

正交约束优化问题的一阶算法*

高斌^{1,2,†} 刘歆^{1,2} 袁亚湘²

摘要 带有正交约束的矩阵优化问题在材料计算、统计及数据分析等领域中有着广泛的应用。由于正交约束的可行域是 Stiefel 流形，一直以来流形上的优化方法是求解这一问题的主要方法。近年来，随着实际应用问题所要求的变量规模的扩大，传统的流形优化方法在计算上的劣势显现出来，而一些迭代简单、收敛快的新算法逐渐被提出。通过收缩方法、非收缩可行方法、不可行方法三个类别分别来介绍求解带有正交约束的矩阵优化问题的最新算法。通过分析这些方法的主要特性，以及应用问题的要求，对这类问题算法设计的研究进行了展望。

关键词 正交约束, Stiefel 流形, 收缩方法

中图分类号 O221.2, O24, O29

2010 数学分类号 15A18, 65F15, 65K05, 90C06

First-order algorithms for optimization problems with orthogonality constraints*

GAO Bin^{1,2,†} LIU Xin^{1,2} YUAN Yaxiang²

Abstract Optimization problems with orthogonality constraints have a wide range of applications in the field of materials science, statistics and data science. Many optimization algorithms on manifold can be applied to this type of problems, since the feasible region of orthogonal constraint is known as Stiefel manifold. In recent years, with the expansion of variable scale required by practical application, the limitations of existing methods on manifold are reflected in practice. On the other hand, some efficient approaches based on new concepts are proposed recently. In this paper, we briefly introduce the main classes of methods for optimization problems with orthogonality constraints including retraction based method, non-retraction based method and infeasible method respectively. We also discuss the main characteristics of these approaches, the scenarios in which these approaches are suitable and the possible directions for further development.

Keywords orthogonality constraints, Stiefel manifold, retraction based method

Chinese Library Classification O221.2, O24, O29

2010 Mathematics Subject Classification 15A18, 65F15, 65K05, 90C06

收稿日期: 2017-09-02

*基金项目: 国家自然科学基金项目(Nos. 11471325, 91530204, 11622112, 11688101, 11331012, 11461161005), 中国科学院前沿科学重点研究计划 (No. QYZDJ-SSW-SYS010), 国家数学与交叉科学中心, 中国科学院科学与工程计算国家重点实验室

1. 中国科学院大学, 北京 100190; University of Chinese Academy of Sciences, Beijing 100190, China

2. 中国科学院数学与系统科学研究院, 科学与工程计算国家重点实验室, 北京 100190; State Key Laboratory of Scientific and Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

† 通信作者 E-mail: gaobin@lsec.cc.ac.cn

0 引言

带有正交约束的矩阵优化问题 (下面简称为正交约束优化问题) 是指具有如下形式的优化问题.

$$\begin{aligned} \min_{X \in \mathbb{R}^{n \times p}} \quad & f(X) \\ \text{s. t.} \quad & X^\top X = I_p, \end{aligned} \quad (0.1)$$

其中 $p \leq n$, I_p 是 p 阶单位矩阵, 变量 X 是 $n \times p$ 的实矩阵. 在本文中, 我们讨论的目标函数 $f: \mathbb{R}^{n \times p} \rightarrow \mathbb{R}$ 一般为连续函数.

正交约束优化问题在科学与工程中有广泛应用. 例如, 在材料科学中, 密度泛函理论[1-3]的电子结构计算的本质就是求解一个带有正交约束的能量极小问题; 除此之外, 与统计与数据科学密切相关的线性特征值问题[4-5]、低秩协方差矩阵的估计[6]、正交 Procrustes 问题[7-8]、稀疏主成分分析[9-10]等问题, 其核心计算问题也是正交约束优化问题. 关于其他应用, 我们推荐读者参考文献[11-13].

由于 $X^\top X$ 的对称性, 问题 (0.1) 的约束是由 $p(p+1)/2$ 个二次等式组成, 其本质是一个非凸问题. 在大部分情况下, 这个问题是 NP-难的[14]. 即使对于目标函数 f 是二次函数的情形, 我们也仅能在非常特殊的情况下, 例如目标函数是齐次二次函数或 $p = 1$ 时, 才能保证求得问题的全局最优解, 对于其他情况, 我们仍不清楚这个问题是否 NP 或 NP-难的. 其次, 保持迭代点列的可行性意味着每一步需要做正交化. 正交化过程需要 $O(np^2)$ 的计算量, 并且因为其不具备可扩展性, 当 p 比较大时, 正交化会成为计算瓶颈. 这些挑战使得越来越多的专家和学者研究正交约束优化问题.

本文将回顾求解正交约束优化问题的主要方法. 可行方法包括收缩方法 (retraction based method) 和非收缩方法, 将分别在第 1、2 节中介绍. 不可行方法的简要介绍和正交约束方法研究的展望在第 3 节中.

为了方便起见, 我们首先给出一些常用的记号定义. 两个矩阵 $X, Y \in \mathbb{R}^{n \times p}$ 的欧氏内积定义为 $\langle X, Y \rangle = \text{tr}(X^\top Y)$, 其中 $\text{tr}(A)$ 是矩阵 $A \in \mathbb{R}^{p \times p}$ 的迹. $\|\cdot\|_2$ 和 $\|\cdot\|_F$ 分别表示 2 范数和 Frobenius 范数. X^\dagger 表示 X 的伪逆. 矩阵 $X \in \mathbb{R}^{n \times p}$ 的第 i 列定义为 X_i . $X_{\bar{i}} \in \mathbb{R}^{n \times (p-1)}$ 表示矩阵 X 去掉第 i 列, 也就是 $X_{\bar{i}} = [X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_p]$. 我们用 $X_{i,v} \in \mathbb{R}^{n \times p}$ 表示 X 的第 i 列被替换为给定向量 v 的矩阵. 球心为 $C \in \mathbb{R}^{m_1 \times m_2}$, 半径为 r 的球 $\mathcal{B}(C, r)$ 定义为 $\{X \in \mathbb{R}^{m_1 \times m_2} \mid \|X - C\|_F \leq r\}$. $\mathbf{qr}(X)$ 是列满秩矩阵 X 的部分 QR 分解*的 Q 矩阵. $\mathcal{P}_{\mathcal{S}_{n,p}}(X)$ 表示 X 到 Stiefel 流形 $\mathcal{S}_{n,p} := \{X \in \mathbb{R}^{n \times p} \mid X^\top X = I_p\}$ 的投影†.

