i|\geqslant\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|$，则对任意 $\boldsymbol{x}\in D_B$，$\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{Q}$ 是正半定矩阵。

对任意 $i=1,\cdots,n$，定义 $\gamma_i=\min\{2\alpha_i x_i:\boldsymbol{x}\in D_B\}+a_{ii}-\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|$，$\boldsymbol{G}=\mathrm{diag}(\gamma_1,\cdots,\gamma_n)$，$\widetilde{\gamma}_i=\begin{cases}\min\{0,\gamma_i\},i\in I\\ \gamma_i,i\in J\end{cases}$，$\widetilde{\boldsymbol{G}}=\mathrm{diag}(\widetilde{\gamma}_1,\cdots,\widetilde{\gamma}_n)$。

定理 4　对于问题(CP)，令$\bar{\boldsymbol{x}}\in D_B$，如果[SC3] $\mathrm{diag}(\boldsymbol{p}_{\bar{x}})\leqslant\frac{1}{2}\widetilde{\boldsymbol{G}}$成立，则$\bar{\boldsymbol{x}}$是问题(CP)的全局极小点。其中 $\boldsymbol{G}$ 与$\widetilde{\boldsymbol{G}}$分别如前面所定义。

证明　对$\bar{\boldsymbol{x}}\in D_B$，令 $\varphi(\boldsymbol{x})=f(\boldsymbol{x})-l(\boldsymbol{x})$，$l(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{x}+(\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+(\boldsymbol{A}-\boldsymbol{G})\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}$，$h(\boldsymbol{x})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})+f(\bar{\boldsymbol{x}})$，注意到 $\varphi(\boldsymbol{x})$是 D 上二次连续可微函数。则对任意 $\boldsymbol{x}\in D$，

$$\nabla^2\varphi(\boldsymbol{x})=(v_{ij}(\boldsymbol{x}))_{n\times n}=\begin{pmatrix}2\alpha_1x_1+a_{11}-\gamma_1 & a_{12} & \cdots & a_{1n}\\ a_{21} & 2\alpha_2x_2+a_{22}-\gamma_2 & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & 2\alpha_nx_n+a_{nn}-\gamma_n\end{pmatrix},$$

由 γ_i 的定义可得 $\forall\,\boldsymbol{x}\in D$，$2\alpha_i x_i+a_{ii}-\gamma_i\geqslant\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|\geqslant 0$，$i=1,\cdots,n$，则$\nabla^2\varphi(\boldsymbol{x})$是一个对角元非负的对角占优矩阵。因此，对 $\boldsymbol{x}\in D$，$(\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{H})\geqslant\boldsymbol{0}$，由引理 2 知 $h(\boldsymbol{x})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})+f(\bar{\boldsymbol{x}})$是 f 在 D 上$\bar{\boldsymbol{x}}$处的一个二次下估计函数，且是 f 在 D_B 上$\bar{\boldsymbol{x}}$处的一个二次下估计函数。则有

$$f(\boldsymbol{x})-f(\bar{\boldsymbol{x}})\geqslant h(\boldsymbol{x})-h(\bar{\boldsymbol{x}})=l(\boldsymbol{x})-l(\bar{\boldsymbol{x}})=\sum_{i=1}^{n}\left(\frac{1}{2}\gamma_i\,(x_i-\bar{x}_i)^2+[\alpha_i\,\bar{x}_i^2+b_i+(\boldsymbol{A}\,\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)。$$

若对 $\forall\, i=1,\cdots,n$，有

$$\sum_{i=1}^{n}\left(\frac{1}{2}\gamma_i\,(x_i-\bar{x}_i)^2+[\alpha_i\,\bar{x}_i^2+b_i+(\boldsymbol{A}\,\bar{\boldsymbol{x}})_i](x_i-\bar{x}_i)\right)\geqslant 0\ ,\tag{4}$$

