数理解析研究所講究録
1114 巻 1999 年 13-21

Moderate Nonconvexity = Convexity + Quadratic Concavity

東京工業大学情報理工学研究科 小島政和 (Masakazu Kojima)
帝京科学大学電子・情報科学科 松本敏浩 (Toshihiro Matsumoto)
神奈川大学数学教室 信太正之 (Masayuki Shida)

1 Introduction.

This short note is concerned with two types of optimization problems. The one is a nonlinear program:

$$\text{maximize } \boldsymbol{d}^T\boldsymbol{u} \quad \text{subject to } \boldsymbol{u} \in G_0 \text{ and } f(\boldsymbol{u}) \leq 0 \ (\forall f(\cdot) \in \mathcal{F}) \tag{1}$$

Here

$\boldsymbol{d}$: a constant column vector in the ℓ-dimensional Euclidean space R^ℓ,
$\boldsymbol{d}^T$: the transposition of $\boldsymbol{d}$,
$\boldsymbol{u}$: a variable vector in R^ℓ ,
G_0 : a convex subset of R^ℓ,
$\mathcal{F}$: a class of finitely or infinitely many real valued functions on R^ℓ.

We may start with a more general nonlinear program having a nonlinear objective function $f_0(\boldsymbol{u})$, but we can always reduce such a problem to a nonlinear program with a linear objective function. In fact, if we replace $f_0(\boldsymbol{u})$ by a scalar variable $u_0 \in R$ and add the inequality $u_0 - f_0(\boldsymbol{u}) \leq 0$ to the constraints, we have a nonlinear program of the form (1).

The other problem is a general quadratic program:

$$\text{maximize } \boldsymbol{c}^T\boldsymbol{x} \quad \text{subject to } \boldsymbol{x} \in C_0 \text{ and } q(\boldsymbol{x}) \leq 0. \tag{2}$$

Here

$\boldsymbol{c}$: a constant column vector in R^n,
$\boldsymbol{x}$: a variable vector in R^n,
$q(\cdot)$: a concave quadratic function on R^n ,
C_0 : a convex subset of R^n.

The problem (2) is a special case of d.c. (difference of two convex functions) programs. Conversion of general nonlinear programs into d.c. programs has been extensively studied in the field of global optimization. See, for example, [2, 3, 9]. In theory, it is known that any closed subset G of R^ℓ can be represented as

$$G = \{\boldsymbol{u} \in R^\ell : \phi(\boldsymbol{u}) - \|\boldsymbol{u}\|^2 \leq 0\}$$

using some convex function $\phi(\cdot) : R^\ell \to R$. See Corollary 3.5 of [9]. Hence, given any closed subset G of R^ℓ, we can reduce a maximization of a linear function $\boldsymbol{d}^T\boldsymbol{u}$ over G into the problem (2) by taking

$$\left.\begin{array}{l} n = \ell + 1,\ \boldsymbol{x} = \begin{pmatrix} \boldsymbol{u} \\ t \end{pmatrix} \in R^n,\ \boldsymbol{c} = \begin{pmatrix} \boldsymbol{d} \\ 0 \end{pmatrix} \in R^n, \\ q(\boldsymbol{x}) = t - \|\boldsymbol{u}\|^2,\ C_0 = \{\boldsymbol{x} \in R^n : \phi(\boldsymbol{u}) - t \leq 0\}. \end{array}\right\} \tag{3}$$

In general, however, this conversion is neither practical nor implementable because an explicit algebraic representation of such a convex function $\phi(\cdot) : R^\ell \to R$ is not available.

The purpose of this short note is to add an observation that if the nonconvexity of every $f(\cdot) \in \mathcal{F}$ is moderate, then we can reduce the nonlinear program (1) to a general quadratic program of the form (2). Recently Kojima and Tunçel [4] proposed a class of successive convex relaxation methods for nonconvex quadratic programs. See also [5, 8]. In applying their methods to an optimization problem, it is necessary and also essential to transform the problem into an optimization problem having a linear objective function, finitely or infinitely many quadratic inequality constraints and an additional compact convex constraint set. Since (2) is a special case of such optimization problems whenever the constraint set C_0 is compact, our observation opens up a possibility of applying their methods to general nonlinear programs. See also Remark (B) of Section 4.

