浙江理工大学学报，2021，45(2)：249-255
Journal of Zhejiang Sci-Tech University
DOI：10. 3969/j.issn.1673－3851(n).2021. 02.013

求解单二次约束非凸二次规划问题的全局最优 DC 算法

王建国，郑芳英，胡觉亮
（浙江理工大学理学院，杭州 310018）

摘　要：针对单二次约束的非凸二次规划问题，首先提出一种 DC 算法，并证明了该算法收敛到问题的 Karush-Kuhn-Tucker(KKT)点；其次利用 KKT 点提出了寻找新的初始可行点的方法；最后结合此方法，设计了一个求单二次约束非凸二次规划问题全局最优解的 DC 算法。数值结果表明，该全局算法能有效找到大规模单二次约束非凸二次规划问题的全局最优解。

关键词：非凸二次规划；DC 算法；KKT 点；全局最优解
中图分类号：O224　　**文献标志码**：A　　**文章编号**：1673-3851 (2021) 03-0249-07

A globally optimal DC algorithm for solving single quadratic constrained non-convex quadratic programming problems

WANG Jianguo，*ZHENG Fangying*，*HU Jueliang*
(School of Science，Zhejiang Sci-Tech University，Hangzhou 310018，China)

Abstract：For the non-convex quadratic programming problem with single quadratic constraints，firstly，a difference of convex function (DC) algorithm is proposed，and it is proved that the algorithm converges to the Karush-Kuhn-Tucker (KKT) point of the problem. Secondly，using the KKT point，a method to find a new initial feasible point is proposed. Finally，combined with this method，a DC algorithm for finding the global optimal solution of the single quadratic constrained non-convex quadratic programming problem is designed. Numerical experiments show that the proposed global algorithm can effectively find the global optimal solution of large-scale single quadratic constrained non-convex quadratic programming problems.

Key words：non-convex quadratic programming；DC algorithm；KKT point；global optimal solution

0　引　言

本文考虑如下单二次约束的非凸二次规划问题：

$$\begin{cases}\min f(\boldsymbol{x})=\dfrac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}_0\boldsymbol{x}+\boldsymbol{b}_0^{\mathrm{T}}\boldsymbol{x}+c_0\\ \text{s.t. } g(\boldsymbol{x})=\dfrac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{x}+\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{x}+c_1\leqslant 0\end{cases}\tag{1}$$

收稿日期：2020－10－12　　**网络出版日期**：2021－01－05
基金项目：浙江省自然科学基金项目(LY19A010025)
作者简介：王建国(1996－　)，男，安徽濉溪人，硕士研究生，主要从事非线性规划理论与算法方面的研究。

通信作者：郑芳英，E-mail：zfy@zstu.edu.cn

其中：$\boldsymbol{A}_0$ 为 $n\times n$ 阶实对称矩阵，$\boldsymbol{A}_1$ 为 $n\times n$ 阶正定对称矩阵，$\boldsymbol{b}_0,\boldsymbol{b}_1\in\mathbb{R}^n$，$c_0,c_1\in\mathbb{R}$，$c_1<0$，而且 $g(\boldsymbol{x})$ 满足 Slater 条件，即存在 $x\in\mathbb{R}^n$，使得 $g(\boldsymbol{x})<0$。

非凸二次规划问题在工程、经济管理和金融等领域中有广泛的应用，如管理科学中的生产计划问题，金融中的投资组合问题和系统性风险估计问题，电器工程中的集成电路芯片设计问题。很多经典的组合优化问题如最大割问题和二次指派问题，都可以归结为非凸二次规划问题。当 $\boldsymbol{A}_1$ 为单位矩阵 I 且 $\boldsymbol{b}_1=(0,0,\cdots,0)^{\mathrm{T}}$ 时，问题(1)即为经典的信赖域子问题，在文献中已有多项式时间算法[1-3]。Moré 等[1]首次提出了求解信赖域子问题的原始-对偶算法，其思想主要是利用 Safeguarding 技术并结合牛顿法来计算拉格朗日函数的鞍点，该算法能在有限步内得到原问题的一个近似解。Martínez[4]研究了信赖域子问题的局部非全局解的性质，证明了该问题最多只有一个局部非全局解，但局部非全局解的充分条件和必要条件之间存在间隙；Wang 等[5]消除了这一间隙，建立了局部最优解的充要条件。Lucidi 等[6]证明了信赖域子问题的目标函数在 Karush-Kuhn-Tucker (KKT)点处最多有 $2m+2$ 个不同的值，其中 m 是矩阵的不同负特征值的个数。Tao 等[7]提出了求解信赖域子问题的 DC 算法，数值结果表明该算法能有效地找到该问题的全局最优解。An 等[8]提出了 DC 算法的最新理论研究成果及实际应用。