则$\bar{\boldsymbol{x}}$是问题(CP)的全局极小点。

类似定理 3 的证明，(4)式与[SC3]等价，因此如果[SC3]成立，$\bar{\boldsymbol{x}}$则是问题(CP)的全局极小点。　　证毕

下面将给出一个例子，并用文中所给出的全局最优性条件来判定某给定的点是否是该问题的全局最优解。

例 1　$\min\ f(\boldsymbol{x})=\frac{1}{3}x_2^3-4x_1^2+2x_1x_2-x_2^2-x_1+4x_2$，

$\mathrm{s.t.}\ \ x_1\in[-2,2]$，$x_2\in\{-2,-1,0,1,2\}$。

考虑点$\bar{\boldsymbol{x}}=(2,-2)^{\mathrm{T}}$，则 $I=\{1\}$，$J=\{2\}$，这里 $\boldsymbol{A}=\begin{pmatrix}-8 & 2\\ 2 & -2\end{pmatrix}$，$\boldsymbol{b}=(-1,-4)^{\mathrm{T}}$，$\boldsymbol{c}(\boldsymbol{x})=(0,x_2^2)^{\mathrm{T}}$，$\boldsymbol{C}(\boldsymbol{x})=\mathrm{diag}(0,2x_2)$，由 γ_i 的定义有：$\gamma_1=-10$，$\gamma_2=\min\{2x_2\,|\,\boldsymbol{x}\in D_B\}-2-2=-8$。令 $\boldsymbol{G}=\mathrm{diag}(-10,-8)$，显然对 $\boldsymbol{x}\in D_B$，$\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{G}\geqslant\boldsymbol{0}$，且$\widetilde{\overline{\boldsymbol{X}}}=\mathrm{diag}(1,-1)$，$\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+\boldsymbol{A}\,\bar{\boldsymbol{x}}=(-21,16)^{\mathrm{T}}$，$\widetilde{\boldsymbol{G}}=\mathrm{diag}(\widetilde{\gamma}_1,\widetilde{\gamma}_2)=\mathrm{diag}(-10,-8)$，由 $\boldsymbol{p}_{\bar{x}}$ 的定义有 $\boldsymbol{p}_{\bar{x}}=\left(-\frac{21}{4},-4\right)^{\mathrm{T}}$，显然有 $\mathrm{diag}(\boldsymbol{p}_{\bar{x}})\leqslant\frac{1}{2}\widetilde{\boldsymbol{G}}$，即$\bar{\boldsymbol{x}}$满足充分条件[SC3]，因此$\bar{\boldsymbol{x}}$是该问题的全局极小点。此外，由 β_i 的定义有：$\beta_1=0-8+2=-6$，$\beta_2=\max\{2x_2\,|\,\boldsymbol{x}\in D_B\}-2+2=4$，令 $\boldsymbol{H}=\mathrm{diag}(\beta_1,\beta_2)=\mathrm{diag}(-6,4)$，显然对 $\boldsymbol{x}\in D_B$，$\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{H}\leqslant\boldsymbol{0}$，$\widetilde{\overline{\boldsymbol{X}}}=\mathrm{diag}(1,-1)$，$\boldsymbol{c}(\bar{\boldsymbol{x}})+\boldsymbol{b}+\boldsymbol{A}\,\bar{\boldsymbol{x}}=(-21,16)^{\mathrm{T}}$，$\widetilde{\boldsymbol{H}}=\mathrm{diag}(\widetilde{\beta}_1,\widetilde{\beta}_2)=\mathrm{diag}(-6,4)$，显然有 $\mathrm{diag}(\boldsymbol{p}_{\bar{x}})\leqslant\frac{1}{2}\widetilde{\boldsymbol{H}}$，即$\bar{\boldsymbol{x}}$满足必要条件[NC4]，这说明所提出的充分条件蕴含了必要条件。

