A two steps Levenberg-Marquardt type method for weighted linear complementarity problems

Abstract: The weighted linear complementarity problem (denoted by WLCP) can be used for modeling a larger class of equilibrium problems in economics. In this paper we propose a two steps Levenberg-Marquardt type method to solve WLCP. The proposed method is globally convergent without requiring WLCP to be monotone. Moreover, under the local error bound condition, our method is cubically convergent which is faster than the existing Levenberg-Marquardt type method for WLCP. Numerical results are reported to confirm the advantage of the method.

Keywords: Weighted linear complementarity problems; Levenberg-Marquardt method; Cubic convergence

1 Introduction

The weighted linear complementarity problem (WLCP) was introduced by Potra [12] which is to find vectors $x \in \mathbb{R}^n, s \in \mathbb{R}^n, y \in \mathbb{R}^m$ such that

$$\text{(WLCP)} \quad x \geq 0, \quad s \geq 0, \quad Px + Qs + Ry = a, \quad xs = w. \tag{1.1}$$

Here $P \in \mathbb{R}^{(n+m)\times n}, Q \in \mathbb{R}^{(n+m)\times n}, R \in \mathbb{R}^{(n+m)\times m}$ are given matrices, $a \in \mathbb{R}^{n+m}$ is a given vector, $w \geq 0$ is a given weight vector (the data of the problem) and xs is the componentwise product of the vectors x and s. The significance of studying the WLCP lies in the fact that a lot of equilibrium problems in economics can be formulated in a natural way as WLCP [12]. Moreover, those formulations lend themselves to the development of highly efficient algorithms for solving the corresponding equilibrium problems. For example, Fisher market equilibrium problem, which can be written as a nonlinear CP, can also be formulated as a WLCP that can be efficiently solved by interior-point methods [12].

Since Potra introduced the notion of WLCP, various numerical algorithms have been proposed to solve this problem (e.g., [1, 5, 6, 7, 10, 11, 14, 15, 16, 17, 18, 19, 20]). Among them, the smoothing Newton-type method is one class of the most effective algorithms (e.g., [14, 15, 18, 19]). The main idea of this class of methods is to use a smoothing function to reformulate the WLCP concerned as a system of smooth nonlinear equations and then solve it by Newton method. Note that in smoothing Newton-type methods, to ensure Newton step be feasible, one usually requires that the WLCP is monotone. Moreover, to obtain the local quadratic convergence, these smoothing Newton-type method need the nonsingularity condition. Lately, by using a smooth weighted complementarity function, Tang and Zhou [17] reformulated the weighted nonlinear complementarity problem as a smooth nonlinear equation and proposed a Levenberg-Marquardt type method to solve it. Their algorithm is globally convergent without any additional condition, and it has the local quadratic convergence under the local error bound condition which is weaker than the nonsingularity condition used in smoothing Newton-type methods.

Recently, Fan [8, 9] studied the modified Levenberg-Marquardt method for nonlinear equations, in which not only a LM step but also an approximate LM step are computed at every iteration. The cubic convergence of the modified Levenberg-Marquardt method was proved under the local error bound condition. In this paper, based on the work [8, 9, 17], we propose a two steps Levenberg-Marquardt type method (denoted by TS-LMM) to solve the WLCP. We prove that TS-LMM is globally convergent without requiring the WLCP to be monotone. Moreover, under the local error bound condition, we show that TS-LMM is cubically convergent which is faster than the Levenberg-Marquardt type method studied in [17]. We also give some numerical results to confirm the advantage of TS-LMM.

This paper is organized as follows. In Sect. 2, we give a weighted complementarity function and use it to reformulate the WLCP as a smooth nonlinear equation. In Sect. 3, we present the two steps Levenberg-Marquardt type method and give its global convergence. In Sect. 4, we analyze the local cubic convergence of the method. Numerical results are reported in Sect. 5. Some conclusions are given in Sect. 6.

2 A weighted complementarity function

To equivalently reformulate the WLCP as a system of the nonlinear equations, we consider the following weighted complementarity function introduced in [17]:

$$\phi_\tau^c(a,b) := (a+b)^3 - \left(\sqrt{a^2+b^2+(\tau-2)ab+(4-\tau)c}\right)^3,\ \forall (a,b)\in\mathbb{R}^2, \tag{2.1}$$

where $c \geq 0$ is a constant and $\tau \in [0,4)$. The following lemma gives some nice properties of ϕ_τ^c whose proof can be found in [17].

Lemma 2.1 *Let ϕ_τ^c be defined by* (2.1)*. Then the following results hold.*
(a) ϕ_τ^c *satisfies*

$$\phi_\tau^c(a,b) = 0 \Longleftrightarrow a\geq 0,\ b\geq 0,\ ab=c.$$

(b) ϕ_τ^c *is continuously differentiable at any* $(a,b)\in\mathbb{R}^2$ *with*

$$\nabla\phi_\tau^c(a,b) = \begin{bmatrix} \frac{\partial\phi_\tau^c}{\partial a} \\ \frac{\partial\phi_\tau^c}{\partial b} \end{bmatrix},$$

where

$$\begin{aligned}\frac{\partial\phi_\tau^c}{\partial a} &= 3\big[(a+b)^2 - h_\tau^c(a,b)(a+(\tau/2-1)b)\big],\\ \frac{\partial\phi_\tau^c}{\partial b} &= 3\big[(a+b)^2 - h_\tau^c(a,b)(b+(\tau/2-1)a)\big],\end{aligned}$$

in which

$$h_\tau^c(a,b) := \sqrt{a^2+b^2+(\tau-2)ab+(4-\tau)c}.$$

Let $z := (x,s,y)$. By using ϕ_τ^c, we define the function $F(z) : \mathbb{R}^{2n+m} \to \mathbb{R}^{2n+m}$ as

$$F(z) := \begin{pmatrix} Px+Qs+Ry-a \\ \phi_\tau^{w_1}(x_1,s_1) \\ \vdots \\ \phi_\tau^{w_n}(x_n,s_n) \end{pmatrix}, \tag{2.2}$$

