Mathematical Programming 69 (1995) 75-88

A sufficient condition for self-concordance, with application to some classes of structured convex programming problems

D. den Hertog [a,1], F. Jarre [b], C. Roos [a,*], T. Terlaky [a,2]

[a] *Faculty of Technical Mathematics and Computer Science, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, Netherlands*

[b] *Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Am Hubland, Würzburg, Germany*

Received 9 June 1993; revised manuscript received 13 August 1993

Abstract

Recently a number of papers were written that present low-complexity interior-point methods for different classes of convex programs. The goal of this article is to show that the logarithmic barrier function associated with these programs is self-concordant. Hence the polynomial complexity results for these convex programs can be derived from the theory of Nesterov and Nemirovsky on self-concordant barrier functions. We also show that the approach can be applied to some other known classes of convex programs.

Keywords: Interior-point method; Barrier function; Dual geometric programming; (Extended) entropy programming; Primal and dual l_p-programming; Relative Lipschitz condition; Scaled Lipschitz condition; Self-concordance

1. Introduction

The efficiency of a barrier method for solving convex programs strongly depends on the properties of the barrier function used. A key property that is sufficient to prove polynomial convergence for barrier methods is the property of self-concordance introduced in [17]. This condition not only allows a proof of polynomial convergence, but numerical experiments in [1,11,14] and others further indicate that numerical algorithms

* Corresponding author. e-mail: c.roos@math.tudelft.nl.

[1] This author's research was supported by a research grant from SHELL.

[2] On leave from the Eötvös University, Budapest, Hungary. This author's research was partially supported by OTKA No. 2116.

SSDI 0025-5610(94)00065-4

based on self-concordant barrier functions are of practical interest and effectively exploit the structure of the underlying problems.

A well-known barrier function for solving convex programs is the logarithmic barrier function, introduced in [5,6]. To describe the logarithmic barrier function more precisely, we will first give a general form for the classes of problems considered in this paper:

$$(\mathcal{CP}) \qquad \begin{cases} \min f_0(x), \\ f_i(x) \leqslant 0, \quad i = 1, \dots, m, \\ Ax = b, \end{cases}$$

where A is a $p \times n$ matrix and b a p-dimensional vector. The logarithmic barrier function for this program is given by

$$\phi(x, \mu) = \frac{f_0(x)}{\mu} - \sum_{i=1}^{m} \ln(-f_i(x)),$$

where $\mu > 0$ is the barrier parameter. We show that for several classes of convex problems for which interior-point methods were presented in the literature the logarithmic barrier function is self-concordant. These classes are: dual geometric programming, (extended) entropy programming, primal and dual l_p-programming. Since for dual geometric programming and dual l_p-programming no complexity results are known in the literature, these self-concordance proofs enlarge the class of problems for which polynomiality can be proved. (In [12] only a convergence analysis is given.) Moreover, we show that some other smoothness conditions used in the literature (relative Lipschitz condition [3,9], scaled Lipschitz condition [13,25], Monteiro and Adler's condition [16]) are also covered by this self-concordance condition. These observations allow a unification of the analyses of interior-point methods for a number of convex problems.

The article is divided in three parts. In Section 2 we give the definition of self-concordance and state some basic lemmas about self-concordant functions. In Sections 3–6 we prove self-concordance for the classes of problems treated in [7,12,23], and in Section 7 we show that the smoothness conditions used in [3,9,13,16,25] imply self-concordance of the barrier function.

2. Some general composition rules

Let us first give the precise definition of self-concordance as given by Nesterov and Nemirovsky [17].

Definition. Let $\mathcal{F}^0$ be an open convex subset of $\mathbb{R}^n$. A function $\varphi : \mathcal{F}^0 \to \mathbb{R}$ is called *κ-self-concordant* on $\mathcal{F}^0$, $\kappa \geqslant 0$, if φ is three times continuously differentiable in $\mathcal{F}^0$ and if for all $x \in \mathcal{F}^0$ and $h \in \mathbb{R}^n$ the following inequality holds:

$$\nabla^3\varphi(x)[h, h, h] \leqslant 2\kappa \left(h^{\mathrm{T}} \nabla^2 \varphi(x) h\right)^{3/2},$$

where $\nabla^3\varphi(x)[h, h, h]$ denotes the third differential of φ at x and h.

Intuitively, since $\nabla^3\varphi$ describes the change in $\nabla^2\varphi$, and since $\nabla^3\varphi$ is bounded by a suitable power of $\nabla^2\varphi$, this condition implies that the *relative* change of $\nabla^2\varphi$ is bounded by 2κ. The associated norm to measure the relative change is given by $\nabla^2\varphi(x)$, i.e., for $h \in \mathbb{R}^n$ the norm associated with the point x is $\|h\|_{\nabla^2\varphi(x)} := (h^{\mathrm{T}}\nabla^2\varphi(x)h)^{1/2}$. (See [10] and [2], for example, where also a brief analysis is given, showing that the property of self-concordance of the barrier function of a convex program is sufficient to prove polynomial convergence. A more detailed analysis that includes certain nonconvex programs and that uses an additional condition relating the first and second derivatives of φ is given in [17].)

The following lemma gives some helpful composition rules for self-concordant functions. The proof follows immediately from the definition of self-concordance.

