Global convergence of sequential injective algorithm for weakly univalent vector equation: application to regularized smoothing Newton algorithm

Shunsuke Hayashi*

Abstract It is known that the complementarity problems and the variational inequality problems are reformulated equivalently as a vector equation by using the natural residual or Fischer-Burmeister function. In this short paper, we first study the global convergence of a sequential injective algorithm for weakly univalent vector equation. Then, we apply the convergence analysis to the regularized smoothing Newton algorithm for mixed nonlinear second-order cone complementarity problems. We prove the global convergence property under the (Cartesian) P_0 assumption, which is strictly weaker than the original monotonicity assumption.

1 Sequential injective algorithm for weakly univalent equation

Consider the following vector equation (VE):

$$H(z) = 0, \tag{1.1}$$

where $H : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuous over the domain $\mathcal{D} \subseteq \mathbb{R}^n$, but need not be linear or differentiable. Notice that many classes of problems such as linear complementarity problem (LCP), nonlinear complementarity problem (NCP), second-order cone complementarity problem (SOCCP) [3, 4], symmetric cone complementarity problem (SCCP) [5], variational inequality problem (VIP) and fixed point problem can be cast as VE (1.1). (For more details, see [2]). In this section, we first study the global convergence of a certain conceptual algorithm for solving VE (1.1). Then, in the next section, we apply the obtained convergence theorem to the regularized smoothing Newton algorithm in a direct manner.

1.1 Weak univalence property

In the convergence analysis, the notion of the weak univalence property plays an important role.

Definition 1.1 (weak univalence property) [2, Sec. 3.6] *Function $H : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is said to be "weakly univalent" if it is continuous and there exists a sequence of continuous and injective functions $\{\tilde{H}_k\}$ converging to H uniformly*[*1] *over any bounded subset of $\mathcal{D}$.*

We can easily see that any weakly univalent function is continuous. Moreover, any P_0 function or monotone function is weakly univalent. For VE (1.1), we suppose that H satisfies the following assumption:

*Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan (s_hayashi@plan.civil.tohoku.ac.jp)

*1 We say that the sequence of functions $\{\tilde{H}_k\}$ converges to H uniformly over the bounded set Ω, if $\sup\{\|\tilde{H}_k(w) - H(w)\| : w \in \Omega\}$ converges to 0 as $k \to \infty$.

Assumption A $H : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *satisfies the following statements*:

(i) H *is weakly univalent*;

(ii) *The solution set* $H^{-1}(0)$ *is nonempty and bounded.*

1.2 Sequential injective algorithm and its global convergence

For solving VE (1.1), we first provide the following conceptual algorithm, which involves the regularized smoothing Newton algorithm as a special case.

Algorithm 1 (Sequential injective algorithm)

> **Step 0** *Choose* $w^0 \in \mathcal{D}$ *and* $\beta_0 \geq 0$ *arbitrarily. Obtain a continuous and injective function* $\tilde{H}_0 : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^n$. *Set* $k := 0$.
>
> **Step 1** *If* $\|H(w^k)\| = 0$, *then terminate and output* w^k *as a solution. Otherwise, go to Step* 2.
>
> **Step 2** *Find a vector* $w^{k+1} \in \mathbb{R}^n$ *such that*
>
> $$\|\tilde{H}_k(w^{k+1})\| \ \leq \ \beta_k. \tag{1.2}$$
>
> **Step 3** *Obtain* $\beta_{k+1} \geq 0$ *and a continuous injective function* $\tilde{H}_{k+1} : \mathbb{R}^n \to \mathbb{R}^n$ *so that they converge to* 0 *and* H *eventually and uniformly. Set* $k := k+1$. *Go back to Step* 1.

To obtain w^{k+1} in Step 2, we may use any suitable unconstrained minimization technique. These issues will be discussed later. In order for Algorithm 1 to be well-defined, there must exist w^{k+1} satisfying (1.2).

Assumption B *For the functions* $\{\tilde{H}_k\}$ *and parameters* $\{\beta_k\}$ *used in Algorithm* 1, $\{w \mid \|\tilde{H}_k(w)\| \leq \beta_k\}$ *is nonempty for for all* k.

Note that Assumption B holds when $\tilde{H}_k^{-1}(0) \neq \emptyset$.

Now, we are to show the global convergence of the algorithm. To this end, we introduce the following lemma, which indicates a property the weakly univalent functions possess.

Lemma 1.1 [2, Cor. 3.6.5] *Let* $H : \mathbb{R}^n \to \mathbb{R}^n$ *be a weakly univalent function such that the inverse image* $H^{-1}(0)$ *is nonempty and compact. Then, for any* $\varepsilon > 0$, *there exists* $\delta = \delta(\varepsilon) > 0$ *such that the following statement holds*: *For any weakly univalent function* $\tilde{H} : \mathbb{R}^n \to \mathbb{R}^n$ *such that*[*2]

$$\sup\left\{\|\tilde{H}(w) - H(w)\| : w \in \mathrm{cl}(H^{-1}(0) + B(0,\varepsilon))\right\} \leq \delta,$$

we have

$$\emptyset \neq \tilde{H}^{-1}(0) \subseteq H^{-1}(0) + B(0,\varepsilon),$$

and $\tilde{H}^{-1}(0)$ *is connected.*

By using this lemma, we establish the global convergence of Algorithm 1.

