文章编号：0253-374X(2013)05-0788-04　　DOI:10.3969/j.issn.0253-374x.2013.05.026

Gurman 摄动变换在非凸全局优化中的应用

朱经浩，陈硕晶
(同济大学 数学系，上海 200092)

摘要：讨论球约束下的一类非凸函数的全局优化问题. 把全局优化问题转化为奇异最优控制问题，通过 Gurman 摄动变换引入 canonical 全局优化方法，得到判别全局优化问题的最优解的等价性条件和必要条件，并证明球约束下非凸二次函数的全局优化问题的最优解的一个充要条件.

关键词：Gurman 摄动变换；球约束下全局优化；奇异最优控制
中图分类号：O224；O232　　**文献标志码**：A

Non-convex Global Optimization with Gurman Perturbation

ZHU Jinghao, *CHEN Shuojing*
(Department of Mathematics, Tongji University, Shanghai 200092, China)

Abstract: The global optimization of a non-convex function over a sphere is investigated by the Gurman perturbation method and the canonical backward flow. The constrained optimization is converted into a singular optimal control problem for solving non-convex global optimization. A sufficient and necessary optimality condition is obtained for the non-convex quadratic optimization over a sphere.

Key words: Gurman perturbation; global optimization; singular optimal control

考虑以下在球约束下的全局优化问题(P)：

$$\begin{aligned}&(P):\min Q(x)\\&\text{s.t.}\ x\in U=\{\|x\|\leqslant 1\}\end{aligned}\tag{1}$$

式中：假设 $Q(x)$ 是二阶连续可微的非凸函数；U 是凸的有界闭区域，不失一般性，认为 U 包含原点；x 为相应的解. 全局优化问题(P)通常隶属于一般优化方法，例如在信赖域法和序列二次规划方法中作为子问题出现，当 $Q(x)$ 为一般非凸函数时这个问题是 N-P 难的(non-deterministic polynomial hard)，即认为这类问题的大型实例不能用精确算法求解，必须寻求这类问题的有效近似算法. 对问题(P)的最优解的存在性条件的研究是优化领域的一个热点. 最优化方法与最优控制理论之间有着紧密的联系. 本文建立与问题(P)等价的最优控制问题，利用 Gurman 摄动方法[1]，得到原问题最优解判别的等价性条件和必要性条件. 当 $Q(x)$ 为一类非凸的二次函数时，本文建立了相应的球约束下的非凸二次优化问题最优解的充要条件. 对于一般非凸球约束的二次规划，一般利用 KKT 条件(Karush-Kuhn-Tucker 最优化条件)进行数值求解，而 canonical 对偶方法则研究非凸二次优化的对偶规划[2-3]，与之不同，本文建了非凸二次多项式在最优点处的梯度和最小的负特征值的模之间的不等式，并用于求解. 给出了球约束下的一般非凸二次函数的全局优化问题最优解的一个充要条件，即当 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$(其中 $\boldsymbol{A}$ 为对称半正定矩阵，且至少有 1 个特征值小于零，$\boldsymbol{f}$ 为非零向量，$\boldsymbol{u}$ 为满足条件的解)时，$\boldsymbol{u}^*$ 是问题(P)的最优解的充要条件是

$$\begin{cases}\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|>-\min\limits_{\lambda<0}\lambda\\ \boldsymbol{u}^*=\dfrac{-(\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f})}{\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|}\end{cases}\tag{2}$$

其中，λ 为 $\boldsymbol{A}$ 的特征值.

1　与全局优化问题等价的最优控制问题

为了应用 Gurman 摄动方法研究全局优化问题(P)，引用文献[3]中的结果建立与原问题等价的最优控制问题.