where $w = (w_1, ..., w_n)^T$ is the weight vector given in the WLCP. Then, due to Lemma 2.1, the function $F(z)$ is continuously differentiable at any $z \in \mathbb{R}^{2n+m}$ and solving the WLCP is equivalent to computing a solution of the nonlinear equations

$$F(z) = 0. \tag{2.3}$$

By Lemma 2.1 (b) and (2.2), the Jacobian of $F(z)$ is given as

$$J(z) = \begin{bmatrix} P & Q & R \\ \text{diag}\big(\frac{\partial\phi_\tau^{w_i}}{\partial x_i}\big) & \text{diag}\big(\frac{\partial\phi_\tau^{w_i}}{\partial s_i}\big) & 0 \end{bmatrix}, \tag{2.4}$$

where

$$\begin{aligned} \frac{\partial\phi_\tau^{w_i}}{\partial x_i} &= 3\big[(x_i+s_i)^2 - h_\tau^{w_i}(x_i,s_i)(x_i+(\tau/2-1)s_i)\big], \\ \frac{\partial\phi_\tau^{w_i}}{\partial s_i} &= 3\big[(x_i+s_i)^2 - h_\tau^{w_i}(x_i,s_i)(s_i+(\tau/2-1)x_i)\big], \end{aligned}$$

in which

$$h_\tau^{w_i}(x_i,s_i) := \sqrt{x_i^2+s_i^2+(\tau-2)x_is_i+(4-\tau)w_i}.$$

The following theorem shows that the Jacobian $J(z)$ satisfies the Lipschitz continuity.

Theorem 2.1 *Let $\tau \in (0,4)$. Then the Jacobian $J(z)$ given in (2.4) is Lipschitz continuous on any closed and convex set $N(z) := \{z \in \mathbb{R}^{2n+m} | \|z\| \le \varrho\}$ where $\varrho > 0$ a constant.*

Proof Obviously, we only need to prove that the gradient $\nabla\phi_\tau^c(a,b)$ given in Lemma 2.1 (b) is Lipschitz continuous on any closed and convex set $\Omega := \{(a,b) \in \mathbb{R}^2 | \|(a,b)\| \le \theta\}$ for any $\theta > 0$. We consider the following three cases.

Case 1. Assume that $c > 0$. Since $h_\tau^c(a,b)$ can be rewritten as

$$\begin{aligned} h_\tau^c(a,b) &= \sqrt{(a+(\tau/2-1)b)^2+\tau(1-\tau/4)b^2+(4-\tau)c} \\ &= \sqrt{(b+(\tau/2-1)a)^2+\tau(1-\tau/4)a^2+(4-\tau)c}, \end{aligned}$$

we have from $\tau \in (0,4)$ that $h_\tau^c(a,b) > 0$ for any $(a,b) \in \Omega$. Thus, ϕ_τ^c is twice continuously differentiable at any $(a,b) \in \Omega$ with

$$\nabla^2\phi_\tau^c(a,b) = \begin{bmatrix} \frac{\partial^2\phi_\tau^c}{\partial a^2} & \frac{\partial^2\phi_\tau^c}{\partial a\partial b} \\ \frac{\partial^2\phi_\tau^c}{\partial b\partial a} & \frac{\partial^2\phi_\tau^c}{\partial b^2} \end{bmatrix},$$

where

$$\begin{aligned}
\frac{\partial^2 \phi_\tau^c}{\partial a^2} &= 3\left\{2(a+b) - \left(\frac{(a+(\tau/2-1)b)^2}{h_\tau^c(a,b)} + h_\tau^c(a,b)\right)\right\},\\
\frac{\partial^2 \phi_\tau^c}{\partial b^2} &= 3\left\{2(a+b) - \left(\frac{(b+(\tau/2-1)a)^2}{h_\tau^c(a,b)} + h_\tau^c(a,b)\right)\right\},\\
\frac{\partial^2 \phi_\tau^c}{\partial a \partial b} &= \frac{\partial^2 \phi_\tau^c}{\partial b \partial a}\\
&= 3\left\{2(a+b) - \left(\frac{(a+(\tau/2-1)b)(b+(\tau/2-1)a)}{h_\tau^c(a,b)} + (\tau/2-1)h_\tau^c(a,b)\right)\right\}.
\end{aligned}$$

Since $h_\tau^c(a,b)$ is continuous on Ω, there exists a constant $\xi > 0$ such that $h_\tau^c(a,b) \le \xi$ for any $(a,b) \in \Omega$. It follows that

$$\left|\frac{(a+(\tau/2-1)b)^2}{h_\tau^c(a,b)}\right| \le h_\tau^c(a,b) \le \xi,$$

$$\left|\frac{(b+(\tau/2-1)a)^2}{h_\tau^c(a,b)}\right| \le h_\tau^c(a,b) \le \xi,$$

$$\left|\frac{(a+(\tau/2-1)b)(b+(\tau/2-1)a)}{h_\tau^c(a,b)}\right| \le h_\tau^c(a,b) \le \xi.$$

Thus there exists a constant $C > 0$ independent of $(a,b) \in \Omega$ such that

$$\|\nabla^2 \phi_\tau^c(a,b)\| \le C, \quad \forall\ (a,b) \in \Omega.$$

By Mean Value Theorem, we have that

$$\|\nabla \phi_\tau^c(a_1,b_1) - \nabla \phi_\tau^c(a_2,b_2)\| \le C\|(a_1,b_1) - (a_2,b_2)\|$$

holds for any $(a_1,b_1),(a_2,b_2) \in \Omega$ and prove the desired result.

Case 2. Assume that $c = 0$ and $(0,0) \notin \Omega$. Since $\tau \in (0,4)$ and

$$\begin{aligned}
h_\tau^0(a,b) &= \sqrt{(a+(\tau/2-1)b)^2 + \tau(1-\tau/4)b^2}\\
&= \sqrt{(b+(\tau/2-1)a)^2 + \tau(1-\tau/4)a^2},
\end{aligned}$$

we have $h_\tau^0(a,b) > 0$ for any $(a,b) \in \Omega$. Thus ϕ_τ^0 is twice continuously differentiable at any $(a,b) \in \Omega$. By following exactly the same steps as in Case 1, we can prove the desired result.

