

Gauss Back Substitution Alternating Direction Method for a Class of Inverse Quadratic Programming

Lidan Li^{1,2}, Hongwei Zhang¹, Liwei Zhang¹

¹School of Mathematical Sciences, Dalian University of Technology, Dalian Liaoning

²College of Science, Liaoning Technical University, Fuxin Liaoning

Email: lilidan2000@126.com

Received: Dec. 25th, 2019; accepted: Jan. 7th, 2020; published: Jan. 14th, 2020

Abstract

In this paper, we solve the inverse problem of quadratic programming whose objective function is the sum of matrix spectrum norm and vector infinite norm. We transform the problem into a convex optimization problem with objective function separable and propose Gauss back substitution alternating direction method to solve it. We find that one of its subproblems is still a convex optimization problem with objective function separable, but it is impossible to solve every variable accurately. So we use the inexact method to solve the subproblem. Finally, the numerical experiment of the problem in this paper is given. The data shows that the method in this paper can solve the inverse quadratic programming problem efficiently and quickly.

Keywords

Spectrum Norm, Infinite Norm, Quadratic Programming, G – ADMM Method

Gauss 回代交替方向法求解一类二次规划逆问题

李丽丹^{1,2}, 张宏伟¹, 张立卫¹

¹ 大连理工大学数学科学学院, 辽宁 大连

² 辽宁工程技术大学理学院, 辽宁 阜新

Email: lilidan2000@126.com

收稿日期: 2019 年 12 月 25 日; 录用日期: 2020 年 1 月 7 日; 发布日期: 2020 年 1 月 14 日

摘要

本文求解了目标函数为矩阵谱范数与向量无穷范数之和的一类二次规划的逆问题。先将该问题转化为目标函数可分离变量的凸优化问题, 提出用 Gauss 回代交替方向法求解该问题。而对于其中一个子问题的求解过程中发现其仍是目标函数可分离变量的凸优化问题, 但无法精确求解每个变量, 所以采用非精确方法求解该子问题。最后给出采用的 Gauss 回代交替方向法求解本文问题的数值实验。数据表明, 本文所采用的方法能够高效快速地解决该二次规划逆问题。

关键词

谱范数, 无穷范数, 二次规划, G – ADMM 法

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在本文里, 我们考虑如下凸二次规划问题的逆问题:

$$\begin{aligned} \min \quad & f(x) = \frac{1}{2}x^\top Gx + c^\top x \\ \text{s.t.} \quad & x \in \Omega_P := \{x' \in \mathbf{R}^n | Ax' \geq b\} \end{aligned} \quad (1.1)$$

其中 $G \in \mathcal{S}_+^n$, \mathcal{S}^n 定义 $n \times n$ 对称矩阵空间, \mathcal{S}_+^n 定义了 $n \times n$ 半正定矩阵锥, $c \in \mathbf{R}^n, A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$ 令

$$A := (a_1, \dots, a_m)^\top, a_i \in \mathbf{R}^n, i = 1, \dots, m,$$

我们用 $\text{SOL}(P)$ 表示问题 (P) 的最优解, 并且在本文中假设 $m \leq n$ 。

给定问题 $\text{QP}(G, c, A, b)$ 的最优解 $x^0 \in \Omega_P$ 以及 (G, c) 的估计值 $(G^0, c^0) \in \mathcal{S}^n \times \mathbf{R}^n$ 。在本文中考虑的逆二次优化问题是找到一对值 $(G, c) \in \mathcal{S}_+^n \times \mathbf{R}^n$ 使

$$\text{IQP}(G^0, A, x^0, c^0)$$

$$\begin{aligned} \min \quad & \|G - G^0\|_2 + \|c - c^0\|_\infty \\ \text{s.t.} \quad & x^0 \in \text{SOL}(\text{QP}(G, c, A, b)), \\ & (G, c) \in \mathcal{S}_+^n \times \mathbf{R}^n, \end{aligned} \quad (1.2)$$

其中 $\|\cdot\|_2, \|\cdot\|_\infty$ 分别表示矩阵的谱范数 (即矩阵的最大奇异值) 和向量的无穷范数。

根据传统的对偶理论, 如果 $G \in \mathcal{S}_+^n$, 那么 $x^0 \in \text{SOL}(\text{QP}(G, c, A, b))$ 的充分必要条件是存在 $u \in \mathbf{R}^m$ 满足

$$\begin{cases} c - A^\top u + Gx^0 = 0, \\ 0 \leq u \perp Ax^0 - b \geq 0. \end{cases} \quad (1.3)$$

不失一般性, 假设前 p 个约束在 x^0 处是有效约束, 或者等价写成 $I(x^0) := \{j | a_j^\top x^0 - b_j = 0, j = 1, \dots, m\} = \{1, \dots, p\}$ 。令 $A_0 := (a_1, \dots, a_p)^\top, a_i \in \mathbf{R}^n, i = 1, \dots, p$, 即 A_0 是 A 的前 p 行。则 (1.3) 可以等价地表示为

$$\begin{cases} c - A_0^\top y + Gx^0 = 0, \\ y \geq 0, \end{cases} \quad (1.4)$$

