浙江理工大学学报，2022，47(4)：601-607
Journal of Zhejiang Sci-Tech University
DOI:10.3969/j.issn.1673-3851(n).2022.04.018

带凸二次约束非凸二次规划的双非负规划松弛及其解法

章显业，罗和治
（浙江理工大学理学院，杭州 310018）

摘　要：针对带有非负变量、线性等式和凸二次约束的非凸二次规划问题，给出了一个带有矩阵非负和半正定约束的紧双非负规划(Doubly nonnegative programming, DNP)松弛，估计了它与原问题之间的间隙，并提出了求DNP松弛最优解的交替方向乘子法。数值实验表明：交替方向乘子法能有效找到DNP松弛问题的最优解，并且计算时间优于求解器CVX。

关键词：非凸二次规划；双非负规划松弛；交替方乘子向法；半定规划；CVX
中图分类号：O221　　**文献标志码**：A　　**文章编号**：1673-3851 (2022) 07-0601-07

Doubly non-negative programming relaxation for non-convex quadratic programming with convex quadratic constraint and its solution

ZHANG Xianye, LUO Hezhi
(School of Science, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: A tight doubly non-negative programming (DNP) relaxation with non-negative and positive semi-definite matrix constraints is developed for non-convex quadratic programming problems with non-negative variables, linear equations, and convex quadratic constraints. The gap between it and the original problem is estimated, and an alternating direction multiplier method is proposed to find an optimal solution of DNP relaxation. The numerical experiments show that the optimal solution of DNP relaxation can be found effectively by applying the alternating direction multiplier method , and the calculation time is better than the solver CVX.

Key words: non-convex quadratic programming; doubly non-negative programming relaxation; alternating direction multiplier method; semi-definite programming; CVX

0 引　言

本文考虑如下带有非负变量、线性等式和凸二次约束的非凸二次规划问题(Quadratically constrained quadratic program, QCQP)：

$$\begin{cases}\min f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.}\quad \boldsymbol{x}\in\mathcal{F};\end{cases}\tag{1}$$

其中：

$$\mathcal{F}:=\left\{\boldsymbol{x}\in\mathbb{R}^n:\left|\begin{array}{l}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_i\boldsymbol{x}+\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{x}\leqslant \boldsymbol{d}_i, i=1,\cdots,m\\ \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geqslant\boldsymbol{0}\end{array}\right.\right\};$$

收稿日期：2022—02—16　　网络出版日期：2022—04—06
基金项目：浙江省自然科学基金重点项目(LZ21A010003)；国家自然科学基金项目(11871433)
作者简介：章显业(1997—　)，男，浙江温州人，硕士研究生，主要从事最优化理论与算法方面的研究。
通信作者：罗和治，E-mail：hzluo@zstu.edu.cn

$\boldsymbol{Q}$ 是具有 r 个负特征值的 n 阶对称不定矩阵或者对称负定矩阵;$\boldsymbol{Q}_i\,(i=1,\cdots,m)$ 是 n 阶半正定矩阵;$\boldsymbol{q}\in\mathbb{R}^n$;$\boldsymbol{A}\in\mathbb{R}^{m\times n}$;$\boldsymbol{b}\in\mathbb{R}^m$。非凸 QCQP 问题在工程和金融领域有广泛应用,如 0-1 二次规划问题、最大割问题、二次背包问题、0-1 最小二乘问题、图形处理问题、多用户检测问题、项目选择和资源分配问题、多传感器波束形成问题以及系统平衡问题等都可以归结为非凸 QCQP 问题。Pardalos 等[1]证明了求解具有一个负特征值的非凸二次规划问题的全局最优解是 NP-难的,甚至寻找其局部最优解也是 NP-难的[2]。故问题(1)是 NP-难的。

寻找非凸 QCQP 问题全局解的主要方法是基于凸松弛或半定规划(Semi-definite programming, SDP)松弛的分支定界算法[3]。Vandenbussche 等[4]对箱子约束非凸 QP 问题提出了基于一阶 Karush-Kuhn-Tucker(KKT)条件的分支割算法。Le Thi 等[5]对箱子约束非凸 QP 问题提出了基于 DC(Difference of convex function,DC) 算法和椭球技术的分支定界算法,并推广到凸二次约束非凸 QP 问题[6]。Barrientos 等[7]对线性约束非凸 QP 问题提出了基于正交化和 Lagrangian 方法的分支定界算法。Burer 等[8]对线性约束非凸 QP 问题提出了基于 SDP 松弛和 KKT 分支技术的分支定界算法。Luo 等[9]对凸二次约束非凸 QP 问题在分支定界框架下提出了基于交替方向法和凸松弛的新全局算法。

众所周知,分支定界算法的有效性主要取决于两个因素,即松弛界的质量及其相关的计算成本[3]。由于 SDP 松弛能提供更紧的下界,学者们对非凸 QCQP 问题提出了大量 SDP 松弛问题。Anstreicher[10]利用重构线性技术(Reformulation-linearization technique,RLT)技术改进了非凸 QCQP 问题的 SDP 松弛。Zheng 等[11]对带凸二次约束非凸 QP 问题提出了基于最佳 DC 分解的 SDP 松弛方法。Zheng 等[12]对一般非凸 QCQP 问题提出了基于矩阵锥分解和多胞形逼近技术的 SDP 松弛方法。Saxena 等[13]讨论了如何利用 disjunctive 割来改进非凸 QCQP 问题的 SDP 松弛。Luo 等[14]利用罚函数方法改进非凸 QCQP 问题的 SDP 松弛,得到了具有多项式时间可解的非线性 SDP 松弛。Chen 等[15-16]对带有非负变量和线性等式约束的非凸 QP 问题,提出了带有矩阵非负和半正定约束的双非负规划(Doubly nonnegative programming, DNP)松弛,并给出了求解 DNP 松弛的增广拉格朗日分解方法和求解子问题的块坐标下降法。Cerulli 等[17]和 Oliveira 等[18]对 DNP 问题分别提出了两个不同的新交替方向乘子法,但这些方法不能用于求解带有关于$(\boldsymbol{x},\boldsymbol{X})$的线性约束的 DNP 问题。

