An. şt. Univ. „Al. I. Cuza" Iaşi
Informatică, 35 (1989)

CONVEX OPTIMIZATION BY THE UNCONSTRAINED MULTIPLICATIVE BARRIER FUNCTION METHOD*

BY

VINCENŢIU DUMITRU and FLORICA LUBAN

1. Introduction. Let be the following constrained minimization problem minimize $f(x)$

subject to

$$g_i(x) \geqslant 0, \quad i=1,\ldots,m, \tag{1}$$

where f and $-g_1,\ldots,-g_m$ are convex functions of class C^2, from the n-dimensional real Euclidian space R^n into R, and let be

$$X=\{x \in R^n \mid g_i(x) \geqslant 0,\ i=1,\ldots,m\}$$

$$\operatorname{Int} X=\{x \in R^n \mid g_i(x)>0,\ i=1,\ldots,m\}$$

$$X^*=\{x \in X \mid f(x)=\mu^*\},$$

where μ^* is a priori known optimum objective value of problem (1), that is $\mu^*=\{\min f(x) \mid x \in X\}$.

We assume that a) $\operatorname{Int} X \neq \emptyset$; b) $X^* \neq X$, $f(x)>\mu^*$ for $x \in \operatorname{Int} X$; c) X^* is a bounded set.

For solving the problem (1) as an unconstrained minimization problem, when $f(x)$ and $g_i(x)$, $i=1,\ldots,m$ are linear and the optimum objective value μ^* of this problem is available, I r i in [5] has proposed the convex multiplicative barrier function

$$F(x)=(f(x)-\mu^*)^{m+1} / \prod_{i=1}^{m} g_i(x) \tag{2}$$

defined for $x \in \operatorname{Int} X$.

It is readily seen that, under the above assumptions, we have $F(x)>0$ for $x \in \operatorname{Int} X$, and if the sequence $\{x^k\} \subset \operatorname{Int} X$ and $x^k \rightarrow y \in X-\operatorname{Int} X$, $y \notin X^*$, where $\|y\|<\infty$ then $\lim_{k \rightarrow \infty} F(x^k)=+\infty$.

In the following we shall study a generalized form of (2) for the convex problem (1). Before that, it is worth to mention that the I r i - I m a i line

---

* *The results of this paper were submitted for presentation to the 14-th Symposium on Operations Research, University of Ulm, September 1989 and delivered in a conference at the Info—Iaşi on the 21-th of October 1989.*

of research is not singular but in the movement generated by the question of Klee and Minty [11] on *How good is the simplex algorithm* (1972), where it has been shown that the simplex method can grow exponentially with the problem dimension in the worst case due to the combinatorial complexity of traversing the boundary of the feasible region. In the attempt to avoid the combinatorial complexities of vertex following methods, from the very early time von Neumann (1947), A. J. Hoffman et al. (1953), K. R. Frish (1957) have proposed methods which traverse across the interior of feasible region.

In 1979, Khachiyan [10] provided the first results in the attempt to develop a linear programming method with polynomial worst-case complexity, by showing how to adopt the 1970 ellipsoid method for convex optimization of Shor [14] to linear programming. In 1984, Karmarkar [9] gave a new polynomial-time algorithm for linear programming, which is an interior feasible direction method of minimizing a potential function by a kind of gradient projection method in a projectively transformed space. One year later, Iri and Imai [5], in 1985, have proposed a multiplicative penalty function method for solving large scale linear programming problems. The method, substantiated after in a series of papers [6], [7], [8] is a simple Newton-like descent algorithm for minimizing a multiplicative penalty convex function defined on the interior of the feasible region. The algorithm is shown to give local superlinear convergence to the optimum and, under some assumption, global linear convergence.

In this paper we are comming back on the path of Iri-Imai, from linear to nonlinear, in an attempt to use their ideas not to solve the linear programming problem but also any nonlinear convex problem of class C^2. Moreover we are able to show that the Iri-Imai multiplicative barrier function approach can be extended over a large range of parameters that covers also the earlier 1963 Huard's method of centres [4] for convex optimization. In this respect in [1] we introduced the generalized Huard-Iri function:

$$B(x)=(f(x)-\mu^k)^p\cdot\left(\prod_{i=1}^{m} g_i(x)\right)^q,\ \text{for } x\in \operatorname{Int} X \tag{3}$$

As we just have mentioned, for $p=1$ we obtain one of the distance function used iteratively by Huard in its method of centres [4] to maximize $(-f(x))$ over X, while for $p=m+1$, $q=-1$ we obtain the Iri's barrier function to solve in one iteration the linear programming problem.

2. Basic theorems. Let us define the function

$$\Phi(x)=(f(x)-\mu^*)^p/\left(\prod_{i=1}^{m} g_i(x)\right)^q \text{ for } x\in \operatorname{Int} X \text{ and } q>0,\ p>1+mq.$$

For this function, as in [7], we can prove the following basic propositions for the use of $\Phi(x)$ in solving convex problems.

Proposition 1. *If $\Phi(x^k)\to 0$ for $\{x^k\}\subset \operatorname{Int} X$ then the distance between x^k and X^* converges to 0 (and hence, if there is an unique optimum, the sequence converges to it).*

Proof. If the set of points $\{x^k\}$ is bounded, so are the $g_i(x^k)$'s. Therefore $\Phi(x^k)\to 0$ implies $(f(x^k)-\mu^*)\to 0$, and, as $X^*\neq X$ and $f(x)-\mu^*>0$ for $x\in \operatorname{Int} X$, the distance between x^k and X^* tends to 0.

Let us assume now that the set of points $\{x^k\}\subset \operatorname{Int} X$ is not bounded. Because $g_i(x^k)$ can become large, $(f(x^k)-\mu^*)$ does not converge to 0 even if $\Phi(x^k)\to 0$.