其中 $y \in \mathbf{R}^p$ 是 u 的前 p 个元素。所以问题 $\text{IQP}(G^0, A, x^0, c^0)$ 可以被简化为 $\text{IQP}(G^0, A_0, x^0, c^0)$

$$\begin{aligned} \min \quad & \|G - G^0\|_2 + \|c - c^0\|_\infty \\ \text{s.t.} \quad & c - A_0^\top y + Gx^0 = 0, \\ & (G, c, y) \in \mathcal{S}_+^n \times \mathbf{R} \times \mathbf{R}_+^p, \end{aligned} \quad (1.5)$$

下面我们假设问题(1.5)有解的前提下研究利用交替方向法求解它。由于交替方向法适用于目标函数是可分离的优化问题, 记 $\delta_{\mathbf{R}_+^p}(\cdot)$ 是 \mathbf{R}_+^p 上的指示函数, 我们把问题(1.5)写成如下可分离形式

$$\begin{aligned} \min \quad & \|G - G^0\|_2 + \|c - c^0\|_\infty + \delta_{\mathbf{R}_+^p}(y) \\ \text{s.t.} \quad & c - A_0^\top y + Gx^0 = 0, \\ & (G, c, y) \in \mathcal{S}_+^n \times \mathbf{R}^n \times \mathbf{R}^p. \end{aligned} \quad (1.6)$$

本文的组织结构如下: 第 2 节给出求逆问题所需要的部分预备知识; 第 3 节给出具体求解问题(1.6) 的 Gauss 回代方法的算法以及各子问题求解的具体步骤; 第 4 节给出 Gauss 回代算法求

解逆问题的数值结果；最后一节为结论。

2. 预备知识

我们需要讨论一下 Moreau-Yosida 正则化的一些性质。设 \mathcal{X} 是一个有限维欧氏空间，其内积 $\langle \cdot, \cdot \rangle_{\mathcal{X}}$ 引导的范数记为 $\|\cdot\|_{\mathcal{X}}$ 。 $f: \mathcal{X} \mapsto (-\infty, +\infty]$ 是定义在 \mathcal{X} 上的闭凸函数，则 f 在 $x \in \mathcal{X}$ 点的 Moreau-Yosida 正则化定义如下：

$$\min_{y \in \mathcal{X}} f(y) + \frac{1}{2} \|y - x\|_{\mathcal{X}}^2.$$

上述问题的唯一最优解记为 $P_f(x)$ 定义 f 的共轭函数

$$f^*(x) = \sup_{y \in \mathcal{X}} \{\langle x, y \rangle_{\mathcal{X}} - f(y)\}.$$

由 Moreau 分解理论 [1] (定理 31.5) 可知，对于 $\forall x \in \mathcal{X}$ ，我们有

$$x = P_f(x) + P_{f^*}(x). \quad (2.1)$$

如果 f 是范数函数，记 $f = \|\cdot\|_{\sharp}$ 则记 $\|\cdot\|_{\sharp}$ 为其对偶范数，定义为

$$\|x\|_{\sharp} = \sup_{y \in \mathcal{X}} \{\langle x, y \rangle_{\mathcal{X}} : \|y\|_{\natural} \leq 1\},$$

则有 $P_{f^*}(x) = \Pi_{\mathcal{B}_{\sharp}^1}(x)$ ，其中闭球 $\mathcal{B}_{\sharp}^1 := \{x \in \mathcal{X} : \|x\|_{\sharp} \leq 1\}$ ，而 $\Pi_{\mathcal{M}}(\cdot)$ 是到闭凸集 \mathcal{M} 的投影算子，即

$$\Pi_{\mathcal{M}}(x) = \underset{y \in \mathcal{M}}{\text{Argmin}} \frac{1}{2} \|x - y\|_{\mathcal{X}}^2, \quad \forall x \in \mathcal{X}.$$

则(2.1)可以写成

$$x = P_f(x) + \Pi_{\mathcal{B}_{\sharp}^1}(x). \quad (2.2)$$

或等价为

$$P_f(x) = x - \Pi_{\mathcal{B}_{\sharp}^1}(x). \quad (2.3)$$

对于具体的矩阵的核范数，取 $\hat{\tau} > 0$ ，令 $f_{\hat{\tau}}: \mathbf{R}^{m \times n} \mapsto \mathbf{R}$ ，即 $f_{\hat{\tau}}(X) := \hat{\tau} \|X\|_*$ 函数 $f_{\hat{\tau}}$ 在点 $X \in \mathbf{R}^{m \times n}$ 处 Moreau-Yosida 正则化定义如下：

$$\min_{Y \in \mathbf{R}^{m \times n}} f_{\hat{\tau}}(Y) + \frac{1}{2} \|Y - X\|_F^2. \quad (2.4)$$

其中 $\|\cdot\|_F$ 为矩阵的 Frobenius 范数。事实上，矩阵的核范数的对偶函数是谱范数 (矩阵的奇异值之和)，定义为 $\|\cdot\|_2$ 根据(2.2) 或(2.3)知

$$P_{f_{\hat{\tau}}}(X) = X - \Pi_{\mathcal{B}_{\hat{\tau}}^1}(X), \quad (2.5)$$

其中闭球 $\mathcal{B}_2^{\hat{\tau}}$ 定义为 $\mathcal{B}_2^{\hat{\tau}} := \{X \in \mathbf{R}^{m \times n} : \|X\|_2 \leq \hat{\tau}\}$.