本文对带有非负变量、线性等式和凸二次约束的非凸 QP 问题给出了一个带有 Secant 割的紧 DNP 松弛,估计它与原问题之间的间隙,提出求解该 DNP 松弛的交替方向乘子法,并证明该算法收敛于 DNP 松弛问题的最优解,同时还提出了一个割平面方法来求解交替方向乘子法中的子问题。本文引入如下记号:定义 $\mathrm{Tr}(\bullet)$为矩阵的迹,$\boldsymbol{S}^n$ 为 n 阶对称矩阵的集合,$\boldsymbol{B}\bullet\boldsymbol{C}=\mathrm{Tr}(\boldsymbol{BC})$为 $\boldsymbol{B},\boldsymbol{C}\in\boldsymbol{S}^n$ 的内积;对于 $\boldsymbol{X}\in\boldsymbol{S}^n$,$\boldsymbol{X}\succeq 0$ 表示矩阵 $\boldsymbol{X}$ 为半正定矩阵,$\boldsymbol{X}\geqslant 0$ 表示矩阵 $\boldsymbol{X}$ 的元素 $\boldsymbol{X}_{ij}\geqslant 0$,$\mathrm{diag}(\boldsymbol{X})$为以矩阵 $\boldsymbol{X}$ 的对角元为分量的向量。

1 双非负规划松弛及其间隙估计

本节给出问题(1)的一个带有 Secant 割的紧 DNP 松弛,并估计它与问题(1)之间的间隙。

设 $\boldsymbol{Q}$ 的负特征值个数为 r,不失一般性,假设 $\boldsymbol{Q}$ 的前 r 个特征值为负。对 $\boldsymbol{Q}$ 特征值分解,得 $\boldsymbol{Q}=\boldsymbol{Q}^+-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$,其中:$\boldsymbol{Q}^+=\sum_{i=r+1}^{n}\boldsymbol{\lambda}_i\boldsymbol{\xi}_i\boldsymbol{\xi}_i^{\mathrm{T}}\succeq\boldsymbol{0}$;$\boldsymbol{C}=(\sqrt{-\boldsymbol{\lambda}_1}\,\boldsymbol{\xi}_1,\cdots,\sqrt{-\boldsymbol{\lambda}_r}\,\boldsymbol{\xi}_r)^{\mathrm{T}}$;且 $\boldsymbol{\lambda}_i<0$,$i=1,\cdots,r$,$\boldsymbol{\lambda}_i\geqslant 0$,$i=r+1,\cdots,n$,为 $\boldsymbol{Q}$ 的特征值;$\boldsymbol{\xi}_i$,$i=1,\cdots,n$ 是对应的正交单位特征向量。设 $\boldsymbol{l},\boldsymbol{u}\in\mathbb{R}^n$ 为 $\boldsymbol{Cx}$ 在$\mathcal{F}$上的上下界,且:

$$\begin{cases}\boldsymbol{l}_i=\min\limits_{x\in\mathcal{F}}\boldsymbol{c}_i^{\mathrm{T}}\boldsymbol{x}, & i=1,\cdots,r;\\ \boldsymbol{u}_i=\max\limits_{x\in\mathcal{F}}\boldsymbol{c}_i^{\mathrm{T}}\boldsymbol{x}, & i=1,\cdots,r\end{cases}\tag{2}$$

其中:$\boldsymbol{c}_i=\sqrt{-\boldsymbol{\lambda}_i}\,\boldsymbol{\xi}_i$,$i=1,\cdots,r$。设 $\boldsymbol{X}=\boldsymbol{xx}^{\mathrm{T}}$,易得 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx}=\boldsymbol{Q}\bullet\boldsymbol{X}$。由于对 $\boldsymbol{x}\in\mathcal{F}$,有 $\boldsymbol{l}_i\leqslant\boldsymbol{c}_i^{\mathrm{T}}\boldsymbol{x}\leqslant\boldsymbol{u}_i$,故得:

$$\boldsymbol{c}_i\boldsymbol{c}_i^{\mathrm{T}}\bullet\boldsymbol{X}=(\boldsymbol{c}_i^{\mathrm{T}}\boldsymbol{x})^2\leqslant(\boldsymbol{l}_i+\boldsymbol{u}_i)^{\mathrm{T}}\boldsymbol{c}_i^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{l}_i\boldsymbol{u}_i,\ i=1,\cdots,r\tag{3}$$