We introduce a class $\mathcal{F}_{\text{p.conv}}(G)$ of peri-convex functions on a convex subset G of R^ℓ. For every $f : R^\ell \to R$, let

$$\sigma(f, G) \equiv \inf\{\sigma \geq 0 : f(\cdot) + \sigma\|\cdot\|^2 \text{ is convex on } G\}.$$

Then $f : R^\ell \to R$ lies in $\mathcal{F}_{\text{p.conv}}(G)$ if and only if $\sigma(f, G) < \infty$. Obviously, if $f(\cdot) : R^\ell \to R$ is convex on G, then it is in the class $\mathcal{F}_{\text{p.conv}}(G)$. In addition, the class $\mathcal{F}_{\text{p.conv}}(G)$ contains "moderately nonconvex" functions. But, for example, the function $-\|\cdot\| : R^\ell \to R$ is not peri-convex for any convex open neighborhood of $\boldsymbol{0} \in R^\ell$. We also note that if $f(\cdot)$ s a peri-convex function on a subset G of R^ℓ , then it is continuous in the interior of G.

In Section 2,

- we show that if every $f(\cdot) \in \mathcal{F}$ is peri-convex on G_0 then we can reduce the nonlinear program (1) to a general quadratic program of the form (2).

In Section 3, we present some properties of peri-convex functions. In particular,

- we characterize peri-convex functions in terms of proximal subdifferentiability [1], and
- we show that "moderately nonlinear" functions are peri-convex.

2 Reduction of the Nonlinear Program (1) to the General Quadratic Program (2).

Assume that $\mathcal{F} \subset \mathcal{F}_{\text{p.conv}}(G_0)$ and that $\sigma(f, G) \leq \sigma_f < +\infty$ for each $f(\cdot) \in \mathcal{F}$. Then the functions

$$g(\cdot; \sigma_f, f) = f(\cdot) + \sigma_f\|\cdot\|^2 : R^\ell \to R \ (f(\cdot) \in \mathcal{F}) \tag{4}$$

Then $g(\cdot;\sigma_f,f): R^\ell \to R$ are convex on G_0. Now we rewrite the nonlinear program (1) as

$$\left.\begin{array}{ll} \text{maximize} & \boldsymbol{d}^T\boldsymbol{u} \\ \text{subject to} & \boldsymbol{u} \in G_0,\ t - \|\boldsymbol{u}\|^2 \leq 0, \\ & g(\boldsymbol{u};\sigma_f,f) - \sigma_f t \leq 0\ (\forall f(\cdot) \in \mathcal{F}). \end{array}\right\}$$

Defining

$$\left.\begin{array}{l} n = \ell + 1,\ \boldsymbol{x} = \begin{pmatrix} \boldsymbol{u} \\ t \end{pmatrix} \in R^n,\ \boldsymbol{c} = \begin{pmatrix} \boldsymbol{d} \\ \boldsymbol{0} \end{pmatrix} \in R^n, \\ q(\boldsymbol{x}) = t - \|\boldsymbol{u}\|^2, \\ h(\boldsymbol{x};\sigma_f,f) = g(\boldsymbol{u};\sigma_f,f) - \sigma_f t\ (f(\cdot) \in \mathcal{F}), \\ C_0 = \left\{\boldsymbol{x} \in R^n :\ \boldsymbol{u} \in G_0,\ h(\boldsymbol{x};\sigma_f,f) \leq 0\ (f(\cdot) \in \mathcal{F})\ \right\}, \end{array}\right\} \qquad (5)$$

we thus obtain the quadratic program (2), which is equivalent to the nonlinear program (1) as we will see in the theorem below. By definition, the functions $h(\cdot;\sigma_f,f): R^n \to R$ $(f(\cdot) \in \mathcal{F})$ are convex on the convex set $G_0 \times R$. This ensures that C_0 is a convex subset of R^n. It should be also noted that the function $h(\cdot;\sigma_f,f): R^\ell \to R$ is C^k (k times continuously differentiable) whenever $f(\cdot): R^\ell \to R$ is so.