我们要想求解(2.4), 关键是如何求解 $\Pi_{\mathcal{B}_2^{\hat{\tau}}}(X)$ 事实上 $\Pi_{\mathcal{B}_2^{\hat{\tau}}}(X)$ 我们可以由下面两个命题得到。

命题 1. ([2]) 对于 $x \in \mathbf{R}^n$, 令 $u_i = |x_i|, i = 1, \dots, n$. 假设 $u_1 \geq \dots \geq u_n \geq 0$, k 是使 $u_k > z_k \geq u_{k+1}$ 成立的最小正整数, 其中

$$z_k = \frac{\sum_{i=1}^k u_i - \hat{\tau}}{k}.$$

令 $\mathcal{C}_1^{\hat{\tau}} := \{v \in \mathbf{R}^n : \|v\|_1 \leq \hat{\tau}\}$, $\bar{x} := \Pi_{\mathcal{C}_1^{\hat{\tau}}}(x)$ 其中 $\|\cdot\|_1$ 表示向量的 1 范数, 即 $\|v\|_1 = \sum_{i=1}^n |v_i|$ 则当 $x \in \mathcal{C}_1^{\hat{\tau}}$ 时, $\bar{x} = x$ 否则

$$\bar{x}_i = \begin{cases} \text{sign}(x_i)(u_i - z_k), & i \leq k, \\ 0, & i > k. \end{cases} \quad (2.6)$$

设 $r = \text{rank}(X), X \in \mathbf{R}^{m \times n}$ 令 X 的简化奇异值分解为

$$X = U_r \text{Diag}(\sigma(X)) V_r^T, \quad (2.7)$$

其中 $\sigma(X) = (\sigma_1, \dots, \sigma_r)^T$ 满足 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, $U_r \in \mathbf{R}^{m \times r}$, $V_r \in \mathbf{R}^{n \times r}$ 为列正交阵。则利用命题 1 和 John von Neumann 迹不等式, 如下命题成立。

命题 2. 假设 $X \in \mathbf{R}^{m \times n}$ 具有形如(2.7)的奇异值分解, 则 X 在闭球 $\mathcal{B}_*^{\hat{\tau}}$ 上的投影

$$\Pi_{\mathcal{B}_*^{\hat{\tau}}}(X) = U_r \text{Diag}(\bar{x}) V_r^T,$$

其中 $\bar{x} = \Pi_{\mathcal{C}_1^{\hat{\tau}}}(\sigma(X))$

若 f 对应的是向量的无穷范数, 令 $\tilde{\tau} > 0$, $f_{\tilde{\tau}} : \mathbf{R}^n \mapsto \mathbf{R}$ 且 $f_{\tilde{\tau}}(x) := \tilde{\tau} \|x\|_{\infty}$ 函数 $f_{\tilde{\tau}}$ 在点 $x \in \mathbf{R}^n$ 处 Moreau-Yosida 正则化定义如下:

$$\min_{y \in \mathbf{R}^n} f_{\tilde{\tau}}(y) + \frac{1}{2} \|y - x\|^2. \quad (2.8)$$

$\|\cdot\|$ 表示向量的欧式范数。事实上, 向量的无穷范数的对偶函数是向量的 1 范数。根据(2.2)或(2.3), 同样有

$$P_{f_{\tilde{\tau}}}(x) = x - \Pi_{\mathcal{B}_1^{\tilde{\tau}}}(x), \quad (2.9)$$

其中闭球 $\mathcal{B}_1^{\tilde{\tau}}$ 定义为 $\mathcal{B}_1^{\tilde{\tau}} := \{x \in \mathbf{R}^n : \|x\|_1 \leq \tilde{\tau}\}$ 其中 $\Pi_{\mathcal{B}_1^{\tilde{\tau}}}(x)$ 可由命题 1 得到。

3. Gauss 回代交替方向法求解逆问题

经典的交替方向法 (ADMM) 适用于线性约束目标可分离变量的优化问题, 但从收敛性分析上看, 有两个变量的可分离变量形式的问题所对应的交替方向法, 其收敛性证明可在文献 [3] [4] [5] 中找到。而对于变量个数超过两个的问题, 收敛性分析证明仍然是开放的。而我们的问题(1.6)中

