Journal of Mathematical Analysis and Applications **234**, 675–694 (1999)
Article ID jmaa.1999.6407, available online at http://www.idealibrary.com on IDEAL®

Convex Constrained Programmes with Unattained Infima

Wiesława T. Obuchowska

Submitted by E. S. Lee

Received November 2, 1998

We consider a problem of minimization of convex function $f(x)$ over the convex region $\mathscr{R}$ where the objective function and the feasible region have a common direction of recession. In cases when one of these directions is not in the constancy space of the objective function, then the minimal solution is not achieved even if the function $f(x)$ is bounded below over the region $\mathscr{R}$. Many algorithms, if applied to this class of programmes, do not guarantee convergence to the global infimum. Our approach to this problem leads to derivation of the equation of the feasible parametrized curve $C(t)$, such that the infimum of the logarithmic penalty function along this curve is equal to the global infimum of the objective function over the region $\mathscr{R}$. We show that if all functions defining the program are analytic, then $C(t)$ is also an analytic function. The equation of the curve can be successfully used to determine the global infimum (in particular, unboundedness) of the convex constrained programmes in cases when the application of classical methods, such as the steepest descent method, fails to converge to the global infimum.

1. INTRODUCTION

We consider the problem

$$\text{minimize } f_0(x)$$
$$\text{subject to: } x \in \mathscr{R} = \{x \in R^n | f_i(x) \le 0, i \in I = \{1, 2, \dots, m\}\} \quad (1.1)$$

where $f_i(x)$, $i \in I \cup \{0\}$, are analytic convex functions with unbounded level sets. For simplicity of notation the objective function will also be

0022-247X/99 $30.00

referred to as $f(x)$. We assume that $\operatorname{int}(\mathscr{R}) \neq \emptyset$. The problem (1.1) either has a finite infimum (although there may be no such point at which $f(x)$ achieves its minimum), or the function $f(x)$ is unbounded from below [2–8].

It is well known [11] that the minimum of the convex problem of the form (1.1) is not achieved only if the objective function and the feasible region have a common direction of recession. In this case the convex program (1.1) is called degenerate. Abrams [1] proposed a method to solve the degenerate problem of the form (1.1), with the assumption that the problem is bounded below and that the set

$$C = \bigcap_{i=0}^{m} 0^{+} f_i$$

of the vectors of recession of the problem is known. The method relies on replacing the original problem with a reduced problem defined by projecting the feasible region $\mathscr{R}$ and the epigraph of $f(x)$ onto the orthogonal complement of the recession cone of the problem. The projection procedure may be repeated in case the reduced problem is still degenerate, until the objective function and the feasible region of a canonical problem have no directions in common. One of the difficulties in the above approach is that it requires knowledge of the cone of recession of the problem (1.1). Also, the reduction procedure is known in only a few cases: posynomial, geometric programming and quadratic programming and l_p-approximation [1, 4]. Furthermore, the procedure fails if the objective function strictly decreases along some direction vector in the cone C. Indeed, if every vector in C belongs to the constancy space of f, then the problem (1.1) achieves its infimum, which may be determined by the method in [1]. Our method of the solution of the problem (1.1) relies on derivation of the equation of the feasible trajectory such that the infimum of the modified logarithmic penalty function along this curve is equal to the infimum of $f(x)$ on $\mathscr{R}$. We will call such a curve an infimal trajectory. We do not impose any restrictions on the cone of recession of the problem (1.1). However, we assume that the direction vector $p \in C$ along which the function $f(x)$ strictly decreases is given. We also do not impose an assumption on boundedness of $f(x)$ over $\mathscr{R}$, which implies, in particular, that the function may be divergent to $-\infty$ along the trajectory. The method proposed in this paper is an extension of the approach proposed in [10] for the unconstrained minimization problem, where we have derived the equation of the parametrized analytic curve, along which the objective function converges monotonically to its global infimum.

2. PRELIMINARY RESULTS

It is well known that if a closed, proper, convex analytic function is constant along some half-line with a direction vector d, then it is constant along any half-line with this direction. The set of vectors with the latter property forms what is called constancy space of $f(x)$, which is denoted by $D_f^=$, [7, 11]. A vector s is called a direction of recession of $f(x)$ if for every x the function $f(x + ts)$ is a nonincreasing function of t [6, 11]. Since any proper lower semicontinuous convex function is a closed function, then Theorem 8.6 [11] implies that if $f(x + ts)$ is nonincreasing for even one $x \in R^n$, then it is nonincreasing for every x. We assume in this paper that the intersection of the cones of recession of f and f_i, $i \in I$, denoted by 0^+f, and 0^+f_i respectively, is nonempty. The lineality space $D_f^=$ of $f(x)$ may be defined in terms of the set 0^+f [11] as

$$D_f^= = \{y \in R^n | y \in 0^+f \wedge -y \in 0^+f\}.$$

DEFINITION 2.1. (i) We say that the function $f(x)$ decreases asymptotically along the half-line $\bar{x}(t) = \bar{x} + ts$, $t \geq 0$ if it has a finite infimum but not a minimum along this half-line.

(ii) We say that the vector $s \in 0^+f$ is an *asymptotic direction* vector if f has a finite infimum but not a minimum along every half-line with the direction s. Let D_f^a denote the set of all vectors s with the above property.

(iii) We say that the vector $s \in 0^+f$ is a direction of *unboundedness*, if the function f is unbounded below along every half-line with the direction s. The set of all vectors possessing this property will be denoted by D_f^u.

LEMMA 2.1 [9]. (i) *If the convex function* $f(x)$ *is unbounded below along the half-line* $x_a(t) = a + ts$, $t \geq 0$, *then for every* $\bar{x} \in R^n$, $f(x)$ *is unbounded below along a half-line* $\bar{x}(t) = \bar{x} + ts$, $t \geq 0$ *and along a half-line* $\bar{x}(t) = \bar{x} + ty$, $t \geq 0$, *for any* $y \in \operatorname{rint}(0^+f)$.

(ii) *Let us assume that* $f(x) \in C^\infty$. *Then, if* $f(x)$ *asymptotically decreases along a half-line* $x_a(t) = a + ts$, $t \geq 0$, *then it also asymptotically decreases along any half-line with the direction* s.

(iii) $0^+f = D_f^= \cup D_f^u \cup D_f^a$, *and* $D_f^u \cap D_f^a = \emptyset$.

(iv) *If* $f \in C^\infty$ *and* $\operatorname{rint}(0^+f) \cap D_f^a \neq \emptyset$, *then every vector* $s \in \operatorname{rint}(0^+f)$ *is an asymptotic direction.*

LEMMA 2.2. *Let us assume that the function* f *is convex, and* $f \in C^\infty$. *Then* f *achieves an infimum over* $\mathscr{R}$ *if and only if* $C \subset D_f^=$.

