ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439

SMOOTHING OF MIXED COMPLEMENTARITY PROBLEMS

Steven A. Gabriel and Jorge J. Moré

Mathematics and Computer Science Division

Preprint MCS-P541-0995

September 1995

This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

SMOOTHING OF MIXED COMPLEMENTARITY PROBLEMS

Steven A. Gabriel and Jorge J. Moré

Abstract

We extend the smoothing approach to the mixed complementarity problem, and study the limiting behavior of a path defined by approximate minimizers of a nonlinear least squares problem. Our main result guarantees that, under a mild regularity condition, limit points of the iterates are solutions to the mixed complementarity problem. The analysis is applicable to a wide variety of algorithms suitable for large-scale mixed complementarity problems.

1 Introduction

A central problem in complementarity is the development of algorithms for the solution of *mixed* complementarity problems. Given a mapping $f : \mathbb{R}^n \mapsto \mathbb{R}^n$, and bounds l and u in $\mathbb{R}^n$, with $l < u$, a solution of the mixed nonlinear complementarity problem is a vector $x \in \mathbb{R}^n$ such that

$$
\begin{aligned}
f_i(x) = 0 \quad & \text{if} \quad x_i \in [l_i, u_i], \\
f_i(x) \geq 0 \quad & \text{if} \quad x_i = l_i, \\
f_i(x) \leq 0 \quad & \text{if} \quad x_i = u_i.
\end{aligned}
\tag{1.1}
$$

A wide variety of algorithms have been proposed for the solution of the mixed complementarity problem, although most algorithms are only applicable to the *classical* nonlinear complementarity problem of finding $x \in \mathbb{R}^n$ such that

$$
x \geq 0, \qquad f(x) \geq 0, \qquad x^T f(x) = 0. \tag{1.2}
$$

Clearly, the mixed complementarity problem (1.1) reduces to the classical problem (1.2) when $l \equiv 0$ and $u \equiv +\infty$.

We consider a recent approach to the solution of the classical problem (1.2), introduced by Chen and Mangasarian [6], that is based on smoothing a residual $r : \mathbb{R}^n \mapsto \mathbb{R}^n$ associated with the complementarity problem. This approach has attracted attention because the numerical results of Billups, Dirkse, and Ferris [2] show that the smooth code, which is based on these ideas, has superior performance.

We extend the smoothing approach to the mixed complementarity problem, and study the limiting behavior of a path $x(\cdot)$, where $x(\lambda)$ is an approximate solution to the least squares problem

$$
\min \left\{ \tfrac{1}{2} \| r_\lambda(x) \|^2 : x \in \mathbb{R}^n \right\}. \tag{1.3}
$$

This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

If $\{\lambda_k\}$ is a sequence converging to zero and $\{x_k\}$ is an approximate minimizer of (1.3) with $\lambda = \lambda_k$, we want to guarantee that any limit point of $\{x_k\}$ solves the mixed complementarity problem (1.1). In related work, Chen and Harker [5] considered a path where $r_\lambda[x(\lambda)] = 0$. This approach requires stronger assumptions since we need to guarantee the existence and uniqueness of the solutions of $r_\lambda(x) = 0$. On the other hand, (1.3) always has approximate minimizers.

We introduce the smoothing approach in Section 2. We relate this approach to the smoothed functionals introduced by Katkovnik (see, for example, Katkovnik and Kulchitskii [12]) to approximate n-dimensional nondifferentiable functions in stochastic optimization. References and generalizations of this work can be found in the work of Rubinstein [19, 20, 21]. Smoothing is also used in global optimization, but with the purpose of removing local minimizers of n-dimensional smooth functions. See Moré and Wu [16] for information and additional references for this work.

Section 3 defines the smooth residual for the mixed complementarity problem (1.1), while Section 4 studies the limiting behavior of a sequence $\{x_k\}$, where x_k is an approximate minimizer of (1.3) with $\lambda = \lambda_k$ and $\{\lambda_k\}$ converges to zero. We show that if a regularity condition is satisfied, then any limit point of $\{x_k\}$ solves the mixed complementarity problem (1.1). Chen and Mangasarian [6] were only able to show that the limit points of the sequence were ϵ-accurate solutions to the classical complementarity problem (1.2).

The notation that we use is fairly standard. Unless otherwise noted, $\|\cdot\|$ is the Euclidean norm. We use $A_{\mathcal{K}}$, or $A_{\mathcal{K},\mathcal{K}}$ for the principal submatrix of $A \in \mathbb{R}^{n\times n}$ with $(i,j) \in \mathcal{K} \times \mathcal{K}$, and $A_{\mathcal{I},\mathcal{J}}$ for the submatrix of A with $(i,j) \in \mathcal{I} \times \mathcal{J}$.

2 Smoothing Functions

We propose a class of smoothing functions that can be used to approximate the mixed nonlinear complementarity problem by a smooth system of nonlinear equations. We first develop the properties for the scalar-valued version of the smoothing function and then extend the results to the vector case in the next section.

The key to our development is a smooth approximation to the median function. We use the notation $\mathrm{mid}(\cdot)$ for the median of three numbers because in this case the median is the number in the middle.

