*Journal of Global Optimization*, **29:** 377–399, 2004.



# Canonical Duality Theory and Solutions to Constrained Nonconvex Quadratic Programming

*Dedicated to Professor Ivar Ekeland on the occasion of his 60th birthday*

DAVID YANG GAO
*Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA* (*e-mail: gao@math.vt.edu*)

(Received and accepted 13 August 2003)

**Abstract.** This paper presents a perfect duality theory and a complete set of solutions to nonconvex quadratic programming problems subjected to inequality constraints. By use of the *canonical dual transformation* developed recently, a canonical dual problem is formulated, which is perfectly dual to the primal problem in the sense that they have the same set of KKT points. It is proved that the KKT points depend on the index of the Hessian matrix of the total cost function. The global and local extrema of the nonconvex quadratic function can be identified by the triality theory [11]. Results show that if the global extrema of the nonconvex quadratic function are located on the boundary of the primal feasible space, the dual solutions should be interior points of the dual feasible set, which can be solved by deterministic methods. Certain nonconvex quadratic programming problems in $\mathbb{R}^n$ can be converted into a dual problem with only one variable. It turns out that a complete set of solutions for quadratic programming over a sphere is obtained as a by-product. Several examples are illustrated.

**Mathematics Subject Classifications.** 90C, 49N.

**Key words.** canonical dual transformation, duality theory, global optimization, NP-hard problems, quadratic programming.

## 1. Primal Problems and Motivation

The primary goal of this paper is to study the complete set of solutions to the following standard quadratic programming problem (primal problem $(\mathcal{P})$ in short).

$$(\mathcal{P}): \quad \min_{\mathbf{x}\in\mathbb{R}^n} \; P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{f}^T\mathbf{x}, \tag{1}$$

$$\text{s.t.} \quad B\mathbf{x} \leqslant \mathbf{b}, \tag{2}$$

where $A = A^T \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times n}$ are given two matrices, $\mathbf{f} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$ are two vectors.

The quadratic programming problem $(\mathcal{P})$ appears in many applications. In the case that the matrix $A$ is not symmetric, it can be converted to symmetric form by replacing $A$ by $\frac{1}{2}(A+A^T)$. Also, if the problem has any other additional linear constraints, such as $\mathbf{x}\geqslant 0\in\mathbb{R}^n$, we can always replace $B$ and $\mathbf{b}$ by certain extended matrix $\bar{B}$ and vector $\bar{\mathbf{b}}$, respectively, such that both constraints $B\mathbf{x}\leqslant\mathbf{b}$ and $\mathbf{x}\geqslant 0$ can be written in a unified form $\bar{B}\mathbf{x}\leqslant\bar{\mathbf{b}}$. Thus the primal problem $(\mathcal{P})$ can be considered as the general quadratic programming problem. The primal feasible space

$$\mathcal{X}_f=\{\mathbf{x}\in\mathbb{R}^n\,|\,B\mathbf{x}\leqslant\mathbf{b}\} \tag{3}$$

is a convex subset of $\mathbb{R}^n$. The problem $(\mathcal{P})$ has at least one solution if the radius $r_0$ of $\mathcal{X}_f$, defined by $|\mathbf{x}|\leqslant r_0\ \forall\mathbf{x}\in\mathcal{X}_f$, is finite.

Introducing the Lagrange multiplier $\nu\in\mathbb{R}^m$ to relax the inequality constraint $B\mathbf{x}\leqslant\mathbf{b}$, the classical Lagrange function for $(\mathcal{P})$ is given by

$$L(\mathbf{x},\nu)=\frac{1}{2}\mathbf{x}^T A\mathbf{x}-\mathbf{f}^T\mathbf{x}+\nu^T(B\mathbf{x}-\mathbf{b}). \tag{4}$$

Thus the first order Karush-Kuhn-Tucker (KKT) optimality conditions for $(\mathcal{P})$ can be written as follows

$$A\mathbf{x}+B^T\nu=\mathbf{f}, \tag{5}$$

$$B\mathbf{x}-\mathbf{b}\leqslant 0,\quad \nu\geqslant 0, \tag{6}$$

$$\nu^T(B\mathbf{x}-\mathbf{b})=0. \tag{7}$$

Equation (7) is also referred as the complementarity condition, which is usually written in the form of $\nu^T\perp(B\mathbf{x}-\mathbf{b})$, i.e. the Lagrange multiplier $\nu\in\mathbb{R}^m$ should be perpendicular to the constraint vector $(B\mathbf{x}-\mathbf{b})\in\mathbb{R}^m$. Any point $\bar{\mathbf{x}}$ which satisfies (5)–(7) is called a KKT stationary point of $(\mathcal{P})$. It is known that the KKT conditions are only necessary for the quadratic programming problem $(\mathcal{P})$, i.e. if $\bar{\mathbf{x}}$ is an optimal solution of $(\mathcal{P})$, then $\bar{\mathbf{x}}$ must be a KKT point. If the matrix $A$ is positive semi-definite, or positive definite, then $(\mathcal{P})$ is a convex programming problem. In this case, a KKT point $\bar{\mathbf{x}}$ is also sufficient for problem $(\mathcal{P})$, which can be solved easily by any of polynomial algorithms. However, when $A$ is not positive semi-definite, the cost function $P(\mathbf{x})$ is nonconvex, and it might possess many local minimizers. In this case, $(\mathcal{P})$ becomes a nonconvex problem, and the application of traditional local optimization procedures for this problem can not guarantee the identification of the global minima (see Floudas and Visweswaran, 1995).

Nonconvex quadratic programming problem has great importance both from the mathematical and application viewpoints. Sahni (1974) first showed that for a negative definite matrix $A$, the problem $(\mathcal{P})$ is NP-hard. This result was also proved by Vavasis (1990, 1991) and by Pardalos (1991). During the last decade, several authors have shown that the general quadratic programming problem $(\mathcal{P})$

is an *NP-hard problem* in global optimization (cf. Murty and Kabadi, 1987; Horst et al., 2000). It was shown by Pardalos and Vavasis (1990) that even when the matrix $A$ is of rank one with exactly one negative eigenvalue, the problem is NP-hard. In order to solve this difficult problem, many efforts have been made during the last decade. A comprehensive survey has been given by Floudas and Visweswaran (1995).

Duality is a fundamental concept that plays a central role in almost all natural science. In convex systems, the mathematical theory of duality has been well studied. In the primal problem $(\mathcal{P})$, if $A$ is positive definite, the classical Lagrangian $L(\mathbf{x}, \nu)$ defined by (4) is a saddle function, and the dual function can be defined as

$$P^*(\nu) = \min_{\mathbf{x}\in\mathbb{R}^n} L(\mathbf{x}, \nu).$$

By the well-known saddle min-max theory, the following duality relation

$$\min P(\mathbf{x}) = \min_{\mathbf{x}} \max_{\nu} L(\mathbf{x}, \nu) = \max_{\nu} \min_{\mathbf{x}} L(\mathbf{x}, \nu) = \max_{\nu} P^*(\nu)$$

holds on certain feasible spaces of $\mathbf{x}$ and $\nu$. Based on this classical saddle-Lagrangian duality, the so-called *primal-dual interior-point method* has been considered as a revolutionary technic in convex programming during the last fifteen years (cf. Wright, 1998). However, if the matrix $A$ is indefinite, the classical Lagrangian $L(\mathbf{x}, \nu)$ is no longer a saddle function. Although in this case, the Fenchel-Rockefeller dual function can still be defined as

$$P^*(\nu) = \inf_{\mathbf{x}\in\mathbb{R}^n} L(\mathbf{x}, \nu) = -P^\sharp(-B^T\nu) - \mathbf{b}^T\nu,$$

where

$$P^\sharp(\mathbf{x}^*) = \sup_{\mathbf{x}\in\mathbb{R}^n} \{\mathbf{x}^T\mathbf{x}^* - P(\mathbf{x})\}$$

is the Fenchel sup-conjugate transformation, the Fenchel-Young inequality leads to a broken duality relation:

$$\inf P(\mathbf{x}) \geqslant \sup P^*(\nu).$$

The no zero $\theta = \inf P(\mathbf{x}) - \sup P^*(\nu) > 0$ is called *duality gap*. Very often $\theta = \infty$ in concave minimization where $A$ is negative definite. This duality gap shows that the traditional Lagrange duality theory can be used mainly for convex problems. In dynamical systems, it is known that the so-called *chaotic phenomena* is mainly due to the nonconvexity of the total potential of the system (see [12]).

In order to recover the duality gap, many efforts have been made during the last decade in global optimization, and some important results have been achieved (see, for example, Penot and Volle, 1990; Thach et al., 1993–96; Tuy, 1991, 1995; Rubinov et al., 2001; Gasimov, 2002; Goh and Yang; 2002, Rubinov and Gasimov, 2003, and much more). These results are based on the augmented Lagrangian theory and penalty function methods, a so-called nonlinear Lagrange theory has been developed very recently for solving nonconvex constrained optimization problems, where the zero duality gap property is equivalent to the lower semi-continuity of a perturbation function (see Rubinov and Yang, 2003). However, how to use the traditional Legendre transformation to formulate perfect dual action has been listed as one of two open problems in the very recent paper by Ivar Ekeland (2003).

Actually, this open problem has been solved very recently by the *canonical dual transformation method* developed by the author in general nonconvex systems (Gao, 2003). The key idea of this method is to choose a certain (geometrically reasonable) operator $\mathbf{y}=\Lambda(\mathbf{x})$: $\mathbb{R}^n \to \mathbb{R}^m$ such that a given nonconvex function $P(\mathbf{x})$ can be written in the canonical form $P(\mathbf{x})=\Phi(\mathbf{x},\Lambda(\mathbf{x}))$, where $\Phi(\mathbf{x},\mathbf{y})$: $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is a canonical function in each of its variables (see [13]). By the definition introduced in [11], a real-valued function $\overline{W}(\mathbf{y})$: $\mathbb{R}^m \to \mathbb{R}$ is said to be a canonical function on $\mathbb{R}^m$ if the duality relation $\mathbf{y}^*=D\overline{W}(\mathbf{y})$ is invertible for all $\mathbf{y}\in\mathbb{R}^m$. Thus, the Legendre conjugate of a canonical function $W(\mathbf{y})$ can be uniquely defined by

$$\overline{W}^*(\mathbf{y}^*)=\{\mathbf{y}^T\mathbf{y}^*-\overline{W}(\mathbf{y})\,|\,\mathbf{y}^*=D\overline{W}(\mathbf{y})\}.$$

