On Stationary Points of the Mangasarian and Solodov's Function for Nonlinear Complementarity Problems

Nobuo Yamashita and Masao Fukushima

Graduate School of Information Science
Nara Institute of Science and Technology
Ikoma, Nara 630-01, Japan

August 17, 1993
Revised December 15, 1993

Abstract. Mangasarian and Solodov have recently introduced an unconstrained optimization problem whose global minima are solutions of the nonlinear complementarity problem (NCP). In this paper, we show that if the mapping involved in NCP has a positive definite Jacobian, then any stationary point of the optimization problem actually solves NCP. We also discuss a descent method for solving the unconstrained optimization problem.

Key Words. Nonlinear complementarity problem, implicit Lagrangian, equivalent differentiable optimization problem, stationary point, descent method.

1 Introduction

The nonlinear complementarity problem [1, 2] is to find an $x \in R^n$ such that

$$[\text{NCP}] \qquad x^T F(x) = 0,\ x \geq 0,\ F(x) \geq 0,$$

where $F : R^n \to R^n$. Recently Mangasarian and Solodov [3] introduced the following implicit Lagrangian function to construct an optimization problem equivalent to NCP:

$$\begin{aligned} M_\alpha(x) = \quad & x^T F(x) + \frac{1}{2\alpha}\Big(\|(-\alpha F(x) + x)_+\|^2 - \|x\|^2 \\ & + \|(-\alpha x + F(x))_+\|^2 - \|F(x)\|^2\Big), \end{aligned} \tag{1}$$

where $\alpha > 0$ is a real parameter and $(z)_+$ denotes the vector with components $\max\{0, z_i\}$, $i = 1, \ldots, n$. Mangasarian and Solodov [3] in particular established the following property of $M_\alpha(x)$.

Theorem 1.1 *[3] For $\alpha \in (1, \infty)$, $M_\alpha(x)$ is nonnegative on R^n. Moreover, for $\alpha \in (1, \infty)$, $M_\alpha(x) = 0$ if and only if x solves NCP.*

This theorem reveals that the set of global minima of M_α such that $M_\alpha(x) = 0$ coincides with that of solutions of NCP. Moreover, since M_α is differentiable whenever F is so, the next result is an immediate consequence of Theorem1.1.

Corollary 1.1 *[3] If F is differentiable at a solution $\bar{x}$ of NCP, then $\nabla M_\alpha(\bar{x}) = 0$ for $\alpha \in (1, \infty)$.*

The above results suggest that NCP may be solved by reformulating it as the unconstrained optimization problem

$$\min_{x \in R^n} M_\alpha(x).$$

However since it is not necessarily easy to find a global minimum of M_α, it is worth investigating under what conditions the converse of Corollary 1.1 holds, namely, under what conditions a stationary point of M_α is a solution of NCP. This paper gives an answer to this question, which was raised by Mangasarian and Solodov [3].

2 Main Results

By the definition of $(z)_+$, we have the following lemma. The proof is easy and omitted here.

Lemma 2.1 *The following three conditions are equivalent :*

(a) $F(x) \geq 0$, $x \geq 0$, $x^T F(x) = 0$;

(b) $x = (-\alpha F(x) + x)_+$;

(c) $F(x) = (-\alpha x + F(x))_+$.

It will be convenient to define the mappings $G : R^n \to R^n$ and $H : R^n \to R^n$ by

$$G(x) = (-\alpha F(x) + x)_+ - x$$

and

$$H(x) = (-\alpha x + F(x))_+ - F(x),$$

respectively. By Lemma 2.1, x is a solution of NCP if and only if $G(x) = 0$ and $H(x) = 0$. By definition, the mappings G and H may be represented component-wise as

$$G_i(x) = \begin{cases} -\alpha F_i(x), & \text{if } -\alpha F_i(x) + x_i \geq 0 \\ -x_i, & \text{if } -\alpha F_i(x) + x_i < 0 \end{cases}$$

and

$$H_i(x) = \begin{cases} -\alpha x_i, & \text{if } -\alpha x_i + F_i(x) \geq 0 \\ -F_i(x), & \text{if } -\alpha x_i + F_i(x) < 0. \end{cases}$$

Lemma 2.2 *For any* $\alpha > 0$ *and* $x \in R^n$, *we have*

$$(-\alpha G(x) + H(x))^T (G(x) - \alpha H(x)) \geq 0.$$

Proof. We consider the following four cases:

Case 1: $-\alpha F_i(x) + x_i \geq 0$ and $-\alpha x_i + F_i(x) \geq 0$.