Definition 2.1 *Let $\rho : \mathbb{R} \mapsto \mathbb{R}_+$ be a density function with bounded absolute mean, that is,*

$$\kappa = \int_{-\infty}^{+\infty} |s|\rho(s)\,ds < +\infty. \tag{2.1}$$

For any $\lambda > 0$ and constants $\alpha_l, \alpha_u \in \mathbb{R}$, with $\alpha_l < \alpha_u$, define the smooth approximation to the mid *function by*

$$p_\lambda(\alpha) = \int_{-\infty}^{+\infty} \mathrm{mid}(\alpha_l, \alpha_u, \alpha - \lambda s)\rho(s)\,ds,$$

where $\mathrm{mid}(\cdot)$ *is the median operator.*

Chen and Mangasarian [6], following the work of Kreimer and Rubinstein [13], introduced a smooth approximation to the standard complementarity problem (1.2); we will show that the above definition extends their work to the mixed complementarity problem.

Lemma 2.2 *If* $p_\lambda : \mathbb{R} \mapsto \mathbb{R}$ *is the smooth approximation to the median function, then*

$$\alpha_l \leq p_\lambda(\alpha) \leq \alpha_u, \tag{2.2}$$

and

$$\lim_{\alpha \to +\infty} p_\lambda(\alpha) = \alpha_u, \qquad \lim_{\alpha \to -\infty} p_\lambda(\alpha) = \alpha_l. \tag{2.3}$$

Proof. We prove (2.2) by noting that since $\alpha_l \leq \mathrm{mid}(\alpha_l, \alpha_u, \alpha) \leq \alpha_u$ and ρ is a density function,

$$\alpha_l \leq \int_{-\infty}^{+\infty} \alpha_l \rho(s)\, ds \leq \int_{-\infty}^{+\infty} \mathrm{mid}(\alpha_l, \alpha_u, \alpha - \lambda s)\rho(s)\, ds \leq \int_{-\infty}^{+\infty} \alpha_u \rho(s)\, ds = \alpha_u.$$

We prove only the first relationship in (2.3) since the proof of the other is similar. Note that

$$|\mathrm{mid}(\alpha_l, \alpha_u, \alpha - \lambda s)\rho(s)| \leq \max\{|\alpha_l|, |\alpha_u|\}\rho(s),$$

and thus Lebesgue's dominated convergence theorem implies that

$$\lim_{\alpha \to +\infty} p_\lambda(\alpha) = \int_{-\infty}^{+\infty} \lim_{\alpha \to +\infty} \mathrm{mid}(\alpha_l, \alpha_u, \alpha - \lambda s)\rho(s)\, ds = \int_{-\infty}^{+\infty} \alpha_u \rho(s)\; ds = \alpha_u.$$

∎

Lemma 2.2 shows that the mapping p_λ is a smooth approximation to the mid function with the same bounds as the mid function. We now show that the derivative of p_λ approximates the derivative of the function $\alpha \mapsto \mathrm{mid}(\alpha_u, \alpha_u, \alpha)$. In particular, when the support of the density function ρ,

$$\mathrm{supp}(\rho) = \{\alpha \in \mathbb{R} : \rho(\alpha) > 0\},$$

is all of $\mathbb{R}$, we show that $p'_\lambda(\alpha) \in (0, 1)$.

Lemma 2.3 *The mapping* $p_\lambda : \mathbb{R} \mapsto \mathbb{R}$ *is continuously differentiable with*

$$p'_\lambda(\alpha) = \int_{(\alpha - \alpha_u)/\lambda}^{(\alpha - \alpha_l)/\lambda} \rho(s)\, ds. \tag{2.4}$$

In particular, $p'_\lambda(\alpha) \in [0, 1]$. *Furthermore, if* $\mathrm{supp}(\rho) = \mathbb{R}$, *then* $p'_\lambda(\alpha) \in (0, 1)$.

Proof. The definition of the mid function implies that

$$p_\lambda(\alpha) = \alpha_u \int_{-\infty}^{(\alpha-\alpha_u)/\lambda} \rho(s)\, ds + \int_{(\alpha-\alpha_u)/\lambda}^{(\alpha-\alpha_l)/\lambda} (\alpha - \lambda s)\rho(s)\, ds + \alpha_l \int_{(\alpha-\alpha_l)/\lambda}^{+\infty} \rho(s)\, ds.$$

Hence, (2.4) follows by direct computation. If we assume that $\mathrm{supp}(\rho) = \mathbb{R}$, then we cannot have $p'_\lambda(\alpha) = 0$ because this implies that ρ vanishes in a nontrivial interval. ∎

An immediate consequence of Lemma 2.3 is that p_λ is an increasing function and that if $\mathrm{supp}(\rho) = \mathbb{R}$, then p_λ is strictly increasing. Also note that p_λ is nonexpansive, that is,

$$|p_\lambda(\alpha) - p_\lambda(\beta)| \le |\alpha - \beta|.$$

This property follows directly from the result $|p'_\lambda(\alpha)| \le 1$ in Lemma 2.3.

The behavior of p'_λ as λ goes to zero is important to the results in Section 4. Note that expression (2.4) shows that

$$\lim_{\lambda\to 0} p'_\lambda(\alpha) = 1, \quad \alpha \in (\alpha_l, \alpha_u), \qquad \lim_{\lambda\to 0} p'_\lambda(\alpha) = 0, \quad \alpha \notin [\alpha_l, \alpha_u]. \tag{2.5}$$

Expression (2.4) also shows that p_λ is twice differentiable with

$$p''_\lambda(\alpha) = \frac{1}{\lambda}\left[\rho\left(\frac{\alpha-\alpha_l}{\lambda}\right) - \rho\left(\frac{\alpha-\alpha_u}{\lambda}\right)\right]. \tag{2.6}$$

As a special case of (2.6) note that if $\alpha_u = +\infty$, then $p''_\lambda(\alpha) \ge 0$, and hence, p_λ is convex. This is reasonable to expect, since in this case, the function $\alpha \mapsto \mathrm{mid}(\alpha_u, \alpha_u, \alpha)$ is also convex. Also note that (2.6) implies that

$$\lim_{\lambda\to 0} p''_\lambda(\alpha) = 0, \qquad \alpha \notin \{\alpha_l, \alpha_u\},$$

but that $p''_\lambda(\alpha)$ can be unbounded for $\alpha \in \{\alpha_l, \alpha_u\}$.

