U.S.S.R. Comput. Maths. Math. Phys., Vol.30, No.1, pp.31–42, 1990 0041-5553/90 \$10.00+0.00
Printed in Great Britain

EXACT AUXILIARY FUNCTIONS IN OPTIMIZATION PROBLEMS[1]

YU.G. YEVTUSHENKO and V.G. ZHADAN

(Revised 20 February 2002)

The concept of an exact auxiliary function such that the problem of minimizing it has the same set of solutions as the original optimization problem. Sufficient conditions are given for the auxiliary functions to be exact and examples of such functions are described. The introduction of exact auxiliary functions makes it possible to reduce the solution of the original problem to single minimization of an auxiliary function. The constrained optimization problem often reduces to unconstrained optimization.

1. Basic definitions

Consider the following nonlinear programming problem: find

$$f_* = \min_{x \in X} f(x), \qquad X = \{x \in \mathbb{E}^n \mid g(x) \le 0\}. \tag{1.1}$$

Here $\mathbb{E}^n$ is n-dimensional Euclidean space, $f(x)$ and $g(x)$ are continuous functions $f : \mathbb{E}^n \to \mathbb{E}^1$, $g : \mathbb{E}^n \to \mathbb{E}^m$. A solution of problem (1.1) is any point x_* from the set

$$X_* = \{x_* \in X | f(x) - f(x_*) \ge 0 \quad \forall x \in X\},$$

which in what follows is assumed to be nonempty. All the results of this paper were obtained without assuming the convexity and differentiability of the functions f and g with respect to x.

We introduce the function $R(x, y)$ that depends on the original variables $x \in \mathbb{E}^n$ and on some vector y from the set Y. The length of y and the form of the set Y are not specified at this stage. Consider the auxiliary minimization problem

$$\min_{x \in P} R(x, y), \tag{1.2}$$

where P is a closed set in $\mathbb{E}^n$ containing X_*. In particular, P may be the entire space $\mathbb{E}^n$, the feasible set X, or a part of the feasible set.

Assuming that a solution of problem (1.2) exists, we define the set-valued mapping

$$X(y) = \operatorname{Arg}\min_{x \in P} R(x, y).$$

Definition 1. *The function $R(x, y)$ is an exact auxiliary function (e.a.f.) for problem* (1.1) *on $P \times Y$ if $X(y) \ne \emptyset$ and $X(y) = X_*$ for any $y \in Y$.*

The study e.a.f.'s is highly important for constructing numerical methods, because a knowledge of such functions produces a solution of the original problem (1.1) via a single minimization (1.2) of the auxiliary function. It is desirable to have the set Y as "large" as possible, to avoid

[1] *Zh. vȳchisl. Mat. mat. Fiz.*, **30**, 1, 43–57, 1990

additional difficulties with the determination of points from Y. Functions for which Y consists of a single point are very inconvenient in this respect.

The first e.a.f.'s were constructed in [1, 2], where so-called exact penalty functions were discovered (these are a variety of e.a.f.'s). Subsequently, these functions were studied by numerous mathematicians, but, as far as we know, no fundamentally new e.a.f.'s have been proposed, with the exception of [3, 4, 5]. In this paper, we develop an approach proposed by the present authors in [6, 7, 8]. Sufficient conditions of e.a.f. are given and examples of such functions are described.

E.a.f.'s are conveniently constructed as a lower bounded function. Let $R(x,y)$ be an e.a.f. for problem (1.1) on the set $P \times Y$. Consider the function

$$M(x,y) = R(x,y) - R(x_*,y), \qquad x_* \in X_* \subseteq P. \tag{1.3}$$

For any $x \in P$, $y \in Y$, $x_* \in X_*$ the function M is nonnegative. The set of points x on which the function $M(x,y)$ reaches a minimum over x on P is identical with the set $X(y)$ and if $y \in Y$ then also with the set X_*. It is also the set of solutions of the equation $M(x,y) = 0$ that belong to P for each fixed vector y from Y. The choice of a specific vector x_* from X_* does not affect the values of the function $M(x,y)$ for $y \in Y$, because $R(x,y)$ takes the same value for all x from X. Therefore,

$$M(x,y) \equiv 0 \qquad \forall x \in X(y) = X_*, \qquad \forall y \in Y.$$

If $R(x,y)$ is an e.a.f., then the function $M(x,y)$ is an e.a.f. The converse is also true: if the function $M(x,y)$ constructed in the form (1.3) is an e.a.f., then $R(x,y)$ is an e.a.f. Therefore, instead of proving that $R(x,y)$ is an e.a.f., it suffices to show that the corresponding $M(x,y)$ is an e.a.f.

For the function $M(x,y)$ (1.3) to be an e.a.f. on the set $P \times Y$, it is necessary and sufficient to have the following two conditions:

$$M(x,y) \geq 0 \qquad \forall x \in P, \qquad \forall y \in Y, \tag{1.4}$$

$$\forall y \in Y \text{ and } M(x,y) = 0 \text{ it follows that } x \in X(y) = X_*. \tag{1.5}$$

If the function $M(x,y)$ is differentiable with respect to y and the set Y is open, then it is necessary that

$$M_y(x,y) = 0 \qquad \forall x \in X(y) = X_*, \qquad \forall y \in Y.$$

This condition may be added to (1.5), and simultaneously using these conditions we can often show that $X(y) = X_*$.

Multiplication of an e.a.f. by any positive number leaves the auxiliary function exact. The sum of exact auxiliary functions on the same set $P \times Y$ is an e.a.f. on the same set. If the sets $P \times Y$ are different, then we should take their intersection.

Let the function $Q(z)$ be defined on the set $Z \subseteq \mathbb{E}^s$; its polar function $Q^0(z_0|Z)$ on the set $Z_0 \subseteq \mathbb{E}^s$ is defined as

$$Q^0(z_0|Z) = \inf_{\mu \in M(z_0)} \mu, \qquad M(z_0) = \{\mu \in \overline{\mathbb{E}^1} | \langle z, z_0 \rangle \leq \mu Q(z) \ \ \forall z \in Z\}, \tag{1.6}$$

where $z_0 \in Z_0$, $\overline{\mathbb{E}^1}$ is the extended real line $\mathbb{E}^1$, i.e. the line $\mathbb{E}^1$ with the elements $\{+\infty\}$ and $\{-\infty\}$. From this definition we obtain the Minkowski–Mahler inequality

$$\langle z, z_0 \rangle \leq Q(z) Q^0(z_0|Z) \qquad \forall z \in Z, \qquad \forall z_0 \in Z_0. \tag{1.7}$$

If the function $Q(z)$ takes only positive values on Z, then alongside (1.6) we may use the following definition:

$$Q^0(z_0|Z) = \sup_{z\in Z} \frac{\langle z, z_0\rangle}{Q(z)}. \tag{1.8}$$

If $Q(z)$ is nonpositive on Z, then

$$Q^0(z_0|Z) = \inf_{z\in Z} \frac{\langle z, z_0\rangle}{Q(z)}. \tag{1.9}$$

Denote by $\mathbb{E}^m_+$ and $\mathbb{E}^m_-$ the nonnegative and nonpositive orthants in $\mathbb{E}^m$, i.e. the collection of all vectors from $\mathbb{E}^m$ whose coordinates are, respectively, nonnegative and nonpositive.