式(3)即为 Saxena 等[13]所提出的 Secant 割。注意到 $\boldsymbol{Ax}=\boldsymbol{b}$ 蕴含 $\mathrm{diag}(\boldsymbol{AXA}^{\mathrm{T}})=\boldsymbol{b}^2$,其中 $\boldsymbol{b}^2=(\boldsymbol{b}_1^2,\cdots,\boldsymbol{b}_m^2)^{\mathrm{T}}$。设$\bar{\boldsymbol{u}}\in\mathbb{R}^n$ 为 $\boldsymbol{x}$ 在$\mathcal{F}$上的上界。由 $\boldsymbol{0}\leqslant\boldsymbol{x}\leqslant\bar{\boldsymbol{u}}$ 得 $\boldsymbol{0}\leqslant\boldsymbol{X}\leqslant\boldsymbol{U}$,其中 $\boldsymbol{U}=\bar{\boldsymbol{u}}\bar{\boldsymbol{u}}^{\mathrm{T}}$。将 $\boldsymbol{X}=\boldsymbol{xx}^{\mathrm{T}}$ 松弛为半正定锥约束 $\boldsymbol{X}-\boldsymbol{xx}^{\mathrm{T}}\succeq 0$,可得问题(1)的 SDP 松弛:

$$\begin{cases}\min \boldsymbol{Q}\cdot\boldsymbol{X}+\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t. } \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \operatorname{diag}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}^{\mathrm{T}})=\boldsymbol{b}^2, \\ \quad \boldsymbol{Q}_i\cdot\boldsymbol{X}+\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{x}\leqslant \boldsymbol{d}_i, i=1,\cdots,m, \\ \quad \boldsymbol{c}_i\boldsymbol{c}_i^{\mathrm{T}}\cdot\boldsymbol{X}-(\boldsymbol{l}_i+\boldsymbol{u}_i)\boldsymbol{c}_i^{\mathrm{T}}\boldsymbol{x}+\boldsymbol{l}_i\boldsymbol{u}_i\leqslant\boldsymbol{0}, i=1,\cdots,r, \\ \quad \boldsymbol{0}\leqslant\boldsymbol{x}\leqslant\bar{\boldsymbol{u}}, \boldsymbol{0}\leqslant\boldsymbol{X}\leqslant\boldsymbol{U}, \\ \quad \boldsymbol{X}-\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\succeq\boldsymbol{0}\end{cases} \tag{4}$$

因为它含有双非负矩阵(即它既是非负矩阵又是正半定矩阵),所以问题(4)称为 DNP 问题[15],问题(4)比 Chen 等[15]提出的 DNP 松弛提供了更紧的下界。

定理 1 给出了问题(1)与问题(4)之间的间隙估计。

定理 1　设 $f^*_{[l,u]}$ 和 $v^*_{[l,u]}$ 分别为问题(1)和问题(4)的最优值,且($\bar{\boldsymbol{x}},\bar{\boldsymbol{X}}$)为问题(4)的最优解,则

$$0\leqslant f^*_{[l,u]}-v^*_{[l,u]}\leqslant f(\bar{\boldsymbol{x}})-v^*_{[l,u]}\leqslant \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\cdot\bar{\boldsymbol{X}}-\|\boldsymbol{C}\bar{\boldsymbol{x}}\|_2^2\leqslant\frac{1}{4}\|\boldsymbol{u}-\boldsymbol{l}\|_2^2。$$

证明　首先观察到$\bar{\boldsymbol{x}}$是问题(1)的一个可行解。利用 $\boldsymbol{Q}=\boldsymbol{Q}^+-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$,可以由 $f^*_{[l,u]}$ 推导出:

$$\begin{aligned}0&\leqslant f^*_{[l,u]}-v^*_{[l,u]}\leqslant f(\bar{\boldsymbol{x}})-v^*_{[l,u]}, \\ &=\boldsymbol{Q}^+\cdot(\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^{\mathrm{T}}-\bar{\boldsymbol{X}})+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\cdot\bar{\boldsymbol{X}}-\|\boldsymbol{C}\bar{\boldsymbol{x}}\|_2^2, \\ &\leqslant\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\cdot\bar{\boldsymbol{X}}-\|\boldsymbol{C}\bar{\boldsymbol{x}}\|_2^2, \\ &\leqslant-\|\boldsymbol{C}\bar{\boldsymbol{x}}\|_2^2+(\boldsymbol{l}+\boldsymbol{u})^{\mathrm{T}}\boldsymbol{C}\bar{\boldsymbol{x}}-\boldsymbol{l}^{\mathrm{T}}\boldsymbol{u}, \\ &\leqslant\frac{1}{4}\|\boldsymbol{u}-\boldsymbol{l}\|_2^2。\end{aligned}$$

其中第三个不等式由 $\boldsymbol{Q}^+\succeq\boldsymbol{0}$ 和 $\boldsymbol{X}-\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\succeq\boldsymbol{0}$ 推得,而第四个不等式由问题(4)中的 Secant 割不等式推得。

证毕。

2　求解 DNP 松弛的交替方向乘子法

Chen 等[15]针对 DNP 松弛问题提出了增广拉格朗日分解方法,其中用块坐标下降法来求解子问题。然而,由于问题(4)含有关于($\boldsymbol{x},\boldsymbol{X}$)如 $\boldsymbol{Q}_i\cdot\boldsymbol{X}+\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{x}\leqslant\boldsymbol{d}_i$ 的线性约束,Chen 等[15]的分解方法不能直接求解问题(4)。为了求解松弛问题(4),本节提出了交替方向乘子法。

首先给出 Burer 等[16]提到的性质。定义矩阵

$$\begin{cases}\boldsymbol{M}:=(\boldsymbol{b},-\boldsymbol{A})\in\mathbb{R}^{m\times(n+1)}, \\ \boldsymbol{N}:=\operatorname{Null}(\boldsymbol{M})\text{ 的标准正交基矩阵}\end{cases} \tag{5}$$