An important property of the function p_λ is that we can bound the error between the smooth function p_λ and the original function $\mathrm{mid}(\alpha_l, \alpha_u, \alpha)$, independent of α.

Lemma 2.4 *If $p_\lambda : \mathbb{R} \mapsto \mathbb{R}$ is the smooth approximation to the* mid *function, and κ is the constant defined by* (2.1)*, then*

$$|p_\lambda(\alpha) - \mathrm{mid}(\alpha_l, \alpha_u, \alpha)| \le \kappa\lambda. \tag{2.7}$$

Proof. If $\phi : \mathbb{R} \mapsto \mathbb{R}$ is defined by $\phi(s) = \mathrm{mid}(\alpha_l, \alpha_u, s)$, then ϕ is the projection operator for the interval $[\alpha_l, \alpha_u]$, and thus

$$|\phi(\alpha_2) - \phi(\alpha_1)| \le |\alpha_2 - \alpha_1|. \tag{2.8}$$

We can also prove this inequality by noting that ϕ is piecewise linear, and that (2.8) holds in each of the pieces. Since

$$|p_\lambda(\alpha) - \operatorname{mid}(\alpha_l, \alpha_u, \alpha)| \le \int_{-\infty}^{+\infty} |\phi(\alpha - \lambda s) - \phi(\alpha)| \rho(s)\, ds \le \int_{-\infty}^{+\infty} \lambda |s| \rho(s)\, ds \le \kappa\lambda,$$

the result follows immediately. ∎

The techniques used in Lemma 2.4 can be used to establish additional properties for the function p_λ. For example, we can show that

$$|p_{\lambda_2}(\alpha) - p_{\lambda_1}(\alpha)| \le \kappa|\lambda_2 - \lambda_1|.$$

This result follows from (2.8) in Lemma 2.4, since

$$|p_{\lambda_2}(\alpha) - p_{\lambda_1}(\alpha)| \le \int_{-\infty}^{+\infty} |\phi(\alpha - \lambda_2 s) - \phi(\alpha - \lambda_2 s)|\, \rho(s)\, ds \le \kappa|\lambda_2 - \lambda_1|.$$

Thus, we have shown that the mapping $(\alpha, \lambda) \mapsto p_\lambda(\alpha)$ is Lipschitz in λ.

Chen and Mangasarian [6] approximated the function $\alpha \mapsto \alpha_+ = \max(0, \alpha)$ with the function $q : \mathbb{R} \mapsto \mathbb{R}$ defined by

$$q_\lambda(\alpha) = \frac{1}{\lambda} \int_{-\infty}^{\alpha} (\alpha - t) \rho\left(\frac{t}{\lambda}\right)\, dt = \frac{1}{\lambda} \int_{-\infty}^{+\infty} (\alpha - t)_+ \rho\left(\frac{t}{\lambda}\right)\, dt, \tag{2.9}$$

where ρ is a density function with bounded absolute mean. We now show that the two definitions agree in this case.

Lemma 2.5 *If $\alpha_l = 0$ and $\alpha_u = +\infty$, then $p_\lambda(\alpha) = q_\lambda(\alpha)$.*

Proof. The substitution $t = \lambda s$ shows that

$$q_\lambda(\alpha) = \frac{1}{\lambda} \int_{-\infty}^{+\infty} (\alpha - t)_+ \rho\left(\frac{t}{\lambda}\right)\, dt = \int_{-\infty}^{+\infty} \max(0, \alpha - \lambda s) \rho(s)\, ds = p_\lambda(\alpha),$$

as desired. ∎

Our development does not impose any restrictions on the density function ρ. In contrast, Chen and Mangasarian [6] based much of their development on the density function

$$\rho(\alpha) = \frac{\exp(\alpha)}{(1 + \exp(\alpha))^2}.$$

In global optimization work, Moré and Wu [16] have shown that the normal density function has superior smoothing properties, but at present it is not clear if similar results will hold in the complementarity area.

3 The Mixed Complementarity Problem

The mixed complementarity problem can be formulated as the solution of a nonsmooth system of nonlinear equations $r(x) = 0$, where $r : \mathbb{R}^n \mapsto \mathbb{R}^n$ is defined by

$$r(x) = x - \text{mid}(l, u, x - f(x)). \tag{3.1}$$

For future reference, we state this equivalence formally.

Lemma 3.1 *Let $r : \mathbb{R}^n \mapsto \mathbb{R}^n$ be the residual* (3.1)*. A vector $x \in \mathbb{R}^n$ solves the mixed complementarity problem* (1.1) *if and only if $r(x) = 0$.*

Lemma 3.1 is a special case of a result of Eaves [8], which formulates variational inequalities as a system of nonsmooth nonlinear equations. In the case of the mixed complementarity problem, the equivalence of (1.1) with $r(x) = 0$ is a direct consequence of the definition of the mid function.

