INFORMATICA, 1990, Vol.1, No.1, 040–058

EXACT AUXILIARY FUNCTIONS

Yuri EVTUSHENKO and Vitali ZHADAN

Computer Center of the USSR Academy of Sciences,
117967 Moscow, GSP-1, Vavilova St.40, USSR

Abstract. A new concept of an exact auxiliary function (EAF) is introduced. A function is said to be EAF, if the set of global minimizers of this function coincides with the global solution set of the initial optimization problem. Sufficient conditions for exact equivalence of the constrained minimization problem and minimization of EAF are provided. The paper presents various classes of EAF for a nonlinear programming problem, which has a saddle point of Lagrange function.

Key words: nonlinear programming, exact penalty function, saddle point, Lagrange function.

1. Introduction. We shall be concerned here with the nonlinear programming problem

$$\text{minimize } f(x) \text{ subject to } x \in X, \tag{1}$$

where f is a continuous function from R^n into R^1, X is a feasible set. Denote the set of global optimal solutions to the problem (1) by X_* and the global optimal objective value to the problem (1) by f_*. Then,

$$X_* = \{x_* \in X | f(x) - f(x_*) \geqslant 0 \text{ for all } x \in X\},$$

$$f_* = f(x_*), \ x_* \in X_*.$$

Everywhere the feasible set X and the set of solutions X_* are supposed to be nonempty.

We deal with the problem of how to transform a constrained program into an equivalent unconstrained minimization problem or a constrained problem with another feasible set. We shall associate with the nonlinear programming problem (1) a class of numerical functions $H(x,y)$, where parameter $y \in Y \sqsubseteq R^l$. Let $X(y)$ denote the set of all points, which globally minimize the function $H(x,y)$ on a set $P \sqsubseteq R^n$:

$$X(y) = \{x_* \in R^n | H(x_*, y) \leqslant H(x,y) \text{ for all } x \in P\}.$$

The function $H(x,y)$ is supposed to be continuous and such that the solution set $X_* \sqsubseteq P$. In many cases P is R^n.

Definition. The function $H(x,y)$ is said to be an exact auxiliary function (EAF) on $P \times Y$ if $X(y) = X_*$ for all $y \in Y$.

The simplest example of EAF is a well-known exact penalty function, which was found simultaneously by Eremin (1967) and Zangwill (1967). Many papers were devoted to the investigation in this field (see e.g., Bertsecas, 1975; Han and Mangasarian, 1979; Evtushenko, 1985, 1987; Zhadan, 1984). Here we will show that besides the exact penalty functions, there are many classes of EAF. Our approach is based on our previous investigations (Evtushenkò and Zhadan, 1988). We will be primarily interested in the case, where the set $P = R^n$ and numerical solution of the constrained problem (1) can be replaced by unconstrained minimization of the function $H(x,y)$. Then the problem (1) could be solved in a relatively easy manner. It is desirable that the set Y should be as "broad" as possible, so that we do not confront a problem of finding a point belonging to the set Y. Therefore, we exclude from consideration the case, where Y is empty or consists of only one point.

Very often EAF is constructed as a difference

$$M(x,y) = H(x,y) - H(x_*,y), \quad x_* \in X_* \sqsubseteq P. \qquad (2)$$

This function is nonnegative for all $x \in P$, $y \in Y$, $x_* \in X_*$. Necessary and sufficient conditions for a function M to be EAF are:

$$i)\ M(x,y) \geqslant 0 \quad \forall x \in P, \quad \forall y \in Y; \tag{3}$$

ii) for any $y \in Y$ from $M(x,y)=0$ follows that

$$x \in X(y) = X_*. \tag{4}$$

For the sake of simplicity, we assume that the feasible set X is defined by inequality type constraints and that a saddle point $[x_*, w_*]$ of Lagrange function exists:

$$X = \{x \in R^n | g(x) \leqslant 0\}, \quad L(x,w) = f(x) + \langle w, g(x) \rangle,$$

$$L(x_*, w) \leqslant L(x_*, w_*) \leqslant L(x, w_*) \quad \forall x \in P, \quad \forall w \in R^m_+, \tag{5}$$

where $g(x)$ is a function from R^n into R^m , $w \in R^m_+$, R^m_+ is a nonnegative orthant of R^m . We denote by $<\cdot,\cdot>$ the usual inner product. We use subscripts to denote different vectors, e.g. x_1 and x_2. Superscripts are used to denote components of a vector.

Throughout the paper we assume that a saddle point $[x_*, w_*]$ of Lagrange function exists in the problem (1). This assumption will not be repeated in the statements of the theorems given in the paper.

