Hindawi
Mathematical Problems in Engineering
Volume 2018, Article ID 5498760, 13 pages
https://doi.org/10.1155/2018/5498760

Research Article

An Implementable SAA Nonlinear Lagrange Algorithm for Constrained Minimax Stochastic Optimization Problems

Chuanmei Wang, Suxiang He, and Haiying Wu

School of Science, Wuhan University of Technology, Wuhan 430070, China

Correspondence should be addressed to Suxiang He; hesux@whut.edu.cn

Received 3 September 2018; Accepted 18 November 2018; Published 9 December 2018

Academic Editor: A. M. Bastos Pereira

This paper proposes an implementable SAA (sample average approximation) nonlinear Lagrange algorithm for the constrained minimax stochastic optimization problem based on the sample average approximation method. A computable nonlinear Lagrange function with sample average approximation functions of original functions is minimized and the Lagrange multiplier is updated based on the sample average approximation functions of original functions in the algorithm. And it is shown that the solution sequences obtained by the novel algorithm for solving subproblem converge to their true counterparts with probability one as the sample size approximates infinity under some moderate assumptions. Finally, numerical experiments are carried out for solving some typical test problems and the obtained numerical results preliminarily demonstrate that the proposed algorithm is promising.

1. Introduction

Minimax stochastic optimization is a kind of important problem in stochastic optimization. Minimax stochastic optimization has drawn much attention in recent years, which has been widely applied in subjects such as inventory theory, finance optimization, control science, and engineering field (see [1–7]).

This paper considers the constrained minimax stochastic optimization problems as follows:

$$\begin{aligned} \min \quad & T(x) \\ \text{s.t.} \quad & \mathbb{E}\left[g_j(x,\xi)\right] \le 0, \quad j = 1,\dots,p, \\ & \mathbb{E}\left[h_l(x,\xi)\right] = 0, \quad l = 1,\dots,q, \end{aligned} \tag{1}$$

where $T(x) = \max_{i\in\{1,2,\dots,m\}} \mathbb{E}[f_i(x,\xi)]$, $\xi : \Omega \longrightarrow \Xi \subseteq \mathfrak{R}^k$ is a random vector that is defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, $\mathbb{E}$ denotes mathematical expectation with respect to the distribution of $\xi \in \Xi$, and $f_i : \mathfrak{R}^n \times \Xi \longrightarrow \mathfrak{R}$ $(i = 1,\dots,m)$, $g_j : \mathfrak{R}^n \times \Xi \longrightarrow \mathfrak{R}$ $(j = 1,\dots,p)$ and $h_l : \mathfrak{R}^n \times \Xi \longrightarrow \mathfrak{R}$ $(l = 1,\dots,q)$ are well-defined. Since the objective function $T(x)$ is not differentiable, the efficient smooth optimization methods cannot be used to solve the problem (1) directly. Moreover, either distribution of random vector ξ is unknown or it is too complex to compute the multidimensional integral, so the exact numerical evaluation of the expected value functions in problem (1) is very difficult, which results in that problem (1) cannot be solved directly by the traditional deterministic optimization methods even though problem (1) is smooth.

On the one hand, as one of the famous smoothing techniques aiming to overcome the nonsmoothness of $T(x)$ (see [1, 8–16]), the nonlinear Lagrange method has many interesting merits, such as no restrictions on the feasibility of variables x, improvement on the convergence rate and the numerical robustness compared with penalty method by introducing the Lagrangian multipliers as the main driving force. On the other hand, the sample average approximation method is one of the well-behaved approaches for solving stochastic programming problems, the basic idea of which is to generate an independent and identically distributed (i.i.d.) sample $\xi^1, \cdots, \xi^N$ of the random variable $\xi \in \Xi$ with sample size N by Monte Carlo sampling method and approximate the involved expected value functions in problem (1) by their corresponding sample average functions. The SAA method has drawn much attention from many authors, see the comprehensive work by Shapiro [17] and the other works in [18–30].

Motivated by the effectiveness of the nonlinear Lagrange method, this paper presents a nonlinear Lagrange function for problem (1) based on the work in [13] as follows:

$$\overline{L}(x,u,v,y,z,t) = t\ln\left\{\sum_{i\in M} u_i e^{(F_i(x)-z)/t} + \sum_{j\in P} v_j\left(e^{G_j(x)/t}-1\right) + \frac{1}{t}\sum_{l\in Q}\left(y_l + \frac{1}{2t}H_l(x)\right)H_l(x)\right\}, \tag{2}$$

where $F_i(x) = \mathbb{E}[f_i(x,\xi)]$, $G_j(x) = \mathbb{E}[g_j(x,\xi)]$, $H_l(x) = \mathbb{E}[h_l(x,\xi)]$, $u = (u_1,\dots,u_m)^T \in U_{(m)}$ ($U_{(m)} := \{u \in \mathfrak{R}^m \mid \sum_{i=1}^m u_i = 1, u_i \geq 0, i \in M = \{1,\dots,m\}\}$), $v = (v_1,\dots,v_p)^T \in \mathfrak{R}_+^p$ ($\mathfrak{R}_+^p := \{v \in \mathfrak{R}^p \mid v_j \geq 0, j \in P = \{1,\dots,p\}\}$, $y = (y_1,\dots,y_q)^T \in \mathfrak{R}^q$, (u,v,y) is Lagrange multiplier, $Q = \{1,\dots,q\}$, $t > 0$ is a penalty parameter, and $z \in \mathfrak{R}$ is an estimate of the objective function $T(x)$. $\overline{L}(x,u,v,y,z,t)$ has good properties and the corresponding nonlinear Lagrangian algorithm is recalled below (see [13]).

Algorithm 1.

Step 1. Choose $t \in (0,\hat{t})$, where $\hat{t} \in (0,1)$, $u^{(0)} \in U_{(m)}$, $v^{(0)} \in \mathfrak{R}_{++}^p(\mathfrak{R}_{++}^p := \{v \in \mathfrak{R}^p | v_j > 0, j \in P\})$, $y^{(0)} \in \mathfrak{R}^q$, $z^{(0)} \in \mathfrak{R}_{++}(\mathfrak{R}_{++} := \{z \in \mathfrak{R} | z > 0\})$, and $\varepsilon \in (0,1)$ being small enough and set $k = 0$.

Step 2. Solve

$$\min_{x\in\mathfrak{R}^n} \overline{L}\left(x,u^{(k)},v^{(k)},y^{(k)},z^{(k)},t\right) \tag{3}$$

and obtain the optimal solution $x^{(k)}$.

Step 3. If $|\sum_{i\in M} u_i^{(k)}(F_i(x^{(k)}) - T(x^{(k)}))| + |\sum_{j\in P} v_j^{(k)} G_j(x^{(k)})| + \|H_l(x^{(k)})\| \leq \varepsilon$, then stop; otherwise go to Step 4.

Step 4. Update the Lagrange multiplier $u_i^{(k)}, v_j^{(k)}, y_l^{(k)}$, and $z^{(k)}$ by

$$\begin{aligned} u_i^{(k+1)} &= \frac{u_i^{(k)} e^{F_i(x^{(k)})/t}}{\sum_{i=1}^m u_i^{(k)} e^{F_i(x^{(k)})/t}}, \quad i \in M, \\ v_j^{(k+1)} &= v_j^{(k)} e^{G_j(x^{(k)})/t}, \quad j \in P, \\ y_l^{(k+1)} &= y_l^{(k)} + \frac{1}{t} H_l\left(x^{(k)}\right), \quad l \in Q, \\ z^{(k+1)} &= t\ln\left(\sum_{i=1}^m u_i^{(k)} e^{F_i(x^{(k)})/t}\right). \end{aligned} \tag{4}$$

Step 5. Set $k = k+1$ and return to Step 2.

In view of the difficulty in the numerical computation of the expected value function in Algorithm 1 and the inspiration from the SAA method, this paper will propose an implementable SAA nonlinear Lagrange algorithm in Section 3. And under some suitable assumptions, the convergence of the proposed algorithm will be analyzed in Section 3. In Section 2, some useful preliminaries will be presented. Furthermore, the numerical results for some typical test problems are reported to verify the feasibility and effectiveness of the proposed algorithm in Section 4. In the last section, the conclusion is drawn.

2. Preliminaries

This section serves as a preparation for the theoretical analysis in the subsequent section. Firstly, this section provides some assumptions on problem (1), and then recalls some related definitions and conclusions.

The Lagrange function for problem (1) is defined by $L(x,u,v,y) = \sum_{i\in M} u_i F_i(x) + \sum_{j\in P} v_j G_j(x) + \sum_{l\in Q} y_l H_l(x)$. Let (x^*,u^*,v^*,y^*) denote the Karush-Kuhn-Tucker (KKT) solution of problem (1) and $z^* = T(x^*)$ (see [13]). Let $\delta > 0$ be small enough. Define $S(x^*,\delta) = \{x \in \mathfrak{R}^n \mid \|x - x^*\| \leq \delta\}$, $S(u^*,\delta) = \{u \in \mathfrak{R}^m \mid \|u - u^*\| \leq \delta\}$, $S(v^*,\delta) = \{v \in \mathfrak{R}^p \mid \|v - v^*\| \leq \delta\}$, $S(y^*,\delta) = \{y \in \mathfrak{R}^q \mid \|y - y^*\| \leq \delta\}$, and $S(z^*,\delta) = \{z \in \mathfrak{R} \mid |z - z^*| \leq \delta\}$. Define $B_\delta(x^*,u^*,v^*,y^*,z^*) = \{(x,u,v,y,z) \in \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R}^p \times \mathfrak{R}^q \times \mathfrak{R} \mid x \in S(x^*,\delta), u \in S(u^*,\delta), v \in S(v^*,\delta), y \in S(y^*,\delta), z \in S(z^*,\delta)\}$. Some assumptions on problem (1) are made as follows:

(a) For any $\xi \in \Xi$, $f(x,\xi)$, $g(x,\xi)$, $h(x,\xi)$ are twice continuously differentiable with respect to x on $S(x^*,\delta)$, and the function values are finite, where $f(x,\xi) = (f_1(x,\xi),\cdots,f_m(x,\xi))^T$, $g(x,\xi) = (g_1(x,\xi),\cdots,g_p(x,\xi))^T$, and $h(x,\xi) = (h_1(x,\xi),\cdots,h_q(x,\xi))^T$.