We introduce a smooth residual for the mixed complementarity problem by first extending the function $p_\lambda : \mathbb{R} \mapsto \mathbb{R}$ to $\mathbb{R}^n$ by defining

$$p_\lambda(x) = (p_\lambda(x_i)).$$

The mapping $p_\lambda : \mathbb{R}^n \mapsto \mathbb{R}^n$ is continuously differentiable with a diagonal Jacobian matrix

$$p'_\lambda(x) = \text{diag}\left(p'_\lambda(x_i)\right).$$

Other properties of p_λ will be developed as needed.

Definition 3.2 *Let $\rho : \mathbb{R} \mapsto \mathbb{R}_+$ be a density function with bounded absolute mean. For any $\lambda > 0$ define the smooth residual $r_\lambda : \mathbb{R}^n \mapsto \mathbb{R}^n$ by*

$$r_\lambda(x) = x - p_\lambda(x - f(x)).$$

The smooth residual r_λ can be used to obtain approximate solutions to the mixed complementarity problem (1.1) by solving $r_\lambda(x) = 0$. On the other hand, Chen and Mangasarian [6] obtain approximate solutions to the mixed complementarity problem (1.1) by solving the system of nonlinear equations

$$\begin{pmatrix} f(x) - w + v \\ x - l - q_\lambda(x - l - w) \\ u - x - q_\lambda(u - x - v) \end{pmatrix} = 0,$$

where q_λ is the smooth function defined by (2.9). Our formulation in terms of $r_\lambda(x) = 0$ requires fewer variables and seems to be more suitable for computations.

We now extend Lemma 2.4 to $\mathbb{R}^n$ by assuming that $\|\cdot\|$ is a monotone norm, that is, $|x| \le |y|$ implies that $\|x\| \le \|y\|$. Clearly, all l_p norms are monotone.

Theorem 3.3 *Let $\|\cdot\|$ be a monotone norm. If $r_\lambda : \mathbb{R}^n \mapsto \mathbb{R}^n$ is the smooth residual, then*

$$\|r_\lambda(x) - r(x)\| \le \kappa\lambda\|e\|,$$

where $e \in \mathbb{R}^n$ is the vector of all ones, and κ is the constant defined by (2.1).

Proof. Lemma 2.4 shows that $\|p_\lambda(x) - \mathrm{mid}(l, u, x)\| \le \kappa\lambda\|e\|$, and thus,

$$\|r_\lambda(x) - r(x)\| = \|p_\lambda(x - f(x)) - \mathrm{mid}(l, u, x - f(x))\| \le \kappa\lambda\|e\|.$$

∎

We now show that approximate solutions to the system $r_\lambda(x) = 0$ are also approximate solutions to $r(x) = 0$ and that the accuracy depends mainly on λ.

Lemma 3.4 *If $r_\lambda : \mathbb{R}^n \mapsto \mathbb{R}^n$ is the smooth residual and $\|r_\lambda(x)\| \le \varepsilon$, then*

$$\|r(x)\| \le \lambda\kappa\|e\| + \varepsilon.$$

Proof. Lemma 3.3 implies that

$$\|r(x)\| \le \|r_\lambda(x) - r(x)\| + \|r_\lambda(x)\| \le \kappa\lambda\|e\| + \varepsilon,$$

as desired.∎

Finally, we show that if we find an $x \in \mathbb{R}^n$ with $\|r(x)\|$ small, then x is an approximate solution of (1.1).

Lemma 3.5 *Let $r : \mathbb{R}^n \mapsto \mathbb{R}^n$ be the residual for the mixed complementarity problem* (1.1). *If $\|r(x)\|_\infty \le \varepsilon$, then $x_i \in [l_i - \varepsilon, u_i + \varepsilon]$ for all indices i, and*

$$\begin{array}{lll} |f_i(x)| \le \varepsilon & \textit{if} & x_i \in (l_i + \varepsilon, u_i - \varepsilon), \\ f_i(x) \ge -\varepsilon & \textit{if} & |x_i - l_i| \le \varepsilon, \\ f_i(x) \le \varepsilon & \textit{if} & |x_i - u_i| \le \varepsilon. \end{array}$$

Proof. Since the mid only has three possible values,

$$r_i(x) = \begin{cases} f_i(x) & \text{if} \quad l_i \le x_i - f_i(x) \le u_i, \\ x_i - l_i & \text{if} \quad x_i - f_i(x) < l_i, \\ x_i - u_i & \text{if} \quad x_i - f_i(x) > u_i. \end{cases}$$

Since $|r_i(x)| \le \varepsilon$, we obtain that $x_i \in [l_i - \varepsilon, u_i + \varepsilon]$ in all three cases. If $x_i \in (l_i + \varepsilon, u_i - \varepsilon)$, then $|x_i - l_i| > \varepsilon$ and $|x_i - u_i| > \varepsilon$, so we must have $r_i(x) = f_i(x)$, and thus $|f_i(x)| \le \varepsilon$. The proofs of the other two cases are similar. ∎

4 Regularity

Given the smooth residual $r_\lambda : \mathbb{R}^n \mapsto \mathbb{R}^n$, we approach the mixed complementarity problem by generating a path $x(\lambda)$ for $\lambda > 0$ in which $x(\lambda)$ is an approximate solution to the least squares problem

$$\min \left\{ \tfrac{1}{2} \|r_\lambda(x)\|^2 : x \in \mathbb{R}^n \right\}. \tag{4.1}$$

Given a sequence $\{\lambda_k\}$ that converges to zero, we require that the approximate minimizer x_k be generated so that

$$\lim_{k \to +\infty} r_k'(x_k)^T r_k(x_k) = 0, \tag{4.2}$$

where

$$r_k(x) \equiv r_{\lambda_k}(x).$$

There are various ways to generate iterates that satisfy (4.2). For example, given the iterate (x_k, λ_k), choose $\lambda_{k+1} > 0$, and use any nonlinear least squares algorithm on (4.1) with $\lambda = \lambda_{k+1}$ to generate an iterate such that

$$\|r_{k+1}'(x_{k+1})^T r_{k+1}(x_{k+1})\| \le \eta \|r_k'(x_k)^T r_k(x_k)\|, \tag{4.3}$$

where $\eta \in (0, 1)$. Clearly, if the sequence $\{(x_k, \lambda_k)\}$ satisfies (4.3), then (4.2) holds.

