A New Self-Dual Embedding Method for Convex Programming

Shuzhong Zhang *

October 26, 2001

Abstract

In this paper we introduce a conic optimization formulation for inequality-constrained convex programming, and propose a self-dual embedding model for solving the resulting conic optimization problem. The primal and dual cones in this formulation are characterized by the original constraint functions and their corresponding conjugate functions respectively. Hence they are completely symmetric. This allows for a standard primal-dual path following approach for solving the embedded problem. Moreover, there are two immediate logarithmic barrier functions for the primal and dual cones. We show that these two logarithmic barrier functions are conjugate to each other. The explicit form of the conjugate functions are in fact not required to be known in the algorithm. An advantage of the new approach is that there is no need to assume an initial feasible solution to start with. To guarantee the polynomiality of the path-following procedure, we may apply the self-concordant barrier theory of Nesterov and Nemirovski. For this purpose, as one application, we prove that the barrier functions constructed this way are indeed self-concordant when the original constraint functions are convex and quadratic.

Keywords: Convex Programming, Convex Cones, Self-Dual Embedding, Self-Concordant Barrier Functions.

*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Research supported by Hong Kong RGC Earmarked Grants CUHK4181/00E and CUHK4233/01E.

1 Introduction

In this paper we propose to solve the constrained convex optimization problems by means of conic self-dual embedding.

The original self-dual embedding method was proposed by Ye, Todd and Mizuno [16] for linear programming. The advantage of this method is twofold. First, it has a strong theoretical appeal, since it displays, and makes use of, the symmetricity of the primal-dual relationship in linear programming. The merit of the symmetric duality becomes explicit, and it is especially well suited for treating infeasible or unbounded problems, as the Farkas type certificate is readily available after solving the self-dual embedded model. Second, in combination of any efficient interior point implementation, the self-dual embedded model is handy to use in practice. As a result, many successful software packages for linear programming are based on this model.

The idea of self-dual embedding was extended to solve more general constrained convex optimization problems in two different ways. For conically constrained convex optimization, including *semidefinite programming*, Luo, Sturm and Zhang [7] proposed a self-dual embedding model; for more details and an overview, see Section 2. The software package of Jos Sturm, SeDuMi, uses the self-dual embedding model for symmetric cone programming. For conventional convex programming with inequality constraints, Andersen and Ye [1, 2] developed a different type of self-dual embedding model based on the simplified model of Xu, Hung and Ye [14] for linear programming. In fact, the method of Andersen and Ye is designed for nonlinear complementarity problems, thus more general. However, it is not exactly a self-dual embedding model due to the simplification made in Xu, Hung and Ye's approach, [14]. As it is well known, for linear programming this simplification does not affect the iterates, but it will make a difference for nonlinear problems. Both codes of Sturm and Andersen are considered the state-of-the-art implementation of interior point method for convex programming.

In this paper we introduce in Section 3 a particular conic formulation for inequality-constrained convex programming. In other words, we reformulate the inequality constraints by a conic constraint. This involves two additional variables. By standard conic duality we also obtain a dual form for the problem. By virtue of this construction, the dual of the cone is completely characterized by the conjugate of the constraint functions. An advantage of this dual form is that it is completely symmetric with respect to the primal problem. Moreover, the barrier functions for the so constructed primal and dual cones are readily available. We show that these natural (logarithmic) barrier functions for the primal and the dual cones are simply conjugate to each other. This exhibits a beautiful symmetricity of duality. After formulating the optimization problem in the conic form and employing the self-dual embedding technique as developed in Section 2, we are in the position to invoke the central path-following interior point method. In order to stay within the polynomial-time complexity realm, we will need to rely on the self-concordant barrier function theory developed by Nesterov and Nemirovski in [8]. For that purpose, as an example, we prove in Section 4 that if all the constraints in the original problem are convex quadratic functions, then the barrier function for the self-dual embedded problem is self-concordant. Hence this class of problems can be solved in $O(\sqrt{m}\log\frac{1}{\epsilon})$ number of iterations, where m is the number of constraints, and $\epsilon > 0$ is the required precision. Certainly, an obvious advantage of the new approach is that it does not require an initial feasible solution of the convex program to start with, which is a generic virtue of the self-dual embedding method.

2 Self-Dual Embedded Conic Optimization

The primal-dual self-dual embedding technique was first proposed by Ye, Todd and Mizuno [16] for solving linear programming problems. The advantage of the method is that the model allows to take any pre-described interior points as initial feasible solutions for the embedded problem, and the embedded problem is guaranteed to have an optimal solution, which can be approximated by using any interior point algorithm. Moreover, by solving the embedded problem, one either obtains an optimal solution for the original problem, or obtains a Farkas type certificate to conclude that the original problem is unsolvable. This technique was independently extended by Luo, Sturm and Zhang [7], De Klerk, Roos and Terlaky [5], and Potra and Sheng [10] to solve semidefinite programming. In fact, the extension of Luo, Sturm and Zhang [7] allowed for a more general conic optimization framework. In this section we shall briefly introduce this method.

Consider the following conic optimization problem

$$\begin{array}{lll} (P) & \text{minimize} & c^T x \\ & \text{subject to} & Ax = b \\ & & x \in \mathcal{K}, \end{array}$$

where $c \in \Re^n$, $b \in \Re^m$, $A \in \Re^{m\times n}$ (assumed to have full row-rank), and $\mathcal{K} \subseteq \Re^n$ is a solid convex cone with its dual cone defined as

$$\mathcal{K}^* = \{s \mid x^T s \geq 0 \text{ for all } x \in \mathcal{K}\}.$$

The problem (P) has an associated dual problem, called (D),

$$\begin{array}{lll}(D) & \text{maximize} & b^T y \\ & \text{subject to} & A^T y + s = c \\ & & s \in \mathcal{K}^*.\end{array}$$

The duality pair, (P) and (D), enjoys a nice symmetric relationship, similar as in the case of linear programming where $\mathcal{K} = \mathcal{K}^* = \Re^n_+$. For a detailed account on the subject, one is referred to either [6] or [13].