可得 $\boldsymbol{N}^{\mathrm{T}}\boldsymbol{N}=\mathbf{I}$。对式(5)中给出的 $\boldsymbol{M}$,定义如下闭凸锥:

$$\mathcal{J}:=\{\boldsymbol{Z}\in\boldsymbol{S}^{n+1}:\boldsymbol{M}\boldsymbol{Z}\boldsymbol{M}^{\mathrm{T}}=\boldsymbol{0},\boldsymbol{Z}\succeq\boldsymbol{0}\} \tag{6}$$

引理 1 和引理 2 见文献[16]。

引理 1　设 $\boldsymbol{Y}:=(1,\boldsymbol{x}^{\mathrm{T}};\boldsymbol{x},\boldsymbol{X})\succeq\boldsymbol{0}$,$\boldsymbol{M}$ 由式(5)给出,则下述两个条件是等价的:

a)$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$,$\operatorname{diag}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}^{\mathrm{T}})=\boldsymbol{b}^2$;

b)$\boldsymbol{M}\boldsymbol{Y}\boldsymbol{M}^{\mathrm{T}}=\boldsymbol{0}$。

引理 2　对任意 $\boldsymbol{A}\in\boldsymbol{S}^{n+1}$,有 $\operatorname{proj}_{\mathcal{J}}\boldsymbol{A}=\boldsymbol{N}[\boldsymbol{N}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{N}]_+\boldsymbol{N}^{\mathrm{T}}$,其中 $\boldsymbol{N}$ 由式(5)给出,$[\boldsymbol{N}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{N}]_+$ 定义为矩阵 $\boldsymbol{N}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{N}$ 在 $\boldsymbol{S}^k_+$ 上的正交投影,k 表示 $\boldsymbol{N}$ 的列数。

为了方便,本文引入了如下的矩阵:

$$\begin{cases}\bar{\boldsymbol{Q}}:=\begin{pmatrix}0 & \frac{1}{2}\boldsymbol{q}^{\mathrm{T}} \\ \frac{1}{2}\boldsymbol{q} & \boldsymbol{Q}\end{pmatrix}, \bar{\boldsymbol{Q}}_i:=\begin{pmatrix}-d_i & \frac{1}{2}\boldsymbol{q}_i^{\mathrm{T}} \\ \frac{1}{2}\boldsymbol{q}_i & \boldsymbol{Q}_i\end{pmatrix}, i=1,\cdots,m, \bar{\boldsymbol{U}}=\begin{pmatrix}1 & \bar{\boldsymbol{u}}^{\mathrm{T}} \\ \bar{\boldsymbol{u}} & \boldsymbol{U}\end{pmatrix}, \\ \hat{\boldsymbol{A}}_{[l_i,u_i]}:=\begin{pmatrix}\boldsymbol{l}_i\boldsymbol{u}_i & -\frac{1}{2}(\boldsymbol{l}_i+\boldsymbol{u}_i)\boldsymbol{c}_i^{\mathrm{T}} \\ -\frac{1}{2}(\boldsymbol{l}_i+\boldsymbol{u}_i)\boldsymbol{c}_i & \boldsymbol{c}_i\boldsymbol{c}_i^{\mathrm{T}}\end{pmatrix}, i=1,\cdots,r, \boldsymbol{Y}:=\begin{pmatrix}1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X}\end{pmatrix}。\end{cases}$$

基于引理 1,可将问题(4)改写为如下等价的问题:

$$\begin{cases}\min & \bar{\boldsymbol{Q}}\cdot\boldsymbol{Y};\\ \text{s.t.} & \boldsymbol{Y}\in\mathcal{Y}_{[l,u]},\boldsymbol{Y}\in\mathcal{J}\end{cases}\tag{7}$$

其中:

$$\mathcal{Y}_{[\boldsymbol{l},\boldsymbol{u}]}=\left\{\boldsymbol{Y}\in\boldsymbol{S}^{n+1}\left|\begin{array}{ll}\bar{\boldsymbol{Q}}_i\cdot\boldsymbol{Y}\leqslant\boldsymbol{0}, & i=1,\cdots,m\\ \hat{\boldsymbol{A}}_{[\boldsymbol{l}_i,\boldsymbol{u}_i]}\cdot\boldsymbol{Y}\leqslant\boldsymbol{0}, & i=1,\cdots,r\\ \boldsymbol{0}\leqslant\boldsymbol{Y}\leqslant\bar{\boldsymbol{U}}, & \boldsymbol{Y}_{00}=1\end{array}\right.\right\}\tag{8}$$

引入辅助变量 $\boldsymbol{Z}$ 和等式 $\boldsymbol{Z}=\boldsymbol{Y}$,可将问题(7)重写为:

$$\begin{cases}\min\bar{\boldsymbol{Q}}\cdot\boldsymbol{Y};\\ \text{s.t.}\quad \boldsymbol{Y}\in\boldsymbol{Y}_{[\boldsymbol{l},\boldsymbol{u}]},\boldsymbol{Y}=\boldsymbol{Z},\boldsymbol{Z}\in\mathcal{J}\end{cases}\tag{9}$$

定义 $\|\boldsymbol{A}\|_F=\sqrt{\mathrm{Tr}(\boldsymbol{A}^2)}$ 为 $\boldsymbol{A}\in\boldsymbol{S}^m$ 的 Frobenius 范数。对应于约束 $\boldsymbol{Y}-\boldsymbol{Z}=\boldsymbol{0}$ 的问题(9)的增广拉格朗日函数定义为:

$$\mathcal{L}_\sigma(\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{\Lambda})=\bar{\boldsymbol{Q}}\cdot\boldsymbol{X}+\boldsymbol{\Lambda}\cdot(\boldsymbol{Y}-\boldsymbol{Z})+\frac{\sigma}{2}\|\boldsymbol{Y}-\boldsymbol{Z}\|_F^2\tag{10}$$