众所周知，带有多个凸二次约束的非凸二次规划问题(Quadratically constrained non -convex quadratic programming，QCQP)一般是 NP-难的[9]，甚至找它的局部解也是 NP-难的[10]。Jeyakumar 等[11]提出并证明了非凸二次约束的非凸二次规划问题的全局最优性条件。QCQP 的几个特殊子类已证明是多项式时间可解的[12-13]。根据经典的 S-引理[14]，问题(1)的最优值与下述半定规划问题(Semi-definite，SD)的最优值相等：

$$\min\left\{\begin{pmatrix} c_0 & \frac{1}{2}\boldsymbol{b}_0^{\mathrm{T}} \\ \frac{1}{2}\boldsymbol{b}_0 & \frac{1}{2}\boldsymbol{A}_0 \end{pmatrix}\cdot\boldsymbol{X}:\begin{pmatrix} c_1 & \frac{1}{2}\boldsymbol{b}_1{}^{\mathrm{T}} \\ \frac{1}{2}\boldsymbol{b}_0 & \frac{1}{2}\boldsymbol{A}_1 \end{pmatrix}\cdot\boldsymbol{X}\leqslant 0,\boldsymbol{X}\succeq 0,\boldsymbol{X}_{11}=1\right\} \tag{2}$$

其中：$\boldsymbol{X}=\begin{pmatrix}1\\ \boldsymbol{x}\end{pmatrix}(1\quad \boldsymbol{x}^{\mathrm{T}})$；$\boldsymbol{X}\succeq 0$ 表示 $\boldsymbol{X}$ 为半正定矩阵；$\boldsymbol{X}_{11}=1$ 表示矩阵 $\boldsymbol{X}$ 的第一行第一列元素为 1。在 Slater 条件下，问题(1)与下述拉格朗日对偶问题半定规划(Semi-definite programming，SDP)有相同的最优值[14]：

$$\begin{cases}\max\ \tau, \\ \text{s.t.}\begin{pmatrix} c_0+(\lambda+\mu)c_1-\tau & \frac{1}{2}(\boldsymbol{b}_0^{\mathrm{T}}+(\lambda+\mu)\boldsymbol{b}_1^{\mathrm{T}}) \\ \frac{1}{2}(\boldsymbol{b}_0+(\lambda+\mu)\boldsymbol{b}_1) & \frac{1}{2}(\boldsymbol{A}_0+(\lambda+\mu)\boldsymbol{A}_1) \end{pmatrix}\succeq 0, \\ \boldsymbol{A}_0+\lambda\,\boldsymbol{A}_1\succeq 0,(\lambda,\mu)\geqslant 0. \end{cases}$$

郑小金[15]证明了通过求解问题(2)不一定能得到原问题的全局最优解，并对单二次约束的非凸二次规划提出了一种基于最优 DC 分解(即把目标函数分解成两个凸函数之差)的精确算法。该算法先对非凸二次目标函数进行 DC 分解，然后对 DC 分解中凹的部分进行线性下逼近得到一个凸二次松弛问题，最后证明了求解凸二次松弛问题可得到原问题的最优解。值得一提的是，该算法涉及求解一个半定规划问题，故不能有效求解中大规模问题。

为了找到大规模单二次约束非凸二次规划问题的全局最优解，本文首先将问题(1)中的凸二次约束变换成球约束，那么问题(1)可等价地转化为经典的信赖域子问题；然后将该问题中的目标函数分解成两个凸函数之差；再将凹函数的部分进行线性化，提出了求解该问题的 DC 算法，并对 DC 算法的收敛性进行分析；最后证明 DC 算法收敛到该问题的 KKT 点，并可以找到原问题的全局最优解。

1 DC 算法

本文根据信赖域子问题的全局最优性条件，针对单二次约束的非凸二次规划问题提出了一个新的 DC 算法。首先对问题(1)中的正定矩阵 $\boldsymbol{A}_1$ 进行 Cholesky 分解，即 $\boldsymbol{A}_1=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$，其中 $\boldsymbol{B}$ 为可逆矩阵；然后作变换

$\boldsymbol{Bx}=\boldsymbol{y}$；最后得到问题(1)与经典的信赖域子问题有相同的最优值，即命题 1。

命题 1　问题(1)与如下经典的信赖域子问题有相同的最优值：

$$\begin{cases}\min f(\boldsymbol{y})=\dfrac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{y}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{y}+c_0\\ \text{s.t.}\ \|\boldsymbol{y}-\boldsymbol{v}\|\leqslant r\end{cases}\tag{3}$$