In the case that the canonical function $\Phi$ can be written in the form of $\Phi(\mathbf{x},\mathbf{y})=\overline{W}(\mathbf{y})-\overline{F}(\mathbf{x})$, where both $\overline{W}$: $\mathbb{R}^m\to\mathbb{R}$ and $\overline{F}$: $\mathbb{R}^n\to\mathbb{R}$ are canonical functions, the extended Lagrangian, i.e. the so-called *total complementary energy* in nonconvex mechanics (See Gao and Strang, 1989a, b)

$$\Xi(\mathbf{x},\mathbf{y}^*)=(\Lambda(\mathbf{x}))^T\mathbf{y}^*-\overline{W}^*(\mathbf{y}^*)-\overline{F}(\mathbf{x}) \tag{8}$$

is well defined on $\mathbb{R}^n\times\mathbb{R}^m$. Then by use of the so-called $\Lambda$-*canonical dual transformation* (see [11])

$$\overline{F}^\Lambda(\mathbf{y}^*)=\{(\Lambda(\mathbf{x}))^T\mathbf{y}^*-\overline{F}(\mathbf{x})\,|\,\Lambda_t^T(\mathbf{x})\mathbf{y}^*-D\overline{F}(\mathbf{x})=0,\ \mathbf{x}\in\mathbb{R}^n\}, \tag{9}$$

where $\Lambda_t(\mathbf{x})=D\Lambda(\mathbf{x})$ is the Gâteaux derivative of $\Lambda(\mathbf{x})$, the canonical dual function of the nonconvex $P(\mathbf{x})$ can be well defined by

$$P^d(\mathbf{y}^*)=\overline{F}^\Lambda(\mathbf{y}^*)-\overline{W}^*(\mathbf{y}^*). \tag{10}$$

It was proved in [11] that if $(\bar{\mathbf{x}},\bar{\mathbf{y}}^*)$ is a critical point of $\Xi$, then the following duality condition holds

$$P(\bar{\mathbf{x}})=P^d(\bar{\mathbf{y}}^*).$$

The canonical dual transformation was originally studied by Gao and Strang in nonsmooth/nonconvex mechanics (1989), where $\Lambda(\mathbf{y})$ is a quadratic partial differential operator, $\overline{W}(\mathbf{y})$ is convex and $\overline{F}(\mathbf{x})$ is a linear functional. The one-to-one duality relation $\mathbf{y}^* = D\overline{W}(\mathbf{y})$ is called the canonical constitutive law. For example, in finite deformation theory and two-phase transitions, the total potential of the systems usually takes the form of $P(\mathbf{x}) = W(\mathbf{x}) - \mathbf{f}^T\mathbf{x}$, where $W(\mathbf{x}) = \frac{1}{2}\left(\frac{1}{2}|\mathbf{x}|^2 - \mu\right)^2$ is the well-known van de Waal double well function, $\mu > 0$ is a given parameter. Since the relation $\mathbf{x}^* = DW(\mathbf{x})$ is not one-to-one, the Legendre conjugate $W^*(\mathbf{x}^*)$ of the nonconvex $W$ is not uniquely defined (cf. Sewell, 1987). However, in terms of the quadratic operator $\mathbf{y} = \Lambda(\mathbf{x}) = \frac{1}{2}|\mathbf{x}|^2$, the function $\overline{W}(\mathbf{y}) = \frac{1}{2}(\mathbf{y} - \mu)^2$ is a canonical (quadratic) function of $\mathbf{y}$, and $\overline{W}(\Lambda(\mathbf{x})) = W(\mathbf{x})$. Since $\mathbf{y}^* = D\overline{W}(\mathbf{y}) = \mathbf{y} - \mu$ is one-to-one, the Legendre conjugate of $\overline{W}(\mathbf{y})$ can be uniquely obtained as

$$\overline{W}^*(\mathbf{y}^*) = \{\mathbf{y}^T\mathbf{y}^* - \overline{W}(\mathbf{y}) \,|\, \mathbf{y}^* = D\overline{W}(\mathbf{y}) = \mathbf{y} - \mu\} = \frac{1}{2}\mathbf{y}^{*2} + \mu\mathbf{y}^*. \tag{11}$$

It was shown in [8, 12] that by use of this canonical dual transformation, a class of nonconvex boundary value problems can be converted into a dual algebraic system, therefore, complete set of solutions have been obtained.

Generally speaking, most of physical variables appear in dual pairs. This one-to-one canonical duality relation serves as a foundation for the canonical dual transformation method. Extensive applications of this general method have been given in [11], and an interesting *triality theory* in nonconvex systems was discovered in post-buckling analysis of a nonlinear beam model, where the total potential $P(\mathbf{x})$ is a nonconvex functional in infinite dimensional space (see [7]). Very recently, this canonical dual transformation method and the triality theory have been generated to solve a class of global optimization problems, where $\overline{F}(\mathbf{x}) = -P(\mathbf{x}) = \mathbf{f}^T\mathbf{x} - \frac{1}{2}\mathbf{x}^TA\mathbf{x}$ and $\overline{W}(\Lambda(\mathbf{x}))$ is a canonical function of $\mathbf{y} = \Lambda(\mathbf{x})$ (see [16]).

The goal of this paper is to present particular application of the general results given in [16] to the nonconvex quadratic problem $(\mathcal{P})$. In the next section, a parametric optimization problem is proposed, which can be considered as a trust region method. By the canonical dual transformation, a perfect dual problem is formulated, which is equivalent to the primal problem in the sense that they have same set of KKT points. The global minima theorem is presented in Section 3. In Section 4 we will show that the canonical dual transformation is naturally linked to the quadratic programming over a sphere. The canonical dual problem can be solved completely, and a complete set of KKT points are obtained. Several examples are illustrated in the last section.

## 2. Parametrization and Canonical Dual Problem

In order to use the canonical dual transformation to solve the nonconvex quadratic programming problem $(\mathcal{P})$, an additional normality constraint $|\mathbf{x}|^2 \leqslant 2\mu$ is introduced, where $\mu > 0$ is a given parameter. Actually, this normality condition is indeed a constraint for many real applications (see [8], and Powell, 2002). By use of this constraint, a parametric optimization problem can be proposed as the following

$$(\mathcal{P}_\mu): \quad \min_{\mathbf{x}\in\mathbb{R}^n} \quad P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{f}^T\mathbf{x}, \tag{12}$$

$$\text{s.t.} \quad B\mathbf{x} \leqslant \mathbf{b}, \quad |\mathbf{x}|^2 \leqslant 2\mu. \tag{13}$$

Since for a given $\mu > 0$, the feasible space

$$\mathcal{X}_\mu = \left\{ \mathbf{x} \in \mathbb{R}^n \,|\, B\mathbf{x} \leqslant \mathbf{b}, \ \frac{1}{2}|\mathbf{x}|^2 \leqslant \mu \right\} \tag{14}$$

is a closed convex subset of $\mathbb{R}^n$, the parametric optimization problem $(\mathcal{P}_\mu)$ has at least one global minimizer $\bar{\mathbf{x}}_\mu$. If $\mu \geqslant \mu_0 = \frac{1}{2}r_0^2$, the radius of the feasible space $\mathcal{X}_f$, then $\bar{\mathbf{x}}_\mu$ solves also the original problem $(\mathcal{P})$. In this section, we will find the canonical dual formulation of the parametric problem $(\mathcal{P}_\mu)$.

Following the standard procedure of the canonical dual transformation developed in [11], the canonical geometrical operator $\Lambda$: $\mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}$ in the primal problem $(\mathcal{P}_\mu)$ can be defined as a vector-valued mapping:

$$\mathbf{y} = \Lambda(\mathbf{x}) = \left( B\mathbf{x}, \frac{1}{2}|\mathbf{x}|^2 \right) = (\epsilon, \rho): \ \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R},$$

where $\epsilon = B\mathbf{x}$ is an m-vector, and $\rho = \frac{1}{2}|\mathbf{x}|^2$ is a scale. Let $\mathcal{Y}_a$ be a convex subset of $\mathcal{Y} = \mathbb{R}^m \times \mathbb{R}$ defined by

$$\mathcal{Y}_a = \{ \mathbf{y} = (\epsilon, \rho) \in \mathbb{R}^m \times \mathbb{R} \,|\, \epsilon \leqslant \mathbf{b}, \ \rho \leqslant \mu \}.$$

Its indicator $\overline{W}$: $\mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$, defined by

$$\overline{W}(\mathbf{y}) = \begin{cases} 0 & \text{if } \ \mathbf{y} \in \mathcal{Y}_a, \\ +\infty & \text{otherwise,} \end{cases}$$

is convex, lower semi-continuous on $\mathcal{Y}$. Thus, the inequality constraints in $(\mathcal{P}_\mu)$ can be relaxed by the indicator of $\mathcal{Y}_a$ and the parametric primal problem $(\mathcal{P}_\mu)$ takes the unconstrained canonical form

$$(\mathcal{P}_\mu): \quad \min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \overline{W}(\Lambda(x)) + \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{f} \right\}. \tag{15}$$

By the fact that $\overline{W}(\mathbf{y})$ is convex, lower semi-continuous on $\mathcal{Y}$, the canonical dual variable $\mathbf{y}^* \in \mathcal{Y}^* = \mathcal{Y} = \mathbb{R}^m \times \mathbb{R}$ is defined by the sub-differential inclusion:

$$\mathbf{y}^* \in \partial^- \overline{W}(\mathbf{y}) = \begin{cases} (\epsilon^*, \rho^*) & \text{if } \ \epsilon^* \geqslant 0 \in \mathbb{R}^m, \ \rho^* \geqslant 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

The canonical conjugate $\overline{W}^\sharp$ of $\overline{W}$ can be obtained by the sup-Fenchel transformation:

$$\overline{W}^\sharp(\mathbf{y}^*) = \sup_{\mathbf{y}\in\mathcal{Y}}\{\mathbf{y}^T\mathbf{y}^* - \overline{W}(\mathbf{y})\} = \sup_{\boldsymbol{\epsilon}\leqslant\mathbf{b}}\sup_{\rho\leqslant\mu}\{\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}^* + \rho\rho^*\}$$
$$= \begin{cases} \mathbf{b}^T\boldsymbol{\epsilon}^* + \mu\rho^* & \text{if } \ \boldsymbol{\epsilon}^* \geqslant 0, \ \rho^* \geqslant 0, \\ +\infty & \text{otherwise.} \end{cases} \tag{16}$$

Its effective domain is a positive cone in $\mathbb{R}^m \times \mathbb{R}$, defined by

$$\mathcal{Y}_a^* = \operatorname{dom} \overline{W}^\sharp(\mathbf{y}^*) = \{(\boldsymbol{\epsilon}^*, \rho^*) \in \mathbb{R}^m \times \mathbb{R} \,|\, \boldsymbol{\epsilon}^* \geqslant 0 \in \mathbb{R}^m, \rho^* \geqslant 0 \in \mathbb{R}\}.$$