Take any $x^0 \in \text{int } \mathcal{K}$, $s^0 \in \text{int } \mathcal{K}^*$, and $y^0 \in \Re^m$. Moreover, define

$$r_p = b - Ax^0,\ r_d = s^0 - c + A^T y^0,\ \text{and } r_g = 1 + c^T x^0 - b^T y^0.$$

Consider the following embedded optimization model

$$\begin{array}{lcccccccc}(SD) \quad \text{minimze} & & & & \beta\theta & & & & \\ \quad \text{subject to} & & Ax & -b\tau & +r_p\theta & & & = & 0 \\ & -A^T y & & +c\tau & +r_d\theta & -s & & = & 0 \\ & b^T y & -c^T x & & +r_g\theta & & -\kappa & = & 0 \\ & -r_p^T y & -r_d^T x & -r_g\tau & & & & = & -\beta \\ & & x \in \mathcal{K}, & \tau \geq 0, & & s \in \mathcal{K}^*, & \kappa \geq 0, & & \end{array}$$

where $\beta = 1 + (x^0)^T s^0 > 1$, and the decision variables are $(y, x, \tau, \theta, s, \kappa)$.

It is elementary to verify that (SD) is *self-dual*, i.e., its dual form coincides with itself. Moreover, (SD) admits a trivial solution

$$(y, x, \tau, \theta, s, \kappa) = (y^0, x^0, 1, 1, s^0, 1)$$

which lies in the interior of the constraint cone

$$\Re^m \times \mathcal{K} \times \Re_+ \times \Re \times \mathcal{K}^* \times \Re_+.$$

This implies that (SD) satisfies the Slater condition. Due to the self-duality, so is true for the dual problem. Hence, (SD) has a non-empty and bounded optimal solution set; see e.g. Nesterov and Nemirovski [8]. Note that for the case when $\mathcal{K}$ is the cone of positive semidefinite matrices, the above self-dual embedding scheme is exactly what is proposed by De Klerk, Roos and Terlaky [5] and Potra and Sheng [10].

The following result is well known; see its analog in [16] for the linear programming case.

Proposition 2.1 *The problem* (SD) *has a maximally complementary optimal solution, denoted by* $(y^*, x^*, \tau^*, \theta^*, s^*, \kappa^*)$, *such that* $\theta^* = 0$ *and* $(x^*)^T s^* + \tau^* \kappa^* = 0$. *Moreover, if* $\tau^* > 0$, *then* x^*/τ^* *is an optimal solution for* (P), *and* $(y^*/\tau^*, s^*/\tau^*)$ *is an optimal solution for* (D). *If* $\kappa^* > 0$ *then either* $c^T x^* < 0$ *or* $b^T y^* > 0$*; in the former case* (D) *is infeasible, and in the latter case* (P) *is infeasible.*

If τ^* and κ^* do not exhibit strict complementarity, namely $\tau^* = \kappa^* = 0$, then in that case we can only conclude that (P) and (D) do not have complementary optimal solutions. In fact, more information can still be deduced, using the notion of, e.g., *weak infeasibility*; for more details see [7].

A barrier (convex) function $F(x)$ for $\mathcal{K}$ is defined to have the property that $F(x) < \infty$ for all $x \in \text{int } \mathcal{K}$ and $F(x^k) \to \infty$ as $x^k \to x$ where x is on the boundary of $\mathcal{K}$. Moreover, it is called *self-concordant* (Section 2.3.1 of [8]) if it further satisfies the property that

$$|\nabla^3 F(x)[h,h,h]| \le 2(\nabla^2 F(x)[h,h])^{3/2} \tag{1}$$

and

$$|\nabla F(x)[h]| \le C(\nabla^2 F(x)[h,h])^{1/2} \tag{2}$$

for any $x \in \text{int } \mathcal{K}$ and any direction $h \in \Re^n$.

Furthermore, we call a barrier function $F(x)$ to be ν-logarithmically homogeneous if

$$F(tx) = F(x) - \nu \log t$$

for all $x \in \text{int } \mathcal{K}$ and $t > 0$. As a fundamental property for convex cones, Nesterov and Nemirovski proved the following important theorem (Section 2.5 of [8]):

Theorem 2.2 *Any closed convex cone admits a self-concordant, logarithmically homogeneous barrier function.*

The following straightforward, but usefully, properties of the ν-logarithmically homogeneous function can be found in Nesterov and Nemirovski [8]; see also [9].

Proposition 2.3 *Suppose that* $F(x)$ *is an* ν*-logarithmically homogeneous barrier function for* $\mathcal{K}$. *Then the following identities hold where* $x \in \text{int } \mathcal{K}$ *and* $t > 0$*:*

$$\nabla F(tx) = \frac{1}{t} \nabla F(x); \tag{3}$$

$$\nabla^2 F(tx) = \frac{1}{t^2} \nabla^2 F(x); \tag{4}$$

$$\nabla^2 F(x) x = -\nabla F(x); \tag{5}$$

$$(\nabla F(x))^T x = -\nu. \tag{6}$$

Related to the duality of convex cones, the *conjugate* of the convex function $f(x)$, is defined as

$$f^*(s) = \sup\{(-s)^T x - f(x) \mid x \in \text{dom } f\},$$

where dom f stands for the domain of the function f. The above operation is known as the *Legendre-Fenchel transformation*. The conjugate of several popularly used functions are well studied. For instance, for $f(x) = \frac{1}{2}x^T Qx + b^T x$ where $Q \succ 0$, we have $f^*(s) = \frac{1}{2}(s+b)^T Q^{-1}(s+b)$. If $f(x)$ is strictly convex and differentiable, then $f^*(s)$ is also strictly convex and differentiable. Moreover, it is easy to see that for $x \in \text{int dom } f$ and $s \in \text{int dom } f^*$ the following three statements are equivalent

$$\begin{aligned} s &= -\nabla f(x) && (7)\\ x &= -\nabla f^*(s) && (8)\\ -x^T s &= f(x) + f^*(s). && (9) \end{aligned}$$