其中：$\boldsymbol{C}=(\boldsymbol{B}^{-1})^{\mathrm{T}}\boldsymbol{A}_0\boldsymbol{B}^{-1}$；$\boldsymbol{d}=(\boldsymbol{B}^{-1})^{\mathrm{T}}\boldsymbol{b}_0$；$\boldsymbol{v}=-(\boldsymbol{B}^{-1})^{\mathrm{T}}\boldsymbol{b}_1$；$r^2=-2c_1+\|(\boldsymbol{B}^{-1})^{\mathrm{T}}\boldsymbol{b}_1\|^2$；$\|\cdot\|^2=\langle\cdot,\cdot\rangle$，表示欧氏空间的二范数；$\langle\cdot,\cdot\rangle$ 表示两个向量的内积。若 $\boldsymbol{y}^*$ 是问题(3) 的全局最优解，则 $\boldsymbol{x}=\boldsymbol{B}^{-1}\boldsymbol{y}^*$ 为问题(1)的全局最优解。

下面通过 DC 算法求解问题(3)得到问题(1)的最优值，先对问题(3)的目标函数 $f(\boldsymbol{y})$ 进行 DC 分解，得

$$f(\boldsymbol{y})=g(\boldsymbol{y})-h(\boldsymbol{y}),$$

其中：$g(\boldsymbol{y})=\frac{1}{2}\rho\ \|\boldsymbol{y}\|^2+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{y}+\chi_E(\boldsymbol{y})$；$h(\boldsymbol{y})=\frac{1}{2}\rho\ \|\boldsymbol{y}\|^2-\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{y}-c_0=\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}(\rho\mathbf{I}-\boldsymbol{C})\boldsymbol{y}-c_0$；$E=\{\boldsymbol{y}\in\mathbb{R}^n:\|\boldsymbol{y}-\boldsymbol{v}\|\leqslant r\}$；$\chi_E$ 表示集合 E 的指示函数；ρ 是一个使 $\rho\mathbf{I}-\boldsymbol{C}$ 为半正定的正实数。则问题(3)可等价转化为如下形式：

$$\min\{g(\boldsymbol{y})-h(\boldsymbol{y}):\boldsymbol{y}\in\mathbb{R}^n\},$$

再选取初始点 $\boldsymbol{y}^0\in\mathbb{R}^n$，迭代次数 $k\geqslant 0$，记 $\boldsymbol{z}^k=\nabla\boldsymbol{h}(\boldsymbol{y}^k)=(\rho\mathbf{I}-\boldsymbol{C})\boldsymbol{y}^k$，则

$$\boldsymbol{y}^{k+1}=\operatorname{argmin}\{g(\boldsymbol{y})-\langle\boldsymbol{y},\boldsymbol{z}^k\rangle\}=\operatorname{argmin}\left\{\frac{1}{2}\rho\ \|\boldsymbol{y}\|^2+\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{d}-\boldsymbol{z}^k)+\chi_E:\boldsymbol{y}\in\mathbb{R}^n\right\},$$

即 $\boldsymbol{y}^{k+1}$ 是下述问题的最优解：

$$\min_{y\in E}\left\|\boldsymbol{y}-\frac{\boldsymbol{z}_k-\boldsymbol{d}}{\rho}\right\|^2.$$

最后可得

$$\boldsymbol{y}^{k+1}=P_E\left(\frac{\boldsymbol{z}^k-\boldsymbol{d}}{\rho}\right),$$

其中：$P_E\left(\frac{\boldsymbol{z}^k-\boldsymbol{d}}{\rho}\right)$ 表示$\frac{\boldsymbol{z}^k-\boldsymbol{d}}{\rho}$ 在集合 E 上的投影。

上述求解问题(3)的 DC 算法可描述如下：

第一步，初始化：输入初始点 $\boldsymbol{y}_0\in\mathbb{R}^n$，迭代次数 $k=0$，终止精度 $\varepsilon=10^{-6}$。

第二步，计算 $\boldsymbol{z}^k=(\rho\mathbf{I}-\boldsymbol{C})\boldsymbol{y}^k$。

第三步，若 $\left\|\frac{\boldsymbol{z}^k-\boldsymbol{d}}{\rho}-\boldsymbol{v}\right\|\leqslant r$，令 $\boldsymbol{y}^{k+1}=\frac{\boldsymbol{z}^k-\boldsymbol{d}}{\rho}$；否则，令

$$\boldsymbol{y}^{k+1}=\frac{\dfrac{r\boldsymbol{z}_k}{\rho}+\left(\left\|\dfrac{\boldsymbol{z}_k}{\rho}-\boldsymbol{v}\right\|-r\right)\boldsymbol{v}}{\left\|\dfrac{\boldsymbol{z}_k}{\rho}-\boldsymbol{v}\right\|}。$$

第四步，若 $\|\boldsymbol{y}^{k+1}-\boldsymbol{y}^k\|\leqslant\varepsilon$，则停止，输出 $\boldsymbol{y}^k$；否则令 $k=k+1$，返回第二步。

