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

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

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 $U =$ 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 $k>0$, $\Omega=\Omega_F=\{x\colon k^{-1}\ln(kf_i(x)+1)\geqslant 0,\ i=1,\dots,m\}=\Omega_C=\{x\colon k^{-1}[kf_i(x)+1)^{-1}-1]\leqslant 0,\ i=1,\dots,m\}$, 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 $u\geqslant 0$ 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 $f_i(x)$, $i=1,\dots,m$, 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 $u^*=(u_1^*,\dots,u_m^*)$ 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 $u_i^*>0$, $i=1,\dots,r$, 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) $(F''_{xx}(x^*, u^*, k)y, y)\geqslant\mu(y, y)\quad\forall y\in\mathbb{R}^n,\ \forall k\geqslant k_0,$

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 $k\geqslant k_0$.

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 $k\geqslant k_0$. 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 $U(\varepsilon) = \{u \in \mathbb{R}^m_+ : u_i \geq \varepsilon,\ i = 1, \ldots, r,\ u_i \geq 0,\ i = r+1, \ldots, n\}$. 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 $S(y, \varepsilon) = \{x \in \mathbb{R}^n : \|x - y\| \leq \varepsilon\}$.

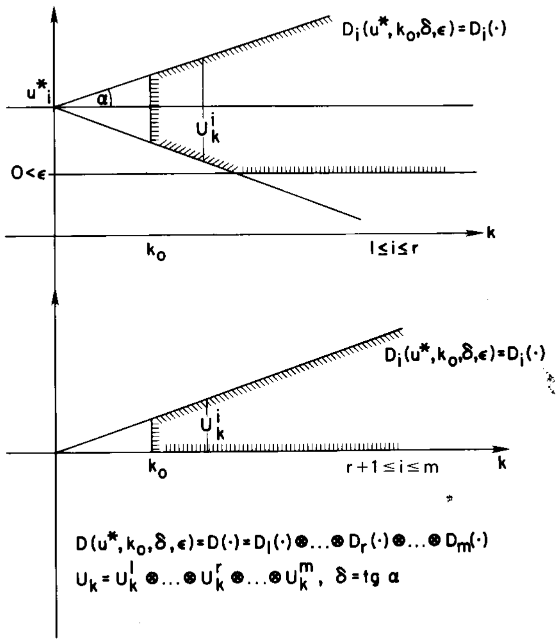

Fig. 1.

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

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

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

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 $\|\Phi_{(\infty)}^{-1}\|\leqslant\lambda_0$.

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

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

and

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

Rewriting (11) we obtain

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

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

and let

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

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

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

So we have

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

and

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

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

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

Then

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

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

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

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

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

Let

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

Then combining (15), (17) we obtain

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

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 $f_{(m-r)}(x(\cdot))=(f_i(x(\cdot)),\ i=r+1,\dots,m)$, $\hat{u}_{(m-r)}(\cdot)=(\hat{u}_i(x(\cdot),\cdot),\ i=r+1,\dots,m)$,

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

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

Now we consider the system (18) for $t=0$ and $k>k_0$. First of all note that $x(0,k)=x^*$, $\hat{u}_{(r)}(0,k)=u_{(r)}^*=(u_1^*,\ldots u_r^*)>0$ and also $\hat{u}_i(x(0,k),0,k)=0$, $i=r+1,\ldots,m$, $f_i(x(0,k))=f_i(x^*)\geq\sigma>0$, $i=r+1,\ldots,m$, $D_r(0,k)=D_r^2(0,k)=I^r$, $D(t_{(m-r)})|_{t_{(m-r)}=0}=O^{m-r,m-r}$, $D(f_{(m-r)}(x^*))=[\text{diag}(f_i(x^*))]_{i=r+1}^m\geq\sigma I^{m-r}$.

Further,

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

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 $\|\Phi_{(k)}^{-1}\|\leq\rho$ 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

$$\|\Phi'^{-1}(x(\tau t,k),\hat{u}_{(r)}(\tau t,k))R(x(\tau t,k);(\tau t,k))\|\leq2\rho[\sigma^{-1}\|f'_{(m-r)}(x^*)\|+1]=c_0 \tag{20}$$

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

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

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

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

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 $\hat{x} = \hat{x}(u,k) = \operatorname{argmin}\{F(x,u,k) \mid x \in \mathbb{R}^n\}$ 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 $\hat{x}(u^*,k) = \operatorname{argmin}\{F(x,u^*,k) \mid x \in \mathbb{R}^n\} = x^*$ 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 $(u,k) \in D(u^*, k_0, \delta, \varepsilon)$.