We want to impose conditions on $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ that guarantee that if $\{(x_k, \lambda_k)\}$ satisfies (4.2), then any limit point of $\{x_k\}$ solves the mixed complementarity problem. Since

$$r_\lambda'(x) = I - D(x)(I - f'(x)), \tag{4.4}$$

where $D(x) = p_\lambda'(x - f(x))$, we are led to the study of matrices of the form $M = I - D + DA$, where $D = \mathrm{diag}(d_i)$ has $d_i \in [0, 1]$. Our results show a strong connection between the nonsingularity of M and the class of P- and P_0-matrices.

A matrix $A \in \mathbb{R}^{n \times n}$ is a P-matrix if for any $x \neq 0$ there is an index i such that $x_i[Ax]_i > 0$. Similarly, a matrix $A \in \mathbb{R}^{n \times n}$ is a P_0-matrix if for each $x \neq 0$ there is an index i such that $x_i \neq 0$ and $x_i[Ax]_i \geq 0$. A P-matrix (P_0-matrix) can also be defined as any matrix for which all principal submatrices have positive (nonnegative) determinants. For additional information on P-matrices, see the book of Cottle, Pang, and Stone [7].

Our analysis of matrices of the form $M = I - D + DA$ hinges on the following result of Sandberg and Willson [22].

Theorem 4.1 *$A \in \mathbb{R}^{n \times n}$ is a P_0-matrix if and only if $A + D$ is nonsingular for each diagonal matrix D with positive diagonal entries.*

Theorem 4.1 has been rediscovered at least twice. This result can be found, for example, in Chen and Harker [3, Theorem 3.3], and in Luca, Facchinei, and Kanzow [14, Lemma 5.1].

Theorem 4.2 *$A \in \mathbb{R}^{n\times n}$ is a P_0-matrix if and only if $M = I - D + DA$ is nonsingular for any diagonal matrix $D = \operatorname{diag}(d_i)$ with $0 \le d_i < 1$.*

Proof. If A is a P_0-matrix, then DA is also a P_0-matrix for any nonnegative diagonal matrix D, and since $I - D$ has positive diagonal entries, Theorem 4.1 implies that M is nonsingular. Conversely, assume that M is nonsingular for any diagonal matrix D with $0 < d_i < 1$. Note that the diagonal matrix $S = D^{-1} - I$ has positive diagonal entries if and only if D has entries with $0 < d_i < 1$, and that $M = D(S + A)$. Hence, M is nonsingular if and only if $S + A$ is nonsingular, and thus the result follows from Theorem 4.1. ∎

If the support of the density function used to generate the smooth residual is $\mathbb{R}$, then Lemma 2.3 guarantees that $p'_\lambda(\alpha) \in (0,1)$, and thus Theorem 4.2 shows that the Jacobian matrix (4.4) of the smooth residual r_λ is nonsingular if $f'(x)$ is a P_0-matrix. If ρ has compact support, then we need to impose stronger conditions.

Theorem 4.3 *$A \in \mathbb{R}^{n\times n}$ is a P-matrix if and only if $M = I - D + DA$ is nonsingular for any diagonal matrix $D = \operatorname{diag}(d_i)$ with $0 \le d_i \le 1$.*

Proof. Assume that M is nonsingular for any diagonal matrix $D = \operatorname{diag}(d_i)$ with $0 \le d_i \le 1$. Theorem 4.2 shows that A must be a P_0-matrix, so we need to prove only that $A_{\mathcal{K},\mathcal{K}}$ is nonsingular for any index set $\mathcal{K}$. If we set $d_i = 1$ for $i \in \mathcal{K}$, and $d_i = 0$ for $i \notin \mathcal{K}$, then

$$M = \begin{pmatrix} A_{\mathcal{K},\mathcal{K}} & A_{\mathcal{K},\mathcal{L}} \\ 0 & I \end{pmatrix},$$

where $\mathcal{L}$ is the complement of $\mathcal{K}$. Hence, $A_{\mathcal{K},\mathcal{K}}$ is nonsingular, as desired. Conversely, if A is a P-matrix, but $Mx = 0$ for some $x \neq 0$, then

$$d_i[Ax]_i = (d_i - 1)x_i, \qquad 1 \le i \le n.$$

If $d_i = 0$, then $x_i = 0$, while if $d_i > 0$, then

$$x_i[Ax]_i = \left(\frac{d_i - 1}{d_i}\right) x_i^2 \le 0.$$

Hence, $x_i[Ax]_i \le 0$ for all indices i, contradicting the assumption that A is a P-matrix. ∎

