系 统 科 学 与 数 学
J. Sys. Sci. & Math. Scis.
16(1) (1996, 1), 1–10

二次规划的直接椭球算法*

吴 士 泉
(中国科学院应用数学研究所， 北京 100080)

郭 田 德
(青岛大学师范学院数学系， 山东 266071)

摘要 本文对凸二次规划问题，给出了一个直接椭球算法，并证明了算法的复杂度为 $O(n^4L)$.

关键词 直接椭球算法，算法的复杂度，内点算法，凸二次规划.

1 引 言

1979 年苏联数学家 Khachiyan[10] 把线性规划问题转化成严格不等式组问题，给出了第一个线性规划的多项式算法 — 椭球算法. 随后，他又将该算法推广到二次规划问题. Khachiyan 的算法有着很高的理论价值，却很难用来求解具体问题. 原因大致有两个方面：其一是初始数据太大，使计算机一开始就产生溢出；其二算法需要把原问题转化成严格不等式组问题，增大了问题的维数. 二次规划的另一个多项式算法是 1987 年 Ye 和 Tse[8] 给出的，该算法是 Karmarkar[1] 投影算法的直接推广，其算法复杂度为 $O(n^4L^2)$. 最近， Goldfarb 和 Liu [7] 又给出了一个二次规划的多项式算法，算法复杂度是 $O(n^3L)$, 但算法的初始点比较难取到. 本文根据 Ye 和 Tse[8] 的思想，将二次规划问题化成一个与之等价的凸二次规划问题，然后用椭球算法直接求解. Khachiyan 椭球算法的收敛性仅是用椭球的体积来衡量的，而直接椭球算法则可以把目标函数值的下降与椭球的收缩密切联系起来，并可以避免上面提到的两个问题.

2 算法及收敛性定理

我们的算法实际上是针对非线性规划问题

$$(\overline{\mathrm{P}})\qquad \begin{aligned}&\min \bar{f}(\bar{x})=\frac{1}{2}\frac{\bar{x}[n-1]^{\mathrm{T}}\overline{Q}\bar{x}[n-1]}{\bar{x}^n}+\bar{c}^{\mathrm{T}}\bar{x}[n-1],\\ &\text{s.t. } \bar{A}\bar{x}=0,\ \bar{e}^{\mathrm{T}}\bar{x}=n,\ \bar{x}[n-1]\geq 0,\ \bar{x}^n>0\end{aligned}$$

而设置的. 其中 $\bar{x}[n-1]$ 表示由 $\bar{x}$ 的前 $n-1$ 个分量而构成的向量， $\bar{c}\in R^{n-1}$, $\bar{A}\in R^{m\times n}$, $\bar{e}=(1,\cdots,1)^{\mathrm{T}}\in R^n$, 且 $(\overline{\mathrm{P}})$ 满足如下条件： i) 最优目标函数值为 0; ii) $\operatorname{rank}\bar{A}=m,\overline{Q}$ 正定或半正定； iii) 存在 $l_0>0$, 使 $\bar{f}(\bar{x})$ 在可行区域内满足 $0\leq\bar{f}(\bar{x})\leq nl_0$.

如何把一个一般形式的凸二次规划问题化成上述问题 $(\overline{\mathrm{P}})$, 将在文章的第四部分讨论.

* 国家自然科学基金资助课题.
1991 年 9 月 19 日收到， 1993 年 8 月 10 日收到最后修改稿.

设 $(A_1, A_2, \cdots, A_{n-m-1})$ 是方程组 $\bar{A}\bar{x}=0$, $\bar{e}^{\mathrm{T}}\bar{x}=0$ 的一组正交基解，记

$$A=(A_1, A_2, \cdots, A_{n-m-1})=\begin{pmatrix} B \\ a^{\mathrm{T}} \end{pmatrix}=\begin{pmatrix} a_1^{\mathrm{T}}, a_2^{\mathrm{T}}, \cdots, a_n^{\mathrm{T}} \end{pmatrix}^{\mathrm{T}}, \tag{2.1}$$

其中 $a, a_i \in R^{n-m-1}$ $(i=1,2,\cdots,n)$. 则

$$\begin{pmatrix} \bar{A} \\ \bar{e}^{\mathrm{T}} \end{pmatrix} A=0, \text{ 且 } A^{\mathrm{T}}A=I, \tag{2.2}$$

其中 I 为单位矩阵. 方程组 $\bar{A}\bar{x}=0, \bar{e}^{\mathrm{T}}\bar{x}=n$ 的解之一般形式为 $\bar{x}=\bar{e}+Ax$, $x \in R^{n-m-1}$. 任给 $x \in R^{n-m-1}$, 作变换 T_0

$$\bar{x}=\bar{e}+Ax. \tag{2.3}$$

在线性变换 T_0 下， $(\overline{\mathrm{P}})$ 变为

$$(\mathrm{P}) \qquad \begin{aligned} &\min f(x)=\frac{1}{2}\frac{x^{\mathrm{T}}Qx+b^{\mathrm{T}}x+\alpha}{1+a^{\mathrm{T}}x}+c^{\mathrm{T}}x+\beta, \\ &\text{s.t. } x \in R, \\ &R=\{x \mid x \in R^{n-m-1}, Ax \geq -\bar{e}\}, \end{aligned} \tag{2.4}$$