The famous bi-conjugate theorem asserts (see e.g. Rockafellar [12]) that $f^{**} = \text{cl } f$. In particular, for the convex barrier function $F(x)$, we simply have $F^{**}(x) = F(x)$. In addition to that, Nesterov and Nemirovski [8] showed that if $F(x)$ is a self-concordant ν-logarithmically homogeneous barrier function for $\mathcal{K}$, then it follows that $F^*(s)$ is a self-concordant ν-logarithmically homogeneous barrier function for $\mathcal{K}^*$.

We consider the following barrier approach for solving (SD) with $\mu > 0$ as the barrier parameter

$$\begin{array}{llccccccccc}
(SD_\mu) & \text{minimze} & & \mu F(x) & -\mu\log\tau & +\beta\theta & +\mu F^*(s) & -\mu\log\kappa & & \\
& \text{subject to} & & Ax & -b\tau & +r_p\theta & & & = & 0 \\
& & -A^T y & & +c\tau & +r_d\theta & -s & & = & 0 \\
& & b^T y & -c^T x & & +r_g\theta & & -\kappa & = & 0 \\
& & -r_p^T y & -r_d^T x & -r_g\tau & & & & = & -\beta \\
& & & x \in \mathcal{K}, & \tau \ge 0, & & s \in \mathcal{K}^*, & \kappa \ge 0. & &
\end{array}$$

Due to the self-duality we derive the following KKT optimality condition for (SD_μ), where the solution is denoted by $(y(\mu), x(\mu), \tau(\mu), \theta(\mu), s(\mu), \kappa(\mu))$,

$$\left.\begin{array}{ccccccccc}
 & Ax(\mu) & -b\tau(\mu) & +r_p\theta(\mu) & & & = & 0 \\
-A^T y(\mu) & & +c\tau(\mu) & +r_d\theta(\mu) & -s(\mu) & & = & 0 \\
b^T y(\mu) & -c^T x(\mu) & & +r_g\theta(\mu) & & -\kappa(\mu) & = & 0 \\
-r_p^T y(\mu) & -r_d^T x(\mu) & -r_g\tau(\mu) & & & & = & -\beta \\
 & & -\mu\frac{1}{\tau(\mu)} & & & & = & -\kappa(\mu) \\
 & & & & & -\mu\frac{1}{\kappa(\mu)} & = & -\tau(\mu) \\
 & \mu\nabla F(x(\mu)) & & & & & = & -s(\mu) \\
 & & & & \mu\nabla F^*(s(\mu)) & & = & -x(\mu).
\end{array}\right\}$$

By (7) and (8) we can simplify the above optimality condition into

$$\left.\begin{array}{rrrrrrcl}
 & Ax(\mu) & -b\tau(\mu) & +r_p\theta(\mu) & & & = & 0 \\
-A^T y(\mu) & & +c\tau(\mu) & +r_d\theta(\mu) & -s(\mu) & & = & 0 \\
b^T y(\mu) & -c^T x(\mu) & & +r_g\theta(\mu) & & -\kappa(\mu) & = & 0 \\
-r_p^T y(\mu) & -r_d^T x(\mu) & -r_g\tau(\mu) & & & & = & -\beta \\
 & & \tau(\mu)\kappa(\mu) & & & & = & \mu \\
 & & & & s(\mu) & & = & -\mu\nabla F(x(\mu)).
\end{array}\right\} \tag{10}$$

The last equation in (10) can be equivalently replaced by

$$x(\mu) = -\mu\nabla F^*(s(\mu)).$$

3 Connection with Convex Programming

We now consider a standard convex programming problem

$$\begin{array}{lll}
(CP) & \text{minimize} & c^T x \\
 & \text{subject to} & Ax = b \\
 & & f_i(x) \le 0,\ i = 1, ..., m,
\end{array}$$

where $f_i(x)$ is smooth and convex, $i = 1, ..., m$.

For simplicity, let us start with the case where $m = 1$. Let $f(x) = f_1(x)$. Let the decision variable now be

$$\bar{x} := \begin{bmatrix} p \\ q \\ x \end{bmatrix} \in \Re^1 \times \Re^1 \times \Re^n,$$

and the problem data as

$$\bar{c} := \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix} \in \Re^1 \times \Re^1 \times \Re^n,\ \bar{b} := \begin{bmatrix} 1 \\ 0 \\ b \end{bmatrix} \in \Re^1 \times \Re^1 \times \Re^m \tag{11}$$

and

$$\bar{A} := \begin{bmatrix} 1 & 0 & 0^T \\ 0 & 1 & 0^T \\ 0 & 0 & A \end{bmatrix} \in \Re^{(m+2)\times(n+2)}. \tag{12}$$

Let

$$\mathcal{K} = \text{cl}\ \{\bar{x} \mid p > 0,\ q - pf(x/p) \ge 0\} \subseteq \Re^{n+2}, \tag{13}$$

which is a closed cone. The lemma below, which was used by Ye in [15], shows that it is also convex. For completeness, we provide a proof here as well.