但是对于点$\bar{\boldsymbol{y}}=(-2,2)^{\mathrm{T}}$，由 β_i 的定义有：$\beta_1=0-8+2=-6$，$\beta_2=\max\{2y_2 \mid \boldsymbol{y}\in D_B\}-2+2=4$，令 $\boldsymbol{H}=\mathrm{diag}(\beta_1,\beta_2)=\mathrm{diag}(-6,4)$，显然对 $\boldsymbol{x}\in D_B$，$\boldsymbol{C}(\boldsymbol{x})+\boldsymbol{A}-\boldsymbol{H}\leqslant \boldsymbol{0}$，$\tilde{\bar{\boldsymbol{X}}}=\mathrm{diag}(1,-1)$，$\boldsymbol{c}(\bar{\boldsymbol{y}})+\boldsymbol{b}+\boldsymbol{A}\,\bar{\boldsymbol{y}}=(19,0)^{\mathrm{T}}$，$\tilde{\boldsymbol{H}}=\mathrm{diag}(\tilde{\beta}_1,\tilde{\beta}_2)=\mathrm{diag}(-6,4)$，由 $\boldsymbol{p}_{\bar{x}}$的定义有 $\boldsymbol{p}_{\bar{x}}=\left(-\frac{19}{4},0\right)^{\mathrm{T}}$，显然有 $\mathrm{diag}(\boldsymbol{p}_{\bar{x}})\leqslant\frac{1}{2}\tilde{\boldsymbol{H}}$，即$\bar{\boldsymbol{y}}$满足必要条件[NC4]。但是$\bar{\boldsymbol{y}}$不是该问题的全局极小点，因为 $f(\bar{\boldsymbol{x}})<f(\bar{\boldsymbol{y}})$，这说明必要条件不具有充分性。

参考文献：

[1] Wang Y J, Liang Z A. Global optimality conditions for cubic minimization problem with box or binary constrains [J]. Journal of Global Optimization, 2010, 47(4): 583-595.

[2] Zhang X M, Wang Y J, Ma W M. Global sufficient optimality conditions for a special cubic minimization problem[J]. Mathematical Problems in Engineering, 2012, Article ID 871741, 1-16.

[3] Quan J, Wu Z Y, Li G Q. Global optimality conditions for some classes of polynomial integer programming problems [J]. Journal of Industrial and Management Optimization, 2011, 7(1): 67-78.

[4] Wu Z Y, Quan J, Li G Q, et al. Necessary optimality conditions and new optimization methods for cubic polynomial optimization problems with mixed variables[J]. Journal of Optimization Theory and Applications, 2012, 153(2): 408-435.

[5] Zhou X G, Gao B Y. New global optimality conditions for cubic minimization subject to box or binary constrains[J]. Pacific Journal of Optimization, 2011, 8(3): 67-78.

[6] Zhang L, Wang Y, Li G Q. Global optimality conditions for non-convex cubic minimization problem with binary constraints[J]. Operations Research Transactions, 2015, 19 (2): 83-90.

[7] Li G Q, Wu Z Y, Quan J. A new local and global optimization method for mixed integer quadratic programming problems[J]. Applied Mathematics and Computation, 2010, 217: 2501-2512.

[8] Wu Z Y, Bai F S. Globaloptimality conditions for mixed non-convex quadratic programs[J]. Optimization, 2009, 58 (1): 39-47.

[9] Jeyakumar V, Rubinov A M, Wu Z Y. Sufficientglobal optimality conditions for non-convex quadratic minimization problems with box constraints[J]. Journal of Global Optimization, 2006, 36(3): 471-481.

[10] Wu Z Y, Li G Q, Quan J. Global optimality conditions and optimization methods for Quadratic integer programs problems[J]. Journal of Global Optimization, 2011, 51 (3): 549-568.

[11] Dahl G. A note on diagonally dominant matrices[J]. Linear Algebra Application, 1972, 317(1): 217-224.

Operations Research and Cybernetics

Global Optimality Conditions for Mixed Integer Cubic Programming Problems

CHEN Lu, LI Guoquan

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

Abstract: In this paper, we establish some necessary and sufficient global optimality conditions for mixed integer cubic programming problems via quadratic overestimators and underestimators. Firstly, by utilizing quadratic overestimators, we derive some necessary global optimality conditions for mixed integer cubic minimization problems where the cubic objective function contains no cross terms. Then, we obtain some sufficient conditions for global minimizers using quadratic underestimators. Finally, an example is given to illustrate how to use the global optimality conditions to check a given point.

Key words: global optimality condition; mixed integer cubic program; quadratic overestimators; quadratic underestimators

(责任编辑 黄 颖)