的变量是三个, 为了保证收敛, 我们采用的 Gauss 回代交替方向法 [6] 求解它。

问题(1.6)的增广 Lagrange 函数为

$$L_\beta(G, c, y, \lambda) = \|G - G^0\|_2 + \|c - c^0\|_\infty + \delta_{\mathbf{R}_+^p}(y) - \langle \lambda, c - A_0^\top y + Gx^0 \rangle + \frac{\beta}{2} \|c - A_0^\top y + Gx^0\|_F^2$$

其中 $\beta > 0$ 是等式约束 $c - A_0^\top y + Gx^0 = 0$ 的罚参数, $\|\cdot\|_F$ 表示矩阵的 Frobenius 范数, $\|\cdot\|$ 表示向量的欧式范数。为了方便下面的讨论, 我们定义如下的符号以及矩阵

$$\text{集合 } \mathcal{W} := \mathcal{S}_+^n \times \mathbf{R}^n \times \mathbf{R}^p \times \mathbf{R}^n, \mathcal{V} := \mathbf{R}^n \times \mathbf{R}^p \times \mathbf{R}^n.$$

$$w = (G, c, y, \lambda) \in \mathcal{W}, v = (c, y, \lambda) \in \mathcal{V}.$$

矩阵

$$H = \text{Diag}(\beta I, \beta A_0 A_0^\top, \frac{1}{\beta} I), M = \begin{bmatrix} \beta I & 0 & 0 \\ -\beta A_0 & \beta A_0 A_0^\top & 0 \\ 0 & 0 & \frac{1}{\beta} I \end{bmatrix}, Q = \begin{bmatrix} \beta I & -\beta A_0^\top & I \\ -\beta A_0 & \beta A_0 A_0^\top & -A_0 \\ I & -A_0^\top & \frac{1}{\beta} I \end{bmatrix}. \quad (3.1)$$

结合上述矩阵, 下面我们给出 Gauss 回代交替方向法的具体步骤。

算法 3.1 (G-ADMM)

取定参数 $\beta > 0, \alpha \in [0, 1)$, 矩阵 M, H 由上述给出, 给定当前迭代步 $v^k = (c^k, y^k, \lambda^k)$, 则新的迭代步 w^{k+1} 生成如下:

步 1 ADMM 步 (预测步) 获得 $\tilde{w}^k = (\tilde{G}^k, \tilde{c}^k, \tilde{y}^k, \tilde{\lambda}^k)$:

$$\tilde{G}^k = \text{Argmin}_{G \in \mathcal{S}_+^n} \left\{ \|G - G^0\|_2 + \frac{\beta}{2} \|c^k - A_0^\top y^k + Gx^0 - \frac{\lambda^k}{\beta}\|^2 \right\} \quad (3.2)$$

$$\tilde{c}^k = \text{Argmin}_{c \in \mathbf{R}^n} \left\{ \|c - c^0\|_\infty + \frac{\beta}{2} \|c - A_0^\top y^k + \tilde{G}^k x^0 - \frac{\lambda^k}{\beta}\|^2 \right\} \quad (3.3)$$

$$\tilde{y}^k = \text{Argmin}_{y \in \mathbf{R}_+^p} \left\{ \frac{\beta}{2} \|\tilde{c}^k - A_0^\top y + \tilde{G}^k x^0 - \frac{\lambda^k}{\beta}\|^2 \right\} \quad (3.4)$$

$$\tilde{\lambda}^k = \lambda^k - \beta(\tilde{c}^k - A_0^\top \tilde{y}^k + \tilde{G}^k x^0) \quad (3.5)$$

步 2 Gauss 回代步 (矫正步)

$$G^{k+1} = \tilde{G}^k, \\ v^{k+1} = v^k - \alpha M^{-\top} H(v^k - \tilde{v}^k).$$

注解 1. (1) 事实上, Gauss 回代步也可按照如下形式进行, 参见 [6];

$$v^{k+1} = v^k - \gamma \alpha_k^* M^{-\top} H(v^k - \tilde{v}^k), \quad (3.6)$$

其中 $\gamma \in (0, 2)$,

$$\alpha_k^* = \frac{\|v^k - \tilde{v}^k\|_H^2 + \|v^k - \tilde{v}^k\|_Q^2}{2\|v^k - \tilde{v}^k\|_H^2}.$$

(2) 便于理论分析, 我们固定参数 β 。关于动态调整 β , 见文献 [7] [8]。

(3) 算法的收敛性分析见 [6], 这里不再赘述。

3.1. 非精确求解子问题(3.2)

子问题(3.2)可知是下述问题的最优解

$$\begin{aligned} \min \quad & \|Y\|_2 + \frac{\beta}{2} \|b^k + Gx^0\|^2 \\ \text{s.t.} \quad & G - G^0 = Y, \\ & G \succeq 0 \end{aligned} \quad (3.7)$$

