Journal of Global Optimization (2005) 31:211–233

Hidden Convex Minimization

DUAN LI[1], ZHI-YOU WU[2,3], HEUNG-WING JOSEPH LEE[4], XIN-MIN YANG[5] and LIAN-SHENG ZHANG[6]

[1]*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong (E-mail: dli@se.cuhk.edu.hk)*
[2]*Department of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, PR China*
[3]*Department of Mathematics, Shanghai University, Baoshan, Shanghai 200436, PR China*
[4]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*
[5]*Department of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, PR China (Current address: Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong)*
[6]*Department of Mathematics, Shanghai University, Baoshan, Shanghai, 200436, PR China.*

(Received 30 March 2004; accepted in revised form 1 April 2004)

Abstract. A class of nonconvex minimization problems can be classified as hidden convex minimization problems. A nonconvex minimization problem is called a hidden convex minimization problem if there exists an equivalent transformation such that the equivalent transformation of it is a convex minimization problem. Sufficient conditions that are independent of transformations are derived in this paper for identifying such a class of seemingly nonconvex minimization problems that are equivalent to convex minimization problems. Thus, a global optimality can be achieved for this class of hidden convex optimization problems by using local search methods. The results presented in this paper extend the reach of convex minimization by identifying its equivalent with a nonconvex representation.

Key words: Convexification, Convex optimization, Global optimization, Hidden convex optimization, Nonconvex optimization

1. Introduction

We consider in this paper the following optimization problem:

$$\begin{aligned} \min\ & g_0(x) \\ \text{s.t. } & g_i(x) \leqslant b_i, \quad i = 1, \ldots, m \\ & x \in X \end{aligned} \tag{1}$$

where $g_i : R^n \to R, i = 0, 1, \ldots, m$, are second-order differentiable functions and

$$X = \{x \in R^n | l_i \leqslant x_i \leqslant u_i,\ i = 1, \ldots, n\}. \tag{2}$$

Without loss of generality, we assume $0 < l_i < u_i$ for $i = 1, \ldots, n$.

If all $g_i(x), i = 0, 1, \ldots, m$, are convex functions, problem (1) is a convex minimization problem whose local minimum is also a global minimum. A very natural question to ask is what are the nonconvex situations of (1) where a local minimum is also guaranteed to be a global minimum.

Convexity plays a central role in mathematical economics, engineering, management science, and optimization theory. One main reason behind its wide applications and cross-board success is that the convexity is a sufficient condition for ensuring that a local optimal solution is also a global one. Convexity, however, is not a necessary condition for a local minimum to be a global one. In the real world, the problems that present a convexity may only cover a small percentage of a large population. Therefore, researchers have been exploring the situations where the convexity condition can be relaxed to certain degree, while at the same time some nice properties similar to those enjoyed by convex functions are preserved. The convexity has been extended to various forms of a generalized convexity in the literature [1, 8]. Examples of the generalized convex functions include pseudoconvex functions, quasiconvex functions and invex functions [3]. Optimality and duality for generalized convex functions are investigated in [6].

Horst [4] discussed convexification of nonconvex functions that can be transformed into convex ones using either a range transformation or a domain transformation. A convex range transformable function must be also quasiconvex. Ben-Tal and Teboulle [2] considered a singly and quadratically constrained quadratic problem and proved that, under certain conditions for simultaneous diagonalization, the dual of the dual problem of the primal problem leads to a convex problem which is completely equivalent to the nonconvex primal problem. Ben-Tal and Teboulle [2] also showed that a special problem of minimizing a concave quadratic function subject to finitely many convex constraints is equivalent to a minimax convex problem. Recent results in [5] revealed that via an equivalent transformation using a pth-power method a monotone nonconvex optimization problem can be always converted into an equivalent better-structured nonconvex optimization problem, e.g., a concave minimization problem or a D.C. programming problem. A general theory of convexification transformations was further developed in [7] in the context of monotone global optimization.

The primal goal of this paper is to identify a class of seemingly nonconvex optimization problems that can be converted into equivalent convex minimization problems. Problem (1) is called a hidden convex minimization problem in this paper if there exists an equivalent transformation such that the equivalent transformation of (1) is a convex minimization problem. Sufficient conditions are derived in this paper to identify such hidden convex problems. One prominent feature of the results in this paper is that these derived conditions are independent of transformations. A global optimality can be achieved for hidden convex minimization problems by using local search methods. The results presented in this paper extend the reach of convex minimization by identifying its equivalent with a nonconvex representation.

This paper is organized as follows. In Section 2, we state a basic theorem that reveals an equivalence between certain hidden convex function and its convex counterpart. In Section 3, we give the equivalence of a primal hidden convex problem and its transformed convex programming problem. Then we obtain some conditions under which a programming problem is hidden convex. In Section 4, we focus on quadratic programming problems and derive some conditions for hidden convexity. The paper concludes in Section 5 with a case study that demonstrates a high possibility of occurrence of hidden convexity in nonconvex situations.

2. Convexification Transformation

Let function $h \in C^2$ be defined on X in (2).

DEFINITION. A function $h: R^n \to R$ is increasing (decreasing) on X with respect to x_i if

$$h(x_1, \ldots, x_{i-1}, x_i^1, x_{i+1}, \ldots, x_n) \leqslant (\geqslant)\ h(x_1, \ldots, x_{i-1}, x_i^2, x_{i+1}, \ldots, x_n)$$

for any $x_i^1 < x_i^2$, where $x_i^1, x_i^2 \in X_i = \{x_i | (x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \in X\}$; A function $h : R^n \to R$ is strictly increasing (decreasing) on X with respect to x_i if

$$h(x_1, \ldots, x_{i-1}, x_i^1, x_{i+1}, \ldots, x_n) < (>)\ h(x_1, \ldots, x_{i-1}, x_i^2, x_{i+1}, \ldots, x_n)$$

for any $x_i^1 < x_i^2$, where $x_i^1, x_i^2 \in X_i$.

DEFINITION. A function $h : R^n \to R$ is said to be monotone on its domain X if h is increasing or decreasing on X with respect to all $x_i, i = 1, \ldots, n$; A function h defined on X is said to be strictly monotone if h is strictly increasing or strictly decreasing on X with respect to all $x_i, i = 1, \ldots, n$.

Consider the following variable transformation of function $h(x)$:

$$h_p(y) = h\Big(t_p(y)\Big) \tag{3}$$

where $p = (p_1, \ldots, p_n)$ is a parameter vector and $t_p(y) : R^n \to R^n$ is a separable mapping, i.e., $t_p(y) = (t_{1,p_1}(y_1), \ldots, t_{n,p_n}(y_n))$ for $y = (y_1, \ldots, y_n)$. We further assume that each of $t_{i,p_i}(y_i) \in C^2, i = 1, \ldots, n$, is a one-to-one mapping. The domain of $h_p(y)$ is

$$Y_p = \{y \in R^n |\, y_i = t_{i,p_i}^{-1}(x_i),\ i = 1, \ldots, n, (x_1, \ldots, x_n) \in X\}. \tag{4}$$

Clearly, Y_p is still a box and is thus convex. Let Ω be an open set satisfying $Y_p \subseteq \Omega$ for all p. Denote

$$\Omega_i = \{y_i \in R | (y_1, \ldots, y_i, \ldots, y_n) \in \Omega\}, \quad i = 1, \ldots, n.$$

Let S^n be the unit sphere in R^n. Let $\min_{x \in \emptyset} x = +\infty$, $\max_{x \in \emptyset} x = -\infty$ for a purpose of convenience.

DEFINITION. If there exists a one-to-one mapping $x = t_p(y)$ such that $h_p(y)$ defined in (3) is convex, then the function $h(x)$ is called a hidden convex function.

Denote by η_i and ζ_i upper and lower bounds of $\partial h / \partial x_i$ over X, $i = 1, \ldots, n$, respectively, i.e.,

$$\eta_i \geqslant \max_{x \in X} \frac{\partial h(x)}{\partial x_i} \tag{5}$$

$$\zeta_i \leqslant \min_{x \in X} \frac{\partial h(x)}{\partial x_i} \tag{6}$$

Denote by λ a lower bound of the minimum eigenvalue of the Hessian of h over X, i.e.,

$$\lambda \leqslant \min_{x \in X} \lambda(x) = \min_{x \in X, d \in S^n} d^T H(x) d, \tag{7}$$

where $H(z)$ is the Hessian of h at z and $\lambda(x)$ is the minimum eigenvalue of $H(x)$.