We have

$$\begin{aligned} -\alpha G_i(x) + H_i(x) &= \alpha^2 F_i(x) - \alpha x_i \\ &= \alpha(\alpha F_i(x) - x_i) \leq 0 \end{aligned}$$

and

$$\begin{aligned} G_i(x) - \alpha H_i(x) &= -\alpha F_i(x) + \alpha^2 x_i \\ &= -\alpha(F_i(x) - \alpha x_i) \leq 0. \end{aligned}$$

Case 2: $-\alpha F_i(x) + x_i < 0$ and $-\alpha x_i + F_i(x) \geq 0$.

We have

$$-\alpha G_i(x) + H_i(x) = -\alpha(-x_i) - \alpha x_i = 0.$$

Case 3: $-\alpha F_i(x) + x_i \geq 0$ and $-\alpha x_i + F_i(x) < 0$.

We have

$$G_i(x) - \alpha H_i(x) = -\alpha F_i(x) + \alpha F_i(x) = 0.$$

Case 4: $-\alpha F_i(x) + x_i < 0$ and $-\alpha x_i + F_i(x) < 0$.

We have

$$-\alpha G_i(x) + H_i(x) = \alpha x_i - F_i(x) > 0$$

and

$$G_i(x) - \alpha H_i(x) = -x_i + \alpha F_i(x) > 0.$$

Thus, we obtain

$$(-\alpha G_i(x) + H_i(x))(G_i(x) - \alpha H_i(x)) \geq 0$$

for each i, and hence,

$$(-\alpha G(x) + H(x))^T (G(x) - \alpha H(x)) \geq 0.$$

The proof is complete. □

By direct calculation, it follows from (1) that

$$\begin{aligned} \alpha \nabla M_\alpha(x) &= \alpha(F(x) + \nabla F(x)^T x) + (-\alpha \nabla F(x)^T + I)(-\alpha F(x) + x)_+ - x \\ &\quad + (-\alpha I + \nabla F(x)^T)(-\alpha x + F(x))_+ - \nabla F(x)^T F(x) \\ &= \nabla F(x)^T \left(-\alpha((-\alpha F(x) + x)_+ - x) + (-\alpha x + F(x))_+ - F(x)\right) \\ &\quad + (-\alpha F(x) + x)_+ - x - \alpha((-\alpha x + F(x))_+ - F(x)) \\ &= \nabla F(x)^T(-\alpha G(x) + H(x)) + G(x) - \alpha H(x). \end{aligned} \tag{2}$$

Lemma 2.3 *Let $\bar{x}$ be a stationary point of M_α and $\alpha \neq 1$. If F is differentiable at $\bar{x}$ and the matrix $\nabla F(\bar{x})$ is nonsingular, then the following three conditions are equivalent:*

(a) $-\alpha G(\bar{x}) + H(\bar{x}) = 0$;

(b) $G(\bar{x}) - \alpha H(\bar{x}) = 0$;

(c) $G(\bar{x}) = 0, H(\bar{x}) = 0$.

Proof. By assumption, $\nabla M_\alpha(\bar{x}) = 0$ and the matrix $\nabla F(\bar{x})$ is nonsingular. We first show the equivalence between (a) and (b). If (a) holds, then it follows immediately from (2) that (b) is satisfied. Conversely, if (b) holds, then (2) yields

$$\nabla F(\bar{x})^T(-\alpha G(\bar{x}) + H(\bar{x})) = 0,$$

which together with the nonsingularity of $\nabla F(\bar{x})$ implies (a). Since (c) obviously implies (a) and (b), it remains to prove that (c) is implied by (a) or (b). Suppose (a) holds. Then (b) also holds as shown above. Moreover, since $\alpha \neq 1$, (a) and (b) are simultaneously satisfied only if $G(\bar{x}) = 0$ and $H(\bar{x}) = 0$. This completes the proof. □

We are now ready to state our main result.