Proof. The backward implication follows from Corollary 27.3.3 [11]. To prove the forward implication, let us assume that x_0 be the point at which the objective function achieves minimum over the region $\mathscr{R}$. Thus for any

$s \in C$, $f(x_0 + ts) = f(x_0)$ for $t \geq 0$. Since $f \in C^\infty$ the latter equation implies that $s \in D_f^=$, and consequently $C \subset D_f^=$. ∎

We assume in the remaining part of this paper that the vector $p \in D_f^a \cap C$ ($\|p\| = 1$) is given. We also assume that $D_f^= \cap C = \emptyset$. Otherwise, if $D_f^= \cap C \neq \emptyset$, we can apply the algorithm in [7] to determine the constancy space $D_f^= \cap C$ (or $D_f^=$), and consequently replace the original problem by its orthogonal projection onto the constancy space $D_f^= \cap C$ (or $D_f^=$, respectively). Since the problem (1.1) has an unattainable infimum, the cone of recession of the new problem will contain only asymptotic directions (D_f^a) and possible directions of unboundedness (D_f^u), for which, by Lemma 2.1, $D_f^a \cap D_f^u = \emptyset$. In the lemma below we assume that $y_i \in R^n$, $i = 1, 2, \ldots, n-1$, are linearly independent, although more specific assumptions on these vectors will be made later. Let us define $Y = (y_1, y_2, \ldots, y_{n-1}) \in R^{(n-1)\times n}$. Let $x_0 \in \mathscr{R}$ be arbitrary but fixed.

LEMMA 2.3 [10]. *Let f: $R^n \to R$ be a convex function assuming finite values for all $x \in R^n$. If there exists a sequence $\{x^i\}$ such that $\lim_{i\to\infty}\|x^i\| = \infty$ and*

$$\limsup_{i\to\infty} f(x^i) < \infty, \tag{2.1}$$

then an arbitrary cluster point of the sequence $\{x^i/\|x^i\|\}$ belongs to 0^+f.

Proof. We will prove the lemma by contradiction. First we note that since f is finite at all $x \in R^n$, then it is a proper, closed convex function of x. Let $\{x^i\}$ satisfy $\lim_{i\to\infty}(x^i/\|x^i\|) = \hat{x}$. Clearly, if $\hat{x} \notin 0^+f$, then it belongs to the open set $R^n \setminus 0^+f$. We will prove first that the latter fact and the above properties of $f(x)$ imply that the function f is unbounded from above along any half-line with the direction $\hat{x}$. To this end let us assume that the opposite is true, that is, $\exists\bar{x}$, such that $f(x)$ is bounded from above along the half-line $\bar{x}(t) = \bar{x} + t\hat{x}$, $t \geq 0$. Since $\hat{x} \notin 0^+f$, then $\exists\bar{t} \geq 0$, such that $f(x)$ is increasing along the half-line $\bar{x}(t)$, for $t \geq \bar{t}$. Therefore we have

$$\liminf_{t\to\infty} f(\bar{x} + t\hat{x}) < \infty.$$

Since f is a proper, closed convex function, then by the second part of Theorem 8.6 in [11], $f(x)$ is a nonincreasing function of t along any line $x + t\hat{x}$ for every $x \in \operatorname{dom} f$. So, $\hat{x} \in 0^+f$, which contradicts the assumption and proves that $f(x)$ is unbounded from above along any half-line with the direction $\hat{x}$.

Now, we observe that

$$\exists \delta_0 > 0, \qquad \left\langle \hat{x}, \frac{x}{\|x\|} \right\rangle \geq 1 - \delta_0 \quad \Rightarrow \quad x \notin 0^+ f.$$

Therefore, any vector x satisfying the latter relation has the property that along any half-line with the direction vector x, f is unbounded from above. Let $f_{\inf} = \liminf_{i \to \infty} f(x^i)$. By assumption (2.1) it follows that there exists a subsequence $\{x^{i_j}\}$ of the sequence $\{x^i\}$ such that $\lim_{j \to \infty} f(x^{i_j}) = f_{\inf}$. Without loss of generality we will use the symbol $\{x^i\}$ for the subsequence satisfying the latter equation. Let us define the half-line $\hat{x}(\tau) = \tau \hat{x}$, $\tau \geq 0$. Since $\lim_{\tau \to \infty} f(\hat{x}(\tau)) = \infty$, then for an arbitrarily large (but fixed) number $W > 0$, such that

$$W > \max\{f(0) + 2, f_{\inf} + 3\}, \tag{2.2}$$

there exists such $\tau_W > 0$ that $f(\hat{x}(\tau_W)) = W$. (In the case when $f_{\inf} = -\infty$, the latter inequality is equivalent to $W > f(0) + 2$.) By continuity of f we have that

$$\exists \epsilon_1 > 0, \qquad \|x - \hat{x}(\tau_W)\| \leq \epsilon_1 \quad \Rightarrow \quad f(x) \geq W - 1. \tag{2.3}$$

We assume that ϵ_1 is a sufficiently small number, so that $x = 0$ does not satisfy the inequality $\|x - \hat{x}(\tau_W)\| \leq \epsilon_1$. In the case when this assumption is not satisfied, we can simply decrease the value ϵ_1. We obtain the following implication:

$$\exists \delta \in [0,1], \qquad \|x - \hat{x}(\tau_W)\| \leq \epsilon_1 \quad \Rightarrow \quad \left\langle \hat{x}, \frac{x}{\|x\|} \right\rangle \geq 1 - \delta. \tag{2.4}$$

The value δ, which satisfies the latter implication, is not unique, and we define

$$\delta_{\inf} = \inf\left\{ \delta \mid \delta \in [0,1], \qquad \|x - \hat{x}(\tau_W)\| \leq \epsilon_1 \quad \Rightarrow \quad \left\langle \hat{x}, \frac{x}{\|x\|} \right\rangle \geq 1 - \delta \right\}.$$

The value $\delta_{\inf}$ is well defined, because the infimum is defined over the set of values δ, which by relation (2.4) is nonempty and bounded. From the fact that the sequence $\{x^i/\|x^i\|\}$ converges to $\hat{x}$, it follows that

$$\exists i_{\delta_{\inf}}, \qquad \forall i > i_{\delta_{\inf}}, \qquad \left\langle \hat{x}, \frac{x^i}{\|x^i\|} \right\rangle \geq 1 - \delta_{\inf}. \tag{2.5}$$

Let i_0 be such that $i_0 > i_{\delta_{\inf}}$ and

$$\|x^{i_0}\| > \|\hat{x}(\tau_W)\| + \epsilon_1, \tag{2.6}$$

(such an i_0 exists because $\{\|x^{i_0}\|\}$ is an unbounded sequence).