由文献[7]中定理 4.1 的证明，可证得 DC 算法的收敛性结果，如下。

定理 1[7]　DC 算法产生的序列 $\{\boldsymbol{y}^k\}$ 的任一聚点 $\boldsymbol{y}^*$ 为问题(3)的 KKT 点，即存在 $\lambda^*\in\mathbb{R}$ 使得 $(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{y}^*=\lambda^*\boldsymbol{v}-\boldsymbol{d}$，$\lambda^*(\|\boldsymbol{y}^*-\boldsymbol{v}\|-r)=0$，$\|\boldsymbol{y}^*-\boldsymbol{v}\|\leqslant r$。

定理 1 说明了 DC 算法收敛于问题(3)的 KKT 点，但其未必是问题(3)的全局最优解。

2　基于 DC 算法的问题(3)全局最优解求解

2.1　问题(3)的全局最优性条件

由文献[1]可得问题(3)的全局最优性条件，即定理 2。

定理 2[1]　$\boldsymbol{y}^*$ 是问题(3)的全局最优解的充要条件为：

a) $(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{y}^{*}=\lambda^{*}\boldsymbol{v}-\boldsymbol{d}$;

b) $\lambda^{*}(\|\boldsymbol{y}^{*}-\boldsymbol{v}\|-r)=0,\ \|\boldsymbol{y}^{*}-\boldsymbol{v}\|\leqslant r$;

c) $\boldsymbol{C}+\lambda^{*}\mathbf{I}$ 半正定。

定理 2 的证明见文献[1],条件 a)、b)即为问题(3)的 KKT 条件。

由定理 1 可知,DC 算法产生的序列 $\{y^{k}\}$ 收敛到问题(3)的 KKT 点 y^{*},为了证明 y^{*} 是否为问题(3)的全局最优解,只需证明 y^{*} 是否满足定理 2 中的三个条件。根据定理 1 可知,y^{*} 满足定理 2 的条件 a)、b),但不一定满足定理 2 的条件 c)。

下面讨论 y^{*} 在什么条件下满足定理 2 中的条件 c)。由定理 2 的 a)、b)可得:

$$\lambda^{*}=\frac{-\langle\boldsymbol{y}^{*}-\boldsymbol{v},\boldsymbol{C}\boldsymbol{y}^{*}\rangle-\langle\boldsymbol{y}^{*}-\boldsymbol{v},\boldsymbol{d}\rangle}{r^{2}},$$

根据矩阵特征值的定义可知 $\boldsymbol{C}\boldsymbol{\eta}=\lambda_{\min}^{C}\boldsymbol{\eta}$,其中:$\lambda_{\min}^{C}$ 为矩阵 C 的最小特征值;$\boldsymbol{\eta}$ 为其对应的特征向量。若 $\lambda^{*}+\lambda_{\min}^{C}\geqslant 0$,则 y^{*} 满足定理 2 的条件 c),即 y^{*} 是问题(3)的全局最优解。否则,

$$\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{\eta}=(\lambda^{*}+\lambda_{\min}^{C})\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta}<0,$$

即定理 2 中的条件 c)不满足,这说明当前点 y^{*} 并不是问题(3) 的全局最优解。此时,需要找到新的初始点 $\bar{y}$ 重新启动 DC 算法,这里 $\bar{y}$ 需满足 $\|\bar{y}-v\|\leqslant r$ 且 $f(\bar{y})<f(y^{*})$。为了找到满足上述条件的 $\bar{y}$,先给出如下引理。

引理 1 设 y^{*} 为问题(3)的 KKT 点,λ^{*} 为相应的拉格朗日乘子,则对任意 $y\in\mathbb{R}^{n}$,

$$f(\boldsymbol{y})=f(\boldsymbol{y}^{*})-\frac{\lambda^{*}}{2}(\|\boldsymbol{y}-\boldsymbol{v}\|^{2}-\|\boldsymbol{y}^{*}-\boldsymbol{v}\|^{2})+\frac{1}{2}\langle\boldsymbol{y}-\boldsymbol{y}^{*},(\boldsymbol{C}+\lambda^{*}\mathbf{I})(\boldsymbol{y}-\boldsymbol{y}^{*})\rangle \tag{4}$$

证明 因 y^{*} 为问题(3)的 KKT 点,且 λ^{*} 为相应的拉格朗日乘子,故有 $(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{y}^{*}=\lambda^{*}\boldsymbol{v}-\boldsymbol{d}$,$\lambda^{*}(\|\boldsymbol{y}^{*}-\boldsymbol{v}\|-r)=0$,于是,

$$f(\boldsymbol{y}^{*})-\frac{\lambda^{*}}{2}(\|\boldsymbol{y}-\boldsymbol{v}\|^{2}-\|\boldsymbol{y}^{*}-\boldsymbol{v}\|^{2})+\frac{1}{2}\langle\boldsymbol{y}-\boldsymbol{y}^{*},(\boldsymbol{C}+\lambda^{*}\mathbf{I})(\boldsymbol{y}-\boldsymbol{y}^{*})\rangle=$$