In the next sections we derive our principle results which relate the global solution set of (1) to the global minimizer set $X(y)$. Our vehicle for deriving many of results will be the Minkowski-Mahler inequality and a notion of polar function which we generalized here. The main advantage of such an approach is that no specific assumptions on the functions appearing in the problem, such as convexity or differentiability, are needed in order to prove equivalence and to find a set Y.

Let $Q(z)$ be a scalar continuous function defined on a set $Z \sqsubseteq R^s$, $z_0 \in R^s$. We define a function $Q^0(z_0|Z)$, which we call a polar function to $Q(z)$ by the following conditions:

$$Q^0(z_0|Z) = \inf_{\mu \in N(z_0)} \mu,$$

$$N(z_0) = \left\{ \mu \in R^1 | < z, z_0 > \leqslant \mu Q(z) \quad \forall z \in Z \right\}.$$

From this definition we obtain the Minkowski-Mahler inequality

$$\begin{aligned} < z, z_0 > \leqslant Q(z) Q^0(z_0|Z), \quad & \forall z \in Z \sqsubseteq R^s, \\ & \forall z_0 \in Z_0 \sqsubseteq R^s. \end{aligned} \tag{6}$$

If a function $Q(z)$ is strictly positive on Z or strictly negative, we can simplify this definition of a polar function. We have, respectively:

$$Q^0(z_0|Z) = \sup_{z \in Z} \frac{< z, z_0 >}{Q(z)}, \quad Q^0(z_0|Z) = \inf_{z \in Z} \frac{< z, z_0 >}{Q(z)}.$$

Let $\|z\|_p$ denote a Hölder vector norm:

$$\|z\|_p = \left(\sum_{i=1}^{s} |z^i|^p \right)^{1/p}. \tag{7}$$

The function $\|z\|_p$ has all properties of a norm only if $p \geqslant 1$. We will use the notation (7) also for cases where $p < 1$. If $p = 2$, then a Holder norm coincides with an Euclidean norm. We mention some other cases:

$$\|z\|_0 = \sqrt{s} \left(\prod_{i=1}^{s} |z^i| \right)^{1/s},$$

$$\|z\|_1 = \sum_{i=1}^{s} |z^i|, \quad \|z\|_{+\infty} = \max_{1 \leqslant i \leqslant s} |z^i|, \quad \|z\|_{-\infty} = \min_{1 \leqslant i \leqslant s} |z^i|.$$

If $p < 0$ then we assume that $\|z\|_p = 0$, if at least one component of z is equal to zero. We give several functions $Q(z)$ and their polar functions:

$$\begin{array}{lll} Q(z) = \|z_+\|_p, & Q^0(z_0|R^s) = \|z_0\|_{p_*}, & p > 1, \quad p_* > 1, \\ Q(z) = -\|z_-\|_p, & Q^0(z_0|R^s_-) = \|z_0\|_{p_*}, & p < 1, \quad p_* < 1, \\ Q(z) = -\|z_-|_0, & Q^0(z_0|R^s_-) = \|z_0\|_0, & \\ Q(z) = \|z_+\|_{+\infty}, & Q^0(z_0|R^s) = \|z_0\|_1, & \\ Q(z) = -\|z_-\|_{-\infty}, & Q^0(z_0|R^s_-) = \|z_0\|_1, & \end{array}$$

where $p^{-1} + p_*^{-1} = 1$, R^s_- is a nonpositive orthant,

$$\begin{array}{ll} z_+ = [z_+^1, z_+^2, \dots, z_+^s], & z_+^i = \max[0, z^i], \\ z_- = [z_-^1, z_-^2, \dots, z_-^s], & z_-^i = \min[0, z^i]. \end{array}$$

The vector norm $\|z_0\|_{p_*}$ is called the dual norm of $\|z\|_p$.

Define the composite numerical function $B(g(x))$. Let $g(P)$ be an image of the set P in mapping $g(x)$. Making use of the right-hand side inequality (5) and (6), we have

$$f_* \leqslant L(x, w_*) \leqslant f(x) + B(g(x))B^0(w_*|g(P)) \quad \forall x \in P. \quad (8)$$

It is easy to show that, if $B(g(x)) \geqslant 0$ for all $x \in R^n$, then we get that

$$B^0(w_*|g(P)) \leqslant B^0(w_*|g(R^n)) \leqslant B^0(w_*|R^m).$$

By using these inequalities, we can rewrite (8) in a slightly less precise but more convenient form, such as

$$f_* \leqslant L(x, w_*) \leqslant f(x) + B(g(x))B^0(w_*|R^m) \qquad \forall x \in P. \quad (9)$$

Similarly, if $B(g(x)) \leqslant 0$ on R^n, then we have

$$B^0(w_*|g(P)) \geqslant B^0(w_*|g(R^n)) \geqslant B^0(w_*|R^m)$$

and the inequality (9) also holds.

