文章编号：0253-374X(2011)09-1373-04
DOI：10.3969/j.issn.0253-374x.2011.09.022

Canonical 对偶理论在一类多项式全局优化中的应用

朱经浩，谭素娥
(同济大学 数学系，上海 200092)

摘要：用 Canonical 对偶理论，讨论一类高阶多项式全局最优化问题的求解. 首先将无约束多项式全局优化问题转换成箱体约束下的多项式全局优化问题，之后通过构造非线性变换对偶函数及相应的共轭函数，得到原问题的 Canonical 对偶问题. 进一步通过求解对偶问题的最优解，导出原多项式全局优化问题的最优解，并给出对偶问题是凹函数的证明. 最后应用所得方法，计算一个二元 6 次多项式全局最优化实例.

关键词：Canonical 对偶理论；全局优化；高阶多元多项式
中图分类号：O 221.2　　**文献标识码**：A

Application of Canonical Duality Theory to Global Optimization with Polynomials

ZHU Jinghao, *TAN Su'e*
(Department of Mathematics, Tongji University, Shanghai 200092, China)

Abstract: A class of global optimization problem with polynomial is investigated with canonical duality theory. The unconstrained polynomial optimization problem is transformed into box constrained global optimization. The canonical dual function is defied for a solution to the original global optimization with polynomial problem by solving the dual problem. In addition, the dual problem proves to be a concave optimization. Finally, an example about binary six-order polynomial global optimization is illustrated.

Key words: Canonical duality theory; global optimization; high-order multivariate polynomials

考虑以下无约束多项式全局优化问题：

$$\min P(\boldsymbol{x}) = \sum_{l=2}^{m}\sum_{i=1}^{n} x_i^{2l} - \sum_{l\leqslant i,j\leqslant n} x_i x_j - \boldsymbol{\gamma}^{\mathrm{T}} - \boldsymbol{x},$$

$$\text{s.t.}\quad \boldsymbol{x} = (x_1,\cdots,x_n)^{\mathrm{T}} \in \mathrm{R}^n. \qquad (1)$$

其中，$\boldsymbol{\gamma}$ 为非零的 n 维向量. 利用参考文献[1]中的结果，可知上述问题(1)的最优点必含于一个以原点为中心的闭球体内，由于闭球体可内切在一个正箱体内，可得到正数 d 使得问题(1)的最优点必含于箱体$\{\boldsymbol{x}\in \mathrm{R}^n \mid x_i^2\leqslant d, i=1,2,\cdots,n\}$内. 于是易知问题(1)等价于以下箱体约束下的多项式全局优化问题：

$$\min P(\boldsymbol{x}) = \sum_{l=2}^{m}\sum_{i=1}^{n} - \sum_{1\leqslant i,j\leqslant n} x_i x_j - \boldsymbol{\gamma}^{\mathrm{T}},$$

$$\text{s.t.}\quad x_i^2 \leqslant d,\quad i = 1,2,\cdots,n \qquad (2)$$

其中：$m, l \in N, m \geqslant 2; \boldsymbol{\gamma}\neq 0; d>0$.

关于多项式 $P(\boldsymbol{x}) = \sum_{i=1}^{n} x_i^{2m} + g(x_1,\cdots,x_n)$的无约束多项式全局优化问题是一个长期困扰数学家的热点问题[2-3]. 这里，$m \geqslant 2$，而 $g(x_1,\cdots,x_n)$是一个最高次数等于或低于 $2m-1$ 的多项式.

下面利用 Canonical 对偶理论求解问题(2). Canonical 对偶理论是一个研究各类非凸问题的重要数学方法[4-6]. 近年来，在全局优化问题的研究中，尤其在处理许多 NP-难的问题中，人们应用 Canonical 对偶理论得到了一系列重要成果[7-9].