其中 $Q=B^{\mathrm{T}}\overline{Q}B, b=2B^{\mathrm{T}}\overline{Q}e, \alpha=e^{\mathrm{T}}\overline{Q}e, c=B\bar{c}, \beta=\bar{c}^{\mathrm{T}}e$, 而 $e=(1,\cdots,1)^{\mathrm{T}} \in R^{n-1}$.

$$f(x)=\bar{f}(\bar{x})=\bar{f}(\bar{e}+Ax). \tag{2.5}$$

从而

$$\nabla f(x)=\frac{1}{2}\frac{2Qx+b}{1+a^{\mathrm{T}}x}-\frac{1}{2}\frac{x^{\mathrm{T}}Qx+b^{\mathrm{T}}x+\alpha}{(1+a^{\mathrm{T}}x)^2}a+c. \tag{2.6}$$

直接椭球算法：

1) 取

$$x_1=0, \; H_1=n(n-1)I, \; k:=1, \tag{2.7}$$

2) 计算

$$d=\begin{cases} \nabla f(x^k), & \text{若 } x_k \in R, \\ -a_i, & \text{若 } x_k \notin R, \text{ 且 } a_i^{\mathrm{T}}x_k<-1, \end{cases} \tag{2.8}$$

3) 令

$$x_{k+1}=x_k-\frac{1}{\sqrt{n-m}}\frac{H_kd}{\sqrt{d^{\mathrm{T}}H_kd}}, \tag{2.9}$$

$$H_{k+1}=\frac{(n-m-1)^2}{(n-m-1)^2-1}\left(H_k-\frac{2}{n-m}\frac{H_kdd^{\mathrm{T}}H_k}{d^{\mathrm{T}}H_kd}\right), \tag{2.10}$$

4) $k:=k+1$, 转回 2).

定理 2.1 设 $\{x_k\}$ 是由上述算法产生的点列，则

$$\min\{f(x_i) \mid x_i \in R, i \leq k\} \leq n(n-1)l_0 \exp\left\{-\frac{k}{2(n-m)^2}\right\}. \tag{2.11}$$

推论 2.2 设 L 为问题 $(\overline{\mathrm{P}})$ 的输入长度，则当 $k \geq k^* = \lceil 2(n-m)^2[L+\log(n(n-1)l_0)]\rceil$ 时必有

$$\min\{f(x_i)|x_i \in R, i \leq k\} \leq 2^{-L}. \tag{2.12}$$

3　收敛性定理的证明

3.1 几个基本引理

引理 3.1 若 $\bar{f}(\bar{x})$ 为凸函数，则经过线性变换 T_0 以后得到的函数 $f(x) = \bar{f}(\bar{e} + Ax)$ 仍为凸函数.

证　设 $\delta \geq 0, \nu \geq 0, \delta + \nu = 1$, 则

$$\begin{aligned} f(\delta x + \nu y) &= \bar{f}(\bar{e} + A(\delta x + \nu y)) = \bar{f}(\delta(\bar{e} + Ax) + \nu(\bar{e} + Ay)) \\ &\leq \delta \bar{f}(\bar{e} + Ax) + \nu \bar{f}(\bar{e} + Ay) = \delta f(x) + \nu f(y). \end{aligned}$$

所以，$f(x)$ 为凸函数. 证毕.

引理 3.2 函数

$$\bar{f}(\bar{x}) = \frac{1}{2}\frac{\bar{x}[n-1]^{\mathrm{T}}\overline{Q}\bar{x}[n-1]}{\bar{x}^n} + \bar{c}^{\mathrm{T}}\bar{x}[n-1]$$

在 $(\overline{\mathrm{P}})$ 的可行区域 $\overline{R} = \{\bar{x} \in R^n | \bar{A}\bar{x} = 0, \bar{e}^{\mathrm{T}}\bar{x} = n, \bar{x}[n-1] \geq 0, x^n > 0\}$ 上为凸函数.

证　用 D 表示正定的对角矩阵，作变换 $T: \bar{x} = T(x)$

$$\bar{x}[n-1] = \frac{nD^{-1}x}{e^{\mathrm{T}}D^{-1}x + 1}, \tag{3.1}$$

$$\bar{x}^n = \frac{n}{e^{\mathrm{T}}D^{-1}x + 1}, \tag{3.2}$$

其逆变换 T^{-1} 为

$$x = T^{-1}(\bar{x}) = \frac{D\bar{x}[n-1]}{\bar{x}^n}. \tag{3.3}$$

特别地，对 $\bar{x} \in \overline{R}$, 其逆变换 $x = T^{-1}(\bar{x})$ 有意义. 易见

$$\bar{f}(\bar{x}) = \bar{x}^n q(T^{-1}(\bar{x})), \tag{3.4}$$

其中 $q(x) = \frac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x$, $Q = D^{-1}\overline{Q}D^{-1}$ 为正定或半正定矩阵， $c = D^{-1}\bar{c}$. 因此， $q(x)$ 为 x 的凸函数.