Lemma 3.1 *The function $-q + pf(x/p)$ is convex in $\Re^1_{++} \times \Re^1 \times \Re^n$.*

Proof. We need only to show that $pf(x/p)$ is convex in $\Re^1_{++} \times \Re^n$. Simply calculation shows that

$$\nabla^2(pf(x/p)) = \frac{1}{p}\begin{bmatrix} (x/p)^T\nabla^2 f(x/p)(x/p) & -(x/p)^T\nabla^2 f(x/p) \\ -\nabla^2 f(x/p)(x/p) & \nabla^2 f(x/p) \end{bmatrix}.$$

Let $H = \nabla^2 f(x/p)$ and $h = x/p$. Then for any $\bar{\xi}^T = (\xi_0, \xi^T) \in \Re^{n+1}$ we have

$$\begin{aligned} \bar{\xi}^T\nabla^2(pf(x/p))\bar{\xi} &= \frac{1}{p}\left[\xi_0^2 h^T H h - 2\xi_0\xi^T H h + \xi^T H \xi\right] \\ &\geq \frac{1}{p}\left(\xi_0\|H^{1/2}h\| - \|H^{1/2}\xi\|\right)^2 \\ &\geq 0. \end{aligned}$$

Therefore, $-q + pf(x/p)$ is convex in $\Re^1_{++} \times \Re^1 \times \Re^n$.

Q.E.D.

An equivalent conic formulation for (CP) is given by

$$\begin{array}{lll} (CCP) & \text{minimize} & \bar{c}^T\bar{x} \\ & \text{subject to} & \bar{A}\bar{x} = \bar{b} \\ & & \bar{x} \in \mathcal{K}. \end{array}$$

Naturally, a 2-logarithmically homogeneous and convex barrier function for $\mathcal{K}$ is

$$F(\bar{x}) = -\log p - \log(q - pf(x/p)). \tag{14}$$

Theorem 3.2 *It holds that*

$$\mathcal{K}^* = \text{cl}\left\{\bar{s} = \begin{bmatrix} u \\ v \\ s \end{bmatrix} \middle| \, v > 0,\, u - vf^*(s/v) \geq 0\right\}$$

and

$$F^*(\bar{s}) = -\log v - \log(u - vf^*(s/v)),$$

which is a 2-logarithmically homogeneous barrier function for $\mathcal{K}^*$.

Proof. For any $\begin{bmatrix} u \\ v \\ s \end{bmatrix}$ with $v > 0$ and $u - vf^*(s/v) \geq 0$, and $\bar{x} = \begin{bmatrix} p \\ q \\ x \end{bmatrix} \in \mathcal{K}$ we have

$$\begin{aligned} pu + qv + x^T s &= pv\left[u/v + q/p + (x/p)^T(s/v)\right] \\ &\geq pv\left[u/v + q/p - f(x/p) - f^*(s/v)\right] \\ &= v\left(q - pf(x/p)\right) + p\left(u - vf^*(s/v)\right) \\ &\geq 0. \end{aligned}$$

Hence,

$$\left\{ \bar{s} = \begin{bmatrix} u \\ v \\ s \end{bmatrix} \middle| \, v > 0,\, u - vf^*(s/v) \geq 0 \right\} \subseteq \mathcal{K}^*.$$

On the other hand, take any $\bar{s} = \begin{bmatrix} u \\ v \\ s \end{bmatrix} \in \text{int}\, \mathcal{K}^*$. Obviously $v \geq 0$. This is because for any fixed $\bar{x} \in \mathcal{K}$ it follows that $\bar{x} + t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \in \mathcal{K}$ for any $t > 0$. Since $\text{int}\,\mathcal{K}^*$ is open, we conclude that $v > 0$. Let $\hat{x} = -\nabla f^*(s/v)$. Consider

$$\begin{bmatrix} 1 \\ f(\hat{x}) \\ \hat{x} \end{bmatrix} \in \mathcal{K}.$$

By the equivalence between (8) and (9), we have $f(\hat{x}) + f^*(s/v) = -\hat{x}^T(s/v)$ and so

$$0 \leq \bar{s}^T \begin{bmatrix} 1 \\ f(\hat{x}) \\ \hat{x} \end{bmatrix} = u - vf^*(s/v).$$

This shows that

$$\mathcal{K}^* = \text{cl}\left\{ \bar{s} = \begin{bmatrix} u \\ v \\ s \end{bmatrix} \middle| \, v > 0,\, u - vf^*(s/v) \geq 0 \right\}.$$

To show that $F^*(\bar{s}) = -\log v - \log(-u - vf^*(s/v))$, we observe that

$$\nabla F(\bar{x}) \;=\; \begin{bmatrix} -1/p - \left[-f(x/p) + \nabla f(x/p)^T(x/p)\right]/(q - pf(x/p)) \\ -1/(q - pf(x/p)) \\ \nabla f(x/p)/(q - pf(x/p)) \end{bmatrix}$$

$$= \begin{bmatrix} -\left[q/p - 2f(x/p) + \nabla f(x/p)^T(x/p)\right]/(q - pf(x/p)) \\ -1/(q - pf(x/p)) \\ \nabla f(x/p)/(q - pf(x/p)) \end{bmatrix}.$$

Let $\bar{s} = \begin{bmatrix} u \\ v \\ s \end{bmatrix} = -\nabla F(\bar{x})$. Hence

$$\left.\begin{aligned} u &= \left[q/p - 2f(x/p) + \nabla f(x/p)^T(x/p)\right]/(q - pf(x/p)), \\ v &= 1/(q - pf(x/p)), \\ s/v &= -\nabla f(x/p). \end{aligned}\right\} \qquad (15)$$