1 收缩方法

正交约束优化问题 (0.1) 的可行域 $\mathcal{S}_{n,p} = \{X \in \mathbb{R}^{n \times p} \mid X^\top X = I_p\}$ 通常被称为 Stiefel 流形[15]. 如果将问题 (0.1) 限制在流形上, 那么这就是一个流形上的无约束优化问

* $Q \in \mathbb{R}^{n \times p}$ 是列满秩矩阵 $X \in \mathbb{R}^{n \times p}$ 的部分 QR 分解的 Q 矩阵, 如果 $X = QR$, $Q \in \mathbb{R}^{n \times p}$ 正交并且 $R \in \mathbb{R}^{p \times p}$ 是具有正对角元的上三角矩阵.

† $\mathcal{P}_{\mathcal{S}_{n,p}}(X) = \tilde{U}\tilde{V}^\top$, 其中 $\tilde{U}\tilde{V}^\top$ 是 X 的部分奇异值分解, $\tilde{U} \in \mathbb{R}^{n \times p}$, $\tilde{V} \in \mathbb{R}^{p \times p}$ 都是正交矩阵, 并且 $\Sigma \in \mathbb{R}^{p \times p}$ 是对角矩阵.

题. 因此只要保证迭代点都在流形上, 各种无约束优化算法都对应着相应的流形上的优化算法. 例如, 梯度类方法^[16–18]、共轭梯度法^[11,19]、信赖域方法^[20–21]、牛顿法^[11]和拟牛顿法^[22–24]等等. 这些流形上的优化算法都可以用于特殊的流形 Stiefel 流形. 从约束优化的角度来看, 这些方法都是可行方法, 也就是都旨在找一个函数值下降的可行点作为下一步迭代点. 感兴趣的读者可以参考专著 [25].

在文献 [25] 中, 作者从优化的角度系统地研究了 Stiefel 流形的几何结构, 并提出了一个新的概念“收缩” (Retraction), 并用它来刻画当时所有已有算法的共性, 给出了一个统一的算法框架. 一个映射 $\mathcal{R}_X : \mathcal{T}_X \mathcal{S}_{n,p} \rightarrow \mathcal{S}_{n,p}$ 被称为收缩, 如果它满足

- (1) $\mathcal{R}_X(0_X) = X$, 其中 0_X 是 $\mathcal{T}_X \mathcal{S}_{n,p}$ 的原点;
- (2) $\frac{d}{dt} \mathcal{R}_X(tZ)|_{t=0} = Z$, 对所有的 $Z \in \mathcal{T}_X \mathcal{S}_{n,p}$ 都成立,

其中 $\mathcal{T}_X \mathcal{S}_{n,p} := \{Y \in \mathbb{R}^{n \times p} \mid Y^\top X + X^\top Y = 0\}$ 是 Stiefel 流形 $\mathcal{S}_{n,p}$ 在点 X 处的切空间. 收缩 \mathcal{R}_X 把切空间中的切向量映射到流形上, 实际上它定义了一种保持正交性的更新准则.

对于正交约束优化问题, 有两大类收缩方法. 第一类方法是沿着流形的测地线搜索合适的试探步, 这类方法叫作测地线类方法^[11,18,25]. 如果选取切空间方向 $D \in \mathcal{T}_X \mathcal{S}_{n,p}$, 则文[11]中提出的具体的可行搜索轨迹如下,

$$Y_{\text{geoe}}(\tau; X) = [X, Q] \exp \left(\tau \begin{bmatrix} -X^\top D & -R^\top \\ R & 0 \end{bmatrix} \right) \begin{bmatrix} I_p \\ 0 \end{bmatrix}, \quad (1.2)$$

其中 $QR = -(I - XX^\top)D$ 表示 QR 分解. 另一类测地线由文 [18] 给出,

$$Y_{\text{geoc}}(\tau; X) = \exp(-\tau W)X, \quad (1.3)$$

其中 $W \in \mathbb{R}^{n \times n}$ 是一个反对称矩阵. 由更新公式(1.2)和(1.3)可知, 计算测地线需要涉及矩阵指数运算, 这会带来计算上的困难. 文献 [17] 基于 Cayley 变换提出了一种拟测地线更新公式,

$$Y_{\text{qgeo}}(\tau; X) = \left(I + \frac{\tau}{2} W \right)^{-1} \left(I - \frac{\tau}{2} W \right) X, \quad (1.4)$$

其中 $W \in \mathbb{R}^{n \times n}$ 是一个反对称矩阵. 在实际计算中, 拟测地线公式每一步迭代的运算量是求解一个 n 阶线性方程组.

另一类方法主要包含两个步骤, 先在切空间进行线搜索然后投影回到 Stiefel 流形. 因此, 这类方法被称为投影类方法^[16,25–26]. 其中奇异值分解, QR 分解和极分解 (polar decomposition) 都可以用来计算到 Stiefel 流形的投影. 给定切空间方向 $D \in \mathcal{T}_X \mathcal{S}_{n,p}$, 文[16]的作者提出了标准的投影方法,

$$Y_{\text{pj}}(\tau; X) = \mathcal{P}_{\mathcal{S}_{n,p}}(X - \tau D). \quad (1.5)$$

基于 QR 分解和极分解, 文献 [25] 提出了两种投影类收缩方法, 具体如下

$$Y_{\text{qr}}(\tau; X) = \text{qr}(X - \tau D), \quad (1.6)$$

$$Y_{\text{pd}}(\tau; X) = (X - \tau D)(I_p + \tau^2 D^\top D)^{-1/2}. \quad (1.7)$$

事实上, 基于极分解的投影方法 (1.7) 和标准投影方法 (1.5) 等价.

