Mathematical Programming 54 (1992) 177-222
North-Holland

Modified barrier functions (theory and methods)

R. Polyak

Mathematical Sciences Department, T.J. Watson Research Center, IBM Research Division, Yorktown Heights, NY 10598, USA

Received 4 November 1987
Revised manuscript received 17 July 1990

The nonlinear rescaling principle employs monotone and sufficiently smooth functions to transform the constraints and/or the objective function into an equivalent problem, the classical Lagrangian which has important properties on the primal and the dual spaces.

The application of the nonlinear rescaling principle to constrained optimization problems leads to a class of modified barrier functions (MBF's) and MBF Methods (MBFM's). Being classical Lagrangians (CL's) for an equivalent problem, the MBF's combine the best properties of the CL's and classical barrier functions (CBF's) but at the same time are free of their most essential deficiencies.

Due to the excellent MBF properties, new characteristics of the dual pair convex programming problems have been found and the duality theory for nonconvex constrained optimization has been developed.

The MBFM have up to a superlinear rate of convergence and are to the classical barrier functions (CBF's) method as the Multipliers Method for Augmented Lagrangians is to the Classical Penalty Function Method. Based on the dual theory associated with MBF, the method for the simultaneous solution of the dual pair convex programming problems with up to quadratic rates of convergence have been developed. The application of the MBF to linear (LP) and quadratic (QP) programming leads to a new type of multipliers methods which have a much better rate of convergence under lower computational complexity at each step as compared to the CBF methods.

The numerical realization of the MBFM leads to the Newton Modified Barrier Method (NMBM). The excellent MBF properties allow us to discover that for any nondegenerate constrained optimization problem, there exists a "hot" start, from which the NMBM has a better rate of convergence, a better complexity bound, and is more stable than the interior point methods, which are based on the classical barrier functions.

Key words: Nonlinear rescaling, modified barrier functions, multipliers method, simultaneous solution, dual problems.

Introduction

In the middle of the 1950's Frisch [8] and at the outset of the 1960's Carroll [3] recommended the classical barrier functions (CBF's) for solving constrained optimization problems. Later these functions were extensively studied by Fiacco and McCormick in [6] (see also [13]) and incorporated in different general solution techniques, so the corresponding methods made up a considerable part of modern optimization theory (see [6, 13, 18, 19]). Interest in these functions and the corresponding methods grew dramatically in connection with the well known progress in Linear Programming (see [5, 7, 10, 12, 15, 16, 17, 28, 29, 34] and bibliography in it).

At the same time the CBF's as well as the methods based on these functions still have their inherent drawbacks. A specific feature of the barrier functions is their unbounded increase in a neighborhood of the boundary. This enables us to start the solution process at any interior point of the feasible set and to remain in the interior without taking particular care of the constraints. It makes it possible to use the smooth optimization methods (see [4, 9, 19]) for solving constrained (nonsmooth) optimization problems. However, this merit of the CBF's becomes a deficiency when the computational process approaches the active constraints boundary.

The CBF's as well as their derivatives do not exist at the solution. The CBF's grow to infinity, the condition number of the Hessian vanishes and the repulsive effect from the active constraints boundary becomes stronger as the computational process approaches the solution. So, while the computations are increasing from step to step, the rate of convergence is rather slow, even when the second order optimality conditions are fulfilled. Furthermore, the CBF's methods obtain the optimal values of the Lagrange multipliers only as a result of a limiting process as the penalty parameter tends to infinity.

On the other hand, the classical Lagrangians, which are fundamental in constrained optimization both for the theoretical analysis (necessary and sufficient condition, duality theory) and computational methods, along with very important qualities have some essential deficiencies.

First of all, generally, the unconstrained optimum of the CL in the primal space under the fixed optimal Lagrange multipliers might not exist even if the second order optimality conditions are fulfilled. The unconstrained optimization CL, which correspond to the Linear Programming problem, under the fixed optimal dual variables is not equivalent to the initial LP problem.

The objective function of the dual problem, which is based on the CL, is in general nonsmooth, independent of the smoothness of the initial functions, even for the convex programming problem when the second order optimality sufficient conditions are fulfilled.

The purpose of this paper is to develop the MBF theory and, based on this theory, to consider MBF methods for solving constrained optimization problems. As will be proven later, the MBF combine the best properties of the CL and CBF, but at the same time are free from their most essential drawbacks and might be considered as interior augmented Lagrangians.

In contrast to the CBF's, the MBF's are defined at the solution. Moreover, these functions keep the smoothness of the order of the initial functions in a neighborhood of the feasible set. They do not grow infinitely and the condition of the Hessian does not vanish when the current approximation approaches the solution.

The most important quality of the MBF is the explicit representation of the Lagrange multipliers. It allows us not only to attach to the MBF, which is in fact a classical Lagrangian, all of the best properties of the augmented Lagrangians (see [2, 11, 14, 20, 27, 31]) but also to find some new important qualities.

In contrast to the CL's, the MBF's is strongly convex in the neighborhood of the solution even in the case of nonconvex programming problems, if the second order optimality conditions are fulfilled. Under the optimal Lagrange multipliers, the unconstrained extremum of the MBF's exists and coincides with the solution of the initial problem. The dual functions, which are based on the MBF's, are as smooth as the initial functions of the primal problem and, the dual problem, which is always convex whether the initial problem is convex or not, has important local (near the solution) properties.

Based on the MBF theory, three versions of MBFM have been developed. The MBFM's have a much better rate of convergence under lower computational complexity at each step compared to the Classical Interior Point Methods (CIPM's) (see [6]), which are based on CBF's. Even under the fixed penalty parameter, the sequence generated by MBFM's converge to the primal and dual solutions linearly. If one increases the penalty parameter from step to step, the MBFM sequence converges to the solution superlinearly, while CIPM have only an arithmetical rate of convergence. In fact, the MBFM is to the CIPM as the multipliers method of the augmented Lagrangians (see [2, 11]) is to the Classical Penalty Functions Method (see [6, 13]).

Moreover, a consideration of the dual problem associated with the MBF's leads to a general method for simultaneous solution of the dual pair of the convex programming problems with up to a quadratic rate of convergence.

The numerical realization of the MBFM leads to the Newton Modified Barrier Method (NMBM). The analysis of MBF's allowed us to discover that for any nondegenerate constrained optimization problem, there exists a "hot" start, from which the NMBM trajectory is much more "powerful" than the Interior Point Methods (IPM's) trajectory. This means that following along the NMBM trajectory, one can obtain the same improvement of the current approximation by using essentially less Newton Method steps. This makes it possible to combine the universal self-concordant properties (see [17]) of the CBF's, which guarantee the polynomial complexity bound of the IPM's, beginning at the "warm" start, with excellent MBF's properties, which guarantee the essential improvement of this bound, beginning at the "hot" start.

Finally, note that in application to a nondegenerate LP, the normal system of equations, which one has to solve at every step of the NMBM, is numerically more stable than the corresponding systems for the IPM which are based on the CBF.

The main results for the nonlinear programming problems were obtained in 1981-1982 as a part of our investigation, which had been undertaken then, concerning the nonlinear rescaling (monotone transformation) principle in external and equilibrium problems with constraints (see [21-24]).

The LP and QP parts were done in 1986. Some results contained in this paper were presented at the 11th and 12th International Mathematical Programming Symposiums (Bonn, 1982, Boston, 1985) (see also [25-26]).

1. Problem formulation and basic assumptions

Let $f_0(x)$ and $f_1(x)$, $i=1,\dots,m$, be C^2-functions in $\mathbb{R}^n$ and let there exist

$$x^* = \operatorname{argmin}\{f_0(x)\,|\,x\in\Omega\}, \tag{1}$$

where $\Omega=\{x: f_i(x)\geqslant 0,\ i=1,\dots,m\}$. If $f_0(x)$ and $-f_1(x)$ are convex and the Slater condition holds, i.e.

$$\exists x_0:\quad f_i(x_0)>0,\quad i=1,\dots,m; \tag{2}$$

then Karush-Kuhn-Tucker's (K-K-T's) theorem holds true, i.e., there exists a vector $u^*=(u_1^*,\dots,u_m^*)\geqslant 0$ such that

$$L'_x(x^*,u^*)=f'_0(x^*)-\sum_{i=1}^m u_i^* f'_i(x^*)=0,\qquad f_i(x^*)u_i^*=0,\quad i=1,\dots,m. \tag{3}$$

Let $I^*=\{i: f_i(x^*)=0\}=\{1,\dots,r\}$ be the active constraint set. In view of (2) the multiplier polyhedron

$$Q=\left\{u=(u_1,\dots,u_r)\geqslant 0: f'_0(x^*)-\sum_{i=1}^r uf'_i(x^*)=0\right\}$$

is nonempty for a convex programming problem and every vertex of this polyhedron is in a one-to-one correspondence with a minimal set of the active constraints, i.e., with an index set $I\subset I^*$ such that

$$\min\left\{\left\|f'_0(x^*)-\sum_{i\in I} u_i f'_i(x^*)\right\|\ \middle|\ u_i\geqslant 0,\ i\in I\right\}=0$$

and

$$\min\left\{\left\|f'_0(x^*)-\sum_{i\in I\setminus j} u_i f'_i(x^*)\right\|\ \middle|\ u_i\geqslant 0,\ i\in I\setminus j\right\}>0\quad \forall j\in I.$$

For convenience, denote $f(x)=(f_i(x),\ i=1,\dots,m)$, $f_{(r)}(x)=(f_i(x),\ i=1,\dots,r)$, and $f'(x)=J(f(x))$, $f'_{(r)}(x)=J(f_{(r)}(x))$ the Jacobi matrix of the vector-functions $f(x)$, $f_{(r)}(x)$ respectively.

If the sufficient regularity conditions are satisfied (for example see [6, p. 30]),

$$\operatorname{rank} f'_{(r)}(x^*)=r,\qquad u_i^*>0,\quad i\in I^*, \tag{4}$$

then the multiplier polyhedron shrinks to a point. Condition (4) together with the sufficient condition for the minimum x^* to be isolated,

$$(L''_{xx}(x^*,u^*)y,y)\geqslant\lambda(y,y),\quad \lambda>0,\quad \forall y\neq 0: f'_{(r)}(x^*)y=0, \tag{5}$$

comprises the standard second-order optimality sufficient conditions for the constrained optimization problem (1).

(Since for any minimal set I conditions (4) are satisfied, it follows that results similar to those established below are valid for convex programming problems, whenever (4) and (5) are replaced by the Slater condition and (5) is satisfied for $L_I(x, u) = f_0(x) - \sum_{i \in I} u_i f_i(x)$, i.e.,

$$(L''_{Ixx}(x^*, u^*)y, y) \geq \lambda(y, y), \quad \lambda > 0, \quad \forall y \neq 0: f'_i(x^*)y = 0, \; i \in I, \tag{5'}$$

hold.)

We shall use the following assertion which is a modification of the Debreu theorem (see [1]) and can be proved in a similar manner.

Assertion 1. *Let A be a symmetric $n \times n$ matrix, let B be an $r \times n$ matrix and $U = \operatorname{diag} u_i : \mathbb{R}^r \to \mathbb{R}^r$, such that $u = (u_1, \ldots, u_r) > 0$ and $By = 0 \Rightarrow (Ay, y) \geq \lambda(y, y), \lambda > 0$. Then there exists a $k_0 > 0$ such that for any $0 < \mu < \lambda$ we have*

$$((A + kB^{\mathrm{T}}UB)x, x) \geq \mu(x, x) \quad \forall x \in \mathbb{R}^n$$

whenever $k \geq k_0$. □

2. Modified barrier functions

The functions $\varphi(x, k) = f_0(x) - k^{-1} \sum_{i=1}^m \ln f_i(x)$ and $c(x, k) = f_0(x) + k^{-1} \sum_{i=1}^m f_i^{-1}(x)$ introduced by Frisch [8] and Carroll [3] are the best-known barrier functions. However, both of these functions have a serious disadvantage because they, as well as their derivatives, do not exist at x^* and the functions grow to infinity when $x \to x^*$.

Let $k > 0$ and the set $\Omega_k = \{x: kf_i(x) + 1 \geq 0, \; i = 1, \ldots, m\}$. Notice that $\Omega \subset \Omega_k$. It is clear that if $f_i(x)$, $i = 1, \ldots, m$ are concave, the compactness of Ω implies the compactness of Ω_k for any $k > 0$ [6, p. 93]. If (1) is a nonconvex programming problem, then the compactness of Ω does not imply the compactness of Ω_k. So in the nonconvex case we will use the following growth condition

$$\exists k_0 > 0 \text{ and } \tau > 0: \quad \max\left\{ \max_{1 \leq i \leq m} f_i(x) \;\middle|\; x \in \Omega_{k_0} \right\} = \theta(k_0) \leq \tau. \tag{6}$$

It is clear that $\theta(k)$ is a monotone decreasing function on $k > 0$. So if (6) is fulfilled for some $k_0 > 0$ the inequality $\theta(k) \leq \tau$ will be fulfilled for any $k \geq k_0$.

We define the Modified Frisch Function $F(x, u, k): \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^1_+ \to \mathbb{R}^1$ by the formula

$$F(x, u, k) = \begin{cases} f_0(x) - k^{-1} \sum_{i=1}^m u_i \ln(kf_i(x) + 1), & \text{if } x \in \operatorname{int} \Omega_k, \\ \infty & \text{if } x \notin \operatorname{int} \Omega_k, \end{cases}$$

and the Modified Carroll Function $C(x, u, k)$ by the formula

$$C(x,u,k)=\begin{cases} f_0(x)+k^{-1}\sum_{i=1}^{m} u_i[(kf_i(x)+1)^{-1}-1], & \text{if } x\in \operatorname{int}\Omega_k,\\ \infty & \text{if } x\notin \operatorname{int}\Omega_k.\end{cases}$$

For every $k>0$, $\Omega=\Omega_F=\{x\colon k^{-1}\ln(kf_i(x)+1)\geq 0,\ i=1,\ldots,m\}=\Omega_C=\{x\colon k^{-1}[kf_i(x)+1)^{-1}-1]\leq 0,\ i=1,\ldots,m\}$, therefore problem (1) is equivalent to the problem

$$x^*=\operatorname{argmin}\{f_0(x)\,|\,x\in\Omega_F\}=\operatorname{argmin}\{f_0(x)\,|\,x\in\Omega_C\}, \tag{7}$$

while $F(x, u, k)$ and $C(x, u, k)$ are classical Lagrangians for the problem (7).

It is easy to see that for every $u\geq 0$ and $k>0$, the functions $F(x, u, k)$ and $C(x, u, k)$ are convex in x provided $f_0(x)$ is convex and $f_i(x)$, $i=1,\ldots,m$, are concave. The critical properties of these MBF's are that:

(P1) $F(x^*, u^*, k)=C(x^*, u^*, k)=f_0(x^*)$ for any $k>0$.

Due to K-K-T's condition (3), for any $k>0$ we have:

(P2) $F'_x(x^*, u^*, k)=C'_x(x^*, u^*, k)=f'_0(x^*)-\sum_{i=1}^{m} u_i^* f'_i(x^*)=0.$

Therefore for any $k>0$ the functions $F(x, u^*, k)$ and $C(x, u^*, k)$ attain their minimum at x^* if (1) is a convex programming problem and thus the knowledge of the Lagrange multipliers $u^*=(u_1^*,\ldots,u_m^*)$ allow us to solve the problem (1) by solving one smooth optimization problem.

(P3) $x^*=\operatorname{argmin}\{F(x, u^*, k)\,|\,x\in\mathbb{R}^n\}=\operatorname{argmin}\{C(x, u^*, k)\,|\,x\in\mathbb{R}^n\}.$

To extend this idea to the nonconvex programming problem we can proceed as follows.

Let $U^*=[\operatorname{diag} u_i^*]_{i=1}^r$, then

(P4) $F''_{xx}(x^*, u^*, k)=C''_{xx}(x^*, u^*, k)=L''_{xx}(x^*, u^*)+kf'^{\mathrm{T}}_{(r)}(x^*)U^*f'_{(r)}(x^*).$

If (5) is fulfilled and $u_i^*>0$, $i=1,\ldots,r$, then for $A=L''_{xx}(x^*, u^*)$, $B=f'_{(r)}(x^*)$ and $U=U^*$ it follows from Assertion 1 that there exists $k_0>0$ and $\lambda>\mu>0$ such that

(P5) $(F''_{xx}(x^*, u^*, k)y, y)\geq\mu(y,y)\quad\forall y\in\mathbb{R}^n,\ \forall k\geq k_0,$

i.e., $F(x, u^*, k)$ and $C(x, u^*, k)$ are strongly convex in $\mathbb{R}^n$ in the neighborhood of x^* for any $k\geq k_0$.

Note that for the CL the property (P5) is not fulfilled even if (1) is a convex programming problem and the second order optimality sufficient conditions are fulfilled in the strict form. On the other hand, the property (P5), hence (P3), holds for the MBF even if the problem (1) is non-convex, whenever (4)–(5) is fulfilled and $k\geq k_0$. For the CL, (P3) is generally false even if (4)–(5) are fulfilled.

So, the Lagrange multipliers, the specific role of the penalty parameter in the construction of the MBF, together with the extension of the feasible set, which is defined by this parameter, give rise to the properties (P1)-(P5) and allow us to establish some new basic facts concerning MBF.

3. Basic theorem

This theorem states the main facts concerning the MBF. For $\varepsilon > 0$ set $U(\varepsilon) = \{u \in \mathbb{R}^m_+ : u_i \geq \varepsilon,\ i = 1, \dots, r,\ u_i \geq 0,\ i = r+1, \dots, n\}$. Suppose ε and $k_0 > 0$ such that for a given vector $u \in U(\varepsilon)$ and parameter $k \geq k_0$ there exists a vector

$$\hat{x} = \hat{x}(u, k) = \operatorname{argmin}\{F(x, u, k) \mid x \in \mathbb{R}^n\}.$$

Together with $\hat{x}$ we consider the vector

$$\hat{u} = \hat{u}(u, k) = [\operatorname{diag}(kf_i(\hat{x}) + 1)^{-1}]_{i=1}^m u.$$

We will say that the vector $u \in U(\varepsilon)$ is *well defined* for the parameter $k \geq k_0$ if $\hat{x}(u, k)$ exists and the estimation

$$\max\{\|\hat{x}(u, k) - x^*\|, \|\hat{u}(u, k) - u^*\|\} \leq ck^{-1}\|u - u^*\| = \gamma_k \|u - u^*\| \tag{8}$$

holds true, c is independent of $k \geq k_0$ and $\gamma_k \leq \frac{1}{2}$. It will be proven later that if $u \in U(\varepsilon)$ is well defined for the parameter $\bar{k} \geq k_0$ then u is well defined for any $k \geq \bar{k}$. For a fixed $k \geq k_0$ consider the set $U_k = \{u \in U(\varepsilon) : u$ is well defined for the parameter $k\} \neq \emptyset$ and define an operator $C_k : U_k \to U_k$ by the formula

$$C_k u = \hat{u}(u, k) = \hat{u}.$$

Then $C_k u^* = u^*$, i.e., u^* is a fixed point of the mapping $u \to \hat{u}(u, k)$.

For a given $k \geq k_0$ also define a transformation $T_k : U_k \to \mathbb{R}^n \times U_k$ by the formula

$$T_k u = (\hat{x}(u, k), \hat{u}(u, k)) = (\hat{x}, \hat{u}).$$

Note that $T_k u^* = (\hat{x}(u^*, k), \hat{u}(u^*, k)) = (x^*, u^*)$, for any $k \geq k_0$. The main results to be established below are the existence of a threshold k_0, such that for every $k \geq k_0$ there exists a nonempty set U_k and a contractive operator C_k with contractibility contr $C_k = \gamma_k$, which tends to zero as $k \to \infty$, i.e.

$$\|C_k u - u^*\| = \|C_k u - C_k u^*\| \leq \gamma_k \|u - u^*\|, \tag{9}$$

holds for $\forall u \in U_k$, $0 < \gamma_k \leq \frac{1}{2}$, $k \geq k_0$ and $\gamma_k \to 0$ if $k \to \infty$. In the course of proving the theorem we will find the estimation for the threshold k_0, which is crucial to the properties of MBF's as well as for the complexity of the MBFM's.

This analysis highlights the most important parameters involved in the computational process which are responsible for the complexity of the constrained optimization problem.

Let $\delta > 0$ be small enough, $0 < \varepsilon < \min\{u_i^* \mid i = 1, \ldots, r\}$ and k_0 large enough (in the course of proof it will be clearer what "small" and "large" mean). Also define sets $D_i(\cdot) = D_i(u^*, k_0, \delta, \varepsilon) = \{u_i : u_i \geq \varepsilon, |u_i - u_i^*| \leq \delta k, k \geq k_0\}$, $i = 1, \ldots, r$, $D_i(u^*, k_0, \delta, \varepsilon) = \{u_i : 0 \leq u_i \leq \delta k, k \geq k_0\}$, $i = r+1, \ldots, m$. $D(u^*, k_0, \delta, \varepsilon) = D_1(\cdot) \otimes \cdots \otimes D_r(\cdot) \otimes \cdots \otimes D_m(\cdot)$ and for any fixed $k \geq k_0$ define sets $U_k^i = \{u_i : \max(\varepsilon, u_i^* - \delta k) \leq u_i \leq u_i^* + \delta k\}$, $i = 1, \ldots, r$, $U_k^i = \{u_i : 0 \leq u_i \leq \delta k\}$, $i = r+1, \ldots, m$, $U_k = U_k^1 \otimes \cdots \otimes U_k^r \otimes \cdots \otimes U_k^m$. So $D(\cdot) = \{u, k): u \in U_k, k \geq k_0\}$ (see Figure 1).