$$\frac{1}{2}(\boldsymbol{y}^{*})^{\mathrm{T}}(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{y}^{*}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{y}^{*}+c_{0}-\frac{\lambda^{*}}{2}\|\boldsymbol{y}\|^{2}+\lambda^{*}(\boldsymbol{y}-\boldsymbol{y}^{*})^{\mathrm{T}}\boldsymbol{v}+\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{y}+\frac{\lambda^{*}}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}-$$

$$\boldsymbol{y}^{\mathrm{T}}(C+\lambda^{*}\mathbf{I})\boldsymbol{y}^{*}+\frac{1}{2}(\boldsymbol{y}^{*})^{\mathrm{T}}(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{y}^{*}=(\boldsymbol{y}^{*})^{\mathrm{T}}(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{y}^{*}+(\lambda^{*}\boldsymbol{v}-(\boldsymbol{C}+$$

$$\lambda^{*}\mathbf{I})\boldsymbol{y}^{*})^{\mathrm{T}}\boldsymbol{y}^{*}-\lambda^{*}(\boldsymbol{y}^{*})^{\mathrm{T}}\boldsymbol{v}+f(\boldsymbol{y})=f(\boldsymbol{y}).$$

证毕。

2.2 新初始点 $\bar{y}$ 的选取

设 y^{*} 为问题(3)的 KKT 点,λ^{*} 为相应的最优拉格朗日乘子,记 $\lambda_{\min}^{C}$ 为矩阵 C 的最小特征值。若 $\lambda^{*}+\lambda_{\min}^{C}\geqslant 0$,则 y^{*} 是问题(3)的全局最优解;否则,

$$\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{\eta}<0 \tag{5}$$

其中 η 为矩阵 C 的最小特征值所对应的特征向量。

令 $y-v=-(y^{*}-v)$,即 $y=-y^{*}+2v$,根据引理 1,式(4)可化简为:

$$f(\boldsymbol{y})=f(\boldsymbol{y}^{*})+2\langle\boldsymbol{y}^{*}-\boldsymbol{v},(\boldsymbol{C}+\lambda^{*}\mathbf{I})(\boldsymbol{y}^{*}-\boldsymbol{v})\rangle=f(\boldsymbol{y}^{*})+2\langle\boldsymbol{y}^{*}-\boldsymbol{v},(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{y}^{*}-(\boldsymbol{C}+\lambda^{*}\mathbf{I})v\rangle=f(\boldsymbol{y}^{*})+2\langle\boldsymbol{y}^{*}-\boldsymbol{v},\lambda^{*}\boldsymbol{v}-\boldsymbol{d}-(\boldsymbol{C}+\lambda^{*}\mathbf{I})\boldsymbol{v}\rangle=f(\boldsymbol{y}^{*})-2\langle\boldsymbol{y}^{*}-\boldsymbol{v},\boldsymbol{C}\boldsymbol{v}+\boldsymbol{d}\rangle \tag{6}$$

由式(6)可知,新初始点 $\bar{y}$ 的选取,可考虑如下几种情况:

a)当 $(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^{*}-\boldsymbol{v})>0$ 时,则令 $\bar{y}=-y^{*}+2v$,由式(6) 得 $f(\bar{y})<f(y^{*})$,并且 $\|\bar{y}-v\|=\|-y^{*}+2v-v\|=\|y^{*}-v\|\leqslant r$,即 $\bar{y}$ 为问题(3) 的可行下降点。故可选取 $\bar{y}$ 为新的初始点重新启动 DC 算法。

b)当 $(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^{*}-\boldsymbol{v})\leqslant 0$ 时,则分为以下两种情况:

b1) $\|y^{*}-v\|<r$ 或 $\|y^{*}-v\|=r$ 且 $\boldsymbol{\eta}^{\mathrm{T}}(y^{*}-v)\neq 0$ 的情况。

令 $\bar{y}=y^{*}+\gamma\boldsymbol{\eta}(\gamma\neq 0)$,使 $\|\bar{y}-v\|=r$ 成立,即 γ 是下述方程的解:

$$\|\boldsymbol{\eta}\|^{2}\gamma^{2}+2\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^{*}-\boldsymbol{v})\gamma+\|\boldsymbol{y}^{*}-\boldsymbol{v}\|^{2}-r^{2}=0,$$

从而，

$$\gamma=\begin{cases}\dfrac{-2\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})}{\|\boldsymbol{\eta}\|^2}, & 若\ \|\boldsymbol{y}^*-\boldsymbol{v}\|=r,\\ \dfrac{-\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\pm\sqrt{\Delta}}{\|\boldsymbol{\eta}\|^2}, & 若\ \|\boldsymbol{y}^*-\boldsymbol{v}\|<r,\end{cases}$$