1　全局优化的 Canonical 对偶方法

利用 Canonical 对偶理论[4]，文献[9]讨论了以下有约束的全局最优化问题：

$$(P): \min\ \{P(\boldsymbol{x}) = (1/2)\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{x}\rangle - \langle \boldsymbol{x}, \boldsymbol{f}\rangle + W(\boldsymbol{x})\}$$

$$\text{s.t.}\quad x_i^2 \leqslant 1,\quad i = 1,\cdots,n \qquad (3)$$

其中：$\boldsymbol{A}\in \mathrm{R}^{n\times n}$，是一对称不定矩阵；$\boldsymbol{f}\in \mathrm{R}^n$，$\boldsymbol{f}\neq 0$；$W(\boldsymbol{x})$是一个 2 阶连续可微函数. 其主要想法为[9]：利用几何变换 $\boldsymbol{\xi} = \Lambda(\boldsymbol{x}) = (x_1^2-1,\cdots,x_n^2-1)^{\mathrm{T}}$，$\boldsymbol{\xi}\in \varepsilon_a = \{\boldsymbol{\xi}: -d \leqslant \xi_i \leqslant 0, i=1,\cdots n\}$，并构造 Canonical

收稿日期：2010-06-10
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和数学规划. E-mail：jinghaok@online.sh.cn
通讯作者：谭素娥(1987—)，女，硕士生，主要研究方向为最优控制和数学规划. E-mail：tansue126@126.com

函数 $V(\boldsymbol{\xi})$，拓展 $W(\boldsymbol{x})$成为 R^n 上的广义函数，使得

$$W(\boldsymbol{x}) = \begin{cases} V(\Lambda(\boldsymbol{x})), & -1 \leqslant x_i^2 - 1 \leqslant 0, \\ & \forall\, i \in \{1,2,\cdots,n\} \\ \infty & \text{其他} \end{cases} \quad (4)$$

接着，利用次微分包含定义 Canonical 对偶变量 $\boldsymbol{\zeta} \in \partial - V(\boldsymbol{\xi})$：$\boldsymbol{\xi} \in \varepsilon_a \to \boldsymbol{\zeta} \in \boldsymbol{S}_a$，这等价于满足 KKT 条件

$$\Lambda(\boldsymbol{x}) = (x_1^2 - 1, \cdots, x_n^2 - 1)^{\mathrm{T}} \leqslant (0,\cdots,0)^{\mathrm{T}}$$

$$\boldsymbol{\zeta} = (\zeta_1,\cdots,\zeta_n) \geqslant 0$$

$$(x_1^2 - 1)\zeta_i = 0, \quad i = 1,2,\cdots,n$$

同时，构造 $V(\boldsymbol{\xi})$的 Canonical 共轭函数

$$V^{\#}(\boldsymbol{\zeta}) = \sup_{-1 \leqslant \xi_i \leqslant 0,\, i=1,2,\cdots,n} \{\boldsymbol{\xi}^{\mathrm{T}} - \boldsymbol{\zeta} - V(\boldsymbol{\xi})\}, \quad \boldsymbol{\zeta} \in \boldsymbol{S}_a \quad (5)$$

进而引入全互补函数

$$\boldsymbol{\Xi}(\boldsymbol{x},\boldsymbol{\zeta}) = \boldsymbol{\zeta}^{\mathrm{T}}\Lambda(\boldsymbol{x}) - V^{\#}(\boldsymbol{\zeta}) - U(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}/2(\boldsymbol{A} + 2\,\mathrm{diag}(\boldsymbol{\zeta}))\boldsymbol{x} - \sum_{i=1}^{n}\zeta_i - \boldsymbol{f}^{\mathrm{T}}\boldsymbol{x} - V^{\#}(\boldsymbol{\zeta}) \quad (6)$$

其中，$U(\boldsymbol{x}) = -(\boldsymbol{x}^{\mathrm{T}}/2)\boldsymbol{A}\boldsymbol{x} + \boldsymbol{f}^{\mathrm{T}}\boldsymbol{x}$. 考虑对偶可行空间

$$\boldsymbol{S}^{+} = \{\boldsymbol{\zeta} \in \boldsymbol{S}_a \mid (\boldsymbol{A} + 2(\mathrm{diag}(\boldsymbol{\zeta})) > 0\} \quad (7)$$

导出问题(P)的对偶函数

$$P^d(\boldsymbol{\zeta}) = -(\boldsymbol{f}^{\mathrm{T}}/2)(\boldsymbol{A} + 2\,\mathrm{diag}(\boldsymbol{\zeta}))^{-1}\boldsymbol{f} - \sum_{i=1}^{n}\zeta_i - V^{\#}(\boldsymbol{\zeta}) \quad (8)$$