Denote

$$s(a) = \begin{cases} 1 & a \geqslant 0 \\ 0 & a < 0 \end{cases} \tag{8}$$

Let

$$m_{i,p_i} = \min_{x \in X} \frac{t''_{i,p_i}\left(t_{i,p_i}^{-1}(x_i)\right)}{\left[t'_{i,p_i}\left(t_{i,p_i}^{-1}(x_i)\right)\right]^2}, \tag{9}$$

$$M_{i,p_i} = \max_{x \in X} \frac{t''_{i,p_i}\left(t_{i,p_i}^{-1}(x_i)\right)}{\left[t'_{i,p_i}\left(t_{i,p_i}^{-1}(x_i)\right)\right]^2}, \tag{10}$$

$$a_{i,p_i} = m_{i,p_i}\zeta_i s(\zeta_i) + M_{i,p_i}\zeta_i s(-\zeta_i) \tag{11}$$

$$b_{i,p_i} = m_{i,p_i}\eta_i s(\eta_i) + M_{i,p_i}\eta_i s(-\eta_i) \tag{12}$$

$$\begin{aligned} B_i =& \{p_i | a_{i,p_i} \geqslant -\lambda, \quad m_{i,p_i} \geqslant 0\} \\ & \cup \{p_i | b_{i,p_i} \geqslant -\lambda, \quad M_{i,p_i} \leqslant 0\} \end{aligned} \tag{13}$$

$$\begin{aligned} \bar{B}_i =& \{p_i | a_{i,p_i} > -\lambda, \quad m_{i,p_i} \geqslant 0\} \\ & \cup \{p_i | b_{i,p_i} > -\lambda, \quad M_{i,p_i} \leqslant 0\} \end{aligned} \tag{14}$$

THEOREM 2.1. Assume that $t_{i,p_i}, i = 1, \ldots, n$, are strictly monotone functions on Ω_i and satisfy $t'_{i,p_i}(y_i) \neq 0$ *for any* $y_i \in \Omega_i$. Furthermore, for any $i = 1, \ldots, n, t_{i,p_i}$ is either convex or concave on Ω_i.

If for all $i = 1, \ldots, n, B_i \neq \emptyset$, then $h_p(y)$ is a convex function on Y_p when $p_i \in B_i, i = 1, \ldots, n$; and if for all $i = 1, \ldots, n, \bar{B}_i \neq \emptyset$, then $h_p(y)$ is strictly convex on Y_p when $p_i \in \bar{B}_i, i = 1, \ldots, n$.

Proof. Let $x = t_p(y), \forall y \in Y_p$. By (3), we have

$$\frac{\partial h_p(y)}{\partial y_k} = \frac{\partial h(x)}{\partial x_k} t'_{k,p_k}(y_k), \quad k = 1, 2, \ldots, n$$

$$\begin{aligned} \frac{\partial^2 h_p(y)}{\partial y_k^2} &= \frac{\partial^2 h(x)}{\partial x_k^2}[t'_{k,p_k}(y_k)]^2 + \frac{\partial h(x)}{\partial x_k} t''_{k,p_k}(y_k) \\ &= [t'_{k,p_k}(y_k)]^2 \left[\frac{\partial^2 h(x)}{\partial x_k^2} + \frac{\partial h(x)}{\partial x_k} \frac{t''_{k,p_k}(y_k)}{[t'_{k,p_k}(y_k)]^2} \right], \quad k = 1, 2, \ldots, n \end{aligned}$$

$$\frac{\partial^2 h_p(y)}{\partial y_k \partial y_j} = \frac{\partial^2 h(x)}{\partial x_k \partial x_j} t'_{k,p_k}(y_k) t'_{j,p_j}(y_j), \quad k,\ j = 1, 2, \ldots, n,\ k \neq j.$$

Let

$$A(x) = \text{ diag } \left(t'_{1,p_1}(y_1), \ldots, t'_{n,p_n}(y_n) \right), \tag{15}$$

$$B(x) = \text{ diag } \left(\frac{\partial h(x)}{\partial x_1} \frac{t''_{1,p_1}(y_1)}{[t'_{1,p_1}(y_1)]^2}, \ldots, \frac{\partial h(x)}{\partial x_n} \frac{t''_{n,p_n}(y_n)}{[t'_{n,p_n}(y_n)]^2} \right). \tag{16}$$

Denote the Hessian of $h_p(y)$ by $H_p(y)$. Then

$$H_p(y) = A(x)[H(x) + B(x)]A(x).$$

For all $d \in S^n$,

$$d^T H_p(y) d = d^T A(x)[H(x) + B(x)]A(x)d.$$

It is clear that $H_p(y)$ is positive definite if and only if $H(x) + B(x)$ is positive definite. For any $d \in S^n$ and $x \in X$, we have

$$\begin{aligned} d^T[H(x) + B(x)]d = &\ d^T H(x) d + \sum_{i=1}^{n} \frac{\partial h(x)}{\partial x_i} \frac{t''_{i,p_i}(y_i)}{[t'_{i,p_i}(y_i)]^2} d_i^2 \\ &\geqslant \lambda + \sum_{i=1}^{n} \frac{\partial h(x)}{\partial x_i} \frac{t''_{i,p_i}(y_i)}{[t'_{i,p_i}(y_i)]^2} d_i^2 \end{aligned} \tag{17}$$

Since $t_{i,p_i}(y_i)$ is either convex or concave on Ω_i,

$$\begin{aligned} \frac{\partial h(x)}{\partial x_i} \frac{t''_{i,p_i}(y_i)}{[t'_{i,p_i}(y_i)]^2} &\geqslant \begin{cases} \zeta_i \frac{t''_{i,p_i}(y_i)}{[t'_{i,p_i}(y_i)]^2} & \text{if } m_{i,p_i} \geqslant 0 \\ \eta_i \frac{t''_{i,p_i}(y_i)}{[t'_{i,p_i}(y_i)]^2} & \text{if } M_{i,p_i} \leqslant 0 \end{cases} \\ &\geqslant \begin{cases} a_{i,p_i} & \text{if } m_{i,p_i} \geqslant 0 \\ b_{i,p_i} & \text{if } M_{i,p_i} \leqslant 0 \end{cases} \end{aligned}$$

we further have

$$d^T[H(x) + B(x)]d \geqslant \lambda + \min_{1 \leqslant i \leqslant n} \left(a_{i,p_i} s(m_{i,p_i}) + b_{i,p_i} s(-M_{i,p_i}) \right).$$

Thus, if for every $i = 1, \ldots, n$, one of the following two inequalities

$$a_{i,p_i} \geqslant -\lambda, \quad m_{i,p_i} \geqslant 0$$

or

$$b_{i,p_i} \geqslant -\lambda, \quad M_{i,p_i} \leqslant 0$$

holds, then we have

$$d^T[H(x) + B(x)]d \geqslant 0.$$

Thus, $h_p(y)$ is a convex function on Y_p for $p_i \in B_i, i = 1, 2, \ldots, n$, if all $B_i \neq \emptyset, i = 1, 2, \ldots, n$. Furthermore, $h_p(y)$ is strictly convex on Y_p for $p_i \in \bar{B}_i, i = 1, 2, \ldots, n$, if all $\bar{B}_i \neq \emptyset, \quad i = 1, \ldots, n$. □

Without loss of generality, we can assume that $\zeta_i \neq 0$ and $\eta_i \neq 0$. Let

$$\begin{aligned} I &= \{i | \zeta_i > 0\} \\ J &= \{i | \eta_i < 0\} \\ \bar{I} &= \{1, \ldots, n\} \backslash I \\ \bar{J} &= \{1, \ldots, n\} \backslash J. \end{aligned}$$

Clearly, h is increasing with respect to x_i if $i \in I$, decreasing with respect to x_i if $i \in J$, and neither increasing nor decreasing with respect to x_i if $i \in \bar{I} \cap \bar{J}$. For an $i \in \bar{I} \cap \bar{J}$, we must have $\zeta_i < 0$ and $\eta_i > 0$.

Set B_i in (13) can be specified for $i \in I, J$ and $\bar{I} \cap \bar{J}$,

$$B_i = \begin{cases} \left\{p_i \middle| \max\left\{0, -\frac{\lambda}{\zeta_i}\right\} \leqslant m_{i,p_i}\right\} \cup \left\{p_i \middle| 0 \geqslant M_{i,p_i} \geqslant m_{i,p_i} \geqslant -\frac{\lambda}{\eta_i}\right\} & i \in I \\ \left\{p_i \middle| 0 \leqslant m_{i,p_i} \leqslant M_{i,p_i} \leqslant -\frac{\lambda}{\zeta_i}\right\} \cup \left\{p_i \middle| M_{i,p_i} \leqslant \min\left\{0, -\frac{\lambda}{\eta_i}\right\}\right\} & i \in J \\ \left\{p_i \middle| 0 \leqslant m_{i,p_i} \leqslant M_{i,p_i} \leqslant -\frac{\lambda}{\zeta_i}\right\} \cup \left\{p_i \middle| 0 \geqslant M_{i,p_i} \geqslant m_{i,p_i} \geqslant -\frac{\lambda}{\eta_i}\right\} & i \in \bar{I} \cap \bar{J} \end{cases} \tag{18}$$

We can conclude the following from (18).

Remark 2.1. (i) If $\lambda \geqslant 0$, h is already a convex function. There exist many transformations $x = t_p(y) (e.g.\ t_p(y) = y)$ such that $B_i \neq \emptyset$ for all $i = 1, \ldots, n$. Thus function $h_p(y)$ retains a convexity. It implies that a convex function is always a hidden convex function.