其中 $\Delta=[\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})]^2-\|\boldsymbol{\eta}\|^2(\|\boldsymbol{y}^*-\boldsymbol{v}\|^2-r^2)>0$。则由式(6)可得：

$$f(\overline{\boldsymbol{y}})=f(\boldsymbol{y}^*)-\frac{\lambda^*}{2}(\|\overline{\boldsymbol{y}}-\boldsymbol{v}\|^2-\|\boldsymbol{y}^*-\boldsymbol{v}\|^2)+\frac{1}{2}\langle\overline{\boldsymbol{y}}-\boldsymbol{y}^*,(\boldsymbol{C}+\lambda^*\mathbf{I})(\overline{\boldsymbol{y}}-\boldsymbol{y}^*)\rangle=f(\boldsymbol{y}^*)-$$

$$\frac{\lambda^*}{2}(\|\overline{\boldsymbol{y}}-\boldsymbol{v}\|^2-\|\boldsymbol{y}^*-\boldsymbol{v}\|^2)+\frac{\gamma^2}{2}\langle\boldsymbol{\eta},(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}\rangle\leqslant f(\boldsymbol{y}^*)+\frac{\gamma^2}{2}\langle\boldsymbol{\eta},(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}\rangle<f(\boldsymbol{y}^*) \tag{7}$$

其中最后一个不等式由式(5)推得，故可选取 $\overline{\boldsymbol{y}}$ 为新的初始点重新启动 DC 算法。因为 $\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}<0$，所以由式(7) 可知 γ^2 越大，$f(\overline{\boldsymbol{y}})$ 的值越小，即当 $\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})<0$ 时，取 $\gamma=\dfrac{-\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})+\sqrt{\Delta}}{\|\boldsymbol{\eta}\|^2}$，当 $\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})>0$ 时，取 $\gamma=\dfrac{-\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})-\sqrt{\Delta}}{\|\boldsymbol{\eta}\|^2}$。

b2) $\|\boldsymbol{y}^*-\boldsymbol{v}\|=r$ 且 $\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})=0$ 的情况。

令 $\xi=\boldsymbol{\eta}+\tau(\boldsymbol{y}^*-\boldsymbol{v})$，其中 $\tau<0$，使得 $\xi^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\xi<0$ 且 $\xi^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\neq 0$，用 ξ 代替 b1) 中的 $\boldsymbol{\eta}$ 并返回 b1)。下面讨论如何寻找一个 $\tau<0$ 使 $\xi^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\xi<0$ 且 $\xi^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\neq 0$ 成立。

因 $\tau<0$，故 $\xi^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})=\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})+\tau\ \|\boldsymbol{y}^*-\boldsymbol{v}\|^2=\tau\ \|\boldsymbol{y}^*-\boldsymbol{v}\|^2<0$，又因为

$$\begin{aligned}\xi^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\xi&=\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}+2\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\tau(\boldsymbol{y}^*-\boldsymbol{v})+\\&\quad\tau^2(\boldsymbol{y}^*-\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})(\boldsymbol{y}^*-\boldsymbol{v})\\&=\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}-2\tau(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}-\tau^2(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\end{aligned} \tag{8}$$

故由式(5)、式(8)可知，要使 $\xi^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\xi<0$，可分为以下几种情况：

b21)若 $(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})=0$ 且$(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}\leqslant 0$，则当 $\tau<0$ 时 $\xi^{\mathrm{T}}(\boldsymbol{C}+\lambda*\mathbf{I})\xi<0$；

b22)若 $(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})=0$ 且$(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}>0$，则令

$$\frac{\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}}{2(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}}<\tau<0,$$

由式(8)得 $\xi^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\xi<0$；

b23)若 $(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})<0$，记

$$g(\tau)=\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}-2\tau(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}-\tau^2(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v}),$$

由式(5)知 $\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}<0$，又因为$(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})<0$，故

$$\Delta=4[(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}]^2+4(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}>0,$$

则 $g(\tau)=0$ 的根为：

$$\tau_1=\frac{(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}+\sqrt{[(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}]^2+(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}}}{(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})},$$

$$\tau_2=\frac{(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}-\sqrt{[(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}]^2+(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}}}{(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})},$$

因为 $|(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}|<\sqrt{[(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}]^2+(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\boldsymbol{\eta}}$，所以

$$(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}+\sqrt{[(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}]^2+(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\boldsymbol{\lambda}^*\mathbf{I})\boldsymbol{\eta}}>0,$$

$$(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}-\sqrt{[(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\eta}]^2+(\boldsymbol{d}+\boldsymbol{C}\boldsymbol{v})^{\mathrm{T}}(\boldsymbol{y}^*-\boldsymbol{v})\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}+\boldsymbol{\lambda}^*\mathbf{I})\boldsymbol{\eta}}<0,$$

则 $\tau_1<0<\tau_2$，即对任意 $\tau\in(\tau_1,0)$，可得 $\xi^{\mathrm{T}}(\boldsymbol{C}+\lambda^*\mathbf{I})\xi<0$。