设 $\delta \geq 0, \nu \geq 0, \delta + \nu = 1, \bar{x}, \bar{y} \in \overline{R}$, 且 $x = T^{-1}(\bar{x}), y = T^{-1}(\bar{y})$. 由 $q(x)$ 的凸性及 (3.4) 式知

$$\begin{aligned} \bar{f}(\delta\bar{x} + \nu\bar{y}) &= (\delta\bar{x}^n + \nu\bar{y}^n) q\left(\frac{D(\delta\bar{x}[n-1] + \nu\bar{y}[n-1])}{\delta\bar{x}^n + \nu\bar{y}^n}\right) = (\delta\bar{x}^n + \nu\bar{y}^n) q\left(\frac{(\delta\bar{x}^n)x + (\nu\bar{y}^n)y}{\delta\bar{x}^n + \nu\bar{y}^n}\right) \\ &\leq \delta\bar{x}^n q(x) + \nu\bar{y}^n q(y) = \delta\bar{x}^n q(T^{-1}(\bar{x}) + \nu\bar{y}^n q(T^{-1}(\bar{y})) = \delta\bar{f}(\bar{x}) + \nu\bar{f}(\bar{y}). \end{aligned}$$

所以， $\bar{f}(\bar{x})$ 为 $\overline{R}$ 上的凸函数. 证毕.

设 $\overline{H}$ 为正定矩阵，记 $E(\bar{x}, \overline{H}) = \{x | x \in R^{n-m-1} : (x-\bar{x})^{\mathrm{T}}\overline{H}^{-1}(x-\bar{x}) \leq 1\}$, $\mathrm{Vol}(E(\bar{x}, \overline{H}))$ 为其体积， π_0 为 $(n-m-1)$ 维单位椭球的体积，则有

引理 3.3

$$\mathrm{Vol}\,(E(\bar{x}, \overline{H})) = \pi_0(\det(\overline{H}))^{\frac{1}{2}}. \tag{3.5}$$

证 作变换 $y=\overline{H}^{-1/2}(x-\bar{x})$, 则

$$\mathrm{Vol}\,(E(\bar{x},\overline{H}))=\int_{E(\bar{x},\overline{H})}\cdots\int \mathrm{d}x^1\cdots\mathrm{d}x^{n-m-1}=[\det(\overline{H})]^{\frac{1}{2}}\int_{E(0,I)}\cdots\int \mathrm{d}y^1\cdots\mathrm{d}y^{n-m-1}$$
$$=\pi_0[\det(\overline{H})]^{\frac{1}{2}}.$$

证毕.

引理 3.4 设 $d\in R^{n-m-1}$

$$\widehat{x}=\bar{x}-\frac{1}{n-m}\frac{\overline{H}d}{\sqrt{d^{\mathrm{T}}\overline{H}d}}, \tag{3.6}$$

$$\widehat{H}=\frac{(n-m-1)^2}{(n-m-1)^2-1}\left(\overline{H}-\frac{2}{n-m}\frac{\overline{H}dd^{\mathrm{T}}\overline{H}}{d^{\mathrm{T}}H}\right), \tag{3.7}$$

则有

$$E(\bar{x},\overline{H})\cap\{x|d^{\mathrm{T}}(x-\bar{x})\le 0\}\subseteq E(\widehat{x},\widehat{H}), \tag{3.8}$$

$$\frac{\mathrm{Vol}(E(\widehat{x},\widehat{H}))}{\mathrm{Vol}(E(\bar{x},\overline{H}))}\le e^{\frac{1}{2(n-m)}}. \tag{3.9}$$

证 首先证明 (3.8) 式成立. 不失一般性, 不妨假设 $\bar{x}=0, d=e_1=(1,0,\cdots,0)^{\mathrm{T}}, \overline{H}=I$, 则

$$\widehat{x}=-\frac{e_1}{n-m},\ \widehat{H}=\frac{(n-m-1)^2}{(n-m-1)^2-1}\left(I-\frac{1}{n-m}e_1e_1^{\mathrm{T}}\right),$$
$$\widehat{H}^{-1}=\frac{(n-m-1)^2-1}{(n-m-1)^2}\left(I+\frac{2}{n-m-2}e_1e_1^{\mathrm{T}}\right).$$

若 $x\in E(\bar{x},\overline{H})\cap\{x|d^{\mathrm{T}}(x-\bar{x})\le 0\}$, 则 $x^{\mathrm{T}}x\le 1, -1\le x^1\le 0$, 从而

$$\left(x+\frac{e_1}{n-m}\right)^{\mathrm{T}}\left(\frac{(n-m-1)^2-1}{(n-m-1)^2}\left(I+\frac{2}{n-m-2}e_1e_1^{\mathrm{T}}\right)\right)\left(x+\frac{e_1}{n-m}\right)$$
$$=\frac{(n-m-1)^2-1}{(n-m-1)^2}\left[\left(x+\frac{e_1}{n-m}\right)^{\mathrm{T}}\left(x+\frac{e_1}{n-m}\right)+\frac{2}{n-m-2}\left(x^1+\frac{1}{n-m}\right)^2\right]$$
$$=\frac{(n-m-1)^2-1}{(n-m-1)^2}\left[x^{\mathrm{T}}x+\frac{2}{n-m-2}x^1(x^1+1)+\frac{1}{(n-m-2)(n-m)}\right]$$
$$\le\frac{(n-m-1)^2-1}{(n-m-1)^2}\left[1+\frac{1}{(n-m-2)(n-m)}\right]=1.$$