For each scalar function $\varphi(z)$ of a vector argument z, define

$$\varphi_+(z) = \max[0, \varphi(z)], \qquad \varphi_-(z) = \min[0, \varphi(z)].$$

Similarly, for the vectors $z \in \mathbb{E}^s$

$$\begin{aligned} z_+ &= [z^1_+, \dots, z^s_+], & z^i_+ &= \max[0, z^i], \\ z_- &= [z^1_-, \dots, z^s_-], & z^i_- &= \min[0, z^i]. \end{aligned}$$

Define the p-th Hölder norm of the vector z:

$$\|z\|_p = \left(\sum_{i=1}^{s} |z^i|^p\right)^{1/p}, \qquad p \geq 1. \tag{1.10}$$

The conjugate norm is $\|z\|_{p_*}$, where $p^{-1} + p_*^{-1} = 1$.

The function (1.10) will be considered for nonzero $p < 1$ and also for $p = 0$ and $p = \pm\infty$, where it is defined as follows:

$$\|z\|_0 = s^{1/2}\left(\prod_{i=1}^{s} |z^i|\right)^{1/s}, \quad \|z\|_{+\infty} = \max_{1\leq i\leq s} |z^i|, \quad \|z\|_{-\infty} = \min_{1\leq i\leq s} |z^i|.$$

For $p < 0$ we assume that the function (1.10) equals zero if at least one of the coordinates of the vector z is zero.

Consider some examples for $Z_0 = \mathbb{E}^s_+$:

$$\begin{array}{llllll} p > 1, & p_* > 1, & Q(z) = & \|z_+\|_p, & Q^0(z_0|\mathbb{E}^s) & = & \|z_0\|_{p_*}, \\ p < 1, & p_* < 1, & Q(z) = & -\|z_-\|_p, & Q^0(z_0|\mathbb{E}^s_-) & = & \|z_0\|_{p_*}. \end{array}$$

Here p and p_* are related by the same dependence as before, $p^{-1} + p_*^{-1} = 1$. Some special cases are the following:

$$\begin{array}{llllll} p = 0, & p_* = 0, & Q(z) = & -\|z_-\|_0, & Q^0(z_0|\mathbb{E}^s_-) & = & \|z_0\|_0, \\ p = +\infty, & p_* = 1, & Q(z) = & \|z_+\|_{+\infty}, & Q^0(z_0|\mathbb{E}^s) & = & \|z_0\|_1, \\ p = -\infty, & p_* = 1, & Q(z) = & -\|z_-\|_{-\infty}, & Q^0(z_0|\mathbb{E}^s_-) & = & \|z_0\|_1. \end{array}$$

Introduce the vector $w \in \mathbb{E}^m_+$ and form the Lagrange function

$$L(x, w) = f(x) + \langle w, g(x)\rangle.$$

We say that the Lagrange function has a saddle point $[x_*, w_*] \in P \times \mathbb{E}^m_+$ if

$$L(x_*, w) \leq L(x_*, w_*) \leq L(x, w_*) \qquad \forall x \in P, \quad \forall w \in \mathbb{E}^m_+. \tag{1.11}$$

If a vector $w_* \in \mathbb{E}^m_+$ exists such that x_* and w_* form a saddle point of the Lagrange function, then $x_* \in X_*$ [9].

Define the composite function $B(g(x))$. We denote by $g(P)$ by the image of the set P under the mapping g. Then, using inequality (1.7), we obtain from the right-hand side of (1.11)

$$f_* \le L(x, w_*) \le f(x) + B(g(x))B^0(w_*|g(P)) \quad \forall x \in P. \tag{1.12}$$

If $B(g(x)) \ge 0$ for all $x \in \mathbb{E}^n$, then this bound can be relaxed by introducing the polar functions $B^0(w_*|g(\mathbb{E}^n))$ and $B^0(w_*|\mathbb{E}^m)$, that are easier to compute, because by (1.8)

$$B^0(w_*|g(P)) \le B^0(w_*|g(\mathbb{E}^n)) \le B^0(w_*|\mathbb{E}^m).$$

Therefore, in what follows inequality (1.12) is written in the form

$$f_* \le L(x, w_*) \le f(x) + B(g(x))B^0(w_*|\mathbb{E}^m) \qquad \forall x \in P. \tag{1.13}$$

Similarly, if $B(g(x)) \le 0$ on $\mathbb{E}^n$, then from (1.9) we obtain

$$B^0(w_*|g(P)) \ge B^0(w_*|g(\mathbb{E}^n)) \ge B^0(w_*|\mathbb{E}^m)$$

and inequality (1.13) remains valid as before.

2. Additive exact auxiliary functions

We will construct e.a.f.'s in the form

$$R(x, y) = A(f(x), y) + B(g(x)), \tag{2.1}$$

where $A(f, y)$ is an arbitrary continuous function of two arguments, and $B(g(x))$ is a strictly exterior penalty function, i.e. $B(g(x))$ is continuous and takes nonnegative values everywhere on $\mathbb{E}^n$; moreover, $B(g(x)) = 0$ if and only if $x \in X$. A well-known example of a strictly exterior penalty function is

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

We will impose two further conditions on these functions.

Condition A. The functions A and B are such that

$$A(f, y) - A(f_*, y) \ge [(f - f_*)/B^0(w_*|\mathbb{E}^m)]_- \qquad \forall y \in Y, \quad \forall f \in \mathbb{E}^1. \tag{2.2}$$

Condition B. For every point y from the set Y, the set of solutions of the system

$$A(f(x), y) + B(g(x)) = A(f_*, y), \qquad x \in P, \tag{2.3}$$

is identical with X_*.

If A is differentiable with respect to y and Y is an open set, then from (2.3) we obtain the system

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

which is often easier to solve than the original system (2.3).