Theorem 2.2 *If F is differentiable and its Jacobian $\nabla F(x)$ is positive definite at a stationary point $\bar{x}$ of M_α, then $\bar{x}$ is a solution of NCP for any $\alpha > 1$.*

Proof. Let $\bar{x}$ be a stationary point of M_α. Premultiplying (2) by $(-\alpha G(\bar{x}) + H(\bar{x}))^T$, we obtain

$$\begin{aligned}(-\alpha G(\bar{x}) + H(\bar{x}))^T \nabla F(\bar{x})^T(-\alpha G(\bar{x}) + H(\bar{x})) & \\ +(-\alpha G(\bar{x}) + H(\bar{x}))^T(G(\bar{x}) - \alpha H(\bar{x})) \quad &= 0.\end{aligned}$$

By the positive-definiteness of $\nabla F(\bar{x})$, the first term is nonnegative, while by Lemma 2.2, the second term is also nonnegative. Thus, the both terms have to vanish simultaneously. Since $\nabla F(\bar{x})$ is positive definite, it then follows that

$$-\alpha G(\bar{x}) + H(\bar{x}) = 0.$$

Hence, by Lemma 2.3, we obtain $G(\bar{x}) = H(\bar{x}) = 0$. Consequently, Lemma 2.1 and the definitions of G and H imply that $\bar{x}$ solves NCP. □

From this theorem, the next result follows immediately.

Theorem 2.3 *If F is differentiable and its Jacobian $\nabla F(x)$ is positive definite everywhere, then any stationary point of M_α is a solution of NCP for any $\alpha > 1$.*

Recall [4, Theorem 5.4.3] that, when F is continuously differentiable and $\nabla F(x)$ is positive definite for all x, F is strictly monotone, i.e.,

$$(F(x) - F(y))^T(x - y) > 0, \qquad \forall x, y \in R^n \quad (x \neq y).$$

But, the converse of this fact is not true in general. It may be worth noting that only strict monotonicity assumption on F cannot replace the positive-definiteness assumption on ∇F in Theorem 2.2. In fact, consider the simple example with $n = 1$, where $F(x) = (x-3)^3+1$ for all $x \in R$. Note that F is strictly monotone but $\nabla F(x)$ fails to be positive definite at $x = 3$. This NCP has the unique solution $x = 2$. But, for $\alpha = 2$, $\nabla M_\alpha(x) = 0$ is also satisfied by $x = 3$, which is not a solution of NCP.

In practice, positive definiteness of the Jacobian $\nabla F(x)$ may appear to be a restrictive condition. For example, consider a linear complementality problem (LCP) derived from a quadratic programming problem (QP). If QP is convex, LCP is monotone [5]. However, even if QP is strictly convex, LCP cannot be strictly monotone. Therefore a stationary point of M_α may not be a solution of LCP. A possible remedy is to apply the proximal point algorithm [6] to LCP, thereby constructing strictly monotone subproblems.

Next, we investigate an asymptotic behavior of M_α as $\|x\| \to \infty$. To this end, let

$$M_\alpha(x) = \sum_{i=1}^{n} M_{\alpha,i}(x),$$

where

$$M_{\alpha,i}(x) = x_i F_i(x) + \frac{1}{2\alpha}\left((x_i - \alpha F_i(x))_+^2 - x_i^2 + (F_i(x) - \alpha x_i)_+^2 - F_i(x)^2\right).$$

Note that $M_{\alpha,i}(x)$ is nonnegative at any x for all i. (See the proof of Theorem 2.1 in [3].)