From now on the proof follows the proof lines of Lemma 12 from [3]. Let x^0 be any point in Int X and let $D=\{x|f(x)\leq f(x^0)+\varepsilon,\ x\in X\}$ for some $\varepsilon>0$, be a bounded set and hence a compact set [3].

By our assumption $X^*=\{x|f(x)\leq \mu^*=\inf f(x),\ x\in X\}$ is nonempty and bounded.

Consider the point x^k far enough out in the sequence so that $x^k\notin D$. Let z^k be the point on the boundary B of D where the line connecting x^0 and x^k intersects B. (Note that B is compact also). By definition of D, $x^k\neq x^0$. Let λ^k be the scalar where, for $0<\lambda^k<1$,

$$z^k=\lambda^k x^k+(1-\lambda^k)x^0\sim \lambda^k=\frac{\|z^k-x^0\|}{\|x^k-x^0\|}.$$

Clearly $\lambda^k\to 0$ as $k\to\infty$.

The functions $g_i(x)$, $i=1,\dots,m$, are concave functions of $x\in R^n$, thus

$$g_i(z^k)\geq \lambda^k g_i(x^k)+(1-\lambda^k)g_i(x^0)>0,\ i=1,\dots,m. \tag{4}$$

Hence $z^k\in$ Int X, but since z^k is a boundary point of D, it must lie on that portion of the boundary $B\subset$ Int X and such that $f(z^k)=f(x^0)+\varepsilon$.

But $f(x)$ is a convex function of $x\in R^n$:

$$f(z^k)\leq\lambda^k f(x^k)+(1-\lambda^k)f(x^0)\Rightarrow f(x^k)\geq\frac{f(z^k)-(1-\lambda^k)f(x^0)}{\lambda^k}=$$

$$=\frac{f(x^0)+\varepsilon-f(x^0)+\lambda^k f(x^0)}{\lambda^k}=f(x^0)+\frac{\varepsilon}{\lambda^k}.$$

Hence

$$f(x^k)-\mu^*\geq f(x^0)-\mu^*+\frac{\varepsilon}{\lambda^k}. \tag{5}$$

Since $\varepsilon>0$, $(f(x^k)-\mu^*)\to+\infty$ as $\lambda^k\to 0$ $(k\to\infty)$. Because $x^k\in$ Int X and from (4) we have

$$0<g_i(x^k)\leq\frac{g_i(z^k)-(1-\lambda^k)g_i(x^0)}{\lambda^k}=g_i(x^0)+\frac{g_i(z^k)-g_i(x^0)}{\lambda^k}\Rightarrow$$

$$\Rightarrow g_i(z^k)-g_i(x^0)>0,\ i=1,\dots m.$$

Let us denote $0<d_i^0=\max\limits_{z^k\in B}(g_i(x^k)-g_i(x^0))$, $i=1,\dots,m$. (This exists because B is compact). Hence

$$0<g_i(x^k)\leq g_i(x^0)+\frac{d_i^0}{\lambda^k},\ i=1,\dots,m. \tag{6}$$

Finally, using (5) and (6) we have

$$\Phi(x^k)=(f(x^k)-\mu^*)^p/\left(\prod_{i=1}^{m} g_i(x^k)\right)^q \geq \left(f(x^0)-\mu^*+\frac{\varepsilon}{\lambda^k}\right)^p / \left(\prod_{i=1}^{m}\left(g_i(x^0)+\frac{d_i^0}{\lambda^k}\right)\right)^q ,$$

$$\frac{\lim\limits_{k\to\infty}\left(f(x^0)-\mu^*+\frac{\varepsilon}{\lambda^k}\right)^p}{\lim\limits_{k\to\infty}\left(\prod\limits_{i=1}^{m}\left(g_i(x^0)+\frac{d_i^0}{\lambda^k}\right)\right)^q} = \frac{\varepsilon^p}{(\prod\limits_{i=1}^{m} d_i^0)^q}\lim_{k\to\infty}\frac{(\lambda^k)^{mq}}{(\lambda^k)^p} =$$

$$\frac{\varepsilon^p}{(\prod\limits_{i=1}^{m} d_i^0)^q}\lim_{k\to\infty}\frac{1}{(\lambda^k)^{p-mq}} = +\infty,$$

because $\lambda^k\to 0$ as $k\to\infty$ and $p>1+mq$. Thus $\lim\limits_{k\to\infty}\Phi(x^k)=+\infty$ contradicting our assumption that $\Phi(x^k)\to 0$ when $\{\bar{x}^k\}\subset \operatorname{Int} X$ is not bounded. The proposition 1 is proved.

Proposition 2. *If the sequence $\{x^k\}$ converges to the optimum of problem* (1) *in a certain closed polyhedron S such that $X^*\subset S\subset X^*\cup \operatorname{Int} X$, then $\lim\limits_{k\to\infty}\Phi(x^k)=0$.*

Proof. For any $x\in S$ with $(f(x)-\mu^*=\varepsilon$, $\varepsilon>0$ and small enough, and $g_i(x)>0$, $i=1,\ldots,m$, since S is a closed polyhedron there exists a fixed $\gamma>0$, such that $g_i(x)\geq\gamma$ for each i. Then $0<\Phi(x^k)<\varepsilon^{p+mq}/\gamma^{mq}$. The proposition 2 is proved.

