第49卷　第2期　　　吉林大学学报(理学版)　　　Vol.49　No.2
2011年3月　　　Journal of Jilin University (Science Edition)　　　Mar　2011

求解最小体积轴向椭球问题的线性收敛算法

丛伟杰，刘红卫
(西安电子科技大学 理学院，西安 710071)

摘要：通过定义求解最小体积轴向椭球问题的两个近似最优性条件，计算满足第二个近似最优性条件的一个新的近似解，给出一种求解最小体积轴向椭球问题的近似算法，并证明了算法具有线性收敛性. 实验结果证实了算法的有效性.

关键词：最小体积轴向椭球；最优性条件；近似算法；线性收敛
中图分类号：O221.2　　**文献标志码**：A　　**文章编号**：1671-5489(2011)02-0173-06

Linearly Convergent Algorithm for Solving the Minimum Volume Axis-Aligned Ellipsoid Problem

CONG Wei-jie, LIU Hong-wei
(*School of Science, Xidian University, Xi'an 710071, China*)

Abstract: Firstly, two approximate optimality conditions of the minimum volume axis-aligned ellipsoid (MVAE) problem were defined. Secondly, a new approximate solution satisfying the second approximate optimality condition was computed. Furthermore, an approximation algorithm for the MVAE problem was presented, which has the linear convergence. The numerical results show the efficiency of the algorithm.

Key words: minimum volume axis-aligned ellipsoid; optimality conditions; approximation algorithm; linear convergence

0　引　言

一个轴向椭球 $E_{\boldsymbol{c},\boldsymbol{D}}\subset\mathbb{R}^n$ 的表达式[1]如下：

$$E_{\boldsymbol{c},\boldsymbol{D}}:=\{\boldsymbol{x}\in\mathbb{R}^n:(\boldsymbol{x}-\boldsymbol{c})^{\mathrm{T}}\boldsymbol{D}(\boldsymbol{x}-\boldsymbol{c})\leqslant 1\}=\Big\{\boldsymbol{x}\in\mathbb{R}^n:\sum_{j=1}^{n}[\sqrt{d_j}(x_j-c_j)]^2\leqslant 1\Big\},\tag{1}$$

其中 $\boldsymbol{c}\in\mathbb{R}^n$ 为其中心，形状由正定对角矩阵 $\boldsymbol{D}=\mathrm{diag}(d_1,d_2,\cdots,d_n)\in\mathbb{R}^{n\times n}$ 确定. 由于它的每个轴长分别为 $1/\sqrt{d_j}(j=1,2,\cdots,n)$，故其体积可表示为

$$\mathrm{Vol}(E_{\boldsymbol{c},\boldsymbol{D}})=\eta_n\det\boldsymbol{D}^{-1/2}=\eta_n\prod_{j=1}^{n}\frac{1}{\sqrt{d_j}},\tag{2}$$

其中 η_n 是 $\mathbb{R}^n$ 中单位球的体积.

给定点集 $S=\{x^1,x^2,\cdots,x^m\}\subset\mathbb{R}^n$，最小体积轴向椭球(minimum volume axis-aligned ellipsoid，简记为 MVAE)[1-2]问题就是寻找一个体积最小的轴向椭球使其能包含 S 中的 m 个点. 点集 S 的最小体积轴向椭球记为 MVAE(S). MVAE 问题在计算机图形学、碰撞检测和机器学习等领域[3-4]应用广泛. 目前，

收稿日期：2010-06-02.

作者简介：丛伟杰(1981—)，男，汉族，博士研究生，从事最优化理论与算法的研究，E-mail：cong24518@163.com. 通讯作者：刘红卫(1967—)，男，汉族，教授，博士生导师，从事最优化方法及机器学习中优化问题的研究，E-mail：hwliu@mail.xidian.edu.cn.

基金项目：国家自然科学基金(批准号：60703118；61072144)和中央高校基本科研业务费资助项目(批准号：JY10000970004).

对于最小闭包球(minimum enclosing ball, 简记为 MEB)[5]和最小体积闭包椭球(minimum volume enclosing ellipsoid, 简记为 MVEE)[6-7]问题的研究已取得许多成果. 而 MVAE 问题是 MEB 问题和 MVEE 问题的中间状态, 具有特别的意义, 目前对 MVAE 问题的研究尚处于初级阶段. 因此, 寻找和设计有效的求解最小体积轴向椭球算法十分必要.