Theorem 1. *Assume that the Lagrange function of problem* (1.1) *has a saddle point* $[x_*, w_*] \in P \times \mathbb{E}^m_+$. *Also assume that* $B(g(x))$ *is a strictly exterior penalty function,* $0 <$

$< B^0(w_*|\mathbb{E}^m) < +\infty$, *and the functions A and B satisfy* **conditions A** *and* **B**. *Then the function $R(x,y)$ defined by* (2.1) *is an e.a.f. for problem* (1.1) *on the set $P \times Y$.*

Proof. Seeing that $B(g(x))$ vanishes on X_*, we obtain

$$M(x,y) = A(f(x),y) - A(f_*,y) + B(g(x)). \tag{2.4}$$

Condition (1.4) may be written in this case in the form

$$A(f(x)) + B(g(x)) \geq A(f_*,y) \qquad \forall x \in P, \quad \forall y \in Y. \tag{2.5}$$

By the conditions of the theorem, $B^0(w_*|\mathbb{E}^m) > 0$, and, therefore, from (1.13) we obtain

$$B(g(x)) \geq [f_* - f(x)]/B^0(w_*|\mathbb{E}^m) \qquad \forall x \in P.$$

Seeing that the function B takes only nonnegative values, we can refine its lower bound:

$$B(g(x)) \geq [[f_* - f(x)]/B^0(w_*|\mathbb{E}^m)]_+ \qquad \forall x \in P. \tag{2.6}$$

We rewrite this inequality in the form

$$B(g(x)) \geq -[[f(x) - f_*]/B^0(w_*|\mathbb{E}^m)]_- \qquad \forall x \in P.$$

Applying it together with (2.2) to the right-hand side of (2.4), we obtain (1.4). **Condition B** ensures that condition (1.5) is satisfied. Thus, the function (2.4) is an e.a.f. on the set $P \times Y$ and, therefore, the function (2.1) is also an e.a.f. on the same set $P \times Y$. The theorem is proved. □

A wide class of functions satisfies the conditions of the theorem. We do not assume an exact knowledge of f_*, and, therefore, the function A should be constructed so that **conditions A** and **B** are satisfied for any f_*. From (2.2) it follows that to this end $A(f,y)$ should at least be an increasing function of f. Moreover, its graph should not lie below the convex cone with origin at the point $N = [f_*, A(f_*,y)]$ and with two boundary rays, the first of which originates from the point N and points in the positive f-direction and the second originates from N and points along the halfline

$$D = (f - f_*)/B^0(w_*|\mathbb{E}^m), \qquad f \leq f_*.$$

Also note that the assertion of the theorem remains unchanged if **condition A** is satisfied not for all $f \in \mathbb{E}^1$ but only for f from the image $f(P)$ of the set P under the mapping $f(x)$.

Let $A(f,y)$ be convex in f for each $y \in Y$. Then

$$A(f,y) - A(f_*,y) \geq \xi(f - f_*) \qquad \forall y \in Y, \quad \forall f \in \mathbb{E}^1, \quad \forall \xi \in \partial_f A(f_*,y). \tag{2.7}$$

Here $\partial_f A(f_*,y)$ is the subdifferential of the function $A(f,y)$ with respect to f at the point f_*. Condition (2.2) is satisfied if

$$\xi(f - f_*) \geq [(f - f_*)/B^0(w_*|\mathbb{E}^m)]_- \qquad \forall y \in Y, \quad \forall f \in \mathbb{E}^1, \quad \forall \xi \in \partial_f A(f_*,y).$$

Together with the assumption that $A(f,y)$ is nondecreasing in f, this condition is equivalent to the following:

$$0 \leq \inf_{\xi \in \partial A_f(f_*,y)} \xi \leq \sup_{\xi \in \partial A_f(f_*,y)} \xi \leq 1/B^0(w_*|\mathbb{E}^m). \tag{2.8}$$

These inequalities may be used to determine the set Y, additionally considering the boundary points of Y. Note that if, in addition to (2.8), the function $B(g)$ is convex and nondecreasing in g and the functions $f(x)$ and $g(x)$ are convex in x, then the auxiliary function $R(x, y)$ is also convex in x.

As an example of e.a.f.'s that satisfy the conditions of Theorem 1 and conditions (2.8) with $P = \mathbb{E}^n$, $Y \subset \mathbb{E}^1$, we give the following functions and the corresponding sets Y:

$$\begin{array}{ll} R_1(x,y) = y^{-1}f(x) + B(g(x)), & Y_1 = \{y : y > B^0(w_*|\mathbb{E}^m)\}, \\ R_2(x,y) = y^{-1}e^{f(x)} + B(g(x)), & Y_2 = \{y : y > B^0(w_*|\mathbb{E}^m)e^{f_*}\}, \\ R_3(x,y) = [y - f(x)]^{\alpha} + B(g(x)), & Y_3 = \{y : y > f_* + [-\alpha B^0(w_*|\mathbb{E}^m)]^{1/(1-\alpha)}\}, \\ R_4(x,y) = [f(x) - y]_+^{\beta} + B(g(x)), & Y_4 = \{y : f_* \geq y > f_* - [\beta B^0(w_*|\mathbb{E}^m)]^{1/(1-\beta)}\}, \end{array}$$

where $\alpha < 0$, $\beta > 1$. For R_3 we should take $R_3(x, y) = +\infty$ when $f(x) \geq y$. If we know that $f(x) > 0$ for any $x \in \mathbb{E}^n$, then we may use the functions

$$\begin{array}{ll} R_5(x,y) = y^{-1}\mathrm{sh} f(x) + B(g(x)), & Y_5 = \{y : y > B^0(w_*|\mathbb{E}^m)\mathrm{ch} f_*\}, \\ R_6(x,y) = y^{-1}[f(x)]^{\gamma} + B(g(x)), & Y_6 = \{y : y > \gamma f_*^{\gamma-1}B^0(w_*|\mathbb{E}^m)\}. \end{array}$$

If $f(x)$ is nonpositive on $\mathbb{E}^n$, we may take the function

$$R_7(x,y) = y^{-1}\mathrm{arctg} f(x) + B(g(x)), \quad Y_7 = \{y : y > B^0(w_*|\mathbb{E}^m)/(1 + f_*^2)\}.$$

The function R_7 is an e.a.f. also when $f(x)$ takes arbitrary values, but then the set Y_7 should be replaced with $Y_7 = \{y : y > B^0(w_*|\mathbb{E}^m)\}$. Also note that the parameters α, β and γ entering R_3, R_4, and R_6 may be considered as second components of the vector y.

The first of these functions is the well-known exact penalty function [1, 2, 9, 10]. The function R_4 has been used by many mathematicians for successive unconstrained minimization, constructing a sequence of vectors y_k that converge to f_* (see [9, 11]). If we take the closure of the sets Y_1 – Y_7, then the functions R_1 – R_7 are no longer e.a.f.'s because instead of (1.5) they satisfy a weaker condition: $M(x, y) = 0$ implies that $X_* \subseteq X(y)$.