命题 1[3]　原全局优化问题(P)等价于以下最优控制问题(P_1)：

收稿日期：2012-04-28
基金项目：国家自然科学基金(10671145)
第一作者：朱经浩(1949—)，男，教授，博士生导师，理学博士，主要研究方向为最优控制、数学规划. E-mail:jinghaok@tongji.edu.cn
通讯作者：陈硕晶(1988—)，女，硕士生，主要研究方向为最优控制、数学规划. E-mail:daisyshuoshuo@sina.com

$$(P_1):\min\{Q(\boldsymbol{x}(1))\}$$
$$\text{s.t. } \dot{\boldsymbol{x}}(t)=\boldsymbol{u}(t),\boldsymbol{u}(t)\in U,t\in[0,1],\boldsymbol{x}(0)=0 \tag{3}$$

而且最优控制可取常向量形式．式中：$\boldsymbol{x}(t)$是在某一时刻 t 时的轨道；$\boldsymbol{u}(t)$是相应的控制．这里的最优控制问题是寻找允许控制系统，使得当 $t=1$ 时，$Q(\boldsymbol{x}(1))$最小．

命题 2[3]　最优控制问题(P_1)等价于以下最优控制问题(P_2)：

$$(P_2):J(u(t))=\min\left\{\int_0^1\nabla^{\mathrm{T}}\boldsymbol{Q}(x)\boldsymbol{u}(t)\mathrm{d}t\right\}$$
$$\text{s.t. } \dot{\boldsymbol{x}}(t)=\boldsymbol{u}(t),\boldsymbol{u}(t)\in U,t\in[0,1],\boldsymbol{x}(0)=0 \tag{4}$$

式中：$\boldsymbol{u}(t)$为允许的控制；$\boldsymbol{x}(t)$为相应的轨道．这里的最优控制问题的允许控制系统是[0,1]区间上的按段连续的向量函数．

2　Gurman 摄动变换

对于最优控制问题(P_2)引入 Gurman 摄动变换[1,4]，对给定的正实数 $\alpha\in(0,1)$，构造在 α 下的摄动函数 $J_\alpha(u)$，引入以下最优控制问题(P_3)：

$$(P_3):\min J_\alpha(\boldsymbol{u})=\left\{\int_0^1\left[(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{x})\boldsymbol{u}(t)+\frac{\alpha}{2}\boldsymbol{u}(t)^{\mathrm{T}}\boldsymbol{u}(t)\right]\mathrm{d}t\right\}$$
$$\text{s.t. } \dot{\boldsymbol{x}}(t)=\boldsymbol{u}(t),t\in[0,1],\boldsymbol{x}(0)=0,\boldsymbol{u}(t)\in U. \tag{5}$$

引理 1　最优控制问题(P_3)的最优控制是一个常向量控制，而相应的最优轨道是一个线性向量函数．

证明　最优控制问题(P_3)的 Hamilton 函数为 $H(t,\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\eta}(t))=\boldsymbol{\eta}(t)^{\mathrm{T}}\boldsymbol{u}(t)+(1-\alpha)\cdot\nabla^{\mathrm{T}}\cdot\boldsymbol{Q}(\boldsymbol{x}(t))\boldsymbol{u}(t)+\frac{\alpha}{2}\boldsymbol{u}(t)^{\mathrm{T}}\boldsymbol{u}(t)$，根据 Pontryagin 原理，如果$(\hat{\boldsymbol{x}}(t),\hat{\boldsymbol{u}}(t),\hat{\boldsymbol{\eta}}(t))$是 t 时刻(P_3)的最优三元组，则有 $\dot{\hat{\boldsymbol{x}}}(t)=H_\eta=\hat{\boldsymbol{u}}(t),\hat{\boldsymbol{x}}(0)=0,\dot{\hat{\boldsymbol{\eta}}}(t)=-H_x=-(1-\alpha)\nabla^2\boldsymbol{Q}(\hat{\boldsymbol{x}}(t))\hat{\boldsymbol{u}}(t),\hat{\boldsymbol{\eta}}(1)=0$，由此得 $\dot{\hat{\boldsymbol{\eta}}}(t)=-(1-\alpha)\nabla^2\boldsymbol{Q}(\hat{\boldsymbol{x}}(t))\dot{\hat{\boldsymbol{x}}}(t)$，从而 $\hat{\boldsymbol{\eta}}(t)+(1-\alpha)\nabla\boldsymbol{Q}(\hat{\boldsymbol{x}}(t))$恒为常数，对于 $t\in[0,1]$几乎处处成立．注意到$\hat{\boldsymbol{\eta}}(1)=0$，就有

$$\hat{\boldsymbol{\eta}}(t)+(1-\alpha)\nabla\boldsymbol{Q}(\hat{\boldsymbol{x}}(t))\equiv(1-\alpha)\nabla\boldsymbol{Q}(\hat{\boldsymbol{x}}(1)) \tag{6}$$