From the last equation, using (7), (8) and (9) we obtain $x/p = -\nabla f^*(s/v)$, and $f(x/p) + f^*(s/v) = -(x/p)^T(s/v)$. Moreover, $F(\bar{x})$ is 2-logarithmically homogeneous, implying that

$$\bar{x}^T\bar{s} = up + vq + x^Ts = 2.$$

Hence,

$$\begin{aligned} 1 &= vq - vpf(x/p) = vq - pv\left(-(s/v)^T(x/p) - f^*(s/v)\right) \\ &= vq + s^Tx + pvf^*(s/v) = 2 - up + pvf^*(s/v) \end{aligned}$$

and so

$$p = 1/(u - vf^*(s/v)).$$

Using the first equation in (15) we have

$$\begin{aligned} u &= v\left[q/p - 2f(x/p) - (s/v)^T(x/p)\right] \\ &= v\left[q/p - f(x/p) + f^*(s/v)\right] \\ &= v\left[q/p + 2f^*(s/v) - \nabla f^*(s/v)^T(s/v)\right] \end{aligned}$$

and so

$$q/p = u/v - 2f^*(s/v) + \nabla f^*(s/v)^T(s/v).$$

This yields

$$\begin{aligned} \nabla F^*(\bar{s}) &= -\bar{x} = -\begin{bmatrix} p \\ q \\ x \end{bmatrix} \\ &= \begin{bmatrix} -1/(u - vf^*(s/v)) \\ \left[u/v - 2f^*(s/v) + \nabla f^*(s/v)^T(s/v)\right]/(u - vf^*(s/v)) \\ \nabla f^*(s/v)/(u - vf^*(s/v)) \end{bmatrix}. \end{aligned}$$

Consequently,

$$F^*(\bar{s}) = -\log v - \log(u - vf^*(s/v)).$$

Q.E.D.

(I like to thank Jan Brinkhuis of Erasmus University for pointing out to me the form of the dual cone $\mathcal{K}^*$ as described in the above theorem, during a private conversation [4].)

Now we consider the conic form self-dual embedding path following scheme as stipulated by Equation (10) for (CCP), where the data of the problem, $(\bar{A}, \bar{b}, \bar{c})$, is given by (11) and (12), and the barrier function is given according to (14). This results in the following system of equations

$$\left.\begin{array}{rrrrrrcl}
 & \bar{A}\bar{x}(\mu) & -\bar{b}\tau(\mu) & +\bar{r}_p\theta(\mu) & & & = & 0 \\
-\bar{A}^T\bar{y}(\mu) & & +\bar{c}\tau(\mu) & +\bar{r}_d\theta(\mu) & -\bar{s}(\mu) & & = & 0 \\
\bar{b}^T\bar{y}(\mu) & -\bar{c}^T\bar{x}(\mu) & & +\bar{r}_g\theta(\mu) & & -\kappa(\mu) & = & 0 \\
-\bar{r}_p^T\bar{y}(\mu) & -\bar{r}_d^T\bar{x}(\mu) & -\bar{r}_g\tau(\mu) & & & & = & -\beta
\end{array}\right.$$
$$\left.\begin{array}{rcl}
\tau(\mu)\kappa(\mu) & = & \mu \\
u(\mu)\left[q(\mu) - p(\mu)f(x(\mu)/p(\mu))\right] & = & \mu\left[q(\mu)/p(\mu) - 2f(x(\mu)/p(\mu)) + \nabla f(x(\mu)/p(\mu))\right] \\
v(\mu)\left[q(\mu) - p(\mu)f(x(\mu)/p(\mu))\right] & = & \mu \\
s(\mu)\left[q(\mu) - p(\mu)f(x(\mu)/p(\mu))\right] & = & -\mu\nabla f(x(\mu)/p(\mu)).
\end{array}\right\} \quad (16)$$

We remark here that whenever the initial p and τ are set to be 1, the first component of $\bar{r}_p$ will be zero. Hence $p = \tau$. This implies that these two normalizing variables can be combined into one.

Next we consider the general formulation of (CP) where $m \geq 1$. Similarly we have its conic representation, (CCP), with

$$\mathcal{K} = \bigcap_{i=1}^m \mathcal{K}_i,$$

where

$$\mathcal{K}_i = \text{cl } \{\bar{x} \mid p > 0,\ q - pf_i(x/p) \geq 0\} \subseteq \Re^{n+2},\ i = 1, ..., m.$$

The natural $2m$-logarithmically homogeneous barrier function for $\mathcal{K}$ is

$$F(\bar{x}) = -m\log p - \sum_{i=1}^m \log(q - pf_i(x/p)).$$

The dual cone of $\mathcal{K}$ is

$$\mathcal{K}^* = \text{cl } (\mathcal{K}_1^* \oplus \cdots \oplus \mathcal{K}_m^*) = \text{cl } \left\{ \sum_{i=1}^m \bar{s}_i = \sum_{i=1}^m \begin{bmatrix} u_i \\ v_i \\ s_i \end{bmatrix} \middle|\ v_i > 0,\ u_i - v_i f_i^*(s_i/v_i) \geq 0,\ i = 1, ..., m \right\}.$$

Using Theorem 3.2, we know that the dual barrier function for $\mathcal{K}^*$, which is the conjugate function of $F(\bar{x})$, hence also $2m$-logarithmically homogeneous, is given as follows

$$F^*(\bar{s}_1, \cdots, \bar{s}_m) = -\sum_{i=1}^{m} \left[\log v_i + \log(u_i - v_i f_i^*(s_i/v_i))\right].$$