Lemma 1 (Nesterov and Nemirovsky [17]).

- (addition and scaling) *Let φ_i be κ_i-self-concordant on $\mathcal{F}_i^0$, $i = 1, 2$, and $\rho_1, \rho_2 \in \mathbb{R}_+$; then $\rho_1\varphi_1 + \rho_2\varphi_2$ is κ-self-concordant on $\mathcal{F}_1^0 \cap \mathcal{F}_2^0$, where* $\kappa = \max\{\kappa_1/\sqrt{\rho_1}, \kappa_2/\sqrt{\rho_2}\}$.
- (affine invariance) *Let φ be κ-self-concordant on $\mathcal{F}^0$ and let $\mathcal{B}(x) = Bx + b : \mathbb{R}^k \to \mathbb{R}^n$ be an affine mapping such that $\mathcal{B}(\mathbb{R}^k) \cap \mathcal{F}^0 \neq \emptyset$. Then $\varphi(\mathcal{B}(\cdot))$ is κ-self-concordant on $\{x\colon\ \mathcal{B}(x) \in \mathcal{F}^0\}$.*

The next lemma gives a sufficient condition for an objective function f to guarantee that f "combined" with the logarithmic barrier function for the positive orthant $\mathbb{R}_+^n$ of $\mathbb{R}^n$ is self-concordant. This lemma will help to simplify self-concordance proofs in the sequel.

Lemma 2. *Let $f(x) \in C^3(\mathcal{F}^0)$ be convex, with $\mathcal{F}^0 \subset \mathbb{R}_+^n$. If there exists a β such that*

$$|\nabla^3 f(x)[h,h,h]| \leqslant \beta h^{\mathrm{T}}\nabla^2 f(x)h\sqrt{\sum_{i=1}^n \frac{h_i^2}{x_i^2}}, \tag{1}$$

$\forall x \in \mathcal{F}^0$ and $\forall h \in \mathbb{R}^n$, then

$$\varphi(x) := f(x) - \sum_{i=1}^n \ln x_i$$

is $(1 + \frac{1}{3}\beta)$-self-concordant on $\mathcal{F}^0$, and

$$\psi(\nu, x) := -\ln\left(\nu - f(x)\right) - \sum_{i=1}^n \ln x_i$$

is $(1 + \frac{1}{3}\beta)$-self-concordant on $\tilde{\mathcal{F}}^0$. Here, $\tilde{\mathcal{F}}^0 \subset \mathbb{R} \times \mathcal{F}^0$ is the set $\{(\nu, x) | x \in \mathcal{F}^0,\ \nu > f(x)\}$.

At a first glance, condition (1) may look somewhat arbitrary. We give a brief motivation right after the following proof, and we will see that the lemma has indeed useful applications.

Proof. We start by proving the first part of the lemma. Straightforward calculations yield

$$\nabla\varphi(x)^{\mathrm{T}}h = \nabla f(x)^{\mathrm{T}}h - \sum_{i=1}^{n}\frac{h_i}{x_i}, \tag{2}$$

$$h^{\mathrm{T}}\nabla^2\varphi(x)h = h^{\mathrm{T}}\nabla^2 f(x)h + \sum_{i=1}^{n}\frac{h_i^2}{x_i^2}, \tag{3}$$

$$\nabla^3\varphi(x)[h,h,h] = \nabla^3 f(x)[h,h,h] - 2\sum_{i=1}^{n}\frac{h_i^3}{x_i^3}. \tag{4}$$

We show that

$$\left(\nabla^3\varphi(x)[h,h,h]\right)^2 \leqslant 4(1+\tfrac{1}{3}\beta)^2\left(h^{\mathrm{T}}\nabla^2\varphi(x)h\right)^3, \tag{5}$$

from which the lemma follows. Since f is convex, the two terms on the right-hand side of (3) are nonnegative, i.e., the right-hand side can be abbreviated by

$$h^{\mathrm{T}}\nabla^2\varphi(x)h = a^2 + b^2, \tag{6}$$

with $a, b \geqslant 0$. Because of (1) we have that

$$|\nabla^3 f(x)[h,h,h]| \leqslant \beta a^2 b.$$

Obviously,

$$\left|\sum_{i=1}^{n}\frac{h_i^3}{x_i^3}\right| \leqslant \sum_{i=1}^{n}\frac{h_i^2}{x_i^2}\sqrt{\sum_{i=1}^{n}\frac{h_i^2}{x_i^2}} = b^3.$$

So we can bound the right-hand side of (4) by

$$|\nabla^3\varphi(x)[h,h,h]| \leqslant \beta a^2 b + 2b^3. \tag{7}$$

It is straightforward to verify that

$$\left(\beta a^2 b + 2b^3\right)^2 \leqslant 4(1+\tfrac{1}{3}\beta)^2\left(a^2+b^2\right)^3.$$

Together with (6) and (7) our claim (5) follows and hence the first part of the lemma.