当 $p = 1$ 时, 投影类方法与测地线类方法有相同的迭代轨迹. 上面提到的这些属于收缩的方法, 包括测地线类和投影类方法, 都需要结合具体的线搜索策略, 在测地线或切空间上搜索, 例如 Armijo 非精确线搜索^[27]或者非单调线搜索^[28]. 线搜索策略用来保证算法的全局收敛性, 但与此同时也会增加额外的函数值, 梯度值估计, 从而增加计算代价.

1.1 Cayley变换的简约计算

文献 [12] 针对正交约束优化问题提出了一个保持约束可行的有效方法. 该文献作者在之前的工作^[29]中, 利用求解热方程或者其他偏微分方程组的 Crank-Nicholson 技术, 针对球面约束提出了曲线搜索方法. 在文 [12] 中, 他们将这个思想推广到了矩阵变量, 同样提出了一类曲线搜索方法. 事实上, Crank-Nicholson 技术在矩阵计算中也被称为 Cayley 变换, 作者提出了一种与 Cayley 变换 (1.4) 完全等价的更新公式.

$$Y_{wy}(\tau; X) = X - \tau U \left(I_{2p} + \frac{\tau}{2} V^T U \right)^{-1} V^T X, \quad (1.8)$$

其中 $U = [P_X D, X]$, $V = [X, -P_X D]$, $D \in \mathcal{T}_X \mathcal{S}_{n,p}$ 是任意给定的切空间方向, $P_X = (I_n - \frac{1}{2} X X^T)$. 值得说明的是, 这个公式也是收缩方法, 由于文[17]根据 Cayley 变换提出了拟测地线方法, 所以更新公式 (1.8) 也是一种测地线类方法. 这个公式每步需求解一个 $2p$ 阶的线性方程组, 而不是 Cayley 变换 (1.4) 中的 n 阶线性方程组, 这极大地降低了计算量. 其次, 通过结合 Barzilai-Borwein (BB) ^[20]非单调线搜索策略^[31], 文献 [12] 提出的方法相较于其他收缩方法, 具有更低计算复杂度, 并且在实际数值计算中表现更加优异.

1.2 保正交性迭代框架

文献 [13] 的作者推广了文 [12] 的想法, 通过在可行点 $X \in \mathcal{S}_{n,p}$ 处对全空间进行子空间分解, 定义搜索轨迹为

$$Y_{jd}(\tau; X) = XR(\tau) + WN(\tau), \quad (1.9)$$

其中 $XR(\tau)$ 在 X 的值空间中, $WN(\tau)$ 在 X^T 的零空间中. 文[13] 的主要思想是寻找合适的 $R(\tau)$ 和 $N(\tau)$, 使得轨迹 $Y_{jd}(\tau; X)$ 始终保持可行性. 也就是对于任意的 $\tau > 0$, 都有 $Y_{jd}(\tau; X) \in \mathcal{S}_{n,p}$. 作者根据不定系数法提出了一大类收缩的算法框架, 并且上述的某些测地线类和投影类的方法都可以看做是他们算法框架的特例, 例如 Y_{geoe} , Y_{wy} , Y_{pd} , Y_{pj} , Y_{qr} 以及文献[13]中的 Y_{gp} .

通过结合 BB 方法^[20]达到加速收敛的目的, 同时为了保证全局收敛性, 作者针对正交约束优化问题提出了一种自适应可行的类 BB 方法 (daptive feasible BB-like (AFBB)) 如下.

$$\begin{cases} W = -(I_n - X X^T)D, \\ J(\tau) = I_p + \frac{\tau^2}{4} W W^T + g(\tau) X^T D, \\ Y_{jd}(\tau; X) = (2X + \tau W)J(\tau)^{-1} - X. \end{cases}$$

其中 $D \in \mathcal{T}_X \mathcal{S}_{n,p}$ 是任意给定的切空间方向, $g(\tau)$ 是任一满足 $g(0) = 0, g'(0) = 1/2$ 的函数. 值得一提的是, 满足这个更新公式的方法也是收缩方法. 作者还给出了上述算法框架结合 BB 非单调线搜索的全局收敛性.

1.3 随机方差下降梯度算法

最近, 文献 [32] 的作者将文献[13]的工作推广到了如下带有正交约束的经验风险最小化问题:

$$\begin{aligned} \min_{X \in \mathbb{R}^{n \times p}} \quad & f(X) := \frac{1}{m} \sum_{i=1}^m f_i(X) \\ \text{s. t.} \quad & X^\top X = I_p. \end{aligned} \quad (1.10)$$

针对目标函数带有大规模训练数据集的经验风险最小化问题, 文献 [33] 提出的随机方差下降梯度方法, 是求解这类问题的非常有效的方法. 目前存在的大部分黎曼流形随机梯度类方法, 都需要考虑特定的收缩方法, 并且需要引入平行移动 (parallel transport) 或者向量移动 (vector transport) 来保持随机搜索方向在切空间中, 这需要大量计算. 在文献 [32] 中, 作者提出了一种不需要引入向量移动的一般流形上的随机方差下降梯度方法, 并且这种算法能够用于求解一般的目标函数, 也能适用于一般的收缩方法. 作者分析了算法的收敛性并给出了算法的复杂度. 通过利用 Lojasiewicz 不等式, 作者证明了算法的局部线性收敛速度, 并且说明了矩阵填充问题以高概率满足 Lojasiewicz 不等式. 通过结合 BB 步长, 文献 [32] 提出了一个实际计算非常有效的算法. 大量的数值实验表明, 针对主成分分析问题 and 矩阵完整化问题, 该算法十分有效.

1.4 其他算法及软件包

近两年来, 在收缩框架下, 还提出了一些有效的改进牛顿方法^[34,35], 这些方法既利用了牛顿方法在局部的二次收敛速度, 又结合了一些全局化的修正技术, 确保了算法的全局收敛性.