并建立问题(P)的 Canonical 对偶问题

$$(P^d): \max_{\zeta \in S^{+}} P^d(\boldsymbol{\zeta}) \quad (9)$$

引用如下 Canonical 对偶定理[9]：

定理 1 如果 $\hat{\boldsymbol{\zeta}}$ 是 Canonical 对偶问题(P^d)的一个最优解，则 $\hat{\boldsymbol{x}} = (\boldsymbol{A} + 2\,\mathrm{diag}(\boldsymbol{\zeta}))^{-1}\boldsymbol{f}$ 是原问题(P)的最优解，并且有

$$\min_{x_i^2 \leqslant 1,\, i=1,2,\cdots,n} P(\boldsymbol{x}) = P(\hat{\boldsymbol{x}}) = P^d(\hat{\boldsymbol{\zeta}}) = \max_{\zeta \in S^{+}} P^d(\boldsymbol{\zeta}) \quad (10)$$

2 一类箱体约束下的多项式最优化问题的求解

下面应用 Canonical 对偶理论求解箱体约束下的多项式最优化问题(2). 首先，目标函数可写为

$$P(\boldsymbol{x}) = \sum_{l=2}^{m}\sum_{i=1}^{n} x_i^{2l} - \sum_{l \leqslant i,j \leqslant n} x_i x_j - \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x} = W(\boldsymbol{x}) - U(\boldsymbol{x}) \quad (11)$$

其中：$U(\boldsymbol{x}) = \sum_{l \leqslant i,j \leqslant n} x_i x_j + \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x}$，$W(\boldsymbol{x}) = \sum_{l=2}^{m}\sum_{i=1}^{n} x_i^{2l}$.

引入非线性几何变换

$$\Lambda(\boldsymbol{x}): \mathrm{R}^n \to \varepsilon_a = \{\boldsymbol{\xi}: -d \leqslant \xi_i \leqslant 0, i = 1,\cdots n\}$$

$$\boldsymbol{\xi} = \Lambda(\boldsymbol{x}) = (x_1^2 - d, \cdots, x_n^2 - d)^{\mathrm{T}} \quad (12)$$

并在 ε_a上定义函数 $V(\boldsymbol{\xi}) = \sum_{l=2}^{m}\sum_{i=1}^{n}(\xi_i + d)^l$，则在箱体 $\Lambda(\boldsymbol{x}) = (x_1^2 - d,\cdots,x_n^2 - d)^{\mathrm{T}} \leqslant (0,\cdots,0)^{\mathrm{T}}$ 上有 $W(\boldsymbol{x}) = V(\Lambda\boldsymbol{x})$拓展$W(\boldsymbol{x})$成为 R^n 上的函数如下：

$$W(\boldsymbol{x}) = \begin{cases} V(\Lambda(\boldsymbol{x})), & x_i^2 \leqslant d, i = 1,2,\cdots,n \\ \infty, & \text{其他} \end{cases} = \begin{cases} \displaystyle\sum_{l=2}^{m}\sum_{i=1}^{n}(\xi_i + d)^l, & \boldsymbol{\xi} \leqslant 0 \\ \infty, & \text{其他} \end{cases} \quad (13)$$

另一方面，在 $\varepsilon_a = \{\boldsymbol{\xi}: -d \leqslant \xi_i \leqslant 0, i = 1,\cdots n\}$ 内，定义如下对偶变量：

$$\boldsymbol{\zeta} = \nabla V(\boldsymbol{\xi}) = \begin{pmatrix} \displaystyle\sum_{l=2}^{m} l(\xi_1 + d)^{l-1} \\ \vdots \\ \displaystyle\sum_{l=2}^{m} l(\xi_n + d)^{l-1} \end{pmatrix} \quad (14)$$

和对偶空间 $\boldsymbol{S}_a = \{(\zeta_1,\cdots,\zeta_n)^{\mathrm{T}} \mid 0 \leqslant \zeta_1 \leqslant \sum_{l=2}^{m} l\, d^{l-1}\}$ 形成对偶映射

$$\boldsymbol{\zeta} = g(\boldsymbol{\xi}) = \nabla V(\boldsymbol{\xi}): \varepsilon_a \to \boldsymbol{S}_a \quad (15)$$

定理 2 对偶映射 $\boldsymbol{\zeta} = g(\boldsymbol{\xi}): \varepsilon_a \to \boldsymbol{S}_a = \{(\zeta_1,\cdots,\zeta_N)^{\mathrm{T}} \mid 0 \leqslant \zeta_1 \leqslant \sum_{l=2}^{m} l\, d^{l-1}\}$ 是可逆的.