(ii) *If* $\lambda < 0$, and $\bar{I} \cap \bar{J} \neq \emptyset$, then $B_i = \emptyset$ for any $i \in \bar{I} \cap \bar{J}$. It implies that a non-monotone nonconvex function h on X cannot be convexified (using the method proposed in this paper).

(iii) *If* $\lambda < 0$, then $B_i = \left\{p_i \middle| -\frac{\lambda}{\zeta_i} \leqslant m_{i,p_i}\right\}$ for $i \in I$.

(iv) *If* $\lambda < 0$, then $B_i = \left\{p_i \middle| M_{i,p_i} \leqslant -\frac{\lambda}{\eta_i}\right\}$ for $i \in J$.

Remark 2.2. It is obvious that the resulted set B_i depends on a chosen transformation $x = t_p(y)$. If we take $x = y^{1/p}$, i.e., $x_i = y_i^{1/p_i}, i = 1, \ldots, n$, then we have

$$m_{i,p_i} = \begin{cases} \frac{1-p_i}{u_i} (\geqslant 0), & p_i \leqslant 1 \\ \frac{1-p_i}{l_i} (\leqslant 0), & p_i \geqslant 1 \end{cases} \tag{19}$$

$$M_{i,p_i} = \begin{cases} \frac{1-p_i}{l_i} (\geqslant 0), & p_i \leqslant 1 \\ \frac{1-p_i}{u_i} (\leqslant 0), & p_i \geqslant 1 \end{cases}, \tag{20}$$

and

$$B_i = \begin{cases} \left(-\infty, \min\left\{1, 1 + \frac{\lambda u_i}{\zeta_i}\right\}\right] \cup \left[1, 1 + \frac{\lambda l_i}{\eta_i}\right] & i \in I \\ \left[1 + \frac{\lambda l_i}{\zeta_i}, 1\right] \cup \left[\max\left\{1, 1 + \frac{\lambda u_i}{\eta_i}\right\}, +\infty\right) & i \in J. \\ \left[1 + \frac{\lambda l_i}{\zeta_i}, 1\right] \cup \left[1, 1 + \frac{\lambda l_i}{\eta_i}\right] & i \in \bar{I} \cap \bar{J} \end{cases}$$

If we take $x = -\frac{1}{p}\ln(1+y)$, *i.e.*, $x_i = -\frac{1}{p_i}\ln(1+y_i), i = 1, \ldots, n$, then we have

$$m_{i,p_i} = M_{i,p_i} = p_i.$$

and

$$B_i = \begin{cases} \left[\max\left\{0, -\frac{\lambda}{\zeta_i}\right\}, +\infty\right) \cup \left[-\frac{\lambda}{\eta_i}, 0\right] & i \in I \\ \left[0, -\frac{\lambda}{\zeta_i}\right] \cup \left(-\infty, \min\left\{0, -\frac{\lambda}{\eta_i}\right\},\right] & i \in J. \\ \left[0, -\frac{\lambda}{\zeta_i}\right] \cup \left[-\frac{\lambda}{\eta_i}, 0\right] & i \in \bar{I} \cap \bar{J} \end{cases}$$

Remark 2.3. There are many nonconvex functions which are hidden convex. If h is strictly monotone and $I \cup J = \{1, \ldots, n\}$, then h must be a hidden convex function. In fact, when h is strictly monotone and $I \cup J = \{1, \ldots, n\}$, if we take a transformation $x = t_p(y)$ that satisfies

$$B_i = \left\{p_i | -\frac{\lambda}{\zeta_i} \leqslant m_{i,p_i}\right\} \neq \emptyset, \quad \text{when } i \in I,$$

and

$$B_i = \left\{p_i | M_{i,p_i} \leqslant -\frac{\lambda}{\eta_i}\right\} \neq \emptyset, \quad \text{when } i \in J,$$

then the transformed function $h(t_p(y))$ must be convex when $p_i \in B_i, i = 1, \ldots, n$. For example, if we take $x_i = t_{i,p_i}(y_i) = -\frac{1}{p_i}\ln(1+y_i)$ *for* $i \in I$ *and take* $x_i = \frac{1}{p_i}\ln(1+y_i)$ *for* $i \in J$, *then* $B_i \neq \emptyset$ when p_i is large enough, $i = 1, \ldots, n$. Therefore, $h(t_p(y))$ is convex when p_i, $i = 1, 2, \ldots, n$, are large enough. Thus, a strictly monotone function is always hidden convex. An illustrative example is that $h(x) = \sin(x)$, $x \in (-\pi/2, +\pi/2)$, is not convex, but hidden convex.

3. Equivalent Convex Programming Problem

Theorem 2.1 provides us a basis to identify a class of hidden convex programming problems. By adopting a selected transformation (3), we can convert the primal problem (1) into the following formulation:

$$\begin{aligned} \min \quad & g_{0,p}(y) = g_0\Big(t_p(y)\Big) \\ \text{s.t.} \quad & g_{i,p}(y) = g_i\Big(t_p(y)\Big) \leqslant b_i, \quad i = 1, \ldots, m \\ & y \in Y_p, \end{aligned} \tag{21}$$

where $Y_p = \{y \in R^n | l_i \leqslant t_{i,p_i}(y_i) \leqslant u_i, i = 1, \dots, n\}$ is a box and t_p satisfies the conditions of Theorem 2.1.

The equivalence between (1) and (21) is clear.

THEOREM 3.1. Assume that $t_p(y)$ is a one-to-one mapping with $X = t_p(Y_p)$ and both t_p and t_p^{-1} are continuous. Then y^* is a global or local optimal solution to (21) if and only if $x^* = t_p(y^*)$ is a global or local optimal solution to (1).

Proof. By the facts that t_p is a one-to-one mapping and both t_p and t_p^{-1} are continuous, we can obtain the above conclusion easily (see [7]). □

DEFINITION. If there exists a one-to-one mapping $x = t_p(y)$ such that (21), an equivalent transformation of a programming problem (1), is convex, then (1) is called a hidden convex minimization problem.

Let η_i^k and $\zeta_i^k, k = 0, 1, \dots, m$, be upper and lower bounds of $\frac{\partial g_k(x)}{\partial x_i}$ over $X, i = 1, \dots, n,$ respectively, i.e.

$$\zeta_i^k \leqslant \min_{x \in X} \frac{\partial g_k(x)}{\partial x_i}, \tag{22}$$

$$\eta_i^k \geqslant \max_{x \in X} \frac{\partial g_k(x)}{\partial x_i}. \tag{23}$$

For $i = 1, 2, \dots, n$ and $k = 0, 1, \dots, m$, let

$$a_{i,p_i}^k = m_{i,p_i} \zeta_i^k s(\zeta_i^k) + M_{i,p_i} \zeta_i^k s(-\zeta_i^k) \tag{24}$$

$$b_{i,p_i}^k = m_{i,p_i} \eta_i^k s(\eta_i^k) + M_{i,p_i} \eta_i^k s(-\eta_i^k) \tag{25}$$

$$\lambda_k \leqslant \min_{x \in X, ||d||=1} d^T H_k(x) d \tag{26}$$

$$A_i^k = \{p_i | a_{i,p_i}^k \geqslant -\lambda_k, \quad m_{i,p_i} \geqslant 0\} \cup \{p_i | b_{i,p_i}^k \geqslant -\lambda_k, \quad M_{i,p_i} \leqslant 0\} \tag{27}$$

$$\bar{A}_i^k = \{p_i | a_{i,p_i}^k > -\lambda_k, \quad m_{i,p_i} \geqslant 0\} \cup \{p_i | b_{i,p_i}^k > -\lambda_k, \quad M_{i,p_i} \leqslant 0\} \tag{28}$$

$$A_i = \cap_{k=0}^m A_i^k \tag{29}$$

$$\bar{A}_i = \cap_{k=0}^m \bar{A}_i^k \tag{30}$$

$$A \;= \{p = (p_1, \ldots, p_n) | p_i \in A_i, i = 1, \ldots, n\} \tag{31}$$

$$\bar{A} \;= \{p = (p_1, \ldots, p_n) | p_i \in \bar{A}_i, i = 1, \ldots, n\}, \tag{32}$$

where $H_k(z)$ is the Hessian of g_k at $z, d \in R^n, s(\cdot), m_{i,p_i}$, and M_{i,p_i} are defined by (8), (9), and (10), respectively.

THEOREM 3.2. Assume that in (21) $t_{i,p_i}, i = 1, \ldots, n$, are strictly monotone functions on Ω_i and satisfy $t'_{i,p_i}(y_i) \neq 0$ for any $y_i \in \Omega_i$. Furthermore, for any $i = 1, \ldots, n, t_{i,p_i}$ is either convex or concave on Ω_i.