文献[1]在给定 $\varepsilon>0$ 的情况下, 计算了 MVAE(S)的一个$(1+\varepsilon)$-近似 E, 定义如下:

$$S\subseteq E,\qquad \mathrm{Vol}(E)\leqslant(1+\varepsilon)\mathrm{Vol}(\mathrm{MVAE}(S)).\tag{3}$$

此外, 还得到了 S 的一个子集 $X\subseteq S$, 称为核心集, 即如果存在一个轴向椭球 $E_{c,D}\subset\mathbb{R}^n$, 则 $S\subseteq E_{c,D}$, 并且 $E_{c,D}$是 MVAE(X)的一个$(1+\varepsilon)$-近似. 文献[1]将求解 MVEE 问题的算法[6]应用于求解 MVAE 问题, 但线性搜索步没有显式最优解, 从而用一个显式近似解代替, 得到了算法的计算复杂度为 $O(mn^5/\varepsilon)$, 核心集大小为 $O(n^4/\varepsilon)$, 对于 $\varepsilon\in(0,1)$. 文献[2]使用切线方程的方法得到了另一个显式近似解来改进算法, 得到改进的计算复杂度为$O(mn^3/\varepsilon)$, 核心集大小为 $O(n^2/\varepsilon)$, 对于$\varepsilon\in(0,1)$, 结果与 MVEE 问题的计算复杂度和核心集结果一致.

本文定义了求解 MVAE 问题的两个近似最优性条件, 并计算了线性搜索步的一个新近似解(满足所定义的强近似最优性条件). 将这个新的近似解结合文献[2]中的算法 3.1, 得到求解 MVAE 问题的一个$(1+\varepsilon)$-近似的新算法, 该算法仍可获得与求解 MVEE 问题一致的计算复杂度和核心集大小. 此外, 证明了该算法是线性收敛的.

1 优化公式及近似最优性条件

由文献[1]可得, 计算 MVAE(S)的问题可以转化为如下凸优化问题:

$$\min_{\boldsymbol{\gamma},\boldsymbol{\mu}}\ -\sum_{j=1}^n\log\gamma_j:\sum_{j=1}^n(\gamma_jx_j^i-\mu_j)^2\leqslant1,\qquad i=1,2,\cdots,m,\tag{4}$$

其中 $\boldsymbol{\gamma}=(\gamma_1,\gamma_2,\cdots,\gamma_n)^{\mathrm{T}}\in\mathbb{R}^n$ 和$\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_n)^{\mathrm{T}}\in\mathbb{R}^n$ 是原始变量. 式(4)的 Lagrangian 对偶问题为

$$\max_{\boldsymbol{\sigma}}\ \frac{n}{2}\log n+\frac{1}{2}\sum_{j=1}^n\log(u_j(\boldsymbol{\sigma})-v_j^2(\boldsymbol{\sigma})):\sum_{i=1}^m\sigma_i=1,\qquad\boldsymbol{\sigma}\geqslant0,\tag{5}$$

其中: $\boldsymbol{\sigma}=(\sigma_1,\sigma_2,\cdots,\sigma_m)^{\mathrm{T}}\in\mathbb{R}^m$ 是对偶变量; $u_j(\boldsymbol{\sigma}):=\sum_{i=1}^m\sigma_i(x_j^i)^2$; $v_j(\boldsymbol{\sigma}):=\sum_{i=1}^m\sigma_ix_j^i$; $j=1,2,\cdots,n$.

由文献[1]中引理 2.1 可知, 如果$(\boldsymbol{\gamma}^*,\boldsymbol{\mu}^*)$和 $\boldsymbol{\sigma}^*$ 分别表示问题(4)和(5)的最优解, 则 MVAE(S)可表示为

$$\mathrm{MVAE}(S)=\{\boldsymbol{x}\in\mathbb{R}^n:(\boldsymbol{x}-\boldsymbol{c}^*)^{\mathrm{T}}\boldsymbol{D}^*(\boldsymbol{x}-\boldsymbol{c}^*)\leqslant1\}=\Big\{\boldsymbol{x}\in\mathbb{R}^n:\sum_{j=1}^n\big[\sqrt{d_j^*}(x_j-c_j^*)\big]^2\leqslant1\Big\},$$

其中: $c_j^*:=\mu_j^*/\gamma_j^*=v_j(\boldsymbol{\sigma}^*)$; $d_j^*:=(\gamma_j^*)^2=1/[n(u_j(\boldsymbol{\sigma}^*)-v_j^2(\boldsymbol{\sigma}^*))]$; $j=1,2,\cdots,n$. 因此, MVAE(S)问题可以简单地通过计算对偶问题(5)求解.

由问题的 KKT 最优性条件容易得到互补条件为 $\sigma_i^*\Big(1-\sum_{j=1}^nw_j^i(\boldsymbol{\sigma}^*)\Big)=0(i=1,2,\cdots,m)$, 其中: $w_j^i(\boldsymbol{\sigma}):=\dfrac{(x_j^i-v_j(\boldsymbol{\sigma}))^2}{n(u_j(\boldsymbol{\sigma})-v_j^2(\boldsymbol{\sigma}))}$; $i=1,2,\cdots,m$; $j=1,2,\cdots,n$. 进而互补条件可以改写为

$$\sigma_i^*>0\Rightarrow\sum_{j=1}^nw_j^i(\boldsymbol{\sigma}^*)=1,\qquad i=1,2,\cdots,m.\tag{6}$$

下面类似于 MVEE 问题的近似最优性条件[7-8], 给出 MVAE 问题的两个近似最优性条件的定义.