(b) There exist nonnegative measurable functions $p_i(\xi)$ $(i = 1,2,3)$ such that $\mathbb{E}[p_i(\xi)]$ $(i = 1,2,3)$ being finite and for every $x \in S(x^*,\delta)$ the following inequalities are true with probability one:

$$\begin{aligned} \sup_{x\in S(x^*,\delta)} \|f(x,\xi)\| &< p_1(\xi), \\ \sup_{x\in S(x^*,\delta)} \|g(x,\xi)\| &< p_2(\xi), \\ \sup_{x\in S(x^*,\delta)} \|h(x,\xi)\| &< p_3(\xi). \end{aligned} \tag{5}$$

(c) The random sample $\xi_1,\dots,\xi_N$ is independent and identically distributed, and obeys the law of large numbers.

(d) For ease of presentation, assume

$$\begin{aligned} I_1(x^*) &= \{i \mid F_i(x^*) = T_i(x^*),\ i \in M\} = \{1,\dots,r_1\}, \\ I_2(x^*) &= \left\{j \mid G_j(x^*) = 0,\ j \in P\right\} = \{1,\dots,r_2\}. \end{aligned} \tag{6}$$

(e) KKT condition holds. That is,

$$\nabla_x L(x^*, u^*, v^*, y^*) = 0,$$
$$u^* \in U_{(m)},\ v^* \in \Re^p_+,\ y^* \in \Re^q,$$
$$u_i^*(T(x^*) - F_i(x^*)) = 0,\quad i \in M, \tag{7}$$
$$v_j^* G_j(x^*) = 0,\quad G_j(x^*) \le 0,\ j \in P,$$
$$H_l(x^*) = 0,\quad l \in Q.$$

(f) Strict complementarity condition holds, i.e., $v_j^* > 0$ for $i \in I_1(x^*)$ and $u_i^* > 0$ for $j \in I_2(x^*)$.

(g) The vectors of $\{\nabla F_i(x^*) \mid i \in I_1(x^*)\} \cup \{\nabla G_j(x^*) \mid i \in I_2(x^*)\} \cup \{\nabla H_l(x^*) \mid l \in Q\}$ are linearly independent.

(h) There exists a constant $\lambda > 0$ such that, for all $\overline{y}$ in $\Re^n$ satisfying $\nabla F_i(x^*)^T \overline{y} = 0$, $i \in I_1(x^*)$, $\nabla G_j(x^*)^T \overline{y} = 0$, $j \in I_2(x^*)$, and $\nabla H_l(x^*)^T \overline{y} = 0$, $l \in Q$, it holds that

$$\overline{y}^T \nabla_x^2 L(x^*, u^*, v^*, y^*)\, \overline{y} \ge \lambda \|\overline{y}\|^2. \tag{8}$$

The following definition (see [17]) is recalled.

Definition 2. For nonempty sets A and B in $\Re^n$, we denote by $dist(x, A) = \inf_{x' \in A} \|x - x'\|$ the distance from $x \in \Re^n$ to A, and by $\mathbb{D}(A, B) = \sup_{x \in A} dist(x, B)$ the deviation of the set A from the set B.

To present the basic lemma, we now consider the following stochastic optimization problem:

$$\min_{x \in C} \{d(x) := \mathbb{E}[D(x, \xi)]\}, \tag{9}$$

where C is a nonempty and compact subset on $\Re^n$, ξ is a random vector on Ξ, $D : C \times \Xi \longrightarrow \Re$. For any $\xi \in \Xi$, $D(x, \xi)$ is finite and continuous for all $x \in C$. The sample average approximation problem of (9) can be expressed as

$$\min_{x \in C} \left\{ \widehat{d}_N(x) := \frac{1}{N} \sum_{j=1}^{N} D(x, \xi_j) \right\}, \tag{10}$$

where $\xi_1, \ldots,$ and ξ_N are N independent sample observations and obey the law of large numbers. Let v^* and S denote the optimal value and the optimal solution set of problem (9), $\widehat{v}_N$ and $\widehat{S}_N$ indicate the optimal value and the optimal solution set of problem (10). Then the resulted essential conclusion is obtained in Lemma 3.

Lemma 3 (see [17]). *Suppose that there exists a nonnegative measurable function $\overline{p}(\xi)$ independent of x for $D(x, \xi)$ such that $\sup_{x \in C} \|D(x, \xi)\| \le \overline{p}(\xi)$ with probability one. Then the following conclusions are true:*

(i) *$d(x)$ is continuous and finite on C;*

(ii) *$\widehat{d}_N(x)$ converges to $d(x)$ with probability one uniformly on C as $N \longrightarrow \infty$;*

(iii) *$\widehat{v}_N$ converges to v^* with probability one as $N \longrightarrow \infty$;*

(iv) *$D(\widehat{S}_N, S)$ converges to 0 with probability one as $N \longrightarrow \infty$.*

3. The SAA Nonlinear Lagrange Algorithm and Its Convergence

This section presents an implementable SAA nonlinear Lagrange algorithm based on the SAA nonlinear Lagrange function of nonlinear Lagrange function (2) and then analyzes its convergence by means of the preliminaries in Section 2.

Firstly, we construct a SAA nonlinear Lagrange function of nonlinear Lagrange function (2) below:

$$\overline{L}_N(.) = t \ln \left\{ \sum_{i \in M} u_i e^{(\widehat{F}_{N,i}(x) - z)/t} + \sum_{j \in P} v_j \left(e^{\widehat{G}_{N,j}(x)/t} - 1 \right) + \frac{1}{t} \sum_{l \in Q} \left(y_l + \frac{1}{2t} \widehat{H}_{N,l}(x) \right) \widehat{H}_{N,l}(x) \right\}, \tag{11}$$

where (.) denotes (x, u, v, y, z, t), $\widehat{F}_{N,i}(x) = (1/N) \sum_{k=1}^{N} f_i(x, \xi_k)$, $\widehat{G}_{N,j}(x) = (1/N) \sum_{k=1}^{N} g_j(x, \xi_k)$, $\widehat{H}_{N,l}(x) = (1/N) \sum_{k=1}^{N} h_l(x, \xi_k)$, and $\xi_1, \ldots, \xi_N$ is a random sample.

Based on the SAA nonlinear Lagrange function (11) and Algorithm 1, an implementable SAA nonlinear Lagrange algorithm is presented as follows.

Algorithm 4.

Step 1. Choose $t \in (0, \widehat{t})$, where $\widehat{t} \in (0, 1)$, $\epsilon \in (0, 1)$ being small enough, $\widehat{u}_N^{(0)} \in U_{(m)}$, $\widehat{v}_N^{(0)} \in \Re^p_{++}$, $\widehat{y}_N^{(0)} \in \Re^q$, $\widehat{z}_N^{(0)} \in \Re_{++}$, and N is large enough. Set $k = 0$.

Step 2. Solve

$$\min_{x \in \Re^n} \overline{L}_N\left(x, \widehat{u}_N^{(k)}, \widehat{v}_N^{(k)}, \widehat{y}_N^{(k)}, \widehat{z}_N^{(k)}, t\right) \tag{12}$$

and obtain the optimal solution $\widehat{x}_N^{(k)}$.

Step 3. If $|\sum_{i \in M} \widehat{u}_{N,i}^{(k)} (\widehat{F}_{N,i}(\widehat{x}_N^{(k)}) - T(\widehat{x}_N^{(k)}))| + |\sum_{j \in P} \widehat{v}_{N,j}^{(k)} \widehat{G}_{N,j}(\widehat{x}_N^{(k)})| + \|\widehat{H}_N(\widehat{x}_N^{(k)})\| \le \varepsilon$, then stop; otherwise go to.

Step 4. Update the Lagrange multiplier $\widehat{u}_{N,i}^{(k)}$, $\widehat{v}_{N,j}^{(k)}$, $\widehat{y}_{N,l}^{(k)}$, and $\widehat{z}_N^{(k)}$ by

$$\begin{aligned}
\widehat{u}_{N,i}^{(k+1)} &= \frac{\widehat{u}_{N,i}^{(k)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(k)})/t}}{\sum_{i=1}^{m} \widehat{u}_{N,i}^{(k)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(k)})/t}}, \quad i \in M, \\
\widehat{v}_{N,j}^{(k+1)} &= \widehat{v}_{N,j}^{(k)} e^{\widehat{G}_{N,j}(\widehat{x}_N^{(k)})/t}, \quad j \in P, \\
\widehat{y}_{N,l}^{(k+1)} &= \widehat{y}_{N,l}^{(k)} + \frac{1}{t} \widehat{H}_{N,l}\left(\widehat{x}_N^{(k)}\right), \quad l \in Q, \\
\widehat{z}_N^{(k+1)} &= t \ln \left(\sum_{i=1}^{m} \widehat{u}_{N,i}^{(k)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(k)})/t} \right).
\end{aligned} \tag{13}$$

Step 5. Set $k = k + 1$ and return to Step 2.

Next, we study the convergence of Algorithm 4 on $B_\delta(x^*, u^*, v^*, y^*, z^*)$ based on the assumptions (a)-(h) and Lemma 3 in Section 2. Let $v_N^{(k)}$ and $\widehat{S}_N^{(k)}$ denote the optimal value and optimal solution set of problem (12), $v^{(k)}$ and $S^{(k)}$ denote the optimal value and optimal solution set of problem (3).

Theorem 5. *If assumptions (a)-(c) hold and for k ($k = 0, 1, \ldots$), $\widehat{u}_{N,i}^{(k)}$ converges to $u_i^{(k)}$ ($i = 1, \ldots, m$) with probability one, $\widehat{v}_{N,j}^{(k)}$ converges to $v_j^{(k)}$ ($j = 1, \ldots, p$) with probability one, $\widehat{y}_{N,l}^{(k)}$ converges to $y_l^{(k)}$ ($l = 1, \ldots, q$) with probability one, and $\widehat{z}_N^{(k)}$ converges to $z^{(k)}$ with probability one as $N \longrightarrow \infty$, then the following statements are true.*

(i) $\overline{L}_N(x, \widehat{u}_N^{(k)}, \widehat{v}_N^{(k)}, \widehat{y}_N^{(k)}, \widehat{z}_N^{(k)}, t)$ *converges to* $\overline{L}(x, u^{(k)}, v^{(k)}, y^{(k)}, z^{(k)}, t)$ *with probability one uniformly as* $N \longrightarrow \infty$ *on* $S(x^*, \delta)$;

(ii) $v_N^{(k)}$ *converges to* $v^{(k)}$ *with probability one and* $D(\widehat{S}_N^{(k)}, S^{(k)})$ *converges to 0 with probability one as* $N \longrightarrow \infty$.