其中记 $b^k := c^k - A_0^\top y^k - \lambda^k/\beta$. 我们注意到问题(3.7)是一个目标函数可分离的凸规划问题, 且其未知变量为两个, 所以可以采用经典的交替方向法求解. 此问题与 [9] 中求解问题非常相似, 所以我们也采用它的求解方法. 问题(3.7) 的增广 Lagrange 函数

$$L_\rho(Y, G, \Xi) = \|Y\|_2 + \frac{\rho}{2} \|b^k + Gx^0\|^2 - \langle \Xi, G - G^0 - Y \rangle + \frac{\rho}{2} \|G - G^0 - Y\|_F^2$$

其中 ρ 是对应等式约束的罚参数. 同样给出其交替方向法的迭代形式:

$$Y^{l+1} = \text{Argmin}\{\|Y\|_2 + \frac{\rho}{2} \|G^l - G^0 - Y - \frac{\Xi^l}{\rho}\|_F^2\} \quad (3.8)$$

$$G^{l+1} = \text{Argmin}_{G \in \mathcal{S}_+^n} \{\|Gx^0 + b^k\|^2 + \frac{\rho}{2} \|G - (G^0 + Y^{l+1} + \frac{\Xi^l}{\rho})\|_F^2\} \quad (3.9)$$

$$\Xi^{l+1} = \Xi^l - \rho(G^{l+1} - Y^{l+1} - G^0). \quad (3.10)$$

对于(3.8), 令 $\bar{Y} := G^l - G^0 - \Xi^l/\rho$. 由预备知识的公式(2.4)中的可得

$$Y^{l+1} = \bar{Y} - \Pi_{\mathcal{B}_*^{1/\rho}}(\bar{Y}), \quad (3.11)$$

其中 $\Pi_{\mathcal{B}_*^{1/\rho}}(\cdot)$ 为到闭球 $\mathcal{B}_*^{1/\rho} := \{X \in \mathbf{R}^{m \times n} \mid \|X\|_* \leq 1/\rho\}$ 内的投影算子.

关于(3.9), 我们注意到其没有显示解, 所以讨论其非精确求解. 为保证解的稳定性, 通常对相应子问题进行加二次临近正则化项的正则化处理, 即考虑如下极小值问题

$$G^{l+1} = \text{Argmin}_{G \in \mathcal{S}_+^n} \left\{ \frac{\beta}{2} \|Gx^0 + b^k\|^2 + \frac{\rho}{2} \|G - (G^0 + Y^{l+1} + \frac{\Xi^l}{\rho})\|_F^2 + \frac{1}{2\tau} \|G - G^l\|_F^2 \right\} \quad (3.12)$$

其中 $1/\tau > 0$ 为正则化临近参数. 为了方便讨论, 记(3.12)中的目标函数为 $f(G)$ 则 $f(G)$ 关于 G 的梯度为

$$\nabla f(G) = \frac{\beta}{2} ((Gx^0 + b^k)x^{0\top} + x^0(Gx^0 + b^k)^\top) + \rho \left(G - (G^0 + Y^{l+1} + \frac{\Xi^l}{\rho}) \right) + \frac{1}{\tau} (G - G^l) \quad (3.13)$$

令 G_{exact}^{l+1} 为带正则化项子问题 (3.12) 的精确解, 则有 G_{exact}^{l+1} 满足投影方程

$$G_{exact}^{l+1} = P_{\mathcal{S}_+^n} \left(G_{exact}^{l+1} - \tau \nabla f(G_{exact}^{l+1}) \right).$$

值得注意的是, 对于非精确标准的选取, 应当考虑在具体计算中容易执行的不包含精确解 G_{exact}^{l+1} 信息的非精确标准。为此本文选取 [10] 中给出的非精确标准来计算近似解 G^{l+1} 。其方法是引入误差项 ξ_G^l , 使得近似解满足新的带误差项 ξ_G^l 的投影方程

$$G^{l+1} = \Pi_{S_+^n} (G^{l+1} - \tau \nabla f(G^{l+1}) + \xi_G^l) \quad (3.14)$$

其中误差项 ξ_G^l 满足 [10] 中给的下列两种非精确标准之一,

$$\text{绝对误差标准: } \|\xi_G^l\|_F \leq \varepsilon_l, \text{ 其中 } \sum_0^\infty \varepsilon_l < +\infty, \quad (3.15)$$

或

$$\text{相对误差标准: } \|\xi_G^l\|_F \leq \varepsilon_l \|G^{l+1} - G^l\|_F, \text{ 其中 } \sum_0^\infty \varepsilon_l^2 < +\infty. \quad (3.16)$$

我们选择按照如下方式求解满足(3.15)或(3.16)的误差项 ξ_G^l 和近似解 G^{l+1} 。非精确迭代方法的基本思想是在某种适当的非精确标准下寻求 G_{exact}^{l+1} 的近似解, 记为 G^{l+1} 。首先选择一个收敛到精确解 G_{exact}^{l+1} 的近似序列 $\{G_j^{l+1}\}$ 即

$$G_j^{l+1} \rightarrow \Pi_{S_+^n} (G_j^{l+1} - \tau \nabla f(G_j^{l+1})), \text{ 当 } j \rightarrow \infty. \quad (3.17)$$