[*2] Here, $\mathrm{cl}\,(\cdot)$ and $B(0,\varepsilon)$ denote the closure and the open ball with radius $\varepsilon > 0$, i.e., $B(0,\varepsilon) := \{w \in \mathbb{R}^n \mid \|w\| < \varepsilon\}$, respectively.

Theorem 1.1 *Suppose that Assumptions A and B holds. Let $\{w^k\}$ be the sequence generated by Algorithm 1. Then, $\{w^k\}$ is bounded and any accumulation point solves VE(1.1).*[*3]

Proof. We first show the boundedness of $\{w^k\}$. Choose $\varepsilon > 0$ arbitrarily. Then, there exists $\delta = \delta(\varepsilon) > 0$ such that Lemma 1.1 holds. Let $\Omega_\varepsilon := H^{-1}(0) + B(0, \varepsilon)$, which is nonempty and bounded by Assumption A(ii). Let $\tilde{G}_k : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be defined by $\tilde{G}_k(w) := \tilde{H}_k(w) - \tilde{H}_k(w^{k+1})$ for each k. Then, there exists $\bar{k}$ such that

$$\begin{aligned} \|\tilde{G}_k(w) - H(w)\| &= \|\tilde{H}_k(w) - \tilde{H}_k(w^{k+1}) - H(w)\| \\ &\leq \|\tilde{H}_k(w) - H(w)\| + \|\tilde{H}_k(w^{k+1})\| \\ &\leq \sup\left\{\|\tilde{H}_k(w') - H(w')\| : w' \in \operatorname{cl}\Omega_\varepsilon\right\} + \beta_k \\ &\leq \delta \end{aligned}$$

for any $k \geq \bar{k}$ and $w \in \operatorname{cl}\Omega_\varepsilon$. Here, the first inequality is due to the triangular inequality, the second inequality holds from $w \in \operatorname{cl}\Omega_\varepsilon$ and (1.2), and the last inequality follows since β_k converges to 0 and $\tilde{H}_k(w)$ converges to $H(w)$ uniformly over the compact set $\operatorname{cl}\Omega_\varepsilon$. Note that $\tilde{G}_k$ is weakly univalent since it is continuous and injective. Thus, by Lemma 1.1 with $\tilde{H} := \tilde{G}_k$, we have

$$\emptyset \neq \tilde{G}_k^{-1}(0) \subseteq \Omega_\varepsilon$$

for all $k \geq \bar{k}$. Since $w^{k+1} \in \tilde{G}_k^{-1}(0)$ and Ω_ε is bounded, we have the boundedness of $\{w^k\}$.

Next we show the latter part. Since $\{w^k\}$ is bounded, there exists a bounded set W satisfying $\{w^k\} \subseteq W$. By the triangular inequality and Step 2 of Algorithm 1, we have

$$\begin{aligned} \|H(w^{k+1})\| &\leq \|\tilde{H}_k(w^{k+1}) - H(w^{k+1})\| + \|H_k(w^{k+1})\| \\ &\leq \sup\left\{\|\tilde{H}_k(w') - H(w')\| : w' \in W\right\} + \beta_k. \end{aligned}$$

Thus we have $\|H(w^{k+1})\| \to 0$. Since H is continuous, arbitrary accumulation point w^* of $\{w^k\}$ satisfies $H(w^*) = 0$. ■

Corollary 1.1 *Let $H_{\mu,\varepsilon} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a function with parameters $\mu \geq 0$ and $\varepsilon \geq 0$ such that*

(i) *$H_{\mu,\varepsilon}$ is continuously differentiable over $\mathcal{D}$ for any $\mu > 0$ and $\varepsilon > 0$,*

(ii) *$\nabla H_{\mu,\varepsilon}(w)$ is nonsingular for any $\mu > 0$, $\varepsilon > 0$ and $w \in \mathcal{D}$,*

(iii) *$H_{\mu,\varepsilon}$ converges to H uniformly over an arbitrary bounded subset of $\mathcal{D}$ as $(\mu, \varepsilon) \searrow (0, 0)$.*

Let $\{(\mu_k, \varepsilon_k)\} \subseteq \mathbb{R}_+ \times \mathbb{R}_+$ be arbitrary sequences such that $\mu_k > 0$ and $\varepsilon_k > 0$ for any k, and $(\mu_k, \varepsilon_k) \searrow (0, 0)$ as $k \to \infty$. Suppose that Assumptions A and B hold. Then, any accumulation point of the sequence $\{w^k\}$ generated by Algorithm 1 with $\tilde{H}_k := H_{\mu_k,\varepsilon_k}$ solves VE(1.1).

Proof. By (i) and (ii), the function H_{μ_k,ε_k} is continuous and injective for each k. Thus, by (iii) and Theorem 1.1, we have the corollary. ■

[*3]This implies that the distance from w^k to $H^{-1}(0)$ converges to 0 as $k \to \infty$.