证明 由于$\boldsymbol{\xi} \in \varepsilon_a$，所以其每个分量满足 $\xi_i \leqslant 0$，且 $\xi_i + d \geqslant 0$. 于是，当相异的 $\xi', \xi'' \in \varepsilon_a$，必有某个 $i \in \{1,2,\cdots,n\}$使得 $\xi'_i \neq \xi''_i$，则有 $\zeta'_i = \sum_{l=2}^{m} l(\xi'_i + d)^{l-1} \neq \sum_{l=2}^{m} l(\xi''_i + d)^{l-1} = \zeta''_i$. 另外易见，对每个$\boldsymbol{\xi} \in \varepsilon_a$，在对偶映射下，其像 $\boldsymbol{\zeta} = \nabla V(\boldsymbol{\xi})$的每个分量$\zeta_i$满足 $0 \leqslant \zeta_i = \sum_{l=2}^{m} l(\xi_i + d)^{l-1} \leqslant \sum_{l=2}^{m} l\, d^{l-1}$. 因此，对偶映射是 $\varepsilon_a \to \boldsymbol{S}_a$ 的可逆映射. 证毕.

由定理 2 和文献[9]中的有关论述可知，$V(\boldsymbol{\xi}) = \sum_{l=2}^{m}\sum_{i=1}^{n}(\xi_i + d)^l$是 $\varepsilon_a = \{\boldsymbol{\xi}: -d \leqslant \xi_i \leqslant 0, i = 1,\cdots n\}$上的 Canonical 函数. 下面寻求 $V(\boldsymbol{\xi})$的 Canonical 共轭 $V^{\#}(\boldsymbol{\zeta})$，$\boldsymbol{\zeta} \in \boldsymbol{S}_a$. 由第 1 节所引述的定义，对每个 $\boldsymbol{\zeta} \in \boldsymbol{S}_a$，有

$$V^{\#}(\boldsymbol{\zeta}) = \sup_{\xi \in \varepsilon_a}\{\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\zeta} - V(\boldsymbol{\xi})\} = \sup_{\xi \in \varepsilon_a}\left\{\sum_{i=1}^{n}\left[(\zeta_i\xi_i) - \sum_{l=2}^{m}(\xi_i + d)^l\right]\right\} \quad (16)$$

由于对偶映射 $g(\boldsymbol{\xi})$是从 $\varepsilon_a \to \boldsymbol{S}_a$ 的可逆映射(定

理 2），所以，对每个 $\boldsymbol{\zeta}\in\boldsymbol{S}_a$

$$\nabla\boldsymbol{\xi}\left\{\sum_{i=1}^{n}(\zeta_i\xi_i)-\sum_{l=2}^{m}\sum_{i=1}^{n}(\xi_i+d)^l\right\}=\begin{pmatrix}\zeta_1-\sum_{l=2}^{m}l(\xi_1+d)^{l-1}\\ \vdots\\ \zeta_n-\sum_{l=2}^{m}l(\xi_n+d)^{l-1}\end{pmatrix}=0 \tag{17}$$

在 ε_a 上有唯一解 $\boldsymbol{\xi}=g^{-1}(\boldsymbol{\zeta})$. 且因

$$\nabla^2\boldsymbol{\xi}\left\{\sum_{i=1}^{n}(\zeta_i\xi_i)-\sum_{l=2}^{m}\sum_{i=1}^{n}(\xi_i+d)^l\right\}=\operatorname{diag}\left(-\sum_{l=2}^{m}l(l-1)(\xi_1+d)^{l-2},\cdots,-\sum_{l=2}^{m}l(l-1)(\xi_n+d)^{l-2}\right)\leqslant 0 \tag{18}$$

又，对每个 $i\in\{1,2,\cdots,n\}$，只要 $\xi_i>-d$，就有 $-\sum_{l=2}^{m}l(l-1)(\xi_1+d)^{l-2}<0$. 所以，当 $\boldsymbol{\xi}=g^{-1}(\boldsymbol{\zeta})\in\varepsilon_a$ 时，函数 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\zeta}-V(\boldsymbol{\xi})$ 取到最大值. 代入式(16)，得

$$V^{\#}(\boldsymbol{\zeta})=\sup_{\boldsymbol{\xi}\in\varepsilon_a}\{\langle\boldsymbol{\xi},\boldsymbol{\zeta}\rangle-V(\boldsymbol{\xi})\}=\boldsymbol{\zeta}^{\mathrm{T}}g^{-1}(\boldsymbol{\zeta})-V(g^{-1}(\boldsymbol{\zeta})) \tag{19}$$