定义 1 给定 $\eta>0$, 若

$$\sum_{j=1}^nw_j^i(\boldsymbol{\sigma})\leqslant1+\eta,\qquad i=1,2,\cdots,m,\tag{7}$$

则称问题(5)的可行解 $\boldsymbol{\sigma}$ 满足弱 η-近似最优性条件.

定义 2 给定 $\eta\in(0,1)$, 若式(7)成立, 并且有

$$\sigma_i > 0 \Rightarrow \sum_{j=1}^{n} w_j^i(\boldsymbol{\sigma}) \geqslant 1-\eta, \qquad i=1,2,\cdots,m, \tag{8}$$

则称问题(5)的可行解 $\boldsymbol{\sigma}$ 满足强 η-近似最优性条件.

由定义 2 可知，当 $\sigma_i>0$ 时，有 $1-\eta \leqslant \sum_{j=1}^{n} w_j^i(\boldsymbol{\sigma}) \leqslant 1+\eta(i=1,2,\cdots,m)$，表明对于较小的 η，定义 2 是比定义 1 对式(6)更好的近似.

2　算法及收敛性分析

下面给出求解 MVAE(S)问题的近似算法.

算法 1　给定 $S=\{x^1,x^2,\cdots,x^m\}\subset\mathbb{R}^n$，$\eta>0$. 置 $k=0$.

1）执行初始体积近似算法 3.1[1]得到初始核心集 X_0，令对偶问题(5)的初始可行解 $\sigma^0\in\mathbb{R}^m$ 满足：

$$\sigma_i^0=\begin{cases}1/|X_0|, & x^i\in X_0;\\ 0, & \text{否则}.\end{cases}$$

2）令每次迭代构造的轴向椭球为 $E_k=\left\{\boldsymbol{x}\in\mathbb{R}^n:\ \sum_{j=1}^{n}\frac{(x_j-v_j(\boldsymbol{\sigma}^k))^2}{n(u_j(\boldsymbol{\sigma}^k)-v_j^2(\boldsymbol{\sigma}^k))}\leqslant 1\right\}$，并令 $w_j^i(\boldsymbol{\sigma}^k)=\frac{(x_j^i-v_j(\boldsymbol{\sigma}^k))^2}{n(u_j(\boldsymbol{\sigma}^k)-v_j^2(\boldsymbol{\sigma}^k))}$，$i=1,2,\cdots,m$，$j=1,2,\cdots,n$. 计算指标

$$i_+=\arg\max\Big\{\sum_{j=1}^{n}w_j^i(\boldsymbol{\sigma}^k):\ i=1,2,\cdots,m\Big\},\quad i_-=\arg\min\Big\{\sum_{j=1}^{n}w_j^i(\boldsymbol{\sigma}^k):\ i=1,2,\cdots,m,\ \sigma_i^k>0\Big\},$$

并令 $\varepsilon_+=\sum_{j=1}^{n}w_j^{i_+}(\boldsymbol{\sigma}^k)-1$，$\varepsilon_-=1-\sum_{j=1}^{n}w_j^{i_-}(\boldsymbol{\sigma}^k)$，$\varepsilon_k=\max\{\varepsilon_+,\varepsilon_-\}$.

3）若 $\varepsilon_k\leqslant\eta$，则算法终止，输出 $\boldsymbol{\sigma}^k$ 和核心集 $X_k=\{\boldsymbol{x}^i\in S:\ \sigma_i^k>0\}$；否则，转 4).

4）若 $\varepsilon_k=\varepsilon_+$，则令 $w_{j_+}^{i_+}(\boldsymbol{\sigma}^k)=\max\{w_j^{i_+}(\boldsymbol{\sigma}^k):j=1,2,\cdots,n\}$，$\tilde{\beta}_k=\frac{\varepsilon_k}{1+\varepsilon_k+nw_{j_+}^{i_+}(\boldsymbol{\sigma}^k)}$，更新可行解为 $\boldsymbol{\sigma}^{k+1}=(1-\tilde{\beta}_k)\boldsymbol{\sigma}^k+\tilde{\beta}_ke^{i_+}$，置 $k=k+1$，转 2)；否则，令 $w_{j_+}^{i_-}(\boldsymbol{\sigma}^k)=\max\{w_j^{i_-}(\boldsymbol{\sigma}^k):j=1,2,\cdots,n\}$，$\tilde{\beta}_k=\min\left\{\frac{\varepsilon_k}{1-\varepsilon_k+nw_{j_+}^{i_-}(\boldsymbol{\sigma}^k)},\frac{\sigma_{i_-}^k}{1-\sigma_{i_-}^k}\right\}$，更新可行解为 $\boldsymbol{\sigma}^{k+1}=(1+\tilde{\beta}_k)\boldsymbol{\sigma}^k-\tilde{\beta}_ke^{i_-}$，置 $k=k+1$，转 2).