Now we prove the second part of the lemma. Let

$$\tilde{x} = \begin{pmatrix}\nu\\ x\end{pmatrix}, \qquad h = \begin{pmatrix}h_0\\ \vdots\\ h_n\end{pmatrix}, \qquad g(\tilde{x}) = \nu - f(x) > 0; \tag{8}$$

then,

$$\psi(\tilde{x}) = -\ln g(\tilde{x}) - \sum_{i=1}^{n}\ln x_i, \tag{9}$$

$$\nabla\psi(\bar{x})^{\mathrm{T}}h = -\frac{\nabla g(\bar{x})^{\mathrm{T}}h}{g(\bar{x})} - \sum_{i=1}^{n}\frac{h_i}{x_i}, \tag{10}$$

$$h^{\mathrm{T}}\nabla^2\psi(\bar{x})h = -\frac{h^{\mathrm{T}}\nabla^2 g(\bar{x})h}{g(\bar{x})} + \frac{\left(\nabla g(\bar{x})^{\mathrm{T}}h\right)^2}{g(\bar{x})^2} + \sum_{i=1}^{n}\frac{h_i^2}{x_i^2}, \tag{11}$$

$$\begin{aligned}\nabla^3\psi(\bar{x})[h,h,h] = &-\frac{\nabla^3 g(\bar{x})[h,h,h]}{g(\bar{x})} + 3\frac{\left(h^{\mathrm{T}}\nabla^2 g(\bar{x})h\right)\nabla g(\bar{x})^{\mathrm{T}}h}{g(\bar{x})^2}\\ &-2\frac{\left(\nabla g(\bar{x})^{\mathrm{T}}h\right)^3}{g(\bar{x})^3} - 2\sum_{i=1}^{n}\frac{h_i^3}{x_i^3}.\end{aligned} \tag{12}$$

We show that

$$\left(\nabla^3\psi(\bar{x})[h,h,h]\right)^2 \leqslant 4(1+\tfrac{1}{3}\beta)^2\left(h^{\mathrm{T}}\nabla^2\psi(\bar{x})h\right)^3, \tag{13}$$

which will prove the lemma. Since g is concave, all three terms on the right-hand side of (11) are nonnegative, i.e., the right-hand side can be abbreviated by

$$h^{\mathrm{T}}\nabla^2\psi(\bar{x})h = a^2 + b^2 + c^2, \tag{14}$$

with $a, b, c \geqslant 0$. Due to (1) we have

$$\left|\frac{\nabla^3 g(\bar{x})[h,h,h]}{g(\bar{x})}\right| \leqslant \beta a^2 c,$$

so that we can bound the right-hand side of (12) by

$$|\nabla^3\psi(\bar{x})[h,h,h]| \leqslant \beta a^2 c + 3a^2 b + 2b^3 + 2c^3. \tag{15}$$

It is straightforward to verify that

$$\left(\beta a^2 c + 3a^2 b + 2b^3 + 2c^3\right)^2 \leqslant 4(1+\tfrac{1}{3}\beta)^2\left(a^2+b^2+c^2\right)^3,$$

by eliminating all odd powers in the left-hand side via inequalities of the type $2ab \leqslant a^2 + b^2$. Together with (14) and (15) our claim (13) follows. This proves the lemma. □

We now explain property (1) in more detail. Let $\phi(x) = -\sum_{i=1}^{n}\ln x_i$ be the logarithmic barrier for $\mathbb{R}^n_+$. Observe that

$$\sqrt{\sum_{i=1}^{n}\frac{h_i^2}{x_i^2}} = \sqrt{h^{\mathrm{T}}\nabla^2\phi(x)h} = \|h\|_{\nabla^2\phi(x)}.$$

We recall that (as mentioned above) the canonical norm associated with some barrier function ϕ at a point x is given by $\nabla^2\phi(x)$. Loosely speaking, property (1) tells us that for $\|h\|_{\nabla^2\phi(x)} = 1$, the spectral norm of the third derivative $\nabla^3 f$ is bounded by a multiple β of the spectral norm of the second derivative $\nabla^2 f$. This property is defined in [17] as f being compatible with ϕ, and, as we have seen, it implies self-concordance of the combined barrier functions φ and ψ.

Clearly, if f satisfies (1), then so does f/μ for any (fixed) parameter $\mu > 0$. In particular, this implies that also the function $f(x)/\mu - \sum \ln x_i$ is $(1+\frac{1}{3}\beta)$-self-concordant. Finally we note that for any parameter $q \geqslant 1$ the above proof also holds true for $-q\ln(\nu - f(x)) - \sum_{i=1}^n \ln x_i$. This observation can be used to prove that for the classes of problems considered in this paper not only the logarithmic barrier function but also the center function of [8] (also used in, e.g., [2,9,10,21]) is self-concordant.

3. The dual geometric programming problem

Let $\{I_k\}_{k=1,\dots,r}$ be a partition of $\{1,\dots,n\}$ (i.e., $\bigcup_{k=1}^r I_k = \{1,\dots,n\}$ and $I_k \cap I_l = \emptyset$ for $k \neq l$). The dual geometric programming problem [4] is then given by

$$(\mathcal{DGP}) \qquad \begin{cases} \min c^{\mathrm{T}}x + \sum_{k=1}^r \left[\sum_{i\in I_k} x_i \ln x_i - \left(\sum_{i\in I_k} x_i\right)\ln\left(\sum_{i\in I_k} x_i\right)\right], \\ Ax = b, \\ x \geqslant 0, \end{cases}$$

where A is an $m \times n$ matrix and c and b are n- and m-dimensional vectors, respectively. For this problem we have the following lemma.