In the case when $f_{\inf} = -\infty$, we determine $\bar{i}_0 \geq i_0$, such that $x^{\bar{i}_0}$ satisfies

$$f(x^{\bar{i}_0}) < 0. \tag{2.7}$$

In the case when $f_{\inf} > -\infty$, the index $\bar{i}_0 \geq i_0$ is determined so that $x^{\bar{i}_0}$ satisfies

$$f(x^{\bar{i}_0}) < f_{\inf} + 1. \tag{2.8}$$

From (2.5), we conclude that

$$\left\langle \hat{x}, \frac{x^{\bar{i}_0}}{\|x^{\bar{i}_0}\|} \right\rangle \geq 1 - \delta_{\inf}. \tag{2.9}$$

The latter inequality implies that any point x defined by $x = \tau x^{\bar{i}_0}$, $0 < \tau \leq 1$, satisfies

$$\left\langle \hat{x}, \frac{x}{\|x\|} \right\rangle \geq 1 - \delta_{\inf}.$$

Let us denote

$$B(\hat{x}(\tau_W), \epsilon_1) = \{x \mid \|x - \hat{x}(\tau_W)\| \leq \epsilon_1\}.$$

From the definition of $\delta_{\inf}$, it follows that

$$B(\hat{x}(\tau_W), \epsilon_1) \subset \left\{ x \,\Big|\, \left\langle \hat{x}, \frac{x}{\|x\|} \right\rangle \geq 1 - \delta_{\inf} \right\} \tag{2.10}$$

and that for $\epsilon > \epsilon_1$

$$B(\hat{x}(\tau_W), \epsilon) \not\subset \left\{ x \,\Big|\, \left\langle \hat{x}, \frac{x}{\|x\|} \right\rangle \geq 1 - \delta_{\inf} \right\}. \tag{2.11}$$

The inequalities (2.6) and $\|x^{\bar{i}_0}\| \geq \|x^{i_0}\|$ imply $x^{\bar{i}_0} \notin B(\hat{x}(\tau_W), \epsilon_1)$. The inclusions (2.9), (2.10), and (2.11) imply that

$$(0, x^{\bar{i}_0}) \cap B(\hat{x}(\tau_W), \epsilon_1) \neq \emptyset.$$

The latter relation can be written equivalently as

$$\exists \tau_0 \in (0,1), \quad \text{such that } x_0 = \tau_0 x^{\bar{i}_0}, \qquad \|x_0 - \hat{x}(\tau_W)\| \leq \epsilon_1.$$

So, relation (2.3) yields

$$f(x_0) \geq W - 1. \tag{2.12}$$

From the convexity of f we obtain

$$f\left(\tau_0 x^{\bar{i}_0}\right) \leq (1 - \tau_0) f(0) + \tau_0 f\left(x^{\bar{i}_0}\right). \tag{2.13}$$

Let us first consider the case when $f_{\inf} = -\infty$. Taking into account Eqs. (2.7), (2.2), and (2.12) in the inequality (2.13) results in the following double inequality

$$W - 1 \leq f(x_0) \leq (1 - \tau_0)(W - 2),$$

which yields $\tau_0 < 0$. The latter conclusion is inconsistent with our earlier assumption that $\tau_0 \in (0, 1)$. This ends the proof of the lemma in the case when $f_{\inf} = -\infty$.

Now suppose that $f_{\inf}$ is finite. Taking into account Eqs. (2.8), (2.2), and (2.12) in the inequality (2.13) results in the following inequality:

$$W - 1 \leq f(x_0) \leq (1 - \tau_0)(W - 2) + \tau_0 (f_{\inf} + 1).$$

Simple calculations along with the use of the inequality (2.2) yields the contradiction $W - 1 \leq W - 2$, which establishes the proof of Lemma 2.3. ∎

LEMMA 2.4. *Let f_i: $R^n \to R$, $i \in I \cup \{0\}$ be convex functions. If there exists an unbounded sequence $\{x^i\} \in \mathscr{R}$, such that*

$$\limsup_{i \to \infty} f(x^i) < \infty, \tag{2.14}$$

then an arbitrary cluster point of the sequence $\{x^i/\|x^i\|\}$ belongs to $\bigcap_{i=0}^m 0^+ f_i$.

Proof. We will prove the lemma by contradiction. Let $\{x^i\}$ satisfy $\lim_{i \to \infty}(x^i/\|x^i\|) = \hat{x}$. We note that $\{x^i\} \in \mathscr{R}$ implies that $\limsup_{i \to \infty} f_j(x^i) < \infty$, $j \in I$. Therefore all functions f_j, $j \in I \cup \{0\}$, satisfy hypotheses of Lemma 2.3 and consequently $\hat{x} \in \bigcap_{j=0}^m 0^+ f_j$. ∎

Remark 1. We note that the result of Lemma 2.4 may be used to generate a vector $p \in D_f^a \cap C$ assuming that the sequence $\{x_i\} \in \mathscr{R}$ along which the objective function decrease is given.

From now on we assume that y_i, $i = 1, 2, \ldots, n - 1$, is the orthonormal basis for the subspace $\nabla f(x_0)^{\perp}$, which will be denoted by Y. The symbol Y will be also used to denote the matrix in $R^{(n-1)\times n}$ formed by the vectors y_i, $i = 1, \ldots, n - 1$, and $\xi = (\xi_1, \ldots, \xi_{n-1})^T$.

LEMMA 2.5. *Let us assume that f is convex and $f \in C^\infty$.*

(i) *Then the function*

$$G_t(\xi) = f(x_0 + tp + Y^T\xi) \tag{2.15}$$

achieves a minimum over the region

$$\mathscr{R}_t = \{\xi \mid G_t^i(\xi) = f_i(x_0 + tp + Y^T\xi) \le 0, \qquad i = 1, \dots, m\} \tag{2.16}$$

for any $t \ge 0$.

If in addition $D_f^= = \varnothing$, then the function $G_t(\xi)$ achieves a unique minimum over $\mathscr{R}_t$.

(ii) *If $C \cap D_f^= = \varnothing$, then the set of the optimal solutions to the problem* (2.15)–(2.16) *is bounded.*

(iii) *If $\xi(t)$ is an optimal solution to the problem* (2.15)–(2.16), *then*

$$\lim_{t \to \infty} G_t(\xi(t)) = \inf_{x \in \mathscr{R}} f(x).$$

Proof. (i) We observe that C_t, the cone of recession of the problem (2.15)–(2.16), is given by $C_t = \{\xi \in R^{n-1} | Y^T\xi \in C\}$. Since $y_i \perp \nabla f(x_0)$, $\forall i = 1, \dots, n-1$, then $Y^T\xi \in C$ implies $Y^T\xi \in D_f^=$ and $\xi \in D_G^=$. The latter conclusion can be written as $C_t \subset D_G^=$, so from Lemma 2.2 we obtain that the problem (2.15)–(2.16) attains its minimum. We note that this part of the lemma remains valid without assumption that $f \in C^\infty$. This is because the Lemma 2.2 can be easily restated without assumption that f is an analytic function.