2 Regularized smoothing Newton algorithm for mixed nonlinear second-order cone complementarity problem

In this section, we focus on the regularized smoothing Newton algorithm [1, 4] for solving the following mixed nonlinear second-order cone complementarity problem (MNSOCCP):

$$\text{(MNSOCCP)} \qquad \begin{array}{rl} \text{Find} & (x,y,p) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^\ell \\ \text{such that} & x \in \mathcal{K},\ y \in \mathcal{K},\ x^\top y = 0, \\ & y = F_1(x,p),\ F_2(x,p) = 0, \end{array} \tag{2.1}$$

where $F_1 : \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}^n$ and $F_2 : \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}^\ell$ are given continuously differentiable functions, and $\mathcal{K}$ is a Cartesian product of several second-order cones (SOCs), i.e.,

$$\mathcal{K} := \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_m} \tag{2.2}$$

with $n_1 + n_2 + \cdots + n_m = n$ and

$$\mathcal{K}^{n_i} := \begin{cases} \left\{ z \in \mathbb{R} \;\middle|\; z \geq 0 \right\} & (n_i = 1) \\ \left\{ z \in \mathbb{R}^{n_i} \;\middle|\; z_1 \geq \sqrt{z_2^2 + \cdots + z_{n_i}^2} \right\} & (n_i \geq 2). \end{cases}$$

We therefore have

$$\mathbb{R}^n_+ = \mathcal{K}^1 \times \mathcal{K}^1 \times \cdots \times \mathcal{K}^1,$$

where $\mathbb{R}^n_+$ denotes the nonnegative orthant in $\mathbb{R}^n$.

MNSOCCP (2.1) involves many kinds of problems as special cases as follows.

(i) When $\mathcal{K} = \mathbb{R}^n_+$ (i.e., $n_1 = n_2 = \cdots = n_m = 1$), MNSOCCP (2.1) reduces to the following nonlinear mixed complementary problem (MCP):

$$\text{(MCP)} \qquad \begin{array}{rl} \text{Find} & (x,p) \in \mathbb{R}^n \times \mathbb{R}^\ell \\ \text{such that} & x \geq 0,\ y \geq 0,\ x^\top y = 0, \\ & y = F_1(x,p),\ F_2(x,p) = 0. \end{array} \tag{2.3}$$

(ii) When $\mathcal{K} = \mathbb{R}^n_+$ and $\ell = 0$, MNSOCCP (2.1) reduces to the following nonlinear complementary problem (NCP):

$$\text{(NCP)} \qquad \begin{array}{rl} \text{Find} & x \in \mathbb{R}^n \\ \text{such that} & x \geq 0,\ y \geq 0,\ x^\top y = 0, \\ & y = F_1(x). \end{array}$$

(iii) When $\mathcal{K} = \mathbb{R}^n_+$ and $F_1(x) = Mx + q$ for some $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$, MNSOCCP (2.1) reduces to the following linear complementary problem (LCP):

$$\text{(LCP)} \qquad \begin{array}{rl} \text{Find} & x \in \mathbb{R}^n \\ \text{such that} & x \geq 0,\ y \geq 0,\ x^\top y = 0, \\ & y = Mx + q. \end{array}$$

(iv) For the nonlinear second-order cone program (NSOCP)

$$\begin{array}{rl} \text{Minimize} & \theta(z) \\ \text{subject to} & G(z) \in \mathcal{K}, \;\; H(z) = 0 \end{array} \tag{2.4}$$

with $\theta : \mathbb{R}^{\ell_1} \to \mathbb{R}$, $G : \mathbb{R}^{\ell_1} \to \mathbb{R}^n$ and $H : \mathbb{R}^{\ell_1} \to \mathbb{R}^{\ell_2}$, its KKT conditions are given as follows:

$$\text{(SOCP-KKT)} \qquad \begin{array}{rl} \text{Find} & (x, y, z, w) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{\ell_1} \times \mathbb{R}^{\ell_2} \\ \text{such that} & x \in \mathcal{K}, \; y \in \mathcal{K}, \; x^\top y = 0, \\ & y = G(z), \quad H(z) = 0, \\ & \nabla\theta(z) - \nabla G(z)x - \nabla H(z)w = 0. \end{array}$$

This is of the form MNSOCCP (2.1).

(v) When $\mathcal{K} = \mathcal{K}^n$ and $\ell = 0$, MNSOCCP (2.1) reduces to the following non-mixed nonlinear SOCCP with a single cone:

$$\begin{array}{rl} \text{Find} & x \in \mathbb{R}^n \\ \text{such that} & x \in \mathcal{K}^n, \; y \in \mathcal{K}^n, \; x^\top y = 0, \\ & y = F_1(x). \end{array} \tag{2.5}$$

In [4], the regularized smoothing Newton algorithm was proposed for solving SOCCP (2.5), and the global and quadratic convergence was proved. In this section, we extend the algorithm proposed in [4] to MNSOCCP (2.1) in a direct manner. Then, in the next section, we prove that the generated sequence is globally convergent by using the sequential injective approach. In what follows, we often leave the detailed mathematical discussions on SOCs and related functions to the existing references [3, 4] since we use similar techniques.