Proposition 3. *If $(f(x)-\mu^*)>0$ is a convex function of $x\in\operatorname{Int} X$, where μ^* is a given constant, and if $g=(g_1,\ldots,g_m)>0$ are concave functions of $x\in\operatorname{Int} X$ then for every $q>0$ and $p>1+mq$ the function $\Phi(x)=(f(x)-\mu^*)^p/(\prod\limits_{i=1}^{m} g_i(x))^q$ is a strict convex function of $x\in\operatorname{Int} X$.*

Proof. Let us denote $\Phi=\prod\limits_{i=1}^{m+1} h_j^{\alpha_j}>0$. From [2], the function Φ is a strict convex function of $h_j>0, j=1,\ldots,m+1$, if $\sum\limits_{j=1}^{m+1}\alpha_j>1$ and m numbers from $\alpha_1,\ldots,\alpha_{m+1}$ are negative numbers. It follows that if we take, for instance $\alpha_1=p>0$, $\alpha_j=-q<0$, $j=2,\ldots,m+1$, $h_1=f(x)-\mu^*$, $h_j=g_{j-1}$, $j=2,\ldots,m+1$, then for $p-mq>1$ the function $\Phi=(f(x)-\mu^*)^p/(\prod\limits_{i=1}^{m} g_i(x))^q$ is a strict convex function of $(f(x)-\mu^*)>0$ and $g_i(x)>0$, $i=1,\ldots,m$.

As Φ is nondecreasing of the convex function $(f(x)-\mu^*)>0$ for $x\in\operatorname{Int} X$, and Φ is nonincreasing of the concave functions $g_i(x)>0$, $i=1,\ldots m$ for $x\in\operatorname{Int} X$ from the convexity of composite functions theorem it results that the function Φ is a strict convex function of $x\in\operatorname{Int} X$. See for instance [12]. The Proposition 3 is proved.

From the above propositions it results that starting from a given initial interior point $x^0\in\operatorname{Int} X$ and minimizing the strict convex function $\Phi(x)$ with a conjugate-gradient type algorithm or a quasi-Newton type algorithm we may find a sequence of points $x^1, x^2,\ldots$ in $\operatorname{Int} X$ such that the sequence $\Phi(x^0)$, $\Phi(x^1),\ldots$ converges to zero. Then the limit point x^* of the sequence $\{x^k\}$ is a solution of problem (1).

3. The substantiation of the multiplicative barrier function approach to the nonlinear convex problem. In this section we study the case of solving the nonlinear problem (1) with unknown optimum value using the multiplicative barrier function

$$\Phi(x,\mu^k)=(f(x)-\mu^k)^p/(\prod_{i=1}^{m} g_i(x))^q \tag{7}$$

defined for $x\in\operatorname{Int} X$, $q>0$, $p>1+mq$, $\mu^k<\mu^*=\{\min f(x)|x\in X\}$

The function (7) has the following properties:

a) $\Phi(x,\mu^k)>0$ is a strict convex function for $x\in\operatorname{Int} X$, for any $\mu^k<\mu^*$;

b) If $\{x^k\}\subset\operatorname{Int} X$, and $x^k\to y\in X-\operatorname{Int} X$, $y\notin X^*$, where $\|y\|<\infty$, then $\lim\limits_{k\to\infty}\Phi(x^k,\mu^k)=+\infty$, for any $\mu^k<\mu^*$;

c) If $\{x^k\}\subset\operatorname{Int} X$, $\mu^k<\mu^*$ for every k, and $(x^k,\mu^k)\to(y,\mu^*)$, where $y\in S$, and S is a polyhedron such that $X^*\subset S\subset X^*\cup\operatorname{Int} X$, then $\liminf\limits_{k\to\infty}\Phi(x^k,\mu^k)\geq 0$.

The property (a) follows from Proposition 3, the property (b) follows from the fact that $\Phi(x,\mu^k)$ is a barrier function, and the property (c) follows from Proposition 2.

In the next we shall show that for the convex programming problem (1), if the set X^* of solution points is bounded, a finite minimum of $\Phi(x,\mu^k)$ exists for every $\mu^k<\mu^*$.

Proposition 4. *For $\mu^k<\mu^*$, $k=0,1,\ldots$, there exists an unique optimum solution $x^*(\mu^k)$ which minimizes $\Phi(x,\mu^k)$ in $\operatorname{Int} X$ and $[\nabla_x\Phi(x,\mu^k)]_{x=x^*(\mu^k)}=0$.*

Proof. If X, the set of feasible points, is bounded, then it is obviously that in $\operatorname{Int} X$ for any $\mu^k<\mu^*$, the isocontours of $\Phi(x,\mu^k)$ are bounded, since $\Phi(x,\mu^k)$ is infinite at every boundary point of X, except for $x\in X^*$, and $\Phi(x,\mu^k)$ is continuous in $\operatorname{Int} X$.

If the set of point $\{x^k\}\subset\operatorname{Int} X$ is not bounded the proof follows the proof of Proposition 1 from this paper.

Thus a finite minimum of $\Phi(x,\mu^k)$ exists for every $\mu^k<\mu^*$.

As $\Phi(x,\mu^k)$ is a strict convex function of $x\in\operatorname{Int} X$ for every $\mu^k<\mu^*$ then the finite point $x^*(\mu^k)$ is an unique minimum solution of $\Phi(x,\mu^k)$ in $\operatorname{Int} X$ and $[\nabla_x\Phi(x,\mu^k)]_{x=x^*(\mu^k)}=0$.