基于以上讨论，若由 DC 算法得到的 KKT 点 $\boldsymbol{y}^*$ 不是问题(3)的全局最优解，则可以根据 2.2 节的分析找到新的初始点 $\overline{\boldsymbol{y}}$($\overline{\boldsymbol{y}}$ 满足 $\|\overline{\boldsymbol{y}}-\boldsymbol{v}\|\leqslant r$ 且 $f(\overline{\boldsymbol{y}})<f(\boldsymbol{y}^*)$)重新启动 DC 算法，得到一个更好的解，重复这

个过程,可以找到原问题的全局解。

2.3 基于 DC 算法的全局最优算法

下面给出求解问题(3)的全局最优 DC(Global difference of convex function,GDC)算法。

第一步,初始化:计算矩阵 $\boldsymbol{C}$ 的最小特征值 $\lambda_{\min}^{\boldsymbol{C}}$ 及其对应的特征向量 $\boldsymbol{\eta}$,并计算矩阵 $\boldsymbol{C}$ 的最大特征值 $\lambda_{\max}^{\boldsymbol{C}}$,输入参数 ρ。

第二步,运行 DC 算法得到 $\boldsymbol{y}^*$。

第三步,计算 $\lambda^* = \dfrac{-\langle \boldsymbol{y}^* - \boldsymbol{v}, \boldsymbol{C}\boldsymbol{y}^* \rangle - \langle \boldsymbol{y}^* - \boldsymbol{v}, \boldsymbol{d} \rangle}{r^2}$。

第四步,如果 $\lambda^* \geqslant -\lambda_{\min}^{\boldsymbol{C}}$,则 $\boldsymbol{y}^*$ 为问题(3)的全局解,停止,输出 $\boldsymbol{y}^*$;否则,先按照 2.2 节计算 $\overline{y}$ 使得 $f(\overline{y}) < f(\boldsymbol{y}^*)$,再将 $\overline{y}$ 作为 DC 算法新的初始点,返回第二步。

3 数值实验

本文通过数值实验验证 GDC 算法的有效性。其中算法 GDC 用 Matlab(R2017a)编程,电脑的操作系统为 Windows 10,处理器为 Intel(R) Core (TM)i5-2400F CPU(3.10 GHz),运行内存为 4 GB。

本文在数值实验中将 GDC 算法同半定规划问题 SDP 及文献[15]中的精确算法 3.1(Exact algorithm,EA)进行了比较,SDP 用求解器 CVX 求解,DC 算法的初始点选为 $y_i^0 = \dfrac{r}{\sqrt{n}}, i = 1,2,\cdots,n$,参数 $\rho = \lambda_{\max}^{\boldsymbol{C}} + 0.1$。

数值实验中,对随机生成的 10 个问题进行测试,测试问题中的参数 $\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{b}_0, \boldsymbol{b}_1, c_0, c_1$ 随机产生,产生方式类似于文献[15],矩阵 $\boldsymbol{A}_0$ 和 $\boldsymbol{A}_1$ 的形式为:$\boldsymbol{A}_i = \boldsymbol{P}_i \boldsymbol{D}_i \boldsymbol{P}_i^{\mathrm{T}}, i = 0,1$,其中 $\boldsymbol{P}_i$ 为正交矩阵,$\boldsymbol{D}_i$ 为对角矩阵。正交矩阵 $\boldsymbol{P}_i = \boldsymbol{Q}_1 \boldsymbol{Q}_2 \boldsymbol{Q}_3, i = 0,1$,其中 $\boldsymbol{Q}_j = \mathbf{I} - 2\dfrac{\boldsymbol{w}_j \boldsymbol{w}_j^{\mathrm{T}}}{\| \boldsymbol{w}_j \|^2}, j = 1,2,3$,$\boldsymbol{w}_j$ 中的元素为 $(-1,1)$ 中的随机数,对角矩阵 $\boldsymbol{D}_0 = \mathrm{diag}(d_{01}, d_{02}, \cdots, d_{0n})$,$d_{0i} \in (-5,5)$,$\boldsymbol{D}_1 = \mathrm{diag}(d_{11}, d_{12}, \cdots, d_{1n})$,$d_{1i} \in (0,20)$,$\boldsymbol{b}_i = (b_{i1}, b_{i2}, \cdots, b_{in})^{\mathrm{T}}$,$b_{ij} \in (-5,5)$,$i = 0,1$,$c_0 = 0$,$c_1 \in (-10,0)$。

表 1 给出了 GDC 算法与 SDP、EA 的平均数值结果,其中:n 表示问题的维数;v_{opt} 表示求解 10 个测试问题所得到的平均最优值;t_{cpu} 表示求解 10 个测试问题所需的平均 CPU 时间,s;n_{iter} 表示求解 10 个测试问题所需的平均迭代次数。