We can weaken the conditions needed for nonsingularity of M when we know that the submatrix $A_{\mathcal{K},\mathcal{K}}$ with $\mathcal{K} = \{i : d_i = 1\}$ is nonsingular. For this result, recall that if the principal submatrix $A_{\mathcal{K},\mathcal{K}}$ of A is nonsingular, and $\mathcal{L}$ is the complement of $\mathcal{K}$ in $\{1, 2, \ldots, n\}$, then

$$A/A_{\mathcal{K},\mathcal{K}} = A_{\mathcal{L},\mathcal{L}} - A_{\mathcal{L},\mathcal{K}} A_{\mathcal{K},\mathcal{K}}^{-1} A_{\mathcal{K},\mathcal{L}}$$

is the Schur complement of $A_{\mathcal{K},\mathcal{K}}$ in A, and that the fundamental relationship

$$\begin{pmatrix} I & 0 \\ -A_{\mathcal{L},\mathcal{K}}A_{\mathcal{K},\mathcal{K}}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{\mathcal{K},\mathcal{K}} & A_{\mathcal{K},\mathcal{L}} \\ A_{\mathcal{L},\mathcal{K}} & A_{\mathcal{L},\mathcal{L}} \end{pmatrix} = \begin{pmatrix} A_{\mathcal{K},\mathcal{K}} & A_{\mathcal{K},\mathcal{L}} \\ 0 & A/A_{\mathcal{K},\mathcal{K}} \end{pmatrix}, \tag{4.5}$$

shows that A is nonsingular if and only if the Schur complement $A/A_{\mathcal{K},\mathcal{K}}$ is nonsingular.

Theorem 4.4 *Let $D = \mathrm{diag}(d_i)$ be a diagonal matrix with $d_i \in [0,1]$, and set*

$$\mathcal{K} = \{i : d_i = 1\}, \qquad \mathcal{L} = \{i : 0 < d_i < 1\}.$$

If the principal submatrix $A_{\mathcal{K}}$ of $A \in \mathbb{R}^{n\times n}$ is nonsingular and $A_{\mathcal{K}\cup\mathcal{L}}/A_{\mathcal{K}}$ is a P_0-matrix, then $M = I - D + DA$ is nonsingular.

Proof. Define $\mathcal{S} = \{i : d_i = 0\}$, and note that we can permute the rows and columns of A so that

$$M = \begin{pmatrix} I & 0 & 0 \\ A_{\mathcal{K},\mathcal{S}} & A_{\mathcal{K},\mathcal{K}} & A_{\mathcal{K},\mathcal{L}} \\ EA_{\mathcal{L},\mathcal{S}} & EA_{\mathcal{L},\mathcal{K}} & I - E + EA_{\mathcal{L},\mathcal{L}} \end{pmatrix},$$

where $E = \mathrm{diag}(e_i)$ and $0 < e_i < 1$. Hence, M is nonsingular if and only if the lower 2×2 block principal submatrix $\widehat{M}$ of M is nonsingular. A computation shows that

$$\widehat{M}/A_{\mathcal{K}} = I - E + E(A_{\mathcal{K}\cup\mathcal{L}}/A_{\mathcal{K}}),$$

and thus Theorem 4.1 shows that $\widehat{M}/A_{\mathcal{K}}$ is nonsingular. Hence, M is nonsingular, as desired. ∎

Theorem 4.4 can be used to prove that every limit point of $\{r_k'(x_k)\}$ is nonsingular provided a regularity condition is imposed. For any $x \in \mathbb{R}^n$ define the index sets

$$\begin{aligned} \mathcal{I} &= \{i : l_i < x_i - f_i(x) < u_i\}, \\ \mathcal{B} &= \{i : x_i - f_i(x) = l_i\} \cup \{i : x_i - f_i(x) = u_i\}, \\ \mathcal{E} &= \{i : x_i - f_i(x) < l_i\} \cup \{i : x_i - f_i(x) > u_i\}. \end{aligned} \tag{4.6}$$

The notation is suggestive because $\mathcal{I}, \mathcal{B}, \mathcal{E}$ are the indices of $x_i - f_i(x)$ that are, respectively, on the interior, boundary, and exterior of the interval $[l_i, u_i]$. Also note that we have suppressed the dependence on x because the vector x in question will always be clear from the discussion.

Definition 4.5 *A vector $x \in \mathbb{R}^n$ is* regular *if $[f'(x)]_{\mathcal{K}}$ is nonsingular for all $\mathcal{K}$ such that $\mathcal{I} \subset \mathcal{K} \subset \mathcal{I}\cup\mathcal{B}$, and*

$$[f'(x)]_{\mathcal{I}\cup\mathcal{B}}/[f'(x)]_{\mathcal{I}}$$

is a P_0-matrix.

Regularity conditions typically are used to show that if a complementarity problem is formulated as an optimization problem, then any local minimizer that satisfies the regularity condition is actually a global minimizer, and thus a solution to the complementarity problem. Pang and Gabriel [17], Moré [15], Xiao and Harker [23, 24], Ferris and Ralph [10], and De Luca, Facchinei, and Kanzow [14] have introduced regularity conditions for the classical complementarity problem (1.2), while Gabriel [11], Billups [1], and Billups, Dirkse, and Ferris [2] seem to be the only researchers that have used this type of regularity condition for mixed complementarity problems. Comparisons between these regularity conditions are difficult because they depend on the formulation of the complementarity problem as an optimization problem. We could use our regularity condition to show that any stationary point of

$$\min\left\{\tfrac{1}{2}\|r(x)\|_2^2 : x \in \mathbb{R}^n\right\}$$

is a solution to the mixed complementarity problem (1.1), but we use it to guarantee that every limit point of $\{r_k'(x_k)\}$ is nonsingular.