最后，引入全互补函数：对于 $(\boldsymbol{x},\boldsymbol{\zeta})\in\mathrm{R}^n\times\boldsymbol{S}_a$，有

$$\boldsymbol{\Xi}(\boldsymbol{x},\boldsymbol{\zeta})=\boldsymbol{\zeta}^{\mathrm{T}}\Lambda(\boldsymbol{x})-V^{\#}(\boldsymbol{\zeta})-U(\boldsymbol{x})=\sum_{i=1}^{n}\zeta_i(x_i^2-d)-V^{\#}(\boldsymbol{\zeta})-\frac{1}{2}\sum_{1\leqslant i,j\leqslant n}2x_ix_j-\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x} \tag{20}$$

并确定 Canonical 对偶函数

$$P^d(\boldsymbol{\zeta})=\{\boldsymbol{\Xi}(\boldsymbol{x},\boldsymbol{\zeta}):\nabla_x\boldsymbol{\Xi}(\boldsymbol{x},\boldsymbol{\zeta})=0\}=\left\{\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}(-\boldsymbol{G}+2\operatorname{diag}(\boldsymbol{\zeta}))\boldsymbol{x}-\sum_{i=1}^{n}d\zeta_i-V^{\#}(\boldsymbol{\zeta})-\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x}\,\middle|\,\text{s.}(-\boldsymbol{G}+2\operatorname{diag}(\boldsymbol{\zeta}))\boldsymbol{x}=\boldsymbol{\gamma}\right\} \tag{21}$$

其中，$\boldsymbol{G}$ 为 $n\times n$ 矩阵，$\boldsymbol{G}=\begin{bmatrix}2&\cdots&2\\ \vdots&&\vdots\\ 2&\cdots&2\end{bmatrix}$. 得到当 $\boldsymbol{\zeta}\in\boldsymbol{S}^+=\{\boldsymbol{\zeta}\in\boldsymbol{S}_a\mid(-\boldsymbol{G}+2(\operatorname{diag}(\boldsymbol{\zeta}))>0\}$ 时，Canonical 对偶函数

$$P^d(\boldsymbol{\zeta})=(-\boldsymbol{\gamma}^{\mathrm{T}}/2)(-\boldsymbol{G}+2\operatorname{diag}(\boldsymbol{\zeta}))^{-1}\boldsymbol{\gamma}-\sum_{i=1}^{n}d\zeta_i-V^{\#}(\boldsymbol{\zeta}) \tag{22}$$

由定理 1，若 $\hat{\boldsymbol{\zeta}}$ 是下列 Canonical 对偶问题的解：

$$(P^d):\max_{\boldsymbol{\zeta}\in\boldsymbol{S}^+}(-\boldsymbol{\gamma}^{\mathrm{T}}/2)(-\boldsymbol{G}+2\operatorname{diag}(\boldsymbol{\zeta}))^{-1}\boldsymbol{\gamma}-\sum_{i=1}^{n}d\zeta_i-V^{\#}(\boldsymbol{\zeta}) \tag{23}$$

则 $\hat{\boldsymbol{x}}=(-\boldsymbol{G}+2\operatorname{diag}(\hat{\boldsymbol{\zeta}}))^{-1}\boldsymbol{\gamma}$ 是问题(2)的最优解. 由此可得以下结果：

推论 1　若 $\hat{\boldsymbol{\zeta}}\in\boldsymbol{S}^+=\{\boldsymbol{\zeta}\in\boldsymbol{S}_a\mid(-\boldsymbol{G}+2(\operatorname{diag}(\boldsymbol{\zeta}))>0\}$ 是 $P^d(\boldsymbol{\zeta})=(-\boldsymbol{\gamma}^{\mathrm{T}}/2)(-\boldsymbol{G}+2\operatorname{Diag}(\boldsymbol{\zeta}))^{-1}\boldsymbol{\gamma}-\sum_{i=1}^{n}d\zeta_i-V^{\#}(\boldsymbol{\zeta})$ 的稳定点，则 $\hat{\boldsymbol{\zeta}}$ 必为式(22)最优解.

定理 3　$P^d(\boldsymbol{\zeta})=(-\boldsymbol{\gamma}^{\mathrm{T}}/2)(-\boldsymbol{G}+2\operatorname{diag}(\boldsymbol{\zeta}))^{-1}\boldsymbol{\gamma}-\sum_{i=1}^{n}d\zeta_i-V^{\#}(\boldsymbol{\zeta})$ 在 $\boldsymbol{\zeta}\in\boldsymbol{S}^+=\{\boldsymbol{\zeta}\in\boldsymbol{S}_a\mid(-\boldsymbol{G}+2(\operatorname{diag}(\boldsymbol{\zeta}))>0$ 是严格凹函数.