Put $X^0 = \{x \in \mathbb{E}^n | g(x) < 0\}$ and consider the case when $B(g(x))$ is an interior penalty function, i.e. $B(g(x))$ is continuous and nonpositive on $\mathbb{E}^m$ and $B(g(x)) < 0$ for $x \in X^0$. If we additionally assume that $B(g(x)) < 0$ if and only if $g(x) < 0$, then such a function will be called a strictly interior penalty function. As an example of interior penalty functions, consider

$$B(g(x)) = -\|g_-(x)\|_p, \qquad -\infty \leq p < 1. \tag{2.9}$$

For $p \leq 0$ this function is a strictly interior penalty function.

Below for all interior penalty functions we assume that the constraints are regular, i.e. the set X^0 is nonempty and its closure coincides with X.

We construct an e.a.f. in the form (2.1), taking P as the feasible set X. Condition (2.2) is replaced by

$$A(f, y) - A(f_*, y) \geq (f - f_*)/B^0(w_*|\mathbb{E}_-^m) \qquad \forall y \in Y, \quad \forall f \geq f_*. \tag{2.10}$$

Theorem 1 in this case can be restated as follows.

Theorem 2. *Assume that the Lagrange function in problem* (1.1) *has a saddle point* $[x_*, w_*]$. *Also assume that the function* $B(g(x))$ *in* (2.1) *is an interior penalty function and* $0 < B^0(w_*|\mathbb{E}_-^m) < +\infty$. *For* $R(x, y)$ *to be an e.a.f. on* $X \times Y$ *it is sufficient that it satisfies condition* (2.10) *and that any feasible solution* x *of system* (2.3) *is such that* $x \in X_*$.

Proof. Since the Lagrange function has a saddle point, $B^0(w_*|\mathbb{E}_-^m) > 0$ and $B(g(x)) \leq 0$ on X, we have

$$f_* \leq f(x) + B(g(x))B^0(w_*|\mathbb{E}_-^m) \leq f(x) \qquad \forall x \in X.$$

Hence, like (2.6),

$$[f_* - f(x)]/B^0(w_*|\mathbb{E}^m_-) \leq B(g(x)) \leq 0 \qquad \forall x \in X. \tag{2.11}$$

Taking an arbitrary point $x \in X_*$, on the left-hand side, we obtain $B(g(x)) = 0$. Thus, $B(g(x))$ equals zero everywhere on X_*. Therefore, to have (1.4) it is sufficient that inequality (2.5) holds for all $x \in X$. But by (2.11), (2.5) implies (2.10).

The validity of (1.5) follows from the assumption that any feasible solution $x_* \in X$ of system (2.3) is such that $x_* \in X_*$, because, as we have established, $B(g(x_*)) = 0$ when $x_* \in X_*$. The theorem is proved. $\square$

We see from the proof of Theorem 2 that the condition $B^0(w_*|\mathbb{E}^m_-) > 0$ leads to the equality $B(g(x_*)) = 0$ at the point $x_* \in X_*$. Thus, by the regularity of the constraints, the point x_* is necessarily a boundary point of the set X in this case. If $B(g(x))$ is the function (2.9) with $0 < p < 1$, then $B(g(x))$ is an interior (but not strictly interior) penalty function and the inequality $B^0(w_*|\mathbb{E}^m_-) = \|w_*\|_{p_*} > 0$ holds if and only if all the components of the vector w_* are strictly positive. This means that all the constraints simultaneously vanish at the point x_*, or in other words are active. If (2.9) is a strictly interior penalty function (which is so when $p \leq 0$), then to have the inequality $\|w_*\|_{p_*} > 0$ it is sufficient that at least one constraint is active.

Note that if A is a convex function of f for each $y \in Y$, then to have inequality (2.10) it is sufficient that

$$\xi \geq 1/B^0(w_*|\mathbb{E}^m_-) \qquad \forall \xi \in \partial_f A(f_*, y), \quad \forall y \in Y.$$

The functions R_1 – R_4, satisfy the conditions of Theorem 2, but instead of a strictly exterior penalty function $B(g(x))$ we should use an interior penalty function (e.g. (2.9)) and replace Y_i with the following sets:

$$\begin{array}{l}
Y_1 = \{y : 0 < y < B^0(w_*|\mathbb{E}^m_-)\}, \\
Y_2 = \{y : 0 < y < e^{f_*} B^0(w_*|\mathbb{E}^m_-)\}, \\
Y_3 = \{y : f_* < y < f_* + [-\alpha B^0(w_*|\mathbb{E}^m_-)]^{1/(1-\alpha)}\}, \quad \alpha < 0, \\
Y_4 = \{y : y < f_* - [\beta B^0(w_*|\mathbb{E}^m_-)]^{1/(1-\beta)}\}, \quad \beta > 1.
\end{array}$$

Also note that since under the assumptions of Theorem 2 the function $B(g(x))$ is nonpositive and vanishes on X_*, then the auxiliary function $R(x, y) = f(x) - yB(g(x))$ is also an e.a.f. on $X \times \mathbb{E}^1_+$.

We will introduce a new class of penalty functions which have the properties of both strictly exterior and strictly interior penalty functions. A continuous function $B(g(x))$ is called a strictly mixed penalty function if $B(g(x)) > 0$ when $x \notin X$, and $B(g(x)) < 0$ when $x \in X^0$.

As an example of strictly mixed penalty functions, consider the following two functions:

$$B(g(x)) = \|g_+(x)\|_p - \|g_-(x)\|_{-p}, \quad 1 < p < +\infty, \tag{2.12}$$

$$B(g(x)) = \max_{1 \leq i \leq m} g^i(x). \tag{2.13}$$

We alter the conditions (2.2) and (2.10) taking for any $y \in Y$

$$A(f, y) - A(f_*, y) \geq \begin{cases} (f - f_*)/B^0(w_*|\mathbb{E}^m_-), & \text{if } f \geq f_*, \\ (f - f_*)/B^0(w_*|\mathbb{E}^m), & \text{if } f < f_*. \end{cases} \tag{2.14}$$

We have the following theorem.

Theorem 3. *Assume that the Lagrange function in problem* (1.1) *has a saddle point* $[x_*, w_*]$. *Also assume that* $B(g(x))$ *is a strictly mixed penalty function and* $0 < B^0(w_*|\mathbb{E}^m) <$

$< B^0(w_*|\mathbb{E}^m_-) < +\infty$. *For $R(x,y)$ to be an e.a.f. on $P \times Y$, it is sufficient that conditions* (2.14) *and B hold.*

Proof. Since $B^0(w_*|\mathbb{E}^m_-) > 0$, then like the proof of Theorem 2 we show that $B(g(x_*)) = 0$. Therefore, (1.4) holds by (2.14). Condition (1.5) follows from (2.3). The theorem is proved. □

Condition (2.14) for a smooth function holds if and only if $B^0(w_*|\mathbb{E}^m) \le B^0(w_*|\mathbb{E}^m_-)$. Thus, for instance, the function $R_1(x,y)$ with a strictly mixed penalty function $B(g(x))$ is an e.a.f. The corresponding set Y_1 for this function is

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

In particular, if $B(g(x))$ has the form (2.12), we obtain

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

where $p^{-1} + p_*^{-1} = 1$, $-p^{-1} + r_*^{-1} = 1$. Since $p_* > 1$, $1 > r_* > 1/2$, then $\|w_*\|_{r_*} > \|w_*\|_{p_*}$. For the function (2.13) the set Y_1 is empty.