Further, let $\sigma = \min\{f_i(x^*) \mid r+1 \leq i \leq m\} > 0$, I^r is the $r \times r$ identity matrix, $O^{r,r}$ is the $r \times r$ zero matrix, $M > 0$ large enough, $\|x\| = \max_{1 \leq i \leq n} |x_i|$, $\|u\| \leq M$ and $S(y, \varepsilon) = \{x \in \mathbb{R}^n : \|x - y\| \leq \varepsilon\}$.

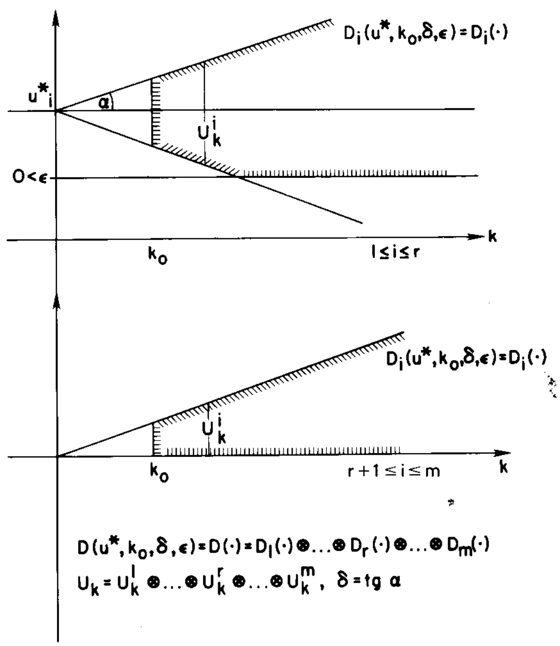

Fig. 1.

Theorem 1. (i) *Let $f_i(x) \in C^2$, $i = 0, \ldots, m$, and the conditions* (3)–(6) *hold. Then there exist $k_0 > 0$ and small enough $\delta > 0$ that for any $0 < \varepsilon < \min_{1 \leq i \leq r} u_i^*$ and any $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$ the following statements hold*:

(a) *There exists a vector*

$$\hat{x} = \hat{x}(u, k) = \operatorname{argmin}\{F(x, u, k) \mid x \in \mathbb{R}^n\}$$

such that $F'_x(\hat{x}, u, k) = 0$.

(b) *For the pair of vectors* $\hat{x}$ *and* $\hat{u} \equiv \hat{u}(u, k) = [\operatorname{diag}(kf_i(\hat{x}(u, k)) + 1)^{-1}]_{i=1}^{m} u$ *the estimate*

$$\max\{\|\hat{x} - x^*\|, \|\hat{u} - u^*\|\} \leq ck^{-1}\|u - u^*\| \tag{10}$$

holds, with constant c independent of k.

(c) $\hat{x}(u^*, k) = x^*$, $\hat{u}(u^*, k) = u^*$, *i.e.* u^* *is the fixed point of the mapping* $u \to \hat{u}(u, k)$.

(d) *The function* $F(x, u, k)$ *is strongly convex in a neighborhood of* $\hat{x} = \hat{x}(u, k)$.

(ii) *Let* $f_0(x)$ *and* $-f_i(x)$, $i = 1, \ldots, m$, *be convex and* $f_i(x) \in C^2$.

(a) *If* $\Omega^* = \{x \in \Omega : f_0(x) = f_0(x^*)\}$ *is a compact, then for any* $(u, k) \in \mathbb{R}_+^{m+1}$ *there exist* $\hat{x} = x(u, k)$ *such that* $F'_x(\hat{x}, u, k) = 0$.

(b) $\hat{x}(u^*, k) = x^*$, $\hat{u}(u^*, k) = u^*$ *for any* $k > 0$.

(c) *If conditions* (3)–(5) *are fulfilled, then for any* $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$ *the estimation* (10) *holds and* $F(x, u, k)$ *is strongly convex in a neighborhood of* $\hat{x}$.

Proof (i) (a) Let $t_i = (u_i - u_i^*)k^{-1}$, $i = 1, \ldots, m$, $t = (t_i, i = 1, \ldots, m)$, $S(0, \delta) = \{t = \{t_1, \ldots, t_m\} : |t_i| \leq \delta, i = 1, \ldots, m\}$, $\hat{u}_{(r)} = (\hat{u}_i, i = 1, \ldots, r)$, $\hat{u}_i(x, t, k) = kt_i(kf_i(x) + 1)^{-1}$, $i = r+1, \ldots, m$, $h(x, t, k) = \sum_{i=r+1}^{m} \hat{u}_i(x, t, k) f_i'^{\mathrm{T}}(x) = k \sum_{i=r+1}^{m} t_i(kf_i(x) + 1)^{-1} f_i'^{\mathrm{T}}(x)$. Then for any $k > 0$ and $x \in S(x^*, \varepsilon_0)$, $t \in S(O, \delta)$ the vector function $h(x, t, k)$ is smooth enough and $h(x^*, 0, k) = O \in \mathbb{R}^n$, $h'_k(x^*, 0, k) = O^{n,n}$, $h'_{\hat{u}_{(r)}}(x^*, 0, k) = O^{n,r}$. On the $S(x^*, \varepsilon_0) \times S(u^*_{(r)}, \varepsilon_0) \times S(0, \delta) \times (0, \infty)$ we consider the map $\Phi(x, \hat{u}_{(r)}, t, k) : \mathbb{R}^{n+r+m+1} \to \mathbb{R}^{n+r}$ defined by

$$\Phi(x, \hat{u}_{(r)}, t, k) = \Bigg(f_0'^{\mathrm{T}}(x) - \sum_{i=1}^{r} \hat{u}_i f_i'^{\mathrm{T}}(x) - h(x, t, k);$$

$$k^{-1}(kt_i + u_i^*)(kf_i(x) + 1)^{-1} - k^{-1}\hat{u}_i, i = 1, \ldots, r\Bigg).$$

Taking into account (3) and $h(x^*, 0, k) = 0$ we obtain $\Phi(x^*, u^*_{(r)}, 0, k) = 0$ for $\forall k > 0$. Let $\Phi'_{x\hat{u}_{(r)}} \equiv \Phi'_{x\hat{u}}(x^*, u^*_{(r)}, 0, k)$; $L''_{xx} = L''_{xx}(x^*, u^*)$, $f' \equiv f'(x^*)$, $f'_{(r)} = f'_{(r)}(x^*)$, $U^*_{(r)} = [\operatorname{diag} u_i^*]_{i=1}^{r} : \mathbb{R}^r \to \mathbb{R}^r$, $u_i^* > 0$, $i = 1, \ldots, r$.

In view of $h'_x(x^*, 0, k) = O^{n,n}$, $h'_{\hat{u}_{(r)}}(x^*, 0, k) = O^{n,r}$ we obtain

$$\Phi'_{(k)} = \Phi'_{x\hat{u}_{(r)}}(x^*, u^*_{(r)}, 0, k) = \begin{pmatrix} L''_{xx} & -(f'_{(r)})^{\mathrm{T}} \\ -U^*_{(r)} f'_{(r)} & -k^{-1} I^r \end{pmatrix}.$$

Along with $\Phi'_{(k)}$ we consider the matrix

$$\Phi'_{(\infty)} \equiv \Phi'_{x\hat{u}}(x^*, u^*, 0, \infty) = \begin{pmatrix} L''_{xx} & -(f'_{(r)})^{\mathrm{T}} \\ -U^*_{(r)} f'_{(r)} & O^{r,r} \end{pmatrix}.$$

The matrix $\Phi'_{(\infty)}$ is nonsingular, because for any vector $w = (y, v) \in \mathbb{R}^{n+r}$ the system $\Phi'_{(\infty)} w = 0$ implies $L''_{xx} y - (f'_{(r)})^{\mathrm{T}} v = 0$ and $U^*_{(r)} f'_{(r)} y = 0$. Since $u_i^* > 0$, $i = 1, \ldots, r$ the second set of equations implies $f'_{(r)} y = 0$. So multiplying the first set of equations by y we obtain $(L''_{xx} y, y) - (f'_{(r)} y, v) = 0$. Therefore $f'_{(r)} y = 0$ implies $(L''_{xx} y, y) = 0$. By virtue of (5) this is possible only if $y = 0$, but then $f'_{(r)} v = 0$ in view of (4) we obtain $v = 0$, i.e. it follows from $\Phi'_{(\infty)} w =$ [illegible] $= 0$. So the matrix $\Phi'_{(\infty)}$ is nonsingular.

Consequently, there exists a constant $\lambda_0>0$ (independent of $k>0$) such that $\|\Phi_{(\infty)}^{-1}\|\leq\lambda_0$.

Moreover for the Gram matrix $G_{(\infty)}=\Phi_{(\infty)}'^{\mathrm{T}}\Phi_{(\infty)}'$ there exists a scalar $\mu_0>0$ such that $(G_{(\infty)}w, w)\geq\mu_0(w, w)\ \forall w\in\mathbb{R}^{n+r}$. Therefore there exists a $k_0>0$ such that for every $k\geq k_0$ and for the matrix $G_k=\Phi_{(k)}'^{\mathrm{T}}\Phi_{(k)}'$ we have $(G_k w, w)\geq\frac{1}{2}\mu_0(w, w)\ \forall w\in\mathbb{R}^{n+r}$ and $\mu_0>0$ is independent of $k\geq k_0$. So the matrix $\Phi_{(k)}'$ is not only nonsingular, but there exists a constant $\rho>0$ which is independent of $k\geq k_0$ such that $\|\Phi_{(k)}^{-1}\|\leq\rho$. Let $k_1>k_0$ be any large enough number and $K=\{O\in\mathbb{R}^n\}\times[k_0, k_1]$. Since $\Phi(x^*, u_{(r)}^*, 0, k)=0$, $f_i(x)\in C^2$, $i=0,\ldots,m$, and the matrix $\Phi_{(k)}'$ is nonsingular for any $k\in[k_0, k_1]$ it follows from the second implicit function theorem (see [2, p. 12]) that there exist $\varepsilon_0>0$, $\delta>0$ and smooth vector-functions $x(\cdot)=x(t,k)=(x_1(t,k),\ldots,x_n(t,k))$, $\hat{u}_{(r)}(\cdot)=\hat{u}_{(r)}(t,k)=(\hat{u}_1(t,k),\ldots,\hat{u}_r(t,k))$ defined uniquely in a neighborhood $S(K,\delta)=\{(t,k)\colon |t_i|\leq\delta,\ i=1,\ldots,m,\ k\in[k_0,k_1]\}$ of the compact K such that $x(0,k)=x^*$, $\hat{u}_{(r)}(0,k)=u_{(r)}^*=(u_1^*,\ldots,u_r^*)$ for any $k\in[k_0,k_1]$.

(b) Now we are going to prove the estimate (10). There exist $\varepsilon_0>0$ such that $\max\{\|x(t,k)-x^*\|, \|\hat{u}_{(r)}(t,k)-u^*\|\}\leq\varepsilon_0$,

$$\Phi(x(t,k), \hat{u}_{(r)}(t,k), t, k)\equiv\Phi(x(\cdot), \hat{u}_r(\cdot), \cdot)\equiv 0 \tag{11}$$

and

$$\|(\Phi_{x\hat{u}_{(r)}}'(x(t,k), \hat{u}_{(r)}(t,k), t, k))^{-1}\|\leq 2\rho\quad\forall(t,k)\in S(K,\delta).$$

Rewriting (11) we obtain

$$f_0'^{\mathrm{T}}(x(t,k))-\sum_{i=1}^{r}\hat{u}_i(t,k)f_i'^{\mathrm{T}}(x(t,k))-h(x(t,k),t,k)=0, \tag{12}$$

$$\hat{u}_i(t,k)=(kt_i+u_i^*)(kf_i(x(t,k))+1)^{-1},\quad i=1,\ldots,r, \tag{13}$$

and let

$$\hat{u}_i(t,k)=kt_i(kf_i(x(t,k))+1)^{-1},\quad i=r+1,\ldots,m. \tag{14}$$

We recall that $u_{(m-r)}^*=(u_{r+1}^*,\ldots,u_m^*)=O\in\mathbb{R}^{m-r}$. First let us estimate the $\|\hat{u}_{(m-r)}(\cdot)\|$ where $\hat{u}_{(m-r)}(\cdot)=(\hat{u}_i(\cdot), i=r+1,\ldots,m)$. If $\delta>0$ small enough then for any $(t,k)\in S(K,\delta)$ we have $\|x(t,k)-x(0,k)\|=\|x(\cdot)-x^*\|\leq\varepsilon$ and $f_i(x(t,k))\geq\frac{1}{2}\sigma$ therefore

$$\hat{u}_i(\cdot)=\frac{u_i-u_i^*}{k}\cdot\frac{1}{f_i(x(\cdot))+k^{-1}},\quad i=r+1,\ldots,m.$$

So we have

$$\hat{u}_i(\cdot)\leq\frac{2}{\sigma}\,\frac{u_i-u_i^*}{k}=\frac{2u_i}{k\sigma}$$

and

$$\|\hat{u}_{(m-r)}(\cdot)\|=\|\hat{u}_{(m-r)}(\cdot)-u_{(m-r)}^*\|\leq 2\sigma^{-1}k^{-1}\|u_{(m-r)}-u_{(m-r)}^*\|.$$

Now we are going to show that the estimation (10) holds for $x(t,k)=x(\cdot)$ and $\hat u_{(r)}(t,k)=(\hat u_i(t,k),\ i=1,\dots,r)\equiv \hat u_{(r)}(\cdot)$.

To this end we differentiate the identities (12) and (13) with respect to t.

From (12) we obtain

$$f''_{0xx}(x(\cdot))x'_t(\cdot)-\sum_{i=1}^{r}\hat u_i(\cdot)f''_{ixx}(x(\cdot))x'_t(\cdot)$$
$$-(f'_{(r)}(x(\cdot)))^{\mathrm T}\hat u'_{(r),t}(\cdot)-h'_t(x(\cdot),\cdot)\equiv 0,$$

i.e.

$$\bar L''_{xx}(x(\cdot),\hat u_{(r)}(\cdot))x'_t(\cdot)-(f'_{(r)}(\cdot))^{\mathrm T}\hat u'_{(r),t}(\cdot)\equiv h'_t(x(\cdot),\cdot) \tag{15}$$

where

$$\bar L''_{xx}(x(\cdot),\hat u_{(r)}(\cdot))=f''_{0xx}(x(\cdot))-\sum_{i=1}^{r}\hat u_i(\cdot)f''_{ixx}(x(\cdot)),$$

$$x'_t(\cdot)=J_t(x(\cdot))=(x'_{j,t}(\cdot),j=1,\dots,n):\mathbb R^m\to\mathbb R^n,$$

$$\hat u'_{(r),t}(\cdot)=J_t(\hat u_{(r)}(\cdot))=(\hat u'_{i,t}(\cdot),\ i=1,\dots,r):\mathbb R^m\to\mathbb R^r.$$

Let

$$D_r(\cdot)=[\operatorname{diag}(kf_i(x(\cdot))+1)]_{i=1}^r:\mathbb R^r\to\mathbb R^r.$$

Then

$$D_r^{-1}(\cdot)=[\operatorname{diag}(kf_i(x(\cdot)+1)^{-1})]_{i=1}^r.$$

Differentiating (13) with respect to t and multiplying both sides to k^{-1} we obtain

$$-\operatorname{diag}(kt_i+u_i^*)D_r^{-2}(\cdot)f'_{(r)}(x(\cdot))x'_t(\cdot)-k^{-1}\hat u'_{(r),t}(\cdot)=[-D_r^{-1}(\cdot);\ O^{r,m-r}]. \tag{16}$$

Multiplying both sides of the system (16) to $D_r^2(\cdot)$ we obtain

$$-\operatorname{diag}(kt_i+u_i^*)f'_{(r)}(x(\cdot))x'_t(\cdot)-k^{-1}D_r^2(\cdot)\hat u'_{(r),t}(\cdot)=[-D_r(\cdot),\ O^{r,m-r}]. \tag{17}$$

Let

$$\Phi'(\cdot)=\begin{array}{c}\\ n\\ r\end{array}\!\!\begin{array}{c}\begin{array}{cc}\qquad n\qquad\qquad & \qquad r\qquad\end{array}\\ \begin{bmatrix}\bar L''_{xx}(x(\cdot),\hat u_{(r)}(\cdot)) & -(f'_{(r)}(x(\cdot)))^{\mathrm T}\\ -\operatorname{diag}(kt_i+u_i^*)f'_{(r)}(x(\cdot)) & -k^{-1}D_r^2(\cdot)\end{bmatrix}\end{array}.$$

Then combining (15), (17) we obtain

$$\begin{array}{c}\\ n\\ r\end{array}\!\!\begin{array}{c} m\\ \begin{bmatrix}x'_t(\cdot)\\ \hat u'_{(r),t}(\cdot)\end{bmatrix}\end{array}=\Phi'^{-1}(\cdot)\times\begin{array}{c}\\ n\\ r\end{array}\!\!\begin{array}{c} m\\ \begin{bmatrix}h'_t(x)(\cdot),\cdot)\\ [-D_r(\cdot);\ O^{r,m-r}]\end{bmatrix}\end{array}=\Phi'^{-1}(\cdot)R(\cdot). \tag{18}$$

In order to estimate the norm of the $(n+r)\times(n+r)$ matrix $\Phi'^{-1}(\cdot)$ and the $(n+r)\times m$ matrix $R(\cdot)$ we will consider the $n\times m$ and $r\times r$ matrices $h'_t(x(\cdot),\cdot)$ and $D_r(\cdot)$ in more detail. We recall that $h(x(\cdot),\cdot)=\sum_{i=r+1}^{m}\hat u_i(x(\cdot),\cdot)f_i'^{\mathrm T}(x(\cdot))$. Further let $f_{(m-r)}(x(\cdot))=(f_i(x(\cdot)),\ i=r+1,\dots,m)$, $\hat u_{(m-r)}(\cdot)=(\hat u_i(x(\cdot),\cdot),\ i=r+1,\dots,m)$,

$D_{m-r}(\cdot)=[\operatorname{diag}(kf_i(x(\cdot))+1)]_{i=r+1}^m$, $t_{(m-r)}=(t_i,\ i=r+1,\ldots,m)$, $D(t_{(m-r)})=[\operatorname{diag} t_i]_{i=r+1}^m$, $D(f_{(m-r)}(x(\cdot)))=[\operatorname{diag} f_i(x(\cdot))]_{i=r+1}^m$. Then

$$\begin{aligned} h'_t(x(\cdot)\cdot) &= \sum_{i=r+1}^{m} \hat{u}_i(x(\cdot),\cdot) f''_i(x(\cdot))x'_t(\cdot) \\ &\quad + (f'_{(m-r)}(x(\cdot)))^{\mathrm{T}} \hat{u}'_{(m-r),t}(x(\cdot),\cdot), \\ \hat{u}'_{(m-r),t}(\cdot) &= (\hat{u}'_{it}(\cdot),\ i=r+1,\ldots,m) \\ &= [O^{m-r,r};\ kD^{-1}_{m-r}(\cdot)] - k^2 D(t_{(m-r)})D^{-2}_{m-r}(\cdot) f'_{(m-r)}(x(\cdot))x'_t(\cdot). \end{aligned}$$

Now we consider the system (18) for $t=0$ and $k>k_0$. First of all note that $x(0,k)=x^*$, $\hat{u}_{(r)}(0,k)=u^*_{(r)}=(u^*_i,\ldots u^*_r)>0$ and also $\hat{u}_i(x(0,k),0,k)=0$, $i=r+1,\ldots,m$, $f_i(x(0,k))=f_i(x^*)\geq\sigma>0$, $i=r+1,\ldots,m$, $D_r(0,k)=D_r^2(0,k)=I^r$, $D(t_{(m-r)})|_{t_{(m-r)}=0}=O^{m-r,m-r}$, $D(f_{(m-r)}(x^*))=[\operatorname{diag}(f_i(x^*))]_{i=r+1}^m\geq\sigma I^{m-r}$.