证明　设 $Q(\boldsymbol{\zeta})=(-\boldsymbol{\gamma}^{\mathrm{T}}/2)(-\boldsymbol{G}+2\operatorname{diag}(\boldsymbol{\zeta}))^{-1}\boldsymbol{\gamma}-\sum_{i=1}^{n}d\zeta_i$，$P^d(\boldsymbol{\zeta})=Q(\boldsymbol{\zeta})-V^{\#}(\boldsymbol{\zeta})$，易知 $Q(\boldsymbol{\zeta})$ 在 $\boldsymbol{\zeta}\in\boldsymbol{S}^+=\{\boldsymbol{\zeta}\in\boldsymbol{S}_a\mid(-\boldsymbol{G}+2(\operatorname{diag}(\boldsymbol{\zeta}))>0$ 是凹函数，而 $V^{\#}(\boldsymbol{\zeta})=\sup_{\boldsymbol{\xi}\in\mathrm{R}^n}\{\langle\boldsymbol{\xi},\boldsymbol{\zeta}\rangle-V(\boldsymbol{\xi})\}=\boldsymbol{\zeta}^{\mathrm{T}}g^{-1}(\boldsymbol{\zeta})-V(g^{-1}(\boldsymbol{\zeta}))$，则有

$$\begin{aligned}\nabla V^{\#}(\boldsymbol{\zeta})&=g^{-1}(\boldsymbol{\zeta})+\frac{\partial g^{-1}(\boldsymbol{\zeta})^{\mathrm{T}}}{\partial\boldsymbol{\zeta}}\cdot\boldsymbol{\zeta}-\left[\nabla_{g^{-1}(\boldsymbol{\zeta})}V(g^{-1}(\boldsymbol{\zeta}))^{\mathrm{T}}\cdot\frac{\partial(g^{-1}(\boldsymbol{\zeta}))}{\partial\boldsymbol{\zeta}}\right]^{\mathrm{T}}=\\ &g^{-1}(\boldsymbol{\zeta})+\frac{\partial g^{-1}(\boldsymbol{\zeta})^{\mathrm{T}}}{\partial\boldsymbol{\zeta}}\cdot\boldsymbol{\zeta}-\left[\nabla_{\boldsymbol{\xi}}V(\boldsymbol{\xi})^{\mathrm{T}}\cdot\frac{\partial(\boldsymbol{\xi})}{\partial\boldsymbol{\zeta}}\right]^{\mathrm{T}}=\\ &g^{-1}(\boldsymbol{\zeta})+\frac{\partial\boldsymbol{\xi}^{\mathrm{T}}}{\partial\boldsymbol{\zeta}}\cdot\boldsymbol{\zeta}-\left(\boldsymbol{\zeta}^{\mathrm{T}}\cdot\frac{\partial\boldsymbol{\xi}}{\partial\boldsymbol{\zeta}}\right)^{\mathrm{T}}=\\ &g^{-1}(\boldsymbol{\zeta})+\frac{\partial\boldsymbol{\xi}^{\mathrm{T}}}{\partial\boldsymbol{\zeta}}\cdot\boldsymbol{\zeta}-\frac{\partial\boldsymbol{\xi}^{\mathrm{T}}}{\partial\boldsymbol{\zeta}}\cdot\boldsymbol{\zeta}=g^{-1}(\boldsymbol{\zeta})\end{aligned}$$

则

$$\nabla^2V^{\#}(\boldsymbol{\zeta})=\frac{\partial g^{-1}(\boldsymbol{\zeta})}{\partial\boldsymbol{\zeta}}$$

