

一类具有连续变量的全局最优化问题的凸化方法

李博, 杜杰, 万立娟
(青岛科技大学数理学院, 山东 青岛 266061)

摘要: 本文研究了一类非凸最优化问题的凸化方法与最优性条件的问题. 利用构造含有参数的函数变换方法, 将具有次正定性质的目标函数凸化, 并获得了这一类非凸优化问题全局最优解的充要条件, 推广了凸化方法在求解全局最优化问题方面的应用.

关键词: 全局优化; 次正定函数; 凸化; 充要条件

MR(2010) 主题分类号: 90C30; 65K05

中图分类号: O221.2

文献标识码: A

文章编号: 0255-7797(2016)04-0851-08

1 引言

全局最优化问题在工程设计、生产管理、资源分配等领域的应用越来越广泛. 近些年来, 许多全局最优化理论与方法相继出现, 并得到广泛应用. 其中关于全局最优解的研究已经取得一定的进展^[1-3], 得到了多种求解方法. 对于目标函数具有一定单调性的非线性规划问题, 文献[4-6]给出了一些凸化、凹化法, 将单调非线性规划变换成相应的凸规划或凹规划. 对于目标函数或约束函数非单调、非凸凹的非线性规划问题, 文献[7-9]给出了一些变换方法, 通过对目标函数凸化或凹化, 将原问题转化为等价的凹极小问题、反凸规划问题或标准 D.C. 规划问题.

本文对于目标函数是次正定函数的全局最优化问题, 给出一种新的凸化法, 通过将目标函数凸化, 从而将原问题转化为相应的凸规划, 并论证了次正定函数全局最优解的一些最优性条件. 本文提出的凸化方法对目标函数在定义域上的值大于零或小于零的情形都可进行凸化, 从而扩大了方法的应用范围.

2 基本定义

考虑下述全局最优化问题:

$$\begin{aligned} \min & f(x), \\ \text{s.t. } & x \in X, \end{aligned} \quad (2.1)$$

其中 $f: R^n \rightarrow R$, $f(x) \in C^2$, 有界闭凸集 $X = \{x \in R^n \mid l_i \leq x_i \leq u_i, i = 1, 2, \dots, n\}$. 用 S^n 表示 R^n 中的单位球面, $S^n = \{d \in R^n \mid \|d\| = 1\}$.

记 $g(x)$ 为函数 $f(x)$ 在 x 处的梯度, $G(x)$ 为函数 $f(x)$ 在 x 处的 Hesse 矩阵, $G(x) > 0$ 表示函数 $f(x)$ 在 x 处的 Hesse 矩阵正定, $G(x) < 0$ 表示函数 $f(x)$ 在 x 处的 Hesse 矩阵负定.

*收稿日期: 2015-01-15

接收日期: 2015-05-06

基金项目: 山东省优秀中青年科学家科研奖励基金资助 (BS2013SF014).

作者简介: 李博 (1957-), 男, 山东桓台, 教授, 主要研究方向: 最优化理论与应用.

定义^[5] 设函数 $f: R^n \rightarrow R$, $f(x) \in C^2$, 若 $\forall x \in X, d \in S^n$,

$$d^T[G(x) + g(x)g(x)^T]d > 0 \quad (2.2)$$

成立, 则称 $f(x)$ 是 X 上的次正定函数.

例 1 函数 $f(x) = \frac{1}{2}x_1^2 - x_1x_2 - \frac{1}{2}x_2^2$, 其中 $x \in X = \{0 \leq x_1 \leq 3, 3 \leq x_2 \leq 6\}$.

函数 $f(x)$ 在 X 上的梯度: $g(x) = (x_1 - x_2, -x_1 - x_2)^T$, 函数 $f(x)$ 在 X 上的 Hesse 矩阵 $G(x) = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$, 因此得到

$$H_f(x) = G(x) + g(x)g(x)^T = \begin{bmatrix} (x_1 - x_2)^2 + 1 & -(x_1^2 - x_2^2) - 1 \\ -(x_1^2 - x_2^2) - 1 & (x_1 + x_2)^2 - 1 \end{bmatrix},$$

则由 $H_f(x)$ 的特征多项式 $|\lambda E - H_f(x)| = 0$ 得

$$\lambda = (x_1^2 + x_2^2) \pm \sqrt{(x_1^2 + x_2^2)^2 - (2x_2^2 - 2x_1^2 + 4x_1x_2 - 2)},$$

则 $\forall x \in X, \lambda > 0$ 恒成立, 故 $\forall x \in X, d \in S^2, d^T H_f(x) d > 0$ 成立, $H_f(x)$ 是 X 上的正定矩阵, $f(x)$ 是 X 上的次正定函数.