2.1　算法复杂度分析

算法 1 的主要部分与文献[2]中算法 3.1 类似，区别在于第三步中使用的终止条件为 $\varepsilon_k\leqslant\eta$，第四步中当 $\varepsilon_k=\varepsilon_+$ 时选取线性搜索步的一个新的可行解 $\tilde{\beta}_k=\varepsilon_k/[1+\varepsilon_k+nw_{j_+}^{i_+}(\boldsymbol{\sigma}^k)]$. 下面先分别对这两点进行分析说明，然后给出计算复杂度和核心集结果.

在第 k 次迭代，由第二步可得

$$\sum_{j=1}^{n}w_j^{i_+}(\boldsymbol{\sigma}^k)=1+\varepsilon_+\leqslant 1+\varepsilon_k,\qquad \sum_{j=1}^{n}w_j^{i_-}(\boldsymbol{\sigma}^k)=1-\varepsilon_-\geqslant 1-\varepsilon_k. \tag{9}$$

因此，每次迭代算法 1 总能得到式(5)的一个近似可行解 $\boldsymbol{\sigma}^k$ 满足强 ε_k-近似最优性条件，并且当算法终止($\varepsilon_k\leqslant\eta$)时，得到的近似可行解 $\boldsymbol{\sigma}^k$ 满足强 η-近似最优性条件(7)和(8). 若令

$$\eta:=(1+\varepsilon)^{2/n}-1, \tag{10}$$

则由文献[2]中算法 3.1 可知，本文算法 1 终止时，它得到 MVAE(S)的一个$(1+\varepsilon)$-近似为 $\sqrt{1+\varepsilon_k}E_k$. 此外，由式(10)可知，实际上文献[1]中算法 3.1 和文献[2]中算法 2.1 终止时，仅得到式(5)的一个近似可行解 $\boldsymbol{\sigma}^k$ 满足弱 η-似最优性条件(7). 所以，实际上本文算法 1 能获得更精确的近似解.

令式(5)的目标函数为 $g(\boldsymbol{\sigma}):=\frac{n}{2}\log n+\frac{1}{2}\sum_{j=1}^{n}\log(u_j(\boldsymbol{\sigma})-v_j^2(\boldsymbol{\sigma}))$. 由文献[1]的式(17)可得 $g((1-\beta)\boldsymbol{\sigma}^k+\beta e^{i_+})=g(\boldsymbol{\sigma}^k)+\Delta_k(\beta)$，其中

$$\Delta_k(\beta) = \frac{n}{2}\log(1-\beta) + \frac{1}{2}\sum_{j=1}^{n}\log(1+\beta n w_j^{i_+}(\boldsymbol{\sigma}^k)), \qquad k=0,1,\cdots. \tag{11}$$

由文献[1]中引理4.3可知, 函数$\Delta_k(\beta)$取得最大值时有唯一的最优解$\beta_k^*\in[0,1)$, 但没有显式表达式. 因此, 文献[1]给出了最优解的一个下界$\beta_k=\varepsilon_k/[(n+1)(1+\varepsilon_k)]$代替$\beta_k^*$. 下面的引理给出$\beta_k^*$的一个更好的下界近似$\tilde{\beta}_k$, 也即算法1第四步中当$\varepsilon_k=\varepsilon_+$时选取的$\tilde{\beta}_k$.

引理1 设$\beta_k^*\in[0,1)$是函数$\Delta_k(\beta)$取得最大值时唯一的最优解, 则有

$$\beta_k^* \geqslant \tilde{\beta}_k := \frac{\varepsilon_k}{1+\varepsilon_k+n w_{j_+}^{i_+}(\boldsymbol{\sigma}^k)} \geqslant \beta_k, \qquad k=0,1,\cdots, \tag{12}$$

其中, ε_k和$w_{j_+}^{i_+}(\boldsymbol{\sigma}^k)$定义于算法1.