Proof. (i) Define $\varphi^{(1)}(a, b, z, t) = ae^{(b-z)/t}$, $\varphi^{(2)}(c, d, t) = d(e^{c/t} - 1)$, and $\varphi^{(3)}(e, f, t) = (1/t)(e + (1/2t)f)f$, where $a, b, z, c, d, e, f \in \mathfrak{R}$ and t is defined as in (2). And set

$$\begin{aligned}\phi(x,t) &= \sum_{i\in M}\varphi^{(1)}\left(u_i^{(k)}, F_i(x), z^{(k)}, t\right) + \sum_{j\in P}\varphi^{(2)}\left(v_j^{(k)}, G_j(x), t\right) + \sum_{l\in Q}\varphi^{(3)}\left(y_l^{(k)}, H_l(x), t\right), \\ \phi_N(x,t) &= \sum_{i\in M}\varphi^{(1)}\left(\widehat{u}_{N,i}^{(k)}, \widehat{F}_{N,i}(x), \widehat{z}_N^{(k)}, t\right) + \sum_{j\in P}\varphi^{(2)}\left(\widehat{v}_{N,j}^{(k)}, \widehat{G}_{N,j}(x), t\right) + \sum_{l\in Q}\varphi^{(3)}\left(\widehat{y}_{N,l}^{(k)}, \widehat{H}_{N,l}(x), t\right).\end{aligned} \tag{14}$$

At first, we are to prove that $\phi_N(x)$ converges to $\phi(x)$ with probability one uniformly as $N \longrightarrow \infty$, for which we need to prove that as $N \longrightarrow \infty$, $\varphi^{(1)}(u_{N,i}^{(k)}, \widehat{F}_{N,i}(x), \widehat{z}_N^{(k)}, t)$ converges to $\varphi^{(1)}(u_i^{(k)}, F_i(x), z^{(k)}, t)$ ($i = 1, \ldots, m$) with probability one uniformly, $\varphi^{(2)}(v_{N,j}^{(k)}, \widehat{G}_{N,j}(x), t)$ converges to $\varphi^{(2)}(v_j^{(k)}, G_j(x), t)$ ($j = 1, \ldots, p$) with probability one uniformly, and $\varphi^{(3)}(y_{N,l}^{(k)}, \widehat{H}_{N,l}(x), t)$ converges to $\varphi^{(3)}(y_l^{(k)}, H_l(x), t)$ ($l = 1, \ldots, q$) with probability one uniformly on $S(x^*, \delta)$, respectively. The proof for it is divided into the following three parts.

(A) First we shall prove that as $N \longrightarrow \infty$, $\varphi^{(1)}(u_{N,i}^{(k)}, \widehat{F}_{N,i}(x), z_N^{(k)})$ converges to $\varphi^{(1)}(u_i^{(k)}, F_i(x), z^{(k)})$ ($i = 1, \ldots, m$) with probability one uniformly on $S(x^*, \delta)$.

According to the definition of $\varphi^{(1)}(a, b, z)$, we have

$$\begin{aligned}&\left|\varphi^{(1)}\left(\widehat{u}_{N,i}^{(k)}, \widehat{F}_{N,i}(x), \widehat{z}_N^{(k)}, t\right) - \varphi^{(1)}\left(u_i^{(k)}, F_i(x), z^{(k)}, t\right)\right| \\ &\quad = \left|\widehat{u}_{N,i}^{(k)} e^{(\widehat{F}_{N,i}(x)-\widehat{z}_N^{(k)})/t} - u_i^{(k)} e^{(\widehat{F}_{N,i}(x)-\widehat{z}_N^{(k)})/t} + u_i^{(k)} e^{(\widehat{F}_{N,i}(x)-\widehat{z}_N^{(k)})/t} - u_i^{(k)} e^{(F_i(x)-z^{(k)})/t}\right| \le \left|\widehat{u}_{N,i}^{(k)} - u_i^{(k)}\right| e^{(\widehat{F}_{N,i}(x)-\widehat{z}_N^{(k)})/t} + u_i^{(k)}\left|e^{(\widehat{F}_{N,i}(x)-\widehat{z}_N^{(k)})/t} - e^{(F_i(x)-z^{(k)})/t}\right|.\end{aligned} \tag{15}$$

One has that $\widehat{F}_{N,i}(x)$ and $F_i(x)$ ($i = 1, \ldots, m$) are continuous on $S(x^*, \delta)$ with respect to x from the assumption (a) and Lemma 3, so there exists a closed interval $[c_1, d_1]$ ($c_1, d_1 \in R$) such that $\widehat{F}_{N,i}(x) \in [c_1, d_1]$, and $F_i(x) \in [c_1, d_1]$ for $x \in S(x^*, \delta)$. Since $\widehat{z}_N^{(k)}$ and $z^{(k)}$ are bounded, there exist $c_1{}'$, $d_1{}' \in R$, such that $\widehat{F}_{N,i}(x) - z_N^{(k)} \in [c_1{}', d_1{}']$ and $F_i(x) - z^{(k)} \in [c_1{}', d_1{}']$. Thus, we have that $e^{(\widehat{F}_{N,i}(x)-z_N^{(k)})/t}$ is bounded on $S(x^*, \delta)$ with respect to x. Let $P(y) = e^{y/t} (y \in R)$. Then $P(y)$ is continuous in $[c_1{}', d_1{}']$ with respect to y. It follows that $P(y)$ is continuous uniformly in $[c_1{}', d_1{}']$ from the property of continuous function. That is, for any $\varepsilon > 0$, there exists $\widehat{\delta}_1$, for $x, y \in [c_1{}', d_1{}']$; if $|x - y| < \widehat{\delta}_1$, it holds that

$$\left|e^{x/t} - e^{y/t}\right| < \varepsilon. \tag{16}$$

Note that $\widehat{F}_{N,i}(x)$ converges to $F_i(x)$ ($i = 1, \ldots, m$) with probability one uniformly on $S(x^*, \delta)$ from Lemma 3 and $\widehat{z}_N^{(k)}$ converges to $z^{(k)}$ with probability one; hence for $\widehat{\delta}_1 > 0$, there exists N_1, when $N > N_1$, for any $x \in S(x^*, \delta)$, we have

$$\begin{aligned}&\left|\left(\widehat{F}_{N,i}(x) - \widehat{z}_N^{(k)}\right) - \left(F_i(x) - z^{(k)}\right)\right| \\ &\quad < \left|\widehat{F}_{N,i}(x) - F_i(x)\right| + \left|\widehat{z}_N^{(k)} - z^{(k)}\right| < \frac{\widehat{\delta}_1}{2} + \frac{\widehat{\delta}_1}{2} = \widehat{\delta}_1.\end{aligned} \tag{17}$$

It follows from (16) and (17) that for any $\varepsilon > 0$, there exists N_1, when $N > N_1$; for any $x \in S(x^*, \delta)$, it holds that

$$\left|e^{(\widehat{F}_{N,i}(x)-z_N^{(k)})/t} - e^{(F_i(x)-z^{(k)})/t}\right| < \varepsilon. \tag{18}$$

Thus, in view of (15), $u_i^{(k)} \in S(u^*, \delta)$, and the condition that $\widehat{u}_{N,i}^{(k)}$ converges to $u_i^{(k)}$ ($N \longrightarrow \infty$) with probability one, we obtain that $\varphi^{(1)}(u_{N,i}^{(k)}, \widehat{F}_{N,i}(x), \widehat{z}_N^{(k)}, t)$ converges to $\varphi^{(1)}(u_i^{(k)}, F_i(x), \widehat{z}^{(k)}, t)$ with probability one uniformly on $S(x^*, \delta)$ as $N \longrightarrow \infty$.

(B) Next, we prove that $\varphi^{(2)}(v_{N,j}^{(k)}, \widehat{G}_{N,j}(x), t)$ converges to $\varphi^{(2)}(v_j^{(k)}, G_j(x), t)$ ($j = 1, \ldots, p$) with probability one uniformly on $S(x^*, \delta)$ as $N \longrightarrow \infty$.

From the definition of $\varphi^{(2)}(c,d,t)$, we have

$$\begin{aligned}&\left|\varphi^{(2)}\left(\widehat{v}_{N,j}^{(k)},\widehat{G}_{N,j}(x),t\right)-\varphi^{(2)}\left(v_j^{(k)},G_j(x),t\right)\right|\\&=\left|\widehat{v}_{N,j}^{(k)}\left(e^{\widehat{G}_{N,j}(x)/t}-1\right)-v_j^{(k)}\left(e^{\widehat{G}_{N,j}(x)/t}-1\right)\right.\\&\left.+v_j^{(k)}\left(e^{\widehat{G}_{N,j}(x)/t}-1\right)-v_j^{(k)}\left(e^{G_j(x)/t}-1\right)\right|\le\left|\widehat{v}_{N,j}^{(k)}\right.\\&\left.-v_j^{(k)}\right|\left(e^{\widehat{G}_{N,j}(x)/t}-1\right)+v_j^{(k)}\left|e^{\widehat{G}_{N,j}(x)/t}-e^{G_j(x)/t}\right|.\end{aligned}\tag{19}$$

Since $\widehat{G}_{N,j}(x)$ and $G_j(x)$ $(j=1,\ldots,p)$ are continuous on $S(x^*,\delta)$ with respect to x from the assumption (a) and Lemma 3, there exits a closed interval $[c_2,d_2]$ $(c_2,d_2\in R)$ such that $\widehat{G}_{N,j}(x)\in[c_2,d_2]$ and $G_i(x)\in[c_2,d_2]$ for $x\in S(x^*,\delta)$, which means that $e^{\widehat{G}_{N,j}(x)/t}$ is bounded on $S(x^*,\delta)$ with respect to x. From the proof process of (A), we get that, for any $\varepsilon>0$, there exists a positive integer N_2, when $N>N_2$, for $x\in S(x^*,\delta)$, it holds that

$$\left|e^{\widehat{G}_{N,j}(x)/t}-e^{G_j(x)/t}\right|<\varepsilon.\tag{20}$$

Therefore, combined with (19), it follows from $\widehat{v}_{N,j}^{(k)}$ converging to $v_j^{(k)}$ with probability one as $N\longrightarrow\infty$ and $v_j^{(k)}\in S(v^*,\delta)$ that $\varphi^{(2)}(\widehat{v}_{N,j}^{(k)},\widehat{G}_{N,j}(x),t)$ converges to $\varphi^{(2)}(v_j^{(k)},G_j(x),t)$ $(j=1,\ldots,p)$ uniformly with probability one on $S(x^*,\delta)$ as $N\longrightarrow\infty$.