同时有

$$\hat{\boldsymbol{u}}\equiv\arg\min_{u\in U}\left\{\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{u}+(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(\hat{x})\boldsymbol{u}+\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}\right\} \tag{7}$$

对 $t\in[0,1]$几乎处处成立．由式(6)、式(7)，有

$$\hat{\boldsymbol{u}}(\cdot)\equiv\arg\min_{u\in U}\left\{(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(\hat{\boldsymbol{x}}(1))\boldsymbol{u}+\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}\right\} \tag{8}$$

即(P_3)的最优控制 $\hat{u}(\cdot)$恒取二次优化问题 $\min\limits_{u\in U}\left\{(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(\hat{\boldsymbol{x}}(1))\boldsymbol{u}+\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}\right\}$的唯一最优点而不依赖于 t，所以最优控制函数为一常向量，记为 $\hat{\boldsymbol{u}}$．注意到所论系统为 $\dot{\boldsymbol{x}}=\boldsymbol{u},t\in[0,1],\boldsymbol{x}(0)=0$，则易见相应的最优轨道为 $\hat{\boldsymbol{x}}(t)=t\hat{\boldsymbol{u}}$，同时有

$$\hat{\boldsymbol{u}}=\hat{\boldsymbol{x}}(1) \tag{9}$$

证毕．

于是可在常向量控制函数范围内寻找最优控制，而相应的评价函数则可表为

$$J_\alpha(\boldsymbol{u})=\left\{\int_0^1\left[(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(t\boldsymbol{u})\boldsymbol{u}+\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}\right]\mathrm{d}t\right\}=\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}+\left[\int_0^1(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(t\boldsymbol{u})\boldsymbol{u}\mathrm{d}t\right] \tag{10}$$

于是求解最优控制问题(P_3)相当于求解以下优化问题：

$$\min_{\|u\|\leqslant1}J_\alpha(\boldsymbol{u})=\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}+\left[\int_0^1(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(t\boldsymbol{u})\boldsymbol{u}\mathrm{d}t\right] \tag{11}$$

定理 1(等价性条件)　如果对$\|\boldsymbol{u}\|\leqslant1$上的任一点，$\nabla^2J_\alpha(\boldsymbol{u})$非正定，则 $J_\alpha(\boldsymbol{u})=\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}+\left[\int_0^1(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(t\boldsymbol{u})\mathrm{d}t\right]u$ 在$\|\boldsymbol{u}\|\leqslant1$上的最小点位于$\|\boldsymbol{u}\|=1$上，且($P_3$)的常向量最优控制与($P_2$)的常向量最优控制一致．

证明　因为对$\|\boldsymbol{u}\|\leqslant1$上的任一点，$J_\alpha(\boldsymbol{u})$非凸，从而$\nabla^2J_\alpha(\boldsymbol{u})$至少有 1 个特征值小于零，由最优化理论[5]可知，$J_\alpha(\boldsymbol{u})$ 在$\|\boldsymbol{u}\|\leqslant1$上的最小点 $\hat{\boldsymbol{u}}$，满足 $\hat{\boldsymbol{u}}^{\mathrm{T}}\hat{\boldsymbol{u}}=1$，否则$\nabla^2J_\alpha(\hat{\boldsymbol{u}})\geqslant0$，与 $J_\alpha(\hat{\boldsymbol{u}})$非凸矛盾．另外由命题 1 和 2，($P_2$)的最优控制可取常向量形式．这样，对于任一取值于单位球面上的常数控制 $\boldsymbol{u}$，有 $J_\alpha(\boldsymbol{u})\geqslant J_\alpha(\hat{\boldsymbol{u}})$以及 $\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}=\frac{\alpha}{2}=\frac{\alpha}{2}\hat{\boldsymbol{u}}^{\mathrm{T}}\hat{\boldsymbol{u}}$．注意到 $\alpha\in(0,1)$，于是由式(10)可得