另一方面，$\zeta_i=\sum_{l=2}^{m}l(\xi_i+d)^{l-1}=g(\xi_i)$，$i=1,\cdots n$. 在 $\boldsymbol{\zeta}=g(\boldsymbol{\xi})$，两边同时对 $\boldsymbol{\zeta}$ 求导，得到 Jacobi 矩阵，$\boldsymbol{E}=\frac{\partial(g(\boldsymbol{\xi}))^{\mathrm{T}}}{\partial\boldsymbol{\xi}}\cdot\frac{\partial\boldsymbol{\xi}}{\partial\boldsymbol{\zeta}}=\frac{\partial(g(\boldsymbol{\xi}))^{\mathrm{T}}}{\partial\boldsymbol{\xi}}\cdot\frac{\partial(g^{-1}(\boldsymbol{\zeta}))}{\partial\boldsymbol{\zeta}}$，而 $\frac{\partial(g(\boldsymbol{\xi}))}{\partial\boldsymbol{\xi}}=\operatorname{diag}\left(\sum_{l=2}^{m}l(l-1)(\xi_i+d)^{l-2}\right)>0$. 不妨设 $\frac{\partial(g(\boldsymbol{\xi}))}{\partial\boldsymbol{\xi}}=\operatorname{diag}(d_1,d_2,\cdots,d_n)$，其中 $d_i=l(l-1)(\xi_i+d)^{l-2}>0$，$i=1,\cdots,n$. 由于对角阵的逆矩阵仍为对角阵，所以 $\frac{\partial(g^{-1}(\boldsymbol{\zeta}))}{\partial\boldsymbol{\zeta}}$ 也为对角阵，且 $\frac{\partial(g^{-1}(\boldsymbol{\zeta}))}{\partial\boldsymbol{\zeta}}=\operatorname{diag}\left(\frac{1}{d_1},\frac{1}{d_2},\cdots,\frac{1}{d_n}\right)>0$. 再结合前式 $\nabla^2V^{\#}(\boldsymbol{\zeta})=\frac{\partial g^{-1}(\boldsymbol{\zeta})}{\partial\boldsymbol{\zeta}}$，

得到$\nabla^2 V^{\#}(\boldsymbol{\zeta})>0$.可见,$V^{\#}(\boldsymbol{\zeta})$是严格凸函数,从而$-V^{\#}(\boldsymbol{\zeta})$是严格凹函数.所以,$P^d(\boldsymbol{\zeta})=Q(\boldsymbol{\zeta})-V^{\#}(\boldsymbol{\zeta})$是严格凹函数.

3　例子

下面以 $m=3$, $n=2$, $\boldsymbol{\gamma}=(1,2)^{\mathrm{T}}$ 为例进行讨论.此时问题(2)为

$$\min P(\boldsymbol{x}) = x_1^6+x_2^6+x_1^4+x_2^4-x_1^2-x_2^2-2x_1x_2-x_1-2x_2$$
$$\text{s.t.}\quad x_i^2\leqslant d,\quad i=1,2 \tag{24}$$

其中,d 的选择可依文献[1]得到.按其中记号,由于 $I_6=180$,$M_5=24$,可得 $d=\sigma_6=601.6$.这时

$$\boldsymbol{S}_{\mathrm{a}}=\{(\zeta_1,\zeta_2)^{\mathrm{T}}\mid 0\leqslant\zeta_i\leqslant 2d+3d^2\}=\{(\zeta_1,\zeta_2)^{\mathrm{T}}\mid 0\leqslant\zeta_i\leqslant 6.532\,0\times10^8\}$$

由式(20)可得$\boldsymbol{\Xi}(\boldsymbol{x},\boldsymbol{\zeta})=\sum_{i=1}^{n}\zeta_i(x_i^2-d)-V^{\#}(\boldsymbol{\zeta})-\frac{1}{2}\sum_{l\leqslant i,j\leqslant n}2x_ix_j-(1,2)^{\mathrm{T}}\boldsymbol{x}$.再由$\nabla_x\boldsymbol{\Xi}(\boldsymbol{x},\boldsymbol{\zeta})=0$可得到 Canonical 方程即为

$$\left.\begin{aligned}2\zeta_1x_1-2(x_1+x_2)=1\\2\zeta_1x_2-2(x_1+x_2)=2\end{aligned}\right\}\tag{25}$$

解此方程组得

$$\left.\begin{aligned}x_1&=(1+\zeta_2)/2(\zeta_1\zeta_2-\zeta_1-\zeta_2),\\x_2&=(2\zeta_2-1)/2(\zeta_1\zeta_2-\zeta_1-\zeta_1)\end{aligned}\right\}\tag{26}$$

而通过$-\boldsymbol{G}+2\operatorname{diag}(\boldsymbol{\zeta})=\begin{bmatrix}2\zeta_1-2&-2\\-2&2\zeta_2-2\end{bmatrix}$,可知$\boldsymbol{S}^+=\{(\zeta_1,\zeta_2)\mid\zeta_1>1,\zeta_1\zeta_2-\zeta_1-\zeta_2>0\}$.