Lemma 3. *The logarithmic barrier function of the dual geometric programming problem* ($\mathcal{DGP}$) *is 2-self-concordant*[3].

Proof. Because of Lemma 1, it suffices to verify 2-self-concordance for the logarithmic barrier function

$$\varphi(x) = \sum_{i\in I_k} x_i \ln x_i - \left(\sum_{i\in I_k} x_i\right)\ln\left(\sum_{i\in I_k} x_i\right) - \sum_{i\in I_k} \ln x_i, \tag{16}$$

for some fixed k. For simplicity, we will drop the subscript $i \in I_k$. Now we can use Lemma 2, so that we only have to verify that (1) holds for

$$f(x) := \sum x_i \ln x_i - \left(\sum x_i\right)\ln\left(\sum x_i\right),$$

and $\beta = 3$, which is equivalent to the following inequality:

$$\left|\sum \frac{h_i^3}{x_i^2} - \frac{(\sum h_i)^3}{(\sum x_i)^2}\right| \leqslant 3\left(\sum \frac{h_i^2}{x_i} - \frac{(\sum h_i)^2}{\sum x_i}\right)\sqrt{\sum \frac{h_i^2}{x_i^2}}. \tag{17}$$

Here $x_i > 0$ and h_i arbitrary. Dividing the whole inequality by $\sum x_i$ and then substituting first $h_i = y_i x_i$ and thereafter $t_i = x_i / \sum x_j$ we get the equivalent inequality

$$y^{3^{\mathrm{T}}}t - (y^{\mathrm{T}}t)^3 \leqslant 3\left(y^{2^{\mathrm{T}}}t - (y^{\mathrm{T}}t)^2\right)\sqrt{y^{\mathrm{T}}y},$$

[3] This corrects a remark in [12], in which it is claimed that the self-concordance property does not hold for this problem.

where y_i are arbitrary, t_i positive and $\sum t_i = 1$. (Here y^3, e.g., is the vector with entries y_i^3.) Since $y^{\mathrm{T}}t = E(y)$ can be interpreted as the expected value of some random variable y, the last inequality is equivalently rewritten as

$$E(y^3) - E(y)^3 \leqslant 3\big(E(y^2) - E(y)^2\big)\sqrt{\sum y_i^2},$$

relating the variance of y to some third moment. By adding

$$\begin{aligned}&\big(E(y^2) - E(y)^2\big)\sqrt{\sum y_i^2}\\ &\quad\geqslant \big(E(y^2) - E(y)^2\big)\max y_i = E\big((y - E(y))^2 \max y_i\big)\\ &\quad\geqslant E\big((y - E(y))^2 y\big) = E(y^3) - 2E(y)E(y^2) + E(y)^3\end{aligned}$$

and

$$2\big(E(y^2) - E(y)^2\big)\sqrt{\sum y_i^2} \geqslant 2\big(E(y^2) - E(y)^2\big)E(y) = 2E(y)E(y^2) - 2E(y)^3,$$

we get

$$3\big(E(y^2) - E(y)^2\big)\sqrt{\sum y_i^2} \geqslant E(y^3) - E(y)^3,$$

i.e., inequality (17) follows. □

4. The extended entropy programming problem

The extended entropy programming problem is defined as

$$(\mathcal{EEP}) \qquad \begin{cases} \min c^{\mathrm{T}}x + \sum_{i=1}^{n} g_i(x_i), \\ Ax = b, \\ x \geqslant 0, \end{cases}$$

where A is an $m \times n$ matrix and c and b are n- and m-dimensional vectors, respectively. Moreover, it is assumed that the scalar functions $g_i \in C^3$ satisfy $|g_i'''(x_i)| \leqslant \kappa_i g_i''(x_i)/x_i$, $i = 1, \dots, n$. This class of problems is studied in [7,23] [4]. In the case of entropy programming we have $g_i(x_i) = x_i \ln x_i$, for all i, and $\kappa_i = 1$. Self-concordance for the logarithmic barrier function of this problem simply follows from the following lemma.

Lemma 4. *Suppose that* $|g_i'''(x_i)| \leqslant \kappa_i g_i''(x_i)/x_i$, $i = 1, \dots, n$; *then the logarithmic barrier function for the extended entropy programming problem* $(\mathcal{EEP})$ *is* $(1 + \frac{1}{3}\max_i \kappa_i)$-*self-concordant.*

Proof. Using Lemma 1 it suffices to show that

$$g_i(x_i) - \ln x_i$$

[4] In [7] it is conjectured that these problems do not satisfy the self-concordance condition. The lemma shows that this conjecture is not true.