$$\int_0^1(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(t\boldsymbol{u})\boldsymbol{u}\mathrm{d}t=J_\alpha(\boldsymbol{u})-\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}\geqslant J_\alpha(\hat{\boldsymbol{u}})-\frac{\alpha}{2}\hat{\boldsymbol{u}}^{\mathrm{T}}\hat{\boldsymbol{u}}=\int_0^1(1-\alpha)\nabla^{\mathrm{T}}\boldsymbol{Q}(t\hat{\boldsymbol{u}})\hat{\boldsymbol{u}}\mathrm{d}t$$

所以(P_3)的常向量最优控制即为(P_2)的最优控制．

反过来的情形可同样证明，即(P_2)的常向量最优控制即为(P_3)的最优控制，证毕.

定理2(必要条件) 给定的正实数 $\alpha\in(0,1)$，若对 $\|\boldsymbol{u}\|\leqslant 1$ 上的任一点，$\nabla^2 J_\alpha(\boldsymbol{u})$ 非正定，且 $\hat{\boldsymbol{u}}$ 是(P_3)的常向量最优控制，则有

$$(1-\alpha)\|\nabla Q(\hat{\boldsymbol{u}})\|\geqslant\alpha \tag{12}$$

证明 考虑关于变量 $\boldsymbol{u}$ 在单位球 U 上的正定二次优化问题

$$\min_{u\in U}\left\{(1-\alpha)\nabla^{\mathrm{T}}Q(\hat{\boldsymbol{x}}(1))\boldsymbol{u}+\frac{\alpha}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}\right\} \tag{13}$$

因 $\hat{\boldsymbol{u}}$ 是(P_3)的常向量最优控制，则由式(8)知，$\hat{\boldsymbol{u}}$ 是正定二次优化式(13)的唯一的最优解. 又因 $\nabla^2 J_\alpha(\hat{\boldsymbol{u}})$ 有至少1个特征值不大于零，所以 $\hat{\boldsymbol{u}}$ 位于单位球面上. 这样，由 canonical 倒向微分流方法[3]可推出 $(1-\alpha)\|\nabla Q(\hat{\boldsymbol{x}}(1))\|\geqslant\alpha$. 而由式(9)有 $\hat{\boldsymbol{u}}=\hat{\boldsymbol{x}}(1)$，所以有 $(1-\alpha)\|\nabla Q(\hat{\boldsymbol{u}})\|\geqslant\alpha$，证毕.

以下把优化问题的状态变量和与之等价的控制变量统一用字母 u 表示，举例说明上述定理在求解约束全局优化中的应用.

例1 考虑以下 $\mathbf{R}^1$ 上的非凸的全局优化问题：$\min Q(u)=-\frac{1}{12}u^4-u^2+u$, s. t. $u^2\leqslant 1$.

解 首先，对于摄动参数 α，由式(10)有 $J_\alpha(u)=\frac{\alpha}{2}u^2+(1-\alpha)\left(-\frac{u^4}{12}-u^2+u\right)$ 以及 $J''_\alpha(u)=\alpha+(1-\alpha)(-u^2-2)$. 接着，选取 α，满足 $\alpha+(1-\alpha)(-u^2-2)<0$, $\forall u^2\leqslant 1$，得到 $\alpha=\frac{2}{3}$. 运用定理1可知此问题的最优点满足 $|u|=1$. 再由定理2给出的最优点的必要性条件应有 $|Q'(u)|=\left|-\frac{1}{3}u^3-2u+1\right|\geqslant\frac{\frac{2}{3}}{1-\frac{2}{3}}=2$，由此立刻得到此问题的最优点 $u=-1$.

事实上 $u=1$ 是 $Q(u)=-\frac{1}{12}u^4-u^2+u$ 的局部极小点. 所以，适当选择摄动参数，利用必要性条件 $(1-\alpha)\|\nabla Q(\hat{u})\|\geqslant\alpha$ 可剔除一些原问题的局部极小点.