其中:$\boldsymbol{Y}\in\mathcal{Y}_{[\boldsymbol{l},\boldsymbol{u}]}$,$\boldsymbol{Z}\in\mathcal{J}$,$\boldsymbol{\Lambda}\in\boldsymbol{S}^{n+1}$,$\sigma>0$ 是一个惩罚参数。设 $\Lambda_0\geq\boldsymbol{0}$,增广拉格朗日方法在第 k 步迭代中求解:

$$\begin{cases}\min\mathcal{L}_\sigma(\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{\Lambda}^k);\\ \text{s.t.}\quad \boldsymbol{Y}\in\mathcal{Y}_{[\boldsymbol{l},\boldsymbol{u}]},\boldsymbol{Z}\in\mathcal{J}\end{cases}\tag{11}$$

得到最优解$(\boldsymbol{Y}^{k+1},\boldsymbol{Z}^{k+1})$,更新乘子 $\boldsymbol{\Lambda}^{k+1}$:

$$\boldsymbol{\Lambda}^{k+1}=\boldsymbol{\Lambda}^k+\sigma(\boldsymbol{Y}^{k+1}-\boldsymbol{Z}^{k+1})\tag{12}$$

由于求解问题(11)需要对 $\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 联合极小化$\mathcal{L}_\sigma(\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{\Lambda}^k)$,因此求解非常耗时。本文将交替极小化关于 $\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 的拉格朗日函数,不再精确地求解问题(11),即

$$\boldsymbol{Y}^{k+1}=\arg\min_{\boldsymbol{Y}\in\mathcal{Y}_{[\boldsymbol{l},\boldsymbol{u}]}}\mathcal{L}_\sigma(\boldsymbol{Y},\boldsymbol{Z}^k,\boldsymbol{\Lambda}^k)\tag{13}$$

$$\boldsymbol{Z}^{k+1}=\arg\min_{\boldsymbol{Z}\in\mathcal{J}}\mathcal{L}_\sigma(\boldsymbol{Y}^{k+1},\boldsymbol{Z},\boldsymbol{\Lambda}^k)\tag{14}$$

通过对$\mathcal{L}_\sigma(\boldsymbol{Y}^{k+1},\boldsymbol{Z},\boldsymbol{\Lambda}^k)$重新整理得到,问题(14)等价于

$$\min_{\boldsymbol{Z}\in\mathcal{J}}\|\boldsymbol{Z}-\boldsymbol{V}^{k+1}\|_F^2\tag{15}$$

其中:$\boldsymbol{V}^{k+1}=\dfrac{1}{\sigma}\boldsymbol{\Lambda}^k+\boldsymbol{Y}^{k+1}$。问题(15)意味着 $\boldsymbol{Z}^{k+1}$ 为 $\boldsymbol{V}^{k+1}$ 到$\mathcal{J}$上的投影,则有 $\boldsymbol{Z}^{k+1}=\mathrm{proj}_{\boldsymbol{J}}\boldsymbol{V}^{k+1}$。利用引理 2,可得 $\boldsymbol{Z}^{k+1}=\boldsymbol{N}[\boldsymbol{N}^{\mathrm{T}}\boldsymbol{V}^{k+1}\boldsymbol{N}]_+\boldsymbol{N}^{\mathrm{T}}$,其中 $\boldsymbol{N}$ 为(5)中给出的列矩阵,并且$[\boldsymbol{N}^{\mathrm{T}}\boldsymbol{V}^{k+1}\boldsymbol{N}]_+$ 为 $\boldsymbol{N}^{\mathrm{T}}\boldsymbol{V}^{k+1}\boldsymbol{N}\in\boldsymbol{S}^l$ 在 $\boldsymbol{S}_+^l$ 上的正交投影,定义为:

$$[\boldsymbol{N}^{\mathrm{T}}\boldsymbol{V}^{k+1}\boldsymbol{N}]_+=\boldsymbol{P}^k\,\mathrm{diag}([\boldsymbol{\lambda}_1^k]_+,\cdots,[\boldsymbol{\lambda}_l^k]_+)(\boldsymbol{P}^k)^{\mathrm{T}}\tag{16}$$

其中:$[a]_+\coloneqq\max\{0,a\}$,$a\in\mathbb{R}$,$\boldsymbol{\lambda}_1^k,\cdots,\boldsymbol{\lambda}_l^k$ 是$\boldsymbol{N}^{\mathrm{T}}\boldsymbol{V}^{k+1}\boldsymbol{N}$ 按递增排序的特征值,$\boldsymbol{P}^k$是正交分解中的正交矩阵。由(12)可以得出:

$$\boldsymbol{\Lambda}^{k+1}=\boldsymbol{\Lambda}^k+\sigma(\boldsymbol{Y}^{k+1}-\boldsymbol{Z}^{k+1})=\sigma(\boldsymbol{V}^{k+1}-\boldsymbol{Z}^{k+1})\tag{17}$$

根据上述的讨论,可得到求解问题(4)的如下交替方向乘子法(Alternating Direction Method of Multiplier, ADMM)。

算法 1 (ADMM 算法)