Theorem 2.4 *If F is strongly monotone with modulus $\mu > 0$, i.e.,*

$$(F(x) - F(y))^T(x - y) \geq \mu\|x - y\|^2, \qquad \forall x, y \in R^n,$$

and if $\alpha > \max\{1, 1/\mu\}$, *then*

$$\lim_{\|x\|\to\infty} M_\alpha(x) = +\infty.$$

Proof. Since $M_{\alpha,i}$ is nonnegative for all i, it is sufficient to show that, for some i, $M_{\alpha,i}(x^k) \to \infty$, where $x^k = t_k p$ with p being an arbitrary nonzero vector and t_k being any positive sequence such that $t_k \to \infty$. It follows from the strongly monotonicity of F that

$$(F(x^k) - F(0))^T x^k \geq \mu \|x^k\|^2.$$

Hence, we have

$$\sum_{i\in J} x_i^k F_i(x^k) \geq \sum_{i\in J} \left(\mu (x_i^k)^2 + F_i(0) x_i^k\right), \quad k = 1, 2, \ldots,$$

where $J = \{i | p_i \neq 0\}$. Thus, for each k there exists $i_k \in J$ such that

$$t_k p_{i_k} F_{i_k}(t_k p) \geq \mu t_k^2 p_{i_k}^2 + t_k p_{i_k} F_{i_k}(0).$$

Since the number of elements in J is finite, we may assume without loss of generality that $i_k = i$, $k = 1, 2, \ldots$, for some fixed $i \in J$. That is,

$$t_k p_i F_i(t_k p) \geq \mu t_k^2 p_i^2 + t_k p_i F_i(0) \tag{3}$$

for all k. Since $t_k > 0$,

$$p_i F_i(t_k p) \geq \mu t_k p_i^2 + p_i F_i(0). \tag{4}$$

Since the right-hand side of (4) is positive when t_k is sufficiently large, squaring the both sides and dividing them by p_i^2 yield

$$F_i(t_k p)^2 \geq (\mu t_k p_i + F(0))^2. \tag{5}$$

Suppose that t_k is sufficiently large and consider the following four cases. Remember that $\alpha > 1$ and $\alpha\mu > 1$ by assumption.

Case 1: $-\alpha F_i(x^k) + x_i^k \geq 0$ and $-\alpha x_i^k + F_i(x^k) \geq 0$.

We have

$$\begin{aligned} 2\alpha M_{\alpha,i}(x^k) &= (\alpha F_i(x^k) - \alpha x_i^k)^2 + 2\alpha(\alpha-1)x_i^k F_i(x^k) \\ &\geq 2\alpha(\alpha-1)x_i^k F_i(x^k) \\ &\geq 2\alpha(\alpha-1)t_k(\mu t_k p_i^2 + p_i F_i(0)), \end{aligned}$$

where the last inequality follows from (4).

Case 2: $-\alpha F_i(x^k) + x_i^k < 0$ and $-\alpha x_i^k + F_i(x^k) \geq 0$.

We have

$$\begin{aligned} 2\alpha M_{\alpha,i}(x^k) &= (\alpha^2-1)F_i(x^k)^2 \\ &\geq (\alpha^2-1)(\mu t_k p_i + F_i(0))^2, \end{aligned}$$

where the inequality follows from (5).

Case 3: $-\alpha F_i(x^k) + x_i^k \geq 0$ and $-\alpha x_i^k + F_i(x^k) < 0$.

We have

$$\begin{aligned} 2\alpha M_{\alpha,i}(x^k) &= (\alpha^2-1)(x_i^k)^2 \\ &= (\alpha^2-1)t_k^2 p_i^2. \end{aligned}$$

Case 4: $-\alpha F_i(x^k) + x_i^k < 0$ and $-\alpha x_i^k + F_i(x^k) < 0$.

In this case, it is easy to see $F_i(x^k) > 0$. Thus,

$$F_i(x^k)(-F_i(x^k) + \alpha x^k) > 0. \tag{6}$$

Hence, we have

$$\begin{aligned} 2\alpha M_{\alpha,i}(x^k) &= 2\alpha x_i^k F_i(x^k) - (x_i^k)^2 - F_i(x^k)^2 \\ &\geq \alpha x_i^k F_i(x^k) - (x_i^k)^2 \\ &\geq \alpha\mu(x_i^k)^2 + \alpha F_i(0)x_i^k - (x_i^k)^2 \\ &= t_k\left((\alpha\mu-1)t_k p_i^2 + \alpha p_i F_i(0)\right), \end{aligned}$$

where the first and second inequalities follow from (6) and (3), respectively.