3 球约束下非凸二次优化的最优解的充要条件

考虑以下球约束下非凸二次全局优化问题

$$\min Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$$
$$\text{s. t. }\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}\leqslant 1 \tag{14}$$

式中，$\boldsymbol{A}$ 是半正定矩阵且 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$，$\boldsymbol{f}\neq 0$. 并设 A 至少有1个特征根 λ 小于零. 这样 $\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}\leqslant 1$ 的极小点只能在 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}=1$ 上取到，否则 $\boldsymbol{A}>0$. 取定 $h^*=-\min\limits_{\lambda<0}\lambda$，$\alpha^*=h^*/(1+h^*)$，则有 $h^*>0$，$0<\alpha^*<1$. 则与上述二次全局优化问题[6]等价的最优控制问题(P_2)的关于 α^* 的 Gurman 摄动变换为

$$(P_3):\min J_{\alpha^*}(\boldsymbol{u})=\left\{\int_0^1[(1-\alpha^*)(\boldsymbol{A}\boldsymbol{u}t-\boldsymbol{f})^{\mathrm{T}}\boldsymbol{u}+(\alpha^*/2)\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}]\mathrm{d}t\right\}=(1-\alpha^*)\cdot\left(\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}\right)+(\alpha^*/2)\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u} \tag{15}$$

因为 $\alpha^*\boldsymbol{I}+(1-\alpha^*)\boldsymbol{A}$(其中 $\boldsymbol{I}$ 为单位矩阵)有1个特征根为 $\alpha^*+(1-\alpha^*)\min\limits_{\lambda<0}\lambda$，而 $\alpha^*=\frac{-\min\limits_{\lambda<0}\lambda}{1-\min\limits_{\lambda<0}\lambda}$，所以 $\alpha^*+(1-\alpha^*)\min\limits_{\lambda<0}\lambda=0$. 这样 $\nabla^2 J_{\alpha^*}(\boldsymbol{u})=\alpha^*\boldsymbol{I}+(1-\alpha^*)\boldsymbol{A}$ 就为不定矩阵. 而 $J_{\alpha^*}(\boldsymbol{u})$ 的最小点 $\boldsymbol{u}^*$ 位于 $\|\boldsymbol{u}\|=1$ 上. 由等价性(定理1)可知，$\boldsymbol{u}^*$ 也是 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 上的最小点.

由于 $\boldsymbol{f}\neq 0$，假设以下2组向量所成矩阵的秩不相等，即

$$\operatorname{rank}(\boldsymbol{A}+(-\min_{\lambda<0}\lambda)\boldsymbol{I})\neq\operatorname{rank}(\boldsymbol{A}+(-\min_{\lambda<0}\lambda)\boldsymbol{I},\boldsymbol{f}) \tag{16}$$

反之则可利用正交变换降低决策变量的维数，得到新问题以求解原问题.

在假设式(16)满足的情形下，对上述 $\boldsymbol{u}^*$ 由定理2得到的不等式是严格的，即 $\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|>\frac{\alpha^*}{1-\alpha^*}=h^*=-\min\limits_{\lambda<0}\lambda$. 否则则有 $\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|=-\min\limits_{\lambda<0}\lambda$. 由于 $\boldsymbol{u}^*$ 也是 $Q(\boldsymbol{u})$ 的最小点(定理1)，从而 $\boldsymbol{u}^*$ 是球约束下非凸二次全局优化问题的 KKT 点，于是可推知，存在 $\hat{\rho}=-\min\limits_{\lambda<0}\lambda$，满足 $\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}+\hat{\rho}\boldsymbol{u}^*=0$，$(\boldsymbol{u}^*)^{\mathrm{T}}\cdot\boldsymbol{u}^*=1$，而这与式(16)矛盾，因为式(16)表明 $[\boldsymbol{A}+(-\min\limits_{\lambda<0}\lambda)\boldsymbol{I}]\boldsymbol{u}^*=\boldsymbol{f}$ 无解. 所以在式(16)条件下，若记 $\rho^*=\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|$，应有 $\rho^*>-\min\limits_{\lambda<0}\lambda$. 此时必要性条件也是充分的，因为 $\rho^*>-\min\limits_{\lambda<0}\lambda$ 蕴含 $\boldsymbol{A}+\rho^*\boldsymbol{I}>0$，而由对偶微分流理论[3]，若存在 $\rho^*>-\min\limits_{\lambda<0}\lambda$，使得