Case 3. Assume that $c = 0$ and $(0,0) \in \Omega$. Similarly as Case 1, we can prove that there exists a constant $\bar{C} > 0$ independent of (a,b) such that

$$\|\nabla^2 \phi_\tau^0(a,b)\| \le \bar{C}, \quad \forall\ (a,b) \neq (0,0) \in \Omega.$$

By [4, Lemma 2.6], we have

$$\|\nabla \phi_\tau^0(a_1,b_1) - \nabla \phi_\tau^0(a_2,b_2)\| \le \bar{C}\|(a_1,b_1) - (a_2,b_2)\| \tag{2.5}$$

holds for all $(a_1,b_1),(a_2,b_2) \in \Omega$ with $(0,0) \notin [(a_1,b_1),(a_2,b_2)]$. Moreover, since $\nabla \phi_\tau^0(0,0) = (0,0)$, the inequality (2.5) also holds in case $(a_1,b_1) = (a_2,b_2) = (0,0)$. Therefore, we can assume $(a_1,b_1) \neq (0,0) \in \Omega$. Since ϕ_τ^0 is continuously differentiable for all $(a,b) \in \mathbb{R}^2$ with $\nabla \phi_\tau^0(0,0) = (0,0)$, by using a continuity argument, we obtain that the inequality (2.5) remains true for all $(a_2,b_2) \in \Omega$. Thus, the inequality (2.5) holds for all $(a_1,b_1),(a_2,b_2) \in \Omega$ which proves the desired result. $\square$

3 A two steps Levenberg-Marquardt type method

Our two steps Levenberg-Marquardt type method (TS-LMM) is described as follows.

Algorithm TS-LMM: Choose constants $\mu > 0$, $\theta, \rho, \sigma \in (0,1)$ and $z_0 \in \mathbb{R}^n$. Set $k := 0$.

Step 1: If $\|J(z_k)^T F(z_k)\| = 0$, then stop.

Step 2: Choose $\delta \in [1,2]$ and set

$$\lambda_k := \mu \|F(z_k)\|^\delta. \tag{3.1}$$

Compute $\bar{d}_k \in \mathbb{R}^{2n+m}$ by solving

$$[J(z_k)^T J(z_k) + \lambda_k I]\bar{d}_k = -J(z_k)^T F(z_k). \tag{3.2}$$

Set $t_k := z_k + \bar{d}_k$. Then, compute $\hat{d}_k \in \mathbb{R}^{2n+m}$ by solving

$$[J(z_k)^T J(z_k) + \lambda_k I]\hat{d}_k = -J(z_k)^T F(t_k). \tag{3.3}$$

Step 3: If

$$\|F(z_k + \bar{d}_k + \hat{d}_k)\| \le \theta \|F(z_k)\|, \tag{3.4}$$

then set $z_{k+1} := z_k + \bar{d}_k + \hat{d}_k$. Otherwise, set $\alpha_k := \rho^{l_k}$, where l_k is the smallest nonnegative integer l satisfying

$$\|F(z_k + \rho^l \bar{d}_k)\|^2 \le \|F(z_k)\|^2 + \sigma \rho^l F(z_k)^T J(z_k) \bar{d}_k. \tag{3.5}$$

Set $z_{k+1} := z_k + \alpha_k \bar{d}_k$.

Step 4: Set $k := k+1$ and go to Step 1.

Remark 3.1 We denote the merit function

$$\psi(z) := \frac{1}{2}\|F(z)\|^2. \tag{3.6}$$

Then $\psi(z) = 0$ if and only if $z = (x, s, y)$ is a solution of the WLCP. Moreover, $\psi(z)$ is continuously differentiable at any $z \in \mathbb{R}^{2n+m}$ with its gradient

$$\nabla \psi(z) = J(z)^T F(z).$$

Note that $\bar{d}_k + \hat{d}_k$ is not a descent direction of the merit function ψ at z_k. However, the LM step $\bar{d}_k$ is a descent direction. In fact, if $J(z_k)^T F(z_k) \neq 0$ for some k, then $\bar{d}_k \neq 0$ by (3.2), and $F(z_k) \neq 0$ which yields $\lambda_k = \mu \|F(z_k)\|^\delta > 0$. Hence the matrix $J(z_k)^T J(z_k) + \lambda_k I$ is positive definite which gives

$$\begin{aligned} \nabla \psi(z_k)^T \bar{d}_k &= F(z_k)^T J(z_k) \bar{d}_k \\ &= -(\bar{d}_k)^T [J(z_k)^T J(z_k) + \lambda_k I] \bar{d}_k \\ &< 0. \end{aligned} \tag{3.7}$$

Thus in Step 3 we only use the LM step $\bar{d}_k$ so that the Armijo line search procedure can be carried out.

In the following, we assume $J(z_k)^T F(z_k) \neq 0$ for all $k \geq 0$ so that Algorithm TS-LMM generates an infinite sequence $\{z_k\}$. We have the following global convergence property.