Further,

$$\begin{aligned} kD^{-1}_{m-r}(x(0,k)) &= k[\operatorname{diag}(kf_i(x^*)+1)^{-1}]_{i=r+1}^m \leq \sigma^{-1} I^{m-r}, \\ \hat{u}'_{(m-r)}(0,k) &= [O^{m-r,r}, [\operatorname{diag}(f_i(x^*)+k^{-1})^{-1}]_{i=r+1}^m] \leq [O^{m-r,r};\ \sigma^{-1}I^{m-r}], \\ \Phi'_{x\hat{u}_r}(0,k) &= \Phi'_{(k)},\ h'_t(x(0,k);0,k) \\ &= (f'_{(m-r)}(x^*))^{\mathrm{T}}\cdot\hat{u}'_{(m-r)}(0,k) \\ &= (f'_{(m-r)}(x^*))^{\mathrm{T}}[O^{m-r,r};\ [\operatorname{diag}(f_i(x^*)+k^{-1})^{-1}]_{i=r+1}^m]. \end{aligned}$$

Then for the norm of the matrix $h'_t(x(0,k);0,k)$ we obtain the estimate $\|h'_t(x(0,k),0,k)\|\leq\sigma^{-1}\|f'_{(m-r)}(x^*)\|$. So for the matrix $x'_t(0,k)$, and $\hat{u}'_{(r),t}(0,k)$ we have

$$\begin{bmatrix} x'_t(0,k) \\ \hat{u}'_{(r),t}(0,k)\end{bmatrix} = (\Phi'_{(k)})^{-1}, \qquad \begin{bmatrix} h'_t(x(0,k)0,k) \\ [-I^r, O^{r,m-r}]\end{bmatrix} = (\Phi'_{(k)})^{-1}R_0. \tag{19}$$

Taking into account the estimate $\|\Phi^{-1}_{(k)}\|\leq\rho$ and $\|h'_t(x(0,k),0,k)\|\leq\sigma^{-1}\|f'_{(m-r)}(x^*)\|$ from (19) we obtain

$$\max\{\|x'_t(0,k)\|\ \|\hat{u}'_{(r),t}(0,k)\|\}\leq\rho(\sigma^{-1}\|f'_{(m-r)}(x^*)\|+\|I^r\|)=\rho[\sigma^{-1}\|f'_{(m-r)}(x^*)\|+1].$$

So for a small enough $\delta>0$ and any $(t,k)\in S(K,\delta)$ the inequality

$$\|\Phi'^{-1}(x(\tau t,k),\hat{u}_{(r)}(\tau t,k))R(x(\tau t,k);(\tau t,k))\|\leq 2\rho[\sigma^{-1}\|f'_{(m-r)}(x^*)\|+1]=c_0 \tag{20}$$

holds for any $0\leq\tau\leq1$ and any $k\geq k_0$. Also we have

$$\begin{aligned} \begin{bmatrix} x(t,k)-x^* \\ \hat{u}_{(r)}(t,k)-u^*\end{bmatrix} &= \begin{bmatrix} x(t,k)-x(0,k) \\ \hat{u}_{(r)}(t,k)-\hat{u}_{(r)}(0,k)\end{bmatrix} \\ &= \int_0^1 \Phi'^{-1}(x(\tau t,k),\hat{u}_{(r)}(\tau t,k))R(x(\tau t,k);(\tau t,k))[t]\,\mathrm{d}\tau. \end{aligned} \tag{21}$$

So taking into account the estimate (20) and (21) we obtain

$$\max\{\|x(t,k)-x^*\|, \|\hat{u}(t,k)-u^*\|\} \leq c_0\|t\| = c_0 k^{-1}\|u-u^*\|.$$

Let

$$\hat{x}(u,k)=x\left(\frac{u-u^*}{k},k\right), \qquad \hat{u}(u,k)=\left(\hat{u}_{(r)}\left(\frac{u-u^*}{k},k\right), \hat{u}_{(m-r)}\left(\frac{u-u^*}{k},k\right)\right).$$

Then for $c=\max\{2\sigma^{-1}, c_0\}$ we obtain

$$\max\{\|\hat{x}(u,k)-x^*\|, \|\hat{u}(u,k)-u^*\|\} \leq ck^{-1}\|u-u^*\| = \gamma_k\|u-u^*\|$$

i.e. the estimate (10) holds true.

(c) Using the estimate (10) we will prove later that $F(x,u,k)$ is strongly convex in the neighborhood of $\hat{x}=\hat{x}(u,k)=\operatorname{argmin}\{F(x,u,k)\,|\,x\in\mathbb{R}^n\}$ uniformly in $(u,k)\in D(u^*,k_0,\delta,\varepsilon)$. Meanwhile note that due to (P2), we have $F'_x(\hat{x}(u^*,k),u^*,k)=0$ and due to (P5), the function $F(x,u^*,k)$ is strongly convex at $\hat{x}(u^*,k)$. So $\hat{x}(u^*,k)=\operatorname{argmin}\{F(x,u^*,k)\,|\,x\in\mathbb{R}^n\}=x^*$ and $\hat{u}(u^*,k)=[\operatorname{diag}(kf_i(x^*)+1)^{-1}]_{i=1}^m u^*=u^*$.

(d) Equalities (12)-(14) show that $\hat{x}=\hat{x}(u,k)$ satisfies the necessary optimality condition for the function $F(x,\hat{u},k)$. This condition, along with the strongly convex $F(x,\hat{u},k)$, in a neighborhood of $\hat{x}$ enables us to prove that $\hat{x}$ is a local minimum $F(x,\hat{u},k)$ in a neighborhood of $\hat{x}$. First let us prove that $F(x,u,k)$ is strongly convex in a neighborhood of $\hat{x}$. We have

$$F'_x(x,u,k)=f'_0(x)-\sum_{i=1}^m u_i(kf_i(x)+1)^{-1}f'_i(x)$$

and

$$F''_{xx}(x,u,k)=f''_0(x)-\sum_{i=1}^m u_i(kf_i(x)+1)^{-1}f''_i(x)$$
$$+k\sum_{i=1}^m u_i(kf_i(x)+1)^{-2}f'^{\mathrm{T}}_i(x)f'_i(x).$$

Therefore, in view of $\hat{u}_i=\hat{u}_i(u,k)=u_i(kf_i(\hat{x})+1)^{-1}$ we obtain

$$F''_{xx}(\hat{x},u,k)=f''_0(\hat{x})-\sum_{i=1}^r \hat{u}_i f''_i(\hat{x})-\sum_{i=r+1}^m u_i(kf_i(\hat{x})+1)^{-1}f''_i(\hat{x})$$
$$+k\sum_{i=1}^r u_i(kf_i(\hat{x})+1)^{-2}f'^{\mathrm{T}}_i(\hat{x})f'_i(\hat{x})$$
$$+k\sum_{i=r+1}^m u_i(kf_i(\hat{x})+1)^{-2}f'^{\mathrm{T}}_i(\hat{x})f'_i(\hat{x})$$
$$\forall (u,k)\in D(u^*,k_0,\delta,\varepsilon).$$

By (10) for a sufficiently large k_0 we have $\hat{x}(u,k)$ near x^* and $\hat{u}(u,k)$ near u^* uniformly in $(u,k)\in D(u^*,k_0,\delta,\varepsilon)$.

So $\hat{x} \to x^*$ and $\hat{u} \to u^*$ lead to

$$f_0''(\hat{x}) - \sum_{i=1}^{r} \hat{u}_i f_i''(\hat{x}) \to L_{xx}''(x^*, u^*),$$

$$k \sum_{i=1}^{r} u_i (kf_i(\hat{x})+1)^{-2} f_i'^{\mathrm{T}}(\hat{x}) f_i'(\hat{x}) \to k f_{(r)}'^{\mathrm{T}}(x^*) U_{(r)}^* f_{(r)}'(x^*).$$

Furthermore, $f_i(\hat{x}) \to f_i(x^*) \geq \sigma \geq 0$, $i = r+1, \ldots, m$, hence

$$\sum_{i=r+1}^{m} u_i (kf_i(\hat{x})+1)^{-1} f_i''(\hat{x}) \to O^{n,n},$$

$$k \sum_{i=r+1}^{m} u_i (kf_i(\hat{x})+1)^{-2} f_i'^{\mathrm{T}}(\hat{x}) f_i(\hat{x}) \to O^{n,n}.$$

So for a large enough k_0 we have

$$F_{xx}''(\hat{x}, u, k) \cong L_{xx}''(x^*, u^*) + k f_{(r)}'^{\mathrm{T}}(x^*) U_{(r)}^* f_{(r)}'(x^*)$$
$$= F_{xx}''(x^*, u^*, k) \quad \forall (u, k) \in D(u^*, k_0, \delta, \varepsilon).$$

In view of (5) and Assertion 1 for $A = L_{xx}''(x^*, u^*)$ and $B = f^{(r)'}(x^*)$ there exists $\mu > 0$ that mineigval $F_{xx}''(x^*, u^*, k) \geq \mu$. Therefore for large enough k_0 and small enough δ the inequality

$$(F_{xx}''(\hat{x}, u, k) y, y) \geq \tfrac{1}{2}\mu(y, y),$$

holds true uniformly in $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$. In view of (12) we have $F_x'(\hat{x}, u, k) = 0$. So the strong convexity of $F(x, u, k)$ in x at the neighborhood of $\hat{x}$ implies that $\hat{x} = \operatorname{argmin}\{F(x, u, k) \mid x \in S(\hat{x}, \varepsilon_0)\}$. Due to (10) $\hat{x}$ is a local minimum of $F(x, u, k)$ in $S(x^*, \varepsilon_0)$.

To complete the proof of part (i) of the theorem we should extend the neighborhood $S(x^*, \varepsilon_0)$ to Ω_k, hence, due to the definition of $F(x, u, k)$, to $\mathbb{R}^n$.

First of all note that

$$F(\hat{x}, u, k) \leq F(x^*, u, k) = f_0(x^*) - k^{-1} \sum_{i=1}^{m} u_i \ln(kf_i(x^*)+1)$$

$$\leq f_0(x^*) - k^{-1} \sum_{i=1}^{r} u_i^* \ln(kf_i(x^*)+1) = f_0(x^*). \tag{22}$$

Suppose that there exists a vector $\tilde{x} \in \Omega_k$ and a number $\tilde{\lambda} > 0$ such that $F(\tilde{x}, u, k) \leq F(\hat{x}, u, k) - \tilde{\lambda}$. Then from (22) we obtain

$$F(\tilde{x}, u, k) \leq f_0(x^*) - \tilde{\lambda}.$$

Let $I_+(\tilde{x}) = \{i : f_i(\tilde{x}) > 0\}$. Then from the last inequality we obtain

$$f_0(\tilde{x}) \leq f_0(x^*) + k^{-1} \sum_{i \in I_+(\tilde{x})} u_i \ln(kf_i(\tilde{x})+1) - \tilde{\lambda}.$$

So from the assumption (6) and $\|u\| \leq M$ for the large enough k_0 and any $k \geq k_0$ we have $f_0(\tilde{x}) \leq f_0(x^*) - \tfrac{1}{2}\tilde{\lambda}$.

From the other side we have $f_0(\tilde{x}) \geq \min\{f_0(x) \mid x \in \Omega_k\}$, then by assumptions (3)-(5) from Theorem 6 of [6, p. 34] we obtain $f_0(\tilde{x}) \geq f_0(x^*) - k^{-1}\sum_{i=1}^{r} u_i^*$, therefore taking large enough k_0 we will get for any $k \geq k_0$ that $f_0(\tilde{x}) \geq f_0(x^*) - \frac{1}{4}\tilde{\lambda}$.

This contradiction completes the proof of part (i) of the theorem.

(ii) (a) If $f_0(x)$ and $-f_i(x)$, $i = 1, \ldots, m$, are convex functions and Ω^* is a compact, the existence of the minimum of $F(x, u, k)$ in $\hat{x}$ over Ω_k for any $u > 0$ and any $k > 0$ follows from Lemma 12 of [6, p. 95]. Moreover, $F'_x(\hat{x}, u, k) = 0$ because $F(x, u, k) \to \infty$ when $x \to \partial\Omega_k$.

(b) The convexity of $F(x, u, k)$ on x and (P3) ensure $\hat{x}(u^*, k) = x^*$, $\hat{u}(u^*, k) = u^*$ for any $k > 0$.

(c) If the assumptions (4) and (5) are fulfilled then this statement can be proved in the same manner as in the nonconvex case, but we don't need the assumption (6), because in this case $\Omega^* = \{x^*\}$ and the level set of $F(x, u, k)$ is bounded for any fixed $u > 0$ and $k > 0$.

The theorem is proved. □

Remark 1. Theorem 1 can be proved similarly for the function $C(x, u, k)$ if we consider a point-to-point mapping

$$\Phi_C(x, \hat{u}, t, k) = \Big(f_0'^{\mathrm{T}}(x) - \sum_{i=1}^{r} \hat{u}_i f_i'^{\mathrm{T}}(x) - h(x, t, k),$$
$$(kt_i + u_i^*)(kf_i(x) + 1)^{-2} - \hat{u}_i,\ i = 1, \ldots, r\Big).$$

Remark 2. Theorem 1 is generally invalid if, instead of $F(x, u, k)$, one considers the classical Lagrangian $L(x, u)$ for the problem.

Example. Let us consider a problem

$$x^* = \operatorname{argmin}\{x_1^2 - x_2^2 \mid f_1(x) = 2 - x_2 \geq 0, f_2(x) = x_2 \geq 0\} = (0, 2).$$

The corresponding classical Lagrangian $L(x, u) = x_1^2 - x_2^2 - u_1(2 - x_2) - u_2 x_2$. Then $u_1^* = 4$, $u_2^* = 0$, $f_{(r)}(x) = f_1(x)$, $L''_{xx}(x^*, u^*) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, $f'_{(r)}(x^*) = f'_1(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ and $f'_{(r)}(x^*)y = 0 \Rightarrow y = \begin{pmatrix} y_1 \\ 0 \end{pmatrix}$, so $(L''_{xx}(x^*, u^*)y, y) = 2y_1^2\ \forall y: f'_{(r)}y = 0$, i.e. the second order optimality conditions (4)-(5), are fulfilled. But $\inf\{L(x, u^*) \mid x \in \mathbb{R}^2\} = \inf\{x_1^2 - x_2^2 + 4x_2 - 8 \mid x \in \mathbb{R}^2\} = -\infty$, and moreover, $\inf\{L(x, u) \mid x \in \mathbb{R}^2\} = -\infty$ for any $u = (u_1, u_2) > 0$. Now let us consider the equivalent problem $x^* = \operatorname{argmin}\{x_1^2 - x_2^2 \mid k^{-1}\ln(k(2 - x_2) + 1) \geq 0,\ k^{-1}\ln(kx_2 + 1) \geq 0\}$ and the corresponding classical Lagrangian

$$F(x, u, k) = x_1^2 - x_2^2 - k^{-1}u_1 \ln(k(2 - x_2) + 1) - k^{-1}u_2 \ln(kx_2 + 1).$$

Then

$$F''_{xx}(x^*, u^*, k) = L''_{xx}(x^*, u^*) - \frac{4}{k}(\ln(k(2 - x_2^*) + 1))''_{xx}$$
$$= \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} - \frac{4}{k}\begin{pmatrix} 0 & 0 \\ 0 & -k^2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4k - 2 \end{pmatrix}.$$

So $F''_{xx}(x^*, u^*, k)$ is positive definite and $x^* = (0, 2) = \text{argmin}\{F(x, u^*, k) \mid x \in \mathbb{R}^2\}$ for any $k > \frac{1}{2}$.

Remark 3. All the facts of Theorem 1 remain in force in case of a convex programming problem if only (2) and (5′) hold and instead of $D(u^*, k_0, \delta, \varepsilon)$ we take

$$D_I(u^*, k_0, \delta, \varepsilon) = \{(u, k): u_i \geq \varepsilon > 0,\ i \in I \mid u_i - u_i^* \mid \leq \delta k,\ i \in I,\ u_i = 0,\ i \notin I,\ k \geq k_0\}$$

for any minimal set I.

4. Shifted barrier functions

To use Theorem 1 one has to know a pair $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$. But a priori we don't know such pairs (u, k), as well as $x^0 \in \text{int}\, \Omega$.

To find $x^0 \in \text{int}\, \Omega$ we can use the multipliers method (see [24]) for solving the problem

$$\bar{x} = \text{argmax}\left\{ \min_{1 \leq i \leq m} f_i(x) \mid x \in \mathbb{R}^n \right\}.$$

If $\text{int}\, \Omega \neq \emptyset$ then after some steps of the method of [24] we will get $x^0: f_i(x^0) > 0$ $i = 1, \ldots, m$.

Let $e = (1, \ldots, 1) \in \mathbb{R}^n$. In order to find the pair $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$ we will consider the shifted barrier function

$$M(x, k) = \begin{cases} F(x, e, k) = f_0(x) - k^{-1} \sum_{i=1}^{m} \ln(kf_i(x) + 1), & \text{if } x \in \text{int}\, \Omega_k, \\ \infty, & \text{if } x \notin \Omega_k. \end{cases}$$

Note that if the condition (6) holds then $\exists k_0 > 0$ such that the next inequalities

$$f_0(x) \geq M(x, k) \geq f_0(x) - O(k^{-1} \ln k) \quad \forall x \in \Omega \text{ hold for any } k \geq k_0. \tag{23}$$

The next theorem allows us to find $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$.

Theorem 2. *If functions $f_i(x)$, $i = 0, \ldots, m$, are continuous, and there exists $k_0 > 0$ such that Ω_{k_0} is a compact then*:

(i) *For any $k \geq k_0$ there exists $x(k) = \text{argmin}\{M(x, k) \mid x \in \mathbb{R}^n\}$ such that*

$$M'_x(x(k), k) = 0$$

and

$$\lim_{k \to \infty} f_0(x(k)) = \lim_{k \to \infty} M(x(k), k) = f_0(x^*).$$

(ii) *If $f_i(x) \in C^2$, conditions* (4) *and* (5) *are fulfilled then there exists k_0 such that for any $k \geq k_0$ the vector $x(k)$ exists and for the pair $(x(k), u(k))$, where $u(x(k)) = [\text{diag}(kf_i(x(k)) + 1)^{-1}]_{i=1}^{m} e$, the estimate*

$$\max\{\|x(k) - x^*\|, \|u(k) - u^*\|\} \leq ck^{-1} \tag{24}$$

holds true with $c > 0$ independent of $k \geq k_0$.

(iii) *Under condition* (ii) *the function $M(x, k)$ is strongly convex in a neighborhood of $x(k)$.*

Proof. (i) The $x(k)=\operatorname{argmin}\{M(x,k)\,|\,x\in\mathbb{R}^n\}=\operatorname{argmin}\{M(x,k)\,|\,x\in\Omega_k\}$ exist because Ω_k is a compact, $M(x,k)$ is continuous in $\operatorname{int}\Omega_k$ and increases infinitely as x approaches the boundary of Ω_k. Because of the last property we obtain $M'_x(x(k),k)=0$.

To prove $\lim_{k\to\infty} f_0(x(k))=\lim_{k\to\infty} M(x(k),k)=f_0(x^*)$, we consider any converging subsequence $\{x(k_s)\}\subset\{x(k)\}$ and let $\lim_{k_s\to\infty} x(k_s)=\bar{x}$. It is easy to see that $\bar{x}\in\Omega$, so due to (23) we obtain $f_0(\bar{x})\leqslant M(\bar{x},k_s)+O(k_s^{-1}\ln k_s)$. Therefore for any small $\varepsilon>0$ we can find large enough k_s that $f_0(\bar{x})\leqslant M(\bar{x},k_s)+\varepsilon$. Then, taking large enough k_s one obtains

$$f_0(x(k_s))-\varepsilon\leqslant f_0(\bar{x})\leqslant M(\bar{x},k_s)+\varepsilon\leqslant M(x(k_s),k_s)+\varepsilon+\varepsilon,$$

i.e., $f_0(x(k_s))\leqslant M(x(k_s))+3\varepsilon\leqslant M(x^*,k_s)+3\varepsilon\leqslant f_0(x^*)+3\varepsilon$. Taking into account that $\varepsilon>0$ is arbitrarily small, we obtain

$$f_0(\bar{x})=\lim_{k_s\to\infty} f_0(x(k_s))\leqslant f_0(x^*),\quad \text{so } f_0(\bar{x})=f_0(x^*).$$

Therefore for any subsequence $k_s\to\infty$ we have $\lim_{k_s\to\infty} f_0(x(k_s))=f_0(\bar{x})=f_0(x^*)$. Hence $\lim_{k\to\infty} f_0(x(k))=f_0(x^*)$ and, by (23), $\lim_{k\to\infty} M(x(k),k)=f_0(x^*)$.

(ii) If (4) and (5) are fulfilled then (x^*,u^*) is a unique K-K-T pair. Therefore $\lim_{k\to\infty} f(x(k))=\lim_{k\to\infty} M(x(k),k)=f_0(x^*)$ and $\lim_{k\to\infty} x(k)=x^*$, $\lim_{k\to\infty} u(k)=u^*$. In addition $M(x,k)\to\infty$ if $x\to\partial\Omega_k$, so

$$M'_x(x(k),k)=f'_0(x(k))-\sum_{i=1}^{m} u_i(k)f'_i(x(k))=0. \tag{25}$$

Now we are going to estimate $\|\Delta x\|=\|x(k)-x^*\|$, $\|\Delta u\|=\|u(k)-u^*\|$. For $f_i(x)\in C^2$ we have $f'_i(x(k))=f'_i(x^*)+f''_i(x^*)\Delta x+h^f_i(\Delta x)$, and $h^f_i(0)=0$, $i=0,\ldots,m$, $\|h^r_i(\Delta x)\|\leqslant\alpha_i(\Delta x)\Delta x$, $\alpha_i(\Delta x)\to 0$ as $\Delta x\to 0$, $1=0,\ldots,m$. Then, $u_i(k)=u_i(x^*,k)+u'_i(x^*,k)\Delta x+h^u_i(\Delta x)$ and $h^u_i(0)=0$, $\|h^u_i(\Delta x)\|\leqslant\beta_i(\Delta x)\Delta x$, where $\beta_i(\Delta x)\to 0$ as $\Delta x\to 0$, $i=1,\ldots,m$.

Note that $u_i(x^*,k)=1$, $i\in I^*=\{1,\ldots,r\}$, while $f_i(x^*)\geqslant\sigma>0$ for $i\notin I^*$. Therefore, $u_i(x^*,k)=O(k^{-1})$, $i\notin I^*$. Then, we have $u'_i(x,k)=-kf'_i(x)(kf_i(x)+1)^{-2}$, so, for $i\in I^*$ we obtain $u'_i(x^*,k)=-kf'_i(x^*)$. And in view of $f_i(x^*)\geqslant\sigma>0$ for $i\notin I^*$ we have $\|u'_i(x^*,k)\|\leqslant k\|f'_i(x^*)\|(k\sigma+1)^{-2}=O(k^{-1})$.