比较著名的正交约束软件包, 包括 Manopt^[36], 这是一个 Matlab 语言的软件包, 整合了时下最新的黎曼流形优化算法. 常用微分几何的运算都内嵌在软件包中, 这使得整个算法包更便于研究者以及低门槛用户使用.

另一个比较好用的软件包 OptM 是基于文献 [12] 开发的, 针对一般的正交约束优化问题的 Matlab 语言软件包. 该软件包调用简单并且在大部分正交约束优化问题中表现稳定而高效.

2 非收缩可行方法

非收缩可行方法是指迭代点满足正交约束但不对应于一个收缩的方法. 这类方法是文献 [37] 的作者所提出的. 首先这类方法适用于目标函数满足下述条件的情形:

假设 2.1 (i) f 二次可微, 定义

$$\rho := \sup_{X \in \tilde{\mathcal{S}}} \|\nabla^2 f(X)\|_2,$$

其中 $\tilde{\mathcal{S}} := \{Y \mid \|Y\|_F^2 < p + 1\}^\ddagger$.

(ii) $f(X)$ 能够表示为 $h(X) + \text{tr}(G^\top X)$, 其中 $G \in \mathbb{R}^{n \times p}$, 并且 $h(X)$ 正交不变, 也就是说, 对于任意的 $Q \in \mathcal{S}_{p,p}$, $h(XQ) = h(X)$ 都成立. 同时, $\nabla h(X) = H(X)X$, 其中 $H: \mathbb{R}^{n \times p} \rightarrow S^n$ 是矩阵函数.

[‡] $\tilde{\mathcal{S}}$ 可以为任意包含 $S_{n,p}$ 的有界开集

这里, 集合 S^n 表示所有 n 阶对称矩阵全体. 当 $\rho = 0$ 时, 目标函数 $f(X)$ 退化成一个线性函数. 在这种情况下, 问题 (0.1) 有显式解 $X = -RQ^\top$, 这里 RSQ^\top 是 G 的部分奇异值分解. 所以下面假定 $\rho \neq 0$.

文献[37]中提出的非收缩算法框架的全局收敛性分析需要假设 2.1, 但并不是说这样的限制是必须的. 幸运的是, 有应用背景的正交约束优化问题, 比如电子结构计算中的 Kohn-Sham 能量极小^[44], 线性特征值问题等都满足这个假设. 其主要思想是重新考虑了问题 (0.1) 的一阶最优性条件, 并给出了如下的等价形式:

$$\begin{cases} (I_n - XX^\top)\nabla f(X) = 0; & \text{次稳定性} \\ X^\top \nabla f(X) = \nabla f(X)^\top X; & \text{对称性} \\ X^\top X = I_p. & \text{可行性} \end{cases} \quad (2.1)$$

为了方便起见, 式(2.1)的三个条件分别称为次稳定性[§], 对称性和可行性. 基于一阶最优性条件, 作者提出了一个包含两个主要步骤的算法框架.

第一步是函数值下降, 也就是说找到一个可行点使得函数值下降, 并且下降量和投影梯度 (在当前迭代点转置的零空间里) 的范数平方成一个确定比例. 在文 [37] 中, 作者提出了两类具体实现方法. 梯度反射法 (GR) 和梯度投影法 (GP) 是第一类方法的代表, 它们分别用不同的方法将梯度下降步拉回到 Stiefel 流形上. 第二类算法利用了以列为块的块坐标下降迭代 (CBCD), 并且针对子问题, 提出了有效的非精确算法, 并证明了 CBCD 的框架是符合下降性要求的.

第二步是寻找一个可行点使其满足对称性. 作者根据假设 2.1 提出了一个新的校正步, 它的主要计算是一个 p 阶矩阵的奇异值分解. 这个校正步可以被看作是由第一步得到的试探步的一个旋转. 当在特殊情况 $p = 1$ 和 $G = 0$ 时, 对称性 (2.1) 总是满足, 此时这一步可以省略.

下面我们介绍文 [37] 中提出的非收缩方法的算法框架, 两类具体实现和理论结果.

2.1 非收缩算法框架

注意到问题 (0.1) 的一阶最优性条件 (2.1) 有三个性质: 次稳定性, 对称性和可行性. 受到关系

$$\|\nabla f(X) - X\nabla f(X)^\top X\|_F^2 = \|\nabla f(X) - XX^\top \nabla f(X)\|_F^2 + \|X^\top \nabla f(X) - \nabla f(X)^\top X\|_F^2$$

的启发, 为了使得在切空间中的梯度等于零, 可以采取如下两个步骤.

首先从当前迭代点 X^k 出发, 寻找一个中间点 $\bar{X} \in \mathcal{S}_{n,p}$ 使其满足充分下降性:

$$f(X^k) - f(\bar{X}) \geq C_1 \cdot \|(I_n - X^k X^{k\top})\nabla f(X^k)\|_F^2, \quad (2.2)$$

其中 $C_1 > 0$ 是一个正常数. 不等式 (2.2) 的右边度量了在点 X^k 处, 欧式空间中的投影梯度的 Frobenius 范数的平方.

尽管中间点 $\bar{X} \in \mathcal{S}_{n,p}$ 满足式 (2.2), 但它并不满足式 (2.1) 中的对称性. 于是考虑构造如下校正步使得对称性得以满足:

$$X^{k+1} = \begin{cases} \bar{X}, & \text{如果 } \bar{X}^\top G = G^\top \bar{X}; \\ -\bar{X}UT^\top, & \text{否则.} \end{cases} \quad (2.3)$$

[§]这里, 次稳定性表示目标函数的梯度在 X^\top 的零空间的稳定性.

这里 $\bar{X}^\top G = U\Sigma T^\top$ 是 $\bar{X}^\top G$ 的奇异值分解. 更进一步, 可以证明

$$X^{k+1} \in \mathcal{S}_{n,p}, \quad X^{k+1\top} \nabla f(X^{k+1})$$

是对称的, 并且函数值不增

$$f(\bar{X}) - f(X^{k+1}) \geq \frac{1}{8\theta + 1} \cdot \|\bar{X}^\top \nabla f(\bar{X}) - \nabla f(\bar{X})^\top \bar{X}\|_{\mathbb{F}}^2, \quad (2.4)$$

其中 $\theta := \|G\|_2$.