由此可知当函数 $f(x)$ 在 X 上是非凸非凹函数时, 若 $G(x) + g(x)g(x)^T$ 是正定矩阵, 则 $f(x)$ 是 X 上的次正定函数. 显然, 正定函数必为次正定函数, 但反之不然.

3 凸化

定理 3.1 设函数 $f: R^n \rightarrow R, f(x) \in C^2$, 且 $f(x)$ 为 X 上的次正定函数, $\forall x \in X, f(x) > 0$, 令

$$F(x) = qe^{f(x)} + \frac{1}{2}f^2(x), \quad (3.1)$$

则存在 $q_1 \geq 0$, 使得当 $q > q_1$ 时, $F(x)$ 为 X 上的凸函数.

证 $\forall x \in X$ 和 $q > 0$, 由 (3.1) 式得

$$\frac{\partial F(x)}{\partial x_i} = qe^{f(x)} \frac{\partial f(x)}{\partial x_i} + f(x) \frac{\partial f(x)}{\partial x_i},$$

计算得

$$\begin{aligned} \frac{\partial^2 F(x)}{\partial x_i^2} &= qe^{f(x)} \frac{\partial^2 f(x)}{\partial x_i^2} + qe^{f(x)} \left(\frac{\partial f(x)}{\partial x_i}\right)^2 + f(x) \frac{\partial^2 f(x)}{\partial x_i^2} + \left(\frac{\partial f(x)}{\partial x_i}\right)^2, \\ \frac{\partial^2 F(x)}{\partial x_i \partial x_j} &= qe^{f(x)} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + qe^{f(x)} \frac{\partial f(x)}{\partial x_i} \frac{\partial f(x)}{\partial x_j} + f(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \frac{\partial f(x)}{\partial x_i} \frac{\partial f(x)}{\partial x_j} \\ &(i, j = 1, 2, \dots, n. i \neq j). \end{aligned}$$

用 $H_F(x)$ 表示函数 $F(x)$ 在点 x 处的 Hesse 矩阵, 则 $\forall x \in X$,

$$H_F(x) = qe^{f(x)}[G(x) + g(x)g(x)^T] + f(x)G(x) + g(x)g(x)^T.$$

令

$$Q(x) = e^{f(x)}[G(x) + g(x)g(x)^T], \quad \lambda = \min_{\substack{x \in X \\ 1 \leq i \leq n}} \lambda_i(Q(x)),$$

其中 $\lambda_i(Q(x))$ 表示 $Q(x)$ 的特征值,

$$R(x) = f(x)G(x) + g(x)g(x)^T, \quad \gamma = \min_{\substack{x \in X \\ 1 \leq i \leq n}} \gamma_i(R(x)),$$

其中 $\gamma_i(R(x))$ 表示 $R(x)$ 的特征值. 因此 $d^T Q(x)d \geq \lambda$, $d^T R(x)d \geq \gamma$. 故

$$d^T H_F(x)d = qd^T Q(x)d + d^T R(x)d \geq q\lambda + \gamma,$$

又由 $G(x) + g(x)g(x)^T$ 是正定矩阵且 $e^{f(x)} > 0$ 可知 $\lambda > 0$. 于是, 若令

$$q_1 = \max\{0, \frac{-\gamma}{\lambda}\},$$

当 $q > q_1$ 时, $\forall x \in X, d \in S^n, d^T H_F(x)d > 0$ 成立, $H_F(x)$ 为 X 上的正定矩阵, $F(x)$ 为 X 上的凸函数.

定理 3.2 设函数 $f: R^n \rightarrow R, f(x) \in C^2$, 且 $f(x)$ 为 X 上的次正定函数, $\forall x \in X, f(x) < 0$, 令

$$D(x) = qe^{f(x)} - \ln(-f(x)), \quad (3.2)$$

则存在 $q_2 \geq 0$, 使得当 $q > q_2$ 时, $D(x)$ 为 X 上的凸函数.