Theorem 2.1. *Define the functions $q(\cdot): R^n \to R$, $h(\cdot;\sigma_f,f): R^n \to R$ $(f(\cdot) \in \mathcal{F})$ and the convex subset C_0 as in (5). Then $\boldsymbol{u}^* \in R^\ell$ is a maximum solution of (1) if and only if $\boldsymbol{x}^* = \begin{pmatrix} \boldsymbol{u}^* \\ t^* \end{pmatrix} \in R^n$ is a maximum solution of (2) for some $t^* \in R$.*

Proof: Suppose that $\boldsymbol{x}^* = \begin{pmatrix} \boldsymbol{u}^* \\ t^* \end{pmatrix} \in R^n$ is a maximum solution of the problem (2). Then

$$\boldsymbol{u}^* \ \in\ G_0 \ \text{ and } f(\boldsymbol{u}^*) = h(\boldsymbol{x}^*;\sigma_f,f) + \sigma_f q(\boldsymbol{x}^*) \leq 0\ (\forall f(\cdot) \in \mathcal{F}).$$

It follows that $\boldsymbol{u}^*$ is a feasible solution of (1) which attains the same objective value $\boldsymbol{d}^T\boldsymbol{u}^* = \boldsymbol{c}^T\boldsymbol{x}^*$ as the problem (2).

Now assume that $\boldsymbol{u}^* \in R^\ell$ is a maximum solution of the problem (1). Let

$$t^* \ =\ \|\boldsymbol{u}^*\|^2 \ \text{ and } \boldsymbol{x}^* = \begin{pmatrix} \boldsymbol{u}^* \\ t^* \end{pmatrix} \in R^n.$$

Then we see that

$$\begin{array}{rcl} \boldsymbol{u}^* & \in & G_0,\ q(\boldsymbol{x}^*) = t^* - \|\boldsymbol{u}^*\|^2 = 0,\ \boldsymbol{c}^T\boldsymbol{x}^* = \boldsymbol{d}^T\boldsymbol{u}^*, \\ h(\boldsymbol{x}^*;\sigma_f,f) & = & g(\boldsymbol{u}^*;\sigma_f,f) - \sigma_f t^* = f(\boldsymbol{u}^*) \leq 0\ (\forall f(\cdot) \in \mathcal{F}). \end{array}$$

Thus $\boldsymbol{x}^*$ is a feasible solution of (2) that attains the same objective value as the problem (1). ∎

3 Some Characterizations of Peri-Convex Functions.

We first give a characterization of a peri-convex function $f(\cdot) : R^\ell \to R$ in terms of the proximal subdifferentiabilty [1]. We say that a function $f(\cdot) : R^\ell \to R$ is proximal subdifferentiable at $\boldsymbol{u} \in R^\ell$ if there exist an open neighborhood U of $\boldsymbol{u}$, a nonnegative number σ, and a $\boldsymbol{\zeta} \in R^\ell$ such that

$$f(\boldsymbol{v}) - f(\boldsymbol{u}) \geq \boldsymbol{\zeta}^T(\boldsymbol{v} - \boldsymbol{u}) - \sigma\|\boldsymbol{v} - \boldsymbol{u}\|^2 \ (\forall \boldsymbol{v} \in U). \tag{6}$$