(C) Now we prove that $\varphi^{(3)}(y_{N,l}^{(k)},\widehat{H}_{N,l}(x),t)$ converges to $\varphi^{(3)}(y_l^{(k)},H_l(x),t)$ with probability one uniformly on $S(x^*,\delta)$ as $N\longrightarrow\infty$.

From the definition of $\varphi^{(3)}(e,f,t)$, we have

$$\begin{aligned}&\left|\varphi^{(3)}\left(\widehat{y}_{N,l}^{(k)},\widehat{H}_{N,l}(x),t\right)-\varphi^{(3)}\left(y_l^{(k)},H_l(x),t\right)\right|\\&=\left|\frac{1}{t}\left(\widehat{y}_{N,l}^{(k)}+\frac{1}{2t}\widehat{H}_{N,l}(x)\right)\widehat{H}_{N,l}(x)\right.\\&-\frac{1}{t}\left(y_l^{(k)}+\frac{1}{2t}\widehat{H}_{N,l}(x)\right)\widehat{H}_{N,l}(x)\\&+\frac{1}{t}\left(y_l^{(k)}+\frac{1}{2t}\widehat{H}_{N,l}(x)\right)\widehat{H}_{N,l}(x)\\&\left.-\frac{1}{t}\left(y_l^{(k)}+\frac{1}{2t}H_l(x)\right)H_l(x)\right|\le\frac{1}{t}\left|\widehat{y}_{N,l}^{(k)}-y_l^{(k)}\right|\\&\cdot\left|\widehat{H}_{N,l}(x)\right|+\frac{1}{t}\left|\left[y_l^{(k)}+\frac{1}{2t}\left(\widehat{H}_{N,l}(x)+H_l(x)\right)\right]\right|\\&\cdot\left|\widehat{H}_{N,l}(x)-H_l(x)\right|.\end{aligned}\tag{21}$$

Since $\widehat{H}_{N,l}(x)$ and $H_l(x)$ $(l=1,\ldots,q)$ are continuous on $S(x^*,\delta)$ with respect to x according to the assumption (a) and Lemma 3, there exists a closed interval $[c_3,d_3]$ $(c_3,d_3\in R)$ such that $\widehat{H}_{N,l}(x)\in[c_3,d_3]$ and $H_l(x)\in[c_3,d_3]$ for $x\in S(x^*,\delta)$; i.e., $\widehat{H}_{N,l}(x)+H_l(x)$ is bounded. Then from $y_l^{(k)}\in S(y^*,\delta)$ it follows that $[y_l^{(k)}+(1/2t)(\widehat{H}_{N,l}(x)+H_l(x))]$ is bounded on $S(x^*,\delta)$. From Lemma 3, it is true that $\widehat{H}_{N,l}(x)$ converges to $H_l(x)$ $(l=1,\ldots,q)$ with probability one uniformly on $S(x^*,\delta)$. That is, for any $\varepsilon>0$, there exists a positive integer N_3, when $N>N_3$, for $x\in S(x^*,\delta)$, the following inequality holds with probability one:

$$\left|\widehat{H}_{N,l}(x)-H_l(x)\right|<\varepsilon.\tag{22}$$

Moreover, considering (21) and the fact that $\widehat{y}_{N,l}^{(k)}$ converges to $y_l^{(k)}$ with probability one as $N\longrightarrow\infty$, we have that $\varphi^{(3)}(\widehat{y}_{N,l}^{(k)},\widehat{H}_{N,l}(x),t)$ converges to $\varphi^{(3)}(y_l^{(k)},H_l(x),t)$ with probability one uniformly on $S(x^*,\delta)$ as $N\longrightarrow\infty$.

Thus, from the above analyses of (A), (B) and (C), we draw the conclusion that $\phi_N(x)$ converges to $\phi(x)$ with probability one uniformly on $S(x^*,\delta)$ as $N\longrightarrow\infty$. Furthermore, in view of the fact that $\ln x$ is continuous with respect to x on $\mathfrak{R}$, it can be proven that $t\ln\phi_N(x,t)$ converges to $t\ln\phi(x)$ uniformly with probability one as $N\longrightarrow\infty$ for $t>0$, which implies that the conclusion (i) is true.

(ii) From the conclusion (i) and Lemma 3, we can prove that the conclusion (ii) is true. □

Theorem 6. *If assumptions (a)-(c) hold and letting $\widehat{u}_{N,i}^{(0)}=u_i^{(0)}$ $(i=1,\ldots,m)$, $\widehat{v}_{N,j}^{(0)}=v_j^{(0)}$ $(j=1,\ldots,p)$, $\widehat{y}_{N,l}^{(0)}=y_l^{(0)}$ $(l=1,\ldots,q)$, and $\widehat{z}_N^{(0)}=z^{(0)}$, then for any $k\ge0$, the following statements hold:*

(i) *As $N\longrightarrow\infty$, $\widehat{u}_{N,i}^{(k)}$ converges to $u_i^{(k)}$ with probability one for $i=1,\cdots,m$, $\widehat{v}_{N,j}^{(k)}$ converges to $v_j^{(k)}$ with probability one for $j=1,\cdots,p$, $\widehat{y}_{N,l}^{(k)}$ converges to $y_l^{(k)}$ with probability one for $l=1,\cdots,q$, and $\widehat{z}_N^{(k)}$ converges to $z^{(k)}$ with probability one;*

(ii) *$\overline{L}_N(x,\widehat{u}_N^{(k)},\widehat{v}_N^{(k)},\widehat{y}_N^{(k)},\widehat{z}_N^{(k)},t)$ converges to $\overline{L}(x,u^{(k)},v^{(k)},y^{(k)},z^{(k)},t)$ with probability one uniformly on $S(x^*,\delta)$ as $N\longrightarrow\infty$;*

(iii) *$\widehat{v}_N^{(k)}$ converges to $v^{(k)}$ with probability one, and $D(\widehat{S}_N^{(k)},S^{(k)})$ converges to 0 with probability one as $N\longrightarrow\infty$.*

Proof. (i) We use the mathematical induction method to show that the statement (i) is true below. For $k=0$, we know that $\widehat{u}_{N,i}^{(0)}=u_i^{(0)}$ $(i=1,\ldots,m)$, $\widehat{v}_{N,j}^{(0)}=v_j^{(0)}$ $(j=1,\ldots,p)$, $\widehat{y}_{N,l}^{(0)}=y_l^{(0)}$ $(l=1,\ldots,q)$ and $\widehat{z}_N^{(0)}=z^{(0)}$, which means that the conclusion (i) is true for $k=0$. Next we prove that the conclusion (i) is true for $k\ge1$.

For $k=1$, from $\widehat{u}_{N,i}^{(0)}=u_i^{(0)}$, one has that

$$\begin{aligned}\widehat{u}_{N,i}^{(1)}&=\frac{u_i^{(0)}e^{\widehat{F}_{N,i}(\widehat{x}_N^{(0)})/t}}{\sum_{j=1}^m u_j^{(0)}e^{\widehat{F}_{N,j}(\widehat{x}_N^{(0)})/t}},\\u_i^{(1)}&=\frac{u_i^{(0)}e^{F_i(x^{(0)})/t}}{\sum_{j=1}^m u_j^{(0)}e^{F_j(x^{(0)})/t}},\\&\qquad i=1,\ldots,m.\end{aligned}\tag{23}$$

For any i $(i = 1, \dots, m)$, we have

$$\begin{aligned} &\left|\widehat{F}_{N,i}\left(\widehat{x}_N^{(0)}\right) - F_i\left(x^{(0)}\right)\right| \\ &\quad \leq \left|\widehat{F}_{N,i}\left(\widehat{x}_N^{(0)}\right) - F_i\left(\widehat{x}_N^{(0)}\right)\right| + \left|F_i\left(\widehat{x}_N^{(0)}\right) - F_i\left(x^{(0)}\right)\right|. \end{aligned} \tag{24}$$

From Lemma 3, we know that $\widehat{F}_{N,i}(\widehat{x}_N^{(0)})$ converges to $F_i(\widehat{x}_N^{(0)})$ with probability one as $N \longrightarrow \infty$, which implies that the first term on the right side of (24) converges to 0 as $N \longrightarrow \infty$. And we know that $F_i(x)$ is continuous and bounded on $S(x^*, \delta)$ from Lemma 3, and $\widehat{x}_N^{(0)}$ converges to $x^{(0)}$ with probability one as $N \longrightarrow \infty$ from Theorem 5, so the second part on the right side of (24) converges to 0 as $N \longrightarrow \infty$. That is, $\widehat{F}_{N,i}(\widehat{x}_N^{(0)})$ converges to $F_i(x^{(0)})$ with probability one as $N \longrightarrow \infty$. And we know $e^{x/t}$ is continuous with respect to x on $\mathfrak{R}$, hence $e^{\widehat{F}_{N,i}(\widehat{x}_N^{(0)})/t}$ converges to $e^{F_i(x^{(0)})/t}$ with probability one as $N \longrightarrow \infty$. Furthermore, we can prove that $\widehat{u}_{N,i}^{(1)}$ converges to $u_i^{(1)}$ with probability one as $N \longrightarrow \infty$ for $i = 1, \dots, m$.