We now replace $u_i(k)$, $i\in I^*$, and $f'_i(x(k))$, $i\in I^*\cup\{0\}$, in (25) by their values. Taking into account that $u_i(k)=\Delta u_i+u^*_i$ we obtain

$$\begin{aligned}0&=f'_0(x^*)+f''_0(x^*)\Delta x+h^f_0(\Delta x)\\&\quad-\sum_{i=1}^{r}(\Delta u_i+u^*_i)(f'_i(x^*)+f''_i(x^*)\Delta x+h^f_i(\Delta x))+\sum_{i=r+1}^{m}u_i(k)f'_i(x(k))\\&=f'_0(x^*)-\sum_{i=1}^{r}u^*_if'_i(x^*)\\&\quad+\left(f''_0(x^*)-\sum_{i=1}^{r}u^*_if''_i(x^*)\right)\Delta x-\sum_{i=1}^{r}f'_i(x^*)\Delta u_i+h^f(\Delta x),\end{aligned}$$

where

$$h^f(\Delta x) = h_0^f(\Delta x) - \sum_{i=1}^{m} (\Delta u_i + u_i^*) h_i^f(\Delta x) - \sum_{i=1}^{r} \Delta u f_i''(x^*)\Delta x + \sum_{i=r+1}^{m} u_i(k) f_i'(x(k)),$$

$$h^f(0) = 0, \qquad \|h^f(\Delta x)\| \leq \alpha(\Delta x)\|\Delta x\|,$$

and $\alpha(\Delta x) \to 0$ as $\|\Delta x\| \to 0$. In view of K–K–T's condition (3) we can rewrite the latter as

$$L''_{xx}\Delta x - f'^{\mathrm{T}}_{(r)}\Delta u + h^f(\Delta x) = 0. \tag{26}$$

Then, we have

$$\begin{aligned} u_i(k) &= u_i(x^*, k) + u_i'(x^*, k)\Delta x + h_i^u(\Delta x) \\ &= 1 - kf_i'(x^*)\Delta x + h_i^u(\Delta x), \quad i = 1, \ldots, r. \end{aligned} \tag{27}$$

Let $\bar{e} = (1, \ldots, 1) \in \mathbb{R}^r$, $\bar{u}^* = (u_i^*, \ldots, u_r^*)$, $v^* = \bar{e} - \bar{u}^*$, $h^u(\Delta x) = (h_i^u(\Delta x),\ i = 1, \ldots, r)$. Then (27) takes the form

$$\Delta u = v^* - kf'_{(r)}(x^*)\Delta x + h^u(\Delta x) \tag{28}$$

with $h^u(0) = 0$ and $\|h^u(\Delta x)\| \leq \beta(\Delta x)\|\Delta x\|$, where $\beta(\Delta x) \to 0$ as $\|\Delta x\| \to 0$. Combining (26) and (28) we obtain

$$D\begin{pmatrix} \Delta x \\ \Delta u \end{pmatrix} = \begin{pmatrix} L''_{xx} & -f'^{\mathrm{T}}_{(r)} \\ -f'_{(r)} & k^{-1}I^r \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta u \end{pmatrix} = \begin{pmatrix} 0 \\ k^{-1}v^* \end{pmatrix} + \begin{pmatrix} h^f(\Delta x) \\ k^{-1}h^u(\Delta x) \end{pmatrix}.$$

As shown in the proof of Theorem 1, the matrix D is nonsingular and for sufficiently large $k \geq k_0$ there exists a constant $\rho_0 > 0$ independent of k such that $\|D^{-1}\| \leq \rho_0$. Let $p = (0, k^{-1}v^*)$ and $q(\Delta x) = (h^f(\Delta x); k^{-1}h^u(\Delta x))$. Then $\|q(\Delta x)\| \leq \gamma(\Delta x)\|\Delta x\|$ and $\gamma(\Delta x) \to 0$ as $\|\Delta x\| \to 0$. Recalling again that $x(k) \to x^*$ and $u(k) \to u^*$, we obtain for the vector $\Delta z = (\Delta x, \Delta u) = D^{-1}p + D^{-1}q(\Delta x)$ there exists an independent of $k \geq k_0$ constant $c > 0$ such that $\|\Delta z\| \leq ck^{-1}$, i.e., estimates (24) hold.

(iii) Finally, we show that $F(x, k)$ is strongly convex in a neighborhood of $x(k)$. We have

$$F'_x(x(k), k) = F'_x(x, k)|_{x=x(k)} = f_0'(x(k)) - \sum_{i=1}^{m} (kf_i(x(k)) + 1)^{-1} f_i'(x(k)),$$

$$\begin{aligned} F''_{xx}(x(k), k) &= F''_{xx}(x, k)|_{x=x(k)} \\ &= f_0''(x(k)) - \sum_{i=1}^{m} (kf_i(x(k)+1))^{-1} f_i''(x(k)) \\ &\quad + k\sum_{i=1}^{m} (kf_i(x(k)) + 1)^{-2} f_i'^{\mathrm{T}}(x(k)) f_i'(x(k)) \\ &= f_0''(x(k)) - \sum_{i=1}^{m} u_i(k) f_i''(x(k)) \\ &\quad + k\sum_{i=1}^{m} u_i(k)(kf_i(x(k)) + 1)^{-1} f_i'^{\mathrm{T}}(x(k)) f_i'(x(k)). \end{aligned}$$

We have $u_i(k)\to u_i^*, f_i(x(k))\to f_i(x^*)=0$, $i=1,\dots,r$, and there exists $\sigma>0$ such that $f_i(x(k))\to f_i(x^*)\geq\sigma>0$, $i=r+1,\dots,m$. Therefore, $k(kf_i(x(k))+1)^{-2}=O(k^{-1})$, $i=r+1,\dots,m$. Moreover, $u_i(k)=O(k^{-1})$, $i=r+1,\dots,m$. Consequently, if $k\to\infty$, then

$$f_0''(x(k))-\sum_{i=1}^{m}u_i(k)f_i''(x(k))\to L_{xx}''(x^*,u^*),$$

$$k\sum_{i=1}^{r}(kf_i(x(k))+1)^{-2}f_i'^{\mathrm{T}}(x(k))f_i'(x(k))\to kf_{(r)}'^{\mathrm{T}}U^*f_{(r)}'$$

and $\sum_{i=r+1}^{m}u_i(k)f_i''(x(k))\to O^{n,n}$,

$$k\sum_{i=r+1}^{m}u_i(k)(kf_i(x(k))+1)^{-1}f_i'^{\mathrm{T}}(x(k))f_i'(x(k))\to O^{n,n},$$

i.e., for sufficiently large k we have $F_{xx}''(x(k),k)\simeq L_{xx}''(x^*,u^*)+k(f_{(r)}'(x^*))^{\mathrm{T}}U^*f_{(r)}'(x^*)$. Therefore, in view of $f_i(x)\in C^2$ the strong convexity of $F(x,k)$ in a neighborhood of $x(k)$ for sufficiently large k follows from Assertion 1. The theorem is proved. □

Remark 4. The results of Theorem 2 remain in force for the function $N(x,k)=C(x,e,k)$ if we set $x(k)=\operatorname{argmin}\{N(x,k)\mid x\in\mathbb{R}^n\}$, $u(k)=(u_i(k)=(kf_i(x(k))+1)^{-2}$, $i=1,\dots,m)$.

5. Modified Barrier Function Method

In this section we introduce and investigate the MBF method for solving constrained optimization problems. We consider different versions of the MBFM for convex and nonconvex programming problems.

The version with a permanent penalty parameter k has a linear rate of convergence. By increasing the parameter from step to step, one can obtain MBFM with a superlinear rate of convergence. Note that the CIPM generally do not converge to the solution with a permanent penalty parameter. If the penalty parameter increases infinitely CIPM converge to the solution only with arithmetical rate of convergence (24).

For a given $k>0$ we consider a bounded set $U_k\neq\emptyset$, a contraction operator $C_k: U_k\to U_k$ and the transformation $T_k: U_k\to\mathbb{R}^n\times U_k$. Then $C_kU_k=\{\hat{u}=C_ku\colon u\in U_k\}\subset U_k$, so there exists a bounded set $U\supset U_k\supset C_kU_k\supset\cdots\supset C_k^sU_k\supset\cdots$. Also, we consider a sequence of sets

$$T_kU_k=\{(\hat{x},\hat{u})\colon \hat{x}=\hat{x}(u,k),\ \hat{u}=C_ku,\ u\in U_k\},$$

$$T_k^2U_k=T_k(T_kU_k),\dots,T_k^s=T_k(T_k^{s-1}U_k),\dots .$$

There exists a bounded set $X \subset \mathbb{R}^n : x^* \in X$ and $X \times U \supset T_k U_k \supset T_k^2 U_k \cdots \supset T_k^s U_k \supset \cdots$.

For a given $k > 0$ consider a nonnegative function

$$v(y, k) \equiv v(x, u, k) = \max\left\{-\min_{1\le i\le m} f_i(x), \|F'_x(x, u, k)\|, \sum_{i=1}^m u_i|f_i(x)|\right\}$$

defined on $X \times U$. If (1) is a convex programming problem, then the relation

$$v(y, k) = 0 \Leftrightarrow y = y^* = (x^*, u^*)$$

holds true for any $k > 0$.

For a given $k > k_0$ and $u \in U_k$ we consider a sequence $\{y^s = T_k^s u\}_{s=1}^\infty = \{T_k(T_k^{s-1}u)\}_{s=1}^\infty = \{\hat{x}(u^{s-1}, k), \hat{u}(u^{s-1}, k)\}_{s=1}^\infty$. Then

$$v(y^s, k) = \max\left\{-\min_{1\le i\le m} f_i(x^s), \sum_{i=1}^m u_i^s|f_i(x^s)|\right\},$$

because

$$F'_x(x^s, u^{s-1}, k) = f'_0(x^s) - \sum_{i=1}^m u_i^{s-1}(kf_i(x^s)+1)^{-1}f'_i(x^s)$$

$$= f'_0(x^s) - \sum_{i=1}^m u_i^s f'_i(x^s) = L'_x(x^s, u^s) = 0.$$

For any $u \in U_k$ the sequence $\{y^s\}_{s=1}^\infty$ belongs to $X \times U$ therefore there exists such a constant $L > 1$ that

$$v(y^s, k) = v(y^s, k) - v(y^*, k) \le L\|y^s - y^*\| \quad \forall u \in U_k, \; s \ge 1.$$

For a given $0 < \gamma \le \frac{1}{2}$ and $u \in U_k$ we can find such $\bar{k} > k_0$ that $\gamma_{\bar{k}} \le \gamma L^{-1}$. Due to Theorem 1, for any $k \ge \bar{k} > k_0$ and for the sequence $\{y^s = T_k^s u\}_{s=1}^\infty$ the estimate

$$\max\{\|x^s - x^*\|, \|u^s - u^*\|\} \le \gamma_k \|u^{s-1} - u^*\|, \quad \gamma_k \le \gamma_{\bar{k}}, \; s \ge 1,$$

holds true. Therefore $v(y^s, k) \le L\gamma_k\|u^{s-1} - u^*\| \le \gamma\|u^{s-1} - u^*\| \le \gamma^s\|u - u^*\|, 0 < \gamma \le \frac{1}{2}$. So for an a priori given $0 < \gamma \le \frac{1}{2}$ one can find $k \ge k_0$ that for any $u \in U_k$ the sequence $\{y^s = T_k^s u\}_{s=1}^\infty$ exists and the sequence $\{v(y^s, k)\}_{s=0}^\infty$ is bounded by $\{\gamma^s\}_{s=0}^\infty$.

Further for any sequence $\{k_s\}_{s=1}^\infty$, $k_1 > k_0$, $k_{s+1} > k_s$, $\lim_{s\to\infty} k_s = \infty$, there exists such a sequence $\{\gamma_s\}_{s=1}^\infty$, $\gamma_{s+1} < \gamma_s$, $\lim_{s\to\infty} \gamma_s = 0$, that for $\{y^{s+1} = T_{k_s}u^s\}_{s=1}^\infty$ the estimate

$$\max\{\|x^{s+1} - x\|, \|u^{s+1} - u^*\|\} \le \gamma_s\|u^s - u^*\|$$

holds and $v(y^s, k_s) \le \gamma_1 \cdots \gamma_s\|u - u^*\|$, $\gamma_s \to 0$ for any $u \in U_{k_1}$.

(a) *Permanent parameter version* (*PPV*). Let start with $x = x^0 \in \operatorname{int} \Omega$, $u = u^0 = e = (1, \ldots, 1) \in \mathbb{R}^m$, $k \ge \bar{k}$ and suppose (x^s, u^s) have been found already. To find the next approximation, $y^{s+1} = (x^{s+1}, u^{s+1})$, one has to fulfill the next operations:

Step 1. Start with $x := x^s$, $u := u^s$.

Step 2. Find $\hat{x} = \hat{x}(u, k)$, $\hat{u} = \hat{u}(u, k) = C_k u$.

Step 3. Set $x^{s+1} := \hat{x}$, $u^{s+1} := \hat{u}$, $s + 1 := s$ and go to Step 1.

The next assertion is a consequence of Theorems 1 and 2.

Assertion 2. *If* (1) *is a convex programming problem,* $f_i(x) \in C^2$, $i = 0, \ldots, m$, *and conditions* (3)–(5) *are fulfilled, then for any* $0 < \gamma \leq \frac{1}{2}$ *there exists such* $\bar{k} > k_0$ *that for any* $k \geq \bar{k}$ *the sequence* $\{y^{s+1} = T_k^{s+1} u\}_{s=1}^{\infty}$ *converges to* y^* *and the estimate*

$$\max\{\|x^{s+1} - x^*\|, \|u^{s+1} - u^*\|\} \leq \gamma_k \|u^s - u^*\|, \quad \gamma_k \leq \gamma,$$

holds true. □

Note that Assertion 2 is true for any $u^0 = u \in U_k$.

So the MBFM with a fixed penalty parameter converges to the solution with a linear rate of convergence. Normally we don't know the threshold $\bar{k}$ a prior. So the second version, which we are going to describe below, allows the adjustment of the penalty parameter on the level, which guarantees the convergence with at least a linear rate.

(b) *Adjusted parameter version* (*APV*). Along with $\Omega_k = \{x : f_i(x) \geq -k^{-1}, i = 1, \ldots, m\} \supset \Omega$ we will consider a set $\Omega_k^+ = \{x : f_i(x) \geq k^{-1}, i = 1, \ldots, m\} \subset \Omega$.

Let $\{k_s > 0\}_{s=0}^{\infty}$, $k_s < k_{s+1}$, $k_s \to \infty$, $k = k(0) = k_0$, $d(0) = 1$, $0 < \gamma \leq \frac{1}{2}$ is fixed, start with $x := \bar{x}^0 = x^0 \in \Omega_k^+$, $u = u^0 = e = (1, \ldots, 1) \in \mathbb{R}^m$ and suppose $\bar{x}^s$, x^s, u^s, $k(s)$, $d(s)$ have been found already. The APV consists of the next steps:

Step 0. Start with $x := \bar{x}^s = \bar{x}$.

Step 1. Set $u := u^s$, $k := k(s)$, $d := d(s)$.

Step 2. Find $\hat{x} = \hat{x}(u, k)$, $\hat{u} = \hat{u}(u, k) = C_k u$, i.e., $\hat{y} = T_k u = (\hat{x}, \hat{u})$.

Step 3. If $v(\hat{y}, k) \leq \gamma^{d+1}$, set $x^{s+1} = \hat{x}$, $u^{s+1} = \hat{u}$, start with $x = x^{s+1}$, $\bar{x}^{s+1} := \bar{x}$, $d(s+1) = d(s) + 1$, $k(s+1) = k(s)$, $s+1 := s$, go to Step 1.

Step 4. If $v(\hat{y}, k) > \gamma^{d+1}$, set $\bar{x}^{s+1} = \operatorname{argmin}\{f_0(x^i) \mid i = 1, \ldots, s+1\}$, $t_{s+1} = \max\{t \mid \bar{x} + t(\bar{x}^{s+1} - \bar{x}) \in \Omega_k^+\}$, $\bar{x}^{s+1} = t_{s+1}\bar{x}^{s+1} + (1 - t_{s+1})\bar{x}$, $u^{s+1} = u^0$, $k(s+1) = k_{s+1}$, $d(s+1) = 1$, $s+1 := s$, and go to Step 0.

The next assertion is a consequence of Theorems 1 and 2.

Assertion 3. *If* (1) *is a convex programming problem,* $f_i(x) \in C^2$, $i = 0, \ldots, m$ *and conditions* (3)–(5) *are fulfilled, then for any* $0 < \gamma \leq \frac{1}{2}$, *there exists a number* s_0 *such that* $k(s) = k_{s_0} = k$, $s \geq s_0$, *the sequence* $\{y^s\}_{s=0}^{\infty}$ *converges to* $y^* = (x^*, u^*)$ *and the estimate*

$$\max\{\|x^{s+1} - x^*\|, \|u^{s+1} - u^*\|\} \leq \gamma_k \|u^s - u^*\|, \quad s \geq s_0, \ \gamma_k \leq \gamma, \tag{29}$$

holds true. □

Now we are going to consider the variable penalty parameter version of the MBFM for solving convex as well as nonconvex programming problems.

(c) *Varying parameter version (VPV)*. *Let* $\{k_s>0\}_{s=0}^{\infty}$, $k_s<k_{s+1}$, $\lim_{s\to\infty} k_s=\infty$, $u^0=e\in\mathbb{R}^m$, $x^0=\bar{x}^0\in\operatorname{int}\Omega$, $k:=k_0$, suppose x^s, $\bar{x}^s$, u^s have been found already. The VPV step consists of the next operations:

Step 1. Start with $x:=\bar{x}^s$, $u:=u^s$, $k:=k_s$.

Step 2. Find $x^{s+1}=\hat{x}(u,k)$, $u^{s+1}=\hat{u}(u,k)=C_k u$, i.e., $y^{s+1}=T_k u^s$,

Step 3. $\bar{x}^{s+1}=\operatorname{argmin}\{f_0(x^i)\,|\,i=1,\ldots,s+1\}$, $t_{s+1}=\max\{t\,|\,\bar{x}^s+t(\bar{x}^{s+1}-\bar{x}^s)\in\Omega_k^+\}$.

Step 4. Set $\bar{x}^{s+1}:=t_{s+1}\bar{x}^{s+1}+(1-t_{s+1})\bar{x}^s$, $s+1:=s$, go to Step 1.

Assertion 4. *If* $f_i(x)\in C^2$, $i=0,\ldots,m$, *the conditions* (3)-(5) *are fulfilled and* $u^0=e\in\mathbb{R}^m$ *is well defined for the parameter* k_0, *then the sequence* $\{y^{s+1}=T_{k_s}u^s\}_{s=1}^{\infty}$ *converge to* y^* *and the estimate*

$$\max\{\|x^{s+1}-x^*\|,\ \|u^{s+1}-u^*\|\}\leq\gamma_1\cdots\gamma_s\|e-u^*\|,\quad \gamma_s\to 0, \tag{30}$$

holds true. □

Assertion 4 follows from Theorems 1 and 2.

Corollary. *If* $\varepsilon>0$ *is small enough and the conditions of Assertion* 4 *are fulfilled then for any* $k>k_0$, *any* $u\in U_k$ *and any start* $x\in S(\hat{x}(u,k),\varepsilon)$, *the VPV of the MBFM leads to finding a minimum of the strongly convex and smooth function at every step even if the initial problem is nonconvex. In addition, estimation* (30) *holds true.* □

We would like to emphasize that for the CBF method under the same assumptions instead of estimate (30) one can guarantee only the estimate (24).

To realize the above mentioned versions of MBFM numerically, we have to replace the infinite procedure of finding $\hat{x}=\hat{x}(u,k)=\operatorname{argmin}\{F(x,u,k)\,|\,x\in\mathbb{R}^n\}$ by a finite procedure maintaining the properties of the vector $\hat{x}$. In the next section we describe such a method.

6. Newton Modified Barrier Method

To maintain the properties of the MBFM without solving the unconstrained optimization problem at every step, one has to use a finite procedure, which allows to find an approximation for $\hat{x}$ with certain accuracy. Below we describe such a method, which is based on the APV of the MBFM and on the global converging step size version of the Newton Method.[1] Let $\varepsilon>0$ be small enough, $\{k_s\}_{s=0}^{\infty}$, $k_{s+1}>k_s$, $\lim_{s\to\infty}k_s=\infty$, $k=k(0)=k_0$, $d(0)=1$, $0<\gamma\leq\frac{1}{2}$ is fixed.

[1] The step size can be defined by the Goldstein-Armijo rule (see [4]).

Let start with $x := \bar{x}^0 = x^0 \in \Omega_k^+$, $u^0 = e = (1, \dots, 1) \in \mathbb{R}^m$ and let $\bar{x}^s$, x^s, u^s, $k(s)$, $d(s)$ be already found. To find the approximation (x^{s+1}, u^{s+1}) one has to fulfill the next operations:

Step 0. Start with $x := \bar{x}^s = \bar{x}$.

Step 1. Set $u := u^s$, $k := k(s)$, $d := d(s)$.

Step 2. Find $\zeta = \zeta(x, u, k)$ by solving the system

$$F''_{xx}(x, u, k)\zeta = -F'_x(x, u, k)$$

and set $t := 1$.

Step 3. Check $x + t\zeta \in \Omega_k$ and $F(x + t\zeta, u, k) - F(x, u, k) \leq \frac{1}{3}t(F'_x(x, u, k), \zeta)$.

Step 4. If $x + t\zeta \in \Omega_k$, the last inequality is fulfilled and $t = 1$ set $x := x + \zeta$ and go to Step 5; if $x + t\zeta \in \Omega_k$, the last inequality is fulfilled and $t < 1$ set $x := x + t\zeta$ and go to Step 2; if $x + t\zeta \notin \Omega_k$, and/or the inequality is not fulfilled set $t := \frac{1}{2}t$ and go to Step 3.

Step 5. If $\|\zeta\| \leq \varepsilon$ go to Step 6; otherwise go to Step 2.

Step 6. Set $\hat{x} := x$, $\hat{u} = [\mathrm{diag}(kf_i(\hat{x}) + 1)^{-1}]_1^m u$, $\hat{y} = (\hat{x}, \hat{u})$; if $v(\hat{y}, k) \leq \gamma^{d+1}$ set $x^{s+1} = \hat{x}$, $u^{s+1} = \hat{u}$; start $x := x^{s+1}$, $d(s+1) = d(s) + 1$, $k(s+1) = k(s)$, $s + 1 := s$, $\varepsilon := \varepsilon\gamma$, and go to Step 1.