证 $\forall x \in X$ 和 $q > 0$, 由 (3.2) 式得

$$\frac{\partial D(x)}{\partial x_i} = qe^{f(x)} \frac{\partial f(x)}{\partial x_i} - \frac{1}{f(x)} \frac{\partial f(x)}{\partial x_i},$$

计算得

$$\begin{aligned} \frac{\partial^2 D(x)}{\partial x_i^2} &= qe^{f(x)} \frac{\partial^2 f(x)}{\partial x_i^2} + qe^{f(x)} \left(\frac{\partial f(x)}{\partial x_i}\right)^2 - \frac{1}{f(x)} \frac{\partial^2 f(x)}{\partial x_i^2} + \frac{1}{f^2(x)} \left(\frac{\partial f(x)}{\partial x_i}\right)^2, \\ \frac{\partial^2 D(x)}{\partial x_i \partial x_j} &= qe^{f(x)} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + qe^{f(x)} \frac{\partial f(x)}{\partial x_i} \frac{\partial f(x)}{\partial x_j} - \frac{1}{f(x)} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \frac{1}{f^2(x)} \frac{\partial f(x)}{\partial x_i} \frac{\partial f(x)}{\partial x_j} \\ &(i, j = 1, 2, \dots, n. i \neq j). \end{aligned}$$

用 $H_D(x)$ 表示函数 $D(x)$ 在点 x 处的 Hesse 矩阵, 则 $\forall x \in X$,

$$H_D(x) = qe^{f(x)}[G(x) + g(x)g(x)^T] - \frac{1}{f(x)}G(x) + \frac{1}{f^2(x)}g(x)g(x)^T.$$

令

$$Q_1(x) = e^{f(x)}[G(x) + g(x)g(x)^T], \quad \lambda' = \min_{\substack{x \in X \\ 1 \leq i \leq n}} \lambda_i(Q_1(x)),$$

其中 $\lambda_i(Q_1(x))$ 表示 $Q_1(x)$ 的特征值,

$$R_1(x) = -\frac{1}{f(x)}G(x) + \frac{1}{f^2(x)}g(x)g(x)^T, \quad \gamma' = \min_{\substack{x \in X \\ 1 \leq i \leq n}} \gamma_i(R_1(x)),$$

其中 $\gamma_i(R_1(x))$ 表示 $R_1(x)$ 的特征值. 因此 $d^T Q_1(x)d \geq \lambda'$, $d^T R_1(x)d \geq \gamma'$. 故

$$d^T H_D(x)d = qd^T Q_1(x)d + d^T R_1(x)d \geq q\lambda' + \gamma',$$

又由 $G(x) + g(x)g(x)^T$ 是正定矩阵且 $e^{f(x)} > 0$ 可知 $\lambda' > 0$. 于是, 若令

$$q_2 = \max\left\{0, \frac{-\gamma'}{\lambda'}\right\},$$

当 $q > q_2$ 时, $\forall x \in X, d \in S^n, d^T H_D(x)d > 0$ 成立, $H_D(x)$ 为 X 上的正定矩阵, $D(x)$ 为 X 上的凸函数.

4 算例

本节给出一个实例说明可将全局优化问题中具有次正定性质的目标函数转化为相应的凸函数.

例 2 考虑如下全局最优化问题:

$$\begin{aligned} \min f(x) &= \frac{1}{2}x_1^2 - x_1x_2 - \frac{1}{2}x_2^2, \\ \text{s.t. } x \in X &= \{0 \leq x_1 \leq 3, 3 \leq x_2 \leq 6\}. \end{aligned}$$

图 1 给出了函数 $f(x)$ 在 X 上的图像, 由图 1 可知 $f(x)$ 是 X 上的非凸非凹函数.