3. Nonlinear exact auxiliary functions

Now assume that the auxiliary function $R(x,y)$ is constructed in the form

$$R(x,y) = H(A(f(x),y), B(g(x))), \tag{3.1}$$

where $B(g(x))$ is a strictly exterior penalty function, and $H(t,\tau)$ is a continuous nondecreasing function of two arguments, i.e. $H(t_1,\tau_1) \ge H(t,\tau)$ for any $t_1 \ge t$, $\tau_1 \ge \tau$. Regarding the function $A(f,y)$ we will assume that it is a monotone increasing and convex function of the first argument. We will also assume that there exists a constant D such that

$$0 < \sup_{y \in Y} \sup_{\xi \in \partial A_f(f_*,y)} \xi \le D < +\infty. \tag{3.2}$$

We will give sufficient conditions for the function (3.1) to be an e.a.f. for problem (1.1). We put $N_1 = DB^0(w_*|\mathbb{E}^m)$.

Theorem 4. *Assume that the Lagrange function of problem* (1.1) *has a saddle point $[x_*, w_*]$ and $A(f,y)$ is a convex and monotone increasing function of the first argument which satisfies* (3.2). *Also assume that $B(g(x))$ is a strictly exterior penalty function and $0 < N_1 < +\infty$. If there exists a set $T \subseteq \mathbb{E}^1$, such that $A(f_*,y) \in T$ for any $y \in Y$ and*

$$H(t, (t_* - t)_+/N_1) > H(t_*, 0) \qquad \forall t_* \in T, \quad \forall t \ne t_*, \tag{3.3}$$

then (3.1) *is an e.a.f. for problem* (1.1) *on the set $P \times Y$.*

Proof. Take an arbitrary point $x \in P$. Since $H(t,\tau)$ is nondecreasing in the second argument, we obtain from (2.6)

$$R(x,y) = H(A(f(x),y), B(g(x))) \ge H(A(f(x),y), [f_* - f(x)]_+/B^0(w_* \mid \mathbb{E}^m)). \tag{3.4}$$

By the convexity of $A(f,y)$ in f, we have the inequality (2.7), which, combined with (3.2), gives

$$A(f_*,y) - A(f,y) \le D(f_* - f) \qquad \forall f \le f_*. \tag{3.5}$$

But $A(f,y)$ is monotone increasing in f, and, therefore, alongside (3.5) we have the inequality

$$(A(f_*,y) - A(f,y))_+ \le D(f_* - f)_+ \qquad \forall f \in \mathbb{E}^1.$$

Substituting it into (3.4) and using condition (3.3), we obtain for the case when $A(f(x), y) \neq$ $\neq A(f_*, y)$

$$R(x,y) \geq H(A(f(x),y), [A(f_*,y) - A(f(x),y)]_+/N_1) > H(A(f_*,y),0) = R(x_*,y).$$

Now assume that $A(f(x), y) = A(f_*, y)$. In this case, if $x \notin X_*$, then necessarily $x \notin X$, and $B(g(x)) > 0$. Therefore, by the monotonicity of $H(t,\tau)$ in t and (3.3),

$$\begin{aligned} R(x,y) = H(A(f_*,y), B(g(x))) &\geq H(A(f_*,y) - N_1 B(g(x)), B(g(x))) > \\ &> H(A(f_*,y),0) = R(x_*,y). \end{aligned}$$

Thus, in any case $R(x,y) > R(x_*,y)$ if $x \notin X_*$, and, therefore, (3.1) is an e.a.f. on the set $P \times Y$. The theorem is proved. □

Let us consider two important special cases of the function (3.1). First assume $A(f,y)$ has the form

$$A(f,y) = f - y. \tag{3.6}$$

For this function, condition (3.2) is satisfied for any $Y \subseteq \mathbb{E}^1$ and $D = 1$. Therefore, by Theorem 4, if there exists a set $T \subseteq \mathbb{E}^1$, such that (3.3) is satisfied, then the function (3.1) with $A(f,y)$ representable in the form (3.6) is an e.a.f. for problem (1.1) on the set $P \times Y$, where $Y = f_* - T$.

Consider another function;

$$A(f,y) = y^{-1} f. \tag{3.7}$$

In this case, if (3.3) is satisfied for some $0 < D < +\infty$, the function (3.1) with $A(f,y)$ of the form (3.7) is also an e.a.f. on $P \times Y$, but here $Y = \{y \geq D^{-1} : y^{-1} f_* \in T\}$.

Note that if the function $H(t,\tau)$ is linear and has the form $H(t,\tau) = ct + \tau$, then condition (3.3) is satisfied if and only if $c < 1/B^0(w_*|\mathbb{E}^m)$. The set T in this case coincides with the entire real line $\mathbb{E}^1$. Therefore, if we change from y to $\bar{y} = y/c$, we obtain that the function $R(x,\bar{y}) = H(y^{-1}f(x), B(g(x))) = \bar{y}^{-1} f(x) + B(g(x))$ is an e.a.f. if and only if $\bar{y} > B^0(w_*|\mathbb{E}^m)$, which is fully consistent with the formula for Y_1 obtained above.

Denote by $K_1(t_*)$ the cone

$$K_1(t_*) = \{[t,\tau] \in \mathbb{E}^2 | \tau \geq (t_* - t)_+/N_1\}.$$

In geometrical terms, condition (3.3) implies that for any $t_* \in T$ the level line of the function $H(t,\tau)$ corresponding to the value $H(t_*,0)$ should not intersect the cone $K_1(t_*)$, with the exception of the point $[t_*, 0]$.