Luca, Facchinei, and Kanzow [14] used a similar regularity condition for the classical complementarity problem to guarantee that all elements of the generalized Jacobian of the residual $r : \mathbb{R}^n \mapsto \mathbb{R}^n$, where

$$r(x) = \phi(x, f(x)), \qquad \phi(\alpha, \beta) = \sqrt{(\alpha^2 + \beta^2)} - (\alpha + \beta)$$

are nonsingular. Our regularity condition seems to be weaker because if (as expected) $\mathcal{B}$ is empty, then we only require $[f'(x)]_{\mathcal{I}}$ to be nonsingular.

If x is a solution to the mixed complementarity problem (1.1), then the index sets (4.6) can be be expressed in the form

$$\begin{aligned}
\mathcal{I} &= \{i : x_i \in (l_i, u_i),\ f_i(x) = 0\}, \\
\mathcal{B} &= \{i : x_i \in \{l_i, u_i\},\ f_i(x) = 0\}, \\
\mathcal{E} &= \{i : x_i = l_i,\ f_i(x) > 0\} \cup \{i : x_i = u_i,\ f_i(x) < 0\}.
\end{aligned}$$

Facchinei and Kanzow [9] used this condition for the classical complementarity problem, where $l \equiv 0$ and $u \equiv +\infty$, to guarantee superlinear convergence of a truncated Newton method.

Theorem 4.6 *Assume that $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is continuously differentiable. If $\{(x_k, \lambda_k)\}$ converges to $(x^*, 0)$ and x^* is regular, then every limit point of $\{r_k'(x_k)\}$ is nonsingular.*

Proof. Any limit point of $\{r_k'(x_k)\}$ is of the form $I - D + Df'(x^*)$, where $D = \operatorname{diag}(d_i)$ has $d_i \in [0, 1]$, and thus the proof follows from Theorem 4.4 if we show that $[f'(x^*)]_{\mathcal{K}}$ is nonsingular for $\mathcal{K} = \{i : d_i = 1\}$ and that the matrix

$$B = [f'(x^*)]_{\mathcal{K}\cup\mathcal{L}}/[f'(x^*)]_{\mathcal{K}}, \tag{4.7}$$

where $\mathcal{L} = \{i : 0 < d_i < 1\}$, is a P_0-matrix. We first prove that $[f'(x^*)]_{\mathcal{K}}$ is nonsingular.

We claim that $\mathcal{I} \subset \mathcal{K} \subset (\mathcal{I} \cup \mathcal{B})$. If $i \in \mathcal{I}$, then (2.5) shows that $d_i = 1$, and thus $i \in \mathcal{K}$. Hence, $\mathcal{I} \subset \mathcal{K}$. We prove that $\mathcal{K} \subset (\mathcal{I} \cup \mathcal{B})$ by noting that if $i \notin (\mathcal{I} \cup \mathcal{B})$, then $i \in \mathcal{E}$, and thus (2.5) implies that $d_i = 0$. Hence $i \notin \mathcal{K}$. This establishes our claim.

Since $\mathcal{I} \subset \mathcal{K} \subset (\mathcal{I} \cup \mathcal{B})$, the regularity assumption implies that $[f'(x^*)]_{\mathcal{K}}$ is nonsingular. We now prove that the matrix B in (4.7) is a P_0-matrix.

We need to know that a submatrix of a P_0-matrix is also a P_0-matrix. This result is a direct consequence of the definition of a P_0-matrix. We also need to know that the Schur complement of a P_0-matrix is again a P_0-matrix. This result is due to Chen and Harker [4, Lemma 2.3].

The Schur quotient formula (see, for example, Cottle, Pang, and Stone [7, pages 76–77]) shows that

$$[f'(x^*)]_{\mathcal{I}\cup\mathcal{B}}/[f'(x^*)]_{\mathcal{K}} = \left([f'(x^*)]_{\mathcal{I}\cup\mathcal{B}}/[f'(x^*)]_{\mathcal{I}}\right) / \left([f'(x^*)]_{\mathcal{K}}/[f'(x^*)]_{\mathcal{I}}\right).$$

is a Schur complement of the P_0-matrix $[f'(x^*)]_{\mathcal{I}\cup\mathcal{B}}/[f'(x^*)]_{\mathcal{I}}$. Hence, $[f'(x^*)]_{\mathcal{I}\cup\mathcal{B}}/[f'(x^*)]_{\mathcal{K}}$ is also a P_0-matrix.

We prove that $(\mathcal{K} \cup \mathcal{L}) \subset (\mathcal{I} \cup \mathcal{B})$ by noting that if $i \notin (\mathcal{I} \cup \mathcal{B})$, then $i \in \mathcal{E}$, and thus (2.5) implies that $d_i = 0$. Hence, $i \notin (\mathcal{K} \cup \mathcal{L})$, as desired. Since $(\mathcal{K} \cup \mathcal{L}) \subset (\mathcal{I} \cup \mathcal{B})$, the matrix B in (4.7) is a submatrix of the P_0-matrix $[f'(x^*)]_{\mathcal{I}\cup\mathcal{B}}/[f'(x^*)]_{\mathcal{K}}$. Hence, B is a P_0-matrix. ∎

Theorem 4.6 is applicable to any sequence $\{x_k\}$. If we assume that x_k is an approximate minimizer of (4.1), in the sense that (4.2) holds, then we can obtain a result applicable to the mixed complementarity problem (1.1).