证明: 由算法1中ε_k和$w_{j_+}^{i_+}(\boldsymbol{\sigma}^k)$的定义, 得

$$w_j^{i_+}(\boldsymbol{\sigma}^k) \leqslant w_{j_+}^{i_+}(\boldsymbol{\sigma}^k) \leqslant \sum_{j=1}^{n} w_j^{i_+}(\boldsymbol{\sigma}^k) = 1+\varepsilon_k, \qquad j=1,2,\cdots,n. \tag{13}$$

结合式(11)可得

$$\Delta_k'(\beta) = -\frac{n}{2(1-\beta)} + \frac{1}{2}\sum_{j=1}^{n}\frac{n w_j^{i_+}(\boldsymbol{\sigma}^k)}{1+\beta n w_j^{i_+}(\boldsymbol{\sigma}^k)} \geqslant -\frac{n}{2(1-\beta)} + \frac{1}{2}\sum_{j=1}^{n}\frac{n w_j^{i_+}(\boldsymbol{\sigma}^k)}{1+\beta n w_{j_+}^{i_+}(\boldsymbol{\sigma}^k)} = -\frac{n}{2(1-\beta)} + \frac{n(1+\varepsilon_k)}{2(1+\beta n w_{j_+}^{i_+}(\boldsymbol{\sigma}^k))}.$$

从而, 可得

$$\Delta_k'(\tilde{\beta}_k) \geqslant -\frac{n}{2(1-\tilde{\beta}_k)} + \frac{n(1+\varepsilon_k)}{2(1+\tilde{\beta}_k n w_{j_+}^{i_+}(\boldsymbol{\sigma}^k))} = 0 = \Delta_k'(\beta_k^*). \tag{14}$$

此外, 由式(11)得$\Delta_k''(\beta) = -\dfrac{n}{2(1-\beta)^2} - \dfrac{1}{2}\sum_{j=1}^{n}\dfrac{n^2(w_j^{i_+}(\boldsymbol{\sigma}^k))^2}{(1+\beta n w_j^{i_+}(\boldsymbol{\sigma}^k))^2} < 0$, 在$[0,1)$内. 所以, 函数$\Delta_k'(\beta)$在$[0,1)$内是严格减少的. 因此, 通过求解式(14)并结合式(13)可得式(12).

下面的引理给出了函数$\Delta_k(\beta)$通过使用引理1中的$\tilde{\beta}_k$获得的函数值下界改进量.

引理2 设$\beta_k^*\in[0,1)$是函数$\Delta_k(\beta)$取得最大值时唯一的最优解, 且$\tilde{\beta}_k$由式(12)给出, 则有

$$\Delta_k(\beta_k^*) \geqslant \Delta_k(\tilde{\beta}_k) \geqslant \begin{cases} \varepsilon_k^2/64, & \varepsilon_k < 1, \\ 1/64, & \text{否则}. \end{cases} \tag{15}$$

证明: 由文献[1]中引理4.4的证明可知, 函数$f(x)=\log(1+x)-x/(1+x)$在$x\geqslant 0$上是严格增加的, 且对于$x\in[0,1)$, 有$f(x)\geqslant(1/8)x^2$, 此外, 有

$$\Delta_k(\beta_k^*) \geqslant \Delta_k(\tilde{\beta}_k) \geqslant \frac{1}{2}\sum_{j=1}^{n}\left(\log(1+\tilde{\beta}_k n w_j^{i_+}(\boldsymbol{\sigma}^k)) - \frac{\tilde{\beta}_k n w_j^{i_+}(\boldsymbol{\sigma}^k)}{1+\tilde{\beta}_k n w_j^{i_+}(\boldsymbol{\sigma}^k)}\right).$$

因此, 若$\tilde{\beta}_k n w_{j_+}^{i_+}(\boldsymbol{\sigma}^k)\geqslant 1$, 其中$w_{j_+}^{i_+}(\boldsymbol{\sigma}^k)=\max\{w_j^{i_+}(\boldsymbol{\sigma}^k): j=1,2,\cdots,n\}$, 则有$\Delta_k(\tilde{\beta}_k)\geqslant\dfrac{1}{2}\left(\log 2-\dfrac{1}{2}\right)$. 否则, 有

$$\Delta_k(\tilde{\beta}_k) \geqslant \frac{1}{2}\sum_{j=1}^{n}\frac{1}{8}(\tilde{\beta}_k n w_j^{i_+}(u^k))^2 = \frac{n^2}{16}\frac{\varepsilon_k^2\sum_{j=1}^{n}(w_j^{i_+}(\boldsymbol{\sigma}^k))^2}{(1+\varepsilon_k+n w_{j_+}^{i_+}(\boldsymbol{\sigma}^k))^2} = \frac{\varepsilon_k^2}{16}\frac{\sum_{j=1}^{n}(w_j^{i_+}(\boldsymbol{\sigma}^k))^2}{((1+\varepsilon_k)/n+w_{j_+}^{i_+}(\boldsymbol{\sigma}^k))^2}.$$