We call $\boldsymbol{\zeta}$ a proximal subgradient of $f(\cdot)$ at $\boldsymbol{u}$. See [1]. Note that in the definition of the proximal subdifferentiability of $f(\cdot)$ at $\boldsymbol{u} \in R^\ell$ above, not only the proximal subgradient $\boldsymbol{\zeta}$ but also the nonnegative number σ can depend on the point $\boldsymbol{u} \in R^\ell$ under consideration. It follows from definition that if $f(\cdot) : R^\ell \to R$ is proximal subdifferentiable at $\boldsymbol{u} \in R^\ell$ then it is lower semi-continuous at $\boldsymbol{u} \in R^\ell$ but not necessarily continuous at $\boldsymbol{u} \in R^\ell$; for example the function $f(\cdot) : R \to R$ defined by

$$f(u) = \begin{cases} u & \text{if } u \leq 0, \\ u^2 + 1 & \text{otherwise} \end{cases} \tag{7}$$

is proximal differentiable at every $u \in R$ but not continuous at $u = 0$. Let $G \subset R^\ell$. We say that $f(\cdot)$ is uniformly proximal subdifferentiable on G if we can take a common nonnegative number σ independent of points $\boldsymbol{u} \in G$, and that $f(\cdot)$ is uniformly-and-globally proximal subdifferentiable on G if in addition the inequality (6) holds for every $\boldsymbol{v} \in G$. (In both cases, $\boldsymbol{\zeta}$ can be dependent on the point $\boldsymbol{u} \in G$). The function $f(\cdot) : R \to R$ defined in (7) is uniformly proximal subdifferentiable on R but neither peri-convex nor uniformly-and-globally proximal subdifferentiable on any interval containing 0.

Proposition 3.1. *Let G be a convex subset of R^ℓ and $f(\cdot) : R^\ell \to R$.*

(i) If $f(\cdot)$ is continuous and uniformly proximal subdifferentiable on G, then it is peri-convex on G.

(ii) If $f(\cdot)$ is peri-convex on int(G), the interior of G, then it is uniformly-and-globally proximal subdifferentiable on int(G).

Proof: (i) By assumption, there exists a nonnegative number σ such that given any $\boldsymbol{u} \in G$, the inequality

$$f(\boldsymbol{v}) - f(\boldsymbol{u}) \geq \boldsymbol{\zeta}_u^T(\boldsymbol{v} - \boldsymbol{u}) - \sigma\|\boldsymbol{v} - \boldsymbol{u}\|^2 \ (\forall \boldsymbol{v} \in U_u). \tag{8}$$

holds for some $\boldsymbol{\zeta}_u \in R^\ell$ and some open neighborhood U_u of $\boldsymbol{u}$. We will show that $g(\cdot) \equiv f(\cdot) + \sigma\| \cdot \|^2$ is convex on G. Assume on the contrary that there exist $\bar{\boldsymbol{u}}, \ \bar{\boldsymbol{v}} \in G$ and $\bar{\lambda} \in (0, \ 1)$ for which

$$g((1 - \bar{\lambda})\bar{\boldsymbol{u}} + \bar{\lambda}\bar{\boldsymbol{v}}) > (1 - \bar{\lambda})g(\bar{\boldsymbol{u}}) + \bar{\lambda}g(\bar{\boldsymbol{v}}) \tag{9}$$

holds. We see by (8) that for every $\boldsymbol{u} \in G$ and $\boldsymbol{v} \in U_u$,

$$\begin{aligned} g(\boldsymbol{v}) - g(\boldsymbol{u}) &= \left(f(\boldsymbol{v}) + \sigma\|\boldsymbol{v}\|^2\right) - \left(f(\boldsymbol{u}) + \sigma\|\boldsymbol{u}\|^2\right) \\ &= f(\boldsymbol{v}) - f(\boldsymbol{u}) + \sigma\|\boldsymbol{v}\|^2 - \sigma\|\boldsymbol{u}\|^2 \\ &\geq \boldsymbol{\zeta}_u^T(\boldsymbol{v} - \boldsymbol{u}) - \sigma\|\boldsymbol{v} - \boldsymbol{u}\|^2 + \sigma\|\boldsymbol{v}\|^2 - \sigma\|\boldsymbol{u}\|^2 \\ &= (\boldsymbol{\zeta}_u + 2\sigma\boldsymbol{u})^T(\boldsymbol{v} - \boldsymbol{u}) \end{aligned} \tag{10}$$