If $A \neq \emptyset$, *i.e.* $A_i \neq \emptyset$ for all $i = 1, 2, \ldots, n$, then the problem (21) is a convex programming problem when $p \in A$. If $\bar{A} \neq \emptyset$, *i.e.*, $\bar{A}_i \neq \emptyset$ for all $i = 1, 2, \ldots, n$, then the problem (21) is a strictly convex programming problem when $p \in \bar{A}$.

Proof. If $A \neq \emptyset$, i.e. for all $i = 1, \ldots, n, A_i \neq \emptyset$, it implies that for any $k = 0, 1, \ldots, m, i = 1, \ldots, n, A_i^k \neq \emptyset$. From Theorem 2.1 we know that $g_k(t_p(y))$ is convex on Y_p when $p_i \in A_i^k, i = 1, \ldots, n$. Thus, all the functions $g_k(t_p(y)), k = 0, 1, \ldots, m$, are convex on Y_p when $p \in A$. We can conclude that the programming problem (21) is convex on Y_p when $p \in A$. Similarly, the problem (21) is strictly convex when $p \in \bar{A}$ if $\bar{A} \neq \emptyset$. □

From Theorems 3.1 and 3.2, we know that the problem (1) can be converted into an equivalent convex programming problem (21) if $A \neq \emptyset$ when $g_i, i = 0, 1, \ldots, m$, and t_p satisfy the conditions in Theorem 3.2.

Without loss of generality, we can assume that $\zeta_i^k \neq 0$ and $\eta_i^k \neq 0$ for all $k = 0, 1, \ldots, m$ and $i = 1, \ldots, n$. Let for $i = 1, 2, \ldots, n$,

$$I_i = \{k \mid \zeta_i^k > 0, \quad k = 0, 1, \ldots, m\} \tag{33}$$

$$J_i = \{k \mid \eta_i^k < 0, \quad k = 0, 1, \ldots, m\} \tag{34}$$

$$\bar{I}_i = \{0, 1, \ldots, m\} \backslash I_i \tag{35}$$

$$\bar{J}_i = \{0, 1, \ldots, m\} \backslash J_i. \tag{36}$$

If $k \in I_i$, then A_i^k in (27) reduces to

$$A_i^k = \left\{ p_i \mid m_{i,p_i} \geqslant \max\left\{0, -\frac{\lambda_k}{\zeta_i^k}\right\}\right\} \cup \left\{ p_i \mid 0 \geqslant M_{i,p_i} \geqslant m_{i,p_i} \geqslant -\frac{\lambda_k}{\eta_i^k}\right\};$$

If $k \in J_i$, then A_i^k in (27) reduces to

$$A_i^k = \left\{ p_i \mid 0 \leqslant m_{i,p_i} \leqslant M_{i,p_i} \leqslant -\frac{\lambda_k}{\zeta_i^k}\right\} \cup \left\{ p_i \mid M_{i,p_i} \leqslant \min\left\{0, -\frac{\lambda_k}{\eta_i^k}\right\}\right\};$$

If $k \in \bar{I}_i \cap \bar{J}_i$, then A_i^k in (27) reduces to

$$A_i^k = \left\{ p_i \mid 0 \leqslant m_{i,p_i} \leqslant M_{i,p_i} \leqslant -\frac{\lambda_k}{\zeta_i^k} \right\} \cup \left\{ p_i \mid 0 \geqslant M_{i,p_i} \geqslant m_{i,p_i} \geqslant -\frac{\lambda_k}{\eta_i^k} \right\};$$

Then

$$A_i = \left\{ p_i | \max\left\{ 0, \max_{k \in I_i} \left[-\frac{\lambda_k}{\zeta_i^k} \right] \right\} \leqslant m_{i,p_i} \leqslant M_{i,P_i} \leqslant \min_{k \in \bar{I}_i} \left\{ -\frac{\lambda_k}{\zeta_i^k} \right\} \right\}$$
$$\cup \left\{ p_i | \max_{k \in \bar{J}_i} \left\{ -\frac{\lambda_k}{\eta_i^k} \right\} \leqslant m_{i,p_i} \leqslant M_{i,P_i} \leqslant \min 0, \left\{ \min_{k \in J_i} \left[-\frac{\lambda_k}{\eta_i^k} \right] \right\} \right\}. \tag{37}$$

Note from Remark 2.1 that if there exists $k \in \bar{I}_i \cap \bar{J}_i$ such that $\lambda_k < 0$, then $A_i = \emptyset$.

Remark 3.1. The set $A_{\rm i}$ depends on the transformation chosen. For instance, if we take $x_i = t_{i,p_i}(y_i) = y_i^{1/p_i}$, then we have m_{i,p_i} and M_{i,p_i} in (19) and (20), respectively. Thus we have

$$A_i = \left[1 + \max_{k \in \bar{I}_i} \frac{\lambda_k l_i}{\zeta_i^k}, 1 + \min \left\{ 0, \min_{k \in I_i} \left[\frac{\lambda_k u_i}{\zeta_i^k} \right] \right\} \right]$$
$$\cup \left[1 + \max \left\{ 0, \max_{k \in J_i} \left[\frac{\lambda_k u_i}{\eta_i^k} \right] \right\}, 1 + \min_{k \in \bar{J}_i} \frac{\lambda_k l_i}{\eta_i^k} \right].$$

Therefore, *if* $\max_{k \in \bar{I}_i} \frac{\lambda_k l_i}{\zeta_i^k} \leqslant \min\{0, \min_{k \in I_i} [\frac{\lambda_k u_i}{\zeta_i^k}]\}$ or $\max\{0, \max_{k \in J_i} [\frac{\lambda_k u_i}{\eta_i^k}]\} \leqslant \min_{k \in \bar{J}_i} \frac{\lambda_k l_i}{\eta_i^k}$, then the programming problem (21) is a convex programming problem for $p_i \in A_i, i = 1, 2, \ldots, n$.

Obviously, there are many transformations that may convert the primal problem (1) into a convex programming problem. A natural question to ask is if there is a best choice among different transformations. A transformation to serve as a best choice should have the largest freedom for p to perform convexification. From (37), it is clear that if a transformation $x = t_p(y)$ satisfies:

$$m_{i,p_i} = M_{i,p_i}, \quad \text{and} \quad \{M_{i,p_i} | p_i \in R\} = R, \quad \text{for any } i = 1, \ldots, n, \tag{38}$$

then $A_i \neq \emptyset$ when

$$\left\{\max\left\{0, \max_{k\in I_i}\left[-\frac{\lambda_k}{\zeta_i^k}\right]\right\} \leqslant \min_{k\in \bar{I}_i}\left[-\frac{\lambda_k}{\zeta_i^k}\right]\right\}$$

$$\cup\left\{\max_{k\in \bar{J}_i}\left[-\frac{\lambda_k}{\eta_i^k}\right] \leqslant \min\left\{0, \min_{k\in J_i}\left[-\frac{\lambda_k}{\eta_i^k}\right]\right\}\right\} \neq \emptyset.$$

There do exist transformations that satisfy (38). For example, if we take $t_{i,p_i}(y_i) = \frac{1}{p_i}\ln y_i$, or $t_{i,p_i}(y_i) = \frac{1}{p_i}\ln(1+\frac{y_i}{p_i})$, then we have $M_{i,p_i} = m_{i,p_i} = -p_i$ and $\{M_{i,p_i}|p_i \in R\} = R$; if we take $t_{i,p_i}(y_i) = -\frac{1}{p_i}\ln y_i$, or $t_{i,p_i}(y_i) = -\frac{1}{p_i}\ln(1+\frac{y_i}{p_i})$, then we have $M_{i,p_i} = m_{i,p_i} = p_i$ and $\{M_{i,p_i}|p_i \in R\} = R$.

By identifying the best choices among different transformations, we can now obtain the following conditions for hidden convex programming problems that are independent of transformations.

THEOREM 3.3. *Suppose* $g_j \in C^2(X), j = 0, 1, \ldots, m$. *If, in addition, for each* $i = 1, \ldots, n$, *at least one of the following inequalities holds*:

$$\max\left\{0, \max_{k\in I_i}\left[-\frac{\lambda_k}{\zeta_i^k}\right]\right\} \leqslant \min_{k\in \bar{I}_i}\left\{-\frac{\lambda_k}{\zeta_i^k}\right\} \tag{39}$$

or

$$\max_{k\in \bar{J}_i}\left\{-\frac{\lambda_k}{\eta_i^k}\right\} \leqslant \min\left\{0, \min_{k\in J_i}\left[-\frac{\lambda_k}{\eta_i^k}\right]\right\}, \tag{40}$$

then the problem (1) *is a hidden convex programming problem. If, for each* $i = 1, \ldots, n$, *the inequality* (39) *or the inequality* (40) *is strict, then problem* (1) *is a hidden strictly convex programming problem.*

Proof. Since for each $i = 1, \ldots, n$, one of (39) and (40) holds, if we take transformation $x = t_p(y)$ satisfying (38), then $A_i \neq \emptyset$ for all $i = 1, \ldots, n$. By Theorem 3.2, we know that the programming problem (21) is convex on Y_p when $p_i \in A_i, i = 1, \ldots, n$. Thus, the original problem (1) is a hidden convex problem. If (39) or (40) is strict and if we take transformation $x = t_p(y)$ satisfying (38), then $\bar{A}_i \neq \emptyset$. So the programming problem (21) is strictly convex on Y_p when $p_i \in \bar{A}_i, i = 1, \ldots, n$. Thus, the original problem (1) is a hidden strictly convex problem if (39) or (40) is strict for each $i = 1, \ldots, n$. □

We should emphasize here that no actual transformation is needed when judging if a problem is a hidden convex problem or not. If the conditions in Theorem 3.3 are satisfied, the primal problem (1) is a hidden convex problem. We can simply find the global minimum of (1) by using certain existing efficient local search algorithms in the literature.