由算法1中ε_k和$w_{j_+}^{i_+}(\boldsymbol{\sigma}^k)$的定义得

$$1+\varepsilon_k = \sum_{j=1}^{n} w_j^{i_+}(\boldsymbol{\sigma}^k) \leqslant n w_{j_+}^{i_+}(\boldsymbol{\sigma}^k), \qquad (w_{j_+}^{i_+}(\boldsymbol{\sigma}^k))^2 \leqslant \sum_{j=1}^{n}(w_j^{i_+}(\boldsymbol{\sigma}^k))^2.$$

于是, 可得

$$\Delta_k(\tilde{\beta}_k) \geqslant \frac{\varepsilon_k^2}{16}\frac{(w_{j_+}^{i_+}(\boldsymbol{\sigma}^k))^2}{(2w_{j_+}^{i_+}(\boldsymbol{\sigma}^k))^2} = \frac{\varepsilon_k^2}{64}.$$

当 $\varepsilon_k < 1$ 时，$\tilde{\beta}_k n w_{j_+}^{i_+}(u^k) = \varepsilon_k n w_{j_+}^{i_+}(u^k)/(1+\varepsilon_k + n w_{j_+}^{i_+}(u^k)) < 1$. 因此，当 $\varepsilon_k < 1$ 时，有 $\Delta_k(\tilde{\beta}_k) \geqslant \varepsilon_k^2/64$；否则，有 $\Delta_k(\tilde{\beta}_k) \geqslant \min\{(1/2)(\log 2 - 1/2), 1/64\} = 1/64$，从而得到式(15).

在算法 1 中应用式(10)，并结合文献[2]中算法 3.1 的分析，容易得到本文算法 1 计算 MVAE(S)的一个$(1+\varepsilon)$-近似的计算复杂度和核心集结果如下：

定理 1　给定 $\varepsilon \in (0,1)$，算法 1 计算 MVAE(S)的一个$(1+\varepsilon)$-近似需要 $O(mn^3/\varepsilon)$ 次基本运算，并且得到一个大小为 $|X_k| = O(n^2/\varepsilon)$ 的核心集.

2.2　算法的收敛性分析

下面运用文献[8]的方法得到算法 1 的线性收敛性. 首先，给出式(4)的一个扰动：

$$\min_{\gamma,\mu} -\sum_{j=1}^n \log \gamma_j : \sum_{j=1}^n (\gamma_j x_j^i - \mu_j)^2 \leqslant 1 + z_i(\boldsymbol{\sigma}^k, \varepsilon_k), \qquad i = 1,2,\cdots,m, \tag{16}$$

其中，$\boldsymbol{\sigma}^k \in \mathbb{R}^m$ 为式(5)的可行解，且满足强 ε_k-近似最优性条件. 设 $\gamma_j^2(\boldsymbol{\sigma}) = 1/[n(u_j(\boldsymbol{\sigma}) - v_j^2(\boldsymbol{\sigma}))]$，且 $c_j(\boldsymbol{\sigma}) = \mu_j(\boldsymbol{\sigma})/\gamma_j(\boldsymbol{\sigma}) = v_j(\boldsymbol{\sigma})$，$j = 1,2,\cdots,n$，其中 $u_j(\boldsymbol{\sigma}) = \sum_{i=1}^m \sigma_i (x_j^i)^2$，$v_j(\boldsymbol{\sigma}) = \sum_{i=1}^m \sigma_i x_j^i$，$j = 1,2,\cdots,n$. 定义 $\boldsymbol{z}^k := \boldsymbol{z}(\boldsymbol{\sigma}^k, \varepsilon_k) \in \mathbb{R}^m$ 如下：

$$z_i^k := \begin{cases} \varepsilon_k, & \sigma_i^k = 0, \\ \sum_{j=1}^n \gamma_j^2(\boldsymbol{\sigma}^k)(x_j^i - c_j(\boldsymbol{\sigma}^k))^2 - 1, & \text{否则}. \end{cases}$$

由定义 2 可得

$$\sigma_i^k > 0 \Rightarrow \sum_{j=1}^n w_j^i(\boldsymbol{\sigma}^k) - 1 = \sum_{j=1}^n \frac{(x_j^i - v_j(\boldsymbol{\sigma}^k))^2}{n(u_j(\boldsymbol{\sigma}^k) - v_j^2(\boldsymbol{\sigma}^k))} - 1 =$$
$$\sum_{j=1}^n \gamma_j^2(\boldsymbol{\sigma}^k)(x_j^i - c_j(\boldsymbol{\sigma}^k))^2 - 1 \geqslant -\varepsilon_k, \quad i = 1,2,\cdots,m,$$