2.1 Reformulation as a vector equation

Let x and y be partitioned according to the Cartesian structure of $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}$, i.e.,

$$\begin{array}{rcl} x & = & (x^1, \ldots, x^m) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}, \\ y & = & (y^1, \ldots, y^m) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}. \end{array} \tag{2.6}$$

Define function $\Phi_{\text{NR}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, called a *natural residual* [3, 4], by

$$\begin{array}{rcl} \Phi_{\text{NR}}(x, y) & := & \begin{pmatrix} \varphi_{\text{NR}}(x^1, y^1) \\ \vdots \\ \varphi_{\text{NR}}(x^m, y^m) \end{pmatrix}, \\ \varphi_{\text{NR}}(x^i, y^i) & := & x^i - P_{\mathcal{K}^{n_i}}(x^i - y^i), \end{array}$$

where $P_{\mathcal{K}^{n_i}}(x^i - y^i)$ denotes the Euclidean projection of $x^i - y^i$ onto $\mathcal{K}^{n_i}$. Note that, when $n_i = 1$, we have $\varphi_{\text{NR}}(x^i, y^i) = \min(x^i, y^i)$ since $\mathcal{K}^1 = \mathbb{R}_+$ yields

$$\begin{array}{rcl} x^i - P_{\mathcal{K}^1}(x^i - y^i) & = & x^i - \max(0, x^i - y^i) \\ & = & \min(x^i, y^i). \end{array}$$

It is known that the natural residual $\Phi_{\rm NR}$ satisfies

$$\Phi_{\rm NR}(x,y)=0 \iff x\in\mathcal{K},\ y\in\mathcal{K},\ x^\top y=0.$$

Therefore, letting $H_{\rm NR}:\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^\ell\to\mathbb{R}^{2n+\ell}$ be defined by

$$H_{\rm NR}(x,y,p) \ := \ \begin{pmatrix}\Phi_{\rm NR}(x,y)\\ F_1(x,p)-y\\ F_2(x,p)\end{pmatrix},$$

we can reformulate MNSOCCP (2.1) as the following VE equivalently:

$$H_{\rm NR}(x,y,p)=0. \tag{2.7}$$

2.2 Smoothing and regularization

Since MNSOCCP (2.1) is equivalent to (2.7), we have only to solve (2.7) instead of MNSOCCP (2.1). However, function $\Phi_{\rm NR}$ is nondifferentiable, and hence the Newton based method cannot be applied directly. Moreover, even if function $\Phi_{\rm NR}$ is smoothened, its Jacobian matrix may become singular. To overcome those difficulties, we introduce the smoothing method and the regularization method.

Smoothing method

A function Φ_μ parameterized by $\mu\ge 0$ is called a *smoothing function* of $\Phi_{\rm NR}$ if it satisfies the following conditions:

- For any fixed $\mu>0$, Φ_μ is continuously differentiable over $(x,y)\in\mathbb{R}^n\times\mathbb{R}^n$.
- $\lim_{\mu\searrow 0}\Phi_\mu(x,y)=\Phi_{\rm NR}(x,y)$ for any fixed $(x,y)\in\mathbb{R}^n\times\mathbb{R}^n$.

In the smoothing method, we handle Φ_μ instead of $\Phi_{\rm NR}$ with letting $\mu\searrow 0$. This is the basic idea of smoothing method.

In [4], the smoothing function Φ_μ is composed as follows. Consider a continuously differentiable convex function $\hat g:\Re\to\Re$ such that

$$\lim_{\alpha\to-\infty}\hat g(\alpha)=0,\ \ \lim_{\alpha\to\infty}(\hat g(\alpha)-\alpha)=0,\ \ 0<\hat g'(\alpha)<1. \tag{2.8}$$

For example, $\hat g_1(\alpha)=(\sqrt{\alpha^2+4}+\alpha)/2$ and $\hat g_2(\alpha)=\ln(e^\alpha+1)$ satisfy (2.8). Then, we can easily see that $\lim_{\mu\searrow 0}\mu\hat g(\alpha/\mu)=\max\{0,\alpha\}$ for any $\alpha\in\mathbb{R}$. By using this fact, Φ_μ is defined by

$$\Phi_\mu(x,y) \ := \ \begin{pmatrix}\varphi_\mu(x^1,y^1)\\ \vdots\\ \varphi_\mu(x^m,y^m)\end{pmatrix},$$

where

$$\begin{aligned}\varphi_\mu(x^i,y^i) \ &:= \ x^i-P_\mu(x^i-y^i),\\ P_\mu(z) \ &:= \ \begin{cases}\mu\hat g(\lambda_1/\mu)u^{\{1\}}+\mu\hat g(\lambda_2/\mu)u^{\{2\}} & (\dim(z)\ge 2)\\ \mu\hat g(z/\mu) & (\dim(z)=1).\end{cases}\end{aligned} \tag{2.9}$$

In the definition of $P_\mu(z)$, λ_1 and λ_2 denotes the spectral values of z, and $u^{\{1\}}$ and $u^{\{2\}}$ denotes the spectral vectors of z. For more details of the spectral factorization, see [3, 4].

Regularization method

Let the functions $F_{1,\varepsilon} : \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}^n$ and $F_{2,\varepsilon} : \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}^\ell$ be defined by

$$\begin{aligned} F_{1,\varepsilon}(x,p) &:= F_1(x,p) + \varepsilon x, \\ F_{2,\varepsilon}(x,p) &:= F_2(x,p) + \varepsilon p, \end{aligned}$$

respectively, with a positive parameter ε. In general, functions $F_{1,\varepsilon}$ and $F_{2,\varepsilon}$ have better properties than F_1 and F_2 from the viewpoint of global convergence. For example, if $F = \binom{F_1}{F_2}$ is a P_0 function, then $\binom{F_{1,\varepsilon}}{F_{2,\varepsilon}}$ is a uniformly P function for any $\varepsilon > 0$.