(ii) Part (i) of the lemma implies that the function $G_t(\xi)$ achieves its minimum over the set $\mathscr{R}_t$, at some point $\bar{\xi}^t$. We will show that the set T of the optimal solutions to this problem is bounded. If not, then the convexity of the set T implies that there exists a half-line $\eta(\tau) = \bar{\xi}^t + \eta\tau$, $\tau \ge 0$, $\eta \in \bigcap_{i=1}^m 0^+G_t^i$, contained in $T \subset \mathscr{R}_t$. Thus the relations $Y^T\eta \in \bigcap_{i=1}^m 0^+f_i$ and $C \cap D_f^= = \varnothing$ imply that either $Y^T\eta \in D_f^u \cup D_f^a$ or $Y^T\eta \notin 0^+f$. If the first case holds, then $f(x)$ strictly decreases along the half-line $x(\tau) = x_0 + tp + Y^T\eta(\tau)$, which contradicts the assumption that $\bar{\xi}^t \in T$. If $Y^T\eta \notin 0^+f$ then the function $f(x)$ strictly increases along this half-line, which contradicts the inclusion $\eta(\tau) \in T$, $\tau \ge 0$.

(iii) We will show first that $f(x_0 + tp + Y^T\xi(t))$ monotonically decreases with respect to t. Let $t_2 > t_1$. Then, since $p \in D_f^a$, we have

$$f(x_0 + t_1 p + Y^T\xi(t_1)) > f(x_0 + t_2 p + Y^T\xi(t_1)) \ge f(x_0 + t_2 p + Y^T\xi(t_2)).$$

Suppose that the sequence $\{x_i\} \in \mathscr{R}$ is minimizing for $f(x)$, that is,

$$\lim_{i \to \infty} f(x_i) = \inf_{x \in \mathscr{R}} f(x). \tag{2.17}$$

Each x_k can be represented uniquely as $x_k = x_0 + t_k p + Y^T \xi^k$, with $t_k \in R$.

We will show that there exists a subsequence $\{t_{k_j}\}$ of the sequence $\{t_k\}$ such that $\{t_{k_j}\} \to \infty$. Let us suppose that the opposite is true, that is, $M = \sup_k t_k < \infty$. Since $p \in 0^+ f$, then

$$f\left(x_0 + t_k p + Y^T \xi^k\right) \geq f\left(x_0 + Mp + Y^T \xi^k\right) \geq f\left(x_0 + Mp + Y^T \xi(M)\right), \quad \forall k,$$

where $\xi(M) = \operatorname{argmin}_{\xi \in \mathscr{R}_M} f(x_0 + Mp + Y^T \xi)$. Taking $k \to \infty$ yields $\lim_{k \to \infty} f(x_0 + t_k p + Y^T \xi^k) \geq f(x_0 + Mp + Y^T \xi(M))$. Since $p \in D_f^a$, then $f(x_0 + Mp + Y^T \xi(M)) > f(x_0 + (M+1)p + Y^T \xi(M))$. The latter two inequalities and the equation (2.17) give

$$\inf_{x \in \mathscr{R}} f(x) > f\left(x_0 + (M+1)p + Y^T \xi(M)\right).$$

Since $p \in C$, then $x_0 + (M+1)p + Y^T \xi(M)$ belongs to $\mathscr{R}_{M+1} \subset \mathscr{R}$, which contradicts the definition of the infimum of $f(x)$ over $\mathscr{R}$. This completes the proof of the statement that there exists $\{t_{k_j}\} \to \infty$. Without loss of generality, we will use the symbol $\{t_k\}$ to denote the subsequence $\{t_{k_j}\}$ divergent to $+\infty$. It follows from part (i) of the lemma that the problem

$$\min\left\{G_{t_k}(\xi) \mid \xi \in \mathscr{R}_{t_k}\right\},$$

where $G_{t_k}(\xi)$ and $\mathscr{R}_{t_k}$ are defined in (2.15) and (2.16), respectively, has a solution for every t_k. Let us define $\xi(t_k) = \operatorname{argmin}\{G_{t_k} \mid \xi \in \mathscr{R}_{t_k}\}$. We observe that $x_k = x_0 + t_k p + Y^T \xi^k \subset \mathscr{R}_{t_k}$ implies

$$f\left(x_0 + t_k p + Y^T \xi(t_k)\right) \leq f(x_k), \qquad \forall t_k \geq 0. \tag{2.18}$$

Taking the limit with $k \to \infty$, of both sides of the inequality (2.18), proves that

$$\inf_{x \in \mathscr{R}} f(x) = \lim_{k \to \infty} G_{t_k}(\xi(t_k)),$$

which along with the fact that $G_t(\xi(t))$ strictly decreases with respect to t completes the proof of the lemma. ∎

3. INFIMAL TRAJECTORY

The problem of minimizing the function (2.15) with respect to the constraints (2.16) is clearly a parametric program. We observe that to find an approximate infimum of $f(x)$, it is sufficient to solve the problem (2.15)–(2.16) for a sufficiently large value of the parameter t. We will prove that the sequence of the optimal solutions to the problem (2.15)–(2.16) can be approximated by the points on the trajectory $C(t)$ along which the logarithmic penalty function converges to the global infimum of $f(x)$ with respect to the set $\mathscr{R}$.

Our analysis of the infimal trajectory will be based upon the unconstrained minimization technique which exploits the logarithmic penalty function. Let us define the functions

$$L(t, \xi, r) = G_t(\xi) - r \sum_{i=1}^{m} \ln\left(-G_t^i(\xi)\right) \tag{3.1}$$

and

$$\mathscr{L}(t, \xi) = L(t, \xi, t) = G_t(\xi) - \frac{1}{t} \sum_{i=1}^{m} \ln\left(-G_t^i(\xi)\right). \tag{3.2}$$

LEMMA 3.1. *Assume that the functions* f_i: $R^n \to R$, $i \in I \cup \{0\}$ *are convex and analytic and* $C \cap D_f^= = \emptyset$. *Let us define the function*

$$F(t) = \min_{\xi \in R^{n-1}} \{\mathscr{L}(t, \xi)\}, \qquad t \in R_+. \tag{3.3}$$

Then

(i) $F(t)$ *is well defined, that is,* $F(t) > -\infty$, $\forall t \geq 0$, *and the minimum of the function on the right side of* (3.3) *exists and is unique for every* $t \geq 0$.

(ii) $F(t)$ *is a strictly monotonically decreasing function of* t, *for* $t \geq 0$.

Proof. (i) Lemma 2.5 (ii) assures that the set of points that solve the problem of minimizing of $G_t(\xi)$ subject to $\xi \in \mathscr{R}_t$ is bounded (in particular, if $D_f^= = \emptyset$, then part (i) of the lemma states that the set of optimal solutions consists of one point only); thus the conditions of Theorem 25 in [5] are satisfied. Part (i) of this theorem implies that the function $L(t, \xi, r)$ has a finite unconstrained minimum $\xi(r)$ with respect to ξ for every $r > 0$, and $t > 0$. Letting $r = \frac{1}{t}$ proves the existence of the minimum of $\mathscr{L}(t, \xi)$, for every $t \geq 0$.