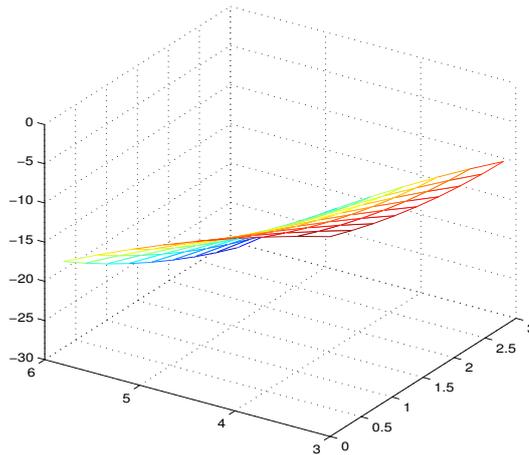


图 1: $f(x)$ 在 X 上的图像

由例 1 知 $f(x)$ 是 X 上的次正定函数, 又 $\forall x \in X, f(x) < 0$, 根据定理 3.2 可知

$$Q_1(x) = e^{f(x)}[G(x) + g(x)g(x)^T],$$

计算得 $\lambda' = \min_{\substack{x \in X \\ 1 \leq i \leq 2}} \lambda_i(Q_1(x)) = 2.92e - 014$,

$$R_1(x) = -\frac{1}{f(x)}G(x) + \frac{1}{f^2(x)}g(x)g(x)^T,$$

计算得 $\gamma' = \min_{\substack{x \in X \\ 1 \leq i \leq 2}} \gamma_i(R_1(x)) = 0.033$. 故 $q_2 = \max\{0, \frac{-\gamma'}{\lambda}\} = 0$, 又 $q > q_2$, 取 $q = 1$. 此时,

$$D(x) = e^{\frac{1}{2}x_1^2 - x_1x_2 - \frac{1}{2}x_2^2} - \ln(\frac{1}{2}x_2^2 + x_1x_2 - \frac{1}{2}x_1^2),$$

因此当 $q = 1$ 时, $\forall x \in X, d \in S^2, d^T H_D(x) d > 0$, 即 $H_D(x)$ 是 X 上的正定矩阵.

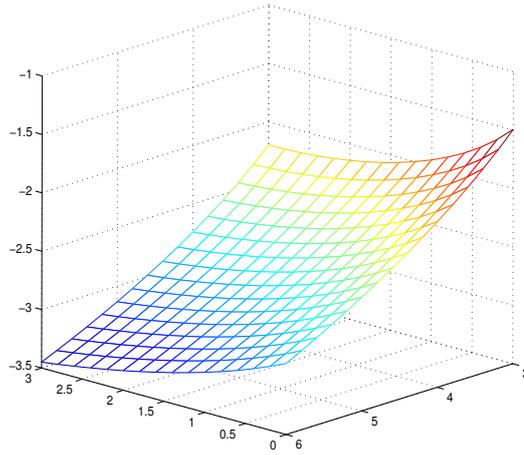


图 2: $D(x)(q = 1)$ 在 X 上的图像

图 2 给出了 $D(x)(q = 1)$ 在 X 上的图像, 由图 2 可知 $\forall x \in X, D(x)$ 是 X 上的凸函数.

5 最优性条件

定理 5.1 设函数 $f: R^n \rightarrow R, f(x) \in C^2$, 且 $f(x)$ 为 X 上的次正定函数, $\forall x \in X, f(x) > 0$, 令

$$F(x) = qe^{f(x)} + \frac{1}{2}f^2(x),$$

若函数 $f(x)$ 在 X 上存在全局极小点, 则存在 $q_1 \geq 0$, 使得当 $q > q_1$ 时, x_0 为 $f(x)$ 的全局极小点的充要条件是 x_0 为 $F(x)$ 的全局极小点.

证 必要性. 设 x_0 是 $f(x)$ 的全局极小点, 则有 $0 < f(x_0) \leq f(x)$, 故

$$\frac{1}{2}f^2(x_0) \leq \frac{1}{2}f^2(x). \quad (5.1)$$

又因为 $q > q_1 \geq 0$, 所以

$$qe^{f(x_0)} \leq qe^{f(x)}. \quad (5.2)$$

由 (5.1), (5.2) 式得

$$qe^{f(x_0)} + \frac{1}{2}f^2(x_0) \leq qe^{f(x)} + \frac{1}{2}f^2(x),$$

因此 x_0 是 $F(x)$ 的全局极小点.

充分性. 设 x_0 是 $F(x)$ 的全局极小点, 有

$$\nabla F(x_0) = [qe^{f(x_0)} + f(x_0)]g(x_0) = 0.$$

又因为 $qe^{f(x_0)} + f(x_0) > 0$, 所以 $g(x_0) = 0$.