Theorem 3.1 *Let $\{z_k\}$ be the iteration sequence generated by Algorithm* TS-LMM*. Then any accumulation point z^* of $\{z_k\}$ satisfies $\nabla\psi(z^*) = 0$.*

Proof We assume that z^* is the limit of the subsequence $\{z_k\}_{k\in K} \subset \{z_k\}$ where $K \subset \{0, 1, ...\}$. So, $\lim\limits_{(K\ni)k\to\infty} z_k = z^*$. By the continuity,

$$\lim_{(K\ni)k\to\infty} F(z_k) = F(z^*), \qquad \lim_{(K\ni)k\to\infty} J(z_k) = J(z^*),$$

and consequently,

$$\lim_{(K\ni)k\to\infty} \lambda_k = \mu\|F(x^*)\|^\delta := \lambda^*.$$

Now we assume $\nabla\psi(z^*) \neq 0$ and will derive a contradiction. For all $k \in K$, since $J(z_k)^T F(z_k) \neq 0$, $F(z_k) \neq 0$ and hence $\lambda_k = \mu\|F(z_k)\|^\delta > 0$. So, the matrix $J(z_k)^T J(z_k) + \lambda_k I$ is positive definite for any $k \in K$. Moreover,

$$\lim_{(K\ni)k\to\infty} [J(z_k)^T J(z_k) + \lambda_k I] = J(z^*)^T J(z^*) + \lambda^* I. \tag{3.8}$$

Since $\nabla\psi(z^*) = J(z^*)^T F(z^*) \neq 0$, $F(z^*) \neq 0$ and so $\lambda^* > 0$. Thus the matrix $J(z^*)^T J(z^*) + \lambda^* I$ is also positive definite. By (3.8) we have

$$\lim_{(K\ni)k\to\infty} [J(z_k)^T J(z_k) + \lambda_k I]^{-1} = [J(x^*)^T J(x^*) + \lambda^* I]^{-1},$$

which together with (3.2) yields

$$\lim_{(K\ni)k\to\infty} \bar{d}_k = -[J(x^*)^T J(x^*) + \lambda^* I]^{-1} J(z^*)^T F(z^*) := d^*. \tag{3.9}$$

By (3.4), (3.5) and (3.7), we have that $\|F(z_{k+1})\| < \|F(z_k)\|$ for all $k \geq 0$. This indicates that there exists a constant $F^* \geq 0$ such that $\lim\limits_{k\to\infty} \|F(z_k)\| = F^*$. Obviously, $F^* = F(z^*)$. Thus, if there are infinitely many k for which the condition (3.4) holds, then $\|F(z_{k+1})\| \leq \theta\|F(z_k)\|$ holds for infinitely many k. This gives $F^* \leq \theta F^*$ and so $F^* = 0$ because $\theta \in (0, 1)$. It follows that $J(z^*)^T F(z^*) = 0$ which contradicts our assumption. Thus there exists an index $\bar{k} > 0$ such that α_k is determined by the Armijo line search (3.5) for all $k \geq \bar{k}$. Now we show that $F(z^*)^T J(z^*) d^* = 0$. We divide the proof into the following two parts.

Part 1. $\alpha_k \geq c > 0$ for all $k \in K$ and $k \geq \bar{k}$ where c is a fixed constant. In this case, it follows from (3.5) and (3.7) that for all $k \in K$ and $k \geq \bar{k}$.

$$0 \leq -\sigma c F(z_k)^T J(z_k) \bar{d}_k \leq -\sigma\alpha_k F(z_k)^T J(z_k) \bar{d}_k \leq \|F(z_k)\|^2 - \|F(z_{k+1})\|^2. \tag{3.10}$$

Since $\lim\limits_{k\to\infty} \|F(z_k)\| = F^*$, by letting $k \to \infty$ with $k \in K$ in (3.10), we have $F(z^*)^T J(z^*) d^* = 0$.

Part 2. $\{\alpha_k\}_{k\in K}$ has a subsequence converging to zero. We may pass to the subsequence and assume that $\lim\limits_{(K\ni)k\to\infty} \alpha_k = 0$. Let $\hat{\alpha}_k := \rho^{-1}\alpha_k$. Then $\lim\limits_{(\bar{K}\ni)k\to\infty} \hat{\alpha}_k = 0$. Moreover, by (3.5), for any sufficiently large $k \in K$,

$$\|F(z_k + \hat{\alpha}_k \bar{d}_k)\|^2 - \|F(z_k)\|^2 > \sigma\hat{\alpha}_k F(z_k)^T J(z_k)\bar{d}_k,$$

which gives

$$\frac{\psi(z_k + \hat{\alpha}_k \bar{d}_k) - \psi(z_k)}{\hat{\alpha}_k} > \frac{\sigma}{2} F(z_k)^T J(z_k)\bar{d}_k, \tag{3.11}$$

where ψ is defined by (3.6). Since ψ is continuously differentiable at any $z \in \mathbb{R}^{2n+m}$, by letting $k \to \infty$ with $k \in K$ in (3.11), we have $\nabla\psi(z^*)^T d^* = F(z^*)^T J(z^*) d^* \geq \frac{\sigma}{2} F(z^*)^T J(z^*) d^*$, i.e., $(1 - \frac{\sigma}{2}) F(z^*)^T J(z^*) d^* \geq 0$. This together with $\sigma \in (0, 1)$ gives $F(z^*)^T J(z^*) d^* \geq 0$. On the other hand, by (3.7) it holds that $F(z^*)^T J(z^*) d^* \leq 0$. Thus $F(z^*)^T J(z^*) d^* = 0$.

By $F(z^*)^T J(z^*) d^* = 0$ and (3.9), we have

$$(d^*)^T [J(z^*)^T J(z^*) + \lambda^* I] d^* = 0.$$

Since the matrix $J(z^*)^T J(z^*) + \lambda^* I$ is positive definite, we have $d^* = 0$ which together (3.9) gives $J(z^*)^T F(z^*) = 0$. A contradiction is derived. We complete the proof. $\square$

Corollary 3.1 *Let z^* be any accumulation point of $\{z_k\}$ generated by Algorithm* TS-LMM. *If $J(z^*)$ is nonsingular, then $F(z^*) = 0$.*

4 Cubic convergence rate

In this section, we assume that the solution set Z of the nonlinear equations (2.3) is nonempty. We further suppose that the sequence $\{z_k\}$ generated by Algorithm TS-LMM converges to a point $z^* \in Z$ and lies in some neighbourhood of z^*. We make the following assumptions.