For $k = 1$, we have

$$\begin{aligned} \widehat{v}_{N,j}^{(1)} &= \widehat{v}_{N,j}^{(0)} e^{\widehat{G}_{N,j}(\widehat{x}_N^{(0)})/t}, \\ v_j^{(1)} &= v_j^{(0)} e^{\widehat{G}_j(x^{(0)})/t}, \\ & \qquad j = 1, \dots, p. \end{aligned} \tag{25}$$

For any j $(j = 1, \dots, p)$, $\widehat{v}_{N,j}^{(0)} = v_j^{(0)}$, so we have

$$\left|\widehat{v}_{N,j}^{(1)} - v_j^{(1)}\right| = v_j^{(0)} \left|e^{\widehat{G}_{N,j}(\widehat{x}_N^{(0)})/t} - e^{G_j(x^{(0)})/t}\right|. \tag{26}$$

Similarly, we can prove that $\widehat{v}_{N,j}^{(1)}$ converges to $v_j^{(1)}$ with probability one as $N \longrightarrow \infty$ for $j = 1, \dots, p$ from Lemma 3 and Theorem 5.

For $k = 1$, since $\widehat{y}_{N,l}^{(0)} = \widehat{y}_l^{(0)}$ $(l = 1, \dots, q)$, it holds that

$$\begin{aligned} \widehat{y}_{N,l}^{(1)} &= y_l^{(0)} + \frac{1}{t}\widehat{H}_{N,l}\left(\widehat{x}_N^{(0)}\right), \\ y_l^{(1)} &= y_l^{(0)} + \frac{1}{t}H_l\left(x^{(0)}\right), \\ & \qquad l = 1, \dots, q. \end{aligned} \tag{27}$$

Hence for any l $(l = 1, \dots, q)$, we have

$$\begin{aligned} \left|\widehat{y}_{N,l}^{(1)} - y_l^{(1)}\right| &= \frac{1}{t}\left|\widehat{H}_{N,l}\left(\widehat{x}_N^{(0)}\right) - H_l\left(x^{(0)}\right)\right| \\ &\leq \frac{1}{t}\left|\widehat{H}_{N,l}\left(\widehat{x}_N^{(0)}\right) - H_l\left(\widehat{x}_N^{(0)}\right)\right| \\ &\quad + \frac{1}{t}\left|H_l\left(\widehat{x}_N^{(0)}\right) - H_l\left(x^{(0)}\right)\right|. \end{aligned} \tag{28}$$

We can also prove that $\widehat{y}_{N,l}^{(1)}$ converges to $y_l^{(1)}$ $(l = 1, \dots, q)$ with probability one as $N \longrightarrow \infty$ for $l = 1, \dots, q$ from Lemma 3 and Theorem 5.

For $k = 1$, it holds that

$$\begin{aligned} \widehat{z}_N^{(1)} &= t \ln\left(\sum_{i=1}^{m} \widehat{u}_{N,i}^{(0)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(0)})/t}\right), \\ z^{(1)} &= t \ln\left(\sum_{i=1}^{m} u_i^{(0)} e^{F_i(\widehat{x}^{(0)})/t}\right). \end{aligned} \tag{29}$$

Similar to the above proof process, we can draw the conclusion that $\widehat{z}_N^{(1)}$ converges to $z^{(1)}$ with probability one as $N \longrightarrow \infty$.

That is, we have proven that the conclusion (i) is true for $k = 1$. Now suppose that the conclusion (i) is true for $k = r$, where r is an integer $(r \geq 2)$. Next, we prove that the conclusion (i) is true for $k = r + 1$.

Since the conclusion (i) is true for $k = r$, we have that $\widehat{x}_N^{(r)}$ converges to $x^{(r)}$ with probability one as $N \longrightarrow \infty$ from Theorem 5. For $k = r + 1$, we have

$$\begin{aligned} \widehat{u}_{N,i}^{(r+1)} &= \frac{\widehat{u}_{N,i}^{(r)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(r)})/t}}{\sum_{j=1}^{m} \widehat{u}_{N,j}^{(r)} e^{\widehat{F}_{N,j}(\widehat{x}_N^{(r)})/t}}, \\ u_i^{(r+1)} &= \frac{u_i^{(r)} e^{F_i(x^{(r)})/t}}{\sum_{j=1}^{m} u_j^{(r)} e^{F_j(x^{(r)})/t}}, \\ & \qquad i = 1, \dots, m. \end{aligned} \tag{30}$$

For any i $(i = 1, \dots, m)$, one has

$$\begin{aligned} &\left|\widehat{u}_{N,i}^{(r)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(r)})/t} - u_i^{(r)} e^{F_i(x^{(r)})/t}\right| \\ &\quad \leq \left|\widehat{u}_{N,i}^{(r)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(r)})/t} - u_i^{(r)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(r)})/t}\right| \\ &\qquad + \left|u_i^{(r)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(r)})/t} - u_i^{(r)} e^{F_i(x^{(r)})/t}\right| \\ &\quad = \left|\widehat{u}_{N,i}^{(r)} - u_i^{(r)}\right| e^{\widehat{F}_{N,i}(\widehat{x}_N^{(r)})/t} \\ &\qquad + u_i^{(r)} \left|e^{\widehat{F}_{N,i}(\widehat{x}_N^{(r)})/t} - e^{F_i(x^{(r)})/t}\right|. \end{aligned} \tag{31}$$

In view of the proof process in Theorem 5, it follows that $\widehat{u}_{N,i}^{(r)} e^{\widehat{F}_{N,i}(\widehat{x}_N^{(r)})/t}$ converges to $u_i^{(r)} e^{F_i(x^{(r)})/t}$ with probability one as $N \longrightarrow \infty$. Moreover, note that the special forms of $\widehat{u}_{N,i}^{(r+1)}$ and $u_i^{(r+1)}$, we can verify that $\widehat{u}_{N,i}^{(r+1)}$ converges to $u_i^{(r+1)}$ with probability one as $N \longrightarrow \infty$ for $i = 1, \cdots, m$.

Similarly, we can prove that $\widehat{v}_{N,j}^{(r+1)}$ converges to $v_j^{(r+1)}$ with probability one as $N \longrightarrow \infty$ for $j = 1, \cdots, p$, $\widehat{y}_{N,l}^{(r+1)}$ converges to $y_l^{(r+1)}$ with probability one as $N \longrightarrow \infty$ for $l = 1, \cdots, q$, and $\widehat{z}_N^{(r+1)}$ converges to $z^{(r+1)}$ with probability one as $N \longrightarrow \infty$ from Lemma 3 and Theorem 5. That is, the conclusion (i) is true for $k = r + 1$. By mathematical induction method, hence we know that conclusion (i) is true for any $k \geq 0$.

(ii) Considering conclusion (i) and Theorem 5, we know that conclusion (ii) is true.

(iv) In view of conclusion (ii) and Lemma 3, we have that conclusion (iii) is true.

Thus, the proof of Theorem 6 is completed. □

Up until now, we have established the relationship between the optimal solution of problem (3) and the optimal solution of problem (12), and the convergence of the SAA Lagrange multiplier sequence in Algorithm 4 with probability one as $N \longrightarrow \infty$. Next we are to prove that the optimal solution sequence and the SAA Lagrange multiplier sequence obtained by Algorithm 4 converge to the optimal solution and the corresponding Lagrange multiplier of problem (1) with probability one as $N \longrightarrow \infty$ under the assumptions (a)-(h).

Theorem 7. *If assumptions (a)-(h) hold, and let* $\widehat{u}_{N,i}^{(0)} = u_i^{(0)}$ $(i = 1, \dots, m)$, $\widehat{v}_{N,j}^{(0)} = v_j^{(0)}$ $(j = 1, \dots, p)$, $\widehat{y}_{N,l}^{(0)} = y_l^{(0)}$ $(l = 1, \dots, q)$, $z_N^{(0)} = z^{(0)}$, *then there exist* $\overline{\delta} > 0$ *and* $\widehat{t} \in (0, 1)$ *such that for any* $(u^{(0)}, v^{(0)}, y^{(0)}, z^{(0)}, t) \in B_{\overline{\delta}}(u^*, v^*, y^*, z^*) \times (0, \widehat{t})$, $\widehat{x}_N^{(k)}$ *converges to* x^* *with probability one,* $\widehat{u}_N^{(k)}$ *converges to* u^* *with probability one,* $\widehat{v}_N^{(k)}$ *converges to* v^* *with probability one, and* $\widehat{y}_N^{(k)}$ *converges to* y^* *with probability one, respectively, when* $N \longrightarrow \infty$ *and* $k \longrightarrow \infty$.

Proof. Based on the property of norm, one has

$$\left\| \widehat{x}_N^{(k)} - x^* \right\| \le \left\| \widehat{x}_N^{(k)} - x^{(k)} \right\| + \left\| x^{(k)} - x^* \right\|, \tag{32}$$

$$\left\| \widehat{u}_N^{(k)} - u^* \right\| \le \left\| \widehat{u}_N^{(k)} - u^{(k)} \right\| + \left\| u^{(k)} - u^* \right\|, \tag{33}$$

$$\left\| \widehat{v}_N^{(k)} - v^* \right\| \le \left\| \widehat{v}_N^{(k)} - v^{(k)} \right\| + \left\| v^{(k)} - v^* \right\|, \tag{34}$$

$$\left\| \widehat{y}_N^{(k)} - y^* \right\| \le \left\| \widehat{y}_N^{(k)} - y^{(k)} \right\| + \left\| y^{(k)} - y^* \right\|, \tag{35}$$

$$\left\| \widehat{z}_N^{(k)} - z^* \right\| \le \left\| \widehat{z}_N^{(k)} - z^{(k)} \right\| + \left\| z^{(k)} - z^* \right\|. \tag{36}$$

If assumptions (d)-(h) are satisfied, then it follows from Theorem 3.1 of [13] that there exist $\overline{\delta} > 0$ $(\overline{\delta} < \delta)$ and $\widehat{t} \in (0, 1)$ such that as $k \longrightarrow \infty$, $x^{(k)} \longrightarrow x^*$, $u^{(k)} \longrightarrow u^*$, $v^{(k)} \longrightarrow v^*$, $y^{(k)} \longrightarrow y^*$, and $z^{(k)} \longrightarrow z^*$ for any $(u^{(0)}, v^{(0)}, y^{(0)}, z^{(0)}, t) \in B_{\overline{\delta}}(u^*, v^*, y^*, z^*) \times (0, \widehat{t})$.