2. Additive EAF. We start from the simplest additive class of EAF. Let

$$H(x,y) = A(f(x),y) + B(g(x)), \tag{10}$$

where $A(f,y)$ is a continuous function, $B(g(x))$ is a strict exterior penalty function. It means that $B(g(x)) \geqslant 0$ everywhere on R^n and $B(g(x)) = 0$ if and only if $x \in X$. The most popular exterior penalty function and its polar function are

$$B(g(x)) = \|g_+(x)\|_p, \quad p \geqslant 1, \quad B^0(w_*|R^m) = \|w_*\|_{p_*}.$$

We impose two additional conditions on A and B:

a) functions A and B are such that the following inequality holds

$$\begin{aligned} A(f,y) - A(f_*,y) \geqslant [(f-f_*)/B^0(w_*|R^m)]_- \\ \forall y \in Y, \quad \forall f \in R^1; \end{aligned} \tag{11}$$

b) for any point $y \in Y$ the set X_* equals to a solution set of the equation

$$A(f(x),y) + B(g(x)) = A(f_*,y), \quad x \in P. \tag{12}$$

If Y is an open set in R^l , A is a differentiable function of f and y, then from (12) we get the system

$$A_y(f(x),y) = A_y(f_*,y), \quad x \in P.$$

Very often it is much easier to solve this system instead of (12).

Theorem 1. Assume that $B(g(x))$ is a strict exterior penalty function, $0 < B^0(w_*|R^m) < +\infty$ and functions A,

B satisfy conditions a) and b). Then the additive function $H(x, y)$ defined by (10) is EAF on $P * Y$.

Proof. Taking into account that the function $B(g(x))$ is equal to zero on X, we get

$$M(x, y) = A(f(x), y) - A(f_*, y) + B(g(x)). \qquad (13)$$

The condition (3) can be rewritten in the form

$$A(f(x)) + B(g(x)) \geqslant A(f_*, y) \quad \forall x \in P, \quad \forall y \in Y.$$

Noting that $B^0(w_*|R^m) > 0$, we find from (9) the following inequality

$$B(g(x)) \geqslant (f_* - f(x))/B^0(w_*|R^m) \quad \forall x \in P.$$

The function B is always nonnegative, therefore we can correct the previous inequality

$$B(g(x)) \geqslant -\left[(f(x) - f_*)/B^0(w_*|R^m)\right]_- \quad \forall x \in P. \qquad (14)$$

Applying this inequality and (11) for estimation of the right-hand side of (13), we get (3). From the condition b) it follows that (4) holds. This ensures that $M(x, y)$ is EAF on $P * Y$, and the proof is now complete.

A lot of functions satisfy this theorem. We suppose that f_* is unknown, hence A must be constructed in such a way, that conditions a) and b) are satisfied for any value of f_*. It follows from (11), that the function $A(f, y)$ must be at least a nondecreasing function of f. Let us restrict the class of EAF under consideration by requiring that $A(f, y)$ is a convex function of f for any $y \in Y$. Then we have

$$A(f, y) - A(f_*, y) \geqslant \xi(f - f_*)$$
$$\forall y \in Y, \quad \forall f \in R^1, \quad \forall \xi \in \partial_f A(f_*, y), \qquad (15)$$

where $\partial_f A(f_*, y)$ is a subdifferential of $A(f, y)$ at a point f_*. The condition (11) holds if

$$\xi(f - f_*) \geqslant \left[(f - f_*)/B^0(w_*|R^m)\right]_-$$

$$\forall y \in Y, \quad \forall f \in R^1, \quad \forall \xi \in \partial_f A(f_*, y).$$

This inequality and nondecreasing condition are equivalent to the following inequalities

$$0 \leqslant \inf_{\xi \in \partial A_f(f_*, y)} \xi \leqslant \sup_{\xi \in \partial A_f(f_*, y)} \xi \leqslant 1/B^0(w_*|R^m). \qquad (16)$$

Here we show some examples of EAF that satisfy Theorem 1 and conditions (15), (16)

$H_1(x, y) = y^{-1} f(x) + B(g(x))$,
$Y_1 = \{y | y > B^0(w_*|R^m)\}$,
$H_2(x, y) = y^{-1} e^{f(x)} + B(g(x))$,
$Y_2 = \{y | y > B^0(w_*|R^m) e^{f_*}\}$,
$H_3(x, y) = (y - f(x))_+^{\alpha} + B(g(x))$,
$Y_3 = \{y | y > f_* + (-\alpha B^0(w_*|R^m))^{1/(1-\alpha)}\}$,
$H_4(x, y) = (f(x) - y)_+^{\beta} + B(g(x))$,
$Y_4 = \{y | f_* \geqslant y > f_* - (\beta B^0(w_*|R^m))^{1/(1-\beta)}\}$,

where $\alpha < 0$, $\beta > 1$. Parameters α or β can be considered also as a second component of the vector y.