is $(1+\frac{1}{3}\kappa_i)$-self-concordant. Since (1) reduces in the present case to

$$|g_i'''(x_i)| \leqslant \kappa_i g_i''(x_i)\frac{1}{x_i},$$

this immediately follows from Lemma 2. $\square$

5. The primal l_p-programming problem

Let $\{I_k\}_{k=1,\dots,r}$ be a partition of $\{1,\dots,m\}$ (i.e., $\bigcup_{k=1}^r I_k = \{1,\dots,m\}$ and $I_k \cap I_l = \emptyset$ for $k \neq l$). Let $p_i \geqslant 1$, $i=1,\dots,m$. Then the primal l_p-programming problem [18,22] can be formulated as

$$(\mathcal{PL}_p) \qquad \begin{cases} \max \eta^{\mathrm{T}}x, \\ \sum_{i\in I_k}(1/p_i)|a_i^{\mathrm{T}}x - c_i|^{p_i} + b_k^{\mathrm{T}}x - d_k \leqslant 0, \quad k=1,\dots,r, \end{cases}$$

where (for all i and k) a_i, b_k and η are n-dimensional vectors, and c_i and d_k are real numbers. Nesterov and Nemirovsky [17] treated a special case of this problem, namely the so-called l_p-approximation problem. We will reformulate $(\mathcal{PL}_p)$ such that all problem functions remain convex, contrary to Nesterov and Nemirovsky's reformulation.

In a first step, the primal l_p-programming problem can be reformulated as:

$$\begin{cases} \max \eta^{\mathrm{T}}x, \\ \sum_{i\in I_k}(1/p_i)t_i + b_k^{\mathrm{T}}x - d_k \leqslant 0, \quad k=1,\dots,r, \\ \left.\begin{array}{l} s_i^{p_i} \leqslant t_i, \\ a_i^{\mathrm{T}}x - c_i \leqslant s_i, \\ -a_i^{\mathrm{T}}x + c_i \leqslant s_i, \end{array}\right\} \quad i=1,\dots,m, \\ s \geqslant 0. \end{cases} \tag{18}$$

In the same way as we will prove Lemma 5, it can be proved that the logarithmic barrier function for this reformulated l_p-programming problem is $(1+\frac{1}{3}\max_i|p_i-2|)$-self-concordant, i.e., the concordance parameter depends on p_i. We can eliminate this dependence as follows. Replace the constraints $s_i^{p_i} \leqslant t_i$ by the equivalent constraints $s_i \leqslant t_i^{\pi_i}$, where $0 < \pi_i := 1/p_i \leqslant 1$, and replace the (redundant) constraints $s \geqslant 0$ by $t \geqslant 0$. So, we obtain the following reformulated l_p-programming problem:

$$(\mathcal{PL}_p') \qquad \begin{cases} \max \eta^{\mathrm{T}}x, \\ \sum_{i\in I_k} t_i/p_i + b_k^{\mathrm{T}}x - d_k \leqslant 0, \quad k=1,\dots,r, \\ \left.\begin{array}{l} s_i \leqslant t_i^{\pi_i}, \\ a_i^{\mathrm{T}}x - c_i \leqslant s_i, \\ -a_i^{\mathrm{T}}x + c_i \leqslant s_i, \end{array}\right\} \quad i=1,\dots,m, \\ t \geqslant 0. \end{cases} \tag{19}$$

Observe that the transformed problem has $4m+r$ constraints, compared with r in the original problem $(\mathcal{PL}_p)$. Now we have the following lemma.

Lemma 5. *The logarithmic barrier function for the reformulated l_p-programming problem $(\mathcal{PL}_p')$ is $\frac{5}{3}$-self-concordant.*

Proof. First note that the logarithmic barrier function for the linear constraints is 1-self-concordant. Moreover, since $f_i(t_i) := -t_i^{\pi_i}$, $0 < \pi_i \leqslant 1$, satisfies (1) with $\beta = 2 - \pi_i$, we have from Lemma 2 that

$$-\ln(t_i^{\pi_i} - s_i) - \ln t_i$$

is $(1 + \frac{1}{3}(2 - \pi_i))$-self-concordant, where $0 < \pi_i \leqslant 1$. From Lemma 1 it follows that the logarithmic barrier function is $\frac{5}{3}$-self-concordant. □

6. The dual l_p-programming problem

Let q_i be such that $1/p_i + 1/q_i = 1$, $1 \leqslant i \leqslant m$, and let the rows of a matrix A be a_i, $i = 1, \ldots, m$, and the rows of a matrix B be b_k, $k = 1, \ldots, r$. Then, the dual of the l_p-programming problem $(\mathcal{PL}_p)$ is (see [18–20,22])

$$(\mathcal{DL}_p) \qquad \begin{cases} \min c^{\mathrm{T}} y + d^{\mathrm{T}} z + \sum_{k=1}^{r} z_k \sum_{i \in I_k} (1/q_i) |y_i / z_k|^{q_i}, \\ A^{\mathrm{T}} y + B^{\mathrm{T}} z = \eta, \\ z \geqslant 0. \end{cases}$$

(If $y_i \neq 0$ and $z_k = 0$, then $z_k |y_i/z_k|^{q_i}$ is defined as ∞.) The above problem is equivalent to

$$\begin{cases} \min c^{\mathrm{T}} y + d^{\mathrm{T}} z + \sum_{i=1}^{n} t_i / q_i, \\ s_i^{q_i} z_k^{-q_i+1} \leqslant t_i, \quad i \in I_k, \quad k = 1, \ldots, r, \\ y \leqslant s, \\ -y \leqslant s, \\ A^{\mathrm{T}} y + B^{\mathrm{T}} z = \eta, \\ z \geqslant 0, \\ s \geqslant 0. \end{cases} \tag{20}$$