又由定理 3.1 可知当 $q > q_1 \geq 0$ 时, $F(x)$ 不仅是 X 上的凸函数, 而且是严格凸函数, 则 x_0 是 $F(x)$ 在 X 上唯一的全局极小点, 由于 $F(x)$ 在 x_0 处二次可微且 $\nabla F(x_0) = 0$, 因此有

$$F(x_0 + \delta d) = F(x_0) + \frac{1}{2}\delta^2 d^T H_F(x_0)d + \delta^2 \|d\|^2 \alpha(x_0, \delta d),$$

其中当 $\delta \rightarrow 0$ 时, $\alpha(x_0, \delta d) \rightarrow 0$, 于是

$$\frac{F(x_0 + \delta d) - F(x_0)}{\delta^2} = \frac{1}{2}d^T H_F(x_0)d + \|d\|^2 \alpha(x_0, \delta d),$$

由于 x_0 是 $F(x)$ 唯一的全局极小点, 当 $|\delta|$ 充分小时, 必有 $F(x_0 + \delta d) > F(x_0)$, 又 $\delta \rightarrow 0$ 时, $\alpha(x_0, \delta d) \rightarrow 0$, 因此 $d^T H_F(x_0)d > 0$. 故

$$H_F(x_0) = [qe^{f(x_0)} + f(x_0)]G(x_0) > 0.$$

因为 $qe^{f(x_0)} + f(x_0) > 0$, 所以 $G(x_0) > 0$.

由局部极小点的二阶充分条件^[10]知 x_0 是 $f(x)$ 的局部极小点.

下证 x_0 是 $f(x)$ 的全局极小点.

用反证法. 假设 x_0 不是 $f(x)$ 的全局极小点, 则存在点 $x' \in X (x' \neq x_0)$, 使得 $0 < f(x') < f(x_0)$, 利用定理 5.1 证明中的必要性证明方法可得

$$qe^{f(x')} + \frac{1}{2}f^2(x') < qe^{f(x_0)} + \frac{1}{2}f^2(x_0),$$

与 x_0 是 $F(x)$ 的全局极小点矛盾, 故 x_0 是 $f(x)$ 的全局极小点.

定理 5.2 设函数 $f: R^n \rightarrow R$, $f(x) \in C^2$, 且 $f(x)$ 为 X 上的次正定函数, $\forall x \in X, f(x) < 0$, 令

$$D(x) = qe^{f(x)} - \ln(-f(x)),$$

若函数 $f(x)$ 在 X 上存在全局极小点, 则存在 $q_2 \geq 0$, 使得当 $q > q_2$ 时, x^* 为 $f(x)$ 的全局极小点的充要条件是 x^* 为 $D(x)$ 的全局极小点.

证 必要性. 设 x^* 是 $f(x)$ 的全局极小点, 则有 $f(x^*) \leq f(x) < 0$, 故

$$-\ln(-f(x^*)) \leq -\ln(-f(x)), \quad (5.3)$$

又因为 $q > q_2 \geq 0$, 所以

$$qe^{f(x^*)} \leq qe^{f(x)}. \quad (5.4)$$

由 (5.3), (5.4) 式得

$$qe^{f(x^*)} - \ln[-f(x^*)] \leq qe^{f(x)} - \ln[-f(x)],$$

因此 x^* 是 $D(x)$ 的全局极小点.

充分性. 设 x^* 是 $D(x)$ 的全局极小点, 有

$$\nabla D(x^*) = [qe^{f(x^*)} - \frac{1}{f(x^*)}]g(x^*) = 0.$$

又因为 $qe^{f(x^*)} - \frac{1}{f(x^*)} > 0$, 所以 $g(x^*) = 0$. 又由定理 3.2 可知当 $q > q_2 \geq 0$ 时, $D(x)$ 不仅是 X 上的凸函数, 而且是严格凸函数, 则 x^* 是 $D(x)$ 在 X 上唯一的全局极小点, 由于 $D(x)$ 在 x^* 处二次可微且 $\nabla D(x^*) = 0$, 因此有

$$D(x^* + \delta d) = D(x^*) + \frac{1}{2}\delta^2 d^T H_D(x^*)d + \delta^2 \|d\|^2 \alpha(x^*, \delta d),$$

其中当 $\delta \rightarrow 0$ 时, $\alpha(x^*, \delta d) \rightarrow 0$, 于是

$$\frac{D(x^* + \delta d) - D(x^*)}{\delta^2} = \frac{1}{2}d^T H_D(x^*)d + \|d\|^2 \alpha(x^*, \delta d).$$