Hence, in any case we have

$$\lim_{k\to\infty} M_{\alpha,i}(x^k) = +\infty.$$

The proof is complete. □

By this theorem, when F is strongly monotone with modulus $\mu > 0$, the level sets of M_α are compact whenever $\alpha > \max\{1, 1/\mu\}$.

3 Descent method

In this section, we discuss an iterative method that uses the vector

$$d = \alpha G(x) - H(x)$$

as a search direction. Since the vector d does not involve the Jacobian matrix $\nabla F(x)$, d may be much easier to evaluate than the gradient $\nabla M_\alpha(x)$. The next lemma shows that, under monotonicity assumption on F, the vector d is actually a descent direction of the function M_α at x.

Lemma 3.1 *Let F be continuously differentiable. If the Jacobian matrix $\nabla F(x)$ is positive definite, then for each x the vector d satisfies the discent condition*

$$d^T \nabla M_\alpha(x) < 0,$$

whenever $d \neq 0$. Moreover, if F is strongly monotone with modulus $\mu > 0$, then for each x, the vector d satisfies

$$d^T \nabla M_\alpha(x) \leq -\mu \|d\|^2.$$

Proof. From (2) and the proof of Theorem 2.2, we have

$$\begin{aligned} d^T \nabla M_\alpha(x) &= d^T \nabla F(x)^T (-\alpha G(x) + H(x)) + d^T (G(x) - \alpha H(x)) \\ &= -d^T \nabla F(x)^T d - (-\alpha G(x) + H(x))^T (G(x) - \alpha H(x))) \\ &\leq -d^T \nabla F(x)^T d, \end{aligned}$$

where the inequality follows from Lemma 2.2. The desired results are then consequences of the well known properties of the Jacobian of (strongly) monotone mapping [4, Theorem 5.4.3]. □

Now we establish a convergence result for an iterative method that uses the search direction $d = \alpha G(x) - H(x)$. Here we adopt the following Armijo–type steplength rule: Let β and γ be given constants such that $\beta > 0$ and $0 < \gamma < 1$. For the current iterate x^k and the search direction d^k, determine the steplength t_k as $t_k = \gamma^{l_k}$, where l_k is the smallest nonnegative integer l such that

$$M_\alpha(x^k + \gamma^l d^k) \leq M_\alpha(x^k) - \beta\gamma^l \|d^k\|^2. \tag{7}$$

Theorem 3.5 *Suppose that the mapping F is differentiable and strongly monotone with modulus $\mu > 0$. Suppose also that $F(x)$ and $\nabla F(x)$ are Lipschitz continuous. Let $\{x^k\}$ be a sequence generated by the iteration*

$$x^{k+1} = x^k + t_k d^k, \quad k = 0, 1, 2, \ldots,$$

where d^k are given by $d^k = \alpha G(x^k) - H(x^k)$ and $t_k \in [0, 1]$ are determined by the Armijo-type rule (7). Then the sequence $\{x^k\}$ converges to the unique solution of NCP for any starting point $x^0 \in R^n$, if the parameter α is chosen large enough to satisfy $\alpha > \max\{1, 1/\mu\}$ and the positive constant β in (7) is chosen sufficiently small so that $\beta < \mu$.