表明 $|z_i^k| \leqslant \varepsilon_k (i = 1,2,\cdots,m)$. 但对于满足弱 ε_k-近似最优性条件的 $\boldsymbol{\sigma}^k$，该不等式并不成立. 此外，可得

$$(\boldsymbol{\sigma}^k)^{\mathrm{T}} \boldsymbol{z}^k = \sum_{i:\sigma_i^k > 0} \sigma_i^k z_i^k = \sum_{j=1}^n \gamma_j^2(\boldsymbol{\sigma}^k)(u_j(\boldsymbol{\sigma}^k) - 2v_j(\boldsymbol{\sigma}^k) c_j(\boldsymbol{\sigma}^k) + c_j^2(\boldsymbol{\sigma}^k)) - 1 =$$
$$\sum_{j=1}^n \gamma_j^2(\boldsymbol{\sigma}^k)\left(\frac{1}{n\gamma_j^2(\boldsymbol{\sigma}^k)}\right) - 1 = 0. \tag{17}$$

引理 3　设 $\boldsymbol{\sigma}^k$ 满足强 ε_k-近似最优性条件，则$(\boldsymbol{\gamma}(\boldsymbol{\sigma}^k), \boldsymbol{\mu}(\boldsymbol{\sigma}^k))$为式(16)的最优解.

证明：由上面的讨论可知，$(\boldsymbol{\gamma}(\boldsymbol{\sigma}^k), \boldsymbol{\mu}(\boldsymbol{\sigma}^k))$为式(16)的可行解. 此外，由于式(16)是一个凸优化问题，并且 $\boldsymbol{\sigma}^k$ 作为拉格朗日乘子$(\boldsymbol{\gamma}(\boldsymbol{\sigma}^k), \boldsymbol{\mu}(\boldsymbol{\sigma}^k))$满足最优性条件，因此$(\boldsymbol{\gamma}(\boldsymbol{\sigma}^k), \boldsymbol{\mu}(\boldsymbol{\sigma}^k))$为最优解.

令 $\phi(\boldsymbol{z}(\boldsymbol{\sigma}^k, \varepsilon_k))$表示式(16)的最优目标值函数. 注意到 $\phi(\boldsymbol{z}(\boldsymbol{\sigma}^k, \varepsilon_k))$是 $\boldsymbol{z}(\boldsymbol{\sigma}^k, \varepsilon_k)$的凸函数，并且如果 $\boldsymbol{\sigma}^*$ 是对应于式(4)最优解的任意拉格朗日乘子向量，则 $\boldsymbol{\sigma}^*$ 是 $\phi(\boldsymbol{z}(\boldsymbol{\sigma}^k, \varepsilon_k))$在 0 处的次梯度. 因此，对于满足强 ε_k-近似最优性条件的任意 $\boldsymbol{\sigma}^k$，结合引理 3 和式(17)有

$$g(\boldsymbol{\sigma}^k) = \phi(\boldsymbol{z}(\boldsymbol{\sigma}^k, \varepsilon_k)) \geqslant \phi(0) + (\boldsymbol{\sigma}^*)^{\mathrm{T}} \boldsymbol{z}^k =$$
$$g(\boldsymbol{\sigma}^*) + (\boldsymbol{\sigma}^* - \boldsymbol{\sigma}^k)^{\mathrm{T}} \boldsymbol{z}^k \geqslant g(\boldsymbol{\sigma}^*) - \|\boldsymbol{\sigma}^k - \boldsymbol{\sigma}^*\| \, \|\boldsymbol{z}^k\|. \tag{18}$$

由 $|z_i^k| \leqslant \varepsilon_k (i = 1,2,\cdots,m)$可知，$\|\boldsymbol{z}^k\| \leqslant \sqrt{m}\varepsilon_k$. 此外，由于 MVAE 问题是 MVEE 问题的一个特例，因此由文献[8]可知，存在一个对偶最优解 $\boldsymbol{\sigma}^*$ 和一个 Lipschitz 常数 L，使得 $\|\boldsymbol{\sigma}^k - \boldsymbol{\sigma}^*\| \leqslant L \|\boldsymbol{z}^k\| \leqslant L\sqrt{m}\varepsilon_k$. 再结合式(18)，可得

$$g(\boldsymbol{\sigma}^*) - g(\boldsymbol{\sigma}^k) \leqslant Lm\varepsilon_k^2. \tag{19}$$

当 $\varepsilon_k < 1$ 时，由引理 2 有 $g(\boldsymbol{\sigma}^{k+1}) - g(\boldsymbol{\sigma}^k) = \Delta_k(\tilde{\beta}_k) \geqslant \varepsilon_k^2/64$，即 $g(\boldsymbol{\sigma}^{k+1}) \geqslant g(\boldsymbol{\sigma}^k) + \varepsilon_k^2/64$. 于是，可得 $g(\boldsymbol{\sigma}^*) - g(\boldsymbol{\sigma}^{k+1}) \leqslant g(\boldsymbol{\sigma}^*) - g(\boldsymbol{\sigma}^k) - \varepsilon_k^2/64$. 结合式(19)，得

$$g(\boldsymbol{\sigma}^*) - g(\boldsymbol{\sigma}^{k+1}) \leqslant \left(1 - \frac{1}{64Lm}\right)[g(\boldsymbol{\sigma}^*) - g(\boldsymbol{\sigma}^k)]. \tag{20}$$

从而建立了算法1的线性收敛性.