Step 7. If $v(\hat{y}, k) > \gamma^{d+1}$, set $\hat{x}^{s+1} = \mathrm{argmin}\{f_0(x^i) \mid i = 1, \dots, s+1\}$, $t_{s+1} = \max\{t \mid \bar{x} + t(\hat{x}^{s+1} - \bar{x}) \in \Omega_k^+\}$, $\bar{x}^{s+1} = t_{s+1}\hat{x}^{s+1} + (1 - t_{s+1})\bar{x}$, $u^{s+1} = u^0$, $\varepsilon := \varepsilon k^{-1}$, $k(s+1) = k_{s+1}$, $d(s+1) = 1$, $s + 1 := s$, and go to Step 0.

Assertion 5. *If* (1) *is a convex programming problem,* $f_i(x) \in C^2$, $i = 0, \dots, m$, *and conditions* (3)–(5) *are fulfilled, then for a small enough* $\varepsilon > 0$ *and* $0 < \gamma \leq \frac{1}{2}$, *there exists such* s_0 *that for* $s \geq s_0$:

(i) *The penalty parameter is permanent, i.e.* $k(s) = k_{s_0} = k$ *and the step size* $t = 1$.

(ii) *Every NMBM step ("large" step), i.e. every updating* u *requires* $O(\log_2 \log_2 \varepsilon^{-1})$ *Newton steps.*

(iii) *The sequence* $\{y^s = (x^s, u^s)\}_{s=0}^{\infty}$ *converges to* $y^* = (x^*, u^*)$ *and the estimate*

$$\max\{\|x^s - x^*\|, \|u^s - u^*\|\} \leq \gamma^s, \quad s \geq s_0,$$

holds true. □

Assertion 5 follows from Theorems 1 and 2 and the Newton method properties (see [32, 33]). We will call the approximation (x^{s_0}, u^{s_0}), i.e., the moment when the NMBM switch to the MBF trajectory, a "hot" start. Beginning at this moment, one can update u, i.e. improve the current approximation twice ($\gamma \leq \frac{1}{2}$) in every $O(\log_2 \log_2 \varepsilon^{-1})$ Newton steps in the worst case.

The number s_0 depends on characteristics of the constrained optimization problem in the solution and can be decreased by increasing k_0.

We have already discussed the unpleasant consequences of increasing k. Therefore we are going to study the possibility of improving the estimation (29) by resorting to some means other than increasing k. It turns out that such possibilities exist and are connected to important properties of the dual to (1) problem, which is based on the MBF.

In the next section we are going to consider the duality theory, which is based on the MBF. Because of the excellent MBF properties (P1)–(P5), the dual function as well as the dual problem have some very important characteristics in the convex as well as in the nonconvex case while MBF is a classical Lagrangian for the equivalent problem.

7. Dual problems

First of all we note that since problems (7) and (1) are equivalent for any $k>0$, it follows that the classical Lagrangians $F(x, u, k)$ and $C(x, u, k)$ for problem (7) preserve all the properties of classical Lagrangians for convex programming problems (see [30]), so the following is true.

Assertion 6. *If $f_0(x)$ and all $-f_i(x)$ are convex and Slater's condition holds, then $x^* \in \Omega$ is a solution of problem (7) for any $k>0$ if and only if:*

(i) *There exists a vector $u^* \geqslant 0$ such that*

$$u_i^* f_i(x^*) = 0, \quad i = 1, \ldots, m, \qquad F(x, u^*, k) \geqslant F(x^*, u^*, k) \quad \forall x \in \mathbb{R}^n. \tag{31}$$

(ii) *The pair (x^*, u^*) is a saddle-point of the Lagrangian, i.e.,*

$$F(x, u^*, k) \geqslant F(x^*, u^*, k) \geqslant F(x^*, u, k) \quad \forall x \in \mathbb{R}^n, \ \forall u \in \mathbb{R}^m_+. \qquad \square \tag{32}$$

Let $\psi_k(x) = \sup_{u \geqslant 0} F(x, u, k)$. Then

$$\psi_k(x) = \begin{cases} f_0(x), & \text{if } f_i(x) \geqslant 0, \ i = 1, \ldots, m, \\ \infty, & \text{otherwise,} \end{cases}$$

and the initial problem (1) reduces to finding

$$x^* = \operatorname{argmin}\{\psi_k(x) \mid x \in \mathbb{R}^n\}. \tag{33}$$

Let $\varphi_k(u) = \inf_{x \in \mathbb{R}^n} F(x, u, k)$. Then the dual problem to (1) consists of finding

$$u^* = \operatorname{argmax}\{\varphi_k(u) \mid u \geqslant 0\}. \tag{34}$$

By the definition of $\psi_k(x)$ and $\varphi_k(u)$ we have $f_0(x) = \psi_k(x) \geqslant \varphi_k(u)$ $\forall x \in \Omega, \forall u \in \mathbb{R}^m_+$. Therefore, if $\bar{x}$ and $\bar{u}$ are feasible solutions of the primal and dual problems and $\psi_k(\bar{x}) = \varphi_k(\bar{u})$, then $\bar{x} = x^*$ and $\bar{u} = u^*$. The smoothness of the dual function $\varphi_k(u)$ depends on the convexity and smoothness of the functions $f_i(x)$, $i = 0, \ldots, m$. If (1) is a convex programming problem then for $u \geqslant 0$ and $k > 0$ the function $\varphi_k(u)$ is as smooth as $f_i(x)$, $i = 0, \ldots, m$, if, for example, $f_0(x)$ is strongly convex or $f_0(x)$ strictly convex and Ω is a compact.

If $f_0(x)$ and $f_i(x)$, $i = 1, \ldots, r$, are nonconvex, the next lemma takes place.

Lemma 1. *Let $f_i \in C^2$, $i=0,\ldots,m$, and conditions* (3)–(6) *hold. Then for any fixed $k \geq k_0$ the concave function $\varphi_k(u)$ is twice continuous differentiable in U_k.*

Proof. First of all note that $\varphi_k(u)$ is a concave function for any $k>0$ whether or not the functions $f_0(x)$ and $-f_0(x)$, $i=1,\ldots,m$, are convex. By Theorem 1 the function $F(x,u,k)$ is strongly convex in a neighborhood of $\hat{x}=\hat{x}(u,k)$ $\forall (u,k) \in D(u^*,k_0,\delta,\varepsilon)$. Therefore $\hat{x}(u,k)=\hat{x}(\cdot)$ is a unique minimum of $F(x,u,k)$ over x, while $\varphi_k(u)=F(\hat{x}(u,k),u,k)=F(\hat{x}(\cdot),\cdot)=F(\cdot)$ is smooth in U_k, i.e., there exists $\varphi'_{ku}(u)=F'_x(\cdot)\hat{x}'_u(\cdot)+F'_u(\cdot)=(\varphi'_{ku_1}(\cdot),\ldots,\varphi'_{ku_m}(\cdot))$.

Since the matrix $F''_{xx}(\hat{x}(\cdot),\cdot)$ is positive definite for $(u,k)\in D(u^*,k_0,\delta,\varepsilon)$ the system $F'_x(x,u,k)=0$ yields a unique vector-function $\hat{x}(u,k)$ such that

$$\hat{x}(u^*,k)=x^* \quad \text{and} \quad \hat{x}'_u(u,k)=\hat{x}'_u(\cdot)=-(F''_{xx}(\hat{x}(\cdot),\cdot))^{-1}\cdot F''_{xu}(\hat{x}(\cdot),\cdot)$$

$\forall (u,k)\in D(u^*,k_0,\delta,\varepsilon)$. Since $F'_x(\cdot)=0$ it follows that $\varphi'_{ku}(\cdot)=F'_u(\cdot)=-k^{-1}(\ln(kf_1(\cdot)+1),\ldots,\ln(kf_m(\cdot)+1))$. Furthermore, $\varphi''_{kuu}(\cdot)=F''_{ux}(\cdot)x'_u(\cdot)=-F''_{ux}(\cdot)\times(F''_{xx}(\cdot))^{-1}\times F''_{xu}(\cdot)$. We set $\tilde{f}_k(\cdot)=[\operatorname{diag}(kf_i(\hat{x}(\cdot))+1)^{-1}]_{i=1}^m:\mathbb{R}^m\to\mathbb{R}^m$ and $\tilde{f}_k=\tilde{f}_k(x^*)=[\operatorname{diag}(kf_i(x^*)+1)^{-1}]_{i=1}^m:\mathbb{R}^m\to\mathbb{R}^m$. Then $F''_{ux}(\cdot)=-\tilde{f}_k(\cdot)f'(\cdot)$, $F''_{xu}(\cdot)=-f'^{\mathrm{T}}(\cdot)\tilde{f}_k(\cdot)$, therefore, $\varphi''_{kuu}(\cdot)=-\tilde{f}_k(\cdot)f'(\cdot)(F''_{xx}(\cdot))^{-1}f'^{\mathrm{T}}(\cdot)\tilde{f}_k(\cdot)$. Note that $\varphi''_{kuu}(u^*)=-\tilde{f}_kf'(x^*)(F''_{xx}(x^*,u^*,k))^{-1}f'^{\mathrm{T}}(x^*)\tilde{f}_k=-\tilde{f}_kf'(F''_{xx})^{-1}f'^{\mathrm{T}}\tilde{f}_k$. □

8. Duality theorems

The dual problems based on MBF's not only possess all the properties well known in convex programming but have some new important features.

Theorem 3 (duality, the convex case). *Let $f_0(x)$ and $-f_i(x)$, $i=1,\ldots,m$, be convex.*

(i) *If the Slater condition holds, then the existence of a solution of problem* (1) *implies that problem* (34) *has a solution and $f_0(x^*)=\psi_k(x^*)=\varphi_k(u^*)$ for all $k>0$.*

(ii) *If $f_0(x)$ is strongly convex or if $f_0(x)$ is strictly convex and Ω is compact, $f_i(x)\in C^2$, $i=0,\ldots,m$, then the solution of the dual problem corresponds to that of the primal problem and the optimal values of the objective functions coincide.*

(iii) *If $f_i(x)\in C^2$, $i=0,\ldots,m$, and conditions* (4)–(5) *are satisfied, then for every $k\geq k_0$ there exists a solution of the dual problem and the second order optimality conditions hold for the dual problem.*

Proof. (i) Let x^* be a solution of problem (1). Then Assertion 6 implies that there exists a vector $u^*\geq 0$ satisfying (31). Therefore,

$$\varphi_k(u^*)=\min_{x\in\mathbb{R}^n}F(x,u^*,k)=F(x^*,u^*,k)=f_0(x^*)\geq F(x^*,u,k)$$

$$\geq \min_{x\in\mathbb{R}^n}F(x,u,k)=\varphi_k(u) \quad \forall u\geq 0,$$

i.e., u^* is a solution of the dual problem and $f_0(x^*)=\varphi_k(u^*)$.

(ii) The assumptions imply that $F(x, u, k)$ is strongly convex in $x \in \Omega_k$ for every $u > 0$ and $k > 0$. Therefore the vector $x(u, k) = \text{argmin}\{F(x, u, k) \mid x \in \mathbb{R}^n\}$ is defined uniquely and because of the smoothness of $f_i(x)$, $i = 0, \ldots, m$, the gradients $\varphi'_k(u)$ of the dual function exist. Let $\bar{u}$ be a solution of problem (34) and $\bar{x} = x(\bar{u}, k)$. Then the optimality conditions for problem (34) are satisfied at $\bar{u}$, i.e.,

$$\varphi'_{ku_i}(\bar{u}) = -k^{-1} \ln(kf_i(\bar{x}) + 1) \leq 0 \quad \text{for } i\text{: } \bar{u}_i = 0,$$

$$\varphi'_{ku_i}(\bar{u}) = -k^{-1} \ln(kf_i(\bar{x}) + 1) = 0 \quad \text{for } i\text{: } \bar{u}_i > 0.$$

Further, $\bar{u}_i > 0$ implies $f_i(\bar{x}) = 0$ and it follows from $\bar{u}_i = 0$ that $f_i(\bar{x}) \geq 0$, i.e., $\bar{x} \in \Omega$ and the complementarity conditions $f_i(\bar{x}) \cdot \bar{u}_i = 0$, $i = 1, \ldots, m$, hold for the pair $(\bar{x}, \bar{u})$. Therefore, $\varphi_k(\bar{u}) = f_0(\bar{x}) - k^{-1} \sum_{i=1}^m \bar{u}_i \ln(kf_i(\bar{x}) + 1) = f_0(\bar{x})$, i.e., for the pair of feasible solutions $(\bar{x}, \bar{u})$ of the primal and dual problems we have $\varphi_k(\bar{u}) = f_0(\bar{x})$, hence $\bar{x} = x^*$ and $\bar{u} = u^*$.

(iii) Since (4)-(5) hold true and $k \geq k_0$, it follows from Theorem 1 that $F(x, u, k)$ is strongly convex in x for any $u \in U_k$, so the first part of the statement can be proven as in (ii).

We now show that the second-order optimality conditions in the strict form hold for problem (34), i.e., the gradients of active constraints are linearly independent, the corresponding Lagrange multipliers are positive, and condition type (5) is satisfied.

We first note that the vectors

$$(\overbrace{0, \ldots, 0}^{r}, 0, \ldots, \overset{i}{1}, \ldots, 0) = e_i, \quad i = r+1, \ldots, m,$$

which are gradients of the active constraints $u_i \geq 0$, $i = r+1, \ldots, m$, of the dual problem, are linearly independent, i.e., condition (4) holds true for problem (34). Now we show that condition (5) is satisfied for problem (34). Let us consider the Lagrangian $L(u, \lambda, k)$ for problem (34). We have $L(u, \lambda, k) = \varphi_k(u) + \sum_{i=1}^n \lambda_i u_i$, therefore $L''_{uu}(u, \lambda, k) = \varphi''_{kuu}(u)$. Let $v = (v_1, \ldots, v_m)$, then $(v, e_i) = 0 \Rightarrow v_i = 0$, therefore, any vector $v \in \mathbb{R}^m$: $(v, e_i) = 0$, $i = r+1, \ldots, m$, has the form $v = (v_1, \ldots, v_r, 0, \ldots, 0)$. Taking into account (4)-(5) we obtain from Theorem 1 that for a fixed $k \geq k_0$ the matrix $F''_{xx} = F''_{xx}(x^*, u^*, k)$ is positive definite and the mineigval $F''_{xx} = \mu > 0$. Let maxeigval $F''_{xx} = M_k > 0$, then $\forall y \in \mathbb{R}^n$ we have $\mu^{-1}(y, y) \geq (F''^{-1}_{xx} y, y) \geq M_k^{-1}(y, y)$, i.e. $-\mu^{-1}(y, y) \leq -(F''^{-1}_{xx} y, y) \leq -M_k^{-1}(y, y)$. So for $L''_{uu}(u^*, \lambda^*, k)$ we obtain

$$\begin{aligned}(L''_{uu}(u^*, \lambda^*, k)v, v) &= (\varphi''_{kuu}(u^*)v, v) = (\tilde{f}_k \tilde{f}'(-(F''_{xx})^{-1}) f'^{\mathrm{T}} \tilde{f}_k v, v) \\ &= (-(F''_{xx})^{-1} f'^{\mathrm{T}} \tilde{v}, f'^{\mathrm{T}} \tilde{v}) \leq -M_k^{-1}(f'^{\mathrm{T}} \tilde{v}, f'^{\mathrm{T}} \tilde{v})\end{aligned}$$

where $\tilde{v} = \tilde{f}_k v = ((kf_1(x^*) + 1)^{-1} v_1, \ldots, (kf_r(x^*) + 1)^{-1} v_r, 0, \ldots, 0) = (v_1, \ldots, v_r, 0, \ldots, 0) = (v_{(r)}, 0, \ldots, 0)$. Thus,

$$\begin{aligned}(L''_{uu}(u^*, \lambda^*, k)v, v) &\leq -M_k^{-1}(f'^{\mathrm{T}}_{(r)}(x^*) v_{(r)}, f'^{\mathrm{T}}_{(r)}(x^*) v_{(r)}) \\ &= -M_k^{-1}(f'_{(r)}(x^*) f'^{\mathrm{T}}_{(r)}(x^*) v_{(r)}, v_{(r)}).\end{aligned}$$

It follows from (4) that the Gram matrix $f'_{(r)}(x^*)f'^{\mathrm{T}}_{(r)}(x^*)$ is non-singular so mineigval $f'_{(r)}(x^*)f'^{\mathrm{T}}_{(r)}(x^*)=\mu_0>0$. Hence, for $\tilde{\mu}=M_k^{-1}\mu_0>0$ we obtain

$$(L''_{uu}(u^*,\lambda^*,k)v,v)\leqslant-\tilde{\mu}\|v\|^2\quad\forall v \text{ such that } (v,e_i)=0,\quad i=r+1,\ldots,m,$$

i.e., condition type (5) is verified for the dual problem. Moreover, $\lambda_i^*=-\varphi'_{ku_i}(u^*)=k^{-1}(kf_i(x^*)+1)^{-1}>0$, $i=r+1,\ldots,m$, that together with the linear independence gradients e_i, $i=r+1,\ldots,m$, of the active constraints $u_i\geqslant0$, $i=r+1,\ldots,m$, of the dual problem complete the proof of the theorem. □

Remark 5. All the facts of the above theorem fail to be true if the convexity of the functions $f_0(x)$ and $-f_i(x)$ is abandoned, moreover statement (iii) is in general invaled even for the convex programming problem if the dual problems are based on the classical Lagrangian $L(x,u)$. However, these results are valid for $k\geqslant k_0$ even in the nonconvex case if the dual problems are based on the functions $F(x,u,k)$ or $C(x,u,k)$.

Theorem 4 (duality, nonconvex case). *Let $f_i(x)\in C^2$, $i=0,\ldots,m$, and conditions* (3)–(6) *hold. Then there exists $k_0>0$ such that for $k\geqslant k_0$ the following is true*:

(i) *The existence of a solution of the primal problem guarantees the existence of the dual problem solution and*

$$f_0(x^*)=\varphi_k(u^*).$$

(ii) *The second-order sufficient optimality conditions are satisfied for the dual problem in the strict form.*

(iii) *The pair (x^*,u^*) is a solution of the primal and dual problems if and only if this pair is a saddle-point of $F(x,u,k)$, i.e., if* (32) *holds true.*

Proof. (i) Let x^* be a solution of problem (1). Then it follows from Theorem 1 that $F(x,u^*,k)$ is strongly convex in a neighborhood of x^* and by (3) we have

$$F'_x(x,u^*,k)|_{x=x^*}=f'_0(x^*)-\sum_{i=1}^m u_i^*(kf_i(x^*)+1)^{-1}f'_i(x^*)$$

$$=f'_0(x^*)-\sum_{i=1}^m u_i^*f'_i(x^*)=0.$$

Therefore,

$$\varphi_k(u^*)=\operatorname{argmin}\{F(x,u^*,k)\,|\,x\in\mathbb{R}^n\}=F(x^*,u^*,k),$$

and there exists

$$\varphi'_{ku}(u^*)=k^{-1}(\ln(kf_1(x^*)+1,\ldots,\ln(kf_r(x^*)+1),$$

$$\ln(kf_{r+1}(x^*)+1),\ldots,\ln(kf_m(x^*)+1))$$

$$=(0,\ldots,0;-k^{-1}\ln(kf_{r+1}(x^*)+1),\ldots,-k^{-1}\ln(kf_m(x^*)+1)).$$

So for i: $u_i^*>0$, we have $\varphi'_{ku_i}(u^*)=0$ and for $i>r$, due to $f_i(x^*)\geq\sigma>0$, we obtain $\varphi'_{ku_i}(u^*)\leq -k^{-1}\ln(k\sigma+1)<0$, i.e., at the point u^* the optimality conditions are satisfied for the dual problem, which is always convex whether or not $f_0(x)$ and $-f_i(x)$, $i=0,\ldots,r$, are convex. Thus u^* is a solution of the dual problem.

(ii) Since the dual problem is convex and its solvability is guaranteed by conditions (3)-(5), the second-order sufficient conditions for the dual problem can be proved just as in the previous theorem.