The function H_1 is a well-known exact penalty function (see e.g., Bertsecas, 1975; Han and Mangasarian, 1979). The function H_4 was considered by Morrison (1966), Lootsma (1974) and many other authors. In these papers a sequence of minimization problems was used with the parameter $y \leqslant f_*$. This parameter was adjusted from one minimization of H_4 to another so that the sequence of y_i converges to the optimal objective value f_*. From Theorem 1 we come to a conclusion, that any single minimization of $H_4(x, y)$ yields an optimal solution of (1) if $y \in Y_4$. If a function $f(x) > 0$ for all $x \in R^n$, then we can use the following EAF

$H_5(x, y) = y^{-1}\mathrm{sh}f(x) + B(g(x)),$
$Y_5 = \{y | y > B^0(w_* | R^m)\mathrm{ch}f_*\},$
$H_6(x, y) = y^{-1}(f(x))^\gamma + B(g(x)), \ \gamma \geqslant 1,$
$Y_6 = \{y | y > \gamma f_*^{\gamma-1} B^0(w_* | R^m)\}.$

If the function $f(x)$ is nonpositive on R_n , then we can introduce another EAF

$H_7(x, y) = y^{-1} \arctan f(x) + B(g(x)),$
$Y_7 = \{y | y > B^0(w_* | R^m)/(1 + f_*^2)\}.$

Denote $X^0 = \{x \in R^n | g(x) < 0\}$. Assume that $B(g(x))$ is an interior penalty function. It means that $B(g)$ is continuous and nonpositive on R^n and $B(g(x)) < 0$ if $x \in X^0$. Suppose moreover that $B(g(x)) < 0$ if and only if $g(x) < 0$, then the function $B(g(x))$ is called a strict interior penalty function. For example,

$$B(g(x)) = -\|g_-(x)\|_p, \quad -\infty \leqslant p < 1 \qquad (17)$$

is an interior penalty function. If $p \leqslant 0$,then this function is a strict interior penalty function.

We consider EAF (10) with $P = X$, the condition (11) is replaced by the following

$$\begin{gathered} A(f, y) - A(f_*, y) \geqslant (f - f_*))/B^0(w_* | R_-^m) \\ \forall y \in Y, \ \forall f \geqslant f_*. \end{gathered} \qquad (18)$$

Theorem 2. Let the set X^0 be nonempty and its closure coincide with X, let $B(g(x))$ be an interior penalty function and $0 < B^0(w_* | R_-^m) < \infty$. Assume that inequality (18) holds and any solution x of the system (12) is such that $x \in X_*$. Then the function $H(x, y)$ defined by (10) is EAF on $X \times Y$.

The proof is almost identical to that of the previous theorem.

Suppose that $A(f, y)$ is a convex function of f. Then inequality (18) holds,if

$$\xi \geqslant 1/B^0(w_* | R_-^m) \quad \forall \xi \in \partial_f A(f_*, y), \quad \forall y \in Y.$$

All functions $H_1 - H_4$ satisfy Theorem 2, but instead of a strict exterior penalty function we must use an interior penalty function (for example (17)) and replace sets Y_i by the following sets

$Y_1 = \{y | 0 < y < B^0(w_* | R^m_-)\}$,
$Y_2 = \{y | 0 < y < e^{f_*} B^0(w_* | R^m_-)\}$,
$Y_3 = \{y | f_* < y < f_* + (-\alpha B^0(w_* | R^m_-))^{1/(1-\alpha)}\}$, $\alpha < 0$,
$Y_4 = \{y | y < f_* - (\beta B^0(w_* | R^m_-))^{1/(1-\beta)}\}$, $\beta > 1$.

A continuous function $B(g(x))$ is called a strict mixed penalty function if $B(g(x)) > 0$ for $x \notin X$ and $B(g(x)) < 0$ for $x \in X^0$. As examples of strict mixed penalty functions we mention the following two functions:

$$B(g(x)) = \|g_+(x)\|_p - \|g_-(x)\|_{-p}, \quad 1 < p < +\infty, \tag{19}$$

$$B(g(x)) = \max_{1 \leqslant i \leqslant m} g^i(x). \tag{20}$$

The function (19) was used earlier by Charalambos (1976).