Assumption A (a) $F(z)$ provides a local error bound in some neighbourhood of z^*, i.e. there exist constants $\kappa > 0$ and $0 < \varepsilon < 1$ such that

$$\kappa \text{dist}(z, Z) \leq \|F(z)\|, \quad \forall\, z \in N(z^*, \varepsilon), \tag{4.1}$$

where $N(z^*, \varepsilon) := \{z \in \mathbb{R}^{2n+m} | \|z - z^*\| \leq \varepsilon\}$.
(b) The Jacobian $J(z)$ is Lipschitz continuous on $N(z^*, \varepsilon)$, i.e. there exists a constant $M > 0$ such that

$$\|J(u) - J(v)\| \leq M\|u - v\|, \ \forall\, u, v \in N(z^*, \varepsilon). \tag{4.2}$$

Note that by Theorem 2.1, the condition (b) in Assumption A holds for $\tau \in (0, 4)$. By (4.2), we have (see [8])

$$\|F(v) - F(u) - J(u)(v - u)\| \leq M\|v - u\|^2, \ \forall\, u, v \in N(z^*, \varepsilon), \tag{4.3}$$

and there exists a constant $L > 0$ such that

$$\|F(v) - F(u)\| \le L\|v - u\|, \ \forall \ u, v \in N(z^*, \varepsilon). \tag{4.4}$$

In the following, we denote $\bar{z}_k$ as the vector in Z that satisfies

$$\|\bar{z}_k - z_k\| = \mathrm{dist}(z_k, Z).$$

Now we suppose the singular value decomposition (SVD) of $J(\bar{z}_k)$ is

$$\begin{aligned} J(\bar{z}_k) &= \bar{U}_k \bar{\Sigma}_k \bar{V}_k^T \\ &= (\bar{U}_{k,1}, \bar{U}_{k,2}) \begin{pmatrix} \bar{\Sigma}_{k,1} & \\ & 0 \end{pmatrix} \begin{pmatrix} \bar{V}_{k,1}^T \\ \bar{V}_{k,2}^T \end{pmatrix} \\ &= \bar{U}_{k,1} \bar{\Sigma}_{k,1} \bar{V}_{k,1}^T, \end{aligned}$$

where $\bar{\Sigma}_{k,1} = \mathrm{diag}(\bar{\sigma}_{k,1}, ..., \bar{\sigma}_{k,r})$ with $\bar{\sigma}_{k,1} \ge \cdots \ge \bar{\sigma}_{k,r} > 0$. The corresponding SVD of $J(z_k)$ is

$$\begin{aligned} J(z_k) &= U_k \Sigma_k V_k^T \\ &= (U_{k,1}, U_{k,2}) \begin{pmatrix} \Sigma_{k,1} & \\ & \Sigma_{k,2} \end{pmatrix} \begin{pmatrix} V_{k,1}^T \\ V_{k,2}^T \end{pmatrix} \\ &= U_{k,1} \Sigma_{k,1} V_{k,1}^T + U_{k,2} \Sigma_{k,2} V_{k,2}^T, \end{aligned}$$

where $\Sigma_{k,1} = \mathrm{diag}(\sigma_{k,1}, ..., \sigma_{k,r})$ with $\sigma_{k,1} \ge \cdots \ge \sigma_{k,r} > 0$ and $\Sigma_{k,2} = \mathrm{diag}(\sigma_{k,r+1}, ..., \sigma_{k,n})$ with $\sigma_{k,r+1} \ge \cdots \ge \sigma_{k,n} \ge 0$. In the following, if the context is clear, we neglect the subscript k in $U_{k,i}, \Sigma_{k,i}, V_{k,i} (i = 1, 2)$ and write $J(\bar{z}_k)$ and $J(z_k)$ as

$$J(\bar{z}_k) = \bar{U}_1 \bar{\Sigma}_1 \bar{V}_1^T, \quad J(z_k) = U_1 \Sigma_1 V_1^T + U_2 \Sigma_2 V_2^T.$$

By the matrix perturbation theory [13] and (4.2), we have

$$\|\mathrm{diag}(\Sigma_1 - \bar{\Sigma}_1, \Sigma_2)\| \le \|J(z_k) - J(\bar{z}_k)\| \le M\|\bar{z}_k - z_k\|,$$

which gives

$$\|\Sigma_1 - \bar{\Sigma}_1\| \le M\|\bar{z}_k - z_k\|, \quad \|\Sigma_2\| \le M\|\bar{z}_k - z_k\|. \tag{4.5}$$

Lemma 4.1 *Under Assumption* A*, there exists a constant* $c_1 > 0$ *such that for all sufficiently large* k*,*

$$\|\bar{d}_k\| \le c_1 \|\bar{z}_k - z_k\|. \tag{4.6}$$

Proof By (3.1) and(4.1), for all sufficiently large k,

$$\lambda_k = \mu\|F(z_k)\|^\delta \ge \mu\kappa^\delta \|\bar{z}_k - z_k\|^\delta. \tag{4.7}$$

Since

$$\|\bar{z}_k - z^*\| \le \|\bar{z}_k - z_k\| + \|z_k - z^*\| \le 2\|z_k - z^*\|,$$

we have $\bar{z}_k \in N(z^*, \varepsilon)$ for all sufficiently large k. Then, by (4.3) and $F(\bar{z}_k) = 0$, for all sufficiently large k,

$$\|F(z_k) + J(z_k)(\bar{z}_k - z_k)\| \le M\|\bar{z}_k - z_k\|^2. \tag{4.8}$$

For any $k \ge 0$, we consider the following minimization problem:

$$\min_{d \in \mathbb{R}^{2n+m}} \varphi_k(d) := \|F(z_k) + J(z_k)d\|^2 + \lambda_k\|d\|^2. \tag{4.9}$$

Then the LM step $\bar{d}_k$ is a solution of (4.9). By (4.7) and (4.8), for all sufficiently large k,

$$\begin{aligned}
\|\bar{d}_k\|^2 &\le \frac{\varphi_k(\bar{d}_k)}{\lambda_k} \\
&\le \frac{\varphi_k(\bar{z}_k - z_k)}{\lambda_k} \\
&= \frac{\|F(z_k) + J(z_k)(\bar{z}_k - z_k)\|^2}{\lambda_k} + \|\bar{z}_k - z_k\|^2 \\
&\le M^2\mu^{-1}\kappa^{-\delta}\|\bar{z}_k - z_k\|^{4-\delta} + \|\bar{z}_k - z_k\|^2 \\
&\le M^2\mu^{-1}\kappa^{-\delta}\|\bar{z}_k - z_k\|^2.
\end{aligned}$$