If assumptions (a)-(c) are satisfied, and $\widehat{u}_{N,i}^{(0)} = u_i^{(0)}$ $(i = 1, \dots, m)$, $\widehat{v}_{N,j}^{(0)} = v_j^{(0)}$ $(j = 1, \dots, p)$, $\widehat{y}_{N,l}^{(0)} = y_l^{(0)}$ $(l = 1, \dots, q)$ and $\widehat{z}_N^{(0)} = z^{(0)}$, it follows from Theorem 6 that $\widehat{u}_N^{(k)}$ converges to $u^{(k)}$ with probability one, $\widehat{v}_N^{(k)}$ converges to $v^{(k)}$ with probability one, $\widehat{y}_N^{(k)}$ converges to $y^{(k)}$ with probability one, $\widehat{z}_N^{(k)}$ converges to $z^{(k)}$ with probability one, and $\widehat{x}_N^{(k)}$ converges to $x^{(k)}$ with probability one as $N \longrightarrow \infty$.

Thus, combined with (32)-(35) and the above analysis, it has been proven that Theorem 7 is true under assumptions (a)-(h). □

Remark 8. Theorem 7 shows that Algorithm 4 is locally convergent under assumptions (a)-(h). That is, when the initial multiplier $(\widehat{u}_N^{(0)}, \widehat{v}_N^{(0)}, \widehat{y}_N^{(0)})$ are close to the optimal multiplier (u^*, v^*, y^*), $z^{(0)}$ is close to z^*, and t is less than a threshold, the solution sequence $\widehat{x}_N^{(k)}$ obtained by Algorithm 4 locally converges to the optimal solution x^* of original problem (1) with probability one as $N \longrightarrow \infty$, $(\widehat{u}_N^{(k)}, \widehat{v}_N^{(k)}, \widehat{y}_N^{(k)})$ converges to (u^*, v^*, y^*) with probability one as $N \longrightarrow \infty$, and $\widehat{z}_N^{(k)}$) converges to z^* with probability one as $N \longrightarrow \infty$ under assumptions (a)-(h).

4. Numerical Results

In this section, the numerical results for eight test examples by using Algorithm 4 are presented. These test examples are compiled based on the deterministic optimization problems in [13] by considering random variable ξ. The numerical experiments are implemented in Matlab2014 on the same computer whose CPU basic parameters are Intel CORE(TM) i5-3337U@1.80GHZ and memory 8GB.

In the experiments, the random variable ξ is set to be uniformly distributed on $\Xi = [0, 1]$ and the random sample $\{\xi_1, \xi_2, \dots, \xi_N\}$ with sample size N is generated by random number generator *unifrnd* in Matlab2014. We choose $N = 10^2$, $N = 10^3$, $N = 10^4$, $N = 10^5$, and $N = 10^6$ to make comparison for each test example. The initial values of $u_N^{(0)}$, $v_N^{(0)}$, $y_N^{(0)}$ and $z_N^{(0)}$ are set as $u_N^{(0)} = (1/m, \dots, 1/m)^T$, $v_N^{(0)} = (1, \cdots, 1)^T$, $y_N^{(0)} = (1, \cdots, 1)^T$ and $z_N^{(0)} = 1$ for each example, and $t^{(0)}$ is chosen small enough and determined by the scale of test problem; Unconstrained minimization problem in Step 2 of Algorithm 4 is solved by BFGS quasi-Newton method combined with Wolf nonexact linear search rule. The stopping precision in Step 3 is $\varepsilon = 10^{-5}$, and the termination condition is

$$\begin{aligned} &\left| \sum_{i \in M} \widehat{u}_{N,i}^{(k)} \left(\widehat{F}_{N,i} \left(\widehat{x}_N^{(k)} \right) - T \left(\widehat{x}_N^{(k)} \right) \right) \right| \\ &\quad + \left| \sum_{j \in P} \widehat{v}_{N,j}^{(k)} \widehat{G}_{N,j} \left(\widehat{x}_N^{(k)} \right) \right| + \left\| \widehat{H}_N \left(\widehat{x}_N^{(k)} \right) \right\| \le \varepsilon. \end{aligned} \tag{37}$$

The obtained numerical results are reported in Tables 1-8, in which N represents sample size; $1/t$ represents the value of $1/t$; *it* represents the number of iterations, i.e., the numbers of the Lagrange multipliers being updated; $\|\widehat{x}_N^{(k)} - x^*\|$ represents the gap between the solution sequence $\widehat{x}_N^{(k)}$ obtained by Algorithm 4 and the optimal solution x^* of the corresponding test problem; and $\|z_N^{(k)} - z^*\|$ represents the gap between the approximate value $z_N^{(k)}$ of objective function $T(x)$ obtained by Algorithm 4 and the optimal value z^* of the corresponding test problem, respectively.

Example 1. In problem (1), $f_i(x, \xi)$ $(i = 1, 2, 3)$ and $h_l(x, \xi)$ $(l = 1, 2)$ are defined as follows:

$$\begin{aligned} f_1(x, \xi) &= 3\xi^2 x_1^2 + x_2^2, \\ f_2(x, \xi) &= (2 - x_1)^2 + 2\xi (2 - x_2)^2, \\ f_3(x, \xi) &= 4\xi \exp(x_2 - x_1), \\ h_1(x, \xi) &= x_1 + 2\xi x_2 - 4\xi, \\ h_2(x, \xi) &= -x_1^2 - x_2^2 + 6.75\xi^2, \end{aligned} \tag{38}$$

Table 1: The numerical results for Example 1.

N	$\frac{1}{t}$	it	$\lVert\hat{x}_N^{(k)} - x^*\rVert$	$\lVert\hat{z}_N^{(k)} - z^*\rVert$
10^2	3	30	0.053618	2.714061e-02
10^3	3	30	0.013345	5.746432e-02
10^4	3	30	0.002998	1.567222e-03
10^5	3	34	0.001830	6.904709e-03
10^6	3	34	0.000273	2.444700e-04

Table 2: The numerical results for Example 2.

N	$\frac{1}{t}$	it	$\lVert\hat{x}_N^{(k)} - x^*\rVert$	$\lVert\hat{z}_N^{(k)} - z^*\rVert$
10^2	0.2	48	0.040829	1.723749e-02
10^3	0.2	73	0.004546	1.666095e-03
10^4	0.2	89	0.001116	3.148451e-03
10^5	0.2	104	0.000239	8.722890e-04
10^6	0.2	107	0.000188	4.128465e-04

where the optimal solution and the optimal value (see [13]) are

$$\begin{aligned} x^* &= (1.35355, 0.64645)^T, \\ z^* &= 2.25. \end{aligned} \tag{39}$$

The numerical results for this example obtained by Algorithm 4 are shown in Table 1.

Example 2. In problem (1), $f_i(x,\xi)$ $(i = 1,2,3)$, $g_j(x,\xi)$ $(j = 1)$, and $h_l(x,\xi)$ $(l = 1)$ are defined as follows:

$$\begin{aligned} f_1(x,\xi) &= 3\xi^2 x_1^4 + x_2^2, \\ f_2(x,\xi) &= (2 - x_1)^2 + 2\xi(2 - x_2)^2, \\ f_3(x,\xi) &= 4\xi \exp(x_2 - x_1), \\ g_1(x,\xi) &= -4\xi x_1^3 - 3\xi^2 x_2^2, \\ h_1(x,\xi) &= x_1^2 - 2\xi x_2^2, \end{aligned} \tag{40}$$

where the optimal solution and the optimal value (see [13]) are

$$\begin{aligned} x^* &= (1,1)^T, \\ z^* &= 2. \end{aligned} \tag{41}$$

The numerical results for this example obtained by Algorithm 4 are shown in Table 2.

Example 3. In problem (1), $f_i(x,\xi)$ $(i = 1,2)$ and $h_l(x,\xi)$ $(l = 1,2)$ are defined as follows:

$$\begin{aligned} f_1(x,\xi) &= 2\xi \exp\left(\frac{x_1^2}{1000} + (x_2 - 1)^2\right), \\ f_2(x,\xi) &= 3\xi^2 \exp\left(\frac{x_1^2}{1000} + (x_2 + 1)^2\right), \\ h_1(x,\xi) &= \frac{x_1^2}{1000} + 2\xi x_2^2 + 3\xi^2 x_1 x_2, \\ h_2(x,\xi) &= -2\xi x_1 + x_2^2, \end{aligned} \tag{42}$$

where the optimal solution and the optimal value (see [13]) are

$$\begin{aligned} x^* &= (0,0)^T, \\ z^* &= 2.71828. \end{aligned} \tag{43}$$

The numerical results for this example obtained by Algorithm 4 are shown in Table 3.

Example 4. In problem (1), $f_i(x,\xi)$ $(i = 1,2,3)$ and $h_l(x,\xi)$ $(l = 1,2)$ are defined as follows:

$$\begin{aligned} f_1(x,\xi) &= 0.5\left(2\xi x_1 + \frac{10x_1}{(x_1 + 0.1)} + 4\xi x_2^2\right), \\ f_2(x,\xi) &= 0.5\left(-2\xi x_1 + \frac{10x_1}{(x_1 + 0.1)} + 4\xi x_2^2\right), \\ f_3(x,\xi) &= 0.5\left(2\xi x_1 - \frac{10x_1}{(x_1 + 0.1)} - 4\xi x_2^2\right), \\ h_1(x,\xi) &= 2\xi x_1^2 + 2\xi x_2^2 + x_1 x_2, \\ h_2(x,\xi) &= -2\xi x_1 + x_2^2, \end{aligned} \tag{44}$$

Table 3: The numerical results for Example 3.

N	$\frac{1}{t}$	it	$\Vert \hat{x}_N^{(k)} - x^* \Vert$	$\Vert \hat{z}_N^{(k)} - z^* \Vert$
10^2	0.1	23	0.000755	4.628155e-04
10^3	0.1	26	0.000408	1.361478e-04
10^4	0.1	26	0.000363	5.530880e-04
10^5	0.1	26	0.000191	2.105646e-04
10^6	0.1	26	0.000051	1.955180e-05

where the optimal solution and the optimal value (see [13]) are

$$\begin{aligned} x^* &= (0,0)^T, \\ z^* &= 2.71828. \end{aligned} \tag{45}$$

The numerical results for this example obtained by Algorithm 4 are shown in Table 4.