Instead of (11) and (18), we introduce the following condition:

$$A(f, y) - A(f_*, y) \geqslant$$

$$\geqslant \begin{cases} (f - f_*)/B^0(w_* | R^m_-), & \text{if } f \geqslant f_*, \\ (f - f_*)/B^0(w_* | R^m), & \text{if } f < f_*. \end{cases} \forall y \in Y. \tag{21}$$

Theorem 3. Assume that $B(g(x))$ is a strict mixed penalty function, $0 < B^0(w_* | R^m) < B^0(w_* | R^m_-) < +\infty$, conditions b) and (21) hold. Then the function $H(x, y)$ defined by (10) is EAF.

If we use the function $H_1(x, y)$ with a strict mixed penalty function $B(g(x))$, then

$$Y_1 = \{y | B^0(w_* | R^m) < y < B^0(w_* | R^m_-)\}.$$

If $B(g(x))$ is defined by (19), then we have

$$Y_1 = \{y | \; \|w_*\|_{p_*} < y < \|w_*\|_{r_*}\},$$

where $p^{-1} + p_*^{-1} = 1,\ -p^{-1} + r_*^{-1} = 1$. Taking into account that $p_* > 1, 1 > r_* > 1/2$, we obtain that $\|w_*\|_{r_*} > \|w_*\|_{p_*}$. For function (20) the set Y_1 is empty.

3. Nonlinear EAF. We introduce a nonadditive EAF

$$H(x,y) = G(A(f(x),y), B(g(x))), \tag{22}$$

where $B(g(x))$ is a strict exterior penalty function, $G(t,\tau)$ is a nondecreasing function of two variables, it means that $G(t_1,\tau_1) \geqslant G(t,\tau)$ for any $t_1 \geqslant t,\ \tau_1 \geqslant \tau$. We assume that $A(f,y)$ is a convex function of f and a scalar D exists such that

$$0 < \sup_{y\in Y} \sup_{\xi\in\partial A_f(f_*,y)} \xi \leqslant D < +\infty. \tag{23}$$

Denote $N_1 = DB^0(w_*|R^m)$.

Theorem 4. Let $B(g(x))$ be a strict exterior penalty function and $0 < N_1 < +\infty$. Assume that $A(f,y)$ is a nondecreasing convex function of f and the inequality (23) holds. Assume that there exists a set $T \in R^1$ such that $A(f_*,y) \in T$ for all $y \in Y$ and

$$G(t,(t_* - t)_+/N_1) > G(t_*,0) \quad \forall t_* \in T,\ \forall t \neq t_*, \tag{24}$$

then (22) is EAF on $P \times Y$.

Proof. Let $x \in P$. Using a nondecreasing property of $G(t,\tau)$,we obtain from (14)

$$H(x,y) \geqslant G(A(f(x),y),(f_* - f(x))_+/B^0(w_*|R^m)). \tag{25}$$

Taking into account the convexity of $A(f,y)$ with respect to f, we have

$$A(f_*,y) - A(f,y) \leqslant D(f_* - f) \quad \forall f \leqslant f_*.$$

Because of the monotonicity of A

$$(A(f_*, y) - A(f, y))_+ \leqslant D(f_* - f)_+ \quad \forall f \in R^1.$$

We substitute this inequality in (25) and use (24). If $A(f(x), y) \neq A(f_*, y)$, then we have

$$H(x, y) \geqslant G(A(f(x), y), (A(f_*, y) - A(f(x), y))_+/N_1) >$$
$$> G(A(f_*, y), 0) = H(x_*, y).$$

Consider the case where $A(f(x), y) = A(f_*, y)$. If $x \notin X_*$, then $x \notin X$, $B(g(x)) > 0$. Using the monotonocity of $G(t, \tau)$ with t and (24), we get

$$H(x, y) = G(A(f_*, y), B(g(x))) \geqslant G(A(f_*, y) -$$

$$-N_1 B(g(x)), B(g(x))) > G(A(f_*, y), 0) = H(x_*, y).$$

We conclude that $H(x, y) > H(x_*, y)$, if $x \notin X_*$ and, therefore, H is EAF on $P \times Y$. The theorem is proved.