表 1 算法 GDC 与 SDP、EA 的数值结果

n	GDC			SDP		EA	
	v_{opt}	t_{cpu}/s	n_{iter}	v_{opt}	t_{cpu}/s	v_{opt}	t_{cpu}/s
50	−1637.685	0.005	177	−1637.685	1.916	−1637.685	1.812
100	−3800.451	0.016	102	−3800.451	6.006	−3800.451	3.858
200	−14342.401	0.033	146	−14342.401	15.186	−14342.401	17.11
300	−287796.483	0.045	136	−287796.483	43.922	−287796.483	45.239
400	−37260.246	0.428	118	−37260.246	92.163	−37260.246	111.735
500	−110940.674	0.427	68	−110940.674	187.613	−110940.674	203.664
1000	−14523097.379	2.597	88	—	—	—	—
2000	−4015766.938	10.379	86	—	—	—	—
3000	−5084523.215	45.744	180	—	—	—	—
4000	−18875663.441	77.775	89	—	—	—	—
5000	−21857572.651	153.288	225	—	—	—	—

注:“—”表示 SDP 与 EA 算法在 1 h 内没找到最优值或最优解。

从表 1 的数值实验结果可以看出,当 $n \leqslant 500$ 时,GDC 算法与求解 SDP 得到的最优值相同,算法 GDC 在 1 s 内找到所有测试问题的全局最优解,而 SDP 与 EA 所需的 CPU 时间比 GDC 算法更长。当 $n = 500$ 时,SDP 与 EA 所用的 CPU 时间均大于 150 s;当 $1000 \leqslant n \leqslant 5000$ 时,GDC 算法在 160 s 内找到所有测试问题的全局最优解,而 SDP 与 EA 均未能找到所有测试问题的全局最优解,尤其当 $n = 1000$ 时,算法 GDC 在

3 s内求得测试问题的全局最优解。因此，算法 GDC 相比于 SDP 与 EA 能有效找到大规模问题的全局最优解。

4 结　论

本文研究了单二次约束的非凸二次规划问题，设计了求解该问题全局最优解的 GDC 算法，对随机生成的问题进行了数值实验，结果表明该算法比 SDP 与 EA 更能有效快速地找到大规模问题的全局最优解。本文的算法仅适用于凸二次约束的非凸二次规划问题，后续将研究非凸二次约束的非凸二次规划问题的全局算法。

参考文献：

[1] Moré J J, Sorensen D C. Computing a trust region step[J]. SIAM Journal on Scientific and Statistical Computing, 1983, 4(3): 553-572.

[2] Adachi S, Iwata S, Nakatsukasa Y, et al. Solving the trust-region subproblem by a generalized eigenvalue problem[J]. SIAM Journal on Optimization, 2017, 27(1): 269-291.

[3] Hazan E, Koren T. A linear-time algorithm for trust region problems[J]. Mathematical Programming, 2016, 158(1/2): 363-381.

[4] Martínez J M. Local minimizers of quadratic functions on euclidean balls and spheres[J]. SIAM Journal on Optimization, 1994, 4(1): 159-176.

[5] Wang J L, Xia Y. Closing the gap between necessary and sufficient conditions for local nonglobal minimizer of trust region subproblem[J]. SIAM Journal on Optimization, 2020, 30(3): 1980-1995.

[6] Lucidi S, Palagi L, Roma M. On some properties of quadratic programs with a convex quadratic constraint[J]. SIAM Journal on Optimization, 1998, 8(1): 105-122.

[7] Tao P D, An L T H. A D.C. optimization algorithm for solving the trust-region subproblem[J]. SIAM Journal on Optimization, 1998, 8(2): 476-505.

[8] An L T H, Tao P D. DC programming and DCA: thirty years of developments[J]. Mathematical Programming, 2018, 169(1): 5-68.

[9] Pardalos P M, Vavasis S A. Quadratic programming with one negative eigenvalue is NP-hard[J]. Journal of Global Optimization, 1991, 1(1): 15-22.

[10] Pardalos P M, Schnitger G. Checking local optimality in constrained quadratic programming is NP-hard[J]. Operations Research Letters, 1988, 7(1): 33-35.

[11] Jeyakumar V, Rubinov A M, Wu Z Y. Non-convex quadratic minimization problems with quadratic constraints: global optimality conditions[J]. Mathematical Programming, 2007, 110(3): 521-541.

[12] Kim S, Kojima M. Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations[J]. Computational Optimization & Applications, 2003, 26(2): 143-154.

[13] Zhang S Z. Quadratic maximization and semidefinite relaxation[J]. Mathematical Programming, 2000, 87(3): 453-465.

[14] Pólik I, Terlaky T. A survey of the S-Lemma[J]. SIAM Review, 2007, 49(3): 371-418.

[15] 郑小金. 基于最优 D.C.分解的单二次约束非凸二次规划精确算法[J]. 运筹学学报, 2009, 13(3): 111-118.

（责任编辑：康　锋）