Example 5. In problem (1), $f_i(x,\xi)$ $(i = 1,\ldots,5)$, $g_j(x,\xi)$ $(j = 1,2,3)$, and $h_l(x,\xi)$ $(l = 1)$ are defined as follows:

$$\begin{aligned} f_1(x,\xi) &= 2x_1^2 + 4\xi x_2^2 + 2\xi x_3^2 + 2x_1x_2 + 4\xi x_1x_3 \\ &\quad - 24\xi^2 x_1 - 6x_2 - 8\xi x_3 + 9, \\ f_2(x,\xi) &= f_1(x,\xi) + 10h_1(x,\xi), \\ f_3(x,\xi) &= f_1(x,\xi) + 10g_1(x,\xi), \\ f_4(x,\xi) &= f_1(x,\xi) + 10g_2(x,\xi), \\ f_5(x,\xi) &= f_1(x,\xi) + 10g_3(x,\xi), \\ g_1(x,\xi) &= -2\xi x_1, \\ g_2(x,\xi) &= -3\xi^2 x_2, \\ g_3(x,\xi) &= -2\xi x_3, \\ h_1(x,\xi) &= x_1 + x_2 + 8\xi^3 x_3 - 3, \end{aligned} \tag{46}$$

where the optimal solution and the optimal value (see [13]) are

$$\begin{aligned} x^* &= \left(\frac{4}{3}, \frac{7}{9}, \frac{4}{9}\right)^T, \\ z^* &= \frac{1}{9}. \end{aligned} \tag{47}$$

The numerical results for this example obtained by Algorithm 4 are shown in Table 5.

Example 6. In problem (1), $f_i(x,\xi)$ $(i = 1,\ldots,9)$, $g_j(x,\xi)$ $(j = 1,2)$ and $h_l(x,\xi)$ $(l = 1,\ldots,6)$ are defined as follows:

$$\begin{aligned} f_1(x,\xi) &= x_1^2 + x_2^2 + 2\xi x_1x_2 - 28\xi x_1 - 16x_2 + (x_3 \\ &\quad - 10)^2 + 12\xi^2(x_4 - 5)^2 + 2\xi(x_5 - 3)^2 + 2(x_6 \\ &\quad - 1)^2 + 20\xi^3 x_7^2 + 7(x_8 - 11)^2 + 4\xi(x_9 - 10)^2 \\ &\quad + (x_{10} - 7)^2 + 90\xi, \\ f_2(x,\xi) &= f_1(x,\xi) + 20h_1(x,\xi), \\ f_3(x,\xi) &= f_1(x,\xi) + 20h_2(x,\xi), \\ f_4(x,\xi) &= f_1(x,\xi) + 20g_1(x,\xi), \\ f_5(x,\xi) &= f_1(x,\xi) + 20h_3(x,\xi), \\ f_6(x,\xi) &= f_1(x,\xi) + 20h_4(x,\xi), \\ f_7(x,\xi) &= f_1(x,\xi) + 20h_5(x,\xi), \\ f_8(x,\xi) &= f_1(x,\xi) + 20g_2(x,\xi), \\ f_9(x,\xi) &= f_1(x,\xi) + 20h_6(x,\xi), \\ g_1(x,\xi) &= -\left(-\frac{1}{2(x_1 - 8)^2} - 4\xi(x_2 - 2)^2 - 9\xi^2 x_5^2 \right. \\ &\quad \left. + 2\xi x_6 + 30\right), \\ g_2(x,\xi) &= -\left(3x_1 - 12\xi x_2 - 12(x_9 - 8)^2 + 7x_{10}\right), \\ h_1(x,\xi) &= -\left(-3(x_1 - 2)^2 - 8\xi(x_2 - 3)^2 - 6\xi^2 x_3^2 \right. \\ &\quad \left. + 7x_4 + 240\xi\right), \\ h_2(x,\xi) &= -\left(-5x_1^2 - 16\xi x_2 - (x_3 - 6)^2 + 4\xi x_4 \right. \\ &\quad \left. + 120\xi^2\right), \\ h_3(x,\xi) &= -\left(-x_1^2 - 4\xi(x_2 - 2)^2 + 6\xi^2 x_1x_2 - 14x_5 \right. \\ &\quad \left. + 6x_6\right), \\ h_4(x,\xi) &= -\left(-4x_1 - 10\xi x_2 + 9\xi^2 x_7 - 9x_8 + 105\right), \end{aligned}$$

TABLE 4: The numerical results for Example 4.

N	$\frac{1}{t}$	it	$\Vert\widehat{x}_N^{(k)} - x^*\Vert$	$\Vert\widehat{z}_N^{(k)} - z^*\Vert$
10^2	0.5	4	0.000000	4.061068e-07
10^3	0.5	4	0.000000	4.053219e-07
10^4	0.5	4	0.000000	4.050488e-07
10^5	0.5	4	0.000000	4.050740e-07
10^6	0.5	4	0.000000	4.050678e-07

TABLE 5: The numerical results for Example 5.

N	$\frac{1}{t}$	it	$\left\Vert\widehat{x}_N^{(k)} - x^*\right\Vert$	$\left\Vert\widehat{z}_N^{(k)} - z^*\right\Vert$
10^2	0.8	18	0.350233	7.913036e-01
10^3	0.8	18	0.199523	4.201455e-01
10^4	0.8	17	0.032177	5.933858e-02
10^5	0.8	17	0.006590	1.221448e-03
10^6	0.8	17	0.000542	8.848920e-04

$$h_3(x,\xi) = -\left(-x_1^2 - 4\xi(x_2-2)^2 + 6\xi^2 x_1 x_2 - 14x_5 + 6x_6\right),$$
$$h_4(x,\xi) = -\left(-4x_1 - 10\xi x_2 + 9\xi^2 x_7 - 9x_8 + 105\right),$$
$$h_6(x,\xi) = -\left(8x_1 - 2x_2 - 10\xi x_9 + 6\xi^2 x_{10} + 12\right), \quad (48)$$

where the optimal solution and the optimal value (see [13]) are

$$x^* = (2.171996, 2.363683, 8.773926, 5.095985, 0.990655, 1.430574, 1.321644, 9.828726, 8.280092, 8.375927)^T,$$
$$z^* = 24.306209. \quad (49)$$

The numerical results for this example obtained by Algorithm 4 are shown in Table 6.

Example 7. In problem (1), $f_i(x,\xi)$ $(i = 1,\dots,5)$, $g_j(x,\xi)$ $(j = 1,2)$, and $h_l(x,\xi)$ $(l = 1,2)$ are defined as follows:

$$\begin{aligned}
f_1(x,\xi) &= (x_1-10)^2 + 10\xi(x_2-12)^2 + 5\xi^4 x_3^4 \\
&\quad + 3(x_4-11)^2 + 20\xi x_5^6 + 7x_6^2 + 3\xi^2 x_7^4 - 4x_6 x_7 \\
&\quad - 20\xi x_6 - 8x_7, \\
f_2(x,\xi) &= f_1(x,\xi) + 10h_1(x,\xi), \\
f_3(x,\xi) &= f_1(x,\xi) + 10g_1(x,\xi), \\
f_4(x,\xi) &= f_1(x,\xi) + 10g_2(x,\xi), \\
f_5(x,\xi) &= f_1(x,\xi) + 10h_2(x,\xi), \\
g_1(x,\xi) &= -\left(-7x_1 - 6\xi x_2 - 10x_3^2 - 2\xi x_4 + 3\xi^2 x_5 + 282\right), \\
g_2(x,\xi) &= -\left(-23x_1 - x_2^2 - 18\xi^2 x_6^2 + 16\xi x_7 + 196\right), \\
h_1(x,\xi) &= -\left(-2x_1^2 - 6\xi x_2^4 - x_3 - 8\xi x_4^2 - 5x_5 + 127\right), \\
h_2(x,\xi) &= -\left(-4x_1^2 - x_2^2 + 3x_1 x_2 - 6\xi^2 x_3^2 - 10\xi x_6 + 11x_7\right),
\end{aligned} \quad (50)$$

where the optimal solution and the optimal value (see [13]) are

$$x^* = (2.33050, 1.95137, -0.47754, 4.36573, -0.62449, 1.03813, 1.59423)^T,$$
$$z^* = 680.63006. \quad (51)$$

The numerical results for this example obtained by Algorithm 4 are shown in Table 7.

Table 6: The numerical results for Example 6.

N	$\frac{1}{t}$	it	$\|\hat{x}_N^{(k)} - x^*\|$	$\|\hat{z}_N^{(k)} - z^*\|$
10^2	0.8	21	0.175122	2.486983e-01
10^3	0.8	21	0.017835	1.832466e-02
10^4	0.8	21	0.014858	5.719731e-02
10^5	0.8	21	0.004011	5.719731e-03
10^6	0.8	21	0.000949	4.484174e-04

Table 7: The numerical results for Example 7.