2.3 Main algorithm

Embedding the smoothing and regularization parameters, we define a function $H_{\mu,\varepsilon} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}^{2n+\ell}$ by

$$H_{\mu,\varepsilon}(x,y,p) := \begin{pmatrix} \Phi_\mu(x,y) \\ F_{1,\varepsilon}(x,p) - y \\ F_{2,\varepsilon}(x,p) \end{pmatrix}. \tag{2.10}$$

Then, we solve the inequality $\|H_{\mu,\varepsilon}(x,y,p)\| \le \beta$ by Newton's method with letting $(\mu,\varepsilon,\beta) \searrow (0,0,0)$. This is the main idea of the regularized smoothing Newton algorithm.

Before providing the algorithm, we give some functions and its related property that will be used in the algorithm. Since the functions are important only for the local quadratic convergence, we omit the detailed explanation here. For more details, see [4].

Definition 2.2

(a) $\tilde{\lambda} : \mathbb{R}^n \to [0,+\infty)$ *is a function defined by*

$$\tilde{\lambda}(z) := \begin{cases} \min_{i \in \mathcal{I}(z)} |\lambda_i(z)| & (\mathcal{I}(z) \neq \emptyset) \\ 0 & (\mathcal{I}(z) = \emptyset), \end{cases} \tag{2.11}$$

where $\lambda_i(z)$ $(i = 1,2)$ *are the spectral values of* z, *and* $\mathcal{I}(z) \subseteq \{1,2\}$ *is the index set defined by* $\mathcal{I}(z) := \{i \mid \lambda_i(z) \neq 0\}$.

(b) *Choose any function* $\hat{g}$ *satisfying* (2.8). *Then,* $\overline{\mu} : \mathbb{R}^n \times \mathbb{R}^n \to [0,+\infty]$ *is an arbitrary function such that*

$$\left| \hat{g}'(\alpha/\mu) - \lim_{\mu \searrow 0} \hat{g}'(\alpha/\mu) \right| < \delta \qquad \forall \mu \in \big(0, \overline{\mu}(\alpha,\delta)\big), \tag{2.12}$$

for any fixed $\alpha \in \mathbb{R}$ *and* $\delta > 0$.

Proposition 2.1 [4, Prop. 4.12] *Let* $\hat{g}$ *be defined by* $\hat{g}(\alpha) = (\sqrt{\alpha^2+4} + \alpha)/2$, *which satisfies* (2.8). *Let* $\overline{\mu} : \mathbb{R}^n \times \mathbb{R}^n \to [0,+\infty]$ *be defined by*

$$\overline{\mu}(\alpha,\delta) := \begin{cases} +\infty & (\delta \ge 1/2 \ \text{ or } \ \alpha = 0) \\ \dfrac{1}{2}|\alpha|\sqrt{\delta} & (\delta < 1/2 \ \text{ and } \ \alpha \neq 0). \end{cases}$$

Then, $\overline{\mu}$ *satisfies the condition* (2.12).

Now, we are in the position to provide the regularized smoothing Newton algorithm for solving MNSOCCP (2.1). In what follows, we use the following notations for convenience:

$$w := \begin{pmatrix} x \\ y \\ p \end{pmatrix}, \quad w^{(k)} := \begin{pmatrix} x^{(k)} \\ y^{(k)} \\ p^{(k)} \end{pmatrix}.$$

Algorithm 2 (Regularized smoothing Newton algorithm)

Step 0 *Choose the parameters* η, $\rho \in (0,1)$, $\overline{\eta} \in (0,\eta]$, $\sigma \in (0,1/2)$, $\kappa > 0$ *and* $\hat{\kappa} > 0$. *Choose the initial values* $w^{(0)} \in \mathbb{R}^{2n+\ell}$ *and* $\beta_0 \in (0,\infty)$. *Let* $\mu_0 := \|H_{\rm NR}(w^{(0)})\|$ *and* $\varepsilon_0 := \|H_{\rm NR}(w^{(0)})\|$. *Set* $k := 0$.

Step 1 *Terminate if* $\|H_{\rm NR}(w^{(k)})\| = 0$.

Step 2

Step 2.0 *Set* $v^{(0)} := w^{(k)} \in \mathbb{R}^{2n+\ell}$ *and* $j := 0$.

Step 2.1 *Find a vector* $\hat{d}^{(j)} \in \mathbb{R}^{2n+\ell}$ *such that*

$$H_{\mu_k,\varepsilon_k}(v^{(j)}) + \nabla H_{\mu_k,\varepsilon_k}(v^{(j)})^\top \hat{d}^{(j)} = 0.$$

Step 2.2 *If* $\|H_{\mu_k,\varepsilon_k}(v^{(j)} + \hat{d}^{(j)})\| \le \beta_k$, *then let* $w^{(k+1)} := v^{(j)} + \hat{d}^{(j)}$ *and go to* Step 3. *Otherwise, go to* Step 2.3.