Proposition 5. *Let $x^i\in\operatorname{Int} X$ be the solution for $i=k, k+1$, of $\min\Phi(x,\mu^i)$ for $x\in\operatorname{Int} X$, with $\mu^k<\mu^{k+1}<\mu^*$, Then $\Phi(x^{k+1},\mu^{k+1})<\Phi(x^k,\mu^k)$ and $f(x^{k+1})<<f(x^k)$.*

Proof. As $x^k\in\operatorname{Int} X$ is a global solution of $\min\limits_{x\in\operatorname{Int} X}\Phi(x,\mu^k)$ it implies

$$\Phi(x^k,\mu^k)<\Phi(x^{k+1},\mu^k). \tag{8}$$

Also $x^{k+1}\in\operatorname{Int} X$ being a global solution of $\min\limits_{x\in\operatorname{Int} X}\Phi(x,\mu^{k+1})$ it implies

$$\Phi(x^{k+1},\mu^{k+1})<\Phi(x^k,\mu^{k+1}). \tag{9}$$

On the other side

$$\Phi(x^k,\mu^{k+1})-\Phi(x^k,\mu^k) = \frac{(f(x^k)-\mu^{k+1})^p-(f(x^k)-\mu^k)^p}{(\prod\limits_{i=1}^{m} g_i(x^k))^q}<0, \tag{10}$$

because $(\prod_{i=1}^{m} g_i(x^k))^q > 0$ for $x^k \in \operatorname{Int} X$, and $\mu^{k+1} > \mu^k$ implies $(f(x^k) - \mu^{k+1})^p < (f(x^k) - \mu^k)^p$. From (9) and (10) we obtain $\Phi(x^{k+1}, \mu^{k+1}) < \Phi(x^k, \mu^{k+1}) < \Phi(x^k, \mu^k)$.

To prove the last part of the Proposition 5 we multiply (8) by (9) and because of positivity of Φ we obtain:

$$\Phi(x^k, \mu^k) \cdot \Phi(x^{k+1}, \mu^{k+1}) < \Phi(x^{k+1}, \mu^k) \cdot \Phi(x^k, \mu^{k+1}) \Rightarrow$$

$$\Rightarrow \frac{(f(x^k) - \mu^k)^p (f(x^{k+1}) - \mu^{k+1})^p}{\prod_{i=1}^{m} g_i(x^k)^q \, g_i(x^{k+1})^q} < \frac{f(x^{k+1}) - \mu^k)^p (f(x^k) - \mu^{k+1})^p}{\prod_{i=1}^{m} g_i(x^k)^q \cdot g_i(x^{k+1})^q} \Rightarrow$$

$$\Rightarrow (f(x^k) - \mu^k)\,(f(x^{k+1}) - \mu^{k+1}) < (f(x^{k+1}) - \mu^k)\,(f(x^k) - \mu^{k+1}) \Rightarrow \mu^k (f(x^k) - f(x^{k+1})) -$$

$$- \mu^{k+1}(f(x^k) - f(x^{k+1})) < 0 \Rightarrow (f(x^{k+1}) - f(x^k))\,(\mu^{k+1} - \mu^k) < 0. \quad \text{But } \mu^{k+1} > \mu^k \text{ thus}$$

$$f(x^{k+1}) < f(x^k).$$

The Proposition 5 is proved.

In the following we shall show how we can obtain a sequence $\{\mu^k\}$ monotonically increasing to μ^*.

Let us define the Wolfe dual associated with problem (1):

$$\max_{(x,u) \in Y} L(x, u) = f(x) - ug(x) \tag{11}$$

where $Y = \{(x, u) \mid x \in R^n,\ u \in R^m_+,\ \nabla_x L(x, u) = 0\}$.

Proposition 6. *If* $\mu^k < \mu^*$, $k = 0, 1, \ldots$, $x^*(\mu^k) = \arg\min_{x \in \operatorname{Int} X} \Phi(x, \mu^k)$.

$$u_i(\mu^k) = q(f(x^*(\mu^k)) - \mu^k)/(p \cdot g_i(x^*(\mu^k))), \tag{12}$$

$i = 1, \ldots, m$, $q > 0$, $p > 1 + mq$, *then* $(x^*(\mu^k), u(\mu^k))$ *is a feasible solution to dual problem* (11) *and*

$$(f(x^*(\mu^k)) - \mu^* \leqslant mq(\mu^* - \mu^k)/(p - mq) \tag{13}$$

Proof. We shall write explicitly

$$[\nabla_x \Phi(x, \mu^k)]_{x = x^*(\mu^k)} = 0 \text{ and } [\nabla_x L(x, u)]_{x = x^*(\mu^k)} = 0.$$

For the gradient of Φ in $x^*(\mu^k)$ we have

$$[\nabla_x \Phi(x, \mu^k)]_{x = x^*(\mu^k)} = \left[\frac{1}{\Phi(x, \mu^k)} \left(\nabla_x f(x) - \sum_{i=1}^{m} \frac{q(f(x) - \mu^k)}{p \cdot g_i(x)} \nabla_x g_i(x)\right]_{x = x^*(\mu^k)} = 0.$$

As $\Phi(x, \mu^k) > 0$ for $x \in \operatorname{Int} X$, $\mu^k < \mu^*$, it results

$$\left[\nabla_x f(x) - \sum_{i=1}^{m} \frac{q(f(x) - \mu^k)}{p g_i(x)} \nabla_x g_i(x)\right]_{x = x^*(\mu^k)} = 0.$$

For the gradient of the Lagrangian of (11) we have

$$[\nabla_x L(x, u)]_{x = x^*(\mu^k)} = [\nabla_x f(x) - \sum_{i=1}^{m} u_i \nabla_x g_i(x)]_{x = x^*(\mu^k)} = 0. \tag{14}$$

Obviously $u_i(\mu^k) = q(f(x) - \mu^k)/p g_i(x)) > 0$, $i = 1, \ldots, m$ for $x = x^*(x^k_\mu)$ is a solution for (14), and $(x^*(\mu^k), u(\mu^k))$ is a feasible solution to dual problem (11). Replacing $x^*(\mu^k)$ and $u(\mu^k)$ in the Lagrangian of (11) we obtain

$$\begin{aligned} L(x^*(\mu^k), u(\mu^k)) &= \mu^k + f(x^*(\mu^k)) - \mu_k - \sum_{i=1}^{m} \frac{q(f(x^*(\mu^k)) - \mu^k)}{p \cdot g_i(x^*(\mu^k))} g_i(x^*(\mu^k)) = \\ &= \mu^k + \frac{p - mq}{p} (f(x^*(\mu^k)) - \mu^k). \end{aligned} \tag{15}$$