Define

$$h(\lambda) \equiv g((1-\lambda)\bar{\boldsymbol{u}} + \lambda\bar{\boldsymbol{v}}) - ((1-\lambda)g(\bar{\boldsymbol{u}}) + \lambda g(\bar{\boldsymbol{v}})) \ \ (\forall \lambda \in [0,1]).$$

Then $h : [0,\ 1] \to R$ turns out to be a continuous function satisfying

$$h(0) = h(1) = 0 \ \text{ and } h(\bar{\lambda}) > 0.$$

Here the last inequality follows from (9). Let

$$\begin{aligned} h* &\equiv \max\{h(\lambda) : \lambda \in [0,\ 1]\},\ \Lambda^* \equiv \{\lambda \in [0,\ 1] : h(\lambda) = h^*\}, \\ \lambda^* &\equiv \min\{\lambda : \lambda \in \Lambda^*\}, \ \text{ and } \ \boldsymbol{u}^* \equiv (1-\lambda^*)\bar{\boldsymbol{u}} + \lambda^*\bar{\boldsymbol{v}}. \end{aligned}$$

Then we know that

$$\begin{aligned} &0 < \lambda^* < 1, \\ &0 > h(\lambda) - h(\lambda^*) \ (\forall \lambda \in [0,\ \lambda^*), \\ &0 \geq h(\lambda) - h(\lambda^*) \ (\forall \lambda \in [\lambda^*,\ 1]), \\ &(1-\lambda)\bar{\boldsymbol{u}} + \lambda\bar{\boldsymbol{v}} \in U_{u^*} \ \text{ if } \lambda \in [0,\ 1] \text{ is sufficiently close to } \lambda^*, \end{aligned} \tag{11}$$

and that for every $\lambda \in [0,\ 1]$ sufficiently close to λ^*,

$$\begin{aligned} 0 &\geq h(\lambda) - h(\lambda^*) \\ &= \quad g((1-\lambda)\bar{\boldsymbol{u}} + \lambda\bar{\boldsymbol{v}}) - g((1-\lambda^*)\bar{\boldsymbol{u}} + \lambda^*\bar{\boldsymbol{v}}) \\ &\quad -(\lambda - \lambda^*)\left(g(\bar{\boldsymbol{v}}) - g(\bar{\boldsymbol{u}})\right) \\ &\geq \quad (\lambda - \lambda^*)(\boldsymbol{\zeta}_{u^*} + 2\sigma\boldsymbol{u}^*)^T(\bar{\boldsymbol{v}} - \bar{\boldsymbol{u}}) \\ &\quad -(\lambda - \lambda^*)\left(g(\bar{\boldsymbol{v}}) - g(\bar{\boldsymbol{u}})\right) \ \ (\text{by } (10)) \\ &= \quad (\lambda - \lambda^*)\left((\boldsymbol{\zeta}_{u^*} + 2\sigma\boldsymbol{u}^*)^T(\bar{\boldsymbol{v}} - \bar{\boldsymbol{u}}) - (g(\bar{\boldsymbol{v}}) - g(\bar{\boldsymbol{u}}))\right). \end{aligned}$$

Therefore we obtain that $0 = h(\lambda) - h(\lambda^*)$ for every $\lambda \in [0,\ 1]$ sufficiently close to λ^*.This contradicts to (11).