注意到类似于文献[8]中的MVEE问题, 这里对于MVAE问题所建立的线性收敛性仍然是局部的. 此外, 由于文献[1]中算法3.1和文献[2]中算法2.1所计算的可行解$\boldsymbol{\sigma}^k$仅满足弱近似最优性条件(7), 因此, 并不能得到线性收敛性结果.

3 数值实验

为了验证本文提出的线性收敛算法的有效性, 对于相同的数据集, 在Matlab中同时执行了文献[1]中的算法3.1、文献[2]中的算法2.1和本文的算法1. 实验中用到的数据集是由函数randn(n,m)随机产生的正态分布数据, 其中维数(n,m)的范围为(10,5 000)~(30,30 000). 对于每组固定的维数(n,m), 随机产生10组不同的数据执行算法, 得到的结果以其算术平均值的形式给出.

表1列出了当精度$\varepsilon=10^{-3}$时, 3个算法执行的CPU时间、核心集大小和迭代次数的比较结果. 由表1可见, 3个算法得到的实际核心集大小均比理论值$O(n^2/\varepsilon)$小很多, 文献[1]中算法3.1和文献[2]中算法2.1的核心集大小几乎相同, 而本文算法1核心集大小相比其他两种算法约减少55.7%~63.2%; 在CPU时间和迭代次数方面, 本文算法1比文献[1]中算法3.1和文献[2]中算法2.1明显更优, 特别对于$n=30$, $m=30\ 000$的大规模数据, 本文算法1仅需要执行约7 s, 充分证实了本文算法1线性收敛性质的优势.

表1 3种算法的数值结果比较

Table 1 Comparison of numerical results for the three algorithms

维数(n,m)	时间/s			核心集大小			迭代次数		
	算法3.1[1]	算法2.1[2]	算法1	算法3.1[1]	算法2.1[2]	算法1	算法3.1[1]	算法2.1[2]	算法1
(10,5 000)	11.56	2.98	0.11	30.5	30.3	12.8	49 539.3	12 148.5	336.7
(10,10 000)	21.14	5.01	0.23	32.7	32.5	14.5	50 098.1	11 483.7	408.3
(20,10 000)	141.47	18.83	0.67	61.0	61.1	24.1	200 751.4	26 293.2	805.9
(20,20 000)	544.50	72.74	2.28	62.6	62.7	24.5	201 033.6	26 478.7	761.6
(30,20 000)	1 732.05	160.67	4.57	90.5	91.6	33.7	452 147.7	41 297.2	1 111.8
(30,30 000)	2 921.13	266.50	7.05	92.4	92.2	34.3	452 030.7	42 817.3	1 090.2

参考文献

[1] Kumar P, Y1ld1r1m E A. Computing Minimum Volume Enclosing Axis-Aligned Ellipsoids [J]. Journal of Optimization Theory and Applications, 2008, 136(2): 211-228.

[2] CONG Wei-jie, LIU Hong-wei. Modified Algorithms for the Minimum Volume Enclosing Axis-Aligned Ellipsoid Problem [J]. Discrete Applied Mathematics, 2010, 158(6): 627-635.

[3] Eberly D H. 3D Game Engine Design: A Practical Approach to Real Time Computer Graphis [M]. San Francisco: Kaufmann, 2001.

[4] Shawe-Taylor J, Cristianini N. Kernel Methods for Pattern Analysis [M]. Cambridge: Cambridge University Press, 2004.

[5] Y1ld1r1m E A. Two Algorithms for the Minimum Enclosing Ball Problem [J]. SIAM Journal on Optimization, 2008, 19(3): 1368-1391.

[6] Kumar P, Y1ld1r1m E A. Minimum Volume Enclosing Ellipsoids and Core Sets [J]. Journal of Optimization Theory and Applications, 2005, 126(1): 1-21.

[7] Todd M J, Y1ld1r1m E A. On Khachiyan's Algorithm for the Computation of Minimum Volume Enclosing Ellipsoids [J]. Discrete Applied Mathematics, 2007, 155(3): 1731-1744.

[8] Ahipasaoglu S D, SUN Peng, Todd M J. Linear Convergence of a Modified Frank-Wolfe Algorithm for Computing Minimum Volume Enclosing Ellipsoids [J]. Optimization Methods and Software, 2008, 23(1): 5-19.