Condition (3.3) is relatively easily checked when $H(t,\tau)$ is a quasi-convex function. Indeed, assume that $H(t,\tau)$ is a continuously differentiable quasiconvex function on $\mathbb{E}^2$ and there exists a set $T_1 \subseteq \mathbb{E}^1$, such that

$$H_t(t_*,0) > 0, \qquad H_\tau(t_*,0)/H_t(t_*,0) > N_1 \qquad \forall t_* \in T_1. \tag{3.8}$$

Then the function of one variable $\varphi(t) = H(t, (t_* - t)_+/N_1)$ is also quasi-convex to the right and to the left of the point t_* and both one-sided derivatives exist at this point; by (3.8), $\varphi_t^+(t_*) = H_t(t_*,0) > 0$, $\varphi_t^-(t_*) = H_t(t_*,0) - H_\tau(t_*,0)/N_1 < 0$. Therefore, the function $\varphi(t)$ attains at the point t_* its strict minimum over t and inequality (3.3) holds. The set T_1, can be augmented with the points of the set

$$T_2 = \{t_* \notin T_1 | t_* = \lim_{t_k \to +t} t_k, \ \ t_k \in T_1\}.$$

If (3.8) holds, then (3.3) holds for any $t_* \in T = T_1 \cup T_2$. Also note that conditions (3.8) are preserved when the function $H(t,\tau)$ is differentiable only at the points of some set that contains the interval $\{z \in \mathbb{E}^2 \mid z^{(1)} \in T_1,\ \ z^{(2)} = 0\}$.

Let us give examples of e.a.f.'s for which the set Y can be found from Theorem 4 and condition (3.8). Let

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

This function is quasiconvex on $\mathbb{E}^2$. For $t < 0$ we have $H_t(t,0) = 1$, $H_\tau(t,0) = t_-^2$. Conditions (3.8) are satisfied if $t < -[B^0(w_*|\mathbb{E}^m)]^{1/2}$. Thus $T = (-\infty, -[B^0(w_*|E^m)]^{1/2})$ and the auxiliary function

$$R_8(x,y) = [f(x) - y]_-/\{1 - [f(x) - y]_- B(g(x))\}$$

is thus an e.a.f. on the set $\mathbb{E}^n \times Y_8$, where $Y_8 = \{y \in \mathbb{E}^1 | y > f_* + [B^0(w_*|\mathbb{E}^m)]^{1/2}\}$.

If $f_* > 0$, then using the function $H(t,\tau) = \tau + (\tau^2 + 4t_+^3)^{1/2}$, we obtain the e.a.f.

$$R_9(x,y) = B(g(x)) + \{[B(g(x))]^2 + 4y^{-3}[f_+(x)]^3\}^{1/2}$$

for problem (1.1) on the set $\mathbb{E}^n \times Y_9$, where

$$Y_9 = \{y : y > \max\{1.9 f_* [B^0(w_*|\mathbb{E}^m)]^2\}\}.$$

Consider the sufficient conditions for the case when $B(g(x))$ in (3.1) is an interior penalty function, assuming as before that $H(t,\tau)$ is nondecreasing on $\mathbb{E}^2$, and $A(f,y)$ is a convex monotone increasing function of the first argument. Instead of (3.2) we assume for $A(f,y)$ the existence of a constant C such that

$$0 < C \le \inf_{y \in Y} \inf_{\xi \in \partial A_f(f_*,y)} \xi < +\infty. \tag{3.9}$$

Put $N_2 = CB^0(w_*|\mathbb{E}_-^m)$.

Theorem 5. *Assume that the Lagrange function of problem* (1.1) *has a saddle point* $[x_*, w_*]$ *and* $A(f,y)$ *is a convex monotone increasing function of the first argument that satisfies* (3.9). *Also assume that* $B(g(x))$ *is an interior penalty function,* $0 < N_2 < +\infty$. *If there exists a set* $T \subseteq \mathbb{E}^1$*, such that* $A(f_*,y) \in T$ *for any* $y \in Y$ *and*

$$H(t, (t_* - t)/N_2) > H(t_*, 0) \qquad \forall t_* \in T, \quad \forall t > t_*, \tag{3.10}$$

then (3.1) *is an e.a.f. for problem* (1.1) *on the set* $X \times Y$.

Proof. Take an arbitrary point $x \in X$. If $x \notin X_*$, using inequalities (2.7), (3.9) and (3.10) we obtain

$$\begin{aligned} R(x,y) &= H(A(f(x),y), B(g(x))) \ge \\ &\ge H(A(f(x),y), [f_* - f(x)]/B^0(w_*|\mathbb{E}_-^m)) \ge \\ &\ge H(A(f(x),y), [A(f_*,y) - A(f(x),y)]/N_2) > \\ &> H(A(f_*,y),0) = R(x_*,y). \end{aligned}$$

If $x \in X_*$, then from (2.11) $B(g(x)) = 0$. Therefore, $R(x,y) = H(A(f(x),y),0) = R(x_*,y)$. Thus $R(x,y)$ is an e.a.f. on $X \times Y$. The theorem is proved. □

If $A(f,y)$ is representable in the form (3.6) or (3.7), then given T in condition (3.10), we can specify the set Y. Thus, for the first function we obtain that $Y = f_* - T$, and for the second $Y = \{y \le C^{-1} : y^{-1} f_* \in T\}$.

In geometrical terms, condition (3.10) indicates that for any $t_* \in T$ the level line of the function $H(t,\tau)$ corresponding to the value $H(t_*,0)$ does not have, for $t > t_*$, common points with the cone

$$K_2(t_*) = \{[t,\tau] \in \mathbb{E}^2 | \tau \geq (t_* - t)/N_2,\ \ t \geq t_*\}.$$

Conditions similar to (3.8) simplify the checking of the inequality (3.10) and relatively easily establish the form of the set T. Assume that $H(t,\tau)$ is a continuously differentiable quasiconvex function on $\mathbb{E}^2$ and there exists a set $T_1 \subseteq \mathbb{E}^1$ such that for any $t_* \in T_1$ we have the first inequality in (3.8) and

$$H_\tau(t_*,0)/H_t(t_*,0) < N_2.$$

Then for all $t_* \in T = T_1 \cup T_2$ inequality (3.10) is satisfied, where

$$T_2 = \{t_* \notin T_1 | t_* = \lim_{t_k \to -t} t_k,\ \ t_k \in T_1\}.$$

As an example of the application of Theorem 5, consider the previous functions $R_8(x,y)$ and $R_9(x,y)$, in which $B(g(x))$ is an interior penalty function. Both these functions are e.a.f.'s on the sets $X \times Y_8$, and $X \times Y_9$, respectively, but the sets Y_8 and Y_9 in this case have the form

$$\begin{aligned} Y_8 &= \{y \in \mathbb{E}^1 | f_* \leq y < f_* + [B^0(w_* | \mathbb{E}^m_-)]^{1/2}\}, \\ Y_9 &= \{y \in \mathbb{E}^1 | 0 < y < \min[1.9 f_* [B^0(w_* | \mathbb{E}^m_-)]^2]\}. \end{aligned}$$

Note that Theorems 4 and 5 do not change if conditions (3.3) and (3.10) hold not for all $t \in \mathbb{E}^1$, but only for those from some subset of $\mathbb{E}^1$ that contains the image of the set $P \times Y$ under the mapping $A(f(x),y)$. This property was used in [8] to show that the function

$$R_{10}(x,y) = (f(x) - y)_+ / \{1 + [f(x) - y]_+ B(g(x))\},$$

where $B(g(x))$ is an interior penalty function, is also an e.a.f. for problem (1.1) on the set $X \times Y_{10}$. Here

$$Y_{10} = \{y : f_* - ([(f^* - f_*)^2 + 4B^0(w_* | \mathbb{E}^m_-)]^{1/2} - (f^* - f_*))/2 < y \leq f_*\}, \quad f^* = \sup_{x \in X} f(x).$$

Consider e.a.f.'s of the form (3.1) with strictly mixed penalty functions $B(g(x))$. The conditions on $H(t,\tau)$ in this case are obtained by combining the corresponding conditions for strictly exterior and strictly interior penalty functions.