采用 $c(X) := (I_n - XX^\top)\nabla f(X)$ 的范数来构建终止准则, 完整的算法框架可以描述如下:

算法 1 正交约束优化问题的一阶算法框架

步骤1 终止误差 $\varepsilon > 0$; 初始化: $X^0 \in \mathcal{S}_{n,p}$; 取 $k := 0$

步骤2 当 $\|c(X^k)\|_{\mathbb{F}} > \varepsilon$ 时

步骤3 基于 X^k , 寻找可行点 \bar{X} 满足式 (2.2);

步骤4 基于 \bar{X} , 通过式 (2.3) 计算可行点 X^{k+1} ;

步骤5 取 $k := k + 1$;

步骤6 返回 X^k .

下面两个小节, 我们将介绍两类具体的方法来实现寻找可行点 \bar{X} 满足式 (2.2).

2.2 梯度类方法

首先, 文献 [37] 的作者提出如下发现:

引理 2.1 对于任意的 $Y \in \mathcal{B}_{X,\tau} := \mathcal{B}(X - \tau\nabla f(X), \tau\|\nabla f(X)\|_{\mathbb{F}})$, 其中 $\tau \in (0, \rho^{-1})$, 我们有

$$f(X) - f(Y) \geq \frac{1 - \rho\tau}{2\tau} \cdot \|X - Y\|_{\mathbb{F}}^2. \quad (2.5)$$

可行域、当前迭代点、梯度步和辅助球的关系如右图所示.

基于引理 2.1, 可以构建下面两种算法.

可行试探点的第一种可能的选择是当前迭代点 X^k 沿 $X^k - \tau\nabla f(X^k)$ 的零空间反射, 事实上, 这个点可以通过 Householder 变换计算得到.

$$\text{GR: } \begin{cases} V = X^k - \tau\nabla f(X^k), & \text{取定 } \tau \in (0, \rho^{-1}); \\ \bar{X}_{\text{GR}} = (-I_n + 2V(V^\top V)^\dagger V^\top)X^k. \end{cases} \quad (2.6)$$

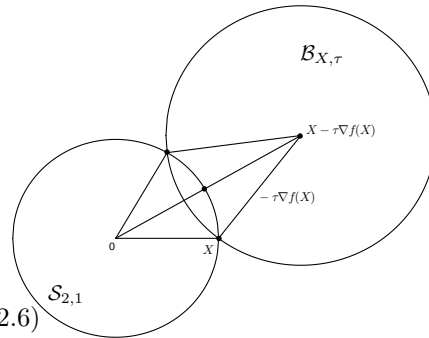


图 1 梯度类方法

因为 \bar{X}_{GR} 是反射点, 称算法 1 的步骤3用式 (2.6) 得到 $\bar{X} := \bar{X}_{\text{GR}}$ 的算法为梯度反射法 (GR). 可以证明由式 (2.6) 定义的中间点 \bar{X}_{GR} 是可行点, 并且满足函数值的充分下降条件 (2.2):

$$f(X^k) - f(\bar{X}_{\text{GR}}) \geq \frac{2(\tau^{-1} - \rho)}{(\tau^{-1} + \rho + \theta)^2} \cdot \|(I_n - X^k X^{k\top}) \nabla f(X^k)\|_{\text{F}}^2. \quad (2.7)$$

另一个可能的可行试探点是直接取 $X^k - \tau \nabla f(X^k)$ 在 Stiefel 流形上的投影点, 也就是

$$\text{GP:} \quad \begin{cases} V = X^k - \tau \nabla f(X^k), & \text{取定 } \tau \in (0, \rho^{-1}); \\ \bar{X}_{\text{GP}} = \mathcal{P}_{\mathcal{S}_{n,p}}(V). \end{cases} \quad (2.8)$$

如果在算法 1 的步骤3选用式 (2.8) 得到 $\bar{X} := \bar{X}_{\text{GP}}$, 那么我们称这个算法为梯度投影法 (GP). 同样可以证明 \bar{X}_{GP} 满足可行性和函数值的充分下降性:

$$f(X^k) - f(\bar{X}_{\text{GP}}) \geq \frac{\tau^{-1} - \rho}{2(\tau^{-1} + \rho + \theta)^2} \cdot \|(I_n - X^k X^{k\top}) \nabla f(X^k)\|_{\text{F}}^2. \quad (2.9)$$

2.3 以列为块的块坐标下降方法

块坐标下降方法 (BCD) 是一类常见的一阶算法. 对于正交约束优化问题, 一个自然的想法是按照变量的列进行划分. 当固定了矩阵 X 的其他 $p-1$ 列, 让第 i 列作为变量, 可以得到如下的子问题:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f_{i,X}(x) \\ \text{s. t.} \quad & \|x\|_2 = 1, \\ & X_i^\top x = 0, \end{aligned} \quad (2.10)$$

其中 $f_{i,X}(x) := f(X_{i,x})$, $X_{i,x}$ 和 X_i 定义在第一节最后.

假设得到了上述子问题的解或者找到了一个可行点 x^+ 使得函数值相比于 $f_{i,X}(X_i)$ 有充分下降, 那么继续用这个可行点以 Gauss-Seidel 的方式按列更新就得到了如下 CBCD 的一个大迭代, 从而得到新的试探步 \bar{X} .

算法 2: 以列为块的块坐标下降法

步骤1 令 $W^0 = X$, $i := 1$;

步骤2 当 $i \leq p$ 时

步骤3 替换 X 为 W^{i-1} , 求解子问题 (2.10), 得到可行点 x^+ 使其满足如下的函数值充分下降性和渐进小的步长保证

$$f_{i,W^{i-1}}(X_i) - f_{i,W^{i-1}}(x^+) \geq k_1 \|X_i - x^+\|_2^2, \quad (2.11)$$

$$\|X_i - x^+\|_2 \geq k_2 \|(I_n - W^{i-1} W^{i-1\top}) \nabla f_{i,W^{i-1}}(X_i)\|_2; \quad (2.12)$$

其中 $k_1, k_2 > 0$ 为任意常数.