因此, $x\in E(\widehat{x},\widehat{H})$, (3.8) 式成立. 由 $\widehat{H}$ 的定义知

$$\det(\widehat{H})=\left[\frac{(n-m-1)^2}{(n-m-1)^2-1}\right]^{n-m-1}\left(1-\frac{2}{n-m}\right)\det(\overline{H}),$$

$$\frac{\mathrm{Vol}(E(\widehat{x},\widehat{H}))}{\mathrm{Vol}(E(\bar{x},\overline{H}))}=\left[\frac{\det(\widehat{H})}{\det(\overline{H})}\right]^{\frac{1}{2}}=\left[\frac{(n-m-1)^2}{(n-m-1)^2-1}\right]^{\frac{1}{2}(n-m-2)}\frac{n-m-1}{n-m}$$
$$=\left[1+\frac{1}{(n-m-1)^2-1}\right]^{\frac{1}{2}(n-m-2)}\frac{n-m-1}{n-m}$$
$$<\left[\exp\{\tfrac{1}{(n-m-1)^2-1}\}\right]^{\frac{1}{2}(n-m-2)}\exp\{-\tfrac{1}{n-m}\}$$
$$\le\exp\{-\tfrac{1}{2(n-m)}\}.$$

证毕.

引理 3.5 记

$$R_1=\left\{x|x\in R^{n-m-1}:\|x\|^2\le\frac{n}{n-1}\right\},\tag{3.10}$$

$$R_2=\{x|x\in R^{n-m-1}:\|x\|^2\le n(n-1)\}.\tag{3.11}$$

则

$$R_1\subseteq R\subseteq R_2.\tag{3.12}$$

证

$$\|Ax+\bar{e}\|^2=x^{\mathrm{T}}A^{\mathrm{T}}Ax+2\bar{e}^{\mathrm{T}}Ax+\bar{e}^{\mathrm{T}}\bar{e}=x^{\mathrm{T}}x+\bar{e}^{\mathrm{T}}\bar{e}=\|x\|^2+n.\tag{3.13}$$

若 $x\in R_1$, 则 $\|x\|^2\le n/(n-1)$. 现设 $x\notin R$, 则存在 i_0, $1\le i_0\le n$, 使 $a_{i_0}^{\mathrm{T}}x<-1$, 那么

$$n<n-(a_{i_0}^{\mathrm{T}}x+1)=\sum_{\substack{j=1\\j\ne i_0}}^{n}(a_j^{\mathrm{T}}x+1)\le(n-1)^{\frac{1}{2}}\|Ax+\bar{e}\|=(n-1)^{\frac{1}{2}}(\|x\|^2+n)^{\frac{1}{2}}$$
$$\le(n-1)^{\frac{1}{2}}\left(\frac{n}{n-1}+n\right)^{\frac{1}{2}}=n,$$

从而矛盾. 因此，$x\in R$, $R_1\subseteq R$ 成立.

再证 $R\subseteq R_2$. 若 $x\in R$, 则 $Ax\ge-\bar{e}$, 故

$$n^2=(\bar{e}^{\mathrm{T}}(Ax+\bar{e}))^2=\left(\sum_{i=1}^{n}(a_i^{\mathrm{T}}x+1)\right)^2=\sum_{i=1}^{n}(a_i^{\mathrm{T}}x+1)^2$$
$$+\sum_{\substack{i=1\\i\ne j}}^{n}\sum_{j=1}^{n}(a_i^{\mathrm{T}}x+1)(a_j^{\mathrm{T}}x+1)\ge\sum_{i=1}^{n}(a_i^{\mathrm{T}}x+1)^2=\|Ax+\bar{e}\|^2=\|x\|^2+n.$$

所以，$\|x\|^2\le n(n-1)$, 即 $x\in R_2$, 故 $R_1\subseteq R\subseteq R_2$. 证毕.

引理 3.6(Brümn-Minkowski) 设 $c_1,c_2\subset R^{n-m-1}$ 是两个凸紧集，则任给 $\lambda\in(0,1)$ 有

$$[\mathrm{Vol}(\lambda c_1+(1-\lambda)c_2)]^{\frac{1}{(n-m-1)}}\ge\lambda[\mathrm{Vol}(c_1)]^{\frac{1}{n-m-1}}+(1-\lambda)[\mathrm{Vol}(c_2)]^{\frac{1}{n-m-1}}.\tag{3.14}$$

证 见文 [11]. 略.

3.2 主要定理的证明

定理 2.1 的证明

设

$$\varepsilon_k=\min\{f(x_i)|x_i\in R,i\le k,k\ge1\}.\tag{3.15}$$

由于 $x_1\in R$, 因此，ε_k 能够取到. 记

$$\Omega_i=\begin{cases}R, & 若\ x_i\notin R,\\ \{x|x\in R,f(x)\le f(x_i)\}, & 若\ x_i\in R.\end{cases}\tag{3.16}$$

对任何 $x_i\in R$, $i\le k$, 必有

$$\varepsilon_k\le f(x_i),\tag{3.17}$$

故

$$\varphi(\varepsilon_k):=\{x|x\in R,f(x)\le\varepsilon_k\}\subset\Omega_i\ (i\le k).\tag{3.18}$$

记

$$U_{k+1}=\bigcap_{i=1}^{k}\Omega_i, \tag{3.19}$$

那么对 $k\ge 1$ 必有

$$U_{k+1}\subset E(x_{k+1},H_{k+1}). \tag{3.20}$$

事实上，当 $k=1$ 时，由 $E(x_2,H_2)$ 的构成及引理 3.4 可知 $E(x_1,H_1)\cap\{x|f(x)\le f(x_1)\}\subset E(x_2,H_2)$. 又由 $R\subset E(x_1,H_1)$ 知， $U_2=\Omega_1=\{x\in R: f(x)\le f(x_1)\}\subset E(x_2,H_2)$. 若 $k=h-1$ 时 (3.20) 式成立，即 $U_h\subset E(x_h,H_h)$.