From the duality theory, for $\mu^* = \{\min f(x) \mid x \in X\}$, $x^*(\mu^k)$-the primal feasible point and $(x^*(\mu^k), u(\mu^k))$-the dual feasible point, we have

$$L(x^*(\mu^k), u(\mu^k)) \leqslant \mu^k \leqslant f(x^*(\mu^k)) \tag{16}$$

and replacing (15) in the left hand side of (16) we get

$$\mu^k + \frac{p - mq}{p} (f(x^*(\mu^k)) - \mu^k) \leqslant \mu^*.$$

As $q > 0$, $p > 1 + mq > 0$ it results

$$(p - mq) \cdot f(x^*(\mu^k)) \leqslant p\mu^* - mq \cdot \mu^k,$$

or

$$f(x^*(\mu^k)) - \mu^* \leqslant \frac{mq}{p - mq} (\mu^* - \mu^k).$$

The Proposition 6 is proved.

We note that by suitable choosing of $q > 0$ and $p > 1 + mq > 0$ we may increase the convergence rate of the algorithm.

Corollary 1. *Starting with* $\mu^\circ < \mu^*$ *and then taking*

$$\mu^{k+1} := \mu^k + (p - mq)\,(f(x^*(\mu^k)) - \mu^k)/p, \tag{17}$$

for $q > 0$, *and* $p > 1 + mq > 0$, *where m is the number of constraints, and* $x^*(\mu^k) = \operatorname{argmin}_{x \in \operatorname{Int} X} \Phi(x, \mu^k)$, *the sequence* $\{\mu^k\}$, $k = 0, 1, \ldots$, *is a strictly increasing sequence to* μ^*.

Proof. As $(p - mq)\,(f(x^*(\mu^k)) - \mu^k)/p > 0$ for any $x^*(\mu^k) \in \operatorname{Int} X$ and $\mu^k < \mu^*$ then $\mu^{k+1} - \mu^k > 0$, and because $\mu^{k+1} = L(x^*(\mu^k), \mu^k) = \mu^k + (p - mq)\,(f(x^*(\mu^k)) - \mu^k)/p$, from (16) it result that $\{\mu^k\}$ converges to μ^* when (μ^k) converges to the optimum of problem (1).

We are now in position to give an algorithm for solving the problem (1) in the framework of multiplicative barrier function with unknown optimum objective value.

Algorithm

STEP 0. Iteration $k := 0$

STEP 1. Let μ° be given such that $\mu^\circ \leqslant \mu^*$. If $f(x)$ is bounded from below on R^n then $\mu^\circ = \{\min f(x) \mid x \in R^n\}$.

If $x^* = \underset{x \in R^n}{\operatorname{argmin}} f(x)$ belongs to X, then x^* is the optimum point of problem (1), STOP.
Else

STEP 2. As starting point determine $x^0 \in \operatorname{Int} X$. If such a point is not readly available it may be obtained as in Section 6.

STEP 3. Find $x^*(\mu^k) = \underset{x \in \operatorname{Int} X}{\operatorname{argmin}} \Phi(x, \mu^k)$, starting from $x^k \in \operatorname{Int} X$.

STEP 4. If $(f(x^*(\mu^k)) - \mu^k) \leq \varepsilon$, $\varepsilon > 0$ then STOP. Else

STEP 5. $x^k := x^*(\mu^k)$
$\mu^k := \mu^k + (p - mq)(f(x^*(\mu^k)) - \mu^k)/p$
$k := k+1$
GO TO STEP 3.

4. A sum type barrier function. The Iri-Imai multiplicative barrier function may cause computational difficulties as m, the number of constraints, may be big. As a consequence we introduce for solving the problem (1) a new type barrier function, namely

(18) $$\psi(x, \mu^k) = \sum_{i=1}^{m} (f(x) - \mu^k)^p/(g_i(x))^q,$$

defined for $x \in \operatorname{Int} X$, and $\mu^k < \mu^* = \{\min f(x) \mid x \in X\}$, $q > 0$, $p > q+1$.

The function (18) is in fact the sum of Iri's barrier function applied to each constraint and (18) has the same properties (a), (b), (c) like function (7). The strict convexity of (18) follows from Proposition 3 for $q > 0$, $p > q+1$, and from the fact that the sum of strict convex functions $(f(x) - \mu^k)^p/(g_i(x))^q > 0$, $i = 1, \ldots, m$, for $x \in \operatorname{Int} X$ is also a strict convex function of $x \in \operatorname{Int} X$.

Proposition 4 remains also valid for $\psi(x, \mu^k)$, and thus for $\mu^k < \mu^*$, $k = 0, 1, \ldots$, there exists an unique optimum solution $\bar{x}(\mu^k)$ which minimizes $\psi(x, \mu^k)$ in $\operatorname{Int} X$ and $[\nabla_x \psi(x, \mu^k)]_{x = \bar{x}(\mu^k)} = 0$.

The sequences $\{\psi(\bar{x}(\mu^k))\}$ $\{f(\bar{x}(\mu^k))\}$ are monotonically decreasing sequences and the proof is the same as the proof of Proposition 5, where $(1/\prod_{i=1}^{m}(g_i(x))^q)$ is replaced by $\sum_{i=1}^{m}(1/(g_i(x))^q)$.