Since the sup-duality relations

$$\mathbf{y}^* \in \partial^- \overline{W}(\mathbf{y}) \ \Leftrightarrow \ \mathbf{y} \in \partial^- \overline{W}^\sharp(\mathbf{y}^*) \ \Leftrightarrow \ \overline{W}(\mathbf{y}) + \overline{W}^\sharp(\mathbf{y}^*) = \mathbf{y}^T\mathbf{y}^* \tag{17}$$

hold on $\mathcal{Y} \times \mathcal{Y}^*$, the duality pair $(\mathbf{y}, \mathbf{y}^*)$ is called the *extended canonical dual pair* on $\mathcal{Y} \times \mathcal{Y}^*$ (see [11]), and the functions $\overline{W}(\mathbf{y})$ and $\overline{W}^\sharp(\mathbf{y}^*)$ are called *canonical functions* (see [11]). Particularly, on $\mathcal{Y}_a \times \mathcal{Y}_a^*$, the sup-duality relations (17) are equivalent to the following KKT conditions:

$$\mathcal{Y}_a \ni \mathbf{y} \ \perp \ \mathbf{y}^* \in \mathcal{Y}_a^*.$$

For a given $\mathbf{y}^* = (\boldsymbol{\epsilon}^*, \rho^*) \in \mathbb{R}^m \times \mathbb{R}$ such that $A + \rho^* I$ is invertible, then the $\Lambda$-canonical conjugate $\overline{F}^\Lambda(\mathbf{y}^*)$ of the canonical function $\overline{F}(\mathbf{x}) = -P(\mathbf{x})$ can be well defined by the $\Lambda$-canonical dual transformation (cf. [11])

$$\overline{F}^\Lambda(\mathbf{y}^*) = \{\Lambda(\mathbf{x})^T\mathbf{y}^* - \overline{F}(\mathbf{x}) \,|\, D\overline{F}(\mathbf{x}) = \Lambda_t^T(\mathbf{x})\mathbf{y}^*, \ \mathbf{x} \in \mathcal{X}_a\}$$
$$= -\frac{1}{2}(\mathbf{f} - B^T\boldsymbol{\epsilon}^*)^T(A + \rho^* I)^{-1}(\mathbf{f} - B^T\boldsymbol{\epsilon}^*).$$

On the dual feasible space defined by

$$\mathcal{Y}_\mu^* = \{(\boldsymbol{\epsilon}^*, \rho^*) \in \mathcal{Y}_a^* \,|\, \det(A + \rho^* I) \neq 0\}$$
$$= \{(\boldsymbol{\epsilon}^*, \rho^*) \in \mathbb{R}^m \times \mathbb{R} \,|\, \boldsymbol{\epsilon}^* \geqslant 0, \rho^* \geqslant 0, \det(A + \rho^* I) \neq 0\},$$

the canonical dual function $P^d(\mathbf{y}^*) = \overline{F}^\Lambda(\mathbf{y}^*) - \overline{W}^\sharp(\mathbf{y}^*)$ takes the following form

$$P^d(\boldsymbol{\epsilon}^*, \rho^*) = -\frac{1}{2}(\mathbf{f} - B^T\boldsymbol{\epsilon}^*)^T(A + \rho^* I)^{-1}(\mathbf{f} - B^T\boldsymbol{\epsilon}^*) - \mu\rho^* - \mathbf{b}^T\boldsymbol{\epsilon}^*. \tag{18}$$

Thus, the canonical dual problem ($(\mathcal{P}_\mu^d)$ in short) associated with the parametric problem $(\mathcal{P}_\mu)$ can be eventually formulated as the following (See Gao, 2003)

$$(\mathcal{P}_\mu^d): \ \operatorname{ext} P^d(\boldsymbol{\epsilon}^*, \rho^*) \tag{19}$$

$$\text{s.t. } \boldsymbol{\epsilon}^* \geqslant 0, \ \rho^* \geqslant 0, \ \det(A + \rho^* I) \neq 0, \tag{20}$$

where ext $P(\mathbf{x})$ stands for finding all the extremum values of $P(\mathbf{x})$.

THEOREM 1 (Perfect duality theorem). *Problem* $(\mathcal{P}^d_\mu)$ *is canonically* (*perfectly*) *dual to the primal parametric optimization problem* $(\mathcal{P}_\mu)$ *in the sense that if* $\bar{y}^* = (\bar{\epsilon}^*, \bar{\rho}^*) \in \mathcal{Y}^*_\mu$ *is a KKT point of* $(\mathcal{P}^d_\mu)$, *then the vector defined by*

$$\bar{\mathbf{x}} = (A + \bar{\rho}^* I)^{-1}(\mathbf{f} - B^T \bar{\epsilon}^*) \tag{21}$$

*is a KKT point of* $(\mathcal{P}_\mu)$, *and*

$$P(\bar{\mathbf{x}}) = P^d(\bar{\mathbf{y}}^*). \tag{22}$$

*S*. uppose that $\bar{\mathbf{y}}^* = (\bar{\epsilon}^*, \bar{\rho}^*) \in \mathcal{Y}^*_\mu$ is a KKT point of $(\mathcal{P}^d_\mu)$, then we have

$$0 \leqslant \bar{\rho}^* \perp \frac{1}{2}(\mathbf{f} - B^T \bar{\epsilon}^*)^T (A + \bar{\rho}^* I)^{-2} (\mathbf{f} - B^T \bar{\epsilon}^*) - \mu \leqslant 0, \tag{23}$$

$$0 \leqslant \bar{\epsilon}^* \perp B(A + \bar{\rho}^* I)^{-1}(\mathbf{f} - B^T \bar{\epsilon}^*) - \mathbf{b} \leqslant 0. \tag{24}$$

In terms of $\bar{\mathbf{x}} = (A + \bar{\rho}^* I)^{-1}(\mathbf{f} - B^T \bar{\epsilon}^*)$, we have

$$0 \leqslant \bar{\rho}^* \perp \frac{1}{2}\bar{\mathbf{x}}^T \bar{\mathbf{x}} - \mu \leqslant 0, \tag{25}$$

$$0 \leqslant \bar{\epsilon}^* \perp B\bar{\mathbf{x}} - \mathbf{b} \leqslant 0. \tag{26}$$

This shows that $\bar{\mathbf{x}} = (A + \bar{\rho}^* I)^{-1}(\mathbf{f} - B^T \bar{\epsilon}^*)$ is a KKT point of the parametric problem $(\mathcal{P}_\mu)$. By the complementarity conditions (25) and (26), we have $\bar{\rho}^* \mu = \frac{1}{2}\bar{\rho}^* \bar{\mathbf{x}}^T \bar{\mathbf{x}}$ and $\mathbf{b}^T \bar{\epsilon}^* = (B\bar{\mathbf{x}})^T \bar{\epsilon}^*$. Thus, in terms of $\bar{\mathbf{x}} = (A + \bar{\rho}^* I)^{-1}(\mathbf{f} - B^T \bar{\epsilon}^*)$, we have

$$P^d(\bar{\mathbf{y}}^*) = -\frac{1}{2}\bar{\mathbf{x}}^T (A + \bar{\rho}^* I)\bar{\mathbf{x}} - \frac{1}{2}\bar{\rho}^* \bar{\mathbf{x}}^T \bar{\mathbf{x}} - (B\bar{\mathbf{x}})^T \bar{\epsilon}^* = \frac{1}{2}\bar{\mathbf{x}}^T A \bar{\mathbf{x}} - \bar{\mathbf{x}}^T \mathbf{f} = P(\bar{\mathbf{x}}),$$

which shows that there is no duality gap between the problems $(\mathcal{P}_\mu)$ and $(\mathcal{P}^d_\mu)$. This proves the theorem. □

Theorem 1 shows that the primal problem $(\mathcal{P}_\mu)$ is equivalent to the canonical dual problem $(\mathcal{P}^d_\mu)$ in the sense that they have the same set of KKT points. It is well known that the KKT conditions are only necessary for nonconvex quadratic programming. The next section will present extremality conditions for the KKT points.

## 3. Triality Theory for Local and Global Extrema

The quadratic programming problem $(\mathcal{P})$ is nonconvex if $A$ has at least one negative eigenvalue. In order to clarify the extremality condition of the KKT points, we need the following definition.

DEFINITION 1 (Index of the matrix $A$). Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. The index $i_d$ of $A$ is defined to be the total number of distinct negative eigenvalues of $A$.

By this definition, the quadratic function $P(\mathbf{x})$ is nonconvex if and only if the index $i_d > 0$. Suppose that the vector $(\bar{\epsilon}_i^*, \bar{\rho}_i^*)$ is a KKT point of $(\mathcal{P}_\mu^d)$, and the matrix $A$ with index $i_d$ has $p \leqslant n$ distinct eigenvalues $\{a_i\}$, $i = 1, \ldots, p \leqslant n$ in the order of

$$a_1 < a_2 < \cdots < a_{i_d} < 0 \leqslant a_{i_d+1} < \cdots < a_p.$$

Then, if $\bar{\rho}_i^* > -a_1$, the matrix $(A + \bar{\rho}_i^* I)$ is positive definite. However, if $a_{i_d+1} = \cdots = a_p = 0$ and the KKT point $\bar{\rho}_i^* < -a_{i_d}$, the matrix $(A + \bar{\rho}_i^* I)$ will be negative definite. Let

$$\mathcal{Y}_{\mu+}^* = \{(\epsilon^*, \rho^*) \in \mathcal{Y}_\mu^* \mid (A + \rho^* I) \text{ is positive definite}\}, \tag{27}$$

$$\mathcal{Y}_{\mu-}^* = \{(\epsilon^*, \rho^*) \in \mathcal{Y}_\mu^* \mid (A + \rho^* I) \text{ is negative definite}\}, \tag{28}$$

$$\mathcal{Y}_{\mu i}^* = \{(\epsilon^*, \rho^*) \in \mathcal{Y}_{\mu-}^* \mid \rho^* > 0)\}, \tag{29}$$

and

$$\mathcal{X}_{\mu s} = \left\{ \mathbf{x} \in \mathcal{X}_\mu \,\middle|\, \frac{1}{2}|\mathbf{x}|^2 = \mu \right\}. \tag{30}$$

Based on the triality theory developed in [11] as well as the recent result (see Gao, 2003), we have the following interest result.