i) 如果 $x_h\in R$, 则 $U_{h+1}=U_h\cap\Omega_h=U_h\cap\{x|f(x)\le f(x_h)\}\subset E(x_h,H_h)\cap\{x|f(x)\le f(x_h)\}\subset E(x_{h+1},H_{h+1})$.

ii) 如果 $x_h\notin R$, 则存在 i 使得 $a_i^{\mathrm T}x_h<-1$, 那么 $R=\{x|Ax\ge -\bar e\}\subset\{x|a_i^{\mathrm T}x\ge -1\}\subset\{x|a_i^{\mathrm T}x\ge a_i^{\mathrm T}x_h\}$, $U_{h+1}=U_h\cap\Omega_h=U_h\cap R\subset E(x_h,H_h)\cap\{x|a_i^{\mathrm T}x\ge a_i^{\mathrm T}x_h\}\subset E(x_{h+1},H_{h+1})$.

因此， $U_{k+1}\subset E(x_{k+1},H_{k+1})$ 对 $k\ge 1$ 成立.

由 (3.18) 式知， $\varphi(\varepsilon_k)\subset\Omega_i, i\le k$, , 故

$$\varphi(\varepsilon_k)\subset\bigcap_{i=1}^{k}\Omega_i=U_{k+1}\subset E(x_{k+1},H_{k+1}), \tag{3.21}$$

$$\begin{aligned}[\mathrm{Vol}(\varphi(\varepsilon_k))]^{\frac{1}{n-m-1}} &\le [\mathrm{Vol}(E(x_{k+1},H_{k+1}))]^{\frac{1}{n-m-1}}\\ &\le \exp\{\tfrac{-k}{2(n-m)^2}\}[\mathrm{Vol}(E(x_1,H_1))]^{\frac{1}{n-m-1}}.\end{aligned} \tag{3.22}$$

另一方面，由于任给 $x\in R, 0\le f(x)\le nl_0$, 所以 $0\le\varepsilon_k\le nl_0$. 记 $\lambda=\varepsilon_k/(nl_0)$ 则 $\lambda\in[0,1]$, 且 $\varepsilon_k=\lambda nl_0$, 那么

$$(1-\lambda)\varphi(0)+\lambda\varphi(nl_0)\subset\varphi(\varepsilon_k) \tag{3.23}$$

必然成立. 事实上，设 $x\in(1-\lambda)\varphi(0)+\lambda\varphi(nl_0)$, 则存在 $x_1\in\varphi(0)$, $x_2\in\varphi(nl_o)$ 满足 $x=(1-\lambda)x_1+\lambda x_2$. 由 $f(x)$ 的凸性知， $f(x)\le(1-\lambda)f(x_1)+\lambda f(x_2)=\lambda f(x_2)\le\lambda nl_0=\varepsilon_k$, 即 $x\in\varphi(\varepsilon_k)$, 从而 (3.23) 式成立. 又由引理 3.6

$$\begin{aligned}[\mathrm{Vol}(\varphi(\varepsilon_k))]^{\frac{1}{n-m-1}} &\ge (1-\lambda)[\mathrm{Vol}(\varphi(0))]^{\frac{1}{n-m-1}}+\lambda[\mathrm{Vol}(\varphi(nl_0))]^{\frac{1}{n-m-1}}\\ &\ge \lambda[\mathrm{Vol}(\varphi(nl_0))]^{\frac{1}{n-m-1}}=\lambda[\mathrm{Vol}(R)]^{\frac{1}{n-m-1}},\end{aligned}$$

所以

$$\begin{aligned}\varepsilon_k=\lambda nl_0 &\le nl_0\left[\frac{\mathrm{Vol}(\varphi(\varepsilon_k))}{\mathrm{Vol}(R)}\right]^{\frac{1}{n-m-1}}\le nl_0\exp\{\tfrac{-k}{2(n-m)^2}\}\left[\frac{\mathrm{Vol}(R_2)}{\mathrm{Vol}(R_1)}\right]^{\frac{1}{n-m-1}}\\ &\le n(n-1)l_0\exp\{\tfrac{-k}{2(n-m)^2}\}.\end{aligned}$$

定理证毕.

4 一般凸二次规划问题

考虑一般的凸二次规划问题

$$(\mathrm{QP})\qquad \begin{aligned}\min\quad & q(x)=\frac{1}{2}x^{\mathrm T}Qx+c^{\mathrm T}x,\\ \mathrm{s.t.}\quad & Ax=b, x\ge 0,\end{aligned}$$

其中 Q 为正定或半正定矩阵， $A \in R^{m\times(n-1)}, x \in R^{n-1}, c \in R^{n-1}, b \in R^m$, 且 A, b 的元素均为整数. 可行区域 $S=\{x : Ax=b, x\geq 0\}$ 有界，即存在 $M_0>0$, 对任给的 $x=(x^1,\cdots,x^{n-1})\in S$ 都满足 $x^i\leq M_0\,(i=1,2,\cdots,n-1)$. 通过简单变换，不妨设 $x^i\leq 1\,(i=1,2,\cdots,n-1)$. 令 $l_0\geq 0$ 为 $q(x)$ 在 S 上的一个上界.