Letting $c_1 := \sqrt{M^2\mu^{-1}\kappa^{-\delta} + 1}$, we have (4.6). □

Lemma 4.2 *Under Assumption* A*, for all sufficiently large k,*
(a) $\|U_1U_1^TF(z_k)\| \le L\|\bar{z}_k - z_k\|$;
(b) $\|U_2U_2^TF(z_k)\| \le 2M\|\bar{z}_k - z_k\|^2$;
(c) $\|U_1^TU_1F(t_k)\| \le c_2\|\bar{z}_k - z_k\|^2$;
(d) $\|U_2U_2^TF(t_k)\| \le c_3\|\bar{z}_k - z_k\|^3$,
where c_2, c_3 are positive constants.

Proof The results can be found in [8, Lemma 3.4 and Lemma 3.5]. □

Lemma 4.3 *Under Assumption* A*, there exists a constant $c_4 > 0$ such that for all sufficiently large k,*

$$\|F(z_k) + J(z_k)\bar{d}_k\| \le c_4\|\bar{z}_k - z_k\|^2. \tag{4.10}$$

Proof By (3.2) and the SVD of $J(z_k)$, we have

$$\begin{aligned}
\bar{d}_k &= [J(z_k)^TJ(z_k) + \lambda_kI]^{-1}J(z_k)^TF(z_k) \\
&= -V_1(\Sigma_1^2 + \lambda_kI)^{-1}\Sigma_1U_1^TF(z_k) - V_2(\Sigma_2^2 + \lambda_kI)^{-1}\Sigma_2U_2^TF(z_k),
\end{aligned} \tag{4.11}$$

and

$$\begin{aligned}
& F(z_k) + J(z_k)\bar{d}_k \\
=\; & F(z_k) - U_1\Sigma_1(\Sigma_1^2 + \lambda_kI)^{-1}\Sigma_1U_1^TF(z_k) - U_2\Sigma_2(\Sigma_2^2 + \lambda_kI)^{-1}\Sigma_2U_2^TF(z_k) \\
=\; & \lambda_kU_1(\Sigma_1^2 + \lambda_kI)^{-1}U_1^TF(z_k) + \lambda_kU_2(\Sigma_2^2 + \lambda_kI)^{-1}U_2^TF(z_k).
\end{aligned} \tag{4.12}$$

Since $\{z_k\}$ converges to z^*, without loss of generality, we assume that $M\|\bar{z}_k - z_k\| \leq \bar{\sigma}/2$ holds for all sufficiently large k. Then it follows from (4.5) that

$$\|(\Sigma_1^2 + \lambda_k I)^{-1}\| \leq \|\Sigma_1^{-2}\| \leq \frac{1}{(\bar{\sigma} - M\|\bar{z}_k - z_k\|)^2} \leq \frac{4}{\bar{\sigma}^2}. \tag{4.13}$$

Moreover, by (3.1) and (4.4), for all sufficiently large k,

$$\lambda_k = \mu\|F(z_k)\|^\delta = \mu\|F(z_k) - F(\bar{z}_k)\|^\delta \leq \mu L^\delta \|\bar{z}_k - z_k\|^\delta. \tag{4.14}$$

Therefore, by Lemma 4.2, (4.12), (4.13), (4.14) and $\|(\Sigma_2^2 + \lambda_k I)^{-1}\| \leq \lambda_k^{-1}$, for all sufficiently large k,

$$\begin{aligned}\|F(z_k) + J(z_k)\bar{d}_k\| &\leq \lambda_k\|(\Sigma_1^2 + \lambda_k I)^{-1}\|\|U_1^T F(z_k)\| + \|U_2^T F(z_k)\| \\ &\leq \frac{4\mu L^{1+\delta}}{\bar{\sigma}^2}\|\bar{z}_k - z_k\|^{1+\delta} + 2M\|\bar{z}_k - z_k\|^2 \\ &\leq \left(\frac{4\mu L^{1+\delta}}{\bar{\sigma}^2} + 2M\right)\|\bar{z}_k - z_k\|^2. \end{aligned} \tag{4.15}$$

Letting $c_4 := 4\mu L^{1+\delta}/\bar{\sigma}^2 + 2M$, we have (4.10). □

Lemma 4.4 *Under Assumption* A*, for all sufficiently large* k*,*

$$z_{k+1} = z_k + \bar{d}_k + \hat{d}_k. \tag{4.16}$$

Proof By (3.3) and the SVD of $J(z_k)$, we have

$$\begin{aligned}\hat{d}_k &= [J(z_k)^T J(z_k) + \lambda_k I]^{-1} J(z_k)^T F(t_k) \\ &= -V_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F(t_k) - V_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F(t_k),\end{aligned}$$

and

$$\begin{aligned}&F(t_k) + J(z_k)\hat{d}_k \\ =\ & F(t_k) - U_1\Sigma_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F(t_k) - U_2\Sigma_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F(t_k) \\ =\ & \lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1} U_1^T F(t_k) + \lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1} U_2^T F(t_k),\end{aligned}$$

which together with (4.5), (4.7), (4.13), (4.14), Lemma 4.2 and $\|(\Sigma_2^2 + \lambda_k I)^{-1}\| \leq \lambda_k^{-1}$ implies that for all sufficiently large k,

$$\begin{aligned}\|\hat{d}_k\| &\leq \|(\Sigma_1^2 + \lambda_k I)^{-1}\|\|U_1^T F(t_k)\| + \|\lambda_k^{-1}\Sigma_2\|\|U_2^T F(t_k)\| \\ &\leq \frac{4c_2}{\bar{\sigma}^2}\|\bar{z}_k - z_k\|^2 + \frac{Mc_3}{\mu\kappa^\delta}\|\bar{z}_k - z_k\|^{4-\delta} \\ &\leq c_5\|\bar{z}_k - z_k\|^2,\end{aligned} \tag{4.17}$$