EXAMPLE 3.1. We now consider the following nonconvex optimization problem:

$$\begin{aligned} \min \quad & g_0(x) = \frac{1}{3}x_1^3 + \frac{3}{2}x_2^2 - 3x_1 - 2x_2 \\ \text{s.t.} \quad & g_1(x) = -x_1^2 - \frac{1}{2}x_2^2 - 14x_1 - 15x_2 + 40 \leqslant 0 \\ & 1 \leqslant x_i \leqslant 2, \quad i = 1, 2. \end{aligned} \tag{41}$$

Obviously, we have that

$$\zeta_1^0 = -2 \leqslant \frac{\partial g_0(x)}{\partial x_1} = x_1^2 - 3 \leqslant 1 = \eta_1^0$$

$$\zeta_2^0 = 1 \leqslant \frac{\partial g_0(x)}{\partial x_2} = 3x_2 - 2 \leqslant 4 = \eta_2^0$$

$$\zeta_1^1 = -18 \leqslant \frac{\partial g_1(x)}{\partial x_1} = -2x_1 - 14 \leqslant -16 = \eta_1^1$$

$$\zeta_2^1 = -17 \leqslant \frac{\partial g_1(x)}{\partial x_2} = -x_2 - 15 \leqslant -16 = \eta_2^1.$$

Let $H_0(x)$ and $H_1(x)$ be the Hessian of g_0 and g_1 at x, respectively,

$$H_0(x) = \begin{bmatrix} 2x_1 & 0 \\ 0 & 3 \end{bmatrix},$$

$$H_1(x) = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}.$$

The minimal eigenvalue of matrix $H_0(x)$ is

$$\lambda_0(x) = \begin{cases} 2x_1 & 1 \leqslant x_1 \leqslant \frac{3}{2} \\ 3 & \frac{3}{2} < x_1 \leqslant 2 \end{cases}$$

for any $1 \leqslant x_2 \leqslant 2$ and the minimal eigenvalue of matrix $H_1(x)$ is $\lambda_1(x) = -2$ for any $x \in X$. So we have $\lambda_0 = \min_{x \in X} \lambda_0(x) = 2$ and $\lambda_1 = \min_{x \in X} \lambda_1(x) = -2$. Notice that the function $g_0(x)$ is convex and the function $g_1(x)$ is concave. It is easy to see that $I_1 = \emptyset, \bar{I}_1 = \{0, 1\},\ J_1 = \{1\}, \bar{J}_1 = \{0\}, I_2 = \{0\}, \bar{I}_2 = \{1\}, J_2 = \{1\}, \bar{J}_2 = \{0\}$. For $i = 1$, we have

$$\max_{k \in \bar{J}_1} \left\{ -\frac{\lambda_k}{\eta_1^k} \right\} = -\frac{\lambda_0}{\eta_1^0} = -2 < -\frac{1}{8} = \min\left\{ 0, -\frac{\lambda_1}{\eta_1^1} \right\} = \min\left\{ 0, \min_{k \in J_1} \left[-\frac{\lambda_k}{\eta_1^k} \right] \right\}.$$

(40) holds for $i = 1$; For $i = 2$, we have

$$\max_{k \in \bar{J}_2} \left\{ -\frac{\lambda_k}{\eta_2^k} \right\} = -\frac{\lambda_0}{\eta_2^0} = -\frac{1}{2} < -\frac{1}{8} = \min\left\{ 0, -\frac{\lambda_1}{\eta_2^1} \right\} = \min\left\{ 0, \min_{k \in J_2} \left[-\frac{\lambda_k}{\eta_2^k} \right] \right\},$$

(40) also holds for $i = 2$. Thus, by Theorem 3.3, the problem (41) is hidden strictly convex. In fact, the primal problem (41) has only one local minimum $x^* = (1.7321, 1.0000)^T$ with $f(x^*) = -3.9641$ and this local minimum is its global minimum.

We emphasize that no exact eigenvalues of the Hessian matrices are required when judging if a nonconvex problem is hidden convex or not. The proposed method only requires some estimated bounds of the minimum eigenvalues and the first-order derivatives. In the literature, there are many methods that estimate a lower bound for the minimum eigenvalue. A classical method is by using the Gerschgorin theorem. It is evident that the tighter the bound, the higher the probability for a problem to be convexified.

4. The Hidden Convex Quadratic Problem

We consider the following general quadratic problem in this section:

$$\begin{aligned}
\min \quad & g_0(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{(0)} x_i x_j + \sum_{i=1}^{n} b_i^{(0)} x_i + c^{(0)} \\
\text{s.t.} \quad & g_k(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{(k)} x_i x_j + \sum_{i=1}^{n} b_i^{(k)} x_i + c^{(k)} \leqslant 0, k = 1, \ldots, m \\
& x \in X = \{(x_1, \ldots, x_n) | l_i \leqslant x_i \leqslant u_i, i = 1, \ldots, n\}, \qquad (42)
\end{aligned}$$

where, without loss of generality, we assume that $a_{ij}^{(k)} = a_{ji}^{(k)}, i,j = 1, \ldots, n$, $k = 0, 1, \ldots, m$. The following are obvious.

$$\begin{aligned}
\frac{\partial g_k(x)}{\partial x_i} &= 2\sum_{j=1}^{n} a_{ij}^{(k)} x_j + b_i^{(k)}, \quad i = 1, \ldots, n, k = 0, 1, \ldots, m \\
\frac{\partial^2 g_k(x)}{\partial x_i \partial x_j} &= 2a_{ij}^{(k)}, \quad i,j = 1, \ldots, n, k = 0, 1, \ldots, m.
\end{aligned}$$

Let

$$\zeta_i^k = \left[2\left(\sum_{j=1}^{n} a_{ij}^{(k)} l_j \cdot s\left(a_{ij}^{(k)}\right) + a_{ij}^{(k)} u_j \cdot s\left(-a_{ij}^{(k)}\right)\right) + b_i^{(k)}\right]$$

$$\eta_i^k = \left[2\left(\sum_{j=1}^{n} a_{ij}^{(k)} u_j \cdot s\left(a_{ij}^{(k)}\right) + a_{ij}^{(k)} l_j \cdot s\left(-a_{ij}^{(k)}\right)\right) + b_i^{(k)}\right]$$

$$I_i = \{k \mid \zeta_i^k > 0, \quad k = 0, 1, \ldots, m\}$$

$$\tilde{I}_i = \{k \mid \zeta_i^k < 0, \quad k = 0, 1, \ldots, m\}$$

$$I_i^0 = \{k \mid \zeta_i^k = 0, \quad k = 0, 1, \ldots, m\}$$

$$J_i = \{k \mid \eta_i^k < 0, \quad k = 0, 1, \ldots, m\}$$

$$\tilde{J}_i = \{k \mid \eta_i^k > 0, \quad k = 0, 1, \ldots, m\}$$

$$J_i^0 = \{k \mid \eta_i^k = 0, \quad k = 0, 1, \ldots, m\},$$

where function $s(\cdot)$ is defined in (8). Let G_k be the Hessian matrix of $g_k(x)$. The following are clear.

$$\zeta_i^k \leqslant \frac{\partial g_k(x)}{\partial x_i} \leqslant \eta_i^k, \quad \forall x \in X, i = 1, \ldots, n, k = 0, 1, \ldots, m$$
$$G_k = 2(a_{ij}^{(k)})_{n \times n}, k = 0, 1, \ldots, m.$$

Let λ_k be the minimal eigenvalue of G_k. Obviously, the following results can be obtained from Theorem 3.3.

(i) If $\lambda_k \geqslant 0, \forall k = 0, 1, \ldots, m$, then the problem (42) is already a convex programming problem.

(ii) If for all $i = 1, \ldots, n$, one of the following two conditions is satisfied:

$$\lambda_k \geqslant 0 \quad \text{for all } k \in I_i^0 \text{ and } \min\left\{0, \min_{k \in I_i} \left[\frac{\lambda_k}{\zeta_i^k}\right]\right\} \geqslant \max_{k \in \tilde{I}_i} \frac{\lambda_k}{\zeta_i^k},$$

or

$$\lambda_k \geqslant 0 \quad \text{for all } k \in J_i^0 \text{ and } \min_{k \in \tilde{J}_i} \frac{\lambda_k}{\eta_i^k} \geqslant \max\left\{0, \max_{k \in J_i} \left[\frac{\lambda_k}{\eta_i^k}\right]\right\},$$

then the problem (42) must be a hidden convex programming problem.