Similarly as in the proof of the next lemma, it can be proved that the logarithmic barrier function of this reformulated dual l_p-programming problem is $(1 + \frac{1}{3}\sqrt{2} \max_i (q_i + 1))$-self-concordant. Again, the dependence on q_i can be eliminated: the constraints $s_i^{q_i} z_k^{-q_i+1} \leqslant t_i$ are replaced by the equivalent constraints $t_i^{\rho_i} z_k^{-\rho_i+1} \geqslant s_i$, where $0 < \rho_i := 1/q_i \leqslant 1$, and the redundant constraints $s \geqslant 0$ are replaced by $t \geqslant 0$. The new reformulated dual l_p-programming problem becomes:

$$(\mathcal{DL}'_p) \qquad \begin{cases} \min c^{\mathrm{T}}y + d^{\mathrm{T}}z + \sum_{i=1}^n t_i/q_i, \\ s_i \leqslant t_i^{\rho_i} z_k^{-\rho_i+1}, \quad i \in I_k, \quad k = 1, \ldots, r, \\ y \leqslant s, \\ -y \leqslant s, \\ A^{\mathrm{T}}y + B^{\mathrm{T}}z = \eta, \\ z \geqslant 0, \\ t \geqslant 0. \end{cases} \tag{21}$$

Note that the original problem $(\mathcal{DL}_p)$ has r inequalities, and the reformulated problem $(\mathcal{DL}'_p)$ $4m + r$. We now have the following lemma.

Lemma 6. *The logarithmic barrier function of the reformulated dual l_p-programming problem $(\mathcal{DL}'_p)$ is 2-self-concordant.*

Proof. By Lemma 1 it suffices to show that

$$-\ln(t_i^{\rho_i} z_k^{-\rho_i+1} - s_i) - \ln z_k - \ln t_i$$

is $(1 + \frac{1}{3}\sqrt{2}(\rho_i + 1))$-self-concordant, or equivalently by Lemma 2 that (we will omit the subscripts i and k in the sequel of this proof) $f(t, z) := -t^{\rho} z^{-\rho+1}$ with $0 < \rho < 1$ satisfies (1) for $\beta = \sqrt{2}(\rho + 1)$, i.e.,

$$|\nabla^3 f(t, z)[h, h, h]| \leqslant \sqrt{2}(\rho + 1) h^{\mathrm{T}} \nabla^2 f(t, z) h \sqrt{\frac{|h_1|^2}{t^2} + \frac{|h_2|^2}{z^2}}, \tag{22}$$

where $h^{\mathrm{T}} = (h_1, h_2)$. After doing some straightforward calculations, we obtain for the second-order term

$$\begin{aligned} h^{\mathrm{T}} \nabla^2 f(t, z) h &= \rho(1 - \rho) t^{\rho-3} z^{-\rho-2} (t z^3 h_1^2 + t^3 z h_2^2 - 2t^2 z^2 h_1 h_2) \\ &= \rho(1 - \rho) t^{\rho-3} z^{-\rho-2} (z h_1 - t h_2)^2 t z, \end{aligned}$$

and for the third-order term

$$\begin{aligned} &|\nabla^3 f(t, z)[h, h, h]| \\ &\quad = \rho(1 - \rho) s^{\rho-3} z^{-\rho-2} \\ &\qquad \times |(\rho - 2) z^3 h_1^3 - (\rho + 1) t^3 h_2^3 - 3(\rho - 1) t z^2 h_1^2 h_2 + 3\rho t^2 z h_1 h_2^2| \\ &\quad = \rho(1 - \rho) t^{\rho-3} z^{-\rho-2} (z h_1 - t h_2)^2 |(\rho - 2) z h_1 - (\rho + 1) t h_2| \\ &\quad \leqslant \rho(1 - \rho)(\rho + 1) t^{\rho-3} z^{-\rho-2} (z h_1 - t h_2)^2 (z|h_1| + t|h_2|). \end{aligned}$$

Now we obtain

$$\frac{|\nabla^3 f(t, z)[h, h, h]|}{h^{\mathrm{T}} \nabla^2 f(t, z) h} \leqslant (\rho + 1) \left(\frac{|h_1|}{t} + \frac{|h_2|}{z} \right) \leqslant \sqrt{2}(\rho + 1) \sqrt{\frac{|h_1|^2}{t^2} + \frac{|h_2|^2}{z^2}}.$$

This proves (22). The lemma follows since $\rho_i \leqslant 1$. $\square$

7. Other smoothness conditions

Relative Lipschitz condition

Jarre [9] introduced the following relative Lipschitz condition (also used in, e.g., [3]) for the Hessian matrix of the problem functions $f_i(x)$, $0 \leqslant i \leqslant m$, of $(\mathcal{CP})$:

$$\exists M > 0: \quad \forall v \in \mathbb{R}^n \quad \forall x, x+h \in \mathcal{F}^0:$$
$$|v^{\mathrm{T}}(\nabla^2 f_i(x+h) - \nabla^2 f_i(x))v| \leqslant M\|h\|_H v^{\mathrm{T}} \nabla^2 f_i(x) v, \tag{23}$$

where H is the Hessian matrix of the corresponding logarithmic barrier function. As shown in [10], if the Hessians of the problem functions f_i of $(\mathcal{CP})$ fulfil this relative Lipschitz condition with parameter M, and if $f_i \in C^3$, then the associated logarithmic barrier function is $(1+M)$-self-concordant. (The converse is not true.) Moreover, in [10] it is shown that the relative Lipschitz condition for the logarithmic barrier function is equivalent to self-concordance if the underlying function is three times continuously differentiable.