The central path for the embedded problem is:

$$\left.\begin{array}{rrrrrrl}
 & \bar{A}\bar{x}(\mu) & -\bar{b}\tau(\mu) & +\bar{r}_p\theta(\mu) & & & = \quad 0 \\
-\bar{A}^T\bar{y}(\mu) & & +\bar{c}\tau(\mu) & +\bar{r}_d\theta(\mu) & -\bar{s}(\mu) & & = \quad 0 \\
\bar{b}^T\bar{y}(\mu) & -\bar{c}^T\bar{x}(\mu) & & +\bar{r}_g\theta(\mu) & & -\kappa(\mu) & = \quad 0 \\
-\bar{r}_p^T\bar{y}(\mu) & -\bar{r}_d^T\bar{x}(\mu) & -\bar{r}_g\tau(\mu) & & & & = \; -\beta \\
\multicolumn{3}{r}{\bar{s}(\mu) - \sum_{i=1}^m \bar{s}_i(\mu) \;=\; 0} \\
\multicolumn{3}{r}{\tau(\mu)\kappa(\mu) \;=\; \mu} \\
\multicolumn{7}{l}{u_i(\mu)\left[q(\mu) - p(\mu)f_i(x(\mu)/p(\mu))\right] = \mu\left[q(\mu)/p(\mu) - 2f_i(x(\mu)/p(\mu)) + \nabla f_i(x(\mu)/p(\mu))\right]} \\
\multicolumn{7}{l}{v_i(\mu)\left[q(\mu) - p(\mu)f_i(x(\mu)/p(\mu))\right] = \mu} \\
\multicolumn{7}{l}{s_i(\mu)\left[q(\mu) - p(\mu)f_i(x(\mu)/p(\mu))\right] = -\mu\nabla f_i(x(\mu)/p(\mu)),\; i = 1, ..., m.}
\end{array}\right\} \quad (17)$$

Observe that there is no need to explicitly involve the conjugate function f_i^* in the above primal-dual framework. However, if the conjugate functions f_i^*, $i = 1, ..., m$, are indeed available, then one may consider applying the standard path-following procedure for the following embedded barrier problem

$$\begin{array}{llrrrrrrl}
(PF_\mu) & \min & & \mu F(\bar{x}) & -\mu\log\tau & +\beta\theta & +\mu F^*(\bar{s}_1, \cdots, \bar{s}_m) & -\mu\log\kappa & \\
 & \text{s.t.} & & & & & & & \\
 & & & \bar{A}\bar{x} & -\bar{b}\tau & +\bar{r}_p\theta & & & = \quad 0 \\
 & & -\bar{A}^T\bar{y} & & +\bar{c}\tau & +\bar{r}_d\theta & -\sum_{i=1}^m \bar{s}_i & & = \quad 0 \\
 & & \bar{b}^T\bar{y} & -\bar{c}^T\bar{x} & & +\bar{r}_g\theta & & -\kappa & = \quad 0 \\
 & & -\bar{r}_p^T\bar{y} & -\bar{r}_d^T\bar{x} & -\bar{r}_g\tau & & & & = \; -\beta.
\end{array}$$

According to Nesterov and Nemirovski's self-concordant barrier theory, as long as $F(\bar{x})$ is self-concordant in the feasible set, applying the standard path following method to the above problem (PF_μ) will lead to a polynomial-time algorithm.

Since the above formulation is based on the conic representation, it is therefore easy to incorporate any other additional genuinely conic constraints. Consider a hybrid type convex optimization

$$\begin{array}{lll}
(H) & \text{minimize} & c^T x \\
 & \text{subject to} & Ax = b \\
 & & f_i(x) \leq 0,\; i = 1, ..., m, \\
 & & x \in \mathcal{C},
\end{array}$$

where $\mathcal{C}$ is a closed convex cone with a known self-concordant barrier $G(x)$. Typically, $\mathcal{C}$ can be the direct product of second order cones, or the cone of positive semidefinite matrices. Let the extended barrier function be

$$G(\bar{x}) := G(x).$$

The corresponding embedded model is

$$\begin{array}{lll}(H_\mu) & \min & \mu F(\bar{x}) + \mu G(\bar{x}) - \mu \log \tau + \beta\theta + \mu F^*(\bar{s}_1, \cdots, \bar{s}_m) + \mu G^*(\bar{s}) - \mu \log \kappa \\ & \text{s.t.} & \bar{A}\bar{x} - \bar{b}\tau + \bar{r}_p\theta = 0 \\ & & -\bar{A}^T\bar{y} + \bar{c}\tau + \bar{r}_d\theta - \sum_{i=1}^m \bar{s}_i - \bar{s} = 0 \\ & & \bar{b}^T\bar{y} - \bar{c}^T\bar{x} + \bar{r}_g\theta - \kappa = 0 \\ & & -\bar{r}_p^T\bar{y} - \bar{r}_d^T\bar{x} - \bar{r}_g\tau = -\beta. \end{array}$$

Again, whenever it is convenient, one may explicitly write out the corresponding KKT system to invoke a Newton-based primal-dual central path-following algorithm.

The extension to the hybrid model being more or less straightforward, in the next section we will return to the pure convex programming formulation, and try to identify some classes of convex programming problems such that the overall barrier function is indeed self-concordant.

4 The Self-Concordant Property

The definition of a self-concordant barrier requires two conditions, (1) and (2). The constant on the right hand side of inequality (1) needs not to be 2 per se; any positive constant can be scaled to 2, and it will only affect C by a constant factor, where C is termed *the parameter of the barrier* by Nesterov and Nemirovski in [8], or *the complexity value* of the barrier function as suggested by Renegar in [11].

Just as the definition of ordinary convexity, self-concordancy is a line-property, i.e., the definition of a self-concordant function can be restricted to any line lying in the domain. To see this, let

$$d(t) := F(x + th). \tag{18}$$

Then,

$$\begin{aligned} d^{(1)}(0) &= \nabla F(x)[h] \\ d^{(2)}(0) &= \nabla^2 F(x)[h, h] \\ d^{(3)}(0) &= \nabla^3 F(x)[h, h, h]. \end{aligned}$$

Therefore, $F(x)$ is a self-concordant function satisfying (1) and (2) if and only if it is a self-concordant function restricted to any line in its domain, i.e.

$$|d^{(3)}(0)| \le 2(d^{(2)}(0))^{3/2} \text{ and } |d^{(1)}(0)| \le C(d^{(2)}(0))^{1/2} \tag{19}$$

for any given x in its domain and any given feasible direction h, where $d(t)$ is defined as in (18).