步骤4 令 $W^i = W_{i,x^+}^{i-1}$, $i := i + 1$;

步骤5 返回 $\bar{X} = W^p$.

可以证明, 由算法 2 计算得到的 \bar{X} 满足 $\bar{X} \in \mathcal{S}_{n,p}$, 并且满足函数值的充分下降条件 (2.2):

$$f(X) - f(\bar{X}) \geq \frac{k_1 k_2^2}{(1 + (p-1)k_2((1 + \sqrt{2})\rho + \sqrt{2}\theta))^2} \cdot \|(I_n - XX^\top)\nabla f(X)\|_F^2. \quad (2.13)$$

另一方面, 文献 [37] 还给出了 CBCD 子问题 (2.10) 的非精确求解方法.

2.4 计算复杂度与收敛性

这一小节中, 我们比较收缩算法和非收缩算法 GR, GP 和 CBCD 的每步迭代的计算复杂度. 首先约定基本的线性代数运算的计算量如下: 给定 $A \in \mathbb{R}^{n \times n}$, $B_1, B_2 \in \mathbb{R}^{n \times p}$, $S_1, S_2 \in \mathbb{R}^{p \times p}$ 和 $x \in \mathbb{R}^n$, 计算矩阵乘法 $B_1^\top B_2$, $B_1^\top B_1$, $B_1 S_1$ 和 $S_1 S_2$ 分别需要 $2np^2$, $np^2 + np$, $2np^2$ 和 $2p^3$ 次浮点运算. 计算 A^{-1} 和 S^{-1} 分别需要 $8n^3/3$ 和 $8p^3/3$ 次浮点运算, 计算 Ax 需要 $2n^2$ 次浮点运算. 我们采用文献 [13] 中表 1 的设定, 并且不考虑 $\nabla f(X)$ 的计算量, 得到每步迭代的计算复杂度如表 1 所示.

表 1 计算量比较

迭代格式	计算复杂度	
	首个 τ	后续 τ
测地线类方法		
$Y_{\text{geoc}}(\tau; X)$ [18]	$O(n^3)$	$O(n^3)$
$Y_{\text{qgeo}}(\tau; X)$ [17]	$O(n^3)$	$O(n^3)$
$Y_{\text{geoe}}(\tau; X)$ [11]	$10np^2 + 2np + O(p^3)$	$4np^2 + O(p^3)$
$Y_{\text{wy}}(\tau; X)$ [12]	$7np^2 + 2np + O(p^3)$	$4np^2 + np + O(p^3)$
投影类方法		
$Y_{\text{qr}}(\tau; X)$ [25]	$6np^2 + 3np + O(p^3)$	$2np^2 + 2np$
$Y_{\text{pd}}(\tau; X)$ [25]	$7np^2 + 4np + O(p^3)$	$2np^2 + 2np + O(p^3)$
$Y_{\text{pj}}(\tau; X)$ [16]	$7np^2 + 4np + O(p^3)$	$3np^2 + 3np + O(p^3)$
$Y_{\text{jd}}(\tau; X)$ [13]	$7np^2 + 3np + O(p^3)$	$2np^2 + 3np + O(p^3)$
非收缩方法		
GR [37]	$9np^2 + 4np + O(p^3)$	
GP [37]	$7np^2 + 3np + O(p^3)$	
CBCD-GR [37]	$4np^2 + 8np + O(p^3)$	
CBCD-GP [37]	$4np^2 + 5np + O(p^3)$	

表 1 中, “首个 τ ” 和 “后续 τ ” 两列分别表示第一个试探步和之后由线搜索导致的其他试探步所需的计算量. 在算法 GP, GR 和 CBCD 中, 并没有额外的函数值计算, 因为这些方法无需线搜索, 其中 GR 和 GP 在固定步长下全局收敛, CBCD 只需要非精确求解一步. 另一方面, CBCD-GR 和 CBCD-GP 表示算法 2 使用 GR 或者 GP 更新公式. 注意到 $\nabla f_{i,X}((I_n - W_i^{i-1} W_i^{i-1 \top})X_i)$ 的计算可以被省略, 因为 $W_i^{i-1 \top} X_i = 0$, 所以它等于 $\nabla f_{i,X}(X_i)$. 如果 $f_{i,X}(X_i)$ ($i = 1, \dots, p$) 都是二次的, 在子空间 $\text{span}\{X_i, \nabla f_{i,X}(X_i)\}$ 上求解子问题 (2.10) 到全局最优, 相应的计算量是 $12np^2 + 3np + O(p^3)$.

下面给出非收缩算法框架 1 的全局收敛性结论. 首先是函数值的收敛性结果.

引理 2.1 假设初始点是 $X^0 \in \mathcal{S}_{n,p}$, 令 $\{X^k\}$ 是由算法 1 生成的迭代序列, 则 $\{f(X^k)\}$ 收敛.

接着是迭代点列的子序列收敛性.

定理 2.1 假设初始点是 $X^0 \in \mathcal{S}_{n,p}$, 令 $\{X^k\}$ 是由算法 1 生成的迭代序列, 则 $\{X^k\}$ 存在收敛子序列. 进一步, 每个 $\{X^k\}$ 的聚点 X^* 都满足问题 (0.1) 的一阶最优性条件.

引理 2.1 和定理 2.1 保证了 $f(X)$ 在 $\{X^k\}$ 的聚点集合上是常数, 我们定义这个常数为 f^* , 并且定义

$$\Omega_{\text{FON}}^{f^*} := \Omega_{\text{FON}} \cap \{X \mid f(X) = f^*\}, \quad (2.14)$$

其中 Ω_{FON} 是包含所有一阶稳定点的集合. 可以证明 X^k 和 $\Omega_{\text{FON}}^{f^*}$ 的距离收敛到 0.