Proposition 7. *For any μ^k such that*

$$\mu^k < \mu^* = \{\min f(x) \mid x \in X\}$$

and

$$\bar{x}(\mu^k) = \underset{x \in \operatorname{Int} X}{\operatorname{argmin}} \psi(x, \mu^k),$$

$\bar{u}_i(\mu^k) = q(f(\bar{x}(\mu^k)) - \mu^k)/(p(g_i(\bar{x}(\mu^k)))^{q+1} \cdot \sum_{i=1}^{m}(1/(g_i(\bar{x}(\mu^k))^q)))$, $i = 1, \ldots, m$ with $q > 0$, $p > 1+q$, the pair $(\bar{x}(\mu^k), u(\mu^k))$ is a feasible solution to dual problem (11) and

(19) $$f(\bar{x}(\mu^k)) - \mu^* \leq q(\mu^* - \mu^k)/(p - q).$$

Proof. Taking the gradient of ψ in $\bar{x}(\mu^k)$, we have:

$$[\nabla_x \psi(x, \mu^k)]_{x = \bar{x}(\mu^k)} =$$
$$= \left[(f(x(\mu^k)) - \mu^k)^{p-1} (p \nabla_x f(x) \sum_{i=1}^{m} \frac{1}{(g_1(x))^q} - (f(x) - \mu^k) q \sum_{i=1}^{m} \frac{\nabla_x g_i(x)}{g_i(x)^{q+1}} \right]_{x = \bar{x}(\mu^k)} = 0.$$

It follows that

$$\left[\nabla_x f(x) - \sum_{i=1}^{m} \frac{q(f(x) - \mu^k)}{p(g_i(x))^{q+1} \cdot \sum_{i=1}^{m} \frac{1}{(g_i(x))^q}} \nabla_x g_i(x) \right]_{x = \bar{x}(\mu^k)} = 0.$$

For the gradient of the Lagrangian (11) we have also

(20) $$[\nabla_x L(x, u)]_{x = \bar{x}(\mu^k)} = [\nabla_x f(x) - \sum_{i=1}^{m} u_i \nabla_x g_i(x)]_{x = \bar{x}(\mu^k)} = 0$$

Obviously $\bar{u}_i(\mu^k) = q(f(x) - \mu^k)/p(g_i(x)^{q+1} \sum_{i=1}^{m} \frac{1}{(g_i(x))^q} > 0$, for $x = \bar{x}(\mu^k)$, $i = 1, \ldots, m$ is a solution for (20), and $\bar{x}(\mu^k)$, $\bar{u}(\mu^k))$ is a feasible solution to dual problem (11). Replacing $(\bar{x}(\mu^k), \bar{u}(\mu^k))$ in the Lagrangian of (11) we obtain

$$L(\bar{x}(\mu^k), \bar{u}(\mu^k)) = \mu^k + f(\bar{x}(\mu^k)) - \mu^k - \frac{q(f(\bar{x}(\mu^k)) - \mu^k)}{p \sum_{i=1}^{m} \frac{1}{(g_i(\bar{x}(\mu^k)))^q}} \cdot$$

(21) $$\sum_{i=1}^{m} \frac{1}{(g_i(\bar{x}(\mu^k)))^{q+1}} (g_i(\bar{x}(\mu^k)) = \mu^k + (f(\bar{x}(\mu^k)) - \mu^k) - \frac{q}{p}(f(\bar{x}(\mu^k)) - \mu^k) =$$
$$= \mu^k + \frac{p-q}{p}(f(\bar{x}(\mu^k)) - \mu^k).$$

From the duality theorem it follows that

(22) $$L(\bar{x}(\mu^k), \bar{u}(\mu^k)) \leq \mu^* \leq f(\bar{x}(\mu^k))$$

and replacing (21) in the left side of (22) we obtain

$$\mu^k + \frac{p-q}{p}(f(\bar{x}(\mu^k)) - \mu^k) \leq \mu^*$$

which implies

$$(p - q) f(\bar{x}(\mu^k)) \leq p\mu^* - q\mu^k$$

and finally we get

$$f(\bar{x}((\mu^k)) \leq \frac{q}{p-q}(\mu^* - \mu^k).$$

The Proposition 7 is proved.

Using (21) and starting from $\mu^0 < \mu^*$ we can obtain a sequence $\{\mu^k\}$ strictly increasing to μ^* with

$$\mu^{k+1} := L(\bar{x}(\mu^k), \bar{u}(\mu^k)) = \mu^k + (p-q)\,(f(\bar{x}(\mu^k)) - \mu^k)/p.$$

Making the corresponding changes in the sequential algorithm proposed in Section 3 this can be used to solve problem (1) by using the unconstrained function $\psi(x, \mu^k)$ given by (18).

Comparing (13) and (19), we remark that the second shows a better convergence for multiplicative sum type function (18) to an optimum solution to problem (1) when $\{\mu^k\}$ is a strictly increasing sequence to μ^*.

5. On the Huard's multiplicative barrier function. If in (3) we take $p \geq 1$, $q=1$ we obtain one of the distance function used iteratively by Huard in his method of centres [4] to maximize $(-f(x))$ over X

$$\theta(x, \mu^k) = (-f(x) - (-\mu^k))^p \prod_{i=1}^{m} (g_i(x))^q \tag{23}$$

where $\mu^0 = f(x^0)$ for $x^0 \in \operatorname{Int} X$.

For $k=1, 2, \ldots$, $\mu^k = f(x^*(\mu^{k-1}))$ and $x^*(\mu^{k-1}) = \operatorname*{argmax}_{x \in \operatorname{Int} X} \theta(x, \mu^{k-1})$ we obtain x^* the problem's solution as a limit point of the sequence $x^*(\mu^k)$ by using the same procedure as in Section 3.

We shall denote as above $\mu^* = \{\min f(x) \mid x \in X\}$ and thus $-\mu^* = \{\max(-f(x)) \mid x \in X\}$.