步骤 1:输入参数 $\boldsymbol{Z}^0\in\mathcal{J}$,$\boldsymbol{\Lambda}^0\geq 0$,罚参数 $\sigma>0$,$\boldsymbol{l}$、$\boldsymbol{u}$ 以及终止精度 $\varepsilon>0$,利用式(5)计算 $\boldsymbol{N}$,令 $k=0$。

步骤 2:求解问题(13),得最优解 $\boldsymbol{Y}^{k+1}$。

步骤 3:计算 $\boldsymbol{V}^{k+1}=\dfrac{1}{\sigma}\boldsymbol{\Lambda}^k+\boldsymbol{Y}^{k+1}$,对 $\boldsymbol{N}^{\mathrm{T}}\boldsymbol{V}^{k+1}\boldsymbol{N}$ 特征值分解。令 $\boldsymbol{Z}^{k+1}=\boldsymbol{N}[\boldsymbol{N}^{\mathrm{T}}\boldsymbol{V}^{k+1}\boldsymbol{N}]_+\boldsymbol{N}^{\mathrm{T}}$。

步骤 4:若 $\|\boldsymbol{Y}^{k+1}-\boldsymbol{Z}^{k+1}\|\leqslant\varepsilon$,停止,$\boldsymbol{Y}^{k+1}$是问题(7)的一个近似最优解。

步骤 5：计算 $\boldsymbol{\Lambda}^{k+1}=\sigma(\boldsymbol{V}^{k+1}-\boldsymbol{Z}^{k+1})$，令 $k=k+1$，转至步骤 2。

类似于文献[19]中定理 2 的证明方法，可以证得算法 1 的如下收敛性结果。

定理 2　算法 1 产生的序列$\{\boldsymbol{Y}^k\}$收敛于问题(7)的最优解。

3　求解问题(13)的割平面算法

本小节讨论如何求解 ADMM 算法中的问题(13)。首先忽略问题(13)中的常数项，将问题(13)重写为以下形式：

$$\begin{cases}\min \hat{\boldsymbol{Q}}^k \bullet \boldsymbol{Y}+\dfrac{\sigma}{2}\|\boldsymbol{Y}\|_F^2;\\ \text{s.t.}\quad \boldsymbol{Y}\in\boldsymbol{\mathcal{Y}}_{[\boldsymbol{l},\boldsymbol{u}]}\end{cases}\tag{18}$$

其中：$\hat{\boldsymbol{Q}}^k:=\bar{\boldsymbol{Q}}+\boldsymbol{\Lambda}^k-\sigma\boldsymbol{Z}^k$。问题(18)的拉格朗日问题为：

$$\begin{cases}d(\boldsymbol{\lambda},\boldsymbol{\mu}):=\min(\hat{\boldsymbol{Q}}^k+\sum\limits_{i=1}^m\boldsymbol{\mu}_i\bar{\boldsymbol{Q}}_i+\sum\limits_{i=1}^r\boldsymbol{\lambda}_i\hat{\boldsymbol{A}}_{[\boldsymbol{l}_i,\boldsymbol{u}_i]})\bullet\boldsymbol{Y}+\dfrac{\sigma}{2}\|\boldsymbol{Y}\|_F^2;\\ \text{s.t.}\quad \boldsymbol{0}\leqslant\boldsymbol{Y}\leqslant\bar{\boldsymbol{U}},\boldsymbol{Y}_{00}=1\end{cases}\tag{19}$$

其中：$\boldsymbol{\mu}\in\mathbb{R}_+^m,\boldsymbol{\lambda}\in\mathbb{R}_+^r$是拉格朗日乘子。问题(19)是一个严格凸二次规划问题。故问题(19)的最优解是唯一的，且由如下公式给出：

$$\boldsymbol{Y}(\boldsymbol{\lambda},\boldsymbol{\mu}):=\min\left\{\max\left[\boldsymbol{0},-\frac{1}{\sigma}(\hat{\boldsymbol{Q}}^k+\sum_{i=1}^m\boldsymbol{\mu}_i\bar{\boldsymbol{Q}}_i+\sum_{i=1}^r\boldsymbol{\lambda}_i\hat{\boldsymbol{A}}_{[\boldsymbol{l}_i,\boldsymbol{u}_i]})\right],\bar{\boldsymbol{U}}\right\}\tag{20}$$

其中分段取最小值和最大值。问题(18)的拉格朗日对偶问题为：

$$\max_{\boldsymbol{\lambda}\geqslant 0,\boldsymbol{\mu}\geqslant 0} d(\boldsymbol{\lambda},\boldsymbol{\mu})\tag{21}$$

对偶问题(21)可以改写为一个等价的线性规划问题：

$$\begin{cases}\max\quad v;\\ \text{s.t.}\quad v\leqslant(\hat{\boldsymbol{Q}}^k+\sum\limits_{i=1}^m\boldsymbol{\mu}_i\bar{\boldsymbol{Q}}_i+\sum\limits_{i=1}^r\boldsymbol{\lambda}_i\hat{\boldsymbol{A}}_{[\boldsymbol{l}_i,\boldsymbol{u}_i]})\bullet\boldsymbol{Y}+\dfrac{\sigma}{2}\|\boldsymbol{Y}\|_F^2,\forall\boldsymbol{Y}\in[\boldsymbol{0},\bar{\boldsymbol{U}}],\boldsymbol{Y}_{00}=1,\\ \qquad\lambda\geqslant\boldsymbol{0},\boldsymbol{\mu}\geqslant\boldsymbol{0}\end{cases}\tag{22}$$

基于问题(22)，可以得到求解问题(18)的割平面算法(Cutting plane，CP)。

算法 2　(CP 算法)