(iii) We first show that if (x^*, u^*) is a solution to the primal and dual problems, then this pair is a saddle-point of the Lagrangian $F(x, u, k)$. Indeed, by Theorem 1 for $k\geq k_0$ the function $F(x, u^*, k)$ is strongly convex in x in a neighborhood of x^* and $F'_x(x^*, u^*, k)=0$, hence $F(x, u^*, k)\geq F(x^*, u^*, k)$ in a neighborhood of x^*. We now extend the latter inequality to all $x\in\Omega_k$, hence, to $\mathbb{R}^n$. Actually, if there is a vector $\tilde{x}\in\Omega_k$, such that $F(\tilde{x}, u^*, k)\leq F(x^*, u^*, k)-\tilde{\lambda}=f_0(x^*)-\tilde{\lambda}$, and $\tilde{\lambda}>0$, then $f_0(\tilde{x})\leq f_0(x^*)+k^{-1}\sum_{i=1}^r u_i^*\ln(kf_i(\tilde{x})+1)-\tilde{\lambda}$. Now repeating the consideration of Theorem 1 (part (i)) we will get from the one side $f_0(\tilde{x})\leq f_0(x^*)-\frac{1}{2}\tilde{\lambda}$ but from the other side $f_0(\tilde{x})\geq f_0(x^*)-\frac{1}{4}\tilde{\lambda}$. This contradiction shows that $F(x, u^*, k)\geq F(x^*, u^*, k)$ for all $x\in\Omega_k$, i.e. $F(x, u^*, k)\geq F(x^*, u^*, k)\ \forall x\in\mathbb{R}^n$ and $k\geq k_0$ whenever k_0 is sufficiently large. Furthermore, since $x^*\in\Omega$, we have $\ln(kf_i(x^*)+1)\geq 0$, $i=1,\ldots,m$, therefore, $f_0(x^*)=F(x^*, u^*, u)\geq F(x^*, u, k)$ for all $u\geq 0$. Therefore

$$F(x, u^*, k)\geq F(x^*, u^*, k)\geq F(x^*, u, k)\quad \forall x\in\mathbb{R}^n,\ \forall u\in\mathbb{R}^m_+.$$

Finally, we show that if $(\bar{x}, \bar{u})$ is a saddle-point of $F(x, u, k)$, i.e., if

$$\begin{aligned} f_0(x)-k^{-1}\sum_{i=1}^m \bar{u}_i\ln(kf_i(x)+1) &\geq f_0(\bar{x})-k^{-1}\sum_{i=1}^m \bar{u}_i\ln(kf_i(\bar{x})+1)\\ &\geq f_0(\bar{x})-k^{-1}\sum_{i=1}^m u_i\ln(kf_i(\bar{x})+1)\\ &\qquad \forall x\in\mathbb{R}^n,\ u\in\mathbb{R}^m_+, \end{aligned}\tag{35}$$

then $\bar{x}=x^*$ and $\bar{u}=u^*$. Indeed, the right inequality in (35) yields

$$\sum_{i=1}^m \bar{u}_i\ln(kf_i(\bar{x})+1)\leq\sum_{i=1}^m u_i\ln(kf_i(\bar{x})+1)\quad \forall u\in\mathbb{R}^m_+.\tag{36}$$

It follows from (36) that $f_i(\bar{x})\geq 0$, $i=1,\ldots,m$, since if there is i_0 such that $f_{i_0}(\bar{x})<0$, then we can take $k>0$ such that $0<kf_{i_0}(\bar{x})+1<1$, i.e. $\ln(kf_{i_0}(\bar{x})+1)<0$ and set $u_i=\bar{u}_i$, $i\neq i_0$, while u_{i_0} can be made sufficiently large to obtain the opposite inequality to (36).

Therefore, it follows from (36) that $\bar{x}$ is a feasible solution to problem (1). Consequently, $\ln(kf_i(\bar{x})+1)\geq 0$, $i=1,\ldots,m$, and $\sum_{i=1}^m \bar{u}_i\ln(kf_i(\bar{x})+1)\geq 0$.

Since (35) holds true for any $u\geq 0$, we set $u=0$ and obtain $\sum_{i=1}^m \bar{u}_i\ln(kf_i(\bar{x})+1)\leq 0$ hence $\sum_{i=1}^m \bar{u}_i\ln(kf_i(\bar{x})+1)=0$ implies $\bar{u}_if_i(\bar{x})=0$, $i=1,\ldots,m$.

The left inequality in (35) yields

$$f_0(x) \geqslant f_0(\bar{x}) + k^{-1} \sum_{i=1}^{m} \bar{u}_i \ln(kf_i(x)+1).$$

For every $x \in \Omega$ we have $\ln(kf_i(x)+1) \geqslant 0$, $i=1,\ldots,m$, therefore, $f_0(x) \geqslant f_0(\bar{x})$ $\forall x \in \Omega$, i.e., $\bar{x}$ is the minimum of $f_0(x)$ over Ω, i.e., $\bar{x} = x^*$. Since $\bar{x} = \operatorname{argmin}\{F(x, \bar{u}, k) \mid x \in \mathbb{R}^n\}$ it follows that

$$F'_x(\bar{x}, \bar{u}, k) = 0 \Rightarrow f'_0(\bar{x}) - \sum_{i \in I^*} \bar{u}_i f'_i(\bar{x}) = f'_0(x^*) - \sum_{i \in I^*} \bar{u}_i f'_i(x^*) = 0,$$

where $I^* = \{i: f_i(\bar{x}) = 0\} = \{i: f_i(x^*) = 0\}$ i.e., $(\bar{x}, \bar{u})$ is a Karush-Kuhn-Tucker pair and hence by (4) we obtain $\bar{u} = u^*$. The theorem is proved. □

Corollary. *The restriction $\bar{\varphi}_k(u) = \varphi_k(u)|_{u_{r+1}=0,\ldots,u_m=0}$ of the dual function to the manifold of the active constraints of the dual problem is strongly concave if the conditions of Theorem 4 are fulfilled.* □

Remark 6. All the facts concerning the dual function $\varphi_k(u)$ and dual problem (34), which have been stated in Theorems 3 and 4 hold for the function $c_k(u) = \min\{C(x, u, k) \mid x \in \Omega_k\}$, for the dual problem

$$u^* = \operatorname{argmax}\{c_k(u) \mid u \geqslant 0\} \tag{37}$$

and for the restriction $\bar{c}_k(u) = c_k(u)|_{u_{r+1}=\cdots=u_m=0}$ of the dual function $c_k(u)$ to the set of active constraints of the dual problem (37).

The convexity and smoothness properties of dual functions can be used for finding nonzero components of u^* by applying smooth optimization methods to $\bar{\varphi}_k(u)$ or $\bar{c}_k(u)$. Let us consider this in detail.

9. Method of controlling sequences for simultaneous solution of primal and dual convex problems

First of all note that the implementation of different versions of the MBFM involve solving unconstrained optimization problems at every step. Therefore, to use these algorithms in practice it is necessary to replace the unconstrained optimization by a finite procedure maintaining the estimates (29).

We consider now the convex programming problem. We choose $\alpha > 0$, find $\tilde{x}$ from the condition

$$\tilde{x} \in \mathbb{R}^n: \quad \|F'_x(\tilde{x}, u, k)\| \leqslant \alpha k^{-1} \|[\operatorname{diag}(kf_i(\tilde{x})+1)^{-1}]_{i=1}^m u - u\|$$

and set

$$\tilde{u} = [\operatorname{diag}(kf_i(\tilde{x})+1)^{-1}]_{i=1}^m u.$$

Let $\eta = \tilde{u} - u^*$. Then

$$\frac{\eta + u^* - u}{k} = k^{-1}([\operatorname{diag}(kf_i(\tilde{x}) + 1)^{-1}]_{i=1}^m u - u).$$

Therefore

$$\|F'_x(\tilde{x}, u, u)\| \leq \alpha k^{-1} \|[\operatorname{diag}(kf_i(\tilde{x}) + 1)^{-1}]_{i=1}^m u - u\| = k^{-1}\alpha(\|\eta\| + \|\Delta u\|) \leq \alpha k^{-1}(\|\Delta u\| + \|\Delta z\|)$$

where $\Delta z = (\Delta x, \Delta u)$, $\Delta x = \tilde{x} - x^*$.

Using arguments as in the proof of Theorem 5 in [19] and taking into account the estimates for Lagrange multipliers corresponding to passive constraints, we obtain from the inequality

$$\|\tilde{x} - x^*\| \leq c(1+\alpha)k^{-1}\|u - u^*\|, \qquad \|\tilde{u} - u^*\| \leq c(1+\alpha)k^{-1}\|u - u^*\|. \tag{38}$$

It gives us the following lemma.

Lemma 2. *If $f_0(x)$ and all $-f_i(x)$, $i = 1, \ldots, m$, are convex and smooth enough and the conditions* (4)–(5) *are fulfilled, then for any $\alpha > 0$ there are such small $\varepsilon > 0$ and $\delta > 0$ that for any $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$ the estimate* (38) *holds.* □

A similar result follows for the function $C(x, u, k)$ if we find $\tilde{x}$ from the condition

$$\|C'_x(\tilde{x}, u, k)\| \leq \alpha k^{-1} \|u[\operatorname{diag}(kf_i(\tilde{x}) + 1)^{-2}]_{i=1}^m - u\|$$

and set $\tilde{u} = [\operatorname{diag}(kf_i(\tilde{x}) + 1)^{-2}]_{i=1}^m u$.

The above arguments enable us to consider the following alternative to the PPV method:

$$x^{s+1} \in \mathbb{R}^n: \quad \|F'_x(x^{s+1}, u^s, k)\| \leq \alpha k^{-1} \|[\operatorname{diag}(kf_i(x^{s+1}) + 1)^{-1}]_{i=1}^m u^s - u^s\|, \tag{39a}$$

$$u^{s+1} = [\operatorname{diag}(kf_i(x^{s+1}) + 1)^{-1}]_{i=1}^m u^s. \tag{39b}$$

So we have the next assertion

Assertion 7. *If $f_0(x)$ and $-f_i(x)$, $i = 1, \ldots, m$, are convex and smooth enough and conditions* (2), (4)–(5) *are fulfilled, then the method* (39) *generates sequence $\{x^s, u^s\}$ such that the estimate*

$$\max\{\|x^{s+1} - x^*\|, \|u^{s+1} - u^*\|\} \leq c(1+\alpha)k^{-1}\|u^s - u^*\|$$

holds and c is independent of $k \geq k_0$ and $\alpha > 0$. □

Now we are going to consider a method for the simultaneous solution of the primal and dual convex programming problems. This method is based on smooth optimization methods and intensively uses the dual problem properties which have been stated above.

Let $\bar{\varphi}_k(u) \equiv \varphi_k(u)|_{u_i=0}$, $i=r+1,\dots,m$, be the restriction of the $\varphi_k(u)$ to the manifold of the active constraints of the dual problem. Then for any $k \geq k_0$,

$$u^* = \operatorname{argmax}\{\bar{\varphi}_k(u) \mid u=(u_1,\dots,u_r,0,\dots,0) \in U_k\}. \tag{40}$$

If the conditions (4)–(5) are fulfilled and $f_i(x)$, $i=1,\dots,m$, smooth enough then $\bar{\varphi}_k(u)$ is strongly concave and smooth enough too. So to solve the problem (40) we can use smooth optimization methods (see [4, 9, 19]).

Based on these methods one can define relaxation operators $R: U_k \to U_k$ with properties

(i) $\|Ru-u^*\| \leq q\|u-u^*\|$, $q<1$,

(ii) $\|Ru-u^*\| \leq q(u)\|u-u^*\|$, where $q(u) \to 0$ as $u \to u^*$,

(iii) $\|Ru-u^*\| \leq q\|u-u^*\|^2$.

We define the gradient relaxation operator by the formula

$$Ru = u + t\bar{\varphi}'_{ku}(u) \tag{41}$$

if the conditions (2)–(5) are fulfilled and $f_i(x)$, $i=0,\dots,m$, is smooth enough that there exists $t>0$ such that the gradient relaxation operator possesses property (i).

We define the Newton operator by the formula

$$Ru = u - (\bar{\varphi}''_{kuu}(u))^{-1}\bar{\varphi}'_{ku}(u). \tag{42}$$

If $f_i(x)$, $i=0,\dots,m$, are smooth enough and (2)–(5) are fulfilled the Newton's operator possesses property (ii) or (iii). Some other relaxation operators with properties (ii) or (iii) can be defined on the basis of smooth optimization methods (see [2, 4, 9]) which require only $\varphi'_{ku}(u)$. To implement the relaxation operators we should find $\hat{x}(u,k)$. Therefore in order to implement these operators numerically we should replace the generally infinite procedure of finding $\hat{x}(u,k)$ by a finite procedure.

Denote by $\tilde{R}$ an operator like R in which $\hat{x}(u,k)$ is replaced by

$$\tilde{x} = \tilde{x}(u,k): \quad \|F'_x(\tilde{x},u,k)\| \leq \delta. \tag{43}$$

We will call the sequence $\{\delta_s\}_{s=0}^{\infty}$ a controlling sequence if $0<\delta_{s+1}<\delta_s$ and $\lim_{s\to\infty}\delta_s = 0$. The controlling sequence $\{\delta_s\}_{s=0}^{\infty}$ is said to be consistent with the operator $\tilde{R}$ if for any $u \in U_k$ and for the sequence $\{\alpha_s = \delta_s\|\tilde{R}^s u - u^*\|^{-1}k^s\}_{s=0}^{\infty}$ the condition

$$\sum \alpha_s < +\infty \tag{44}$$

is fulfilled. The controlling sequences method generates a sequence $\{\tilde{y}^s = (\tilde{x}^s, \tilde{u}^s)\}_{s=0}^{\infty}$ in the next way.

Let $\tilde{u}^0 = \bar{e} = (1,\dots,1) \in \mathbb{R}^r$ and let the vector $\tilde{y}^s = (\tilde{x}^s, \tilde{u}^s)$ have been found already. Set $\tilde{u}^s := \tilde{R}\tilde{u}^s$ and find

$$\tilde{x}^{s+1} \in \Omega_k: \quad \|F'_x(\tilde{x}^{s+1}, \tilde{u}^s, k)\| \leq \delta_s, \tag{45}$$

$$\tilde{u}^{s+1} = [\operatorname{diag}(kf_i(\tilde{x}^{s+1})+1)^{-1}]_{i=1}^{r}\tilde{u}^s. \tag{46}$$

The next assertion takes place.

Assertion 8. *If $f_0(x)$ and all $-f_i(x)$ are convex, $f_i(x) \in C^2$, $i=0,\dots,m$, and the conditions* (2)–(5) *are fulfilled, while the controlling sequence $\{\delta_s\}_{s=0}^{\infty}$ is consistent with the operator $\tilde{R}$, then for sequences $\{\tilde{y}^s = (\tilde{x}^s, \tilde{u}^s)\}_{s=0}^{\infty}$ obtained by* (45)–(46), *the estimates*

(a) $\|\tilde{y}^s - y^*\| \leq (ck^{-1}q)^s, \quad q<1,$

(b) $\|\tilde{y}^s - y^*\| \leq (ck)^{-s} \prod_{i=1}^{s} q_i, \quad q_s \to 0,$

(c) $\|\tilde{y}^s - y^*\| \leq (ck)^{-s} q^{2^s}, \quad q<1,$

hold true provided the operator $\tilde{R}$ possesses one of properties (i)–(iii). □

In the next section we will implement the MBF theory for solving the LP and QP problems.

10. Modified barrier functions in Linear and Quadratic Programming

We start with the implementation of the MBF for solving LP problems.

Let A be an $m \times n$ matrix, $p \in \mathbb{R}^n$, $q \in \mathbb{R}^m$, and there exists the solution of the primal

$$x^* = \operatorname{argmin}\{(p, x) \mid Ax = q, x \geq 0\} \tag{47}$$

and dual

$$v^* = \operatorname{argmax}\{(q, v) \mid vA \leq p\} \tag{48}$$

linear programming problems.

Let $k>0$, $\Omega_k = \{x \in \mathbb{R}^n, x_j \geq -k^{-1}, j=1,\dots,n\}$, $Q = \{x\colon Ax = q\}$, $e=(1,\dots,1) \in \mathbb{R}^n$. The modified barrier functions $F(x,u,k)\colon Q \times \mathbb{R}^n_+ \times \mathbb{R}^1_+ \to \mathbb{R}^1$, $C(x,u,k)\colon Q \times \mathbb{R}^n_+ \times \mathbb{R}^1_+ \to \mathbb{R}^1$, which are correspondent to problem (47) are defined by the formulas

$$F(x,u,k) = \begin{cases} (p,x) - k^{-1} \sum_{j=1}^{n} u_j \ln(kx_j + 1), & \text{if } x \in \operatorname{int} \Omega_k, \\ \infty, & \text{if } x \notin \operatorname{int} \Omega_k, \end{cases}$$

or

$$C(x,u,k) = \begin{cases} (p,x) + k^{-1} \sum u_j((kx_j+1)^{-1} - 1), & \text{if } x \in \operatorname{int} \Omega_k, \\ \infty, & \text{if } x \notin \operatorname{int} \Omega_k. \end{cases}$$

Let $u = p - vA$; we assume that the dual pair (1) and (2) are nondegenerate, i.e., rank $A = m$, $m < n$ and the complementary slackness conditions are fulfilled in strict form with the additional proviso that

$$u_j^* > 0 \text{ and } x_j^* = 0 \quad \text{for } j = 1,\dots,n-m, \tag{49a}$$

$$u_j^* = 0 \text{ and } x_j^* > 0 \quad \text{for } j = n-m+1,\dots,n. \tag{49b}$$

Under the nondegeneracy assumption the optimal solutions x^* and v^* are unique. Let $D(u^*, k_0, \delta, \varepsilon) = \{(u, k) \in \mathbb{R}_+^{n+1}: u = (u_1, \ldots, u_n): u_i \geq \varepsilon > 0,\ |u_i - u_i^*| \leq \delta k,\ i = 1, \ldots, n-m,\ 0 \leq u_i \leq \delta k,\ i = n-m+1, \ldots, n,\ k \geq k_0\}$. The next assertion takes place.

Assertion 9. *If conditions* (49) *are fulfilled and* rank $A = m$, *then there exists* $k_0 > 0$ *and small enough* $\delta > 0$ *that for any* $0 < \varepsilon < \min_{1 \leq i \leq n-m} u_i^*$ *and any* $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$ *the next statements hold*:

(i) *There exists a vector*

$$\hat{x} = \hat{x}(u, k) = \operatorname{argmin}\{F(x, u, k) \mid x \in Q\}.$$

(ii) *The triple* $\hat{x}$, $\hat{u} = \operatorname{diag}(k\hat{x}_j + 1)^{-1} u$, $\hat{v} = (p - \hat{u}) A^{\mathrm{T}} (AA^{\mathrm{T}})^{-1}$, *satisfies the inequality*

$$\max\{\|\hat{x} - x^*\|, \|\hat{u} - u^*\|, \|\hat{v} - v^*\|\} \leq c k^{-1} \|u - u^*\| \tag{50}$$

holds true and $c > 0$ *is independent of* k.

(iii) *The restriction of* $F(x, u, k)$ *to* Q *is strongly convex in a neighborhood of* $\hat{x}$. □

Assertion 9 gives a possibility to realize the PPV of the MBFM if we have a pair $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$. To find such a pair we consider the shifted barrier function $M(x, k) = F(x, e, k)$.

Assertion 10. *If conditions* (49) *are fulfilled and* rank $A = m$, *then there exists* $k_0 > 0$ *such that for every* $k \geq k_0$:

(i) *The vector* $x(k) = \operatorname{argmin}\{M(x, k) \mid x \in Q\}$ *exist.*

(ii) *For the triple* $x(k)$, $u(k) = [\operatorname{diag}(kx_j(k) + 1)^{-1}]_{j=1}^n e$, $v(k) = (p - u(k)) A^{\mathrm{T}} (AA^{\mathrm{T}})^{-1}$ *the estimate*

$$\max\{\|x(k) - x^*\|, \|u(k) - u^*\|, \|v(k) - v^*\|\} \leq c k^{-1} \tag{51}$$

holds and c *is independent of* k.

(iii) *The restriction of* $M(x, k)$ *to* Q *is strongly convex.* □

Remark 7. All of the facts of Assertions 9 and 10 remain in force if instead of $F(x, u, k)$ and $M(x, k)$ we consider $C(x, u, k)$ and $N(x, k) = C(x, e, k)$.

So taking $x^0 \in \operatorname{int} \Omega_k \cap Q$, $u^0 = e$, and large enough $k \geq k_0$, we can develop the PPV of the MBFM for linear programming problems.

The PPV of the MBF method consists of finding the sequence $\{w^s = (x^s, u^s, v^s)\}_{s=1}^{\infty}$ by formulas

$$x^{s+1} = \operatorname{argmin}\{F(x, u^s, k) \mid x \in \Omega\}, \tag{52a}$$

$$u^{s+1} = [\operatorname{diag}(kx_j^{s+1} + 1)^{-1}]_{j=1}^n u^s, \qquad v^{s+1} = (p - u^{s+1}) A^{\mathrm{T}} (AA^{\mathrm{T}})^{-1}. \tag{52b}$$

The next assertion is a consequence of Assertions 9 and 10.

Assertion 11. *If conditions* (49) *are fulfilled and* rank $A=m$ *then there exist such* $k_0>0$ *and* $c>0$, *which are independent of k that the estimate*

$$\max\{\|x^s-x^*\|, \|u^s-u^*\|, \|v^s-v^*\|\}\leqslant ck^{-s}=\gamma_k^s, \quad 0<\gamma_k\leqslant\tfrac{1}{2}, \tag{53}$$

holds for any $k\geqslant k_0$. □

In order to realize method (52) we have to avoid solving problem (52a) at every step, keeping estimate (53).