由于 x^* 是 $D(x)$ 唯一的全局极小点, 当 $|\delta|$ 充分小时, 必有 $D(x^* + \delta d) > D(x^*)$, 又 $\delta \rightarrow 0$ 时, $\alpha(x^*, \delta d) \rightarrow 0$, 因此 $d^T H_D(x^*)d > 0$. 故

$$H_D(x^*) = [qe^{f(x^*)} - \frac{1}{f(x^*)}]G(x^*) > 0.$$

因为 $qe^{f(x^*)} - \frac{1}{f(x^*)} > 0$, 所以 $G(x^*) > 0$.

由局部极小点的二阶充分条件知 x^* 是 $f(x)$ 的局部极小点.

下证 x^* 是 $f(x)$ 的全局极小点.

用反证法. 假设 x^* 不是 $f(x)$ 的全局极小点, 则存在点 $x_1 \in X (x_1 \neq x^*)$, 使得 $f(x_1) < f(x^*) < 0$, 利用定理 5.2 证明中的必要性证明方法可得

$$qe^{f(x_1)} - \ln[-f(x_1)] < qe^{f(x^*)} - \ln[-f(x^*)],$$

与 x^* 是 $D(x)$ 的全局极小点矛盾, 故 x^* 是 $f(x)$ 的全局极小点.

由定理 5.1, 5.2 知可通过凸化方法, 将次正定函数在有界闭凸集上的全局优化问题转化为相应的凸规划.

6 结语

本研究对目标函数为次正定函数的全局最优化问题, 给出了一种新的凸化方法, 通过将目标函数凸化, 从而将原问题转化为相应的凸规划, 然后论证了这一类函数全局最优解的最优性条件. 对于次正定函数的全局最优化问题还需有进一步的研究, 期望得到更好的凸化方法与最优性条件.

参 考 文 献

- [1] 汤京永, 贺国平, 董丽. 一类新的求解无约束优化问题的记忆梯度法 [J]. 数学杂志, 2011, 31(2): 362-368.

- [2] 李博, 鲁殿军. 全局最优化问题的一个无参数的填充函数算法 [J]. 数学杂志, 2014, 34(4): 773–778.
- [3] 徐语伦, 赵德芬, 王薇. 由任意初始点求解离散型约束全局优化问题 [J]. 数学杂志, 2011, 31(3): 539–546.
- [4] 张连生, 邬冬华. 非线性规划的凸化, 凹化和单调化 [J]. 数学年刊, 2002, 23(4): 537–544.
- [5] 李博, 周伊佳. 全局最优化问题的凸凹化法 [J]. 青岛科技大学学报 (自然科学版), 2010, 31(3): 321–324.
- [6] Wu Zhiyou, Bai Fusheng, Zhang Liansheng. Convexification and concavification for a general class of global optimization problems[J]. J. Global Optim., 2005, 31(1): 45–60.
- [7] 何颖. 一类全局优化问题的新的凸化, 凹化法 [J]. 长春大学学报, 2008, 18(1): 1–6.
- [8] Li Duan, Sun Xiaoling. Local convexification of the lagrangian function in nonconvex optimization [J]. J. Optim. The. Appl., 2000, 104(1): 109–120.
- [9] Zhu Wenxing, Fu Qingxiang. A sequential convexification method for continuous global optimization[J]. J. Global Optim., 2003, 26(2): 167–182.
- [10] 陈宝林. 最优化理论与算法 [M]. 北京: 清华大学出版社, 1989.

CONVEXIFICATION APPROACHES FOR A CLASS OF CONTINUOUS GLOBAL OPTIMIZATION PROBLEM

LI Bo, DU Jie, WAN Li-juan

(*School of Mathematics and Physics, Qingdao University of Science and Technology,
Qingdao 266061, China*)

Abstract: In this paper, we develop convexification approaches and optimization criteria for a general class of nonconvex optimization problem. By using the method of function transformations with parameter, a class of novel convexification schemes is presented for solving global optimization problem with positive sub-definite objective function. The general class of nonconvex programming discussed in the paper can be solved to global optimality, which extends applications of convexification schemes in solving global optimization problems.

Keywords: global optimization; positive sub-definite function; convexification; necessary and sufficient condition

2010 MR Subject Classification: 90C30; 65K05