Theorem 6. *Assume that the Lagrange function of problem* (1.1) *has a saddle point* $[x_*, w_*]$ *and* $A(f,y)$ *is a convex monotone increasing function of the first argument that satisfies conditions* (3.2) *and* (3.9). *Also assume that* $B(g(x))$ *is a strictly mixed penalty function and* $0 < N_1 < +\infty$, $0 < N_2 < +\infty$. *If there exists a set* $T \subseteq \mathbb{E}^1$ *such that* $A(f_*,y) \in T$ *for any* $y \in Y$ *and*

$$H(t_*,0) < \begin{cases} H(t,(t_* - t)/N_1), & t < t_*, \\ H(t,(t_* - t)/N_2), & t > t_*, \end{cases} \qquad \forall t_* \in T,$$

then (3.1) *is an e.a.f. for problem* (1.1) *on the set* $P \times Y$.

The proof is omitted, because it is easily obtained by combining the corresponding arguments in the proof of Theorems 4 and 5.

The function $R_8(x,y)$ with a strictly mixed penalty function $B(g(x))$ remains an e.a.f. for problem (1.1) on the set $P \times Y$ if $B(g)$ is such that $B^0(w_* | \mathbb{E}^m) < B^0(w_* | \mathbb{E}^m_-)$. The set Y_8 in this case is

$$Y_8 = \{y : f_* + [B^0(w_* | \mathbb{E}^m)]^{1/2} < y < f_* + [B^0(w_* | \mathbb{E}^m_-)]^{1/2}\}.$$

4. Exact modified Lagrange functions

We will rewrite the right-hand side of inequality (1.11) in the form

$$L(x_*, w_*) \leq L(x, w) + \langle g(x), w_* - w \rangle.$$

If $B^0(w_* - w|g(P)) > 0$, then applying the Minkowski–Mahler inequality (1.7) we obtain

$$B(g(x)) \geq [L(x_*, w_*) - L(x, w)]/B^0(w_* - w|g(P)). \tag{4.1}$$

Inequality (4.1) makes it possible to construct a whole class of e.a.f.'s based on the Lagrange function $L(x, w)$. Let

$$R(x, y) = A(L(x, w), v) + B(g(x)). \tag{4.2}$$

Here $y = [w, v] \in Y \subseteq \mathbb{E}^m_+ \times \mathbb{E}^1$.

Put

$$W_Y = \{w \in \mathbb{E}^m_+ | \exists v \in \mathbb{E}^1,\ \ [w, v] \in Y\}, \quad V_w = \{v \in \mathbb{E}^1 | [w, v] \in Y\}.$$

The set W_Y is the projection of Y on $\mathbb{E}^m_+$, and the set V_w is the section of Y for a fixed $w \in W_Y$.

Assume that the functions A, B and the sets P, Y are such that the following conditions hold.

Condition C.

$$A(L, v) - A(L_*, v) \geq (L - L_*)/B^0(w_* - w|g(P)) \quad \forall L \in \mathbb{E}^1,\ \forall w \in W_Y,\ \forall v \in V_w,$$

where $L_* = L(x_*, w_*) = f_*$.

Condition D. For each point $y \in Y$ the solution set of the system

$$A(L(x, w), v) + B(g(x)) = A(L(x_*, w_*), v), \quad x \in P,$$

is identical with X_*.

We augment these two conditions with the following:

$$L(x_*, w) = L(x_*, w_*) \qquad \forall w \in W_Y. \tag{4.3}$$

Note that if $w_* \neq 0$, condition (4.3) is satisfied for the set

$$W_Y = W_* = \{0 \leq w \leq w_* | w^i < w^i_*,\ \text{if } w^i_* > 0\} \tag{4.4}$$

Theorem 7. *Assume that the Lagrange function of problem* (1.1) *has a saddle point* $[x_*, w_*]$ *and that* **conditions C**, **D**, *and* (4.3) *are satisfied. Also assume that* $B(g(x))$ *is a strictly exterior or interior penalty function (in the latter case,* $P \subseteq X$*) and* $0 < B^0(w_* - w|g(P)) < +\infty$ *for any* $w \in W_Y$*. Then the function* (4.2) *is an e.a.f. for problem* (1.1) *on the set* $P \times Y$*.*

The proof of this theorem is almost a verbatim repetition of the proof of Theorems 1 and 2. Also note that for an exterior penalty function $B(g(x))$ **condition C** can be relaxed: we only need the inequality

$$A(L, v) - A(L_*, v) \geq [(L - L_*)/B^0(w_* - w|g(P))]_- \quad \forall L \in \mathbb{E}^1,\ \forall w \in W_Y,\ \forall v \in V_w.$$

As examples of the function (4.2) which are e.a.f.'s on the set $\mathbb{E}^n \times Y$, where $Y = \{[w, v] | w \in \in W_*,\ \ v \in V_w\}$, we can take the functions R_1 – R_4, replacing $f(x)$ with $L(x, w)$. The sets V_w for these functions are constructed in the same way as the corresponding sets Y for the functions R_1 – R_4. We, therefore, give these e.a.f.'s with their sets V_w only for the case when $B(g(x))$ is a strictly exterior penalty function:

$$\begin{array}{rclrcl}
R_{11}(x,y) &=& v^{-1}L(x,w)+B(g(x)), & V_w &=& \{v : v > B^0(w_*-w|\mathbb{E}^m)\},\\
R_{12}(x,y) &=& v^{-1}e^{L(x,w)}+B(g(x)), & V_w &=& \{v : v > B^0(w_*-w|\mathbb{E}^m)e^{f_*}\},\\
R_{13}(x,y) &=& [v-L(x,w)]_+^{\alpha}+B(g(x)), & V_w &=& \{v : v > f_* + [-\alpha B^0(w_*-w|\mathbb{E}^m)]^{1/(1-\alpha)}\},\\
R_{14}(x,y) &=& [L(x,w)-v]_+^{\beta}+B(g(x)), & V_w &=& \{v : f_* \geq v > f_* - [\beta B^0(w_*-w|\mathbb{E}^m)]^{1/(1-\beta)}\}.
\end{array}$$

We see that as w approaches w_*, the corresponding sets V_w become larger. For e.a.f.'s with interior penalty functions they, conversely, contract. The function R_{11} was proposed in [3].