To solve problem (52a) one can use Newton's method. Now we are going to describe the Newton step for solving (52a). Let

$$U=[\operatorname{diag} u_i]_{j=1}^n, \qquad D_{x,k}=[\operatorname{diag}(kx_j+1)]_{j=1}^n : \mathbb{R}^n\to\mathbb{R}^n.$$

Then

$$F'_x(x,u,k)=F'_x(\cdot)=p-U[\operatorname{diag}(kx_j+1)^{-1}]_{j=1}^n e=p-UD_{x,k}^{-1}e;$$

$$F''_{xx}(x,u,k)=F''_{xx}(\cdot)=kU[\operatorname{diag}(kx_j+1)^{-2}]_{j=1}^n=kUD_{x,k}^{-2}.$$

The Newton's method step for minimizing $F(x,u,k)$ in x consists of finding Newton's direction ζ and updating x by formula

$$x:=x+t\zeta, \quad 0<t\leqslant 1.$$

To find $\zeta=\zeta(x,u,k)$ we have to solve the problem

$$\zeta(x,u,k)=\operatorname{argmin}\{\tfrac{1}{2}(F''_{xx}(x,u,k)\zeta,\zeta)+(F'_x(x,u,k),\zeta)\,|\,A\zeta=0\}, \tag{54}$$

i.e.,

$$\zeta(x,u,k)=\operatorname{argmin}\{\tfrac{1}{2}k(UD_{x,k}^{-2}\zeta,\zeta)+((p-UD_{x,k}^{-1}e),\zeta)\,|\,A\zeta=0\}. \tag{55}$$

Having introduced the Lagrange multipliers $v=(v_1,\ldots,v_m)$ which correspond to the system $A\zeta=0$ we obtain the next system,

$$kUD_{x,k}^{-2}\zeta+(p-UD_{x,k}^{-1}e)=vA, \tag{56a}$$

$$A\zeta=0. \tag{56b}$$

Let $D_{x,k}r=kU\zeta$. Then $\zeta=k^{-1}D_{x,k}U^{-1}r$. Instead of system (56) we obtain the system

$$D_{x,k}^{-1}r=A^{\mathrm{T}}v-p+UD_{k,x}^{-1}e, \tag{57a}$$

$$AU^{-1}D_{x,k}r=0, \tag{57b}$$

or

$$r=D_{x,k}A^{\mathrm{T}}v-D_{x,k}(p-UD_{x,k}^{-1}e), \tag{58a}$$

$$AU^{-1}D_{x,k}r=0. \tag{58b}$$

Putting $u^{-1/2}r = h$ we obtain instead of (58) the system

$$\begin{pmatrix} I^n & -D_{x,k}U^{-1/2}A^{\mathrm{T}} \\ AU^{-1/2}D_{x,k} & O^{m,m} \end{pmatrix}\begin{pmatrix} h \\ v \end{pmatrix} = \begin{pmatrix} U^{1/2}e - U^{-1/2}D_{x,k}p \\ O \end{pmatrix}. \tag{59}$$

The last least square problem can be rewritten as

$$v(x, u, k) = v(\cdot) = \operatorname{argmin}\{\|(AU^{-1/2}D_{x,k})^{\mathrm{T}}v - U^{-1/2}(D_{x,x}p - Ue)\|_2^2 \mid v \in \mathbb{R}^m\}.$$

The last least square problem is equivalent to the next normal system of equations

$$(AD_{x,k}^2U^{-1}A^{\mathrm{T}})v = AD_{x,k}(U^{-1}D_{x,k}p - e). \tag{60}$$

So one can find the Newton direction by the formula

$$\zeta(x, u, k) = k^{-1}D_{x,k}U^{-1}r(x, u, k),$$

where

$$r(x, u, k) = D_{x,k}A^{\mathrm{T}}v(x, u, k) - D_{x,k}p + Ue.$$

The numerical realization of the PPV of the MBFM leads to the Newton Modified Barrier Method (NMBM) for solving the primal problem (47).

The NMBM uses the Newton Method to solve (52a) and update the Lagrange multipliers by formula (52b). Note that instead of solving problem (52a) one can find an approximation for $x^{s+1} = \hat{x}(u^s, k)$ with accuracy 2^{-L}, where L is the input length. If the initial approximation $x^{s,0} = x^s$ is well defined (see [32, 33]), for problem (52a), one can perform the Newton sequence

$$x^{s,j+1} = x^{s,j} + \zeta(x^{s,j}, u^s, k), \quad j = 0, 1, 2, \ldots,$$

which is also well defined (see [29]), i.e. $(F''_{xx}(x^{s,j}, u^s, k))^{-1}$ exist and the sequence $\{x^{s,j}\}_{j=0}^{\infty}$ converges to $\hat{x}(u^s, k)$ quadratically.

So to find an approximation for $\hat{x}(u^s, k)$ with accuracy 2^{-L}, one has to perform $O(\log_2 L)$ Newton steps. If x^s is well defined for problem (52a), and u^s is well defined for the parameter k, one can improve the initial approximation (x^s, u^s) at least twice ($\gamma_k \leqslant \frac{1}{2}$) for $O(\log_2 L)$ steps of the Newton Method.

Now we are going to consider some implementations of the MBF's for solving the QP problems.[2]

Let C be an $n \times n$ symmetric matrix, $A = (a^i)_{i=1}^m$ be an $m \times n$ matrix, $A = \binom{B}{D}$, B an $r \times n$ matrix, D an $(m-r) \times n$ matrix, $p \in \mathbb{R}^n$, $q \in \mathbb{R}^m$, $a^i \in \mathbb{R}^n$, $m > n$. We suppose that there exist

$$x^* = \operatorname{argmin}\{f_0(x) = \tfrac{1}{2}(Cx, x) + (p, x) \mid Ax \geqslant q\} \tag{61}$$

and Karush–Kuhn–Tucker's conditions (3) hold, i.e., there is a vector $u^* \in \mathbb{R}_+^m$ such that

$$Cx^* + p - u^*A = 0 \tag{62}$$

[2] These results we obtained together with B. Yuzefovich (Faculty of Industrial Engineering and Management, Technion, Haifa).

and

$$u_i^*(Ax^*-q)_i = u_i^*[(a^i, x^*) - q_i] = 0, \quad i = 1, \ldots, m. \tag{63}$$

We will suppose that the complementary conditions (63) are fulfilled in the strict form, i.e.

$$u_i^* > 0 \text{ and } (Ax^*-q)_i = 0 \quad \text{for } i = 1, \ldots, r, \tag{64a}$$

$$u_i^* = 0 \text{ and } (Ax^*-q)_i > 0 \quad \text{for } i = r+1, \ldots, m, \tag{64b}$$

and rank A = rank $B = r < n$, i.e. vectors a^i, $i = 1, \ldots, r$, are linearly independent.

Let $L(x, u) = \frac{1}{2}(Cx, x) + (p, x) - (u, Ax - q)$, then $L''_{xx}(x, u) = C$. If condition (5) is fulfilled for the QP problem (61), then there exists $\lambda > 0$,

$$(Cy, y) \geq \lambda(y, y) \quad \forall y\colon By = 0. \tag{65}$$

Let $\Omega_k = \{x\colon r_i(x) = (Ax-q)_i \geq -k^{-1}, i = 1, \ldots, m\}$. On the $\mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^1_+$ we define the MBF's

$$F(x, u, k) = \begin{cases} f_0(x) - k^{-1} \sum_{i=1}^m u_i \ln(kr_i(x)+1), & \text{if } x \in \operatorname{int} \Omega_k, \\ \infty, & \text{if } x \notin \operatorname{int} \Omega_k, \end{cases}$$

and

$$C(x, u, k) = \begin{cases} f_0(x) + k^{-1} \sum_{i=1}^m u_i[(kr_i(x)+1)^{-1} - 1], & \text{if } x \in \Omega_k, \\ \infty, & \text{if } x \notin \Omega_k. \end{cases}$$

Then, for any $k > 0$, we obtain $F(x^*, u^*, k) = C(x^*, u^*, k) = f_0(x^*)$ and

$$F'_x(x, u^*, k)_{|x=x^*} = C'_x(x, u^*, k)_{|x=x^*} = Cx^* + p - u^*A = 0.$$

Let

$$D(u^*, k_0, \delta, \varepsilon) = \{(u, k) \in \mathbb{R}^{m+1}_+\colon u_i \geq \varepsilon > 0, |u_i - u_i^*| \leq \delta k, i = 1, \ldots, r,$$
$$0 \leq u_i \leq \delta \cdot k, i = r+1, \ldots, m, k \geq k_0\}.$$

The next assertion takes place.

Assertion 12. *If* (64) *and* (65) *hold, then there exists such* k_0 *and small enough* $\delta > 0$ *that for any* $0 < \varepsilon < \min\{u_i^* \mid i = 1, \ldots, r\}$ *the next statements hold for any* $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$:

(i) *There exists a vector*

$$\hat{x} \equiv \hat{x}(u, k) = \operatorname{argmin}\{F(x, u, k) \mid x \in \mathbb{R}^n\} \text{ such that } F'_x(\hat{x}, u, k) = 0.$$

(ii) *For the pair* $\hat{x}$, $\hat{u} = [\operatorname{diag}[1/(kr_i(\hat{x})+1)]]_{i=1}^m u$ *the estimate*

$$\max\{\|\hat{x} - x^*\|, \|\hat{u} - \hat{u}^*\|\} \leq ck^{-1}\|u - u^*\|$$

holds and c *is independent of* $k > k_0$.

(iii) *Let* $\hat{U}_B=[\operatorname{diag}\hat{u}_i]_{i=1}^r$, $\hat{U}_D=[\operatorname{diag}\hat{u}_i]_{i=r+1}^m$. *For the matrix*

$$F''_{xx}(\hat{x},u,k)=C+kB^{\mathrm{T}}\hat{U}_B[\operatorname{diag}(kr_i(\hat{x})+1)^{-1}]_{i=1}^r B + kD^{\mathrm{T}}\hat{U}_D[\operatorname{diag}(kr_i(\hat{x})+1)^{-1}]_{i=r+1}^m D,$$

there exist such $0<\mu<\lambda$ *that*

$$(F''_{xx}(\hat{x},u,k)y,y)\geqslant\mu(y,y)\quad\forall y\in\mathbb{R}^n \tag{66}$$

and

$$F''_{xx}(x^*,u^*,k)=C+kB^{\mathrm{T}}U^*_BB.\quad\square$$

Let us consider the shifted barrier function for the QP problem, $M(x,k)=F(x,e,k)$.

The next assertion takes place.

Assertion 13. *If conditions* (64)–(65) *hold, then there exist such* $k_0>0$ *that for any* $k\geqslant k_0$:

(i) *There exist*

$$x(k)=\operatorname{argmin}\{M(x,k)\,|\,x\in\mathbb{R}^n\}:\quad M'_x(x(k),k)=0,$$

and for the pair $(x(k),u(k))$ *where*

$$u(k)=u(x(k))=[\operatorname{diag}(kr_i(x(k))+1)^{-1}]_{i=1}^m e=\Delta^{-1}(x(k))e = (u_1(k),\ldots,u_i(k),\ldots,u_m(k)),$$

$e=(1,\ldots,1)\in\mathbb{R}^m$, *the estimate*

$$\max\{\|x(k)-u^*\|,\|u(k)-u^*\|\}\leqslant ck^{-1} \tag{67}$$

holds and $c>0$ *is independent of* $k\geqslant k_0$.

(ii) *The function* $M(x,k)$ *is strongly convex in a neighborhood of* $x(k)$, *i.e. for the matrix*

$$M''_{xx}(x(k),k)=C+kA^{\mathrm{T}}\Delta^{-2}(x(k))A=C+kA^{\mathrm{T}}U(x(k))\Delta^{-1}(x(k))A,$$

where $U(x(k))=[\operatorname{diag}u_i(k)]_{i=1}^m$, *there exists* $\mu>0$ *independent of* $k\geqslant k_0$ *such that*

$$(M''_{xx}(x(k),k)y,y)\geqslant\mu(y,y)\quad\forall y\in\mathbb{R}^n.\quad\square$$

Assertions 11 and 12 allow us to develop the PPV method for solving QP problems. Let $x^0\in\operatorname{int}\Omega$, $u^0=e=(1,\ldots,1)\in\mathbb{R}^m$, and $k\geqslant k_0$. The sequence $\{x^s,u^s\}_{s=0}^{\infty}$ we obtain by the formulas

$$x^{s+1}=\operatorname{argmin}\{F(x,u^s,k)\,|\,x\in\mathbb{R}^n\}, \tag{68a}$$

$$u^{s+1}=[\operatorname{diag}(kr_i(x^{s+1})+1)^{-1}]_{i=1}^m u^s. \tag{68b}$$

As a consequence of Assertions 12 and 13 we have the next assertion:

Assertion 14. *If conditions* (62), (64), (65) *hold then for the sequence* (68) *the estimate*

$$\max\{\|x^s - x^*\|, \|u^s - u^*\|\} \leqslant (ck^{-1})^s = \gamma_k^s, \quad 0 < \gamma_k \leqslant \tfrac{1}{2}, \tag{69}$$

holds true and $c > 0$ *is independent of* $k \geqslant k_0$. □

Now we consider an important particular case of the Quadratic Programming problem

$$x^* = \operatorname{argmin}\{f_0(x) = \tfrac{1}{2}(Cx, x) - (p, x) \mid x \geqslant 0\}. \tag{70}$$

For the solution x^* we have $u^* = Cx^* - p \geqslant 0$ and the complementary condition

$$u_i^* x_i^* = (Cx^* - p)_i x_i^* = 0, \quad i = 1, \ldots, n, \tag{71}$$

is fulfilled.

We suppose that the complementary condition holds in the strict form, i.e.

$$u_i^* = (Cx^* - p)_i > 0 \quad \text{and} \quad x_i^* = 0, \qquad i = 1, \ldots, r, \tag{72a}$$

$$u_i^* = (Cx^* - p)_i = 0 \quad \text{and} \quad x_i^* > 0, \qquad i = r+1, \ldots, n. \tag{72b}$$

Let $\Omega_k = \{x_i, \ldots, x_n): x_i \geqslant -k^{-1}, i = 1, \ldots, n\}$. Then the MBF, which correspond to the QP problem, (70) is defined on $\mathbb{R}^n \times \mathbb{R}_+^n \times \mathbb{R}_+^1$ by formula

$$F(x, u, k) = \begin{cases} \tfrac{1}{2}(Cx, x) - (p, x) - k^{-1} \sum u_i \ln(kx_i + 1), & x \in \operatorname{int} \Omega_k, \\ \infty, & x \notin \operatorname{int} \Omega_k. \end{cases} \tag{73}$$

So $F(x^*, u^*, k) = f_0(x^*)$, $F'_x(x, u^*, k)|_{x=x^*} = Cx^* - p - u^* = 0$, hence if $f_0(x)$ is a convex function we obtain

$$x^* = \operatorname{argmin}\{F(x, u^*, k) \mid x \in \mathbb{R}^n\}.$$

Let

$$C = \begin{pmatrix} C_{rr} & C_{r,n-r} \\ C_{n-r,r} & C_{n-r,n-r} \end{pmatrix}$$

where C_{rr} and $C_{n-r,n-r}$ are symmetric $r \times r$ and $(n-r) \times (n-r)$ matrices, $C_{r,n-r} = C_{n-r,r}^{\mathrm{T}}$ is an $r \times (n-r)$ matrix. We consider $y \in \mathbb{R}^n$: $(y, e_i) = 0$, $i = 1, \ldots, r$, where

$$e_i = (0, \overset{i}{1}, 0; 0, \ldots, 0).$$

So $y = (0, \ldots, 0, y_{r+1}, \ldots, y_n)$, let $\bar{y} = (y_{r+1}, \ldots, y_n)$. The classical Lagrangian for the problem (70) is $L(x, u) = f_0(x) - \sum_{i=1}^n u_i x_i$, so $L''_{xx}(x, u) = C$. Therefore the condition (65) can be rewritten in the form

$$(Cy, y) \geqslant \lambda(y, y) \quad \forall y: (y, e_i) = 0, \ i = 1, \ldots, r,$$

i.e.

$$(C_{n-r,n-r}\bar{y}, \bar{y}) \geqslant \lambda(\bar{y}, \bar{y}), \quad \lambda > 0. \tag{74}$$

Let $D(u^*, k_0, \delta, \varepsilon) = \{(u, k) \in \mathbb{R}_+^{m+1}: \ u_i \geqslant \varepsilon > 0, \ |u_i - u_i^*| \leqslant \delta \cdot k, i = 1, \ldots, r, \ 0 \leqslant u_i \leqslant \delta \cdot k, \ i = r+1, \ldots, m, \ k \geqslant k_0\}$.

Assertion 15. *If conditions* (72) *and* (74) *are fulfilled, then for any pair* $(u, k) \in D(u_i^*, k_0, \delta, \varepsilon)$ *the next statements hold*:

(*i*) *There exists a vector*

$$\hat{x} = \hat{x}(u, k) = \operatorname{argmin}\left\{ f_0(x) - k^{-1} \sum_{i=1}^{n} u_i \ln(kx_i + 1) \,\middle|\, x \in \mathbb{R}^n \right\}.$$

such that $C\hat{x} - p - [\operatorname{diag}(k\hat{x}_i + 1)^{-1}]_{i=1}^n u = C\hat{x} - p - \hat{u} = 0$.

(ii) *For the pair* $\hat{x}$ *and* $\hat{u}$ *the estimate*

$$\max\{\|\hat{x} - x^*\|, \|\hat{u} - u^*\|\} \leq ck^{-1}\|u - u^*\|$$

holds and $c > 0$ *is independent of* $k > k_0$.

Let $\hat{U} = [\operatorname{diag} \hat{u}_i]_{i=1}^n$. *Then for the matrix*

$$F''_{xx}(\hat{x}, u, k) = C + k\hat{U}[\operatorname{diag}(k\hat{x}_i + 1)^{-1}]_{i=1}^n$$

there exist such $0 < \mu < \lambda$ *that*

$$((C + k\hat{U}[\operatorname{diag}(k\hat{x}_i + 1)^{-1}]_{i=1}^n)y, y) \geq \mu(y, y) \quad \forall y \in \mathbb{R}^n$$

and

$$F''_{xx}(x^*, u^*, k) = C + k\begin{bmatrix} U_{r,r}^* & O^{r,n-r} \\ O^{n-r,r} & O^{n-r,n-r} \end{bmatrix},$$

where $U_{r,r}^* = [\operatorname{diag} u_i^*]_{i=1}^r$. □

Let $N(x, k) = F(x, e, k)$. Then the next assertion takes place.

Assertion 16. *If conditions* (72) *and* (74) *hold then there exists such a* $k_0 > 0$ *that for any* $k \geq k_0$:

(i) *There exists* $x(k) = \operatorname{argmin}\{N(x, k) \mid x \in \Omega_k\}$: $M'_x(x(k), k) = 0$ *and for the vector* $x(k)$, *and* $u(k) = (u_1(k), \ldots, u_n(k)) = [\operatorname{diag}(kx_i(k) + 1)^{-1}]_{i=1}^n e$, *where* $e = (1, \ldots, 1) \in \mathbb{R}^n$, *the estimate*

$$\max\{\|x(k) - x^*\|, \|u(k) - u^*\|\} \leq ck^{-1}$$

holds and $c > 0$ *is independent of k.*

(ii) *The function* $N(x, k)$ *is strongly convex in a neighborhood of* $x(k)$, *and*

$$N''_{xx}(x(k), k) = C + k\begin{pmatrix} U_{r,r}(k) & O^{r,n-r} \\ O^{n-r,r} & O^{n-r,n-r} \end{pmatrix}$$

where $U_{r,r}(k) = [\operatorname{diag} u_i(k)]_{i=1}^r > 0$. □

Assertions 15 and 16 lead to the next multipliers method for solving the QP problem (70).

Let $x^0 > 0$, $u^0 = e = (1, \ldots, 1) \in \mathbb{R}^n$. Then

$$x^{s+1} = \operatorname{argmin}\{\tfrac{1}{2}(Cx, x) - (p, x) - k^{-1} \textstyle\sum u_i^s \ln(kx_i + 1) \mid x \in \mathbb{R}^n\}, \tag{75a}$$

$$u^{s+1} = [\operatorname{diag}(kx_i^{s+1} + 1)^{-1}]_{i=1}^n u^s. \tag{75b}$$

From Assertions 15 and 16 we obtain:

Assertion 17. *If* (72) *and* (74) *hold then for the sequence* $\{x^s, u^s\}_{s=0}^{\infty}$, *which is defined by* (75) *the next estimate holds*:

$$\max\{\|x^s - x^*\|, \|u^s - u^*\|\} \leq (ck^{-1})^s. \qquad \square$$

Remark 8. If $C = G^{\mathrm{T}}G$ and G is an $m \times n$ matrix, then it is sufficient that rank $G = n - r$ (the last $n - r$ columns should be linearly independent) to fulfill condition (74).

If $C = 0$ then problem (61) turns into the LP problem

$$x^* = \operatorname{argmin}\{(p, x) \mid r(x) = Ax - q \geq 0\}, \tag{76}$$

which is equivalent to the dual problem (48). Let $\Omega_k = \{x\colon r_i(x) = (Ax - q)_i \geq -k^{-1}, i = 1, \dots, m\}$. The modified barrier function which corresponds to problem (76) we define by the formula

$$F(x, u, k) = \begin{cases} (p, x) - k^{-1} \sum_{i=1}^{n} u_i \ln(kr_i(x) + 1), & \text{if } x \in \operatorname{int} \Omega_k, \\ \infty, & \text{if } x \notin \operatorname{int} \Omega_k. \end{cases}$$

Let $A = \binom{B}{N}$, B be an $n \times n$ matrix, N be an $(m - n) \times n$ matrix and rank $A =$ rank $B = n$. We also suppose that the complementary conditions are fulfilled in the strict form

$$u_i^* > 0 \quad \text{and} \quad r_i(x^*) = 0, \qquad i = 1, \dots, n, \tag{77a}$$

$$u_i^* = 0 \quad \text{and} \quad r_i(x^*) > 0, \qquad i = n + 1, \dots, m. \tag{77b}$$

Let

$$D(u^*, k_0, \delta, \varepsilon) = \{(u, k) \in \mathbb{R}_+^{m+1}\colon u_i \geq \varepsilon, |u_i - u_i^*| \leq \delta k, i = 1, \dots, n,$$
$$0 \leq u_i \leq \delta k, i = n + 1, \dots, m, k \geq k_0\},$$

$$U = [\operatorname{diag} u_i]_{i=1}^m, \quad U_B = [\operatorname{diag} u_i]_{i=1}^n, \quad U_N = [\operatorname{diag} u_i]_{i=n+1}^m,$$

$$\Delta(x, k) = [\operatorname{diag}(kr_i(x) + 1)]_{i=1}^m, \quad \Delta_B(x, k) = [\operatorname{diag}(kr_i(x) + 1)]_{i=1}^n,$$

$$\Delta_N(x, k) = [\operatorname{diag}(kr_i(x) + 1)]_{i=n+1}^m.$$

So $U_B^* = [\operatorname{diag} u_i^*]_{i=1}^n$, $U_N^* = O^{m-n,m-n}$, $\Delta_B(x^*, k) = I^n$, $U_N^* \Delta_N^{-1}(x^*, k) = O^{m-n,m-n}$. The next assertion takes place.