and

$$\|F(t_k) + J(z_k)\hat{d}_k\| \quad \leq \quad \lambda_k\|(\Sigma_1^2 + \lambda_k I)^{-1}\|\|U_1^T F(t_k)\| + \|U_2^T F(t_k)\|$$

$$
\begin{aligned}
&\le \frac{4\mu L^\delta c_2}{\bar\sigma^2}\|\bar z_k - z_k\|^{2+\delta} + c_3\|\bar z_k - z_k\|^3 \\
&\le c_6\|\bar z_k - z_k\|^3, \qquad (4.18)
\end{aligned}
$$

where $c_5 := 4c_2/\bar\sigma^2 + Mc_3/\mu\kappa^\delta$ and $c_6 := 4\mu L^\delta c_2/\bar\sigma^2 + c_3$. Therefore, by (4.1), (4.2), (4.3), (4.6), (4.17) and (4.18), for all sufficiently large k,

$$
\begin{aligned}
\|F(z_k + \bar d_k + \hat d_k)\| &= \|F(t_k + \hat d_k)\| \\
&\le \|F(t_k + \hat d_k) - F(t_k) - J(t_k)\hat d_k\| \\
&\quad + \|F(t_k) + J(z_k)\hat d_k\| + \|(J(t_k) - J(z_k))\hat d_k\| \\
&\le M\|\hat d_k\|^2 + \|F(t_k) + J(z_k)\hat d_k\| + M\|\bar d_k\|\|\hat d_k\| \\
&\le Mc_5^2\|\bar z_k - z_k\|^4 + c_6\|\bar z_k - z_k\|^3 + Mc_1c_5\|\bar z_k - z_k\|^3 \\
&\le (Mc_5^2 + c_6 + Mc_1c_5)\mathrm{dist}(z_k, Z^*)^3 \\
&\le \frac{Mc_5^2 + c_6 + Mc_1c_5}{\kappa^3}\|F(z_k)\|^3.
\end{aligned}
$$

This implies that

$$\|F(z_k + \bar d_k + \hat d_k)\| \le \theta\|F(z_k)\|$$

holds for all sufficiently large k. Hence, by Step 3 of Algorithm TS-LMM, (4.16) holds for all sufficiently large k. We complete the proof.

Lemma 4.4 indicates that when k is sufficiently large, Algorithm TS-LMM becomes the modified Levenberg-Marquardt method without line search. Thus, by following the arguments in [8], we obtain the cubic convergence of Algorithm TS-LMM as follows.

Theorem 4.1 *Under Assumption* A, *the sequence* $\{z^k\}$ *converges to* z^* *locally cubically.*

5 Numerical experiments

We would like to perform some numerical experiments in this section. We compare the numerical performances of the following two algorithms.

- Algorithm TS-LMM studied in this paper, denoted by **TS-LMM**;
- Levenberg-Marquardt type method studied by Tang and Zhou [17], denoted by **LMM**.

We test the WLCP (1.1) with

$$
P = \begin{pmatrix} A \\ M \end{pmatrix}, \quad Q = \begin{pmatrix} 0 \\ -I \end{pmatrix}, \quad R = \begin{pmatrix} 0 \\ -A^T \end{pmatrix}, \quad a = \begin{pmatrix} b \\ -f \end{pmatrix},
$$

where $A \in \mathbb{R}^{m\times n}$ is a full row rank matrix with $m < n$, $b \in \mathbb{R}^m$, $f \in \mathbb{R}^n$ and $M \in \mathbb{R}^{n\times n}$ is an $n \times n$ symmetric positive semidefinite matrix. This WLCP is the optimality conditions of the quadratic programming and weighted centering problem [12, Theorem 2.1]. In the experiments, we generate a random matrix $A \in \mathbb{R}^{m\times n}$ with full row rank and set $M = BB^T/\|BB^T\|$ with $B = \mathbf{rand}(n, n)$. Then we choose $\hat x = \mathbf{rand}(n, 1)$, $f = \mathbf{rand}(n, 1)$ and set $b := A\hat x$, $\hat s := M\hat x + f$

and $w := \hat{x}\hat{s}$. The parameters used in Algorithm TS-LMM are chosen as $\mu = 10^{-5},\ \sigma = 10^{-6},\ \rho = 0.8,\ \delta = 1,\ \theta = 0.5$.

First, to observe the local convergence behavior of **TS-LMM**, we generate one test problem with $n = 500$ and $m = 250$ and solve it by **TS-LMM** and **LMM** respectively. The starting point is chosen as $x_0 = s_0 = (1, ..., 1)^T$ and $y_0 = (0, ..., 0)^T$. Table 1 gives the value of $\|F(z_k)\|$ at the k-th iteration in which τ is the parameter value used in the function ϕ_τ^c. From Table 1, we can clearly see that **TS-LMM** has at least super-quadratic convergence rate and it converges faster than **LMM** does.

Table 1 The value of $\|F(z_k)\|$ at the k-th iteration

		TS-LMM	**LMM**
$\tau = 0$	$k = 1$	13.1456	28.1305
	$k = 2$	0.8915	13.0428
	$k = 3$	0.0333	3.7961
	$k = 4$	4.3500e-05	0.2935
	$k = 5$	3.0706e-13	0.0063
	$k = 6$	0	5.7924e-06
	$k = 7$	0	3.5507e-11
$\tau = 2$	$k = 1$	7.5373	17.6283
	$k = 2$	0.5989	6.4880
	$k = 3$	0.0273	2.3709
	$k = 4$	8.2111e-05	0.2565
	$k = 5$	8.0648e-12	0.0090
	$k = 6$	0	3.3340e-05
	$k = 7$	0	8.6151e-10

Next, for each problem with different sizes $n(= 2m)$, we generate 10 instances and solve them by **TS-LMM** and **LMM** respectively. The starting point is chosen as before. We use $\|F(z_k)\| < 10^{-8}$ as the stopping criterion. Numerical results are listed in Table 2 where **AIT** and **ACPU** denote the average number of iterations and the average CPU time in seconds respectively. From Table 2, we can see that **TS-LMM** always takes fewer iterations and often uses less CPU time than **LMM** to reach the stopping tolerance.