$$\begin{cases}\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}+\rho^*\boldsymbol{u}^*=0\\(\boldsymbol{u}^*)^{\mathrm{T}}\boldsymbol{u}^*=1\\\boldsymbol{A}+\rho^*\boldsymbol{I}>0\end{cases} \tag{17}$$

则 $\boldsymbol{u}^*=(\boldsymbol{A}+\rho^*\boldsymbol{I})^{-1}\boldsymbol{f}$ 是 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 上的最小点.

定理 3(充分必要条件)　设 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$，$\boldsymbol{f}\neq 0$，并设 $\boldsymbol{A}$ 至少有 1 个特征根 λ 小于零. 在满足式(16)的情形下，$\boldsymbol{u}^*$ 是 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 的最小点的充要条件是

$$\begin{cases}\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|>-\min\limits_{\lambda<0}\lambda\\ \boldsymbol{u}^*=\dfrac{-(\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f})}{\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|}\end{cases}\tag{18}$$

证明　(1) 必要性. 记 $\alpha^*=-\dfrac{\min\limits_{\lambda<0}\lambda}{1-\min\limits_{\lambda<0}\lambda}$. $\boldsymbol{u}^*$ 是 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 上的最小点，则 $\boldsymbol{u}^*$ 也是 $\nabla^2 J_{\alpha^*}(\boldsymbol{u})$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 上的最小点(定理 1)，从而由定理 2 及式(16)和式(17)的说明可得 $\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|>\dfrac{\alpha^*}{1-\alpha^*}=-\min\limits_{\lambda<0}\lambda$.

同时，由 $\boldsymbol{u}^*$ 是 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 上的最小点可推知，$\boldsymbol{u}^*$ 是球约束下非凸二次全局优化问题式(14)的 KKT 点. 因为 $\boldsymbol{A}$ 是非正定矩阵，所以最小点 $\boldsymbol{u}^*$ 在边界上，记 $\rho^*=\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|$，就得到 $\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}+\rho^*\boldsymbol{u}^*=0$，$(\boldsymbol{u}^*)^{\mathrm{T}}\boldsymbol{u}^*=1$，而这表明 $\boldsymbol{u}^*=\dfrac{-(\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f})}{\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|}$.

(2) 充分性. 记 $\rho^*=\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|$，由条件 $\rho^*>-\min\limits_{\lambda<0}\lambda$，$\boldsymbol{u}^*=\dfrac{-(\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f})}{\|\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}\|}$ 可推知 $\boldsymbol{A}+\rho^*\boldsymbol{I}>0$ 及 $\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}+\rho^*\boldsymbol{u}^*=0$，$(\boldsymbol{u}^*)^{\mathrm{T}}\boldsymbol{u}^*=1$. 这样，由对偶微分流理论[3]可推知，$\boldsymbol{u}^*$ 是 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 上的最小点.

推论 1　设 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$，$\boldsymbol{f}\neq 0$，并设 $\boldsymbol{A}$ 至少有 1 个特征根 λ 小于零. 在式(16)满足的情形下，$Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 有唯一的最小点.

证明　由定理 3 的必要性可得到 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 上的最小点 $\boldsymbol{u}^*$ 须满足 $\boldsymbol{A}\boldsymbol{u}^*-\boldsymbol{f}+\boldsymbol{\rho}^*\boldsymbol{u}^*=0$，$(\boldsymbol{u}^*)^{\mathrm{T}}\boldsymbol{u}^*=1$，$\rho^*>-\min\limits_{\lambda<0}\lambda$.