4.1 化最优目标函数值为零

(QP) 的对偶规划是另外一个二次规划问题

$$(\mathrm{QD})\qquad \begin{aligned} &\max\quad d(x,y)=b^{\mathrm T}y-\tfrac{1}{2}x^{\mathrm T}Qx,\\ &\mathrm{s.t.}\quad Ax=b,\ A^{\mathrm T}y-Qx+z=c,\ x\geq 0, z\geq 0. \end{aligned}$$

如果 (QP), (QD) 可行，由对偶定理，下列凸二次规划问题

$$\min q(x)-d(x,y)=x^{\mathrm T}Qx+c^{\mathrm T}x-b^{\mathrm T}y,\ \mathrm{s.t.}Ax=b,\ A^{\mathrm T}y+z-Qx=c,\ x\geq 0,\ z\geq 0$$

有解，且最优值为 0. 因此，不妨就直接假设 (QP) 的最优值为 0.

4.2 求 (QP) 的一个可行内点 x_0

对 (QP), 由假设, 任给 $x=(x^1,x^2,\cdots,x^{n-1})\in S$, 都满足 $0\leq x^i\leq 1\,(i=1,2,\cdots,n-1)$, 因此 (QP) 与下列二次规划问题等价

$$(\mathrm{QP}_1)\qquad \begin{aligned} &\min\quad \tfrac{1}{2}x^{\mathrm T}Qx+c^{\mathrm T}x,\\ &\mathrm{s.t.}\quad Ax=b,\ e^{\mathrm T}x\leq n-1,\ x\geq 0. \end{aligned}$$

记

$$D=\begin{pmatrix} Q & A^{\mathrm T} & -I\\ A & 0 & 0 \end{pmatrix}, \tag{4.1}$$

$$M_1=\max\{|\det(\overline{D})| : \overline{D}\ 为\ D\ 的非奇异子方阵\}. \tag{4.2}$$

显然 M_1 只与 A, Q 有关.

引理 4.1 (QP) 存在最优解 $\widehat{x}$ 满足： $\widehat{x}$ 的任何非零分量 $\widehat{x}^i\geq 1/M_1$.

证　由 Karush-Kuhn-Tucker 条件，x 是 (QP) 之最优解的充要条件为存在 $y\in R^m, v\in R^{n-1}$ 满足

$$Qx+A^{\mathrm T}y-v=-c, \tag{4.3}$$

$$Ax=b, \tag{4.4}$$

$$x^{\mathrm T}v=0,\ x\geq 0,\ v\geq 0. \tag{4.5}$$

记 $w=(x^{\mathrm T},y^{\mathrm T},v^{\mathrm T})^{\mathrm T}, \bar{b}=(-c^{\mathrm T},b^{\mathrm T})^{\mathrm T}$, 则 (4.3), (4.4) 即为

$$DW=\bar{b}. \tag{4.6}$$

由引理的假设及文 [15, 16] 的定理得引理的结论. 证毕.

在 (4.1) 式中，把 A 换成

$$\begin{pmatrix} A & b-Ae & 0\\ e^{\mathrm T} & 1 & 1 \end{pmatrix}, \tag{4.7}$$

由 (4.2) 得到的值仍记为 M_1, 则 M_1 为常数. 取

$$M \geq \left\lceil M_1 \left(\frac{n-1}{2}\|Q\| + (n+2)\sum_{i=1}^{n-1}|c_i| \right) \right\rceil + 1, \tag{4.8}$$

其中 $\lceil \alpha \rceil$ 表示不小于实数 α 的整数. 构造下列二次规划问题

$$(\mathrm{QPM}) \qquad \begin{array}{ll} \min & \frac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x + Mx^n, \\ \text{s.t.} & Ax + (b - Ae)x^n = b, \\ & e^{\mathrm{T}}x + x^n + x^{n+1} = n+1,\ x \geq 0,\ x^n,\ x^{n+1} \geq 0. \end{array}$$

显然 $x_0 = (1, \cdots, 1)^{\mathrm{T}} \in R^{n+1}$ 为 (QPM) 的一个可行内点.

引理 4.2 若 (QP) 可行, 则 (QPM) 存在形如 $(\bar{x}^{\mathrm{T}}, 0, \bar{x}^{n+1})^{\mathrm{T}}$ 的最优解, 且 $\bar{x}$ 为 (QP) 的最优解.

证 由引理 4.1 知, 将 A 换成 (4.7), b, c 分别换成 $(b^{\mathrm{T}}, n+1)^{\mathrm{T}}, (c^{\mathrm{T}}, M, 0)^{\mathrm{T}}$, 那么 (QPM) 存在最优解 $(\bar{x}^{\mathrm{T}}, \bar{x}^n, \bar{x}^{n+1})^{\mathrm{T}}$, 使其任何非零分量 $\bar{x}^i$ 都有 $\bar{x}^i \geq 1/M_1$, 此时必有 $\bar{x}^n = 0$. 否则, 若 $\bar{x}^n > 0$, 则 $\bar{x}^n \geq 1/M_1$. 设 $\widehat{x}$ 为 (QP) 的一个最优解, 取 $\widehat{x}^{n+1} = n+1 - e^{\mathrm{T}}\widehat{x}$, 由 $e^{\mathrm{T}}\widehat{x} \leq n-1$ 知, $\widehat{x}^{n+1} \geq 2 > 0$, 则 $(\widehat{x}^{\mathrm{T}}, 0, \widehat{x}^{n+1})^{\mathrm{T}}$ 为 (QPM) 的一个可行解, 从而

$$\frac{1}{2}\widehat{x}^{\mathrm{T}}Q\widehat{x} + c^{\mathrm{T}}\widehat{x} \geq \frac{1}{2}\bar{x}^{\mathrm{T}}Q\bar{x} + c^{\mathrm{T}}\bar{x} + M\bar{x}^n \geq c^{\mathrm{T}}\bar{x} + M\bar{x}^n.$$