步骤 1：输入参数$\hat{\boldsymbol{Q}},\boldsymbol{\Lambda}^k,\boldsymbol{Z}^k,\hat{\boldsymbol{A}}_{[\boldsymbol{l},\boldsymbol{u}]},\sigma$ 以及终止精度 $\varepsilon>0$。

步骤 2：选择一个可行解 $\boldsymbol{Y}^0\in\boldsymbol{\mathcal{Y}}_{[\boldsymbol{l},\boldsymbol{u}]}$。令 $j=1$。

步骤 3：求解线性规划问题：

$$(LD^j)\begin{cases}\max\quad v;\\ \text{s.t.}\quad v\leqslant(\hat{\boldsymbol{Q}}^k+\sum\limits_{i=1}^m\boldsymbol{\mu}_i\bar{\boldsymbol{Q}}_i+\sum\limits_{i=1}^r\boldsymbol{\lambda}_i\hat{\boldsymbol{A}}_{[\boldsymbol{l}_i,\boldsymbol{u}_i]})\bullet\boldsymbol{Y}^i+\dfrac{\sigma}{2}\|\boldsymbol{Y}^i\|_F^2,i=0,1,\cdots,j-1,\\ \qquad\boldsymbol{\lambda}\geqslant\boldsymbol{0},\boldsymbol{\mu}\geqslant\boldsymbol{0}.\end{cases}$$

令$(\boldsymbol{\lambda}^j,\boldsymbol{\mu}^j,v^j)$是问题$(LD^j)$的最优解。

步骤 4：由式(20)计算 $\boldsymbol{Y}(\boldsymbol{\lambda}^j,\boldsymbol{\mu}^j)$。令

$$d(\boldsymbol{\lambda}^j,\boldsymbol{\mu}^j)=(\hat{\boldsymbol{Q}}^k+\sum_{i=1}^{\boldsymbol{m}}\boldsymbol{\mu}_i\bar{\boldsymbol{Q}}_i+\sum_{i=1}^r\boldsymbol{\lambda}_i\hat{\boldsymbol{A}}_{[\boldsymbol{l}_i,\boldsymbol{u}_i]})\bullet\boldsymbol{Y}^i+\frac{\sigma}{2}\|\boldsymbol{Y}^i\|_F^2。$$

若 $v^j\leqslant d(\boldsymbol{\lambda}^j,\boldsymbol{\mu}^j)-\varepsilon$，终止，$(\boldsymbol{\lambda}^j,\boldsymbol{\mu}^j)$是问题(22)的最优解。否则，令 $j=j+1$，转步骤 2。

4　数值实验

本节给出问题(1)的 DNP 松弛的数值结果，并通过数值实验验证 ADMM 算法的有效性。数值实验在 Matlab R2017a 上实现，在 PC 机(2.90 Gi Hz，16 GiB，ADM)上运行。对具有随机数据的问题(4)测试了

ADMM 算法,并同商业求解器 CVX 进行数值比较。在数值实验中设置 $\varepsilon=10^{-5}$。

测试问题是带有箱子约束 $x\in[0,1]^n$ 的问题(1),测试问题中的参数由与文献[9]相同的方式随机生成,具体如下:

a)$\boldsymbol{Q},\boldsymbol{q}$ 的形式为:$\boldsymbol{Q}=\boldsymbol{PTP}^{\mathrm{T}}$,其中:

$$\begin{cases}\boldsymbol{T}=\mathrm{diag}(\boldsymbol{T}_1,\cdots,\boldsymbol{T}_n),\\ \boldsymbol{P}=\boldsymbol{W}_1\boldsymbol{W}_2\boldsymbol{W}_3,\\ \boldsymbol{W}_j=\mathbf{I}-2\dfrac{\boldsymbol{w}_j\boldsymbol{w}_j^{\mathrm{T}}}{\|\boldsymbol{w}_j\|^2},j=1,2,3,\\ \boldsymbol{w}_j=(\boldsymbol{w}_{j_1},\cdots,\boldsymbol{w}_{j_n})^{\mathrm{T}}。\end{cases}$$

$\boldsymbol{w}_{j_k}(k=1,\cdots,n)$为区间$[-1,1]$上的随机数,$\boldsymbol{T}_k(k=1,\cdots,r)$服从 $\mathrm{U}(-1,0)$,$\boldsymbol{T}_k(k=r+1,\cdots,n)$服从 $\mathrm{U}(0,1)$,$\boldsymbol{q}$ 服从 $\mathrm{U}(-1,1)$。

b)$\boldsymbol{Q}_i$ 的形式:$\boldsymbol{Q}_i=\boldsymbol{M}_i^{\mathrm{T}}\boldsymbol{M}_i(i=1,\cdots,m)$,$\boldsymbol{M}_i\in\mathrm{U}(-1,1)$,且 $\boldsymbol{M}_i\in\mathbb{R}^n$。

c)$\boldsymbol{A}=(a_{ij})\in\mathbb{R}^{l\times n}$,$\boldsymbol{b}=(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_l)^{\mathrm{T}}$,其中 $a_{ij}\in\mathrm{U}(-5,5)$,$\boldsymbol{b}_i\in\mathrm{U}[v_i,v_i+1]$,且 $v_i=0.5\sum_{j=1}^{n}\max(0,a_{ij})$。