The following function satisfies conditions of Theorem 4

$$G(t, \tau) = \begin{cases} t_-/(1 - t_-\tau), & \text{if } 1 - t_-\tau \geqslant 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

This function is quasiconvex on R^2. We have $H_t(t, 0) = 1$, $H_\tau(t, 0) = t_-^2$, where $t < 0$. Hence the condition (24) holds, if $t < -\sqrt{B^0(w_*|R^m)}$ and $D = 1$. Thus,

$$T = (-\infty, -\sqrt{B^0(w_*|R^m)})$$

and the function

$$H_8(x, y) = (f(x) - y)_-/[1 - (f(x) - y)_- B(g(x))]$$

is EAF on the set $R^n \times Y_8$, where $Y_8 = \{y \in E^1 | y > f_* + \sqrt{B^0(w_* | R^m)}\}$.

We consider the case where $B(g(x))$ is an interior penalty function. As before,we suppose that $G(t, \tau)$ is a nondecreasing function on R^n and $A(f, y)$ is a monotonically increasing function of f. We assume that instead of (23) there exists such a constant C that

$$0 < C \leqslant \inf_{y \in Y} \inf_{\xi \in \partial A_f(f_*, y)} \xi < +\infty. \tag{26}$$

The condition (24) is replaced by

$$G(t, (t_* - t)/N_2) > G(t_*, 0) \quad \forall t_* \in T, \forall t > t_*, \tag{27}$$

where $N_2 = CB^0(w_* | R^m_-)$. Theorem 4 for such a case can be reformulated in the following way:

Theorem 5. Suppose that $B(g(x))$ is an interior penalty function and $0 < N_2 < +\infty$. If there exists a set $T \sqsubseteq R^1$ such that (27) holds,then (22) is EAF for the problem (1) on the set $X \times Y$.

Proof. Let us take an arbitrary point $x \in X$. In the case, where $x \notin X_*$, using inequalities (15), (26) and (27), we have

$$\begin{aligned} H(x, y) = G(A(f(x), y), B(g(x))) \geqslant \\ \geqslant G(A(f(x), y), (f_* - f(x))/B^0(w_* | R^m_-)) \geqslant \\ \geqslant G(A(f(x), y), (A(f_*, y) - A(f(x), y))/N_2) > \\ > G(A(f_*, y), 0) = H(x_*, y). \end{aligned}$$

According to (8) $B(g(x)) = 0$ for $x \in X_*$. Hence,$H(x, y) = G(A(f(x), y), 0) = H(x_*, y)$ if $x \in X_*$. Thus,$H(x, y)$ is EAF on $X \times Y$. The theorem is proved.

The function $H_8(x, y)$ satisfies conditions of this theorem, if $B(g(x))$ is an interior penalty function and

$$Y_8 = \{y \in R^1 | f_* \leqslant y < f_* + \sqrt{B^0(w_* | R^m_-)}\}.$$

The function $H_8(x,y)$ with the strict mixed penalty function $B(g(x))$ is EAF on the set $R^n \times Y_8$, where

$$Y_8 = \{y | f_* + \sqrt{B^0(w_*|R^m)} < y < f_* + \sqrt{B^0(w_*|R^m_-)}\}.$$

In particular, if the function $B(g(x))$ has the form (19), we have

$$Y_8 = \{y | f_* + \sqrt{\|w_*\|_{p_*}} < y < f_* + \sqrt{\|w_*\|_{r_*}}\},$$

where $p^{-1} + p_*^{-1} = 1$, $-p^{-1} + r_*^{-1} = 1$. This set is not empty because of $\|w_*\|_{r_*} > \|w_*\|_{p_*}$.

4. Exact modified Lagrange functions. We shall rewrite the right-hand side inequality (5) in the form

$$L(x_*, w_*) \leqslant L(x,w) + < g(x), w_* - w > . \qquad (28)$$

From this inequality and (6), we obtain

$$B(g(x)) \geqslant (L(x_*, w_*) - L(x,w))/B^0(w_* - w|g(P)).$$

This inequality permits us to construct a new class of EAF based on the Lagrange function $L(x,y)$. We introduce the function

$$H(x,y) = A(L(x,w),v) + B(g(x)), \qquad (29)$$

where $y = [w,v] \in Y \sqsubseteq R^m_+ \times R^1$. The function (29) was considered earlier by Skarin (1973) in the case where $A(L,v) = = v^1 L$. Denote

$$W_Y = \{w \in R^m_+ | \exists v \in R^1, [w,v] \in Y\},$$

$$V_w = \{v \in R^1 | [w,v] \in Y\}.$$

The set W_Y is the projection Y onto R_+^m , the set V_w is the section of Y. We impose the following condition

$$L(x_*, w) = L(x_*, w_*) \ \ \forall w \in W_Y. \qquad (30)$$

If $w_* \neq 0$ and $W_Y = W_* = \{0 \leqslant w \leqslant w_* | w^\xi < w_*^\xi \text{ if } w_*^\xi > 0\}$, then the inequality (30) holds.