We now consider the following three special cases of the problem (42):

(a) If $a_{ij}^{(k)} = 0, \quad i, j = 1, \ldots, n, \ k = 1, \ldots, m,$ the problem (42)

reduces to:

$$\begin{aligned} \min \quad & g_0(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^{(0)} x_i x_j + \sum_{i=1}^n b_i^{(0)} x_i + c^{(0)} \\ \text{s.t.} \quad & g_k(x) = \sum_{i=1}^n b_i^{(k)} x_i + c^{(k)} \leqslant 0, \quad k = 1, \ldots, m \\ & x \in X = \{(x_1, \ldots, x_n) | l_i \leqslant x_i \leqslant u_i, i = 1, \ldots, n\}. \end{aligned} \tag{43}$$

This is a quadratic programming problem with linear constraints. Notice that in this situation, $\lambda_k = 0$ for all $k = 1, \ldots, m$, and $\zeta_i^k = \eta_i^k = b_i^{(k)}$ for all $i = 1, \ldots, n$ and $k = 1, \ldots, m$.

If $\lambda_0 \geqslant 0$, the problem (43) is already a convex programming problem. When $\lambda_0 < 0$, if for all $i = 1, \ldots, n$, one of the following two conditions is satisfied:

$$\zeta_i^0 > 0 \text{ and } b_i^{(k)} \cdot \zeta_i^0 > 0, \quad \forall k = 1, \ldots, m \text{ (i.e., } I_i = \{0, 1, \ldots, m\} \text{ and } \tilde{I}_i = \emptyset),$$

or

$$\zeta_i^0 < 0 \text{ and } b_i^{(k)} \cdot \zeta_i^0 > 0, \quad \forall k = 1, \ldots, m \text{ (i.e., } J_i = \{0, 1, \ldots, m\} \text{ and } \tilde{J}_i = \emptyset),$$

then the programming problem (43) must be a hidden convex programming problem.

(b) If $a_{ij}^{(k)} = 0, \ i \neq j, \ i, j = 1, \ldots, n, \ k = 0, 1, \ldots, m$, the problem (42) reduces to:

$$\begin{aligned} \min \quad & g_0(x) = \sum_{i=1}^{n} a_{ii}^{(0)} x_i^2 + \sum_{i=1}^{n} b_i^{(0)} x_i + c^{(0)} \\ \text{s.t.} \quad & g_k(x) = \sum_{i=1}^{n} a_{ii}^{(k)} x_i^2 + \sum_{i=1}^{n} b_i^{(k)} x_i + c^{(k)} \leqslant 0, \quad k = 1, \ldots, m \\ & x \in X = \{(x_1, \ldots, x_n) \mid l_i \leqslant x_i \leqslant u_i, \quad i = 1, \ldots, n\}. \end{aligned} \tag{44}$$

Then we have

$$\lambda_k = \min_{1 \leqslant i \leqslant n} 2a_{ii}^{(k)}.$$

Let

$$I_i = \left\{ k \mid 2a_{ii}^{(k)} l_i \cdot s\left(a_{ii}^{(k)}\right) + 2a_{ii}^{(k)} u_i \cdot s\left(-a_{ii}^{(k)}\right) + b_i^{(k)} > 0, \ k = 0, 1, \ldots, m \right\}$$

$$\tilde{I}_i = \left\{ k \mid 2a_{ii}^{(k)} l_i \cdot s\left(a_{ii}^{(k)}\right) + 2a_{ii}^{(k)} u_i \cdot s\left(-a_{ii}^{(k)}\right) + b_i^{(k)} < 0, \ k = 0, 1, \ldots, m \right\}$$

$$I_i^0 = \left\{ k \mid 2a_{ii}^{(k)} l_i \cdot s\left(a_{ii}^{(k)}\right) + 2a_{ii}^{(k)} u_i \cdot s\left(-a_{ii}^{(k)}\right) + b_i^{(k)} = 0, \ k = 0, 1, \ldots, m \right\}$$

$$J_i = \left\{ k \mid 2a_{ii}^{(k)} u_i \cdot s\left(a_{ii}^{(k)}\right) + 2a_{ii}^{(k)} l_i \cdot s\left(-a_{ii}^{(k)}\right) + b_i^{(k)} < 0, \ k = 0, 1, \ldots, m \right\}$$

$$\tilde{J}_i = \left\{ k \mid 2a_{ii}^{(k)} u_i \cdot s\left(a_{ii}^{(k)}\right) + 2a_{ii}^{(k)} l_i \cdot s\left(-a_{ii}^{(k)}\right) + b_i^{(k)} > 0, \ k = 0, 1, \ldots, m \right\}$$

$$J_i^0 = \left\{ k \mid 2a_{ii}^{(k)} u_i \cdot s\left(a_{ii}^{(k)}\right) + 2a_{ii}^{(k)} l_i \cdot s\left(-a_{ii}^{(k)}\right) + b_i^{(k)} = 0, \ k = 0, 1, \ldots, m \right\}.$$

If for all $i = 1, \ldots, n$, one of the following two conditions is satisfied:
$\lambda_k \geqslant 0$ for all $k \in I_i^0$ and

$$\min\left\{0, \min_{k\in I_i} \frac{\min\limits_{1\leqslant i\leqslant n} a_{ii}^{(k)}}{2a_{ii}^{(k)} l_i \cdot s(a_{ii}^{(k)}) + 2a_{ii}^{(k)} u_i \cdot s(-a_{ii}^{(k)}) + b_i^{(k)}}\right\}$$

$$\geqslant \max_{k\in \tilde{I}_i} \frac{\min\limits_{1\leqslant i\leqslant n} a_{ii}^{(k)}}{2a_{ii}^{(k)} l_i \cdot s(a_{ii}^{(k)}) + 2a_{ii}^{(k)} u_i \cdot s(-a_{ii}^{(k)}) + b_i^{(k)}}, \tag{45}$$

or

$\lambda_k \geqslant 0 \quad$ for all $\quad k \in J_i^0 \quad$ and

$$\min_{k\in \tilde{J}_i} \frac{\min\limits_{1\leqslant i\leqslant n} a_{ii}^{(k)}}{2a_{ii}^{(k)} u_i \cdot s(a_{ii}^{(k)}) + 2a_{ii}^{(k)} l_i \cdot s\big(-a_{ii}^{(k)}\big) + b_i^{(k)}}$$

$$\geqslant \max\left\{0, \max_{k\in J_i} \frac{\min\limits_{1\leqslant i\leqslant n} a_{ii}^{(k)}}{2a_{ii}^{(k)} u_i \cdot s\big(a_{ii}^{(k)}\big) + 2a_{ii}^{(k)} l_i \cdot s\big(-a_{ii}^{(k)}\big) + b_i^{(k)}}\right\}, \tag{46}$$

then the problem (44) must be a hidden convex programming problem.

Consider the singly and quadratically constrained quadratic problem considered in Ben-Tal and Teboulle [2]:

$$\min \quad g_0(x) = \sum_{i=1}^{n} a_{ii}^{(0)} x_i^2 + \sum_{i=1}^{n} b_i^{(0)} x_i + c^{(0)}$$

$$\text{s.t.} \quad c_l \leqslant g_1(x) = \sum_{i=1}^{n} a_{ii}^{(1)} x_i^2 \leqslant c_u \tag{47}$$

As suggested by the double duality approach in [2], problem (47) is a hidden convex problem which is evidenced by a transformation of $x_i = -\text{sgn}(b_i^{(0)})\sqrt{y_i}$. Notice that the transformation in [2] is not an equivalent transformation. Selecting the transformation of $x_i = -\text{sgn}(b_i^{(0)})\sqrt{y_i}$ discards its complement transformation $x_i = \text{sgn}(b_i^{(0)})\sqrt{y_i}$ which will be never optimal in problem (47).