推论 2.1 假设初始点是 $X^0 \in \mathcal{S}_{n,p}$, 令 $\{X^k\}$ 是由算法 1 生成的迭代序列, 则我们有

$$f(X^k) \geq f^*, \quad \forall k = 1, \dots \quad (2.15)$$

并且

$$\lim_{k \rightarrow \infty} \text{dist}(X^k, \Omega_{\text{FON}}^{f^*}) = 0. \quad (2.16)$$

3 不可行方法及总结与展望

前面两节介绍的正交约束优化问题的收缩方法与非收缩可行方法, 这两类方法都属于可行方法, 也即每一个迭代点都必须满足正交约束. 所以这些方法的迭代过程中都隐含了正交化. 当变量列数增加, 也即 p 比较大时, 正交化由于可扩展性低, 成为计算的瓶颈. 这个时候, 就需要设计适当的不可行方法.

文献[38], 将目标函数中的变量与正交约束的变量分裂, 通过等式约束联系在一起, 并设计增广拉格朗日交替方向算法 (ADMM) 来求解. 由于迭代过程中两组分量值不同, 因此这个方法属于不可行方法, 但是关于正交约束分量的子问题仍需计算投影, 因此这个方法仍旧不能解决瓶颈难题. 在文献 [39] 中, 作者对线性特征值问题的 Rayleigh-Ritz 极小化模型提出一种二次罚函数模型, 并证明在求特征子空间的意义上, 当罚参数足够大, 该罚函数就是精确罚函数, 所以只需求解无约束的罚函数模型, 就可以得到原来正交约束问题的最优解. 遗憾的是这类方法利用了线性特征值问题的特性, 不适合推广到一般正交约束优化问题中去. 因此构建有效的正交约束优化问题, 是一个非常值得研究的课题.

另一方面, 张量特征值问题、张量分解问题最近也引起了国内外专家的广泛关注, 其在数据处理等问题中有着广泛的应用. 而张量的特征值问题与张量分解问题也和正交约束优化问题有着密切的关系. 所以与之相关的模型构建与算法设计也是一个值得关注的课题.

参考文献

- [1] Yang C, Meza J C, Wang L W. A constrained optimization algorithm for total energy minimization in electronic structure calculations [J]. *Journal of Computational Physics*, 2006, **217**(2): 709-721.
- [2] Yang C, Meza J C, Wang L W. A trust region direct constrained minimization algorithm for the Kohn - Sham equation [J]. *SIAM Journal on Scientific Computing*, 2007, **29**(5): 1854-1875.
- [3] Yang C, Meza J C, Lee B, et al. KSSOLV—a MATLAB toolbox for solving the Kohn-Sham equations [J]. *ACM Transactions on Mathematical Software (TOMS)*, 2009, **36**(2): 10.

-
- [4] Knyazev A V. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method [J]. *SIAM Journal on Scientific Computing*, 2001, **23**(2): 517-541.
- [5] Liu X, Wen Z, Zhang Y. Limited memory block Krylov subspace optimization for computing dominant singular value decompositions [J]. *SIAM Journal on Scientific Computing*, 2013, **35**(3): A1641-A1668.
- [6] Grubišić I, Pietersz R. Efficient rank reduction of correlation matrices [J]. *Linear Algebra and Its Applications*, 2007, **422**(2-3): 629-653.
- [7] Schönemann P H. A generalized solution of the orthogonal Procrustes problem [J]. *Psychometrika*, 1966, **31**(1): 1-10.
- [8] Eldén L, Park H. A procrustes problem on the Stiefel manifold [J]. *Numerische Mathematik*, 1999, **82**(4): 599-619.
- [9] d'Aspremont A, Ghaoui L E, Jordan M I, et al. A direct formulation for sparse PCA using semidefinite programming [C]// *Advances in neural information processing systems*. 2005, 41-48.
- [10] Zou H, Hastie T, Tibshirani R. Sparse principal component analysis [J]. *Journal of Computational and Graphical Statistics*, 2006, **15**(2): 265-286.
- [11] Edelman A, Arias T A, Smith S T. The geometry of algorithms with orthogonality constraints [J]. *SIAM Journal on Matrix Analysis and Applications*, 1998, **20**(2): 303-353.
- [12] Wen Z, Yin W. A feasible method for optimization with orthogonality constraints [J]. *Mathematical Programming*, 2013, **142**(1-2): 397-434.
- [13] Jiang B, Dai Y H. A framework of constraint preserving update schemes for optimization on Stiefel manifold [J]. *Mathematical Programming*, 2015, **153**(2): 535-575.
- [14] Hu J, Jiang B, Liu X, et al. A note on semidefinite programming relaxations for polynomial optimization over a single sphere [J]. *Science China Mathematics*, 2016, **59**(8): 1543-1560.
- [15] Stiefel E. Richtungsfelder und Fernparallelismus in n -dimensionalen Mannigfaltigkeiten [J]. *Commentarii Mathematici Helvetici*, 1935, **8**(1): 305-353.
- [16] Manton J H. Optimization algorithms exploiting unitary constraints [J]. *IEEE Transactions on Signal Processing*, 2002, **50**(3): 635-650.
- [17] Nishimori Y, Akaho S. Learning algorithms utilizing quasi-geodesic flows on the Stiefel manifold [J]. *Neurocomputing*, 2005, **67**: 106-135.
- [18] Abrudan T E, Eriksson J, Koivunen V. Steepest descent algorithms for optimization under unitary matrix constraint [J]. *IEEE Transactions on Signal Processing*, 2008, **56**(3): 1134-1147.
- [19] Abrudan T, Eriksson J, Koivunen V. Conjugate gradient algorithm for optimization under unitary matrix constraint [J]. *Signal Processing*, 2009, **89**(9): 1704-1714.
- [20] Absil P A, Baker C G, Gallivan K A. Trust-region methods on Riemannian manifolds [J]. *Foundations of Computational Mathematics*, 2007, **7**(3): 303-330.
- [21] Absil P A, Baker C G, Gallivan K A. Convergence analysis of Riemannian trust-region methods [EB/OL]. [2017-08-26]. http://www.optimizationonline.org/DB_HTML/2006/06/1416.html, 2006.
- [22] Savas B, Lim L H. Quasi-Newton methods on Grassmannians and multilinear approximations of tensors [J]. *SIAM Journal on Scientific Computing*, 2010, **32**(6): 3352-3393.
- [23] Huang W, Gallivan K A, Absil P A. A Broyden class of quasi-Newton methods for Riemannian optimization [J]. *SIAM Journal on Optimization*, 2015, **25**(3): 1660-1685.
- [24] Huang W, Absil P A, Gallivan K A. A Riemannian BFGS method for nonconvex optimization problems [C]// *Numerical Mathematics and Advanced Applications ENUMATH 2015*. New York: Springer, 2016, 627-634.
- [25] Absil P A, Mahony R, Sepulchre R. *Optimization Algorithms on Matrix Manifolds* [M]. Princeton: Princeton University Press, 2009.
- [26] Absil P A, Mallick J. Projection-like retractions on matrix manifolds [J]. *SIAM Journal on Optimization*, 2012, **22**(1): 135-158.
- [27] Nocedal J, Wright S J. *Sequential Quadratic Programming* [M]. New York: Springer, 2006.