To prove the uniqueness of the solution we will show that the Hessian matrix of $\mathscr{L}(t, \xi)$ is strictly positive definite. To this end we will use the result proved in [12], that any analytic convex function $f_i(x)$ can be represented as

$$f_i(x) = F_i(c_i + B_i x) + \langle a_i, x \rangle + d_i,$$

where F_i is strictly convex analytic function, $B_i \in R^{p_i \times n}$, $c_i \in R^{p_i}$, and $a_i \in R^n$. It has also been shown in [12] that

$$D_{f_i}^{=} = \mathscr{N}\begin{pmatrix} B_i \\ a_i^T \end{pmatrix},$$

where $\mathscr{N}(\cdot)$ denotes the nullspace of the matrix $(\cdot)$. It follows directly that the Hessian of $\mathscr{L}(t, \xi)$ is given by the formula

$$\begin{aligned} &\nabla_\xi^2 \mathscr{L}(t, \xi) \\ &\quad = YB^T \nabla_y^2 F(c + Bx) BY^T - \frac{1}{t} \sum_{i=1}^{m} \frac{YB_i^T \nabla_{y^i}^2 F_i(c_i + B_i x) B_i Y^T}{f_i(x)} \\ &\qquad + \frac{1}{t} \sum_{i=1}^{m} \frac{\left(\nabla_{y^i} F_i(c_i + B_i x) B_i Y^T + Y a_i\right)^T \left(\nabla_{y^i} F_i(c_i + B_i x) B_i Y^T + Y a_i\right)}{f_i^2(x)}, \end{aligned}$$

where $x = x_0 + pt + Y^T \xi$, and $y^i \in R^{p_i}$, $i \in I \cup \{0\}$. Since $f_i(x) \leq 0$, $\forall x \in \mathscr{R}$, $i \in I$, then all three terms in the expression on the right side of the latter equation are nonnegative. Thus the Hessian matrix $\nabla_\xi^2 \mathscr{L}(t, \xi)$ is not strictly definite (semidefinite) only if there exists a nonzero vector $\bar{\xi}$ satisfying the system of equations

$$B_i Y^T \bar{\xi} = 0, \quad \forall i \in I \cup \{0\}, \qquad \left[\nabla_{y^i} F_i(c_i + B_i x) B_i Y^T + Y a_i\right] \bar{\xi} = 0, \quad \forall i \in I.$$

Note that equation $B_i Y^T \bar{\xi} = 0$ implies that $a_i Y^T \bar{\xi} = 0$, $\forall i$, so

$$Y^T \bar{\xi} \in \bigcap_{i=1}^{m} \mathscr{N}\begin{pmatrix} B_i \\ a_i^T \end{pmatrix} = \bigcap_{i=1}^{m} D_{f_i}^{=} . \tag{3.4}$$

Since the raws of the matrix Y are orthogonal to $\nabla f(x_0)$, then $[\nabla_y F(c + Bx_0)B + a]^T Y^T \bar{\xi} = 0$, which along with the equation $BY^T \bar{\xi} = 0$ yields $a^T Y^T \bar{\xi} = 0$. Thus $Y^T \bar{\xi} \in D_f^{=}$, which along with (3.4) contradicts the assumption that $C \cap D_f^{=} = \varnothing$, and establishes part (i) of the lemma.

(ii) We will show that $F(t_2) < F(t_1)$ if $t_2 > t_1$. It follows from Theorem 25 (vi) in [5] that

$$\min_{x \in R^{n-1}} L\left(t_1, \xi, \frac{1}{t_1}\right) > \min_{\xi \in R^{n-1}} L\left(t_1, \xi, \frac{1}{t_2}\right).$$

Let ξ^1, ξ^2, and ξ^3 be such that

$$L\left(t_1, \xi^1, \frac{1}{t_1}\right) = \min_{\xi \in R^{n-1}} L\left(t_1, \xi, \frac{1}{t_1}\right), \tag{3.5}$$

$$L\left(t_1, \xi^2, \frac{1}{t_2}\right) = \min_{\xi \in R^{n-1}} L\left(t_1, \xi, \frac{1}{t_2}\right) \tag{3.6}$$

and

$$L\left(t_2, \xi^3, \frac{1}{t_2}\right) = \min_{\xi \in R^{n-1}} L\left(t_2, \xi, \frac{1}{t_2}\right). \tag{3.7}$$

We will prove first that

$$L\left(t_2, \xi^2, \frac{1}{t_2}\right) < L\left(t_1, \xi^2, \frac{1}{t_2}\right). \tag{3.8}$$

To this end we will show that

$$\begin{aligned} L\left(t_2, \xi^2, \frac{1}{t_2}\right) &= f\left(x_0 + t_2 p + Y^T\xi^2\right) - \frac{1}{t_2}\sum_{i=1}^{m} \ln\left(-f_i\left(x_0 + t_2 p + Y^T\xi^2\right)\right) \\ &< f\left(x_0 + t_1 p + Y^T\xi^2\right) - \frac{1}{t_2}\sum_{i=1}^{m} \ln\left(-f_i\left(x_0 + t_1 p + Y^T\xi^2\right)\right) \\ &= L\left(t_1, \xi^2, \frac{1}{t_2}\right). \end{aligned} \tag{3.9}$$

We observe that assumption that $p \in C \cap D_f^a$ implies

$$f\left(x_0 + t_2 p + Y^T\xi^2\right) < f\left(x_0 + t_1 p + Y^T\xi^2\right) \tag{3.10}$$

and

$$f_j\left(x_0 + t_2 p + Y^T\xi^2\right) \leq f_j\left(x_0 + t_1 p + Y^T\xi^2\right), \tag{3.11}$$

$i = 1, 2, \ldots, m$. Furthermore, the equation (3.6) and Theorem 25 (i) in [5] imply that $\xi^2 \in \operatorname{int} \mathscr{R}_{t_1}$, which yields

$$f_j\left(x_0 + t_1 p + Y^T\xi^2\right) < 0, \qquad j = 1, \ldots, m. \tag{3.12}$$

The inequalities (3.10)–(3.12) prove that inequality (3.9), and consequently inequality (3.8), holds. Now, by Theorem 25 (vi) in [5], we have

$$L\left(t_1, \xi^2, \frac{1}{t_2}\right) < L\left(t_1, \xi^1, \frac{1}{t_1}\right). \tag{3.13}$$

The definition of ξ^3 gives

$$L\left(t_2, \xi^3, \frac{1}{t_2}\right) \leq L\left(t_2, \xi^2, \frac{1}{t_2}\right). \tag{3.14}$$