表 1 给出了 ADMM 算法与 CVX 对问题(4)的 10 个随机测试例子的平均数值结果,其中 $m=5$,“迭代步数”、“最优值”和“时间”分别表示算法对 10 个测试问题得到的平均迭代次数、平均最优值和求解问题(4)的平均时间(单位:s);“—”表示算法未能在 3600 s 内找到最优解。由表 1 可知,ADMM 算法可以有效地找到问题(4)的近似最优解,而 CVX 只能求解小规模问题,且当问题维数较高、模型较为复杂时 CVX 无法找到问题的最优解,并且求解速度较慢,针对 $n=200$ 以上的问题,CVX 无法在 3600 s 内求到最优解。相比之下,ADMM 算法可以较快地求解问题(4),并且对于 $n=400$ 或 500 维的大规模问题也能有效地求解。

表 1 CVX 与 ADMM 算法对问题(4)的平均数值结果

n	r	CVX		ADMM		
		最优值	时间/s	最优值	时间/s	迭代步数/步
20	5	−4.0291	1.46	−4.0291	0.66	1059
50	5	−6.1667	21.90	−6.1667	3.19	3267
70	5	−5.9413	128.80	−5.9413	4.68	2140
100	5	−5.6863	784.20	−5.6863	7.48	2501
200	5	—	—	−3.8768	34.28	4014
300	5	—	—	−0.3399	94.85	4999
400	5	—	—	2.7906	324.17	6755
500	5	—	—	10.2634	574.31	7488
20	10	−9.3698	1.45	−9.3698	0.93	1019
50	10	−9.1596	23.22	−9.1596	4.23	4145
70	10	−15.4733	131.30	−15.4733	7.68	3259
100	10	−11.7548	796.70	−11.7548	12.73	3774
200	10	—	—	−9.2691	43.88	4520
300	10	—	—	−4.9593	158.22	6866
400	10	—	—	−14.9077	401.80	6956
500	10	—	—	−8.2283	964.71	9782

5 结 语

本文针对一类带凸二次约束的非凸二次规划问题提出了一个紧 DNP 松弛,并给出了求解 DNP 松弛的一个交替方向乘子法以及求解子问题的割平面算法。随机例子的数值实验表明,相比较于商业求解器 CVX,本文提出的交替方向乘子法可以更有效且快速地求解中等规模 DNP 松弛问题。然而,本文所提出的

DNP 松弛方法只能求解带有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二次规划问题的 DNP 松弛方法。

参考文献:

[1] 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.

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

[3] 申培萍. 全局优化方法[M]. 北京: 科学出版社, 2006: 195-241.

[4] Vandenbussche D, Nemhauser G L. A branch-and-cut algorithm for nonconvex quadratic programs with box constraints [J]. Mathematical Programming, 2005, 102(3): 559-575.

[5] Le Thi H A, Dinh T P. A branch and bound method via d.c. optimization algorithms and ellipsoidal technique for box constrained nonconvex quadratic problems[J]. Journal of Global Optimization, 1998, 13: 171-206.

[6] Hoai An L T. An efficient algorithm for globally minimizing a quadratic function under convex quadratic constraints[J]. Mathematical Programming, 2000, 87(3): 401-426.

[7] Barrientos O, Correa R. An algorithm for global minimization of linearly constrained quadratic functions[J]. J Global Optimization, 2000, 16(1): 77-93.

[8] Burer S, Vandenbussche D. A finite branch-and-bound algorithm for nonconvex quadratic programming via semidefinite relaxations[J]. Mathematical Programming, 2008, 113(2): 259-282.

[9] Luo H Z, Bai X D, Lim G, et al. New global algorithms for quadratic programming with a few negative eigenvalues based on alternative direction method and convex relaxation[J]. Mathematical Programming Computation, 2019, 11(1): 119-171.

[10] Anstreicher K M. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming[J]. Journal of Global Optimization, 2009, 43(2/3): 471-484.

[11] Zheng X J, Sun X L, Li D. Nonconvex quadratically constrained quadratic programming: best D.C. decompositions and their SDP representations[J]. Journal of Global Optimization, 2011, 50(4): 695-712.

[12] Zheng X J, Sun X L, Li D. Convex relaxations for nonconvex quadratically constrained quadratic programming: matrix cone decomposition and polyhedral approximation[J]. Mathematical Programming, 2011, 129(2): 301-329.

[13] Saxena A, Bonami P, Lee J. Convex relaxations of non-convex mixed integer quadratically constrained programs: projected formulations[J]. Mathematical Programming, 2011, 130(2): 359-413.

[14] Luo H Z, Bai X D, Peng J M. Enhancing semidefinite relaxation for quadratically constrained quadratic programming via penalty methods[J]. Journal of Optimization Theory and Applications, 2019, 180(3): 964-992.

[15] Chen J Q, Burer S. Globally solving nonconvex quadratic programming problems via completely positive programming[J]. Mathematical Programming Computation, 2012, 4(1): 33-52.

[16] Burer S. Optimizing a polyhedral-semidefinite relaxation of completely positive programs[J]. Mathematical Programming Computation, 2010, 2(1): 1-19.

[17] Cerulli M, Santis M, Gaar E, et al. Improving ADMMs for solving doubly nonnegative programs through dual factorization[J]. 4OR, 2021, 19(3): 415-448.

[18] Oliveira D E, Wolkowicz H, Xu Y Y. ADMM for the SDP relaxation of the QAP[J]. Mathematical Programming Computation, 2018, 10(4): 631-658.

[19] Wen Z W, Goldfarb D, Yin W T. Alternating direction augmented Lagrangian methods for semidefinite programming[J]. Mathematical Programming Computation, 2010, 2(3/4): 203-230.

(责任编辑:康　锋)