又 $0 \leq \widehat{x}^i \leq 1\ (i = 1, 2, \cdots, n-1)$, $e^{\mathrm{T}}\bar{x} + \bar{x}^n + \bar{x}^{n+1} = n+1$, 故 $\bar{x}^i \leq n+1\ (i = 1, 2, \cdots, n+1)$, 且

$$\begin{aligned} \frac{n-1}{2}\|Q\| + \sum_{i=1}^{n-1}|c_i| &\geq \frac{1}{2}\|\widehat{x}\|^2\|Q\| + c^{\mathrm{T}}\widehat{x} \geq \frac{1}{2}\widehat{x}^{\mathrm{T}}Q\widehat{x} + c^{\mathrm{T}}\widehat{x} \\ &\geq c^{\mathrm{T}}\bar{x} + M\bar{x}^n \geq \frac{M}{M_1} - (n+1)\sum_{i=1}^{n-1}|c_i|. \end{aligned}$$

即

$$M \leq M_1 \left(\frac{n-1}{2}\|Q\| + (n+2)\sum_{i=1}^{n-1}|c_i| \right)$$

与 (4.8) 矛盾. 即知 $\bar{x}^n = 0$, 则 $\bar{x}$ 为 (QP_1) 的最优解, 从而 $\bar{x}$ 为 (QP) 之最优解. 证毕.

定理 4.3 设 $(\bar{x}^{\mathrm{T}}, \bar{x}^n\bar{x}^{n+1})^{\mathrm{T}}$ 为 (QPM) 的某一最优解, 若 $\bar{x}^n = 0$, 则 $\bar{x}$ 为 (QP) 之最优解. 若 $\bar{x}^n > 0$, 则可以求得 (QPM) 的另外一个最优解 $(\widehat{x}^{\mathrm{T}}, \widehat{x}^n, \widehat{x}^{n+1})^{\mathrm{T}}$ 满足 $\widehat{x}^n = 0$ 且 $\widehat{x}$ 为 (QP) 之最优解.

证 若 $\bar{x}^n = 0$, 则结论成立. 若 $\widehat{x}^n > 0$, 由上述的讨论可知, 可以求得 (QPM) 的另外一个最优解 $(\widehat{x}^{\mathrm{T}}, 0, \widehat{x}^{n+1})^{\mathrm{T}}$, 从而得到 (QP) 的最优解 $\widehat{x}$. 证毕.

4.3 把 (QP) 化成算法标准型 $(\overline{\mathrm{P}})$

根据上述讨论, 不妨假设 (QP) 满足:

A1) 最优值为 0;

A2) $x_0 > 0$ 为 (QP) 的一个已知可行内点;

A3) Q 正定或半正定, $\operatorname{rank} A = m$;

A4) $l_0 \geq 0$ 是 $q(x)$ 在可行域 S 上的一个上界;

A5) A, b, c 的元素皆为整数.

记 $D=\operatorname{diag}(x_0^1,x_0^2,\cdots,x_0^{n-1})$, 并考虑问题

$$(\overline{\mathrm{P}})\qquad \begin{aligned}&\min\ \bar{f}(\bar{x})=\frac{1}{2}\frac{\bar{x}[n-1]^{\mathrm{T}}\overline{Q}\bar{x}[n-1]}{\bar{x}^n}+\bar{c}^{\mathrm{T}}\bar{x}[n-1],\\&\text{s.t. }\bar{A}\bar{x}=0,\ \bar{e}^{\mathrm{T}}\bar{x}=n,\ \bar{x}[n-1]\geq 0,\bar{x}^n>0,\end{aligned}$$

其中 $\bar{A}=(AD,-b),\overline{Q}=DQD,\bar{c}=Dc$.

引理 4.4 若 $(\widehat{x}[n-1]^{\mathrm{T}},\widehat{x}^n)^{\mathrm{T}}$ 是 $(\overline{\mathrm{P}})$ 的最优解，则 $x^*=\widehat{x}[n-1]/\widehat{x}^n$ 是 (QP) 的最优解；若 x^* 是 (QP) 的最优解，则

$$\widehat{x}[n-1]=\frac{nD^{-1}x^*}{e^{\mathrm{T}}D^{-1}x^*+1},\ \widehat{x}^n=\frac{n}{e^{\mathrm{T}}D^{-1}x^*+1}$$

是 $(\overline{\mathrm{P}})$ 的最优解，且 $(\overline{\mathrm{P}})$ 和 (QP) 的最优值皆为 0.