Proposition 8. *If* $\mu^k = f(x^*(\mu^{k-1}))$, $x^*(\mu^k) = \operatorname*{argmax}_{x \in \operatorname{Int} X} \theta(x, \mu^k)$ $u_i(\mu^k) = q(\mu^k - f(x^*(\mu^k)))/(p \cdot g_i(x^*(\mu^k)))$, $i=1, \ldots, m$, $q>0$, $p \geq 1$ *then* $(x^*(\mu^k), u(\mu^k))$ *is a feasible solution to Wolfe dual associated to problem* $\{\max(-f(x)) \mid x \in X\}$, *and*

$$f(x^*(\mu^k)) - \mu^* \leq mq(\mu^k - \mu^*)/(p + mq) \tag{24}$$

Proof. For the gradient of θ in $x^*(\mu^k)$ we have

$$[\nabla_x \theta(x, \mu^k)]_{x=x^*(\mu^k)} = [p \nabla_x(-f(x)) \prod_{i=1}^{m} g_i + q(-f(x) + \mu^k) (\sum_{i=1}^{m} \nabla_x g_i \prod_{\substack{j=1 \\ j \neq i}}^{m} g_j)]_{x=x^*(\mu^k)} = 0$$

or

$$\left[-\nabla_x f(x) + \frac{q(-f(x) + \mu^k)(\sum_{i=1}^{m} \nabla_x g_i \prod_{\substack{j=i \\ j \neq 1}}^{m} g_j)}{p \cdot \prod_{i=1}^{m} g_i} \right]_{x=x^*(\mu^k)} = 0.$$

It is easily seen that

$$u_i = \frac{q(\mu^k - f(x^*(\mu^k)))}{p \cdot g_i(x^*(\mu^k))} > 0, \quad i = 1, \ldots, m,$$ vanish the gradient of the Lagrangian of the Wolfe dual associated to problem $\{\max(-f(x)) \mid x \in X\}$ and $(x^*(\mu^k), u(\mu^k))$ is a feasible solution to this dual problem. Replacing $x^*(\mu^k)$ and $u(\mu^k)$ in the Lagrangian we obtain

$$-L(x^*(\mu^k), u(\mu^k)) = \mu^k - f(x^*(\mu^k)) - \mu^k - \sum_{i=1}^{m} \frac{q(f(x^*(\mu^k)) - \mu^k)}{p \cdot g_i(x^*(\mu^k))} g_i(x^*(\mu^k)) =$$

$$= -\mu^k - \frac{p + mq}{p}(f(x^*(\mu^k)) - \mu^k) \tag{25}$$

From the duality theory we have

$$L(x^*(\mu^k), u(\mu^k)) \leq \mu^* \leq f(x^*(\mu^k)).$$

and replacing (25) in the left side, we obtain

$$\mu^k + \frac{p + mq}{p}(f(x^*(\mu^k)) - \mu^k) \leq \mu^*.$$

From here we have

$$(p + mq) f(x^*(\mu^k)) \leq p\mu^* + mq\,\mu^k$$

and finally

$$f(x^*(\mu^k)) - \mu^* \leq \frac{mq}{p + mq}(\mu^k - \mu^*).$$

The Proposition 8 is proved.

From (24) it results that the Huard's method of centres has a very good convergence to the optimum when using the multiplicative barrier function (23) and $\{-\mu^k\}$ is a rapidly increasing sequence to $-\mu^*$.

6. Obtaining a strictly feasible starting point. In order to find a strictly feasible point we propose a procedure in which the barrier function is applied to a modified problem with known optimum objective value. For $e=(1, 1, \ldots, 1)$ we set $s_i = g_i(e)$, $i=1, \ldots, m$, $\tilde{g}_i = 1 - s_i$ if $s_i \leq 0$ and $\tilde{g}_i = 0$ if $s_i > 0$, and let us define the modified problem

$$\min_{x, \lambda} \{\lambda \mid g(x) + \tilde{g}\lambda \geq 0,\ \lambda \geq 0\}. \tag{26}$$

Then $x^0 = e$, $\lambda^0 = 1$ is an initial interior feasible point for (26) and problem (1) has feasible solution if the value of objective function λ is zero.

The problem (26) can be solved by minimizing

$$\Phi(x, \lambda) = \lambda^{p-q} / \prod_{i=1}^{m} (g_i(x) + \tilde{g}_i\lambda)^q \quad \text{with } q>0,\ p>1+(m+1)q \quad \text{or}$$

$$\varphi(x, \lambda) = \lambda^{p-q} + \sum_{i=1}^{m} \lambda^p / (g_i(x) + \tilde{g}_i\lambda)^q)$$

with $q>0$, $p>1+q$.

In our experiments we have stopped the procedure when it is found $x \in \operatorname{Int} X$.

7. Computational experiments. The algorithm proposed in Section 3 was coded into Fortran -77 program using quasi-Newton-Broyden-Goldfarb-Fletcher-Shanno method for minimization the barrier functions (7), (18) and (23). All the experiments were executed with double precision arithmetics on Independent 102-F of Department of Cybernetics of the Academy of Economic Studies.

Our purpose was only to show how fast the function values of (7), (18) or (23) modify as iteration proceeds and how the number of necessary iterations depends of the parameters of the barrier functions. Our numerical results with small-size sample problems from some nonlinear programming test problem collections are very promising. In the following we give the numerical results obtained with the Rosen-Suzuki [13] test problem:

$$\min f(x)=x_1^2+x_2^2+2x_3^2+x_4^2-5x_1-5x_2-21x_3+7x_4, \text{ s.t.}$$

$-x_1^2-x_2^2-x_3^2-x_4^2-x_1+x_2-x_3+x_4+8\geqslant 0$; $-x_1^2-2x_2^2-x_3^2-2x_4^2+x_1+x_4+10\geqslant 0$, $-2x_1^2-x_2^2-x_3^2-2x_1+x_2+x_4+5\geqslant 0$. The optimum point $x^*=(0, 1, 2, -1)$ and $\mu^*=f(x^*)=-44$. The starting point $x^0=(0, 0, 0, 0)$ and for (7) and (13) we chose $\mu^0=\{\min f(x) \mid x\in R^4\}=-77$.