This observation allows us to prove the self-concordant property of a function by proving this property for the function restricted to an arbitrary line in its domain.

We note that for an ν-logarithmically homogeneous function, the property (2) is always satisfied. To see this we note that by (5) and (6) we have

$$\begin{aligned}
|\nabla F(x)[h]| &= |h^T \nabla F(x)| = |h^T \nabla^2 F(x) x| \\
&= |[(\nabla^2 F(x))^{1/2} h]^T [(\nabla^2 F(x))^{1/2} x]| \\
&\le \sqrt{h^T \nabla^2 F(x) h} \sqrt{x^T \nabla^2 F(x) x} \\
&= \sqrt{\nu} (\nabla^2 F(x)[h,h])^{1/2},
\end{aligned}$$

where we used the Cauchy-Schwarz inequality.

Therefore, as long as the barrier function is logarithmically homogeneous, the key inequality to be satisfied is (1). Again, we do not have to insist on the constant in (1) being 2. In fact, any universal constant in that inequality suffices, and will lead to an overall complexity value of the barrier function to be in the order of $\sqrt{\nu}$.

The function of interest in this paper is of the type

$$F(\bar{x}) = -\log p - \log(q - p f(x/p)).$$

We have shown that such a function is convex if f is convex. Now we wish to see under what conditions this function is self-concordant. Obviously, $F(\bar{x})$ is 2-logarithmically homogeneous. So we need only to be concerned with the property (1). For this purpose, let us consider the following function in $t \in \Re^1$

$$r(t) = -\log g(t)$$

in the domain $g(t) > 0$.

Simply calculation shows that

$$r^{(1)}(t) = -\frac{g^{(1)}(t)}{g(t)}$$

$$r^{(2)}(t) = -\frac{g^{(2)}(t)}{g(t)} + \frac{(g^{(1)}(t))^2}{g(t)^2}$$

$$r^{(3)}(t) = -\frac{g^{(3)}(t)}{g(t)} + \frac{3g^{(2)}(t)g^{(1)}(t)}{g(t)^2} - \frac{2(g^{(1)}(t))^3}{g(t)^3}.$$

Consequently,

$$(r^{(2)}(t))^3 = -\frac{(g^{(2)}(t))^3}{g(t)^3} + \frac{(g^{(1)}(t))^6}{g(t)^6} + \frac{3(g^{(2)}(t))^2 g^{(1)}(t)^2}{g(t)^4} - \frac{3g^{(2)}(t)g^{(1)}(t)^4}{g(t)^5}, \tag{20}$$

and

$$(r^{(3)}(t))^2 = \frac{(g^{(3)}(t))^2}{g(t)^2} + \frac{9(g^{(1)}(t))^2(g^{(2)}(t))^2}{g(t)^4} + \frac{4(g^{(1)}(t))^6}{g(t)^6} - \frac{6g^{(1)}(t)g^{(2)}(t)g^{(3)}(t)}{g(t)^3} + \frac{4g^{(3)}(t)(g^{(1)}(t))^3}{g(t)^4} - \frac{12(g^{(1)}(t))^4 g^{(2)}(t)}{g(t)^5}. \tag{21}$$

In order to satisfy (1) we wish to bound the quantity in (21) by the quantity in (20).

Lemma 4.1 *If $g(t)$ is a concave quadratic function, then $r(t)$ is convex and self-concordant.*

Proof. In that case, $g^{(2)}(t) \leq 0$ and all the third order terms in (21) disappear, leading to

$$(r^{(3)}(t))^2 = \frac{9(g^{(1)}(t))^2(g^{(2)}(t))^2}{g(t)^4} + \frac{4(g^{(1)}(t))^6}{g(t)^6} - \frac{12(g^{(1)}(t))^4 g^{(2)}(t)}{g(t)^5}.$$

Notice the second and the last term in the above expression are precisely four times the second and the last term in (20) respectively. Further notice that the first term in the above expression is three times the third term in (20). Overall, this yields,

$$(r^{(3)}(t))^2 \leq 4(r^{(2)}(t))^3,$$

and so

$$|r^{(3)}(t)| \leq 2(r^{(2)}(t))^{3/2}. \tag{22}$$

The desired property is proven.

Q.E.D.

Next we proceed to consider the barrier function $F(\bar{x})$. It turns out that this function is also self-concordant, provided that f is convex quadratic.

Theorem 4.2 *Let $f(x)$ be a convex quadratic function. Let* $\bar{x} = \begin{bmatrix} p \\ q \\ x \end{bmatrix} \in \Re^{n+2}$. *Then* $F(\bar{x}) = -\log p - \log(q - pf(x/p))$ *is a self-concordant function in the domain* $p > 0$ *and* $q - pf(x/p) > 0$.

Proof. Consider an arbitrary line in the domain of $F(\bar{x})$.

If the p component remains constant along this line, then by Lemma 4.1, the function $F(\bar{x})$ is self-concordant on this line. In particular, (22) holds.

Let us now consider the case where p changes along the line. We may assume that the line is parameterized by p, i.e.,

$$q(p) = a_0 + b_0 p,$$

and

$$x_i(p) = a_i + b_i p, \text{ for } i = 1, 2, ..., n$$

characterize the line, where p serves as the parameter.