定义新的序列

$$\tilde{G}_j^{l+1} := \Pi_{S_+^n} (G_j^{l+1} - \tau \nabla f(G_j^{l+1})). \quad (3.18)$$

由(3.17)知道,

$$G_j^{l+1} - \tilde{G}_j^{l+1} \rightarrow 0, \text{ 当 } j \rightarrow \infty. \quad (3.19)$$

进而对于任意给定的 ε_l 存在整数 i_{k+1} , 使得当 $j > i_{k+1}$ 时, 恒有

$$\|\tau \nabla f(\tilde{G}_j^{l+1}) - \tau \nabla f(G_j^{l+1}) + G_j^{l+1} - \tilde{G}_j^{l+1}\|_F \leq \varepsilon_l \quad (3.20)$$

或者

$$\|\tau \nabla f(\tilde{G}_j^{l+1}) - \tau \nabla f(G_j^{l+1}) + G_j^{l+1} - \tilde{G}_j^{l+1}\|_F \leq \varepsilon_l \|G_j^{l+1} - \tilde{G}_j^{l+1}\|_F. \quad (3.21)$$

事实上(3.18)可以写成

$$\tilde{G}_j^{l+1} = \Pi_{S_+^n} \left(\tilde{G}_j^{l+1} - \tau \nabla f(\tilde{G}_j^{l+1}) + \tau \nabla f(\tilde{G}_j^{l+1}) - \tau \nabla f(G_j^{l+1}) + G_j^{l+1} - \tilde{G}_j^{l+1} \right). \quad (3.22)$$

定义误差项

$$\begin{aligned} \xi_G^l &:= \tau \nabla f(\tilde{G}_j^{l+1}) - \tau \nabla f(G_j^{l+1}) + G_j^{l+1} - \tilde{G}_j^{l+1} \\ &= \frac{\tau \beta}{2} \left((\tilde{G}_j^{l+1} - G_j^{l+1}) x^0 x^{0\top} + x^0 ((\tilde{G}_j^{l+1} - G_j^{l+1}) x^0)^\top \right) + \tau \rho (\tilde{G}_j^{l+1} - G_j^{l+1}). \end{aligned} \quad (3.23)$$

判断误差项 ξ_G^l 是否满足(3.20), 其中 $\sum_0^\infty \varepsilon_l < +\infty$ 或者(3.21), 其中 $\sum_0^\infty \varepsilon_l^2 < +\infty$ 。如果一个成

立, 则停止迭代, 并记

$$G^{l+1} := \tilde{G}_j^{l+1}$$

基于以上分析, 我们采用非精确交替方向法求解子问题(3.7), 而其子问题(3.8)解 Y 可以给出具体解析表达式, 而(3.9)的解 G 每次迭代都只需求一个满足(3.15)或(3.16)的近似解即可。下面给出求解问题(3.7)的非精确交替方向法的迭代格式, 参见 [9]。

算法 3.2 (部分非精确算法求解子问题(3.7))。

取定参数 $\rho > 0, \tau > 0$ 和常数 ε_l 满足(3.15)或(3.16), 给定当前迭代步 (Y^l, G^l, Ξ^l) , 则 $(Y^{l+1}, G^{l+1}, \Xi^{l+1})$ 生成如下:

步 1 按照(3.11)计算 Y^{l+1}

步 2 计算 G^{l+1} 使其满足投影方程 $G^{l+1} = P_{S_{\mathbb{R}}^n} (G^{l+1} - \tau \nabla f(G^{l+1}) + \xi_G^l)$
其中 ξ_G^l 满足非精确标准(3.15) 或(3.16)。

步 3 $\Xi^{l+1} = \Xi^l - \rho(G^{l+1} - Y^{l+1} - G^0)$

显然, 算法 3.2 的有效性取决于快速地求解 Y 和 G 子问题, 因为 Y 子问题可以精确求解, 故算法 3.2 的关键是选取快速的内迭代求解 G 子问题以获取满足非精确标准的近似解 G^{l+1} 。注意到对于带正则项的 G 子问题(3.12), 其实质是一个标准的 Frobenius 范数下结构约束矩阵最小二乘问题。我们选择谱投影梯度算法 (SPG) [11] 求近似解 G^{l+1} , 利用 SPG 算法在非精确标准(3.15)或(3.16)下计算(3.12)的近似解 G^{l+1} 和 ξ_G^l 的迭代格式如下:

算法 3.3 (SPG 算法在非精确标准 (3.15)或(3.16)下计算(3.12)的近似解 G^{l+1} 和 ξ_G^l)。

步 1 输入正整数 $\bar{M} > 1$, 参数 $\alpha_{min} > 0, \alpha_{max} > \alpha_{min}, \gamma \in (0, 1)$, 以及 $0 < \sigma_1 < \sigma_2 < 1$. 任取 $\alpha_0 \in [\alpha_{min}, \alpha_{max}]$, 并取算法 1 中的第 l 步 G^l 作为初始矩阵, 即令 $G_0 = G^l$, 并令 $j \leftarrow 0$