(c) If $n = 2$, the problem (42) reduces to the form:

$$\min \quad g_0(x) = a_{11}^{(0)} x_1^2 + 2a_{12}^{(0)} x_1 x_2 + a_{22}^{(0)} x_2^2 + b_1^{(0)} x_1 + b_2^{(0)} x_2 + c^{(0)}$$

$$\text{s.t.} \quad g_k(x) = a_{11}^{(k)} x_1^2 + 2a_{12}^{(k)} x_1 x_2 + a_{22}^{(k)} x_2^2 + b_1^{(k)} x_1 + b_2^{(k)} x_2 + c^{(k)} \leqslant 0 \tag{48}$$

$$k = 1, \ldots, m$$

$$x \in X = \{(x_1, x_2) \mid l_i \leqslant x_i \leqslant u_i, i = 1, 2\}.$$

Then we have

$$\lambda_k = a_{11}^{(k)} + a_{22}^{(k)} - \sqrt{\left(a_{11}^{(k)} - a_{22}^{(k)}\right)^2 + 4\left(a_{12}^{(k)}\right)^2}.$$

Let

$$I_i = \left\{ k \,\middle|\, \begin{array}{l} 2\left(a_{ii}^{(k)} l_i \cdot s(a_{ii}^{(k)}) + a_{ii}^{(k)} u_i \cdot s(-a_{ii}^{(k)}) + a_{12}^{(k)} l_i \cdot s(a_{12}^{(k)}) + a_{12}^{(k)} u_i \cdot s(-a_{12}^{(k)})\right) + b_i^{(k)} > 0, \\ k = 0, 1, \ldots, m \end{array} \right\}$$

$$\tilde{I}_i = \left\{ k \,\middle|\, \begin{array}{l} 2\left(a_{ii}^{(k)} l_i \cdot s(a_{ii}^{(k)}) + a_{ii}^{(k)} u_i \cdot s(-a_{ii}^{(k)}) + a_{12}^{(k)} l_i \cdot s(a_{12}^{(k)}) + a_{12}^{(k)} u_i \cdot s(-a_{12}^{(k)})\right) + b_i^{(k)} < 0, \\ k = 0, 1, \ldots, m \end{array} \right\}$$

$$I_i^0 = \left\{ k \,\middle|\, \begin{array}{l} 2\left(a_{ii}^{(k)} l_i \cdot s(a_{ii}^{(k)}) + a_{ii}^{(k)} u_i \cdot s(-a_{ii}^{(k)}) + a_{12}^{(k)} l_i \cdot s(a_{12}^{(k)}) + a_{12}^{(k)} u_i \cdot s(-a_{12}^{(k)})\right) + b_i^{(k)} = 0, \\ k = 0, 1, \ldots, m \end{array} \right\}$$

$$J_i = \left\{ k \,\middle|\, \begin{array}{l} 2\left(a_{ii}^{(k)} u_i \cdot s(a_{ii}^{(k)}) + a_{ii}^{(k)} l_i \cdot s(-a_{ii}^{(k)}) + a_{12}^{(k)} u_i \cdot s(a_{12}^{(k)}) + a_{12}^{(k)} l_i \cdot s(-a_{12}^{(k)})\right) + b_i^{(k)} < 0, \\ k = 0, 1, \ldots, m \end{array} \right\}$$

$$\tilde{J}_i = \left\{ k \,\middle|\, \begin{array}{l} 2\left(a_{ii}^{(k)} u_i \cdot s(a_{ii}^{(k)}) + a_{ii}^{(k)} l_i \cdot s(-a_{ii}^{(k)}) + a_{12}^{(k)} u_i \cdot s(a_{12}^{(k)}) + a_{12}^{(k)} l_i \cdot s(-a_{12}^{(k)})\right) + b_i^{(k)} > 0, \\ k = 0, 1, \ldots, m \end{array} \right\}$$

$$J_i^0 = \left\{ k \,\middle|\, \begin{array}{l} 2\left(a_{ii}^{(k)} u_i \cdot s(a_{ii}^{(k)}) + a_{ii}^{(k)} l_i \cdot s(-a_{ii}^{(k)}) + a_{12}^{(k)} u_i \cdot s(a_{12}^{(k)}) + a_{12}^{(k)} l_i \cdot s(-a_{12}^{(k)})\right) + b_i^{(k)} = 0, \\ k = 0, 1, \ldots, m \end{array} \right\}$$

If for both $i = 1, 2$, one of the following two conditions is satisfied: $\lambda_k \geqslant 0$ for all $k \in I_i^0$ and

$$\min\left\{0,\ \min_{k \in I_i} \frac{a_{11}^{(k)} + a_{22}^{(k)} - \sqrt{\left(a_{11}^{(k)} - a_{22}^{(k)}\right)^2 + 4\left(a_{12}^{(k)}\right)^2}}{2\left(a_{ii}^{(k)} l_i \cdot s(a_{ii}^{(k)}) + a_{ii}^{(k)} u_i \cdot s(-a_{ii}^{(k)}) + a_{12}^{(k)} l_i \cdot s(a_{12}^{(k)}) + a_{12}^{(k)} u_i \cdot s(-a_{12}^{(k)})\right) + b_i^{(k)}}\right\}$$

$$\geqslant \max_{k \in \tilde{I}_i} \frac{a_{11}^{(k)} + a_{22}^{(k)} - \sqrt{\left(a_{11}^{(k)} - a_{22}^{(k)}\right)^2 + 4(a_{12}^{(k)})^2}}{2\left(a_{ii}^{(k)} l_i \cdot s(a_{ii}^{(k)}) + a_{ii}^{(k)} u_i \cdot s(-a_{ii}^{(k)}) + a_{12}^{(k)} l_i \cdot s(a_{12}^{(k)}) + a_{12}^{(k)} u_i \cdot s(-a_{12}^{(k)})\right) + b_i^{(k)}},$$

or

$\lambda_k \geqslant 0$ for all $k \in J_i^0$ and

$$\max\left\{0,\ \max_{k\in J_i}\frac{a_{11}^{(k)}+a_{22}^{(k)}-\sqrt{(a_{11}^{(k)}-a_{22}^{(k)})^2+4(a_{12}^{(k)})^2}}{2\Big(a_{ii}^{(k)}u_i\cdot s(a_{ii}^{(k)})+a_{ii}^{(k)}l_i\cdot s(-a_{ii}^{(k)})+a_{12}^{(k)}u_i\cdot s(a_{12}^{(k)})+a_{12}^{(k)}l_i\cdot s(-a_{12}^{(k)})\Big)+b_i^{(k)}}\right\}$$

$$\leqslant \min_{k\in\bar{J}_i}\frac{a_{11}^{(k)}+a_{22}^{(k)}-\sqrt{(a_{11}^{(k)}-a_{22}^{(k)})^2+4(a_{12}^{(k)})^2}}{2\Big(a_{ii}^{(k)}u_i\cdot s(a_{ii}^{(k)})+a_{ii}^{(k)}l_i\cdot s(-a_{ii}^{(k)})+a_{12}^{(k)}u_i\cdot s(a_{12}^{(k)})+a_{12}^{(k)}l_i\cdot s(-a_{12}^{(k)})\Big)+b_i^{(k)}},$$

then the problem (48) must be a hidden convex programming problem.

In conclusion, for any of the three problems (43), (44) and (48), we can determine if it is a hidden convex problem or not just by examining its coefficients.

EXAMPLE 2:

$$\begin{aligned}
\min\quad & g_0(x) = x_1^2 + \frac{3}{2}x_2^2 - x_1 - 7x_2\\
\text{s.t.}\quad & g_1(x) = -x_1^2 + x_2^2 - 2x_1 + 3x_2 - 5 \leqslant 0\\
& g_2(x) = \frac{1}{2}x_1^2 + \frac{3}{2}x_2^2 - \frac{3}{2}x_1 - 2x_2 \leqslant 0\\
& \quad 1 \leqslant x_i \leqslant 2,\quad i = 1, 2.
\end{aligned} \tag{49}$$

The following can be derived easily.

For $i = 1$,

$$2a_{11}^{(k)}l_1\cdot s(a_{11}^{(k)}) + 2a_{11}^{(k)}u_1\cdot s(-a_{11}^{(k)}) + b_1^{(k)} = \begin{cases} 1 & k=0\\ -6 & k=1\\ -\frac{1}{2} & k=2,\end{cases}$$

$$2a_{11}^{(k)}u_1\cdot s(a_{11}^{(k)}) + 2a_{11}^{(k)}l_1\cdot s(-a_{11}^{(k)}) + b_1^{(k)} = \begin{cases} 3 & k=0\\ -4 & k=1\\ \frac{1}{2} & k=2.\end{cases}$$

For $i = 2$,

$$2a_{22}^{(k)}l_2\cdot s(a_{22}^{(k)}) + 2a_{22}^{(k)}u_2\cdot s(-a_{22}^{(k)}) + b_2^{(k)} = \begin{cases} -4 & k=0\\ 5 & k=1\\ 1 & k=2,\end{cases}$$

$$2a_{22}^{(k)}u_2\cdot s(a_{22}^{(k)}) + 2a_{22}^{(k)}l_2\cdot s(-a_{22}^{(k)}) + b_2^{(k)} = \begin{cases} -1 & k=0\\ 7 & k=1\\ 4 & k=2.\end{cases}$$

It is easy to verify $I_1 = \{0\},\ \tilde{I}_1 = \{1,2\},\ I_1^0 = \emptyset,\ J_1 = \{1\},\ \tilde{J}_1 = \{0,2\}, J_1^0 = \emptyset, I_2 = \{1,2\},\ \tilde{I}_2 = \{0\}, I_2^0 = \emptyset,\ J_2 = \{0\},\ \tilde{J}_2 = \{1,2\},\ J_2^0 = \emptyset$. Furthermore, we have

$$\min_{1\leqslant i\leqslant 2} a_{ii}^k = \begin{cases} 1 & k=0 \\ -1 & k=1 \\ \frac{1}{2} & k=2. \end{cases}$$