- [28] Sun W, Yuan Y X. *Optimization Theory and Methods: Nonlinear Programming* [M]. New York: Springer, 2006.
- [29] Goldfarb D, Wen Z, Yin W. A curvilinear search method for p-harmonic flows on spheres [J]. *SIAM Journal on Imaging Sciences*, 2009, **2**(1): 84-109.
- [30] Barzilai J, Borwein J M. Two-point step size gradient methods [J]. *IMA Journal of Numerical Analysis*, 1988, **8**(1): 141-148.
- [31] Zhang H, Hager W W. A nonmonotone line search technique and its application to unconstrained optimization [J]. *SIAM Journal on Optimization*, 2004, **14**(4): 1043-1056.
- [32] Jiang B, Ma S, So A M C, et al. Vector transport-free SVRG with general retraction for Riemannian optimization: Complexity analysis and practical implementation [J]. arXiv: 1705.09059, 2017.
- [33] Johnson R, Zhang T. Accelerating stochastic gradient descent using predictive variance reduction [C]//*Advances in neural information processing systems*. 2013, 315-323.
- [34] Hu J, Milzarek A, Wen Z, et al. Adaptive Regularized Newton Method for Riemannian Optimization [J]. arXiv: 1708.02016, 2017.
- [35] Li Y, Wen Z, Yang C, et al. A semi-smooth Newton method for solving semidefinite programs in electronic structure calculations [J]. arXiv: 1708.08048, 2017
- [36] Boumal N, Mishra B, Absil P A, et al. Manopt, a matlab toolbox for optimization on manifolds [J]. *Journal of Machine Learning Research*, 2014, **15**(1): 1455-1459.
- [37] Gao B, Liu X, Chen X, et al. A new first-order framework for orthogonal constrained optimization problems [J]. *Optimization*, 2016.
- [38] Lai R, Osher S. A splitting method for orthogonality constrained problems [J]. *Journal of Scientific Computing*, 2014, **58**(2): 431-449.
- [39] Wen Z, Yang C, Liu X, et al. Trace-penalty minimization for large-scale eigenspace computation [J]. *Journal of Scientific Computing*, 2016, **66**(3): 1175-1203.
- [40] Fraikin C, Nesterov Y, Van Dooren P. A gradient-type algorithm optimizing the coupling between matrices [J]. *Linear Algebra and its Applications*, 2008, **429**(5-6): 1229-1242.
- [41] Golub G H, Van Loan C F. *Matrix Computations* [M]. Baltimore: Johns Hopkins University Press, 2012.
- [42] Jiang B, Cui C, Dai Y H. Unconstrained optimization models for computing several extreme eigenpairs of real symmetric matrices [J]. *Pacific Journal of Optimization*, 2014, **10**(1): 55-71.
- [43] Knyazev A V. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method [J]. *SIAM Journal on Scientific Computing*, 2001, **23**(2): 517-541.
- [44] Kohn W, Sham L J. Self-consistent equations including exchange and correlation effects [J]. *Physical Review*, 1965, **140**(4A): A1133.
- [45] Liu X, Wang X, Wen Z, et al. On the convergence of the self-consistent field iteration in Kohn-Sham density functional theory [J]. *SIAM Journal on Matrix Analysis and Applications*, 2014, **35**(2): 546-558.
- [46] Liu X, Wen Z, Wang X, et al. On the analysis of the discretized Kohn-Sham Density functional theory [J]. *SIAM Journal on Numerical Analysis*, 2015, **53**(4): 1758-1785.
- [47] Liu X, Wen Z, Zhang Y. Limited memory block Krylov subspace optimization for computing dominant singular value decompositions [J]. *SIAM Journal on Scientific Computing*, 2013, **35**(3): A1641-A1668.
- [48] Liu X, Wen Z, Zhang Y. An efficient Gauss-Newton algorithm for symmetric low-rank product matrix approximations [J]. *SIAM Journal on Optimization*, 2015, **25**(3): 1571-1608.
- [49] Ulbrich M, Wen Z, Yang C, et al. A proximal gradient method for ensemble density functional theory [J]. *SIAM Journal on Scientific Computing*, 2015, **37**(4): A1975-A2002.
- [49] Zhang X, Zhu J, Wen Z, et al. Gradient type optimization methods for electronic structure calculations [J]. *SIAM Journal on Scientific Computing*, 2014, **36**(3): C265-C289.
- [50] 袁亚湘, 孙文瑜. 最优化理论与方法[M]. 北京: 科学出版社, 1997.
- [51] 袁亚湘. 非线性优化计算方法[M]. 北京: 科学出版社, 2008.