Theorem 4.7 *Assume that $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is continuously differentiable, that $\{\lambda_k\}$ converges to zero, and that $\{(x_k, \lambda_k)\}$ satisfies (4.2). If x^* is a limit point of $\{x_k\}$, and x^* is regular, then x^* solves the mixed complementarity problem.*

Proof. Without loss of generality, assume that $\{x_k\}$ converges to x^*. Since Theorem 4.7 shows that every limit point of $\{r_k'(x_k)\}$ is nonsingular, (4.2) implies that $\{r_k(x_k)\}$ converges to zero. Since Lemma 3.3 shows that

$$\|r_k(x_k) - r(x_k)\| \le \kappa \lambda_k \|e\|,$$

and since $\{\lambda_k\}$ converges to zero, we obtain that $r(x^*) = 0$, and thus x^* solves the mixed complementarity problem. ∎

5 Concluding Remarks

After completing the manuscript, Francisco Facchinei and Christian Kanzow pointed out that $[f'(x)]_{\mathcal{I}\cup\mathcal{B}}/[f'(x)]_{\mathcal{I}}$ is a P-matrix under our regularity assumption. This result shows, for example, that the R-regularity assumption of Robinson [18] is equivalent to our regularity assumption if x is a solution of the classical complementarity problem.

The observation of Facchinei and Kanzow follows by noting that if A is nonsingular, then (4.5) shows that the Schur complement $A/A_{\mathcal{K}}$ is nonsingular with respect to any nonsingular principal submatrix $A_{\mathcal{K}}$. Hence, if $A/A_{\mathcal{K}}$ is a P_0-matrix, then $A/A_{\mathcal{K}}$ has a positive determinant.

Acknowledgments

Zhijun Wu deserves special mention for discussions on the numerical aspects of smoothing, and their use in complementarity theory. We also thank Andy Kydes (Energy Information Agency) for providing motivation and support for this work.

References

[1] S. C. Billups, *Algorithms for complementarity problems and generalized equations*, PhD thesis, University of Wisconsin, Madison, Wisconsin, 1995.

[2] S. C. Billups, S. P. Dirkse, and M. C. Ferris, *A comparison of algorithms for large-scale mixed complementarity problems*, Technical report 95-16, University of Wisconsin, Madison, Wisconsin, 1995.

[3] B. T. Chen and P. T. Harker, *A non-interior-point continuation method for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1168–1190.

[4] ——, *A noninterior continuation method for quadratic and linear programming*, SIAM J. Opt., 3 (1993), pp. 503–515.

[5] ——, *Smooth approximations to nonlinear complementarity problems: Theory and convergence of algorithms*, Technical report, University of Pennsylvania, Philadelphia, Pennsylvania, 1995.

[6] C. H. Chen and O. L. Mangasarian, *A class of smoothing functions for nonlinear and mixed complementarity problems*, Technical report 95-11, University of Wisconsin, Madison, Wisconsin, 1994.

[7] R. W. Cottle, J. S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, 1992.

[8] B. C. Eaves, *On the basic theorem of complementarity*, Math Prog., 1 (1971), pp. 68–75.

[9] F. Facchinei and C. Kanzow, *A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems*, Technical report 03.95, Universita di Roma, La Sapienza, Rome, Italy, 1995.

[10] M. C. Ferris and D. Ralph, *Projected gradient methods for nonlinear complementarity problems via normal maps*, in Recent Advances in Nonsmooth Optimization, D.-Z. Du, L. Qi, and R. S. Wommersley, eds., World Scientifc Publishers, 1995, pp. 57–86.

[11] S. A. Gabriel, *An NE/SQP method for bounded nonlinear complementarity problems*, Preprint MCS–P508–0495, Argonne National Laboratory, Argonne, Illinois, 1995.

[12] V. Y. Katkovnik and O. Y. Kulchitskii, *Convergence of a class of random search algorithms*, Automat. Remote Control, 8 (1972), pp. 81–87.

[13] J. Kreimer and R. Y. Rubinstein, *Nondifferentiable optimization via smooth approximation: General analytical approach*, Math. Oper. Res., 39 (1992), pp. 97–119.

[14] T. D. Luca, F. Facchinei, and C. Kanzow, *A semismooth equation approach to the solution of nonlinear complementarity problems*, Technical report 01.95, Universita di Roma, La Sapienza, Rome, Italy, 1995.

[15] J. J. Moré, *Global methods for nonlinear complementarity problems*, Preprint MCS-P429-0494, Argonne National Laboratory, Argonne, Illinois, 1994.

[16] J. J. Moré and Z. Wu, *Smoothing techniques for macromolecular global optimization*, Preprint MCS-P542-0995, Argonne National Laboratory, Argonne, Illinois, 1995.

[17] J. S. Pang and S. A. Gabriel, *NE/SQP: A robust algorithm for the nonlinear complementarity problem*, Math. Programming, 60 (1993), pp. 295–337.

[18] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.

[19] R. Y. Rubinstein, *Simulation and the Monte Carlo Method*, John Wiley & Sons, 1981.

[20] ——, *Smoothed functionals in stochastic optimization*, Math. Oper. Res., 8 (1983), pp. 26–33.

[21] ——, *Monte Carlo Optimization, Simulation and Sensitivity of Queueing Networks*, John Wiley & Sons, 1986.

[22] I. W. Sandberg and A. N. Willson, *Some theorems on properties of DC equations of nonlinear networks*, AT&T Bell Labs. Tech. J., 48 (1969), pp. 1–34.

[23] B. Xiao and P. T. Harker, *A nonsmooth Newton method for variational inequalities, I: Theory*, Math. Programming, 65 (1994), pp. 151–194.

[24] ——, *A nonsmooth Newton method for variational inequalities, II: Numerical results*, Math. Programming, 65 (1994), pp. 195–216.