THEOREM 2 (Local and global extrema). *Suppose that the matrix A has no zero index $i_d > 0$, and for a given parameter $\mu > 0$, the vector $(\bar{\epsilon}_i^*, \bar{\rho}_i^*)$ is a KKT point of the duals problem $(\mathcal{P}_\mu^d)$, and let $\bar{\mathbf{x}}_i = (A + \bar{\rho}_i^* I)^{-1}(\mathbf{f} - B^T \bar{\epsilon}_i^*)$.*

*If $\bar{\rho}_i^* > -a_1$, then the vector $(\bar{\epsilon}_i^*, \bar{\rho}_i^*)$ is a maximizer of $P^d$ on $\mathcal{Y}_{\mu+}^*$ if and only if the vector $\bar{\mathbf{x}}_i$ is a minimizer of $P$ on $\mathcal{X}_{\mu s}$, and*

$$P(\bar{\mathbf{x}}_i) = \min_{\mathbf{x} \in \mathcal{X}_{\mu s}} P(\mathbf{x}) = \max_{(\epsilon^*, \rho^*) \in \mathcal{Y}_{\mu+}^*} P^d(\epsilon^*, \rho^*) = P^d(\bar{\epsilon}_i^*, \bar{\rho}_i^*). \tag{31}$$

*If $0 \leqslant \bar{\rho}_i^* < -a_{i_d}$, then $(\bar{\epsilon}_i^*, \bar{\rho}_i^*)$ is a maximizer of $P^d$ on $\mathcal{Y}_{\mu-}^*$ if and only if $\bar{\mathbf{x}}_i$ is a global maximizer of $P$ on $\mathcal{X}_\mu$, and*

$$P(\bar{\mathbf{x}}_i) = \max_{\mathbf{x} \in \mathcal{X}_\mu} P(\mathbf{x}) = \max_{(\epsilon^*, \rho^*) \in \mathcal{Y}_{\mu-}^*} P^d(\epsilon^*, \rho^*) = P^d(\bar{\epsilon}_i^*, \bar{\rho}_i^*). \tag{32}$$

*If $0 < \bar{\rho}_i^* < -a_{i_d}$, the KKT point $(\bar{\epsilon}_i^*, \bar{\rho}_i^*)$ is a minimizer of $P^d$ on the open set $\mathcal{Y}_{\mu i}^*$ if and only if $\bar{\mathbf{x}}_i$ is a minimizer of $P$ on the set $\mathcal{X}_{\mu s}$, and*

$$P(\bar{\mathbf{x}}_i) = \min_{\mathbf{x} \in \mathcal{X}_{\mu s}} P(\mathbf{x}) = \min_{(\epsilon^*, \rho^*) \in \mathcal{Y}_{\mu i}^*} P^d(\epsilon^*, \rho^*) = P^d(\bar{\epsilon}_i^*, \bar{\rho}_i^*). \tag{33}$$

*Proof.* By Theorem 1, and the triality theory developed in [11] we know that the vector $\bar{\mathbf{y}}_i^* = (\bar{\epsilon}_i^*, \bar{\rho}_i^*) \in \mathcal{Y}_\mu^*$ is a KKT point of the problem $(\mathcal{P}_\mu^d)$ is and only if $\bar{\mathbf{x}}_i = (A + \bar{\rho}_i^* I)^{-1}(\mathbf{f} - B^T \bar{\epsilon}_i^*)$ is a KKT point of the problem $(\mathcal{P}_\mu)$, and

$$P(\bar{\mathbf{x}}_i) = \Xi(\bar{\mathbf{x}}_i, \bar{\epsilon}_i^*, \bar{\rho}_i^*) = P^d(\bar{\epsilon}_i^*, \bar{\rho}_i^*), \tag{34}$$

where the extended Lagrangian $\Xi(\mathbf{x}, \epsilon^*, \rho^*)$ is the *total complementary function* associated with problem $(\mathcal{P}_\mu)$ (see [11]), defined by

$$\Xi(\mathbf{x}, \epsilon^*, \rho^*) = \frac{1}{2}\mathbf{x}^T (A + \rho^* I)\mathbf{x} - \overline{W}^\sharp(\epsilon^*, \rho^*) + (B\mathbf{x})^T \epsilon^* - \mathbf{x}^T \mathbf{f}. \tag{35}$$

Particularly, if $\bar{\rho}_i^* > -a_1$, the matrix $(A + \rho_i^* I)$ is positive definite, the canonical dual function $P^d(\epsilon^*, \rho^*)$ is concave in each of its components $\epsilon^*$ and $\rho^*$, respectively. In this case, the extended Lagrangian $\Xi$ is convex in $\mathbf{x} \in \mathbb{R}^n$ and concave in each $\epsilon^* \in \mathbb{R}^m$ and $\rho^* \in \mathbb{R}$ ($\Xi$ is concave in each $\epsilon^*$ and $\rho^*$ does note imply that $\Xi$ is concave in the vector $(\epsilon^*, \rho^*)$, see Remark 2.6.1 in [11], p. 82). Thus, we have

$$\begin{aligned}
P^d(\bar{\epsilon}_i^*, \bar{\rho}_i^*) &= \max_{(\epsilon^*, \rho^*) \in \mathcal{Y}_{\mu+}^*} P^d(\epsilon^*, \rho^*) \\
&= \max_{\rho^* > -a_1} \max_{\epsilon^* \geqslant 0} \min_{\mathbf{x} \in \mathbb{R}^n} \Xi(\mathbf{x}, \epsilon^*, \rho^*) \\
&= \max_{\rho^* > -a_1} \min_{\mathbf{x} \in \mathbb{R}^n} \max_{\epsilon^* \geqslant 0} \left\{ \frac{1}{2}\mathbf{x}^T (A + \rho^* I)\mathbf{x} + (B\mathbf{x} - b)^T \epsilon^* - \mu\rho^* - \mathbf{x}^T \mathbf{f} \right\} \\
&= \max_{\rho^* > -a_1} \min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \frac{1}{2}\mathbf{x}^T (A + \rho^* I)\mathbf{x} - \mu\rho^* - \mathbf{x}^T \mathbf{f} \right\} \quad \text{s.t. } B\mathbf{x} \leqslant \mathbf{b} \\
&= \min_{\mathbf{x} \in \mathcal{X}_f} \left\{ P(\mathbf{x}) + \max_{\rho^* > -a_1} \rho^* \left( \frac{1}{2}\mathbf{x}^T \mathbf{x} - \mu \right) \right\} \\
&= \min_{\mathbf{x} \in \mathcal{X}_f} P(\mathbf{x}) \text{ s.t. } \frac{1}{2}\mathbf{x}^T \mathbf{x} = \mu,
\end{aligned}$$

since the linear programming

$$\theta_1 = \max_{\rho^* > -a_1} \rho^* \left( \frac{1}{2}\mathbf{x}^T \mathbf{x} - \mu \right)$$

has a solution in the open domain $(-a_1, +\infty)$ if and only if $\frac{1}{2}\mathbf{x}^T \mathbf{x} - \mu = 0$. By the fact that if $i_d > 0$, then $\bar{\rho}_i^* > -a_1 > 0$, the KKT condition (25) leads to $\frac{1}{2}|\bar{\mathbf{x}}_i|^2 = \mu$ also. From Theorem 1 we have (31).

On the other hand, if $\bar{\rho}_i^* \in [0, -a_{i_d})$, i.e. $0 \leqslant \bar{\rho}_i^* < -a_{i_d}$, and the matrix $A + \bar{\rho}_i^* I$ is negative definite. In this case, the extended Lagrangian $\Xi(\mathbf{x}, \epsilon^*, \rho^*)$ is concave in $\mathbf{x} \in \mathbb{R}^n$ and concave in both $\epsilon^* \in \mathbb{R}_+^m$ and $\rho^* \in [0, -a_{i_d})$. Thus, if $(\bar{\epsilon}_i^*, \bar{\rho}_i^*)$ is a global maximizer of $P^d$ on $\mathcal{Y}^*_{\mu-}$, then by the so-called *super-Lagrange duality theory* developed in [11] and the triality lemma in Section 6, we have

$$\begin{aligned}
P^d(\bar{\epsilon}^*_i, \bar{\rho}^*_i) &= \max_{(\epsilon^*, \rho^*)\in\mathcal{Y}^*_{\mu-}} P^d(\epsilon^*, \rho^*) \\
&= \max_{\rho^*\in[0,-a_{i_d})} \max_{\epsilon^*\geqslant 0} \max_{\mathbf{x}\in\mathbb{R}^n} \Xi(\mathbf{x}, \epsilon^*, \rho^*) \\
&= \max_{\rho^*\in[0,-a_{i_d})} \max_{\mathbf{x}\in\mathbb{R}^n} \max_{\epsilon^*\geqslant 0} \left\{ \frac{1}{2}\mathbf{x}^T(A+\rho^* I)\mathbf{x} + (B\mathbf{x}-\mathbf{b})^T\epsilon^* - \mu\rho^* - \mathbf{x}^T\mathbf{f} \right\} \\
&= \max_{\rho^*\in[0,-a_{i_d})} \max_{\mathbf{x}\in\mathbb{R}^n} \left\{ \frac{1}{2}\mathbf{x}^T(A+\rho^* I)\mathbf{x} - \mu\rho^* - \mathbf{x}^T\mathbf{f} \right\} \quad \text{s.t. } B\mathbf{x}\leqslant\mathbf{b} \\
&= \max_{\mathbf{x}\in\mathcal{X}_f} \left\{ P(\mathbf{x}) + \max_{\rho^*\in[0,-a_{i_d})} \rho^*\left(\frac{1}{2}\mathbf{x}^T\mathbf{x} - \mu\right) \right\} \\
&= \max_{\mathbf{x}\in\mathcal{X}_f} P(\mathbf{x}) \quad \text{s.t. } \frac{1}{2}\mathbf{x}^T\mathbf{x} \leqslant \mu,
\end{aligned}$$

by the fact that the domain $[0, -a_{i_d})$ is closed on the lower bound and open on the upper bound, the problem

$$\theta_2 = \max_{\rho^*\in[0,-a_{i_d})} \left\{ \rho^*\left(\frac{1}{2}\mathbf{x}^T\mathbf{x} - \mu\right) \right\} \tag{36}$$

has a solution if and only if $\frac{1}{2}\mathbf{x}^T\mathbf{x}\leqslant\mu$, and for this solution, $\theta_2 = 0$. Thus, by Theorem 1 and the equality (34) we have (32).