Instead of a), b) we introduce the following conditions:

c) $A(L, v) - A(L_*, v) \geqslant (L - L_*)/B^0(w_* - w|g(P)) \ \forall L \in R^1, \forall w \in W_Y, \forall v \in V_w$, where $L_* = L(x_*, w_*) = f_*$;

d) for any point $y \in Y$ the solution set of the system $A(L(x, w), v) + B(g(x)) = A(L(x_*, w_*), v), \quad x \in P$ coincides with X_*.

Theorem 6. Let conditions (30), c), d) hold. Assume that the function $B(g(x))$ is a strict exterior or interior penalty function (in later case $P = X$) and $0 < B^0(w_* - w|g(P)) < +\infty$ for any $w \in W_Y$. Then the function (29) is EAF on $P \times Y$.

The proof is basically the same as the proof of Theorems 1 and 2.

If $B(g(x))$ is an exterior penalty function,then the condition c) can be simplified and instead of c) we can use the following inequality

$$A(L, v) - A(L_*, v) \geqslant [(L - L_*)/B^0(w_* - w|g(P))]_-$$

$$\forall L \in R^1, \ \forall w \in W_Y, \ \forall v \in V_w.$$

The simplest examples of the EAF, which satisfy Theorem 6,can be obtained from the second section. Let us substitute $L(x, w)$ for $f(x)$ in H_1, H_2, H_3, H_4. Such functions are EAF on $R^n \times Y$,where $Y = \{[w, v]: w \in W_*, \ v \in V_w\}$. If $B(g(x))$ is a strict exterior penalty function,then we have

$H_9(x, y) = v^{-1} L(x, w) + B(g(x))$,

$V_w = \{v | v > B^0(w_* - w|R^m)\}$,

$H_{10}(x, y) = v^{-1} e^{L(x,w)} + B(g(x))$,

$$V_w = \{v | v > B^0(w_* - w | R^m) e^{f_*}\},$$

$$H_{11}(x,y) = (v - L(x,w))_+^\alpha + B(g(x)),$$
$$V_w = \{v | v > f_* + (-\alpha B^0(w_* - w | R^m))^{1/(1-\alpha)}\},$$

$$H_{12}(x,y) = (L(x,w) - y)_+^\beta + B(g(x)),$$
$$V_w = \{v | f_* \geqslant v > f_* - (\beta B^0(w_* - w | R^m))^{1/(1-\beta)}\}$$

Now we consider nonlinear functions. Let

$$H(x,y) = G(A(L(x,w),v), B(g(x))), \tag{31}$$

where $G(t,\tau)$ is a continuous nondecreasing function of two variables, $A(L,v)$ is a monotonically nondecreasing function of L. We assume that the function A and the set Y are such that for any $w \in W_Y$ scalars $C(w)$, $D(w)$ exist for which

$$0 < \sup_{y \in Y} \sup_{\xi \in \partial A_L(L_*,v)} \xi \leqslant D(w) < +\infty, \tag{32}$$

$$0 < C(w) \leqslant \sup_{y \in Y} \sup_{\xi \in \partial A_L(L_*,v)} \xi < +\infty. \tag{33}$$

Moreover, we suppose that

$$L(x,w) > L(x_*,w_*) = L_* \quad \forall w \in W_Y, \quad \forall x \in X \backslash X_*. \tag{34}$$

Instead of (24), (27),we impose on a function G the following conditions: for any $w \in W_Y$ there exists such a set $T(w) \in R^1$, that $A(L_*,v) \in T(w)$ for all $v \in V_w$ and for all $t_* \in T(w)$

$$G(t, (t_* - t)_+ / D(w) B^0(w_* - w | R^m)) > G(t_*, 0) \;\forall t \neq t_*, \tag{35}$$

$$G(t, (t_* - t) / C(w) B^0(w_* - w | R^m_-)) > G(t_*, 0) \;\forall t > t_*. \tag{36}$$

Theorem 7. Let $A(L, v)$ be a convex monotonically increasing function of L and let conditions (30), (34) hold. If $B(g(x))$ is a strict exterior penalty function, if $0 < B^0(w_* - w|R^m) < +\infty$ for any $w \in W_Y$ and (32), (35) hold, then (31) is EAF on $P \times Y$. If $B(g(x))$ is an interior penalty function, if $0 < B^0(w_* - w|R^m_-) < +\infty$ for all $w \in W_Y$ and (33), (36) hold then (31) is EAF on $X \times Y$.