(ii) By assumption, there is a nonnegative number σ such that the function $f(\cdot) + \sigma\|\cdot\|^2$ is convex on $\mathrm{int}(G)$. Hence the function $f(\cdot) + \sigma\|\cdot\|^2$ is subdifferentiable on $\mathrm{int}(G)$; for every $\boldsymbol{u}$ in $\mathrm{int}(G)$, there is a $\boldsymbol{\xi} \in R^\ell$ such that

$$f(\boldsymbol{v}) + \sigma\|\boldsymbol{v}\|^2 - (f(\boldsymbol{u}) + \sigma\|\boldsymbol{u}\|^2) \geq \boldsymbol{\xi}^T(\boldsymbol{v} - \boldsymbol{u}) \ (\forall \boldsymbol{v} \in \mathrm{int}(G)).$$

See, for example, Theorem 23.4 of [7]. It follows that

$$f(\boldsymbol{v}) - f(\boldsymbol{u}) \geq (\boldsymbol{\xi} - 2\sigma\boldsymbol{u})^T(\boldsymbol{v} - \boldsymbol{u}) - \sigma\|\boldsymbol{v} - \boldsymbol{u}\|^2 \ (\forall \boldsymbol{v} \in \mathrm{int}(G)).$$

This implies that $f(\cdot)$ is uniformly proximal subdifferentiable on $\mathrm{int}(G)$. ∎

Corollary 3.2. *Let G be an open convex subset of R^ℓ and $f(\cdot) : R^\ell \to R$ be continuous. Then the following (I), (II) and (III) are equivalent.*

(I) $f(\cdot)$ is uniformly proximal subdifferentiable on G.

(II) $f(\cdot)$ is uniformly-and-globally proximal subdifferentiable on G.

(III) $f(\cdot)$ is peri-convex on G.

Proof: In view of (i) and (ii) of Proposition 3.1, we know that (I) $\Longrightarrow$ (III) and (III) $\Longrightarrow$ (II), respectively. Also the implication (II) $\Longrightarrow$ (I) is obvious from definition. Thus the equivalence of (I), (II) and (III) follows. ∎

When a peri-convex function $f(\cdot) : R^\ell \to R$ is C^1 (continuously differentiable) or C^2 (twice continuously differentiable), we can characterize it in terms of its gradient vector $\nabla f(\cdot)$ or its Hessian matrix $\nabla^2 f(\cdot)$, respectively.

Proposition 3.3. *Let G be a convex subset of R^ℓ. If $f(\cdot) : R^n \to R$ is C^1 on an open neighborhood of G, then the following (a), (b) and (c) are equivalent.*

(a) $f(\cdot) + \sigma\|\cdot\|^2$ is convex on G.

(b) $f(\boldsymbol{v}) - f(\boldsymbol{u}) \geq \nabla f(\boldsymbol{u})^T(\boldsymbol{v} - \boldsymbol{u}) - \sigma\|\boldsymbol{v} - \boldsymbol{u}\|^2$ $(\forall \boldsymbol{u},\ \boldsymbol{v} \in G)$.

(c) $(\nabla f(\boldsymbol{v}) - \nabla f(\boldsymbol{u}))^T(\boldsymbol{v} - \boldsymbol{u}) \geq -2\sigma\|\boldsymbol{v} - \boldsymbol{u}\|^2$ $(\forall \boldsymbol{u},\ \boldsymbol{v} \in G)$.

If $f(\cdot) : R^\ell \to R$ is C^2 on an open neighborhood of G, then (a), (b), (c), and (d) below are equivalent.

(d) For every $\boldsymbol{u} \in G$, $\nabla^2 f(\boldsymbol{u}) + 2\sigma\boldsymbol{I}$ is positive semidefinite on the linear subspace L spanned by $G - \{\boldsymbol{u}^0\}$, where $\boldsymbol{u}^0 \in G$, i.e., *$\boldsymbol{v}^T(\nabla^2 f(\boldsymbol{u}) + 2\sigma\boldsymbol{I})\boldsymbol{v} \geq 0$ for every $\boldsymbol{v}$ in the linear subspace L.*

Proof: It is well-known that if $g(\cdot) : R^\ell \to R$ is C^1 on an open neighborhood of G then the following (a)' (b)' and (c)' are equivalent.

(a)' $g(\cdot)$ is convex on G.

(b)' $g(\boldsymbol{v}) - g(\boldsymbol{u}) \geq \nabla g(\boldsymbol{u})^T(\boldsymbol{v} - \boldsymbol{u})$ $(\forall \boldsymbol{u},\ \boldsymbol{v} \in G)$.