Finally, if $0 < \bar{\rho}^*_i < -a_{i_d}$, and $(\bar{\epsilon}^*_i, \bar{\rho}^*_i)$ is a global minimizer of $P^d$ on $\mathcal{Y}^*_{\mu i}$, then the super-Lagrange duality theory and the triality lemma lead to

$$\begin{aligned}
P^d(\bar{\epsilon}^*_i, \bar{\rho}^*_i) &= \min_{(\epsilon^*, \rho^*)\in\mathcal{Y}^*_{\mu i}} P^d(\epsilon^*, \rho^*) \\
&= \min_{\epsilon^*\geqslant 0} \min_{\rho^*\in(0,-a_{i_d})} \max_{\mathbf{x}\in\mathbb{R}^n} \Xi(\mathbf{x}, \epsilon^*, \rho^*) \\
&= \min_{\rho^*\in(0,-a_{i_d})} \min_{\mathbf{x}\in\mathbb{R}^n} \max_{\epsilon^*\geqslant 0} \Xi(\mathbf{x}, \epsilon^*, \rho^*) \\
&= \min_{\rho^*\in(0,-a_{i_d})} \min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \frac{1}{2}\mathbf{x}^T(A+\rho^* I)\mathbf{x} - \mu\rho^* - \mathbf{x}^T\mathbf{f} \right\} \quad \text{s.t. } B\mathbf{x}\leqslant\mathbf{b} \\
&= \min_{x\in\mathbb{R}^n} \left\{ P(\mathbf{x}) + \min_{\rho^*\in(0,-a_{i_d})} \rho^*\left(\frac{1}{2}\mathbf{x}^T\mathbf{x} - \mu\right) \right\} \quad \text{s.t. } B\mathbf{x}\leqslant\mathbf{b} \\
&= \min_{\mathbf{x}\in\mathbb{R}^n} P(\mathbf{x}) \quad \text{s.t. } B\mathbf{x}\leqslant\mathbf{b},\ \frac{1}{2}|\mathbf{x}|^2 = \mu,
\end{aligned}$$

since the linear minimization

$$\theta_3 = \min_{\rho^* \in (0, -a_{i_d})} \rho^* \left( \frac{1}{2} \mathbf{x}^T \mathbf{x} - \mu \right)$$

has solution on the open domain $(0, -a_{i_d})$ if and only $\frac{1}{2}\mathbf{x}^T\mathbf{x} - \mu = 0$. By the fact that $P^d(\bar{\epsilon}_i^*, \bar{\rho}_i^*) = P(\bar{\mathbf{x}}_i)$ for all KKT points of $(\mathcal{P}_\mu)$, the theorem is proved. □

*Remark* 1. Theorem 2 and its proof show an important fact that if the extrema $\bar{\mathbf{x}}_i$ of the primal problem $(\mathcal{P}_\mu)$ are on the nonconvex set $\mathcal{X}_{\mu s}$, i.e. the boundary of the sphere $|\mathbf{x}|^2 = 2\mu$, then the associated KKT points $\bar{\rho}_i^*$ should be interior points of the dual feasible set, and vice versa. This fact is due to the canonical duality and the complementarity condition. In this case, the minimizers $\bar{\mathbf{x}}_i$ are usually not critical points of $P$. This is one of main reasons why the primal problem is NP-hard. However, for each given $\epsilon^* \in \mathbb{R}^m$, the dual solutions $\bar{\rho}_i^* > 0$ are critical points of the canonical dual function $P^d$ controlled by the dual algebraic equation

$$\frac{1}{2}(\mathbf{f} - B^T\bar{\epsilon}^*)^T (A + \bar{\rho}_i^* I)^{-2} (\mathbf{f} - B^T\bar{\epsilon}^*) = \mu, \tag{37}$$

which can be solved completely by MATHEMATICA. For a given sufficiently large parameter $\mu$, this nonlinear algebraic equation has a unique root $\bar{\rho}^* > -a_1$, which maximizes $P^d$ on the open domain $(-a_1, \infty)$. However, on the open domain $(0, -a_{i_d})$, the dual algebraic equation (37) has at most two roots $\bar{\rho}_{i_d}^* \geqslant \bar{\rho}_{2i_d+1}$. By Theorem 3 in the next section we know that $\bar{\rho}_{2i_d}^*$ is a local minimizer of $P^d$, while $\bar{\rho}_{2i_d+1}^*$ is a local maximizer. Since there is no duality gap between the primal and the dual problems, for a given sufficiently large $\mu > 0$, the triality theory (Theorem 2) can be used to find minimizers of $P(\mathbf{x})$ on the nonconvex set $\mathcal{X}_{\mu s}$. If the feasible set $\mathcal{X}_f$ is bounded, then we can choose $\mu = \mu_0$. In this case, $\mathcal{X}_\mu = \mathcal{X}_f$, and the vector

$$\bar{\mathbf{x}}_i = (A + \bar{\rho}_i^* I)^{-1} (\mathbf{f} - B^T \bar{\epsilon}_i^*)$$

is also a minimizer to the original problem $(\mathcal{P})$. But this minimizer may be not a global minimizer of $P(\mathbf{x})$ on the whole feasible set $\mathcal{X}_f$ since the canonical dual problem $(\mathcal{P}_\mu^d)$ may have some other KKT points $\bar{\rho}_i^*$ located between each open interval $(-a_{j+1}, -a_j), j = i, \ldots, i_d - 1$. (see Theorem 3). If the feasible set $\mathcal{X}_f$ is unbounded, the global minimizer of the problem $(\mathcal{P})$ may even not exist. Physically speaking, the primal problem $(\mathcal{P})$ might be not well-proposed (see page 180, [11]). In this case, the parametrization $(\mathcal{P}_\mu)$ can be used be find certain useful global minimizers within the sphere $\frac{1}{2}|\mathbf{x}|^2 \leqslant \mu$.

## 4. Quadratic Programming Over a Sphere

As a particularly important application of the canonical dual transformation and the parametrization method, let us consider the following quadratic programming with only a quadratic constraint over a sphere:

$$(\mathcal{P}_q) \quad \min \ \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{f}^T\mathbf{x}$$
$$\text{s.t. } \frac{1}{2}\mathbf{x}^T\mathbf{x} \leqslant \mu. \tag{38}$$

This problem often comes up as a subproblem in general optimization algorithms (cf. Powell, 2002). Often, in the model trust region methods, the objective function in nonlinear programming is approximated locally by a quadratic function. In such cases, the approximation is restricted to a small region around the current iterate. If the 2-norm is used to define this region, then these methods ended up with the quadratic programming over a sphere $(\mathcal{P}_q)$.

As indicated by Floudas and Visweswaran (1995), due to the presence of the nonlinear sphere constraint, the solution of $(\mathcal{P}_q)$ is likely to be irrational, which implies that it is not possible to exactly compute the solution. Therefore, many polynomial time algorithms have been suggested to compute the approximate solution to this problem (see, Sorensen, 1982; Karmarkar, 1990; and Ye, 1992). However, by the canonical dual transformation, this problem can be solved completely. Since there is no linear inequality constraint $B\mathbf{x} \leqslant \mathbf{b}$, the canonical dual problem $(\mathcal{P}^d_\mu)$ in this case is simply a concave maximization in $\mathbb{R}$:

$$(\mathcal{P}^d_q): \quad \text{ext } P^d(\rho^*) = -\frac{1}{2}\mathbf{f}^T(A+\rho^* I)^{-1}\mathbf{f} - \mu\rho^*, \tag{39}$$
$$\text{s.t. } \rho^* \geqslant 0, \quad \det(A+\rho^* I) \neq 0. \tag{40}$$

This is a concave maximization with only one degree-of-freedom. The following theorem presents a complete set of solutions for this dual problem.

THEOREM 3 (Complete solution to $(\mathcal{P}_q)$). *Suppose that the symmetric matrix $A$ has $p \leqslant n$ distinct eigenvalues, and $i_d \leqslant p$ of them are negative such that*

$$a_1 < a_2 < \cdots < a_{i_d} < 0 \leqslant a_{i_d+1} < \cdots < a_p.$$

*Then for a given vector $\mathbf{f} \in \mathbb{R}^n$, and a sufficiently large parameter $\mu > 0$, the canonical dual problem $(\mathcal{P}^d_p)$ has at most $2i_d+1$ KKT points $\bar{\rho}^*_i, i = 1, \ldots, 2i_d+1$ satisfying the following distribution law*

$$\bar{\rho}^*_1 > -a_1 > \bar{\rho}^*_2 \geqslant \bar{\rho}^*_3 > -a_2 > \cdots > -a_{i_d} > \bar{\rho}^*_{2i_d} \geqslant \bar{\rho}^*_{2i_d+1} > 0. \tag{41}$$

*For each* $\bar{\rho}_i^*, i=1,\dots,2i_d+1$, *the vector defined by*

$$\bar{\mathbf{x}}_i=(A+\bar{\rho}_i^* I)^{-1}\mathbf{f} \tag{42}$$

*is a KKT spoint of the problem* $(\mathcal{P}_q)$ *and*

$$P(\bar{\mathbf{x}}_i)=P^d(\bar{\rho}_i^*), \quad i=1,2,\dots,2i_d+1. \tag{43}$$

*Moreover, if* $i_d>0$, *then the problem* $(\mathcal{P}_d)$ *has at most* $2i_d+1$ *KKT points on the boundary of the sphere, i.e.*

$$\frac{1}{2}|\bar{\mathbf{x}}_i|^2=\mu, \quad i=1,\dots,2i_d+1, \tag{44}$$

*and* $\bar{\mathbf{x}}_1$ *is a global minimizer of the problem* $(\mathcal{P}_q)$.

*Proof.* For a given $\mathbf{f}\in\mathbb{R}^n$ and $\mu>0$, the scalar $\bar{\rho}^*$ is a KKT point of the dual problem $(\mathcal{P}_q^d)$ if and only if

$$0\leqslant\bar{\rho}^* \perp \frac{1}{2}\mathbf{f}^T(A+\rho^* I)^{-2}\mathbf{f}-\mu\leqslant 0. \tag{45}$$

In term of $\bar{\mathbf{x}}=(A+\rho^* I)^{-1}\mathbf{f}$, (45) can be written as

$$0\leqslant\bar{\rho}^* \perp \frac{1}{2}\bar{\mathbf{x}}^T\bar{\mathbf{x}}-\mu\leqslant 0.$$

This is the KKT condition for $(\mathcal{P}_q)$. Thus for each KKT point $\bar{\rho}^*$ of $(\mathcal{P}_q^d)$, the vector $\bar{\mathbf{x}}=(A+\rho^* I)^{-1}\mathbf{f}$ is a KKT point of the primal problem $(\mathcal{P}_q)$.

Since $A=A^T$, there exists an orthogonal matrix $R^T=R^{-1}$ such that $A=R^T DR$, where $D=(a_i\delta_{ij})$ is a diagonal matrix. For the given vector $\mathbf{f}\in\mathbb{R}^n$, let $\mathbf{g}=R\mathbf{f}=(g_i)$ and

$$\psi(\rho^*)=\frac{1}{2}\mathbf{f}^T(A+\rho^* I)^{-2}\mathbf{f}=\frac{1}{2}\sum_{i=1}^{p}g_i^2(a_i+\rho^*)^{-2}. \tag{46}$$

Clearly, this real valued function $\psi(\rho^*)$ is strictly convex within each interval $-a_{i+1}<\rho^*<-a_i$, as well as the intervals $-\infty<\rho^*<-a_p$ and $-a_1<\rho^*<\infty$ (see Figure 1).