Table 2 Comparison of **TS-LMM** and **LMM**

		TS-LMM		**LMM**	
τ	n	**AIT**	**ACPU**	**AIT**	**ACPU**
0	1000	5.0	2.69	8.5	3.81
	1500	5.0	7.18	7.8	9.71
	2000	5.0	14.60	8.4	21.55
	2500	5.0	25.73	8.1	37.97
	3000	5.0	41.57	8.2	62.27
	3500	5.1	63.94	8.1	93.55
	4000	5.0	89.02	8.1	133.47
2	1000	5.1	2.76	8.1	3.23
	1500	5.2	7.53	8.1	10.43
	2000	5.2	15.35	8.3	21.47
	2500	5.1	21.74	8.3	43.24
	3000	5.3	43.85	9.4	76.48
	3500	5.2	62.50	9.0	102.33
	4000	5.6	95.08	9.1	144.82

6 Conclusions

In this paper, by using a smooth weighted complementarity function, we reformulated the WLCP as a smooth nonlinear equation and proposed a two steps Levenberg-Marquardt type method to solve it. We proved that the proposed method is globally convergent without requiring any additional condition, and it has cubic convergence rate under the local error bound condition. Numerical experiments indicated that our method is more efficient than the existing Levenberg-Marquardt type method for solving WLCP. Lately, many researchers studied high-order Levenberg-Marquardt methods for nonlinear equations (e.g.,[2, 3]). This class of methods computes the LM step and two approximate LM steps at every iteration. Under the local error bound condition, the convergence order of high-order Levenberg-Marquardt methods is fourth. As a future research issue, it is worth designing high-order Levenberg-Marquardt type methods for solving WLCP.

References

[1] S. Asadi, Z. Darvay, G. Lesaja, A full-Newton step interior-point method for monotone weighted linear complementarity problems, *J. Optim. Theory Appl.*, **186** (2020), 864–878.

[2] L. Chen, A high-order modified Levenberg-Marquardt method for systems of nonlinear equations with fourth-order convergence, *Appl. Math. Comput.*, **285** (2016), 79–93.

[3] L. Chen, C.Z. Du, Y.F. Ma, The higher-order Levenberg-Marquardt method with Armijo type line search for nonlinear equations, *Optim. Methods Softw.*, **32** (2017), 516–533.

[4] J.S. Chen, The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem, *J. Global Optim.*, **36** (2006), 565–580.

[5] X.N. Chi, Z.P. Wan, Z.J. Hao, A full-modified-Newton step $O(n)$ infeasible interior-point method for the special weighted linear complementarity problem, *J. Ind. Manag. Optim.*, **18** (2022), 2579–2598.

[6] X.N. Chi, G.Q. Wang, A full-Newton step infeasible interior-point method for the special weighted linear complementarity problem, *J. Optim. Theory Appl.*, **190** (2021), 108–129.

[7] X.N. Chi, Q.L. Yang, Z.P. Wan, S.B. Zhang, The new full-Newton step interior-point algorithm for the Fisher market equilibrium problems based on a kernel function, *J. Ind. Manag. Optim.*, **19** (2023) 7018–7035.

[8] J.Y. Fan, The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence, *Math. Comput.*, **81** (2012), 447–466.

[9] J.Y. Fan, Accelerating the modified Levenberg-Marquardt method for nonlinear equations, *Math. Comput.*, **83** (2014), 1173–1187.

[10] X.R. He, J.Y. Tang, A smooth Levenberg-Marquardt method without nonsingularity condition for wLCP, *AIMS Math.,* **7(5)** (2022), 8914–8932.

[11] B. Kheirfam, Complexity analysis of a full-Newton step interior-point method for monotone weighted linear complementarity problems. *J. Optim. Theory Appl.*, **202** (2024) 133–145.

[12] F.A. Potra, Weighted complementarity problems-a new paradigm for computing equilibria, *SIAM J. Optim.*, **22(4)** (2012), 1634–1654.

[13] G.W. Stewart, J.G. Sun, Matrix Perturbation Theory, Academic Press, San Diego, CA, (1990).

[14] J.Y. Tang, A variant nonmonotone smoothing algorithm with improved numerical results for large-scale LWCPs, *Comput. Appl. Math.*, **37** (2018), 3927–3936.

[15] J.Y. Tang, H.C. Zhang, A nonmonotone smoothing Newton algorithm for weighted complementarity problems, *J. Optim. Theory Appl.*, **189** (2021), 679–715.

[16] J.Y. Tang, J.C. Zhou, A modified damped Gauss-Newton method for non-monotone weighted linear complementarity problems, *Optim. Methods Softw.*, **37(3)** (2022), 1145–1164.

[17] J.Y. Tang, J.C. Zhou, Quadratic convergence analysis of a nonmonotone Levenberg-Marquardt type method for the weighted nonlinear complementarity problem, *Comput. Optim. Appl.*, **80** (2021), 213–244.

[18] J.Y. Tang, J.C. Zhou, H.C. Zhang, An accelerated smoothing Newton method with cubic convergence for weighted complementarity problems, *J. Optim. Theory Appl.*, **196** (2023), 641–665.

[19] J. Zhang, A smoothing Newton algorithm for weighted linear complementarity problem, *Optim. Lett.*, **10** (2016), 499–509.

[20] L. Zhang, X.N. Chi, S.B. Zhang, Y.P. Yang, A predictor-corrector interior-point algorithm for $P_*(\kappa)$-weighted linear complementarity problems. *AIMS Math.*, **8** (2023), 9212–9229.