Step 2.3 *Find the smallest nonnegative integer* m *such that*

$$\|H_{\mu_k,\varepsilon_k}(v^{(j)} + \rho^m \hat{d}^{(j)})\|^2 \le (1 - 2\sigma\rho^m)\|H_{\mu_k,\varepsilon_k}(v^{(j)})\|^2.$$

Let $m_j := m$, $\tau_j := \rho^{m_j}$ *and* $v^{(j+1)} := v^{(j)} + \tau_j \hat{d}^{(j)}$.

Step 2.4 *If*

$$\|H_{\mu_k,\varepsilon_k}(v^{(j+1)})\| \le \beta_k, \tag{2.13}$$

then let $w^{(k+1)} := v^{(j+1)}$ *and go to* Step 3. *Otherwise, set* $j := j+1$ *and go back to* Step 2.1.

Step 3 *Update the parameters as follows*:

$$\begin{aligned}
\mu_{k+1} &:= \min\left\{\kappa\|H_{\rm NR}(w^{(k+1)})\|^2,\ \mu_0\overline{\eta}^{k+1},\ \overline{\mu}\Big(\tilde{\lambda}(x^{(k+1)} - y^{(k+1)}), \hat{\kappa}\|H_{\rm NR}(w^{(k+1)})\|\Big)\right\}, \\
\varepsilon_{k+1} &:= \min\left\{\kappa\|H_{\rm NR}(w^{(k+1)})\|^2, \varepsilon_0\overline{\eta}^{k+1}\right\}, \\
\beta_{k+1} &:= \beta_0\eta^{k+1}.
\end{aligned}$$

Set $k := k+1$. *Go back to* Step 1.

In the inner iteration steps 2.0 − 2.4, a damped Newton method seeks a point $w^{(k+1)}$ such that $\|H_{\mu_k,\varepsilon_k}(w^{(k+1)})\| \le \beta_k$. Especially, Step 2.3 is well-known Armijo's stepsize rule. We note that Algorithm 2 is well-defined in the sense that Steps 2.0 − 2.4 find $v^{(j+1)}$ satisfying (2.13) in a finite number of iterations for each k. (It can be proved easily as in [4].) In Step 3, $\tilde{\lambda}$ and $\overline{\mu}$ are defined as (2.11) and (2.12), respectively. This step specifies the updating rule of the parameters, where $\{\beta_k\}$, $\{\mu_k\}$ and $\{\varepsilon_k\}$ converge to 0 since $0 < \overline{\eta} \le \eta < 1$.

Remark As is shown in the next section, H_{μ_k,ε_k} is injective for each k since its Jacobian is nonsingular. Moreover, H_{μ_k,ε_k} converges to $H_{\rm NR}$ uniformly over any bounded set. Therefore, Algorithm 1 (sequential injective algorithm) can be regarded as a prototype of Algorithm 2. This also means that the convergence analysis for Algorithm 1 can be applied directly to Algorithm 2.

3 Global convergence under Cartesian P_0 assumption

In this section, we prove the global convergence of Algorithm 2 under Cartesian P_0 assumption.

3.1 Cartesian P_0 property

Let

$$\sigma \ := \ (\nu_1, \nu_2, \dots, \nu_r)^\top \in \mathbb{Z}^r \tag{3.1}$$

be an integer vector such that $\nu_i \geq 1$ for i and $\nu = \sum_{i=1}^r \nu_i$. Then, we first consider to decompose the vector $z \in \mathbb{R}^\nu$, matrix $M \in \mathbb{R}^{\nu\times\nu}$ and function $F : \mathbb{R}^\nu \to \mathbb{R}^\nu$ according to the component of σ as follows:

$$z = \begin{bmatrix} z^1 \\ z^2 \\ \vdots \\ z^r \end{bmatrix}, \quad M = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1r} \\ M_{21} & M_{22} & \cdots & M_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ M_{r1} & M_{r2} & \cdots & M_{rr} \end{bmatrix}, \quad F(z) = \begin{bmatrix} F^1(z) \\ F^2(z) \\ \vdots \\ F^r(z) \end{bmatrix},$$

where $z^i \in \mathbb{R}^{\nu_i}$, $M_{ij} \in \mathbb{R}^{\nu_i \times \nu_j}$ and $F^i : \mathbb{R}^\nu \to \mathbb{R}^{\nu_i}$. Then, we define the notion of Cartesian P_0 property, which is a natural extension of well-known P_0 property for NCP or MCP.

Definition 3.3 *Let $\sigma \in \mathbb{Z}^r$ be an integer vector with* (3.1)*. Then, we say that*

(i) *the matrix $M \in \mathbb{R}^{\nu\times\nu}$ satisfies the σ-Cartesian P_0 property, if there exists $i = i(z) \in \{1, 2, \dots, r\}$ such that*

$$(z^i)^\top (Mz)^i \geq 0 \quad \textit{and} \quad z^i \neq 0$$

for any $z \in \mathbb{R}^\nu \setminus \{0\}$;

(ii) *the function $F : \mathbb{R}^\nu \to \mathbb{R}^\nu$ satisfies the σ-Cartesian P_0 property, if there exists $i = i(x, y) \in \{1, 2, \dots, r\}$ such that*

$$(x^i - y^i)^\top (F^i(x) - F^i(y)) \geq 0 \quad \textit{and} \quad x^i \neq y^i$$

for any $(x, y) \in \mathbb{R}^\nu \times \mathbb{R}^\nu$.