Thus, for a given sufficiently large parameter $\mu>0$, the algebraic equation

$$\psi(\rho^*)=\frac{1}{2}\sum_{i=1}^{p}g_i^2(a_i+\rho^*)^{-2}=\mu \tag{47}$$

have at most $2p$ solutions $\{\bar{\rho}_i^*\}$ satisfying $-a_{j+1}<\bar{\rho}_{2j+1}^*\leqslant\bar{\rho}_{2j}^*<-a_j$ for $j=1,\dots,p-1$, and $\bar{\rho}_1^*>-a_1, \bar{\rho}_{2p}^*<-a_p$. Since $A$ has only $i_d$ negative eigenvalues, the equality $\psi(\rho^*)=\mu$ has at most $2i_d+1$ strictly positive roots

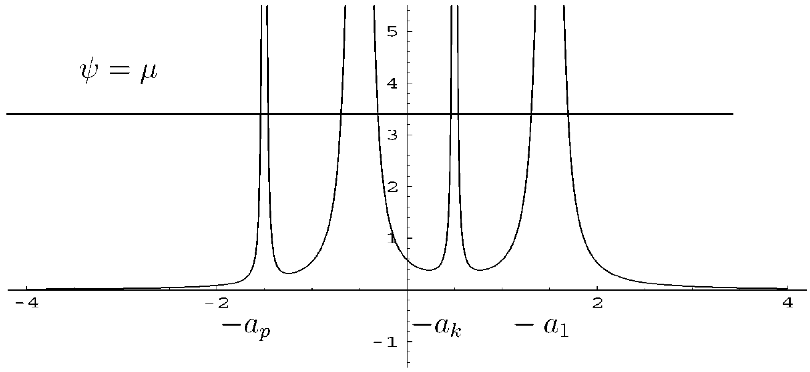


*Figure 1.* Graph of $\psi(\rho^*)$.

$\{\bar{\rho}_i^*\} > 0$, $i = 1, \ldots, 2i_d + 1$. By the complementarity condition $\bar{\rho}_i^* \left(\frac{1}{2}|\bar{\mathbf{x}}_i|^2 - \mu\right) = 0$, we know that the primal problem $(\mathcal{P}_q)$ has at most $2i_d + 1$ KKT points $\bar{\mathbf{x}}_i$ on the sphere $\frac{1}{2}|\bar{\mathbf{x}}_i|^2 = \mu$. If $a_{i_d+1} > 0$, the equality $\psi(\rho^*) = \mu$ may have at most $2i_d$ strictly positive roots. □

Theorem 3 presents a complete set of KKT points to the quadratic programming over the sphere since the dual algebraic equation (47) can be solved completely by MATHEMATICA. From Figure 1 we can see that for a given matrix $A$ and the parameter $\mu > 0$, the canonical dual problem $(\mathcal{P}_q^d)$ has only one solution $\bar{\rho}^* > -a_1$, which leads to a global minimizer $\bar{\mathbf{x}}_1$ of the primal problem $(\mathcal{P}_q)$. This theorem will play an important role in nonconvex quadratic programming.

## 5. Applications

We now list a few examples to illustrate the applications of the theory presented in this paper.

EXAMPLE 1 (One-D concave minimization). First of all, let us consider one dimensional concave minimization problem:

$$\min P(x) = \frac{1}{2}ax^2 - fx, \quad \text{s.t. } |x| \leqslant r. \tag{48}$$

Clearly, if $a < 0$, the global minimizer of $P(x)$ has to be one of boundary points $\bar{\mathbf{x}} = \pm r$. In this case, $\mu = \frac{1}{2}r^2$. The canonical dual problem is

$$\max P^d(\rho^*) = -\frac{1}{2}f^2/(a + \rho^*) - \mu\rho^*, \quad \text{s.t. } (a + \rho^*) > 0. \tag{49}$$

Since $n = 1$, the dual algebraic equation $\frac{1}{2}f^2/(a + \rho^*)^2 - \mu$ has only two roots: $\bar{\rho}_1^* > -a$ is a unique maximizer of $P^d$, and $\bar{\rho}_2^* < -a$ is a local minimizer. If we choose $f = .4$, $a = -.6$ and $r = 1.5$, the global maximizer $\bar{\rho}_1^* = 0.866667$,

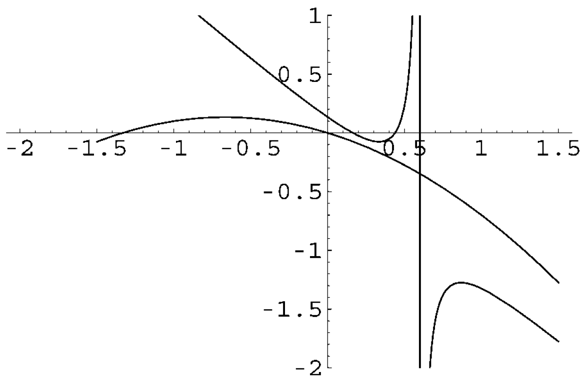


*Figure 2.* Graphs of $P(x)$ and $P^d(\rho^*)$ for one dimensional problem.

which gives the global minimizer $\bar{x}_1 = f/(a+\bar{\rho}_1^*) = 1.5$. It is easy to check that $P(\bar{x}_1) = -1.275 = P^d(\bar{\rho}_1^*)$. While the local minimizer $\bar{\rho}_2^* = 0.3333$, which gives the local minimizer $\bar{x}_2 = -1.5$. Since for $\bar{\rho}_2^* < -a$, the extended Lagrangian (35) is a so-called *super-Lagrangian* (cf. [11]). In this case, the double-min duality theory leads to $P(\bar{x}_2) = -0.075 = P^d(\bar{\rho}_2^*)$. It is interesting to note that for $\bar{\rho}_3^* = 0$, then $\bar{x}_3 = f/(a+\bar{\rho}_3^*) = -0.666667$ is a global maximizer of $P(x)$ and we have $P(\bar{x}_3) = P^d(\bar{\rho}_3^*) = .13333$. The graphs of $P(x)$ and $P^d(\rho^*)$ are shown in Figure 2.

EXAMPLE 2 (Two-D concave minimization within convex set). We now consider the following quadratic programming within a convex set:

$$\min\ P(x_1, x_2) = \frac{1}{2}(a_1 x_1^2 + a_2 x_2^2) - f_1 x_1 - f_2 x_2 \tag{50}$$

$$\text{s.t.}\ \frac{1}{2}x_1 + x_2 \leqslant 1, \quad x_2 \geqslant 0, \quad \frac{1}{2}(x_1^2 + x_2^2) \leqslant 2. \tag{51}$$

In this case, the radius of the feasible set $\mathcal{X}_f = \{(x_1, \mathbf{x}_2) \in \mathbb{R}^2 \,|\, B\mathbf{x} \leqslant \mathbf{b},\ \frac{1}{2}(x_1^2 + x_2^2) \leqslant 2\}$ is $r_0 = 2$, in which, $B = \left\{\left\{\frac{1}{2}, 1\right\}, \{0, -1\}\right\}$ is a $2 \times 2$ matrix, $\mathbf{b} = \{1, 0\}$ is a 2-vector. If both $a_1, a_2 \leqslant 0$, $P$ is concave and its global minima must be located on the boundary of $\mathcal{X}_f$ (see Figure 3). The canonical dual problem in this case is to find $(\epsilon^*, \rho^*) \in \mathbb{R}^2 \times \mathbb{R}$ such that

$$\max\ P^d(\epsilon_1^*, \epsilon_2^*, \rho^*) = -\frac{1}{2}\left\{\frac{\left(f_1 - \frac{1}{2}\epsilon_1^*\right)^2}{a_1 + \rho^*} + \frac{(f_2 - \epsilon_1^* + \epsilon_2^*)^2}{a_2 + \rho^*}\right\} - \mu\rho^* - \epsilon_1^* \tag{52}$$

$$\text{s.t.}\ \epsilon_1^* \geqslant 0, \quad \epsilon_2^* \geqslant 0, \quad \rho^* \geqslant -\min\{a_1, a_2\}. \tag{53}$$

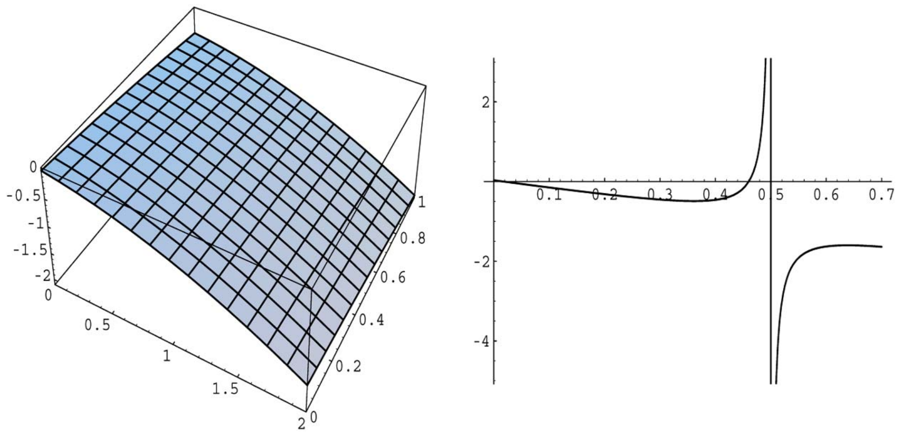


*Figure 3.* Graphs of $P(x_1, x_2)$ and $P^d(\bar{\epsilon}_1^*, \bar{\epsilon}_2^*, \rho^*)$.

If we let $\mathbf{f} = (.3, .3)$, $a_1 = -0.5$, $a_2 = -0.3$, $\mu = \frac{1}{2} r_0^2 = 2$, then this dual problem has a unique solution:

$$\bar{\rho}^* = 0.548375, \quad \bar{\epsilon}_1^* = 0.406502, \quad \bar{\epsilon}_2^* = 0.106507$$

in the domain $\mathcal{Y}^*_{\mu+}$. This leads to a global minimizer $\bar{x}_1 = 2.0$, $\bar{x}_2 = 0$. It is easy to verify that $P(\bar{x}_1, \bar{x}_2) = -1.6 = P^d(\bar{\epsilon}_1^*, \bar{\epsilon}_2^*, \bar{\rho}^*)$.

EXAMPLE 3 (Quadratic programming over a 4-d sphere). We simply let $A$ is a diagonal matrix with four non zero eigenvalues: $\{a_1 = -0.5,\ a_2 = -0.25,\ a_3 = 0.1,\ a_4 = 0.4\}$. If we choose $\mathbf{f} = (.3,\ .4,\ -.2,\ .1)$, and $\mu = 2$, the canonical dual algebraic equation (47) has four real roots (see Figure 4)

$$\rho_4^* = -0.455359 < \rho_3^* = -0.339287 < \rho_2^* = -0.219664 < \rho_1^* = 0.672415.$$

Since $\rho_1^* > 0$ and $(A + \rho_1^* I)$ is positive definite, so $\rho_1^*$ is a global maximizer of $P^d$, which leads to the global minimizer

$$\bar{\mathbf{x}}_1 = (A + \rho_1^* I)^{-1} \mathbf{f} = (1.73999, 0.946937, -0.258928, 0.0932475)$$

on the boundary of the 4-D sphere $|\mathbf{x}| \leqslant 2$, i.e. $(\bar{x}_1^2 + \bar{x}_2^2 + \bar{x}_3^2 + \bar{x}_4^2)^{1/2} = 2$. This is the reason why the primal problem is very difficult. However, the dual problem is a concave maximization programming and the global maximizer is in the interior of the dual feasible set $\mathcal{Y}^*_\mu$.