Relations (3.3), (3.5), and (3.7) yield

$$F(t_1) = \min_{\xi \in R^{n-1}} L\left(t_1, \xi, \frac{1}{t_1}\right) = L\left(t_1, \xi^1, \frac{1}{t_1}\right),$$

$$F(t_2) = \min_{\xi \in R^{n-1}} L\left(t_2, \xi, \frac{1}{t_2}\right) = L\left(t_2, \xi^3, \frac{1}{t_2}\right).$$

Now inequalities (3.14), (3.8), and (3.13) imply

$$L\left(t_2, \xi^3, \frac{1}{t_2}\right) < L\left(t_1, \xi^1, \frac{1}{t_1}\right),$$

which can be written equivalently as $F(t_2) < F(t_1)$. ∎

We note that the equality

$$F(t) = \mathscr{L}(t, \xi) = G_t(\xi) - \frac{1}{t} \sum_{i=1}^{m} \ln\left(-G_t^i(\xi)\right)$$

holds if and only if ξ satisfies the equation

$$\nabla_\xi \mathscr{L}(t, \xi) = Y \nabla_x f\left(x_0 + tp + Y^T \xi\right) - \frac{1}{t} \sum_{i=1}^{m} \frac{Y \nabla_x f_i\left(x_0 + tp + Y^T \xi\right)}{f_i\left(x_0 + tp + Y^T \xi\right)} = 0. \tag{3.15}$$

For every value of t the equation (3.15) consists of $n-1$ equations with $n-1$ unknown variables $\xi_1, \dots, \xi_{n-1}$.

In the remaining part of the paper we assume that the function $\mathscr{L}(t, \xi)$ satisfies the following second order sufficiency condition

$$\forall t \in R_+, \quad \forall \bar{\xi} \in R^n, \quad \nabla_\xi \mathscr{L}\left(t, \bar{\xi}\right) = 0 \quad \Rightarrow \quad \xi^T \nabla_\xi^2 \mathscr{L}\left(t, \bar{\xi}\right) \xi > 0, \quad \forall \xi \neq 0,$$

which can be expressed in terms of the function $f(x)$ and its derivatives as

$$\forall y, \qquad \bar{x} \in R^n, \qquad y \neq 0, \qquad y^T\left[\nabla_x f(\bar{x}) - \frac{1}{t}\sum_{i=1}^{m}\frac{\nabla_x f_i(\bar{x})}{f_i(\bar{x})}\right] = 0 \quad \Rightarrow$$

$$y^T\left[\nabla_x^2 f(\bar{x}) - \frac{1}{t}\sum_{i=1}^{m}\frac{\nabla_x^2 f_i(\bar{x}) f_i(\bar{x}) - \nabla_x f_i(\bar{x})\nabla_x f_i(\bar{x})^T}{f_i^2(\bar{x})}\right] y > 0, \tag{3.16}$$

where $\bar{x} = x_0 + tp + Y^T\xi$.

We note that the Lemma 3.1 remains valid if the assumption that $D_f^= \cap C = \emptyset$ is replaced with the less restrictive condition (3.16).

In the theorem below we will prove that there exists an analytic function $x(t) = C(t)$, such that the infimum of $f(x)$ along the curve is equal to the global infimum of $f(x)$ over $\mathscr{R}$.

THEOREM 3.1. (i) *If f_i, $i \in I \cup \{0\}$ are convex and analytic, then there exists a unique analytic function $h(t)$, which is a trajectory of unconstrained minima of $\mathscr{L}(t, \xi)$. Furthermore, derivatives of all orders of $h(t)$ can be obtained by subsequent differentiation of $\nabla_\xi \mathscr{L}(t, h(t)) = 0$ with respect to t.*

(ii) *The parametrized curve $C(t) = x_0 + tp + Y^T h(t)$ satisfies*

$$\lim_{t\to\infty} F(t) = \lim_{t\to\infty}\left[f(C(t)) - \frac{1}{t}\sum_{i=1}^{m}\ln\left(-f_i(C(t))\right)\right] = \lim_{t\to\infty}\mathscr{L}(t, h(t))$$
$$= \inf_{x\in\mathscr{R}} f(x). \tag{3.17}$$

Proof. (i) The condition (3.16) implies that the Hessian matrix of $\mathscr{L}(t, \xi)$ is positive definite at the point (t, ξ) satisfying Eq. (3.15). Since the system of $(n-1)$ equations in n variables $\xi_1, \ldots, \xi_{n-1}, t$ in (3.15) has an invertible Jacobian matrix $\nabla_\xi^2 \mathscr{L}(t, \xi)$, then by the implicit function theorem [13] there exist open sets $U \in R^n$ and $W \subset R_+$, with $(\bar{\xi}, \bar{t}) \in U$ and $\bar{t} \in W$, having the following property: To every $t \in W$, there corresponds a unique ξ such that $(\xi, t) \in U$ and $\nabla_\xi \mathscr{L}(t, \xi) = 0$. Furthermore, if this ξ is defined to be $h(t)$, then h is a C^1-mapping of W into R^{n-1}, $h(\bar{t}) = \bar{\xi}$, and

$$\nabla_\xi \mathscr{L}(t, h(t)) = 0, \qquad \forall t \in W.$$

Differentiating the latter equation with respect to t yields

$$Y\nabla_x^2 f(h_0(t))\left(Y^T h'(t) + p\right) + \frac{1}{t^2}\sum_{i=1}^{m}\frac{Y\nabla_x f_i(h_0(t))}{f_i(h_0(t))}$$
$$-\frac{1}{t}Y\left[\sum_{i=1}^{m}\frac{\nabla_x^2 f_i(h_0(t)) f_i(h_0(t)) - \nabla_x f_i(h_0(t))\nabla_x f_i(h_0(t))^T}{f_i^2(h_0(t))}\right]$$
$$\times\left(Y^T h'(t) + p\right) = 0, \tag{3.18}$$

where $h_0(t) = x_0 + tp + Y^T h(t)$. By the assumption (3.16) the matrix

$$\nabla_\xi^2 \mathscr{L}(t, \xi) = Y\Bigg[\nabla_x^2 f_i(h_0(t)) - \frac{1}{t}\sum_{i=1}^{m} \frac{\nabla_x^2 f_i(h_0(t)) f_i(h_0(t)) - \nabla_x f(h_0(t))\,\nabla_x f_i(h_0(t))^T}{f_i^2(h_0(t))}\Bigg] Y^T,$$

has an inverse, so the equation (3.18) can be solved for $h'(t)$

$$h'(t) = -\left(\nabla_\xi^2 \mathscr{L}(t, \xi)\right)^{-1}\Bigg[Y\nabla_x^2 f(h_0(t)) - \frac{1}{t}Y\Bigg[\sum_{i=1}^{m} \frac{\nabla_x^2 f_i(h_0(t)) f_i(h_0(t)) - \nabla_x f_i(h_0(t))^T\,\nabla_x f_i(h_0(t))}{f_i^2(h_0(t))}\Bigg] p + \frac{1}{t^2} Y \sum_{i=1}^{m} \frac{\nabla_x f_i(h_0(t))}{f_i(h_0(t))}\Bigg].$$

This proves that the derivative of $h(t)$ with respect to t exists for $t \in W$ and is a C^1-mapping of W into R^n. (We note that the existence and continuity of $h'(t)$ follows also from the implicit function theorem; however, for the completeness of the proof we include this argument as well.)