Remark If $\sigma = (1, 1, \dots, 1)^\top \in \mathbb{R}^\nu$ and $r = \nu$, then the σ-Cartesian P_0 property coincides with the normal P_0 property. On the other hand, if $\sigma = \nu \in \mathbb{R}$ and $r = 1$, then the σ-Cartesian P_0 property coincides with the positive semidefiniteness and the monotonicity. Cartesian P property and uniform Cartesian P property can be defined in a similar way, but we omit the details here.

The following theorem gives the necessary condition for a given matrix M to have Cartesian P_0 property.

Theorem 3.2 *Let σ be given by* (3.1)*, and M be an arbitrary σ-Cartesian P_0 matrix. Let $D = \operatorname{diag}\{D_{ii}\}_{i=1}^r$ be an arbitrary positive definite block diagonal matrix, i.e., $D_{ii} \in \mathbb{R}^{\nu_i\times\nu_i} \succ 0$ for each i. Then, $M + D$ is nonsingular.*

Proof. Let z be a vector such that $(M+D)z=0$. Assume for contradiction that $z\neq 0$. Then, due to the σ-Cartesian P_0 property, there exists i such that $z^i\neq 0$ and $(z^i)^\top(Mz)^i\geq 0$. Thus we have

$$0=(z^i)^\top\big((M+D)z\big)^i=(z^i)^\top(Mz)^i+(z^i)^\top D_{ii}z^i\geq (z^i)^\top D_{ii}z^i,$$

where the first equality is due to $(M+D)z=0$. However, this contradicts the positive definiteness of D_{ii} and $z^i\neq 0$. Thus $M+D$ is nonsingular. ∎

Corollary 3.2 *Let M be an arbitrary P_0 matrix. Then, $M+D$ is nonsingular for any positive definite diagonal matrix D.*

3.2 Nonsingularity analyses for Jacobian matrix

Next we we analyze the nonsingularity of the Jacobian matrix $\nabla H_{\mu,\varepsilon}(x,y,p)\in\mathbb{R}^{(2n+\ell)\times(2n+\ell)}$ for any $\mu>0$. From the definition of $H_{\mu,\varepsilon}$, Φ_μ, $\hat{g}$, etc., $\nabla H_{\mu,\varepsilon}(x,y,p)$ can be calculated as

$$\nabla H_{\mu,\varepsilon}(x,y,p) \;=\; \begin{pmatrix} I-D_\mu(x,y) & \nabla_x F_1(x,p)+\varepsilon I & \nabla_x F_2(x,p) \\ D_\mu(x,y) & -I & 0 \\ 0 & \nabla_p F_1(x,p) & \nabla_p F_2(x,p)+\varepsilon I \end{pmatrix}, \tag{3.2}$$

where

$$D_\mu(x,y) \;:=\; \operatorname{diag}\left\{\nabla P_\mu(x^i-y^i)\right\}_{i=1}^m. \tag{3.3}$$

In (3.3), P_μ is defined by (2.9), and $\operatorname{diag}\left\{\nabla P_\mu(x^i-y^i)\right\}_{i=1}^m$ denotes the block diagonal matrix with entries $\nabla P_\mu(x^i-y^i)\in\mathbb{R}^{n_i\times n_i}$ $(i=1,\ldots,m)$. The explicit expression of $\nabla P_\mu(\cdot)$ is given in [4]. For this Jacobian function, we have the following property.

Proposition 3.2 [3] *Let $P_\mu:\mathbb{R}^{n_i}\to\mathbb{R}^{n_i}$ be defined by* (2.9)*. Then we have*

$$0\prec\nabla P_\mu(z)\prec I$$

for any $z\in\mathbb{R}^{n_i}$, where $A\prec B$ means the positive definiteness of $B-A$.

Corollary 3.3 *Suppose that Algorithm* 2 *is applied to LCP, NCP or MCP, i.e.,* (2.1) *with $n_i=1$ and $\mathcal{K}=\mathbb{R}^n_+$. Then, $D_\mu(x,y)$ in* (3.2) – (3.3) *becomes a diagonal matrix such that every diagonal entry belongs to* $(0,1)$.

By using this fact, we can prove the nonsingularity of $\nabla H_{\mu,\varepsilon}$ under Cartesian P_0 assumption. For the Cartesian structure of $\mathcal{K}$ as in (2.2), we set $\sigma\in\mathbb{Z}^{m+\ell}$ as

$$\sigma:=(n_1,\ldots,n_m,\mathbf{1}_\ell^\top)^\top\in\mathbb{Z}^{m+\ell}. \tag{3.4}$$

Moreover, let $F:\mathbb{R}^{n+\ell}\to\mathbb{R}^{n+\ell}$ be defined by

$$F(x,p):=\begin{pmatrix}F_1(x,p)\\ F_2(x,p)\end{pmatrix}. \tag{3.5}$$

Then, we have the following theorem.