Assertion 18. *If condition* (77) *is fulfilled and* rank $A = n$ *then for any* $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$ *the next statements hold*:

(i) *There exists a vector*

$$\hat{x} = \hat{x}(u, k) = \operatorname{argmin}\{F(x, u, k) \mid x \in \mathbb{R}^n\}$$

such that

$$F'_x(\hat{x}, u, k) = p - A^{\mathrm{T}}\Delta^{-1}(\hat{x}, k),\ u = p - A^{\mathrm{T}}\hat{u} = 0.$$

(ii) *For the pair* $\hat{x}$ *and* $\hat{u} = \Delta^{-1}(\hat{x}, k)u$ *the estimation*

$$\max\{\|\hat{x} - x^*\|, \|\hat{u} - u^*\|\} \leq ck^{-1}\|u - u^*\|$$

holds and $c > 0$ *is independent of* $k \geq k_0$.

(iii) *The matrix*

$$F''_{xx}(\hat{x}, u, k) = kA^{\mathrm{T}}\hat{U}\Delta^{-1}(\hat{x}, k)A = kB^{\mathrm{T}}\hat{U}_B\Delta_B^{-1}(\hat{x}, k)B + kN^{\mathrm{T}}\hat{U}_N\Delta^{-1}(\hat{x}, k)N$$

is positive definite uniformly in $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$, *i.e. there exists* $\mu > 0$,

$$(F''_{xx}(\hat{x}(u, k), u, k)y, y) \geq k\mu(y, y) \qquad \forall y \in \mathbb{R}^n,\ \forall (u, k) \in D(u^*, k_0, \delta, \varepsilon),$$

and

$$F''_{xx}(x^*, u^*, k) = kB^{\mathrm{T}}U_B^*B. \qquad \square$$

Now let us consider the Shifted Barrier Function which corresponds to the LP problem (76). We obtain $M(x, k) = F(x, e, k)$.

Assertion 19. *If condition* (77) *is fulfilled and* rank $A = n$, *then there exists such* $k_0 > 0$ *that for any* $k \geq k_0$:

(i) *There exists*

$$x(k) = \operatorname{argmin}\{M(x, k) \mid x \in \mathbb{R}^n\}$$

such that

$$M'_x(x(k), k) = p - A^{\mathrm{T}}\Delta^{-1}(x(k), k)e = 0.$$

(ii) *For the pair of vectors* $x(k)$ *and* $u(k) = \Delta^{-1}(x(k), k)e = (u_1(k), \ldots, u_m(k))$ *the estimate*

$$\max\{\|x(k) - x^*\|, \|u(k) - u^*\|\} \leq ck^{-1}$$

holds true with $c > 0$ *is independent of* k.

(iii) *Let* $U(k) = [\operatorname{diag} u_i(k)]_{j=1}^m$, *the matrix*

$$F''_{xx}(x(k), k) = kA^{\mathrm{T}}U(k)\Delta^{-1}(x(k), k)A$$

is positive definite. Moreover, there exist $\mu > 0$ *independent of* $k \geq k_0$ *that*

$$(F''_{xx}(x(k), k)y, y) \geq k\mu(y, y) \quad \forall y \in \mathbb{R}^n. \qquad \square$$

Now we consider the PPV of the MBFM for the LP problem (76). Let $x^0 \in \operatorname{int} \Omega_k$, $u^0 = e = (1, \ldots, 1) \in \mathbb{R}^m$ and (x^s, u^s) have been found already. The next approximation (x^{s+1}, u^{s+1}) one finds by formulas

$$x^{s+1} = \operatorname{argmin}\{(p, x) - k^{-1}\Sigma u_i^s \ln(kr_i(x) + 1 \mid x \in \mathbb{R}^n\}, \tag{78a}$$

$$u^{s+1} = \Delta^{-1}(x^{s+1}, k)u^s. \tag{78b}$$

The next assertion is a consequence of Assertions 18 and 19.

Assertion 20. *If the slackness complementary conditions* (77) *are fulfilled in the strict form and* rank A = rank $B = n$, *then for the sequence* $\{x^s, u^s\}_{s=0}^{\infty}$ *the estimate*

$$\max\{\|x^s - x^*\|, \|u^s - u^*\|\} \leq (ck^{-1})^s = \gamma_k^s, \quad 0 < \gamma_k \leq \tfrac{1}{2}, \tag{79}$$

holds and $c > 0$ *is independent of* $k \geq k_0$. □

The numerical realization of method (78) leads to the Newton Modified Barrier Method (NMBM) for simultaneous solution the dual pair LP problem. The NMBM consists of using the Newton method for solving problem (78a) and updating the Lagrange multipliers by formula (78b).

The Newton step for finding

$$\hat{x}(u, k) = \operatorname{argmin}\{F(x, u, k) \mid x \in \mathbb{R}^n\} \tag{80}$$

under fixed $u > 0$ and $k > 0$ leads to finding

$$\bar{x} = x - (F''_{xx}(x, u, k))^{-1} F'_x(x, u, k) = x + \zeta(x, u, k)$$

i.e., $\zeta(x, u, k)$ is the solution of the normal system of equations

$$A_k^{\mathrm{T}} U \Delta^{-2}(x, k) A \zeta = -(p - u\Delta^{-1}(x, k)A). \tag{81}$$

So if x^s is well defined (see [32, 33]) for problem (80) with $u = u^s$ then the sequence

$$x^{s,j+1} = x^{s,j} + \zeta(x^{s,j}, u^s, k), \quad j = 0, 1, 2, \ldots, \qquad x^{s,0} = x^s,$$

is well defined (see [29]), i.e. $(F''_{xx}(x^{s,j}u^s, k))^{-1}$, $j = 0, 1, 2, \ldots$, exists and $\{x^{s,j}\}_{j=0}^{\infty}$ converges to $\hat{x}(u^s, k)$ quadratically. Therefore in $O(\log_2 L)$ Newton method steps one can obtain approximation $\tilde{x}^{s+1}$ for $x^{s+1} = \hat{x}(u^s, k)$ with precision 2^{-L}. If $\max_{1 \leq j \leq m, 1 \leq j \leq n} \{|a_{ij}|, |p_j|, |q_i|\} \leq 2^l$, then, under the natural assumption: $l \ll n < m$, the input length L can be estimated by $l(m+1)^2$. So, in $O(\log_2 m)$ Newton method steps in the worst case, we can obtain the approximation $\tilde{x}^{s+1}$ to $x^{s+1} = x(u^s, k)$ with accuracy 2^{-L}.

Using the approximation $\tilde{x}^{s+1}$ instead of x^{s+1} in formula (78b), we obtain an approximation $\tilde{u}^{s+1}$ for u^{s+1} with property (79). Therefore after $O(\log_2 m)$ Newton steps one can update the Lagrange multipliers, i.e., to realize one "large" NMBM step, which, due to Assertion 20, allows us to improve the current approximation at least twice ($\gamma_k \leq \frac{1}{2}$). In addition the new vector of the Lagrange multipliers $\tilde{u}^{s+1}$, again, due to Assertion 20, is well defined, i.e., $(\tilde{u}^{s+1}, k) \in D(u^*, k_0, \delta, \varepsilon)$. So, in the dual space we are in a position where the basic theorem acts again. As for the primal space, if condition (77) is fulfilled and rank A = rank $B = n$, then

$$\operatorname{mineigval} F''_{xx}(x^*, u^*, k) \geq k\mu_0 \min u^*$$

where $\mu_0 = \operatorname{mineigval} B^{\mathrm{T}}B$, $\min u^* = \min\{u_i^* \mid i = 1, \ldots, n\} > 0$. Therefore there exists $k_0 > 0$ and $1 > \alpha_0 > 0$ that for a fixed $k \geq k_0$,

$$\operatorname{mineigval} F''_{xx}(\hat{x}(u, k), u, k) \geq k\alpha_0\mu_0 \min u^*$$

uniformly in $u \in U_k$. Further, due to estimation (79) we obtain $\|F'_x(\tilde{x}^s, \tilde{u}^s, k)\| \to 0$ for a fixed $k \geq k_0$. Hence there exists such s_0 that for any $s \geq s_0$ if $\tilde{x}^s$ is well defined (see [32]) for the Newton method in conformity to problem (80) with $u = \tilde{u}^s$, then the approximation $\tilde{x}^{s+1}$ will be well defined for the same problem with $u = \tilde{u}^{s+1}$. So, beginning from $\tilde{y}^{s_0} = (\tilde{x}^{s_0}, \tilde{u}^{s_0})$ ("hot" start) in every "large" NMBM step, i.e. after every updating of the Lagrange multipliers, which in the worst case needs $O(\log_2 m)$ Newton method steps, one can improve the current approximation at least twice ($\gamma_k \leq \frac{1}{2}$).

Note that the number s_0 can be decreased by increasing k_0.

11. Some concluding remarks

The properties (P1)-(P5) cause the principal difference between MB and CB functions. Note that the constrained optimization problem (1) is equivalent to the unconstrained nonsmooth problem

$$x^* = \operatorname{argmin}\{\psi(x, x^*) \mid x \in \mathbb{R}^n\} \tag{82}$$

where $\psi(x, x^*) = \max\{f_0(x) - f_0(x^*), -f_i(x), i = 1, \ldots, m\}$, it follows from (P1)-(P5) that $F(x, u^*, k)$ is an exact smooth approximation of the nonsmooth function $\psi(x, x^*)$ for any $k > 0$ in the convex case and $k \geq k_0$ in the nonconvex case. This indicates that in order to solve the constrained optimization problem (1), or the nonsmooth problem (82), one has to solve the smooth unconstrained problem

$$x^* = \operatorname{argmin}\{F(x, u^*, k) \mid x \in \mathbb{R}^n\}$$

where $F(x, u^*, k)$ is strongly convex in the neighborhood of x^*.

On the other hand CBF $\varphi(x, k)$ does not exist at the solution, and cannot be an exact smooth approximation of the $\psi(x, x^*)$ for any $k > 0$.

So together with the penalty parameter $k > 0$, which is the only tool in CBF for improving the smooth approximation of the nonsmooth function ($\psi(x, x^*)$, the MBF has another tool — the vector of Lagrange multipliers. Therefore, the sequence $\{F(x, u^s, k)\}_{s=0}^{\infty}$ under the fixed $k > k_0$ gives a much better approximation to $\psi(x, x^*)$ than $\{\varphi(x, k)\}_{k=k_0}^{\infty}$ and the sequence $\{x^s, u^s\}_{s=1}^{\infty}$ converges to (x^*, u^*) much faster than $\{x(k), u(k)\}_{k=k_0}^{\infty}$.

The difference between CBF trajectory $\{x(k), u(k)\} (k \to \infty)$ and MBF trajectory $\{x(t, k), u(t, k)\}$ $(t = (u - u^*)/k \to 0)$ leads to the principal difference between IPM and NMBM trajectories. The IPM follows along the CBF trajectory turning from one "warm" start to another "warm" start by performing one Newton step and updating the penalty parameter. It allows us, in case of the LP, to improve the current approximation in $(1 - a/\sqrt{m})$ time by one Newton step where α is a universal constant (in [29] the corresponding result has been proved with $\alpha = 41^{-1}$). So to improve the current approximation twice one has to perform $O(\sqrt{m})$ Newton steps.

The NMBM follow along the MBF trajectory turning from one "hot" start to another "hot" start. In the case of the LP, to find the approximation for $\hat{x}(u, k)$ with accuracy 2^{-L}, one has to perform $O(\log_2 m)$ Newton steps. Therefore, to improve the current approximation twice ($\gamma_k \leq \frac{1}{2}$), by following the NMBM trajectory, one has to perform, in the worst case, $O(\log_2 m)$ Newton steps. In the case of the nonlinear programming problem, starting from the "hot" start $\tilde{x}^s$ for the problem $\hat{x}(\tilde{u}^s, k) = \operatorname{argmin}\{F(x, \tilde{u}^s, k) \mid x \in \mathbb{R}^n\}$, one has to perform $O(\log_2 \log_2 \varepsilon^{-1})$ Newton Method steps to find an approximation $\tilde{x}^{s+1}$ for $\hat{x}(\tilde{u}^s, k)$ with accuracy $\varepsilon > 0$ and then to update the vector $\tilde{u}^s$, i.e., to compute $\tilde{u}^{s+1} = [\operatorname{diag}(kf_i(\tilde{x}^{s+1}) + 1)]\tilde{u}^s$. The approximation $\tilde{x}^{s+1}$ will be the "hot" start for the next problem $\hat{x}(\tilde{u}^{s+1}, k) = \operatorname{argmin}\{F(x, \tilde{u}^{s+1}, k) \mid x \in \mathbb{R}^n\}$ i.e., to improve again the current approximation twice $\gamma_k \leq \frac{1}{2}$, one has to perform $O(\log_2 \log_2 \varepsilon^{-1})$ Newton steps. The moment when the NMBM trajectory reaches the "hot" start is crucial for the NMBM complexity. This moment depends on the MBF properties in the solution of the primal and dual problems. For any nondegenerate constrained optimization problem with $f_i(x) \in C^2$, $i = 0, \ldots, m$, due to (P5), the function $F(x, u^*, k)$ is not only strongly convex in the neighborhood of x^* but keeps this property in the neighborhood of $\hat{x} = \hat{x}(u, k) = \hat{x}(\cdot)$ uniformly in $(u, k) \in D(u^*, k_0, \delta, \varepsilon)$. Moreover, the

$$\operatorname{cond} F''_{xx}(\hat{x}(\cdot), \cdot) = \operatorname{mineigval} F''_{xx}(x(\cdot), \cdot)(\operatorname{maxeigval} F''_{xx}(x(\cdot), \cdot))^{-1}$$

is stable for any fixed $k \geq k_0$ and can be estimated uniformly in $u \in U_k$. In other words, let fixed $k \geq k_0$ and

$$\mu = \operatorname{mineigval} F''_{xx}(x^*, u^*, k) = \operatorname{mineigval}(L''_{xx} + kf'^{\mathrm{T}}_{(r)} U^* f'_{(r)}),$$

$$M_k = \operatorname{maxeigval}(L''_{xx} + kf'^{\mathrm{T}}_{(r)} U^* f'_{(r)}),$$

then there exists $0 < \beta_0 < 1$ that

$$\operatorname{cond} F''_{xx}(\hat{x}(u, k), u, k) \geq \beta_0 \mu M_k^{-1} \quad \forall u \in U_k \tag{83}$$

i.e., $\operatorname{cond} F''_{xx}(\hat{x}(\cdot), \cdot)$ is stable for any fixed $k \geq k_0$.

The threshold k_0 is critical for the conditions of the MBF near the solution. Moreover, this parameter is responsible for the contractibility properties of the operator C_k as well as for the transformation of the nonconvex constrained optimization problem (1) into the sequence of strongly convex unconstrained optimization problems. The threshold k_0, which provides the contractibility properties of the operator C_k can be estimated due to (19)–(21) by the value $\pi = \|(\Phi'_{(k)})^{-1} R_0\|$ which, in turn, depends on λ, $\min u^* = \min\{u_i^* \mid i = 1, \ldots, r\}$, $\max u^* = \{\max u_i^* \mid i = 1, \ldots, r\}$, μ, M_k, $\mu_0 = \operatorname{mineigval} f'^{\mathrm{T}}_{(r)}(x^*) f'_{(r)}(x^*)$, $M_0 = \operatorname{maxeigval} f'^{\mathrm{T}}_{(r)}(x^*) f'_{(r)}(x^*)$, and $\sigma = \min\{f_i(x^*) \mid i = r+1, \ldots, m\}$. These parameters characterize the "measure" of the nondegeneracy of the constrained optimization problem.

So the "hot" start very much depends on the measure of nondegeneracy of the constrained optimization problem. Therefore it seems promising to combine the universal self-concordant (see [17]) properties of the CBF which guarantee the polynomial complexity bound of the IPM, with the MBF properties (P1)–(P5), which allow us to speed up the process at the final stage and to reduce from $O(\sqrt{m})$ to $O(\log_2 m)$ the number of Newton steps, which in the worst case, have to be performed to improve the current approximation by a given amount.

Finally, note that in case of nondegenerate dual pair of LP the normal systems of equations (60), (81), which one has to solve at every step of the NMBM, are numerically much more stable than the corresponding systems in the IPM, which are based on the CBF (see [10, 12, 29]).

Acknowledgments

It is my pleasure to express my gratitude to the referees and to Drs. D. Jensen and A. King for their remarks, which contributed to the improvement of this paper. I would like to express my deepest appreciation to Peggy Cargiulo for her excellent work in preparing the manuscript for publication.

References

[1] K. Arrow, L. Hurwicz, H. Uzawa, *Studies on Linear and Nonlinear Programming* (Stanford University Press, Stanford, CA, 1958).

[2] D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic Press, New York, 1982).

[3] C.W. Carroll, "The created response surface technique for optimizing nonlinear restrained systems," *Operations Research* 9(2) (1961) 169–184.

[4] J.E. Dennis Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, NY, 1983).

[5] I.I. Dikin, "Iterative solution of problems of linear and quadratic programming," *Doklady Academii Nauk SSSR* 174 (1967) 747–748.

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

[7] R. Fletcher, "Recent developments in linear and quadratic programming," Report NA/94, Numerical Analysis, University of Dundee (Dundee, UK, 1986).

[8] K.R. Firsch, "The logarithmic potential method of convex programming," Memorandum, University Institute of Economics, Oslo (Oslo, 1955).

[9] P.E. Gill, W. Murray and M.H. Wright, *Practical Optimization* (Academic Press, London, 1981).

[10] P.E. Gill, M. Murray, M.A. Saunders, J.A. Tomlin and M.H. Wright, "On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method," *Mathematical Programming* 36 (1986) 183–209.

[11] E.G. Gol'shtein and N.V. Tret'yakov, *Modified Lagrangian Functions (Theory and Methods Optimization)* (Nauka, Moscow, 1989).

[12] C.C. Gonzaga, "An algorithm for solving Linear Programming problems in $O(n^3L)$ operations," to appear in: N. Megiddo, ed., *Research Issues in Linear Programming, Proceedings of the Asilomar Conference* (Springer, New York, 1988).

[13] K. Grossman and A.A. Kaplan, *Nonlinear Programming Based on Unconstrained Minimization* (Nauka, Novosibirsk, 1981). [In Russian.]

[14] M.R. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications* 4(5) (1969) 303–320.

[15] N.A. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica* 4 (1984) 373–395.

[16] K.A. McShane, C.L. Monma and D.F. Shanno, "An implementation of a primal-dual interior point method for Linear Programming," *ORSA Journal on Computing* 1 (1989) 70–89.

[17] Ju.E. Nesterov and A.S. Nemirovsky, *Self-concordant Functions and Polynomial-Time Methods in Convex Programming* (CEMI Academy of Sciences, Moscow, 1989).

[18] E. Polak, *Computational Methods im Optimization: A Unified Approach* (Academic Press, New York, 1971).

[19] B.T. Polyak, *Introduction to Optimization* (Nauka, Moscow, 1983). [In Russian.]

[20] B.T. Polyak and N.V. Tret'yakov, "The method of penalty estimates for conditional extremum problems," *Computational Mathematics and Mathematical Physics* 13 (1974) 42–58.

[21] R. Polyak, "Smooth optimization methods for solving nonlinear extremal and equilibrium problems with constraints," Abstracts of the papers of *The 11th International Symposium on Mathematical Programming* (Bonn, 1982).

[22] R. Polyak, "Classical Lagrangians with the properties of the modified ones and estimation of complexity for linear and quadratic programming," Abstracts of the papers of *The 12th International Symposium on Mathematical Programming* (Boston, 1985).

[23] R. Polyak, *Controlled Processes in Extremal and Equilibrium Problems* (VINITI, Moscow, 1986), Deposited manuscript. [In Russian.]

[24] R. Polyak, "Smooth optimization methods for minimax problems," *SIAM Journal on Control and Optimization* 26(6) (1988) 1274–1286.

[25] R. Polyak, "Modified barrier functions," IBM Research Report RC14602, IBM T.J. Watson Research Center (Yorktown Heights, NY, 1989).

[26] R. Polyak, "The nonlinear rescaling principle in linear programming," IBM Research Report RC15030, IBM T.J. Watson Research Center (Yorktown Heights, NY, 1989).

[27] M.J.D. Powell, "A method for nonlinear constraints in minimization problems," In: R. Fletcher, ed. *Optimization* (Academic Press, New York, 1969) pp. 283–298.

[28] J. Renegar, "A polynomial-time algorithm, based on Newton's method, for linear programming," *Mathematical Programming* 40 (1988) 59–93.

[29] J. Renegar and M. Shub, "Simplified complexity analysis for Newton LP methods," Technical Report No. 807, School of Operation Research and Industrial Engineering, College of Engineering, Cornell University (Ithaca, NY, 1988).

[30] R.T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).

[31] R.T. Rockafellar, "Augmented Lagrange multiplier functions and duality in nonconvex programming," *SIAM Journal on Control and Optimization* 12(2) (1974) 268–285.

[32] S. Smale, "Newton's method estimates from data at one point," in: R.E. Ewing et al., eds. *The Merging of Disciplines in Pure, Applied, and Computational Mathematics* (Springer, New York–Berlin, 1986) pp. 185–196.

[33] S. Smale, "Algorithms for solving equations," in: *Proceedings of the International Congress of Mathematicians* (Berkeley, CA, 1986) pp. 172–195.

[34] M.J. Todd, "Recent developments and new directions in linear programming," Technical Report No. 827, School of Operation Research and Industrial Engineering, Cornell University (Ithaca, NY, 1988).