Now, the function restricted to the line can be expressed as

$$\begin{aligned} F(\bar{x}(p)) &= -\log(q(p)/p - f(x(p)/p)) - 2\log p \\ &= -\log(a_0/p + b_0 - f(a_1/p + b_1, ..., a_n/p + b_n)) - 2\log p. \end{aligned}$$

Observe that the function

$$a_0 t + b_0 - f(a_1 t + b_1, ..., a_n t + b_n)$$

is a one-dimensional concave quadratic function in t, as f is a convex quadratic function. To simplify, let

$$a_0 t + b_0 - f(a_1 t + b_1, ..., a_n t + b_n) = at^2 + 2bt + c$$

with $a \leq 0$. Hence,

$$\begin{aligned} F(\bar{x}(p)) &= -\log(\frac{a}{p^2} + \frac{2b}{p} + c) - 2\log p \\ &= -\log(a + 2bp + cp^2). \end{aligned}$$

If $c \leq 0$, then by Lemma 4.1, it follows that $F(\bar{x}(p))$ is self-concordant, and we have the desired inequality (22). Let us consider the case where $c > 0$. Due to Lemma 3.1, $F(\bar{x}(p))$ is a convex function in p for all $p > 0$ and $a + 2bp + cp^2 > 0$. Differentiation yields,

$$F^{(2)}(\bar{x}(p)) = 2\frac{(cp + b)^2 + b^2 - ac}{(a + 2bp + cp^2)^2} \tag{23}$$

and

$$F^{(3)}(\bar{x}(p)) = -4\frac{(cp+b)^3 + 3(b^2-ac)(cp+b)}{(a+2bp+cp^2)^3}. \quad (24)$$

Let $\Delta = b^2 - ac \geq 0$. We derive that

$$\begin{aligned}
|F^{(3)}(\bar{x}(p))| &\leq 4\frac{|cp+b|^3}{(a+2bp+cp^2)^3} + 12\frac{\Delta|cp+b|}{(a+2bp+cp^2)^3} \\
&\leq 4(F^{(2)}(\bar{x}(p))/2)^{3/2} + 12(F^{(2)}(\bar{x}(p))/2)^{3/2} \\
&\leq 5.7(F^{(2)}(\bar{x}(p)))^{3/2}.
\end{aligned}$$

This shows that $F(\bar{x})$ is self-concordant in all directions within its domain.

Q.E.D.

This leads to the following result.

Theorem 4.3 *Consider quadratically constrained quadratic programming; that is, problem* (CP) *where* $f_i(\cdot)$ *are all convex quadratic,* $i = 1, ..., m$. *Then, the conic barrier function for the conic formulation* (CCP),

$$F(\bar{x}) = -m \log p - \sum_{i=1}^{m} \log(q - pf_i(x/p)),$$

is self-concordant with a complexity value in the order of m.

A consequence of this result is that, if one applies a standard (short-step) path-following algorithm for the conic formulation of a quadratically constrained quadratic program, based on (PF_μ), then the method terminates within $O(\sqrt{m}\log\frac{1}{\epsilon})$ number of iterations. This bound is independent of the dimension of the decision variables.

References

[1] E.D. Andersen and Y. Ye, *A computational study of the homogeneous algorithm for large-scale convex optimization*, Working Paper, 1997.

[2] E.D. Andersen and Y. Ye, *On a homogeneous algorithm for monotone complementary problem*, Mathematical Programming 84, 375 – 399, 1999.

[3] A. Berkelaar, R. Kouwenberg, and S. Zhang, A Primal-Dual Decomposition Algorithm for Multi-stage Stochastic Convex Programming, Technical Report SEEM2000-07, Department of Systems Engineering & Engineering Management, The Chinese University of Hong Kong, 2000. (Submitted for publication).

[4] J. Brinkhuis, private conversation, Erasmus University Rotterdam, 2001.

[5] E. de Klerk, C. Roos, and T. Terlaky, *Initialization in semidefinite programming via a self-dual skew-symmetric embedding*, Operations Research Letters 20, 213 – 221, 1997.

[6] Z.-Q. Luo, J.F. Sturm, and S. Zhang, *Duality results for conic convex programming*, Technical Report 9719/A, Econometric Institute, Eramus University Rotterdam, The Netherlands, 1997.

[7] Z.-Q. Luo, J.F. Sturm, and S. Zhang, *Conic convex programming and self-dual embedding*, Optimization Methods and Software 14, 169 – 218, 2000.

[8] Yu. Nesterov and A. Nemirovsky, Interior point polynomial methods in convex programming, Studies in Applied Mathematics 13, SIAM, Philadelphia, PA, 1994.

[9] Yu. Nesterov, M.J. Todd, and Y. Ye, *Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems*, Mathematical Programming 84, 227 – 267, 1999.

[10] F.A. Potra and R. Sheng, *On homogeneous interior-point algorithms for semidefinite programming*, Optimization Methods and Software 9, 161 – 184, 1998.

[11] J. Renegar, *A Mathematical View of Interior-Point Methods in Convex Optimization*, MPS-SIAM Series on Optimization, 2001.

[12] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[13] J.F. Sturm, *Theory and Algorithms for Semidefinite Programming*, in Frenk *et al.* eds., High Performance Optimization, Kluwer Academic Publishers, 1999.

[14] X. Xu, P.F. Hung, and Y. Ye, *A simplified homogeneous self-dual linear programming algorithm and its implementation*, Annals of Operations Research 62, 151-171, 1996.

[15] Y. Ye, *Interior Algorithms for Linear, Quadratic and Linearly Constrained Convex Programming*, Ph.D Thesis, Stanford University, 1987.

[16] Y. Ye, M.J. Todd, and S. Mizuno, *An $\mathcal{O}(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm*, Mathematics of Operations Research 19, 53-67, 1994.