证 作变换 T: $\bar{x}=T(x)$,

$$\bar{x}[n-1]=\frac{nD^{-1}x}{e^{\mathrm{T}}D^{-1}x+1},\ \bar{x}^n=\frac{n}{e^{\mathrm{T}}D^{-1}x+1}. \tag{4.9}$$

其逆变换 T^{-1},

$$x=T^{-1}(\bar{x})=\frac{D\bar{x}[n-1]}{\bar{x}^n}. \tag{4.10}$$

在变换 T 下，(QP) 变为

$$(\overline{\mathrm{P}}')\qquad \begin{aligned}&\min\ \ q(T^{-1}(\bar{x})),\\&\text{s.t. }\ \bar{A}\bar{x}=0,\ \bar{e}^{\mathrm{T}}\bar{x}=n,\ \bar{x}[n-1]\geq 0,\ \bar{x}^n>0.\end{aligned}$$

考虑与 $(\overline{\mathrm{P}}')$ 相关的凸规划问题

$$\begin{aligned}&\min \bar{f}(\bar{x})=\bar{x}^n q(T^{-1}(\bar{x})),\\&\text{s.t. }\bar{A}\bar{x}=0,\bar{e}^{\mathrm{T}}\bar{x}=n,\ \bar{x}[n-1]\geq 0,\ \bar{x}^n>0.\end{aligned}$$

通过计算可得

$$\bar{f}(\bar{x})=\frac{1}{2}\frac{\bar{x}[n-1]^{\mathrm{T}}\bar{Q}\bar{x}[n-1]}{\bar{x}^n}+\bar{c}^{\mathrm{T}}\bar{x}[n-1]. \tag{4.11}$$

由于 (QP) 的最优值为 0, 从而 $(\overline{\mathrm{P}}')$ 的最优值为 0, 故 $(\overline{\mathrm{P}})$ 的最优值为 0, 且 $(\overline{\mathrm{P}})$ 与 $(\overline{\mathrm{P}}')$ 的最优解相同，因此引理成立. 证毕.

引理 4.5 在 $(\overline{\mathrm{P}})$ 的可行域内有 $0\leq\bar{f}(\bar{x})\leq nl_0$.

证 因 $0\leq q(x)\leq l_0$, 所以 $0\leq q(T^{-1}(\bar{x}))\leq l_0$, 又由 $\bar{e}^{\mathrm{T}}\bar{x}=n$ 易知， $\bar{x}\leq n$. 从而得到 $0\leq\bar{f}(\bar{x})=\bar{x}^n q(T^{-1}(\bar{x}))\leq nl_0$. 证毕.

参 考 文 献

[1] Karmarkar N. A new polynomial algorithm for linear programming. *Combinatorica*, 1984, 4: 373–395.

[2] Todd M J and Burrell B P. An extension of Karmarkar's algorithm for linear programming using dual variables. *Algorithmica*, 1986, **1**: 409–424.

[3] Roos C and Vail J -ph. A polynomial method of approximate centers for linear programming. Reports of the Faculty of Technical Mathematics and Informatics no. 88–68, Delft, 1988.

[4] Gonzag C C. An algorithm for solving linear programming in $O(n^3L)$ operations. in Megiddo ed., Progress in Mathematical Programming, Springer-Verlag, 1–28.

[5] Renegar J. A polynomial-time algorithm based on Newton's method for linear programming. *Mathematical Programming*, **40**: 59–93.

[6] Sun Jie. A convergence proof of an affine-scaling algorithm for. convex quadratic programming without nondegeneracy assumptions. Report of Department of Industrial Engineering and Management Sciences Northwest University, May, 1990.

[7] Goldfarb D and Liu S. An $O(n^3L)$ primal interior point algorithm for convex quadratic programming. *Mathematical Programming*, 1991, **49**: 325–340.

[8] Ye Yinyu and Edison Tse. An extension of Karmarkar's projective algorithm for convex quadratic programming. *Mathematical Programming*, 1989, **44**: 157–179.

[9] Renato D C Monteiro and Ilan Adler. Interior path following primal-dual algorithms. Part I: linear programming. *Mathematical Programming*, 1989, **44**: 27–41.

[10] Khachiyan L G. A polynomial algorithm in linear programming. *Dokl.Arad.Nauk. SSSR* 244 (English translation: Soviet Math.Dokl , **20**: 191–194.

[11] Wu Shiquan and Wu Fang. A direct ellipsoid method for linear programming. to appear in Bulletin of Australian Mathematical Society.

[12] Eggleston H G. Convexity. Cambridge University Press, 1958.

[13] Martin Grotsches, Lazlo Lovasz and Alexander Schrijver. Geometric Algorithms and Combinatorial Optimization. Spinger-Verlag, 1988.

[14] Renato D C Monteiro and Ilan Adler. Interior path following primal-dual algorithms part II: Convex quadratic programming. *Mathematical Programming*, 1989, **44**: 43–66.

[15] Cottle R W and Dantzig G B. Complementary pivot theory of mathematical programming. *Linear Algebra and Its Applications*, 1968,**1**: 103–125.

[16] Lemke C E. Bimatrix equilibrium points and mathematical programming. *Management Science*, 1965, **11**: 681–689.

DIRECT ELLIPSOID METHOD FOR CONVEX QUADRATIC PROGRAM

Wu Shi-Quan

(*Institute of Applied Mathematics, Academia Sinica Beijing* 100080)

Guo Tian-De

(*Depaartment of Mathematics, Teacher's College of Qingdao University, Shandong* 266071)

Abstract This paper shows how to apply the ellipsoid method directly to the convex quadratic program, and proves that the direct ellipsoid method can be terminated at an approximate optimal point in at most $O(n^2L)$ iterations with a total of $O(n^4L)$ arithmetic operations.

Key words Direct ellipsoid method, computational complexity, interior point method, convex quadratic program.