For $i=1$, we have

$$\begin{aligned}
&\min_{k\in\tilde{J}_i} \frac{\min_{1\leqslant i\leqslant 2} a_{ii}^{(k)}}{2a_{11}^{(k)}u_1\cdot s(a_{11}^{(k)}) + 2a_{11}^{(k)}l_1\cdot s(-a_{11}^{(k)}) + b_1^{(k)}} \\
&= \min\left\{\frac{1}{3},\frac{1/2}{1/2}\right\} = \frac{1}{3} > \frac{1}{4} = \max\left\{0,\frac{-1}{-4}\right\} \\
&= \max\left\{0,\ \max_{k\in J_1} \frac{\min_{1\leqslant i\leqslant 2} a_{11}^{(k)}}{2a_{11}^{(k)}u_1\cdot s(a_{11}^{(k)}) + 2a_{11}^{(k)}l_1\cdot s(-a_{11}^{(k)}) + b_1^{(k)}}\right\}.
\end{aligned}$$

Thus (46) holds. Also for $i=2$, we have

$$\begin{aligned}
&\min\left\{0,\ \min_{k\in I_2} \frac{\min_{1\leqslant i\leqslant 2} a_{ii}^{(k)}}{2a_{22}^{(k)}l_2\cdot s(a_{22}^{(k)}) + 2a_{22}^{(k)}u_2\cdot s(-a_{22}^{(k)}) + b_2^{(k)}}\right\} \\
&= \min\left\{0,\frac{-1}{5},\frac{1/2}{1}\right\} = -\frac{1}{5} > -\frac{1}{4} \\
&= \max_{k\in\tilde{I}_2} \frac{\min_{1\leqslant i\leqslant 2} a_{ii}^{(k)}}{2a_{22}^{(k)}l_2\cdot s(a_{22}^{(k)}) + 2a_{22}^{(k)}u_2\cdot s(-a_{22}^{(k)}) + b_2^{(k)}}.
\end{aligned}$$

Thus (45) holds. The problem (49) is therefore a hidden strictly convex programming problem. Its unique minimum $x^* = (1.0402,\ 1.7268)$ is a global minimal point with the global minimal value of $f^* = -7.5731$.

5. Epilogue

A hidden convex problem has an equivalent counterpart in a form of a convex programming. This paper provides certain sufficient conditions to identify hidden convex programming problems which consist a subclass of nonconvex optimization problems. The results in this paper extend the reach of convex programming to hidden convex programming, thus expanding significantly the problem domain whose local optimum is also global at the same time.

Although no actual transformation is needed when checking the hidden convexity of a problem, we should emphasize that numerical methods for solving convex problems perform much better in general than numerical methods for local minimization. Thus, trade-off exists between whether or not performing an actual transformation when solving a hidden convex problem.

An interesting research subject to explore is how likely a nonconvex minimization problem is hidden convex, or the percentage of hidden convex problems in nonconvex optimization. To illustrate this point, we will discuss in the following a one-dimensional quadratic optimization problem with a single constraint.

$$\begin{aligned} \min \quad & f(x) = e_o x^2 + ax + c \\ \text{s.t.} \quad & g(x) = e_1 x^2 + bx + d \leqslant 0 \\ & 1 \leqslant x \leqslant 2, \end{aligned} \tag{50}$$

where random parameters $e_o, e_1 \in \{1, -1\}$, random parameters $a, b \in [-l, l]$ and l is a positive number. Assume that e_0, e_1, a and b are independent. Assume further that both e_0 and e_1 have the same probabilities of taking either 1 or -1, while both a and b are uniformly distributed in $[-l, l]$. The probability that the quadratic problem (50) is hidden convex will be investigated in the following.

Obviously, if both e_0 and e_1 are equal to one, the problem (50) is a convex programming problem. Otherwise, it is not a convex programming problem. Thus the probability that the problem (50) is a convex programming problem is $p_1 = 1/4$.

Let p_h denote the probability of a hidden convexity of the problem (50). Notice that

$$\begin{aligned} & f'(x) = 2e_0 x + a, \quad f'' = 2e_0, \\ & g'(x) = 2e_1 x + b, \quad g'' = 2e_1. \end{aligned}$$

(Case i) When $e_0 = 1, e_1 = -1$, we have that

$$\begin{aligned} 2 + a \leqslant f'(x) = 2x + a \leqslant 4 + a \\ -4 + b \leqslant g'(x) = -2x + b \leqslant -2 + b \\ f'' = 2,\ g'' = -2. \end{aligned}$$

It can be derived (see [9] for details) that in the case of $e_0 = 1$ and $e_1 = -1$, the probability that the problem (50) is hidden convex is bounded from below by

$$p_2 = p_2^1 + p_2^2 + p_2^3 + p_2^4 \begin{cases} 0 & l \leqslant 1 \\ \dfrac{(2l-2)^2}{8l^2} & 1 < l \leqslant 4 \\ \dfrac{3l^2 - 6l - 6}{4l^2} & l > 4. \end{cases}$$

(Case ii) Similarly, in the case of $e_0 = -1$ and $e_1 = 1$, the probability that the problem (50) is hidden convex is bounded from below by p_3 which has the same form as p_2.

(Case iii) When $e_0 = -1, e_1 = -1$, we have that

$$\begin{aligned} &-4 + a \leqslant f'(x) = -2x + a \leqslant -2 + a \\ &-4 + b \leqslant g'(x) = -2x + b \leqslant -2 + b \\ &f'' = -2, \ g'' = -2. \end{aligned}$$

It can be derived (see [9] for details) that in the case of $e_0 = -1$ and $e_1 = -1$, the probability that the problem (50) is hidden convex is bounded from below by

$$p_4 = p_4^1 + p_4^2 = \begin{cases} 1 & l \leqslant 2 \\ \dfrac{l^2 + 4l + 4}{4l^2} & 2 < l \leqslant 4 \\ \dfrac{l^2 - 2l + 10}{2l^2} & l > 4. \end{cases}$$

Finally, the probability that the problem (50) is hidden convex is bounded from below by

$$p_h = p_1 + \frac{1}{4}p_2 + \frac{1}{4}p_3 + \frac{1}{4}p_4 = \begin{cases} \dfrac{1}{4} + \dfrac{1}{4} & l \leqslant 1 \\ \dfrac{1}{4} + \dfrac{2l^2 - 2l + 1}{4l^2} & 1 < l \leqslant 2 \\ \dfrac{1}{4} + \dfrac{5l^2 - 4l + 8}{16l^2} & 2 < l \leqslant 4 \\ \dfrac{1}{4} + \dfrac{(l-1)^2}{2l^2} & l > 4. \end{cases}$$

It is easy to check that $p_h \geqslant \frac{1}{4} + \frac{9}{32}$ when $l > 4$ since $(l-1)^2/2l^2$ is increasing when $l > 4$. Furthermore, we have that

$$\lim_{l \to +\infty} p_h = \frac{3}{4}.$$

The outcome that 2/3 of the nonconvex problems are hidden convex in the above example is surprisingly revealing! In the real world, there may exist many programming problems which are hidden convex but with a nonconvex representation. One important conclusion is that convexity/nonconvexity is not an inherent property in many situations. Rather it is a character associated with a given representation space. Study of the hidden convexity of a programming problem, thus, is a very prominent research topic to pursue further.

Acknowledgements

This research was partially supported by the Research Grants Council of Hong Kong under Grant CUHK4214/01E and the National Science Foundation of China under Grant 10171118. The authors appreciate very much the comments from the four reviewers.

References

1. Avriel, M., Diewert, W.E., Schaible S. and Zang, I. (1988), *Generalized Concavity*, Plenum Publishing Corporation, New York.
2. Ben-Tal, A. and Teboulle, M. (1996), Hidden Convexity in Some Nonconvex Quadratically Constrained Quadratic Programming, *Mathematical Programming* 72, 51–63.
3. Craven, B.D. (1977), Invex Functions and Constrained Local Minima, *Bulletin of the Australian Mathematical Society* 16, 325–339.
4. Horst, R. (1984), On the Convexification of Nonlinear Programming Problems: An Applications-oriented Survey, *European Journal of Operations Research* 15, 382–392.
5. Li, D., Sun, X.L., Biswal, M.P. and Gao, F. (2001), Convexification, Concavification and Monotonization in Global Optimization, *Annals of Operations Research* 105, 213–226.
6. Mangasarian, O.L. (1994), *Nonlinear Programming*, SIAM, Philadelphia.
7. Sun, X.L., McKinnon K. and Li, D. (2001), A Convexification Method for a Class of Global Optimization Problem with Application to Reliability Optimization, *Journal of Global Optimization* 21, 185–199.
8. Schaible, S. and Ziemba, W.T (1981), (eds.) *Generalized Concavity in Optimization and Economics*, Academic Press, New York.
9. Wu, Z.Y. (2003), Some deterministic methods for global optimization, Ph.D. dissertation, Department of Mathematics, Shanghai University, Shanghai, China.