步 2 计算 $D_j = P_{S_{\mathbb{R}}^n} (G_j - \alpha_j \nabla f(G_j)) - G_j$, 并令 $\lambda \leftarrow 1$ 。

步 3 计算 $\hat{G} = G_j + \lambda D_j$

步 4 若

$$f(\hat{G}) \leq \max_{0 \leq t \leq \min\{j, \bar{M}-1\}} f(G_{j-t}) + \gamma \lambda \langle D_j, \nabla f(G_j) \rangle, \quad (3.24)$$

则令 $G_{j+1} = \hat{G}, s_j = G_{j+1} - G_j, y_j = \nabla f(G_{j+1}) - \nabla f(G_j)$, 进入步 5; 否则定义 $\lambda_{new} \in \{\sigma_1 \lambda, \sigma_2 \lambda\}, \lambda \leftarrow \lambda_{new}$, 返回步 3。

步 5 计算 $b_j = \langle s_j, y_j \rangle$, 若 $b_j \leq 0$, 则令 $\alpha_{j+1} = \alpha_{max}$; 否则计算 $a_j = \langle s_j, s_j \rangle$ 和 $\alpha_{j+1} = \min\{\alpha_{max}, \max\{\alpha_{min}, a_j/b_j\}\}$, 并令 $j \leftarrow j + 1$

步 6 令 $\tilde{G}_j = P_{S_{\mathbb{R}}^n} (G_j - \tau \nabla f(G_j))$, $\xi_G^l = \tau \nabla f(\tilde{G}_j) - \tau \nabla f(G_j) + G_j - \tilde{G}_j$,
若 $\|\xi_G^l\|_F \leq \varepsilon_l$ 成立, 其中 $\sum_{l=0}^{\infty} \varepsilon_l < +\infty$; 或者 $\|\xi_G^l\|_F \leq \varepsilon_l \|G^l - \tilde{G}_j\|_F$, 其中 $\sum_{l=0}^{\infty} \varepsilon_l^2 < +\infty$,
则终止迭代并令 $G^{l+1} = \tilde{G}_j$ 否则返回到步 2。

该算法的收敛性分析参看 [9], 这里不再详述。算法 3.2 的收敛准则标准取为

$$\max (\|Y^{l+1} - Y^l\|_F, \|G^{l+1} - G^l\|_F, \|G^l - G^0 - Y^{l+1}\|_F) \leq \varepsilon,$$

其中 ε 为预设的精度。

3.2. 求解子问题(3.3)

记 $c' = c - c^0, \bar{c} = A_0^\top y^k - \tilde{G}^k x^0 - \frac{\lambda^k}{\beta} - c^0$, 则子问题(3.3)另写为

$$\tilde{c}^k = \operatorname{Argmin}_{c' \in \mathbf{R}^n} \left\{ \|c'\|_\infty + \frac{\beta}{2} \|c' - \bar{c}\|^2 \right\} + c^0.$$

由预备知识中的公式(2.2)或(2.3)可知

$$\operatorname{Argmin}_{c' \in \mathbf{R}^n} \left\{ \|c'\|_\infty + \frac{\beta}{2} \|c' - \bar{c}\|^2 \right\} = \bar{c} - \Pi_{\mathcal{B}_1^{1/\rho}}(\bar{c}),$$

其中 $\Pi_{\mathcal{B}_1^{1/\rho}}(\cdot)$ 为到闭球 $\mathcal{B}_1^{1/\rho} := \{x \in \mathbf{R}^n \mid \|x\|_1 \leq 1/\rho\}$ 内的投影算子。而 $\Pi_{\mathcal{B}_1^{1/\rho}}(\cdot)$ 的具体计算由命题 1 得。

3.3. 求解子问题(3.4)

\tilde{y}^k 本质是下述问题的最优解

$$\begin{aligned} \min \quad & \frac{\beta}{2} \left\| \tilde{c}^k - A_0^\top y + \tilde{G}^k x^0 - \frac{\lambda^k}{\beta} \right\|^2 \\ \text{s.t.} \quad & y \geq 0 \end{aligned}$$

为了方便, 记 $N^k := \tilde{c}^k + \tilde{G}^k x^0 - \frac{\lambda^k}{\beta}, Q_0 := A_0 A_0^\top$ 如果 A_0 是行满秩矩阵, 则 \tilde{y}^k 是下述简单凸二次规划问题的解

$$\begin{aligned} \min \quad & f_k := \frac{1}{2} y^\top Q_0 y - \langle A_0 N^k, y \rangle \\ \text{s.t.} \quad & y \geq 0 \end{aligned} \tag{3.25}$$