If $A(L, v) = L - v$, then it follows from Theorem 7 that $V_{w^*} = f_* - T(w)$ for any $w \in W_Y$, therefore, we have the following EAF

$$H_{13}(x, y) = (L(x, w) - v)_- / [1 - (L(x, w) - v)_- B(g(x))].$$

If $B(g(x))$ is a strict exterior penalty function, then H_{13} is EAF on $P \times Y_{13}$, where

$$Y_{13} = \left\{[w, v] | w \in W_*,\ v > f_* + \sqrt{B^0(w_* - w|R^m)}\right\}.$$

If $B(g(x))$ is an interior penalty function, then H_{13} is EAF on $X \times Y_{13}$, where

$$Y_{13} = \left\{[w, v] | w \in W_*,\ f_* \leqslant v < f_* + \sqrt{B^0(w_* - w|R^m_-)}\right\}.$$

It is possible to consider EAF with a strict mixed penalty function. Sufficient conditions remain the same except for (35), which must be valid only for $t < t_*$. If in the problem (1) exists more than one saddle point of Lagrange function, then in the formulae for Y we can substitute w, which has a minimum norm in the set of Lagrange multipliers.

5. Summary and concluding remarks. In this paper we introduced a notion of an exact auxiliary function and presented classes of EAF. We believe that the investigation of EAF will be extremely useful for numerical methods and theoretical studies. Our preliminary computational results are encouraging. Of course, an efficient implementation of various EAF will require much more work. At present we write a book devoted to nonlinear programming. We plan to describe the main numerical methods on the base of EAF notion and its extensions. The penalty method, method of center, barrier method and some other methods we consider as an investigation of different auxiliary convolution functions.

REFERENCES

Bertsekas, D.P. (1975). Necessary and sufficient conditions for a penalty method to be exact. *Math. Program.,* **9**, 87–99.

Charalambos, Ch. (1976). A negative-positive barrier method for non-linear programming. *Int. J. Systems Sci.,* **7**(5), 557–575.

Eremin, I.I. (1967). The method of penalties in convex programming. *Dokl. Akad. Nauk SSSR,* **173**(4), 748–751 (in Russian).

Han, S.-P., and O.L.Mangasarian (1979). Exact penalty functions in nonlinear programming. *Math. program.,* **17**, 251–269.

Evtushenko, Yu.G. (1987). Accuracy estimations for penalty function methods. In O.M Belotserkovskii (Ed.), *Problemy Prikladnoi Matematiki i Informatiki.* Nauka, Moscow. pp. 199–208 (in Russian).

Evtushenko, Yu.G. (1985). *Numerical Optimization Techniques, Optimization Software.* Inc. Publications Division, New York. 558.pp.

Evtushenko, Yu.G., and V.G. Zhadan (1988). *Systematization of Nonlinear Programming Methods, Sequential Unconstrained Minimization Techniques.* Computer Center of the USSR Academy of Sciences, Moscow. 65pp. (in Russian).

Lootsma, F. (1974). Convergence rate of quadratic exterior penalty function method for solving constrained minimization problems. *Philips research report,* **29**(1), 2–18.

Morrison, D.D. (1968). Optimization by least squares. SIAM J.

Numer. Analysis, **5**(1), 83–88.

Scarin, V.D. (1973). On a modification of penalty function method in convex programming. In I.I.Eremin (Ed.), *Nelineinaja optimizatsija i prilozhenija v planirovanii.* The Ural Research Center of the USSR Academy of Sciences, Sverdlovsk. pp. 51–62 (in Russian).

Zangwill, S.-P. (1967)). Non-linear programming via penalty function. *Manag.Sci.,* **5**, 344–358.

Zhadan, V.G. (1984). On some estimates of the penalty coefficient in methods of exact penalty functions. *USSR Comput. Maths. Math. Phys.,* **24**(4), 124–128.

Received November 1989

Yu. Evtushenko received the Degrees of Candidate and Doctor of Physical and Mathematical Sciences from the Moscow Institute of Physics and Technology in 1965 and 1981, respectively. He is a director of the Computer Center of the USSR Academy of Sciences. His research interests include nonlinear programming, optimal control, global search, multicriteria optimization.

V. Zhadan received the Degree of Candidate of Physical and Mathematical Sciences from the Moscow Institute of Physics and Technology in 1977. He heads the Department of Optimization Methods at the Computer Center of the USSR Academy of Sciences. His research interests include nonlinear programming and multicriteria optimization.