Monteiro and Adler's condition

Monteiro and Adler [16] considered minimization problems with linear equality constraints and a separable convex objective function on the positive orthant of $\mathbb{R}^n$. The objective function $f(x) = \sum_i g_i(x_i)$ must satisfy the following condition:

There exist positive numbers T and p such that for all reals $x > 0$ and $y > 0$ and all $i = 1, \ldots, n$, we have

$$y|g_i'''(y)| \leqslant T \max\left\{\left(\frac{x}{y}\right)^p, \left(\frac{y}{x}\right)^p\right\} g_i''(x).$$

Using Lemma 2 and substituting $y = x$ in the above condition, it is easy to see that g_i satisfies (1) with $\beta = T$, i.e., that the logarithmic barrier function for such a problem is $(1 + \frac{1}{3}T)$-self-concordant. Using Lemma 2 we may simplify the condition of [16] to the (weaker) condition that there exists a positive number T such that for all reals $y > 0$ and all $i = 1, \ldots, n$, we have

$$y|g_i'''(y)| \leqslant T g_i''(x).$$

This condition is not only simpler, also the dependence on some extra parameter p is eliminated.

Scaled Lipschitz condition

In [13,25] interior-point methods are given and analyzed for problems with linear equality constraints and convex objective function $f(x)$ on the positive orthant of $\mathbb{R}^n$. The objective function has to satisfy the following scaled Lipschitz condition:

There exists $M > 0$, such that for any γ, $0 < \gamma < 1$,

$$\|X(\nabla f(x + \Delta x) - \nabla f(x) - \nabla^2 f(x)\Delta x)\| \leqslant M \Delta x^{\mathrm{T}} \nabla^2 f(x) \Delta x, \tag{24}$$

whenever $x > 0$ and $\|X^{-1}\Delta x\| \leqslant \gamma$. (Here, $\|\cdot\|$ is the Euclidean norm.)

This condition is also covered by the self-concordance condition if f is three times continuously differentiable in the interior of the feasible domain. More precisely we will show in the next lemma that the corresponding logarithmic barrier function is $(1+\frac{2}{3}M)$-self-concordant.

Lemma 7. *Suppose $f(x) \in C^3$ fulfils the scaled Lipschitz condition with parameter M. Then the logarithmic barrier functions φ and ψ from Lemma 2 are $(1+\frac{2}{3}M)$-self-concordant.*

Proof. It suffices to prove (1). Set $h = \Delta x$ as in definition (24). First note that

$$\sqrt{\sum_{i=1}^{n} \frac{h_i^2}{x_i^2}} = \|X^{-1}\Delta x\|.$$

Since $f \in C^3$, we may expand ∇f as follows:

$$\nabla f(x+\Delta x) = \nabla f(x) + \nabla^2 f(x)\Delta x + \tfrac{1}{2}\nabla^3 f(x)[\Delta x, \Delta x, \cdot] + \mathrm{o}(\|\Delta x\|^2),$$

where $\nabla^3 f(x)[\Delta x, \Delta x, \cdot]$ is a vector whose ith component is equal to

$$\sum_{j,k} \frac{\partial^3 f(x)}{\partial x_i\, \partial x_j\, \partial x_k} \Delta x_j\, \Delta x_k.$$

Replacing Δx by $\lambda \Delta x$ in definition (24), inserting the above expansion, dividing by λ^2, and taking the limit as λ tends to zero, we obtain

$$\|X\nabla^3 f(x)[\Delta x, \Delta x, \cdot]\| \leqslant 2M \Delta x^{\mathrm{T}} \nabla^2 f(x) \Delta x. \tag{25}$$

Considering $X\nabla^3 f(x)[\Delta x, \Delta x, \cdot]$ as a column vector, we may continue

$$\begin{aligned} \|X\nabla^3 f(x)[\Delta x, \Delta x, \cdot]\| &\geqslant \frac{(X^{-1}\Delta x)^{\mathrm{T}}}{\|X^{-1}\Delta x\|} X\nabla^3 f(x)[\Delta x, \Delta x, \cdot] \\ &= \frac{\nabla^3 f(x)[\Delta x, \Delta x, \Delta x]}{\|X^{-1}\Delta x\|}, \end{aligned}$$

and obtain that

$$\nabla^3 f(x)[\Delta x, \Delta x, \Delta x] \leqslant 2M\|X^{-1}\Delta x\| \Delta x^{\mathrm{T}} \nabla^2 f(x) \Delta x,$$

which is exactly relation (1). □

Before we conclude this work, we would like to briefly point out a class of problems considered in [15] (and also in [24]) which does not have a self-concordant logarithmic barrier function. Mehrotra and Sun [15] introduced a curvature constraint of the following form. There exists a number $\kappa \geqslant 1$ such that for all x, y and h in $\mathbb{R}^n$,

$$h^{\mathrm{T}} \nabla^2 f_i(x) h \leqslant \kappa h^{\mathrm{T}} \nabla^2 f_i(y) h.$$

For constraint functions f_i satisfying this condition, they present a polynomial-time interior-point algorithm (which needs at most $O(\kappa^5\sqrt{m}\ln\epsilon)$ Newton iterations to reduce the error by a factor of ϵ). Clearly, there are constraints with self-concordant barriers that do not satisfy this condition, and, conversely, this condition covers some constraint functions that do not have a self-concordant barrier function. For most applications however, we believe that the self-concordance condition is more practical.