Proof. Since $F(x)$ and $\nabla F(x)$ are Lipschitz continuous, $\nabla M_\alpha(x)$ is also Lipschitz continuous, i.e., there exists a constant $L > 0$ such that

$$\|\nabla M_\alpha(x) - \nabla M_\alpha(x')\| \leq L\|x - x'\| \quad \text{for all } x, x' \in R^n. \tag{8}$$

It follows from Lemma 3.1 and (8) that, for $0 \leq t \leq 1$,

$$\begin{aligned} M_\alpha(x^k + td^k) - M_\alpha(x^k) &= \int_0^t \nabla M_\alpha(x^k + sd^k)^T d^k ds \\ &= t\nabla M_\alpha(x^k)^T d^k + \int_0^t (\nabla M_\alpha(x^k + sd^k) - \nabla M_\alpha(x^k))^T d^k ds \\ &\leq -\mu t\|d^k\|^2 + \int_0^t Ls\|d^k\|^2 ds \\ &= \left(-\mu t + \frac{1}{2}Lt^2\right)\|d^k\|^2. \end{aligned}$$

Then it is not difficult to see that the inequality

$$M_\alpha(x^k + td^k) \leq M_\alpha(x^k) - \beta t \|d^k\|^2$$

holds for all t satisfying $0 \leq t \leq \min\{1, 2(\mu - \beta)/L\}$, provided that $\beta < \mu$. Consequently, from the manner in which the steplength t_k is determined, it follows that $t_k = \gamma^{l_k}$ satisfies

$$t_k > \min\{\gamma, 2\gamma(\mu - \beta)/L\} \tag{9}$$

for all k. On the other hand, (7) implies

$$M_\alpha(x^{k+1}) \leq M_\alpha(x^k) - \beta t_k \|d^k\|^2, \tag{10}$$

and hence $\{M_\alpha(x^k)\}$ is monotonically decreasing. Since $\{M_\alpha(x^k)\}$ is bound from below, it follows from (9) and (10) that

$$\lim_{k \to \infty} \|d^k\| = 0. \tag{11}$$

Moreover, since $\{M_\alpha(x^k)\}$ is decreasing, the sequence $\{x^k\}$ is contained in the level set $\{x | M_\alpha(x) \leq M_\alpha(x^0)\}$, which is compact by Theorem 2.4. Therefore, $\{x^k\}$ has at least one accumulation point. Moreover, since $d^k = \alpha G(x^k) - H(x^k)$, (11) implies that any such accumulation point satisfies $-\alpha G(x) + H(x) = 0$. Thus it follows from Lemma 2.3 that the accumulation point of $\{x^k\}$ solves NCP. Since the strong monotonicity of F guarantees that the solution of NCP is unique, we can conclude that the entire sequence $\{x^k\}$ converges to the solution of NCP. $\square$

4 Conclusion

We have shown that if F is differentiable and its Jacobian $\nabla F(x)$ is positive definite, then any stationary point of M_α is a solution of NCP. Therefore when F satisfies these conditions, a solution of NCP may be obtained by solving the equivalent optimization problem using any algorithm guaranteed to converge to a stationary point of the objective function. We have introduced a globally convergent iterative method using the merit

function M_α. Kanzow [7] recently introduced a class of functions involving the function M_α as a special case. It may be interesting to consider how the results obtained in this paper can be generalized to such functions.

Acknowledgement

We would like to thank a referee for a helpful suggestion that led us to develop the descent method described in Section 3.

References

[1] Harker, P.T. and Pang, J.-S., *Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problem: A Survey of Theory, Algorithms and Applications*, Mathematical Programming, Vol. 48, pp. 161–220, 1990.

[2] Pang, J.-S., *Complementarity Problems*, Handbook of Global Optimization, Edited by R. Horst and P. Pardalos, to appear.

[3] Mangasarian, O.L. and Solodov, M.V., *Nonlinear Complementarity as Unconstrained and Constrained Minimization*, Mathematical Programming, Vol. 62, pp. 277–297, 1993.

[4] Ortega, J.M. and Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, N.Y., 1970.

[5] Cottle, R.W., Pang, J.-S. and Stone, R.E., *The Linear Complementerity Problem*, Academic Press, New York, N.Y., 1992.

[6] Rockafellar, R.T., *Monotone Operators and the Proximal Point Algorithm*, SIAM J. Control and Optimization, Vol. 14, pp. 877–898, 1976.

[7] Kanzow, C., *Nonlinear Complementarity as Unconstrained Optimization*, Preprint 67, Institut für Angewandte Mathematik, Universität Hamburg, Hamburg, Germany, 1993.