另一方面,Canonical 对偶映射可表示为:对 $i=1,2$,$\zeta_i=3(\xi_i+d)^2+2(\xi_i+d)$,其逆映射为 $\xi_i=[-1+(1+3\zeta_i)^{1/2}]/3-d$.于是,有

$$V^{\#}(\boldsymbol{\zeta})=\sum_{i=1}^{2}\left[\zeta_i\cdot\frac{-1+(1+3\zeta_i)^{1/2}}{3}-d\zeta_i\right]-\sum_{l=2}^{3}\sum_{i=1}^{2}\left[\frac{-1+(1+3\zeta_i)^{1/2}}{3}\right]^l\tag{27}$$

将解得的 x_1,x_2代入式(21)中,得到

$$P^d(\boldsymbol{\zeta})=\frac{-\zeta_1\zeta_2^2-4\zeta_1^2\zeta_2+6\zeta_1\zeta_2+4\zeta_1^2+\zeta_2^2-\zeta_1-\zeta_2}{4(\zeta_1\zeta_2-\zeta_1-\zeta_2)^2}-\sum_{i=1}^{2}\zeta_i\frac{-1+(1+3\zeta_i)^{1/2}}{3}+\sum_{l=2}^{3}\sum_{i=1}^{2}\left(\frac{-1+(1+3\zeta_i)^{1/2}}{3}\right)^l\tag{28}$$

则由式(23),得到的相应于式(24)的 Canonical 对偶问题

$$(P^d):\max_{\text{s.t.}\,\boldsymbol{\zeta}\in\boldsymbol{S}^+}P^d(\boldsymbol{\zeta})\tag{29}$$

令$\nabla P^d(\boldsymbol{\zeta})=0$,解方程组得到2个稳定点:$\bar{\boldsymbol{\zeta}}=(1.264\,5,0.444\,4)^{\mathrm{T}}$,$\tilde{\boldsymbol{\zeta}}=(2.619\,3,3.224\,8)^{\mathrm{T}}$.易见$\bar{\boldsymbol{\zeta}}$ 和 $\tilde{\boldsymbol{\zeta}}$ 均属于 $\boldsymbol{S}_{\mathrm{a}}$.但是 $\bar{\boldsymbol{\zeta}}\notin\boldsymbol{S}^+$,故舍去.而 $\tilde{\boldsymbol{\zeta}}\in\boldsymbol{S}^+$.由定理3可知,$P^d(\boldsymbol{\zeta})$是严格凹函数,从而 $\tilde{\boldsymbol{\zeta}}=(2.619\,3,3.224\,8)^{\mathrm{T}}$ 为对偶问题(29)的最优解.再由定理1可知,$\hat{\boldsymbol{x}}=(0.811\,6,0.854\,3)^{\mathrm{T}}$ 是原问题的最优解,且得到最优值 $P(\hat{\boldsymbol{x}})=P^d(\hat{\boldsymbol{\zeta}})=-3.660\,1$.

参考文献:

[1] ZHU Jinghao, ZHANG Xi. On global optimizations with polynomials[J]. Optimization Letters, 2008(2): 239.

[2] Lasserre J B. Global optimization with polynomials and the problem of moments[J]. SIAM J Optim, 2001, 11(3): 796.

[3] Shor N Z. Non-differentiable optimization and polynomial problems[M]. Dordrecht: Kluwer, 1998.

[4] Gao D Y. Duality principles in nonconvex systems: theory, methods and applications [M]. London: Kluwer Academic Publishers, 2000.

[5] Gao D Y. Bi-duality in nonconvex optimization [C] // Encyclopedia of Optimization. London: Kluwer Academic Publishers, 2001: 477-482.

[6] Thach P T. Global optimality criterion and a duality with a zero gap in nonconvex optimization[J]. SIAM J Math Anal, 1993, 24(6): 1537.

[7] Floudas C A, Visweswaran V. A primal-relaxed dual global optimization approach[J]. Journal of Optimization Theory and Its Applications, 1993, 78(2): 187.

[8] Rockafellar R T. Conjugate duality and optimization. SIAM[M]. Bristol: Arrowsmith J W Ltd, 1974.

[9] Gao D Y. Adances in canonical duality theory with applications to global optimization[C] // Proceedings Foundations of Computer-Aided Process Operations. Cambridge: Omni Press, 2008: 73-82.