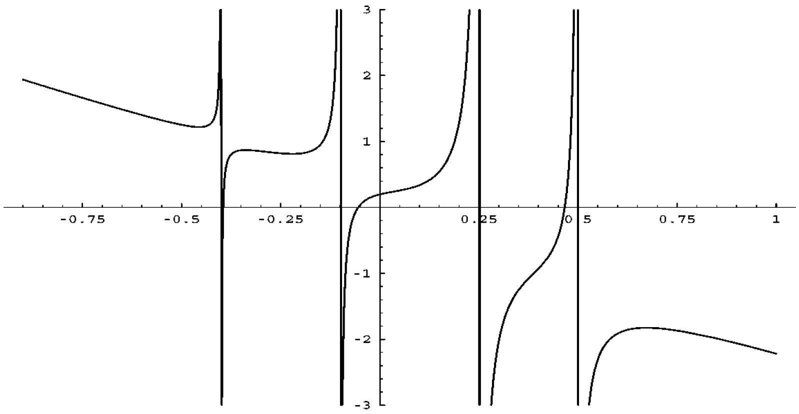


*Figure 4.* Graphs of $P^d(\rho^*)$ in four dimensional problem.

## 6. Triality Lemma for Quadratic Programming

The triality theory was originally discovered in large deformation mechanics (see [7]), where the total potential $P(\mathbf{u})$ is a nonconvex functional in infinite dimensional space and the variable $\mathbf{u}(\mathbf{x})$ is a vector field function. At each material point $\mathbf{x}$ of the system, the Hessian matrix $D^2P(\bar{\mathbf{u}})$ usually has non zero index $i_d > 0$ at its critical points $\bar{\mathbf{u}}(\mathbf{x})$. Applications of this triality theory to global optimization problems in finite dimensional systems were presented in [13]. For the nonconvex quadratic programming problem $(\mathcal{P})$ studied in this paper, if the total cost $P(\mathbf{x})$ is a canonical function, i.e. the matrix $A$ is invertible, the triality theory has a particular simple format.

Recall the classical Lagrangian $L$: $\mathbb{R}^n \times \mathbb{R}^m_+ \to \mathbb{R}$ associated with the primal problem $(\mathcal{P})$, i.e. the equation (4)

$$L(\mathbf{x}, \nu) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{f}^T\mathbf{x} + \nu^T(B\mathbf{x} - \mathbf{b}).$$

Since the matrix $A$ is invertible, for any given $\nu \in \mathbb{R}^m_+$, the canonical dual function $P^d_\nu(\nu)$ can be defined by the canonical transformation

$$\begin{aligned} P^d_\nu(\nu) &= \text{ext}_{\mathbf{x}\in\mathbb{R}^n} L(\mathbf{x}, \nu) \\ &= -\frac{1}{2}(\mathbf{f} - B^T\nu)^T A^{-1}(\mathbf{f} - B^T\nu) - \mathbf{b}^T\nu, \end{aligned}$$

which is well defined on $\mathbb{R}^m_+$. Thus, the canonical dual problem ($(\mathcal{P}^d{}_\nu)$ in short) is to find the extrema $\bar{\nu} \in \mathbb{R}^m$ such that

$$(\mathcal{P}^d{}_\nu):\ P^d_\nu(\bar{\nu}) = \text{ext}P^d_\nu(\nu) \quad \forall \nu \geqslant 0. \tag{54}$$

It is easy to prove that the primal problem $(\mathcal{P})$ is canonically dual to $(\mathcal{P}^d{}_\nu)$ in the sense that they have the same set of KKT points, and at each KKT point $(\bar{\mathbf{x}}, \bar{\nu})$,

$$P(\bar{\mathbf{x}}) = L(\bar{\mathbf{x}}, \bar{\nu}) = P^d_\nu(\bar{\nu}). \tag{55}$$

Clearly if the symmetrical matrix $A$ is positive definite, then the quadratic function $P(\mathbf{x})$ is convex over $\mathcal{X}_f$, and $P^d_\nu(\nu)$ is concave over $\mathbb{R}^m_+$. In this case, the Lagrangian $L(\mathbf{x},\nu)$ is a saddle function. However, if $A$ is negative definite, then $P(\mathbf{x})$ is concave over $\mathcal{X}_f$, and $P^d_\nu(\nu)$ is convex over $\mathbb{R}^m_+$. In this case, $L(\mathbf{x},\nu)$ is a *super-Lagrangian*, i.e. $L$ is concave in each $\mathbf{x}$ and $\nu$[1]. Then by the general triality theory developed in [11], we have

LEMMA 1 (Triality for canonical quadratic programming). *Suppose that the vector $(\bar{\mathbf{x}},\bar{\nu})$ is a KKT point of the quadratic programming problem $(\mathcal{P})$.*

*If A is positive definite, then the saddle minmax theorem in the form*

$$\min_{\mathbf{x}\in\mathcal{X}_f}\max_{\nu\geqslant 0} L(\mathbf{x},\nu)=L(\bar{\mathbf{x}},\bar{\nu})=\max_{\nu\geqslant 0}\min_{\mathbf{x}\in\mathbb{R}^n} L(\mathbf{x},\nu) \tag{56}$$

*holds.*

*If A is negative definite, then either the super-maximum theorem in the form*

$$\max_{\mathbf{x}\in\mathcal{X}_f}\max_{\nu\geqslant 0} L(\mathbf{x},\nu)=L(\bar{\mathbf{x}},\bar{\nu}=\max_{\nu\geqslant 0}\max_{\mathbf{x}\in\mathbb{R}^n} L(\mathbf{x},\nu) \tag{57}$$

*holds, or the super-minimum theorem in the form*

$$\min_{\mathbf{x}\in\mathcal{X}_f}\max_{\nu\geqslant 0} L(\mathbf{x},\nu)=L(\bar{\mathbf{x}},\bar{\nu})=\min_{\nu\geqslant 0}\max_{\mathbf{x}\in\mathbb{R}^n} L(\mathbf{x},\nu) \tag{58}$$

*holds.*

*Proof.* The statement (56) is the classical saddle Lagrangian minmax theorem. We need to prove only (57) and (58). Since $L$ is concave in each $\mathbf{x}$ and $\nu$, if $(\bar{\mathbf{x}},\bar{\nu})\in\mathcal{X}_f\times\mathbb{R}^m_+$ is a KKT point, we have

$$\max_{\mathbf{x}\in\mathcal{X}_f} L(\mathbf{x},\bar{\nu})=L(\bar{\mathbf{x}},\bar{\nu})=\max_{\nu\in\mathbb{R}^m_+} L(\bar{\mathbf{x}},\nu). \tag{59}$$

Since the sets $\mathcal{X}_f$ and $\mathbb{R}^m_+$ are not empty, for any given $\mathbf{x}\in\mathcal{X}_f$, we have

$$L(\mathbf{x},\bar{\nu})=\max_{\nu\geqslant 0} L(\mathbf{x},\nu)=P(\mathbf{x}), \tag{60}$$

and also, for any given $\nu\geqslant 0$.

$$L(\bar{\mathbf{x}},\nu)=\max_{\mathbf{x}\in\mathbb{R}^n} L(\mathbf{x},\nu)=P^d(\nu). \tag{61}$$

Thus, by substituting both (61) and (60) into (59), the statement (57) is proved.

[1]The super-Lagrangian $L$ is concave in each $\mathbf{x}$ and $\nu$ does not imply that $L$ is concave in the vector $(\mathbf{x},\nu)$, see the example given in [11], p. 82.

Now let us consider the statement (58). If for a fixed $\bar{\nu}\in\mathbb{R}^m_+$, the vector $\bar{\mathbf{x}}$ minimizes $L(\mathbf{x},\bar{\nu})$ on the primal feasible set $\mathcal{X}_f$, then,

$$L(\bar{\mathbf{x}},\bar{\nu})=\min_{\mathbf{x}\in\mathcal{X}_f}L(\mathbf{x},\bar{\nu})=\min_{\mathbf{x}\in\mathcal{X}_f}\max_{\nu\geqslant 0}L(\mathbf{x},\nu). \tag{62}$$

On the other hand, for a fixed $\bar{\mathbf{x}}\in\mathcal{X}_f$, the KKT point $\bar{\nu}$ should be an extremer of $L(\bar{\mathbf{x}},\nu)=P^d_\nu(\nu)$ over $\mathbb{R}^m_+$. This means that $\bar{\nu}$ is either a global min or global max of the continuous dual function $P^d_\nu$. If $\bar{\nu}$ maximizes $P^d_\nu$, then

$$L(\bar{\mathbf{x}},\bar{\nu})=\max_{\nu\geqslant 0}L(\bar{\mathbf{x}},\nu)=\max_{\nu\geqslant 0}\max_{\mathbf{x}\in\mathbb{R}^n}L(\mathbf{x},\nu)=\max_{\mathbf{x}\in\mathcal{X}_f}\max_{\nu\geqslant 0}L(\mathbf{x},\nu).$$

This contradicts to (62), i.e. the vector $\bar{\nu}$ can not be a global maximizer of $L(\bar{\mathbf{x}},\nu)=P^d_\nu(\nu)$ when the vector $\bar{\mathbf{x}}$ minimizes $L(\mathbf{x},\bar{\nu})$ over $\mathcal{X}_f$. Thus, $\bar{\nu}$ should be a global minimizer of $L(\bar{\mathbf{x}},\nu)$ on $\mathbb{R}^m_+$. This proved that if $(\bar{\mathbf{x}},\bar{\nu})$ is a KKT point of $L(\mathbf{x},\nu)$, the super-minimax theorem in the form of (58) holds. □

The triality lemma for quadratic programming can also be proved by the so-called *bi-duality theory* developed in [11]. Actually, by introducing the indicator function

$$\mathcal{I}_{\mathcal{X}_f}(B\mathbf{x})=\begin{cases}0 & \text{if } \; B\mathbf{x}\leqslant\mathbf{b},\\ +\infty & \text{otherwise}\end{cases}$$

the constrained primal problem $(\mathcal{P})$ can be written as an unconstrained programming

$$\min_{\mathbf{x}\in\mathbb{R}^n}P(\mathbf{x})=\mathcal{I}_{\mathcal{X}_f}(B\mathbf{x})+P(\mathbf{x}) \tag{63}$$

This is the so-called d.c. programming when the matrix $A$ is negative. In this case, the statement (57) is the so-called double-max duality theory (see [11])

$$\max_{\mathbf{x}\in\mathcal{X}_f}P(\mathbf{x})=\max_{\nu\geqslant 0}P^d(\nu), \tag{64}$$

while the statement (58) is the well-known double-min duality theory

$$\min_{\mathbf{x}\in\mathcal{X}_f}P(\mathbf{x})=\min_{\nu\geqslant 0}P^d(\nu) \tag{65}$$

in d.c. programming. The bi-duality theory plays an important role in convex Hamilton systems, where the total action of the systems is a d.c. functional. However, in nonconvex mechanics and chaotic dynamic systems, where the total energy $P$ is usually a nonconvex functional, the triality theory is needed to clarify local and global extrema (see [8, 12]).