Now, differentiation of Eq. (3.18) with respect to t yields

$$\frac{d\nabla_\xi^2 \mathscr{L}(t, h(t))}{dt} h'(t) + \nabla_\xi^2 \mathscr{L}(t, h(t)) h''(t) + \frac{d\left[Y\nabla_x^2 f(h_0(t))\right]}{dt} - \frac{1}{t}\Bigg[d\Bigg[Y\sum_{i=1}^{m} \frac{\nabla_x^2 f_i(h_0(t)) f_i(h_0(t)) - \nabla_x f_i(h_0(t))^T\,\nabla_x f_i(h_0(t))}{f_i^2(h_0(t))} p\Bigg]\Big/dt\Bigg] + d\Bigg[\frac{1}{t^2}\sum_{i=1}^{m} \frac{Y\nabla_x f_i(h_0(t))}{f_i(h_0(t))}\Bigg]\Big/dt = 0$$

The existence of $h''(t)$ is assured by the existence of the inverse of the Hessian matrix $\nabla_\xi^2 \mathscr{L}(t, h(t))$. Continuing the above process it is possible to obtain explicitly all derivatives $h^{(k)}(t)$ in terms of the derivatives $h^{(i)}$ ($i = 1, \dots, k-1$), and partial derivatives of the functions f_j, $j \in I \cup \{0\}$, of degrees $i = 1, \dots, k+1$. Since $f_j \in C^\infty$, then part (i) of the theorem follows.

(ii) In Lemma 2.5 (ii) we proved that the set of optimal solutions to the problem (2.15)–(2.16) is bounded for every t. Thus the assumptions of Theorem 25 (iv) of [5] are satisfied, which yields that for any fixed $t_k > 0$

$$\lim_{t\to\infty} L\left(t_k, \xi^k(t), \frac{1}{t}\right) = \min_{\xi\in\mathscr{R}_{t_k}} G_{t_k}(\xi), \tag{3.19}$$

where $\xi^k(t)$ denotes the point of unconstrained minimum of $L(t_k, \xi, \frac{1}{t})$ with respect to ξ. Lemma 2.5 (iii) shows that

$$\lim_{k\to\infty} f\left(x_0 + t_k p + Y^T\xi(t_k)\right) = \inf_{x\in\mathscr{R}} f(x),$$

where $\xi(t)$ is an optimal solution to the problem (2.15)–(2.16).

It follows from the latter equation that

$$\inf_{x\in\mathscr{R}} f(x) = \lim_{k\to\infty} \min_{\xi\in\mathscr{R}_{t_k}} G_{t_k}(\xi), \tag{3.20}$$

which, using Eq. (3.19), yields

$$\lim_{k\to\infty} \lim_{t\to\infty} L\left(t_k, \xi^k(t), \frac{1}{t}\right) = \inf_{x\in\mathscr{R}} f(x).$$

It follows from the equation (3.20) that

$$\forall\epsilon > 0, \qquad \exists\bar{k}, \qquad \forall k \geq \bar{k}, \qquad 0 < \min_{\xi\in\mathscr{R}_{t_k}} G_{t_k} - \inf_{x\in\mathscr{R}} f(x) < \epsilon. \tag{3.21}$$

Equation (3.19), along with the fact that for every fixed t_k, $L(t_k, \xi^k(t), \frac{1}{t})$ is monotonically decreasing with respect to t (see Theorem 25 (vi) in [5]), yields

$$\forall\delta > 0, \qquad \exists t_\delta, \qquad \forall t > t_\delta, \qquad 0 \leq L\left(t_k, \xi^k(t), \frac{1}{t}\right) - \min_{\xi\in\mathscr{R}_{t_k}} G_{t_k} < \delta. \tag{3.22}$$

Taking $\delta = \epsilon$ in (3.22) and adding inequalities in (3.21) and (3.22) gives

$$\forall\epsilon > 0, \exists\bar{k}, \exists t_\epsilon, \forall k \geq \bar{k}, \forall t \geq t_\epsilon, 0 \leq L\left(t_k, \xi^k(t), \frac{1}{t}\right) - \inf_{x\in\mathscr{R}} f(x) < 2\epsilon.$$

If for given $\epsilon > 0$, $t_{\bar{k}} > t_\epsilon$, then we substitute $t_\epsilon := t_{\bar{k}}$; otherwise, i.e., if $t_{\bar{k}} < t_\epsilon$, we find the smallest index $\tilde{k}$, such that $t_{\tilde{k}} > t_\epsilon$, and take $\bar{k} := \tilde{k}$.

The latter modification leads to the relation

$$\forall \epsilon > 0, \qquad \exists \bar{k}, \qquad \forall k \geq \bar{k}, \qquad 0 \leq L\left(t_k, \xi^k(t_k), \frac{1}{t_k}\right) - \inf_{x \in \mathscr{R}} f(x) < 2\epsilon,$$

which proves that

$$\lim_{k \to \infty} \mathscr{L}\big(t_k, \xi^k(t_k)\big) = \inf_{x \in \mathscr{R}} f(x).$$

Since $F(t_k) = \mathscr{L}(t_k, \xi^k(t_k))$ and $F(t)$ is a strictly decreasing function of t (see Lemma 3.1(ii)), and $\lim_{k \to \infty} t_k = \infty$, then $\lim_{k \to \infty} F(t_k) = \lim_{t \to \infty} F(t)$, which implies

$$\lim_{t \to \infty} F(t) = \inf_{x \in \mathscr{R}} f(x).$$

This completes the proof of the equalities in (3.17). ∎

In the lemma below we will show that the curve $x = C(t)$ does not depend on the choice of the vector p, as long as the vector p is in the cone of recession C and it belongs to D_f^a.