Theorem 3.3 *Let $\sigma \in \mathbb{Z}^{m+\ell}$ and $F : \mathbb{R}^{n+\ell} \to \mathbb{R}^{n+\ell}$ be given by* (3.4) *and* (3.5)*, respectively. Suppose that F has a σ-Cartesian P_0 property. Then, the matrix $\nabla H_{\mu,\varepsilon}(x, y, p)$ given by* (3.2) *is nonsingular for any $\mu > 0$, $\varepsilon > 0$ and $(x, y, p) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^\ell$.*

Proof. Let $\xi := (\xi_x^\top, \xi_y^\top, \xi_p^\top)^\top$ be a vector satisfying $\nabla H_{\mu,\varepsilon}(x, y, p)\xi = 0$. Then, by (3.2), we have

$$(I - D_\mu)\xi_x + (\nabla_x F_1 + \varepsilon I)\xi_y + \nabla_x F_2 \xi_p = 0, \tag{3.6}$$

$$D_\mu \xi_x - \xi_y = 0, \tag{3.7}$$

$$\nabla_p F_1 \xi_y + (\nabla_p F_2 + \varepsilon I)\xi_p = 0. \tag{3.8}$$

From (3.7), we have $\xi_x = D_\mu^{-1}\xi_y$. Substituting this into (3.6) and (3.8), we have $(D_\mu^{-1} - I + \nabla_x F_1 + \varepsilon I)\xi_y + \nabla_x F_2 \xi_p = 0$ and $\nabla_p F_1 \xi_y + (\nabla_p F_2 + \varepsilon I)\xi_p = 0$, that is,

$$\begin{aligned} 0 &= \left(\begin{bmatrix} \nabla_x F_1 & \nabla_x F_2 \\ \nabla_p F_1 & \nabla_p F_2 \end{bmatrix} + \begin{bmatrix} D_\mu^{-1} - I + \varepsilon I & 0 \\ 0 & \varepsilon I \end{bmatrix} \right) \begin{bmatrix} \xi_y \\ \xi_p \end{bmatrix} \\ &= \left(\nabla F + \begin{bmatrix} \operatorname{diag}\left\{ \nabla P_\mu(x^i - y^i)^{-1} - I \right\}_{i=1}^m + \varepsilon I & 0 \\ 0 & \varepsilon I \end{bmatrix} \right) \begin{bmatrix} \xi_y \\ \xi_p \end{bmatrix}. \end{aligned} \tag{3.9}$$

Notice that $\nabla P_\mu(x^i - y^i)^{-1} - I \succ 0$ since $0 \prec \nabla P_\mu(x^i - y^i) \prec I$. Moreover, $\nabla F(x, p)$ is a σ-Cartesian P_0 matrix since F is a σ-Cartesian P_0 function. Hence, by Theorem 3.2, the matrix

$$\nabla F + \begin{bmatrix} \operatorname{diag}\left\{ \nabla P_\mu(x^i - y^i)^{-1} - I \right\}_{i=1}^m + \varepsilon I & 0 \\ 0 & \varepsilon I \end{bmatrix}$$

is nonsingular. By this together with (3.9), we have $\xi_y = 0$, $\xi_p = 0$ and $\xi_x = D_\mu^{-1}\xi_y = 0$, which implies the nonsingularity of $\nabla H_{\mu,\varepsilon}(x, y, p)$. ■

3.3 Global convergence

Finally, we show the global convergence of Algorithm 2.

Theorem 3.4 *Let $\sigma \in \mathbb{Z}^{m+\ell}$ and $F : \mathbb{R}^{n+\ell} \to \mathbb{R}^{n+\ell}$ be given by* (3.4) *and* (3.5)*, respectively. Suppose that* (a) *the solution set of MNSOCCP* (2.1) *is nonempty and bounded, and* (b) *F is a σ-Cartesian P_0 function. Then, any accumulation point of the sequence $\{w^k\}$ generated by Algorithm* 2 *solves MNSOCCP* (2.1)*.*

Proof. By the definition (2.10) of $H_{\mu,\varepsilon}$ and Theorem 3.3, we can easily see that the assumptions (i)–(iii) of Corollary 1.1 holds. Moreover, by the same argument in [4], we have $v^{(j+1)}$ satisfying (2.13) with a finite j, i.e., Assumption B holds. Hence, by Corollary 1.1, we obtain the result. ■

When Algorithm 2 is applied to MCP (2.3), we readily have the following corollary.

Corollary 3.4 *Let $F : \mathbb{R}^{n+\ell} \to \mathbb{R}^{n+\ell}$ be given by* (3.5)*. Suppose that* (a) *the solution set of MCP* (2.3) *is nonempty and bounded, and* (b) *F is a P_0 function. Then, any accumulation point of the sequence $\{w^k\}$ generated by Algorithm* 2 *solves MCP* (2.3)*.*

References

[1] Website of ReSNA, http://www.plan.civil.tohoku.ac.jp/opt/hayashi/ReSNA/

[2] Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer-Verlag, New York (2003)

[3] Fukushima, M., Luo, Z.Q., Tseng, P.: Smoothing functions for second-order cone complementarity problems. SIAM Journal on Optimization **12**, 436–460 (2001)

[4] Hayashi, S., Yamashita, N., Fukushima, M.: A combined smoothing and regularization method for monotone second-order cone complementarity problems. SIAM Journal on Optimization **15**, 593–615 (2005)

[5] Kong, L., Sun, J., Xiu, N.: A regularized smoothing Newton method for symmetric cone complementarity problems. SIAM Journal on Optimization **19**, 1028–1047 (2008)