但由偶微分流理论[3]可推知，所需参数 ρ^* 是唯一存在的，因而 $Q(\boldsymbol{u})=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}-\boldsymbol{f}^{\mathrm{T}}\boldsymbol{u}$ 在 $\|\boldsymbol{u}\|\leqslant 1$ 有唯一的最小点 $\boldsymbol{u}^*=(\boldsymbol{A}+\rho^*\boldsymbol{I})^{-1}\boldsymbol{f}$.

例 2　考虑以下球约束下关于 u_1，u_2 的二维非凸全局优化问题：(Q^*)：$\min\ Q(\boldsymbol{u})=-u_1-u_2-u_1u_2-u_1-u_2$，s. t. $u_1^2+u_2^2\leqslant 1$.

解　由于 $\nabla Q=\begin{pmatrix}-2u_1-u_2+1\\-2u_2-u_{21}+1\end{pmatrix}$，$\nabla^2 Q=\begin{pmatrix}-2&-1\\-1&-2\end{pmatrix}<0$，可见 $Q(\boldsymbol{u})$ 为凹函数，上述全局优化问题的最优点必在边界 $u_1^2+u_2^2=1$ 上取得. 记 $\boldsymbol{A}=\begin{pmatrix}-2&-1\\-1&-2\end{pmatrix}$，$\boldsymbol{f}=\begin{pmatrix}1\\1\end{pmatrix}$. $\boldsymbol{A}$ 的特征值 $\lambda_1=-3$，$\lambda_2=-1$，选取 $\alpha^*=\dfrac{-(-3)}{1-(-3)}=\dfrac{3}{4}$. 易知 $\mathrm{rank}(\boldsymbol{A}+3\boldsymbol{I})=1\neq 2=\mathrm{rank}(\boldsymbol{A}+3\boldsymbol{I},\boldsymbol{f})$. 应用定理 3，需要求取 $\rho^*>-\lambda_1=3$，使得

$$\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{A}+\rho^*\boldsymbol{I})^{-2}\boldsymbol{f}=1.\tag{19}$$

以下 $\boldsymbol{A}$ 进行正交分解，把矩阵等式(19)转化为代数等式. 由于

$$\boldsymbol{A}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{Q}=\begin{pmatrix}\frac{\sqrt{2}}{2}&\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}-3&0\\0&-1\end{pmatrix}\begin{pmatrix}\frac{\sqrt{2}}{2}&\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}\end{pmatrix},$$

$$\hat{\boldsymbol{f}}^{\mathrm{T}}=\boldsymbol{f}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}=(1\quad 1)\begin{pmatrix}\frac{\sqrt{2}}{2}&\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}\end{pmatrix}=(\sqrt{2},0),$$

式中，$\boldsymbol{\Lambda}$ 为 $\boldsymbol{A}$ 正交分解后得到的由 2 个特征根 λ_1，λ_2 组成的对角阵. 这样，由

$$\begin{aligned}\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{A}+\rho^*\boldsymbol{I})^{-2}\boldsymbol{f}&=\hat{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{\Lambda}+\rho\boldsymbol{I})^{-2}\hat{\boldsymbol{f}}\\&=(\sqrt{2},0)\begin{pmatrix}\frac{1}{(\rho-3)^2}&0\\0&\frac{1}{(\rho-1)^2}\end{pmatrix}\begin{pmatrix}\sqrt{2}\\0\end{pmatrix}\\&=\frac{2}{(\rho-3)^2}=1\end{aligned}$$

式中：2 个参数解 $\rho=3\pm\sqrt{2}$，由于 $\rho=3+\sqrt{2}>-\lambda_1=3$，应取

$$\rho^*=3+\sqrt{2}\tag{20}$$

于是得到全局优化问题(Q^*)的最优点 $\boldsymbol{u}^*=(\boldsymbol{A}+\rho^*\boldsymbol{I})^{-1}\boldsymbol{f}=\begin{pmatrix}\sqrt{2}+1&-1\\-1&\sqrt{2}+1\end{pmatrix}^{-1}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{pmatrix}$，所以根据定理 3 给出的充要条件以及推论 1 知 $\boldsymbol{u}^*=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)^{\mathrm{T}}$ 是原问题(Q^*)的唯一最优点.

(下转第 798 页)