## 7. Concluding Remarks

We have presented a concrete application of the canonical dual transformation and triality theory to constrained quadratic programming. Results shown that by use of this method, the nonconvex constrained problem $(\mathcal{P}_\mu)$ in $\mathbb{R}^n$ can be reformulated as a perfect dual problem in $\mathbb{R}^{m+1}$, while the quadratic programming problem $(\mathcal{P}_q)$ over a sphere in $\mathbb{R}^n$ is equivalent to an one dimensional dual problem $(\mathcal{P}_q^d)$, which can be solved completely. The *KKT* points and extremality conditions of these originally difficult problems are identified by Theorem 2 and Theorem 3, which are actually applications of the general triality theory.

The canonical dual transformation method and triality theory were originally developed from nonconvex mechanics (see [7]). Mathematically speaking, numerical discretizations of nonconvex problems in infinite dimensional space usually lead to very complicated nonconvex optimization problems with many local minimizers (see [13]). Physically speaking, each local minimizer represents a local stable equilibrium state of the system. Triality theory reveals the intrinsic pattern of duality relations of these local critical points, and plays an important role in nonconvex analysis. Detailed study and comprehensive applications of this interesting theory, as well as the associated bi-duality and polarity theories, were presented in the monograph [11]. Generalization and applications have been made into global optimization problems (see [13]). A general *triality algorithm* was proposed recently for solving nonconvex dynamical problems and phase transitions in nonconvex mechanics (see [15, 17, 18]).

The present paper shows again that the canonical dual transformation and associated triality theory may possess important computational impacts on global optimization. The extremality conditions presented in Theorem 2 are only for the open domains $\rho^* < -a_{i_d}$ and $\rho^* > -a_1$. The dual KKT points could be also located in each open domain $(-a_{i+1}, -a_i)$, $i = 1, \ldots, i_d - 1$, which might leads to a global minimizer of the primal problem $(\mathcal{P}_\mu)$. The triality theory should play essential role in the further study on this very important problem.

**Acknowledgements**

The author would like to offer his sincere thanks to Professor C.J. Goh at the University of Western Australia for his very detailed comments and extremely important suggestions, which make this paper readable and more understandable. Comments from Professor Panos Pardalos and his student O. Prokopyev at the University of Florida are also acknowledged.

## References

1. Auchmuty, G. (2001), Variational principles for self-adjoint elliptic eigenproblems. In: Gao, D.Y., Ogden, R.W. and Stavroulakis, G. (eds.), *Nonconvex/Nonsmooth Mechanics: Modelling, Methods and Algorithms*, Kluwer Academic Publishers, 2000, p. 478.
2. Benson, H. (1995), Concave minimization: theory, applications and algorithms. In: Horst, R. and Pardalos, P. (eds.), *Handbook of Global Optimization*, Kluwer Academic Publishers, pp. 43–148.
3. Ekeland, I. (1977), Legendre duality in nonconvex optimization and calculus of variations, *SIAM J. Control and Optimization*, 15, 905–934.

4. Ekeland, I. (2003), Nonconvex duality. In: Gao, D.Y. (ed.), *Proceedings of IUTAM Symposium on Duality, Complementarity and Symmetry in Nonlinear Mechanics*, Kluwer Academic Publishers, Dordrecht/Boston/London, to appear.
5. Ekeland, I. and Temam, R. (1976), *Convex Analysis and Variational Problems*, North-Holland.
6. Floudas, C.A. and Visweswaran, V. (1995), Quadratic optimization. In: Horst, R. and Pardalos, P.M. (eds.), *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht/Boston/London, pp. 217–270.
7. Gao, D.Y. (1997), Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory, *Applied Mechanics Reviews*, 50(11), November 1997, S64–S71.
8. Gao, D.Y. (1998), Duality, triality and complementary extremum principles in nonconvex parametric variational problems with applications, *IMA J. Appl. Math.*, 61, 199–235.
9. Gao, D.Y. (1999), Duality-Mathematics, *Wiley Encyclopedia of Electronical and Electronical Engineering*, 6, 68–77.
10. Gao, D.Y. (1999), General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics, *Meccanica*, 34, 169–198.
11. Gao, D.Y. (2000), *Duality Principles in Nonconvex Systems: Theory, Methods and Applications*, Kluwer Academic Publishers, Dordrecht/Boston/London, pp. xviii+454.
12. Gao, D.Y. (2000), Analytic solution and triality theory for nonconvex and nonsmooth variational problems with applications, *Nonlinear Analysis*, 42(7), 1161–1193.
13. Gao, D.Y. (2000), Canonical dual transformation method and generalized triality theory in nonsmooth global optimization, *J. Global Optimization*, 17(1/4), 127–160.
14. Gao, D.Y. (2001), Complementarity, polarity and triality in nonsmooth, nonconvex and nonconservative Hamilton systems, *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences*, 359, 2347–2367.
15. Gao, D.Y. (2002), Duality and triality in non-smooth, nonconvex and nonconservative systems: A survey, new phenomena and new results. In: Baniotopoulos, C. (ed.), *Nonsmooth/Nonconvex Mechanics with Applications in Engineering*, Thessaloniki, Greece, pp. 1–14.
16. Gao, D.Y. (2003a), Perfect duality theory and complete solutions to a class of global optimization problems, to be published in *Optimisation*, special issue edited by A. Rubinov.
17. Gao, D.Y. (2003b), Canonical dual principle, algorithm, and complete solutions to Landau-Ginzburg equation with applications, *Journal of Mathematics and Mechanics of Solids*, special issue dedicated to Professor Ray Ogden for the occasion of his 60th birthday, edited by D. Steigmann.
18. Gao, D.Y. and Lin, P. (2002), Calculating global minimizers of a nonconvex energy potential. In: Lee, H.P. and Kumar, K. (eds.), *Recent Advances in Computational Science Engineering*, Imperial College Press, pp. 696–700.
19. Gao, D.Y. and Strang, G. (1989), Geometric nonlinearity: potential energy, complementary energy, and the gap function, *Quart. Appl. Math.*, 47(3), 487–504.
20. Gasimov, R.N. (2002), Augmented Lagrangian duality and nondifferentiable optimization methods in nonconvex programming, *J. Global Optimization*, 24, 187–203.
21. Goh, C.J. and Yang, X.Q. (2002), *Duality in Optimization and Variational Inequalities*, Taylor and Francis, p. 329.
22. Karmarkar, N. (1990), An interior-point approach to NP-complete problems. I, *Comtemporary Mathematics*, 114, 297–308.
23. Murty, K.G. and Kabadi, S.N. (1987), Some NP-complete problems in quadratic and nonlinear programmings, *Math. Progr.*, 39, 117–129.
24. Horst, R., Pardalos, P.M. and Nguyen Van Thoai (2000), *Introduction to Global Optimization*, Kluwer Academic Publishers.
25. Pardalos, P.M. (1991), Global optimization algorithms for linearly constrained indefinite quadratic problems, *Comput. Math. Appl.*, 21, 87–97.

26. Pardalos, P.M. and Vavasis, S.A. (1991), Quadratic programming with one negative eigenvalue is NP-hard, *J. Global Optimization*, 21, 843–855.
27. Powell, M.J.D. (2002), UOBYQA: unconstrained optimization by quadratic approximation, *Mathematical Programming, Series B*, 92(3), 555–582.
28. Rubinov, A.M. and Gasimov, R.N. (2003), Scalarization and nonlinear scalar duality for vector optimization with preferences that are not necessarily a pre-order relation, *J. Global Optimization* (special issue on Duality edited by D.Y. Gao and K.L. Teo), to appear.
29. Rubinov, A.M. and Yang, X.Q. (2003), *Lagrange-Type Functions in Constrained Non-Convex Optimization*, Kluwer Academic Publishers, Boston/Dordrecht/London, p. 285.
30. Rubinov, A.M., Yang, X.Q. and Glover, B.M. (2001), Extended Lagrange and penalty functions in optimization, *J. Optim. Theory Appl.*, 111(2), 381–405.
31. Sewell, M.J. (1987), *Maximum and Minimum Principles*, Cambridge Univ. Press, p. 468.
32. Sahni, S. (1974), Computationally related problems, *SIAM J. Comp.*, 3, 262–279.
33. Singer, I. (1998), Duality for optimization and best approximation over finite intersections, *Numer. Funct. Anal. Optim.*, 19(7–8), 903–915.
34. Strang, G. (1986), *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, p. 758.
35. Thach, P.T. (1993), Global optimality criterion and a duality with a zero gap in non-convex optimization, *SIAM J. Math. Anal.*, 24(6), 1537–1556.
36. Thach, P.T. (1995), Diewert-Crouzeix conjugation for general quasiconvex duality and applications, *J. Optim. Theory Appl.*, 86(3), 719–743.
37. Thach, P.T., Konno, H. and Yokota, D. (1996), Dual approach to minimization on the set of Pareto-optimal solutions, *J. Optim. Theory Appl.*, 88(3), 689–707.
38. Toland, J.F. (1978), Duality in nonconvex optimization, *J. Mathematical Analysis and Applications*, 66, 399–415.
39. Tuy, H. (1995), D.C. optimization: theory, methods and algorithms. In: Horst, R. and Pardalos, P. (eds.), *Handbook of Global Optimization*, Kluwer Academic Publishers, pp. 149–216.
40. Vavasis, S. (1990), Quadratic programming is in NP, *Info. Proc. Lett.*, 36, 73–77.
41. Vavasis, S. (1991), *Nonlinear Optimization: Complexity Issues*, Oxford University press, New York.
42. Walk, M. (1989), *Theory of Duality in Mathematical Programming*, Springer-Verlag, Wien/New York.
43. Wright, M.H. (1998), The interior-point revolution in constrained optimization. In: DeLeone, R., Murli, A., Pardalos, P.M. and Toraldo, G. (eds.), *High-Performance Algorithms and Software in Nonlinear Optimization*, Kluwer Academic Publishers, Dordrecht, pp. 359–381.
44. Ye, Y.Y. (1992), On affine scalling algorithms for nonconvex quadratic programming, *Mathematical Programming*, 56, 285–300.