(c)' $(\nabla g(\boldsymbol{v}) - \nabla g(\boldsymbol{u}))^T(\boldsymbol{v} - \boldsymbol{u}) \geq 0$ $(\forall \boldsymbol{u},\ \boldsymbol{v} \in G)$.

If $f(\cdot) : R^n \to R$ is C^2 on an open neighborhood of G, then (a)', (b)', (c)', and (d)' below are equivalent.

(d)' For every $\boldsymbol{u} \in G$, $\nabla^2 g(\boldsymbol{u})$ is positive semidefinite on the linear space spanned by $G - \{\boldsymbol{u}^0\}$, where $\boldsymbol{u}^0 \in G$.

See, for example, Chapter 6 of [6]. When $g(\cdot) = f(\cdot) + \|\cdot\|^2 : R^\ell \to R$, we can rewrite the conditions (a)', (b)', (c)' and (d)' as (a), (b), (c) and (d), respectively. Thus the desired results follow. ∎

Suppose that a function $f : R^\ell \to R$ is C^1. If its gradient vector is Lipschitz continuous on a convex subset G of R^ℓ, *i.e.*,

$$\alpha(f,G) \equiv \sup\left\{ \frac{\|\nabla f(\boldsymbol{u}) - \nabla f(\boldsymbol{v})\|}{\|\boldsymbol{u}-\boldsymbol{v}\|} : \boldsymbol{u} \in G,\ \boldsymbol{v} \in G,\ \boldsymbol{u} \neq \boldsymbol{v} \right\} < \infty,$$

then we may regard such a $f(\cdot)$ as a "moderately nonlinear" function on G. The corollary below ensures that if $f(\cdot) : R^\ell \to R$ is moderately nonlinear on G then it belongs to $\mathcal{F}_{\mathrm{p.conv}}(G)$.

Corollary 3.4. *Let G be a convex subset of R^ℓ. Assume that $f(\cdot) : R^\ell \to R$ is C^1 on an open neighborhood of G and that $\dfrac{\alpha(f,G)}{2} \leq \sigma < +\infty$. Then the function $f(\cdot) + \sigma\|\cdot\|^2$ is convex on G, and $f(\cdot) - \sigma\|\cdot\|^2$ is concave on G .*

Proof: To prove the convexity of $f(\cdot) + \sigma\|\cdot\|^2$ on G, we will derive the relation in (c) of Proposition 3.3. Let $\boldsymbol{u},\ \boldsymbol{v} \in G$. If $\boldsymbol{u} = \boldsymbol{v}$ then the relation obviously holds. Suppose that $\boldsymbol{u} \neq \boldsymbol{v}$. Then

$$\begin{aligned}
(\nabla f(\boldsymbol{v}) - \nabla f(\boldsymbol{u}))^T (\boldsymbol{v}-\boldsymbol{u}) &\geq -\|\nabla f(\boldsymbol{v}) - \nabla f(\boldsymbol{u})\|\ \|\boldsymbol{v}-\boldsymbol{u}\| \\
&= -\frac{\|\nabla f(\boldsymbol{v}) - \nabla f(\boldsymbol{u})\|}{\|\boldsymbol{v}-\boldsymbol{u}\|}\|\boldsymbol{v}-\boldsymbol{u}\|^2 \\
&\geq -\alpha(f,G)\|\boldsymbol{v}-\boldsymbol{u}\|^2 \\
&\geq -2\sigma\|\boldsymbol{v}-\boldsymbol{u}\|^2
\end{aligned}$$

Thus the relation in (c) of Proposition 3.3 holds for any $\boldsymbol{u},\ \boldsymbol{v} \in G$.