Now consider nonlinear e.a.f.'s constructed using a Lagrange function;

$$R(x,y) = H(A(L(x,w),v),B(g(x))), \tag{4.5}$$

where, as before, $H(t,\tau)$ is a continuous nondecreasing function of two variables, $A(L,v)$ is a monotone increasing convex function of the first argument. We also assume that the function A and the set Y are such that for any $w \in W_Y$ constants $C(w)$ and $D(w)$ exist for which

$$0 < \sup_{y\in Y} \sup_{\xi\in\partial A_L(L_*,v)} \xi \leq D(w) < +\infty, \tag{4.6}$$

$$0 < C(w) < \sup_{y\in Y} \sup_{\xi\in\partial A_L(L_*,v)} \xi < +\infty. \tag{4.7}$$

We also assume that

$$L(x,w) > L(x_*,w_*) = L_* \qquad \forall w \in W_Y, \quad \forall x \in X\backslash X_*, \tag{4.8}$$

and instead of (3.3) and (3.10) the function $H(t,\tau)$ satisfies the following conditions: for each $w \in W_Y$ there exists a set $T(w) \subseteq \mathbb{E}^1$ such that $A(L_*,v) \in T(w)$ for all $v \in V_w$ and for any $t_* \in T(w)$ we have

$$H(t,(t_*-t)_+/D(w)B^0(w_*-w|\mathbb{E}^m)) > H(t_*,0) \qquad \forall t \neq t_*, \tag{4.9}$$

$$H(t,(t_*-t)/C(w)B^0(w_*-w|\mathbb{E}^m_-)) > H(t_*,0) \qquad \forall t > t_*. \tag{4.10}$$

Note that if the set Y is such that the corresponding set W_Y is of the form (4.4), then (4.3) implies (4.8). Indeed, if $x \in X$ and $f(x) > f_*$, then for the case when $\langle g(x),w\rangle = 0$, we obtain $L(x,w) = f(x) > f_* = L_*$. If $\langle g(x),w\rangle < 0$, then $\langle g(x),w\rangle > \langle g(x),w_*\rangle$ and, therefore, $L(x,w) > L(x,w_*) \geq L_*$.

Theorem 8. *Assume that the Lagrange function in problem* (1.1) *has a saddle point* $[x_*,w_*]$, $A(L,v)$ *is a convex monotone increasing function of the first argument, and* (4.3) *and* (4.4) *hold. If* $B(g(x))$ *is a strictly exterior penalty function,* $0 < B^0(w_*-w|\mathbb{E}^m) < +\infty$ *for any* $w \in W_Y$ *and* (4.6) *and* (4.9) *hold, then* (4.5) *is an e.a.f. for problem* (1.1) *on the set* $P \times Y$. *If* $B(g(x))$ *is an interior penalty function,* $0 < B^0(w_*-w|\mathbb{E}^m_-) < +\infty$ *for any* $w \in W_Y$ *and* (4.7) *and* (4.10) *hold, then* (4.5) *is an e.a.f. for problem* (1.1) *on the set* $X \times Y$.

The proof is similar to the proof of Theorems 4 and 5.

Assume that the function $A(L,v)$ has the form $A(L,v) = L - v$. Then Theorem 8 implies that $V_w = f_* - T(w)$ for any $w \in W_Y$. This property leads to the following e.a.f.:

$$R_{15}(x,y) = [L(x,w)-v]_-/\{1-[L(x,w)-v]_-B(g(x))\}.$$

If $B(g(x))$ is a strictly exterior penalty function, then R_{15} is an e.a.f. on the set $P \times Y_{15}$, where

$$Y_{15} = \{[w,v] : w \in W_*, \ \ v > f_* + [B^0(w_*-w|\mathbb{E}^m)]^{1/2}\}. \tag{4.11}$$

If $B(g(x))$ is an interior penalty function, R_{15} is an e.a.f. on the set $X \times Y_{15}$, but the set (4.11) has the form

$$Y_{15} = \{[w, v] : w \in W_*, \;\; f_* \leq v < f_* + [B^0(w_* - w|\mathbb{E}^m_-)]^{1/2}\}.$$

We can also consider e.a.f.'s of the form (4.5) with strictly mixed penalty functions $B(g(x))$. The sufficient conditions remain as before, with the exception of condition (4.9), which is now only required to hold for $t < t_*$.

In conclusion note that the functions (2.1) and (3.1) are special cases of the functions (4.2) and (4.5). In order to obtain these functions, it suffices to set $w = 0$ in (4.2) and (4.5).

Also note that if problem (1.1) has more than one saddle point $[x_*, w_*]$, then the sets Y should be bounded for augmentation using upper or lower bounds over all possible values of the Lagrange multipliers w_*.

Bibliography

1. EREMIN I.I., The "penalty" method in convex programming, Dokl. Akad. Nauk SSSR, 173, 4, 748–751, 1967.

2. ZANGWILL W., Nonlinear programming via penalty functions, Manag. Sci., 344–358, 1967.

3. SKARIN V.D., On one modification of the penalty function method in convex programming, in: Nonlinear Optimization and Applications in Planning, UNTs Akad. Nauk SSSR, Sverdlovsk, 51–62, 1973.

4. EVANS J.P., GOULD F.J. and TOLLE J.W., Exact penalty functions in nonlinear programming, Math. Progr., 4, 72–97, 1973.

5. GOLIKOV A.I., The method of exterior centres and nonsmooth penalty functions, in: Software Systems for Optimal Planning Problems, 7th All-Union Symposium, Moscow, 125–126, 1982.

6. YEVTUSHENKO YU.G., Accuracy bounds in penalty function methods, in: Problems of Applied Mathematics and Information, Theory, Nauka, Moscow, 199–208, 1987.

7. ZHADAN V.G., On some bounds of the penalty coefficient in exact penalty function methods, Zh. Vȳchisl. Matem. i Mat. Fiz., 24, 8, 1164–1171, 1984.

8. YEVTUSHENKO YU.G. and ZHADAN V.G., On systematization of numerical methods of nonlinear programming. Successive unconstrained minimization methods, in: Communications in Applied Mathematics, VTs AN SSSR, Moscow, 1988.

9. YEVTUSHENKO YU.G., Methods of Solving Extremal Problems and Their Application in Optimization Systems, Nauka, Moscow, 1982.

10. HAN S.-P. and MANGASARIAN O.L., Exact penalty functions in nonlinear programming, Math. Progr., 17, 251–269, 1979.

11. MORRISON D.D., Optimization by least squares, SIAM J. Numer. Anal., 5, 1, 83–88, 1968.