我们只需利用 matlab 自带的函数 quadprog 求解此问题即可。

4. 数值实验

本节给出利用算法 3.1 求解逆问题(1.6)的数值实验。验的测试环境为 Win7 旗舰版操作系统, 存 4.0G, 频 3.40Hz, MATLAB R2014a。

1. 初始参数 $A_0 \in \mathbf{R}^{p \times n}, c^0 \in \mathbf{R}^n, G^0 \in \mathcal{S}^n$ 均为随机生成, $x^0 = \text{ones}(n, 1)$ 。
2. 初始点的选取为

$$\begin{aligned} c^1 &= \text{zeros}(n, 1), \\ y^1 &= \text{ones}(p, 1), \\ \lambda^1 &= \text{zeros}(n, 1). \end{aligned}$$

3. 算法 3.1 的终止准则为

$$\text{stopc} = \max\left\{\frac{\|G^{k+1} - G^k\|_F}{\|G^k\|_F + 1}, \frac{\|c^{k+1} - c^k\|_1}{\|c^2 - c^1\|_1}, \frac{\|y^{k+1} - y^k\|_1}{\|y^2 - y^1\|_1}\right\} \leq \epsilon,$$

ϵ 为预设精度, 实验中取 $\epsilon = 10^{-4}$ 。

4. $\beta = 5$ 。

表 1 中各个变量说明: p , n 表示矩阵 A_0 的行数和列数, time(s) 和 iter 分别表示算法 1 的运行时间和迭代次数, res 表示终止准则 stopc 的最终值。

Table 1. Numerical results of algorithms 3.1

表 1. 算法 3.1 的数值结果

p	n	time(s)	iter	res
20	20	8.36	46	0.925e-04
20	50	25.02	36	0.873e-04
50	50	126.63	38	0.569e-04
50	100	545.24	67	0.654e-04
100	100	1134.14	98	0.176e-04
100	200	2197.03	31	0.54e-04
200	200	6231.06	42	0.427e-04

5. 结论

本文求解了一类二次规划逆问题, 该问题可转化为目标函数可分离变量的线性约束问题, 由于其未知变量超过两个, 为了保证收敛性, 故采用 Gauss 回代交替方向法求解, 最后给出数值实验, 从数值实验中我们可以看出, 法 3.1 对于求解本文提出的二次规划逆问题(1.2)在维数不是很高的情况下还是非常有效的。

参考文献

- [1] Rockafellar, R.T. (1970) Convex Analysis. Princeton University Press, Princeton.
<https://doi.org/10.1515/9781400873173>
- [2] John, D., Shai, S.S., Yoram, S., et al. (2008) Efficient Projections onto the ℓ_1 -Ball for Learning in High Dimensions. *Proceedings of the 25th International Conference on Machine Learning*, Helsinki, Finland, July 2008, 272-279.
- [3] Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2011) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning*, **3**, 1-122. <https://doi.org/10.1561/2200000016>

-
- [4] Eckstein, J. and Bertsekas, D.P. (1992) On the Douglas—Rachford Splitting Method and the Proximal Point Algorithm for Maximal Monotone Operators. *Mathematical Programming*, **55**, 293-318. <https://doi.org/10.1007/BF01581204>
- [5] Gabay, D. and Mercier, B. (1976) A Dual Algorithm for the Solution of Nonlinear Variational Problems via Finite Element Approximation. *Computers Mathematics with Applications*, **2**, 17-40. [https://doi.org/10.1016/0898-1221\(76\)90003-1](https://doi.org/10.1016/0898-1221(76)90003-1)
- [6] He, B., Tao, M. and Yuan, X. (2012) Alternating Direction Method with Gaussian Back Substitution for Separable Convex Programming. *SIAM Journal on Optimization*, **22**, 313-340. <https://doi.org/10.1137/110822347>
- [7] He, B., Wang, S. and Yang, H. (2003) A Modified Variable-Penalty Alternating Directions Method for Monotone Variational Inequalities. *Journal of Computational Mathematics*, **21**, 495-504.
- [8] He, B.S., Yang, H. and Wang, S.L. (2000) Alternating Direction Method with Self-Adaptive Penalty Parameters for Monotone Variational Inequalities. *Journal of Optimization Theory and Applications*, **106**, 337-356. <https://doi.org/10.1023/A:1004603514434>
- [9] 李姣芬, 宋丹丹, 李涛, 等. 核范数和谱范数下广义 Sylvester 方程最小二乘问题的有效算法 [J]. 计算数学, 2017, 39(2): 129-150.
- [10] Ng, M.K., Wang, F. and Yuan, X. (2011) Inexact Alternating Direction Methods for Image Recovery. *SIAM Journal on Scientific Computing*, **33**, 1643-1668. <https://doi.org/10.1137/100807697>
- [11] Birgin, E.G., Mario, J. and Raydan, M.M. (2000) Nonmonotone Spectral Projected Gradient Methods on Convex Sets. *SIAM Journal on Optimization*, **10**, 1196-1211. <https://doi.org/10.1137/S1052623497330963>