With the purpose of showing the behaviour of the algorithm in relation to each barrier function, we choose $p=3$, $q=0.2$. The computational results are given in Table 1. From Table 2 it can be seen that the parameters p and q influence the convergence of the algorithm such that a great difference between them gives a better convergence to the optimum.

Table 1

Barrier functions	No. seq.	μ^k	$f(x^*(\mu))$	$f(x^*(\mu^k))-\mu^*$
The generalized Iri's function $p=3$, $q=0.2$ Final point: (0.0004, 0.9997, 2.00007, −0.9998)	0 1 2 3 4	−77. −45.8235 −44.1141 −44.0086 −44.0007	−38.0294 −43.6867 −43.9823 −43.9986 −43.9998	5.9705 0.3132 0.0177 0.0013 0.0002
The sum type function $p=3$, $q=0.2$ Final point: (0.0004, 0.9997, 2.0003, −1.00002)	0 1 2	−77. −44.4113 −44.0046	−42.0835 −43.9755 −43.9999	1.9164 0.0244 0.00008
The Huard's function $p=3$, $q=0.2$ Final point: (−0.0013, 0.9843, 2.0054, −0.9958)	0 1 2 3 4	0. −38.1471 −43.2192 −43.9051 −43.9891	−38.1471 −43.2192 −43.9051 −43.9891 −43.9956	5.8529 0.7807 −0.0949 0.0109 0.0044

Table 2

Numerical results with sum type function (For $p = 3$, $q = 0.2$ see Table 1)

Parameters	No. seq.	μ^k	$f(x^*(\mu^k))$	$f(x^*(\mu^k))-\mu^*$
$p=3$, $q=0.5$ Final point: (0,00005, 0.9999, 2.00002, −0.9998)	0 1 2 3	−77. −45.1484 −44.0228 −44.0001	−38.7781 −43.7976 −43.9955 −43.9995	5.2218 0.2023 0.0044 0.00016
$p=3$; $q=1$ Final point: (0.0003, 0.9994, 2.0001, −0.9981)	0 1 2 3 4	−77. −47.5975 −44.1609 −44.0011 −43.9966	−32.8963 −42.4426 −43.9212 −43.9943 −43.9962	11.1036 1.5574 0.0788 0.0056 0.0037

Final Conclusion. Our generalized barrier function $B(x, \mu^k)$ can be used for any $p>1+m\,|q|$ and $q\neq 0$ to solve the convex problem as an unconstrained one.

If $q>0$ then we have to use the Huard's sequential technique of method of centres. If $q<0$ then p is restricted by the inequality $p>1+m\,|q|$ and still it can be used in the sequential algorithm proposed in our paper. So, the main difference between the two methods seems to be a stronger condition on p, implying a better quality of $B(x, \mu^k)$, i.e. convexity versus quasiconcavity.

However the convexity versus quasiconcavity quality of the barrier function did not provide better computational results on the tested problems.

REFERENCES

1. Dumitru, V., Luban, F. — *A generalized Huard-Iri function for unconstrained optimization*, Ec. Comp. and Ec. Cyber. Stud. and Res. No. 1, 1989, pp. 77—80.
2. Dragomirescu, M., Malita, M. — *Programare neliniară*, Editura Ştiinţifică, Bucureşti, 1972.
3. Fiacco, A. V., McCormick, G. P. — *Nonlinear programming sequential unconstrained minimization technique*, John Wiley and Sons, New York, 1968.
4. Huard, P., J. Abadie (ed) — *Resolution of mathematical programming with nonlinear constraints by the methods of centres*, Nonlinear Programming. John Wiley and Sons, New York, 1967, pp. 209—219.
5. Iri, M., Imai, H. — *A method of solving linear programming — with reference to the Karmarkar method and the penalty function method*. Research Meeting of the MP Research Group of the OR Society of Japan, February 16, 1985.
6. Iri, M., Imai, H. — *A multiplicative penalty function method for linear programming*, Proceedings of the 6-th Mathematical Programming Symposium, Japan, 1985 pp. 97—120.
7. Iri, M., Imai, H. — *A multiplicative barrier function method for linear programming*, Algorithmica 1, 1986, pp. 455—482.
8. Iri, M., Imai, H. — *Theory of the multiplicative penalty function method for linear programming*, Discrete Algorithm and Convexity, Academic Press, 1987 pp. 417—435

9. Karmarkar, N. K. — *A new polynominal-time algorithm for linear programming*, Combinatorica 4, 1984, pp. 373—395.
10. Khachiyan, L. G. — *A polynomial algorithm in linear programming*. Soviet Mathematics Doklady 20, 1979 pp. 191—194.
11. Klee, V., Minty G. J., — *How good is the simplex algorithm*, in Inequalities III. Academic Press, New York, 1972.
12. Martos, B. — *Nonlinear programming theory and methods*, Akademiai Kiado, Budapest, 1975.
13. Rosen, J. B., Suzuki, S. — *Construction of nonlinear programming test problems*, Comm. ACM, 1965, pp. 113.
14. Shor, N. Z. — *Utilization of space dilatation operation in minimization of convex functions*, Kibernetica, 1970 No. 1, pp. 6—12.

Received 15.IX.1989

The Academy of Economic Studies
Dept. of Cybernetics
15—17 Calea Dorobanţilor
71131 Bucharest, România