References

[1] F. Alizadeh, "Optimization over the positive definite cone: interior-point methods and combinatorial applications," in: P. Pardalos, ed., *Advances in Optimization and Parallel Computing* (North-Holland, Amsterdam, 1992) pp. 1–25.

[2] D. den Hertog, *Interior Point Approach to Linear, Quadratic and Convex Programming: Algorithms and Complexity* (Kluwer Academic Publishers, Dordrecht, 1994).

[3] D. den Hertog, C. Roos and T. Terlaky, "On the classical logarithmic barrier method for a class of smooth convex programming problems," *Journal of Optimization Theory and Applications* 73 (1) (1992) 1–25.

[4] R.J. Duffin, E.L. Peterson and C. Zener, *Geometric Programming* (Wiley, New York, 1967).

[5] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming, Sequential Unconstrained Minimization Techniques* (Wiley, New York, 1968).

[6] R. Frisch, "The logarithmic potential method for solving linear programming problems," Memorandum, University Institute of Economics, Oslo, Norway, 1955.

[7] C.-G. Han, P.M. Pardalos and Y. Ye, "On interior-point algorithms for some entropy optimization problems," Technical Report CS 91-02, Computer Science Department, Pennsylvania State University, University Park, PA, 1991.

[8] P. Huard, "Resolution of mathematical programming with nonlinear constraints by the methods of centres," in: J. Abadie, ed., *Nonlinear Programming* (North-Holland, Amsterdam, 1967) pp. 207–219.

[9] F. Jarre, "The method of analytic centers for smooth convex programs," Ph.D. Thesis, Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Germany, 1989.

[10] F. Jarre, "Interior-point methods for convex programming," *Applied Mathematics and Optimization* 26 (1992) 287–311.

[11] F. Jarre and M.A. Saunders, "Practical aspects of an interior-point method for convex programming," *SIAM Journal on Optimization*, to appear.

[12] K.O. Kortanek and H. No, "A second order affine scaling algorithm for the geometric programming dual with logarithmic barrier," *Optimization* 23 (1990) 501–507.

[13] K.O. Kortanek and J. Zhu, "A polynomial barrier algorithm for linearly constrained convex programming problems," *Mathematics of Operations Research* 18 (1993) 116–128.

[14] I.J. Lustig, R.E. Marsten and D.F. Shanno, "On implementing Mehrotra's predictor–corrector interior point method for linear programming," *SIAM Journal on Optimization* 2 (1992) 435–449.

[15] S. Mehrotra and J. Sun, "An interior point algorithm for solving smooth convex programs based on Newton's method," *Contemporary Mathematics* 114 (1990) 265–284.

[16] R.D.C. Monteiro and I. Adler, "An extension of Karmarkar type algorithms to a class of convex separable programming problems with global rate of convergence," *Mathematics of Operations Research* 15 (3) (1989) 408–422.

[17] Y.E. Nesterov and A.S. Nemirovsky, *Interior Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, Vol. 13 (SIAM, Philadelphia, PA, 1994).

[18] E.L. Peterson and J.G. Ecker, "Geometric programming: duality in quadratic programming and l_p approximation I," in: H.W. Kuhn and A.W. Tucker, eds., *Proceedings of the International Symposium of Mathematical Programming* (Princeton University Press, Princeton, NJ, 1970) pp. 445–479.

[19] E.L. Peterson and J.G. Ecker, "Geometric programming: duality in quadratic programming and l_p approximation II," *SIAM Journal on Applied Mathematics* 17 (1969) 317–340.

[20] E.L. Peterson and J.G. Ecker, "Geometric programming: duality in quadratic programming and l_p approximation III," *Journal of Mathematical Analysis and Applications* 29 (1970) 365–383.

[21] J. Renegar, "A polynomial-time algorithm, based on Newton's method, for linear programming," *Mathematical Programming* 40 (1) (1988) 59–93.
[22] T. Terlaky, "On l_p programming," *European Journal of Operational Research* 22 (1985) 70–100.
[23] Y. Ye and F. Potra, "An interior-point algorithm for solving entropy optimization problems with globally linear and locally quadratic convergence rate," Working Paper Series No. 90-22, Department of Management Sciences, University of Iowa, Iowa City, IA, 1990; also: *SIAM Journal on Optimization*, to appear.
[24] S. Zhang, "The convergence property of Iri-Imai's algorithm for some smooth convex programming problems," Research Memorandum no. 485, Institute of Economic Research, Faculty of Economics, University of Groningen, Netherlands, 1992.
[25] J. Zhu, "A path following algorithm for a class of convex programming problems," *Zeitschrift für Operations Research* 36 (4) (1992) 359–377.