Since $-f(\cdot) : R^\ell \to R$ satisfies the same assumption in the corollary, we know that $-f(\cdot) + \sigma\|\cdot\|^2$ is convex on G, which implies that $f(\cdot) - \sigma\|\cdot\|^2$ is concave on G ∎

It is known that any twice continuously differentiable function $f(\cdot) : R^\ell \to R$ has a *d.c. decomposition*: $f(\boldsymbol{u}) = g(\boldsymbol{u}) - h(\boldsymbol{u})$ ($\forall \boldsymbol{u} \in R^\ell$). Here $g(\cdot) : R^\ell \to R$ and $h(\cdot) : R^\ell \to R$ are convex functions. See, for example, Corollary I.1 of [3]. The assertion of Corollary 3.4 may be regarded as a variation of this result.

4 Concluding Remarks.

(A) When we utilize the conversion from $f(\cdot) \in \mathcal{F}$ to $g(\cdot;\sigma_f, f)$ in (4) in practice, we need to know $\sigma_f \geq \sigma(f,G)$ ($f(\cdot) \in \mathcal{F}$) in advance. In addition, the conversion may not be practical unless σ_f ($f(\cdot) \in \mathcal{F}$) are bounded above from a not too large positive number.

(B) Suppose that the constraint set G_0 is not only convex but also compact and that every $f(\cdot) \in \mathcal{F}$ is continuous on G_0. We then see that $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{u} \\ t \end{pmatrix}$ is a feasible solution of the problem (2) if and only if $\boldsymbol{u} \in G_0$ and $g(\boldsymbol{u}; \sigma_f, f) \leq \sigma_f t \leq \sigma_f \|\boldsymbol{u}\|^2$ $(\forall f(\cdot) \in \mathcal{F})$ hold. Therefore, taking a $\bar{t} \geq \max\{\|\boldsymbol{u}\|^2 : \boldsymbol{u} \in G_0\}$, we can add the inequality $t \leq \bar{t}$ to the representation of C_0 in (5) such that

$$C_0 = \{\boldsymbol{x} \in R^n \ : \ \boldsymbol{u} \in G_0, \ h(\boldsymbol{x}; \sigma_f, f) \ (f(\cdot) \in \mathcal{F}), \ t \leq \bar{t}\}.$$

Whenever at least one σ_f is positive, this modification makes the constraint set C_0 of the problem (2) compact. When all σ_f's are zero, then the nonlinear program (1) itself is a convex program, and we can eliminate the variable t and the inequality constraint $t - \|\boldsymbol{u}\|^2$ in (5). In this case, the nonlinear program (1) and the induced quadratic programs (2) coincide with each other. Specifically

$$C_0 = \{\boldsymbol{u} \in G_0 : f(\boldsymbol{u}) \leq 0 \ (\forall f(\cdot) \in \mathcal{F})\}$$

is a compact convex subset of R^ℓ.

References

[1] F. H. Clarke, Yu. S. Ledysaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, (Springer-Verlag, Berlin, 1998).

[2] R. Horst, P. M. Pardalos, and N. V. Thoai, *Introduction to Global Optimization*, (Kluwer, Dordrecht, 1995).

[3] R. Horst and H. Tuy, *Global Optimization, Second, Revised Edition*, (Springer-Verlag, Berlin, 1992).

[4] M. Kojima and L. Tunçel, "Cones of matrices and successive convex relaxations of nonconvex sets," Technical Report B-338, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, March 1998. Also issued as CORR 98-6, Dept. of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

[5] M. Kojima and L. Tunçel, "Discretization and localization in successive convex relaxation methods for nonconvex quadratic optimization problems," Technical Report B-341, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, July 1998. Also issued as CORR 98-34, Dept. of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

[6] O.L. Mangasaian, *Nonlinear Programming*, (McGraw-Hill, New York, 1969).

[7] R. T. Rockafellar, *Convex Analysis*, (Princeton University Press, Princeton, 1970).

[8] A. Takeda, Y. Dai, M. Fukuda, and M. Kojima, "Towards the Implementation of Successive Convex Relaxation Method for Nonconvex Quadratic Optimization Problems," Technical Report B-347, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, March 1999.

[9] H. Tuy, *Convex Analysis and Global Optimization*, (Kluwer, Dordrecht, 1998).