Let

$$H_i(t) = \left\{x \in R^n | x = x_0 + tp_i + Y^T\xi\right\}, \qquad i = 1, 2, \qquad t \geq 0.$$

Let $C_1(t)$ and $C_2(t)$ denote two parametrized curves constructed for two different vectors $p_1, p_2 \in C \cap D_f^a$. It follows that

$$C_i(t) = x_0 + tp_i + Y^T h^i(t), \qquad i = 1, 2, \tag{3.23}$$

where

$$h^i(t) = \operatorname{argmin}\left\{\mathscr{L}_i(t, \xi) | \xi \in H_i(t) \cap \operatorname{int} \mathscr{R}\right\} = \operatorname{argmin}\left\{\mathscr{L}_i(t, \xi) | \xi \in R^{n-1}\right\}, \tag{3.24}$$

where

$$\mathscr{L}_i(t, \xi) = f\big(x_0 + tp_i + Y^T\xi\big) - \frac{1}{t} \sum_{j=1}^{m} \ln\big(-f_j\big(x_0 + tp_i + Y^T\xi\big)\big).$$

LEMMA 3.2. (i) *If $p_1, p_2 \in C \cap D_f^a$, then $C_1(t)$ and $C_2(t)$ defined in* (3.23)–(3.24) *are equivalent representations of the same parametrized curve.*

(ii) *For a given* $x_0 \in \mathscr{R}$ *and* $p \in C \cap D_f^a$ *there is exactly one curve satisfying* $C(t) = x_0 + tp + Y^T h(t) \in \mathscr{R}$ *where*

$$h(t) = \operatorname{argmin}\{\mathscr{L}(t, \xi) \mid \xi \in H_i(t) \cap \operatorname{int} \mathscr{R}\} = \operatorname{argmin}\{\mathscr{L}(t, \xi) \mid \xi \in R^{n-1}\},$$

with $H(t) = \{x \in R^n | x = x_0 + tp + Y^T\xi,\ \xi \in R^{n-1}\}$, $t \geq 0$, *passing through* x_0. *The curve satisfies Eq.* (3.17).

Proof. (i) Note that for a fixed value of t, $H_i(t)$, $i = 1, 2$, represents a linear manifold and

$$\forall t_1 \in R_+, \qquad \exists t_2 \in R_+, \qquad H_1(t_1) = H_2(t_2). \tag{3.25}$$

Equalities (3.23)–(3.25) yield

$$C_1(t_1) = C_2(t_2)$$

so $C_1(t)$ and $C_2(t)$ are equivalent representations of the same parametrized curve. Since relation (3.25) is symmetric with respect to t_1 and t_2, it follows that there is a unique correspondence between the points on the curves $C_1(t)$ and $C_2(t)$. This ends the proof of part (i) of the lemma.

(ii) Lemma 2.1 (ii) implies that $f(x)$ is strictly decreasing along a half-line $x_0(t) = x_0 + tp$, $t \geq 0$, for any $x_0 \in R^n$. Therefore replacing the initial point x_0 with another point x_1 requires one to redefine the vectors $y_1, \ldots, y_{n-1}$ as the orthonormal basis for the subspace $\{\nabla f(x_1)\}^\perp$. Thus all results proved for x_0 remain valid for the function $G_t(\xi) = f(x_1 + tp + Y^T\xi)$ and respectively modified functions $\mathscr{L}(t, \xi)$, $F(t)$ and $C(t)$, and the linear manifold $H(t)$. Furthermore, since $C \cap D_f^= = \varnothing$, Lemma 3.1 implies that the curve $C(t)$ is uniquely defined. ∎

4. EXAMPLE AND CONCLUDING REMARKS

We will illustrate the method proposed in Section 3 with the example of the problem with two constraints in two variables.

Let us consider the problem

$$\min f(x) = \frac{1}{x_1 + 1}$$

subject to

$$f_1(x) = -x_1 + x_2 \leq 1$$

$$f_2(x) = -x_2 \leq -2.$$

We have $p = (1,0)^T$, $x_0 = (0,0)^T$, $\nabla f(0) = (-1,0)^T$, $Y = (0,1)$. Therefore,

$$\min_{\mathrm{g}x \in R^{n-1}} \mathscr{L}(t, \xi) = \min_{\xi} \left(\frac{1}{t+1} - \frac{1}{t}\ln(1 + t - \xi) - \frac{1}{t}\ln(\xi - 2) \right),$$

$$\nabla_{\xi} \mathscr{L}(t, \xi) = \frac{1}{t(1 + t - \xi)} - \frac{1}{t(\xi - 2)} = 0.$$

The latter equation yields

$$\xi = h(t) = \frac{1}{2}(t + 3).$$

Finally $C(t) = (t, \frac{1}{2}(t + 3))^T$, which yields $\lim_{t \to \infty} \mathscr{L}(t, h(t)) = 0$ and consequently $\inf_{x \in \mathscr{R}}(1/(x_1 + 1)) = 0$.

Concluding, the aim of this paper is to show that for a convex program (1.1) with unattained infimum and for a given initial feasible point, there exists a unique parametrized curve $C(t)$, along which the logarithmic penalty function converges to the infimum of the function $f(x)$ over the region $\mathscr{R}$. We show that if all functions defining the program are analytic, then $C(t)$ is an analytic function as well. We also prove that an approximate solution to the degenerate program (1.1) can be determined by solving the auxiliary parametric problem, defined in Lemma 2.5, for a sufficiently large value of the parameter t.

The performance of the proposed method was investigated on a number of convex constrained problems with unattained minimum. The application of the classical methods, such as the steepest descent or the gradient projection methods, fails to provide a sequence convergent to the constrained infimum in these problems. The results obtained for all tested problems seem to be promising, which suggests that the method proposed in this paper deserves more study both from the theoretical and the computational points of view.

REFERENCES

1. R. A. Abrams, Projections of convex programs with unattained infima, *SIAM J. Control* **13**, No. 3 (1975), 706–718.
2. A. Auslender, R. Cominetti, and J.-P. Crouzeix, Convex functions with unbounded level sets and applications to duality theory, *SIAM J. Optim.* **3** (1993), 669–687.
3. R. J. Caron and W. T. Obuchowska, Unboundedness of a convex quadratic function subject to concave and convex quadratic constraints, *European J. Operational Res.* **63** (1992), 114–123.

4. R. J. Caron and W. T. Obuchowska, An algorithm to determine boundedness of quadratically constrained convex quadratic programmes, *European J. Operational Res.* **80** (1995), 431–438.
5. A. V. Fiacco and G. P. McCormick, "Nonlinear Programming; Sequential Unconstrained Minimization Techniques," Wiley, New York, 1968.
6. J.-B. Hiriart-Urutty and C. Lemarechal, "Convex Analysis and Minimization Algorithms I," Springer-Verlag, Berlin/New York, 1993.
7. M. Lamoureux and H. Wolkowicz, Numerical decomposition of a convex function, *J. Optim. Theory Appl.* **47** (1985), 51–64.
8. K. G. Murty, "Linear Complementarity and Nonlinear Programming," Helderman, West Berlin, 1988.
9. W. T. Obuchowska and K. G. Murty, Cone of recession and unboundedness of convex functions, submitted for publication.
10. W. T. Obuchowska, On the infimal trajectory of convex functions with unbounded level sets, submitted for publication.
11. R. T. Rockafeller, "Convex Analysis," Princeton Univ. Press, Princeton, NJ, 1979.
12. R. T. Rockafeller, Some convex programs whose duals are linearly constrained, *in* "Nonlinear Programming" (J. B. Rosen, O. L. Mangasarian, and K. Ritter, Eds.), Academic Press, New York, 1970.
13. W. Rudin, "Principles of Mathematical Analysis," McGraw-Hill, New York, 1976.