N	$\frac{1}{t}$	it	$\|\hat{x}_N^{(k)} - x^*\|$	$\|\hat{z}_N^{(k)} - z^*\|$
10^2	1	17	0.105795	2.061338e-01
10^3	1	17	0.015291	2.527593e-02
10^4	1	17	0.002059	4.118053e-03
10^5	1	13	0.001702	1.004443e-03
10^6	1	21	0.000676	9.903274e-04

Example 8. In problem (1), $f_i(x,\xi)$ $(i = 1,\ldots,18)$, $g_j(x,\xi)$ $(j = 1,\ldots,5)$, and $h_l(x,\xi)$ $(l = 1,\ldots,12)$ are defined as follows:

$$\begin{aligned}
f_1(x,\xi) &= x_1^2 + x_2^2 + 2\xi x_1 x_2 - 28\xi x_1 - 16x_2 + (x_3 - 10)^2 + 12\xi^2 (x_4 - 5)^2 + 2\xi (x_5 - 3)^2 + 2(x_6 - 1)^2 + 20\xi^3 x_7^2 + 7(x_8 - 11)^2 + 4\xi (x_9 - 10)^2 \\
&\quad + (x_{10} - 7)^2 + 2\xi (x_{11} - 9)^2 + 10(x_{12} - 1)^2 + 10\xi (x_{13} - 7)^2 + 4(x_{14} - 14)^2 + 27(x_{15} - 1)^2 + 5\xi^4 x_{16}^4 + 2\xi (x_{17} - 2)^2 + 13(x_{18} - 2)^2 + (x_{19} - 2)^2 + 2\xi x_{20}^2 + 95,
\end{aligned}$$

$$\begin{aligned}
f_2(x,\xi) &= f_1(x,\xi) + 20h_1(x,\xi), \\
f_3(x,\xi) &= f_1(x,\xi) + 20h_2(x,\xi), \\
f_4(x,\xi) &= f_1(x,\xi) + 20g_1(x,\xi), \\
f_5(x,\xi) &= f_1(x,\xi) + 20h_3(x,\xi), \\
f_6(x,\xi) &= f_1(x,\xi) + 20h_4(x,\xi, \\
f_7(x,\xi) &= f_1(x,\xi) + 20h_5(x,\xi), \\
f_8(x,\xi) &= f_1(x,\xi) + 20g_2(x,\xi), \\
f_9(x,\xi) &= f_1(x,\xi) + 20h_6(x,\xi), \\
f_{10}(x,\xi) &= f_1(x,\xi) + 20g_3(x,\xi), \\
f_{11}(x,\xi) &= f_1(x,\xi) + 20h_7(x,\xi), \\
f_{12}(x,\xi) &= f_1(x,\xi) + 20h_8(x,\xi), \\
f_{13}(x,\xi) &= f_1(x,\xi) + 20g_4(x,\xi), \\
f_{14}(x,\xi) &= f_1(x,\xi) + 20g_5(x,\xi), \\
f_{15}(x,\xi) &= f_1(x,\xi) + 20h_9(x,\xi), \\
f_{16}(x,\xi) &= f_1(x,\xi) + 20h_{10}(x,\xi), \\
f_{17}(x,\xi) &= f_1(x,\xi) + 20h_{11}(x,\xi), \\
f_{18}(x,\xi) &= f_1(x,\xi) + 20h_{12}(x,\xi),
\end{aligned}$$

$$g_1(x,\xi) = -\left(-\frac{1}{2(x_1 - 8)^2} - 4\xi (x_2 - 2)^2 - 9\xi^2 x_5^2 + 2\xi x_6 + 30\right),$$

$$\begin{aligned}
g_2(x,\xi) &= -\left(3x_1 - 12\xi x_2 - 12(x_9 - 8)^2 + 7x_{10}\right), \\
g_3(x,\xi) &= -\left(-2\xi x_1 - x_2 - 8\xi x_{11} + 21x_{12}\right), \\
g_4(x,\xi) &= -\left(-3x_1 - 12\xi^2 x_2 - 6\xi (x_{13} - 6)^2 + 14x_{14} + 10\right), \\
g_5(x,\xi) &= -\left(-14x_1^2 - 70\xi x_{15} + 79x_{16} + 92\right), \\
h_1(x,\xi) &= -\left(-3(x_1 - 2)^2 - 8\xi (x_2 - 3)^2 - 6\xi^2 x_3^2 + 7x_4 + 240\xi\right), \\
h_2(x,\xi) &= -\left(-5x_1^2 - 16\xi x_2 - (x_3 - 6)^2 + 4\xi x_4 + 120\xi^2\right), \\
h_3(x,\xi) &= -\left(-x_1^2 - 4\xi (x_2 - 2)^2 + 6\xi^2 x_1 x_2 - 14x_5 + 6x_6\right), \\
h_4(x,\xi) &= -\left(-4x_1 - 10\xi x_2 + 9\xi^2 x_7 - 9x_8 + 105\right), \\
h_5(x,\xi) &= -\left(-10x_1 + 16\xi x_2 + 17x_7 - 4\xi x_8\right),
\end{aligned}$$

Table 8: The numerical results for Example 8.

N	$\frac{1}{t}$	it	$\|\hat{x}_N^{(k)} - x^*\|$	$\|\hat{z}_N^{(k)} - z^*\|$
10^2	0.5	30	0.236371	2.061338e-01
10^3	0.5	30	0.107028	2.527593e-01
10^4	0.5	30	0.031119	4.118053e-02
10^5	0.5	30	0.003603	5.506888e-03
10^6	0.5	30	0.001569	2.177321e-03

$$\begin{aligned}
h_6(x,\xi) &= -\left(8x_1 - 2x_2 - 10\xi x_9 + 6\xi^2 x_{10} + 12\right),\\
h_7(x,\xi) &= -\left(-x_1^2 - 15x_{11} + 16\xi x_{12} + 28\right),\\
h_8(x,\xi) &= -\left(-4x_1 - 18\xi x_2 - 10\xi x_{13}^2 + 27\xi^2 x_{14} + 87\right),\\
h_9(x,\xi) &= -\left(-15x_2^2 - 22\xi x_{15} + 183\xi^2 x_{16} + 54\right),\\
h_{10}(x,\xi) &= -\left(-5x_1^2 - 2x_2 - 9x_{17}^4 + 4\xi^3 x_{18} + 68\right),\\
h_{11}(x,\xi) &= -\left(-x_1^2 + x_2 - 38\xi x_{19} + 40\xi x_{20} - 19\right),\\
h_{12}(x,\xi) &= -\left(-7x_1^2 - 5x_2^2 - 2\xi x_{19}^2 + 30x_{20}\right),
\end{aligned} \quad (52)$$

where the optimal solution and the optimal value (see [13]) are

$$\begin{aligned}
x^* &= (2.175216, 2.352850, 8.766448, 5.066932, 0.988667, 1.431000, 1.329483,\\
&\quad 9.835926, 8.287277, 8.370178, 2.275828, 1.358623, 6.077186, 14.170830, 0.9962345, 0.655691, 1.466590, 2.000361, 1.046588, 2.063194)^T,\\
z^* &= 133.7283.
\end{aligned} \quad (53)$$

The numerical results for this example obtained by Algorithm 4 are shown in Table 8.

From the numerical results in Tables 1-8, the following remarks are made.

Remark 9. The preliminary numerical results show that Algorithm 4 is feasible and promising.

Remark 10. Compared with the numerical results for the same test example with the different sample size N, the numerical results in Tables 1-8 show that the precisions of the optimal solution and the optimal value by Algorithm 4 become higher as the sample size is chosen larger, which coincides with the convergence result of Algorithm 4 in Section 3.

5. Conclusions

An implementable SAA nonlinear Lagrange algorithm for solving constrained minimax stochastic optimization problems is presented by this paper. And the convergence theory of the proposed algorithm is established under some assumptions, in which the KKT solution sequence obtained by the algorithm is demonstrated to converge to the optimal KKT solution of the original problem with probability one as the sample size approaches to infinity. Furthermore, numerical experiments are implemented by using the proposed SAA nonlinear Lagrange algorithm for solving eight typical test examples, and the results of numerical experiment verify the convergence theory and indicate that the new algorithm is promising. Moreover, the numerical experiments for obtaining the solutions with higher precision and solving large scale problems deserve our future attention. And applying this proposed algorithm to solve some practical problems is also interesting.

Data Availability

The [numerical examples and results] data used to support the findings of this study are included within the article, which are given in Section 4 of the manuscript.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research is supported by the National Natural Science Foundation of China under project (no. 11201357) and the Fundamental Research Funds for the Central Universities under project (no. 2018IB016).

References

[1] M. Breton and S. E. Hachem, "Algorithms for the solution of stochastic dynamic minimax problems," *Computational optimization and applications*, vol. 4, no. 4, pp. 317–345, 1995.

[2] R. C. Hu, H. Xiong, W. L. Jin, and W. Q. Zhu, "Stochastic minimax semi-active control for MDOF nonlinear uncertain systems under combined harmonic and wide-band noise excitations using MR dampers," *International Journal of Non-Linear Mechanics*, vol. 83, pp. 26–38, 2016.

[3] Y. Ermoliev, A. Gaivoronski, and C. Nedeva, "Stochastic optimization problems with partially known distribution functions," *SIAM Journal on Control and Optimization*, vol. 23, no. 5, pp. 697–716, 1985.

[4] K. Frauendorfer and P. Kall, "A solution method for SLP recourse problems with arbitrary multivariate distributions-the independent case," *Problems of Control and Information Theory. Problemy Upravlenija i Teorii Informacii*, vol. 17, no. 4, pp. 177–205, 1988.

[5] M. Riis and K. A. Andersen, "Applying the minimax criterion in stochastic recourse programs," *European Journal of Operational Research*, vol. 165, no. 3, pp. 569–584, 2005.

[6] A. Shapiro, "Minimax and risk averse multistage stochastic programming," *European Journal of Operational Research*, vol. 219, no. 3, pp. 719–726, 2012.

[7] Y. Wang, Z. Ying, and W. Zhu, "Stochastic minimax control for stabilizing uncertain quasi-integrable Hamiltonian systems," *Automatica*, vol. 45, no. 8, pp. 1847–1853, 2009.

[8] A. M. Bagirov, A. Al Nuaimat, and N. Sultanova, "Hyperbolic smoothing function method for minimax problems," *Optimization. A Journal of Mathematical Programming and Operations Research*, vol. 62, no. 6, pp. 759–782, 2013.

[9] C. Charalambous, "Acceleration of the least pth algorithm for minimax optimization with engineering applications," *Mathematical Programming*, vol. 19, pp. 270–297, 1979.

[10] G. Di Pillo, L. Grippo, and S. Lucidi, "A smooth method for the finite minimax problem," *Mathematical Programming*, vol. 60, no. 2, pp. 187–214, 1993.

[11] S. He and S. Zhou, "A nonlinear augmented Lagrangian for constrained minimax problems," *Applied Mathematics and Computation*, vol. 218, no. 8, pp. 4567–4579, 2011.

[12] S. He and Y. Nie, "A class of nonlinear Lagrangian algorithms for minimax problems," *Journal of Industrial and Management Optimization*, vol. 9, no. 1, pp. 75–97, 2013.

[13] S. He, X. Liu, and C. Wang, "A nonlinear Lagrange algorithm for minimax problems with general constraints," *Numerical Functional Analysis and Optimization*, vol. 37, no. 6, pp. 680–698, 2016.

[14] J. Jian, X. Zhang, R. Quan, and Q. Ma, "Generalized monotone line search SQP algorithm for constrained minimax problems," *Optimization. A Journal of Mathematical Programming and Operations Research*, vol. 58, no. 1, pp. 101–131, 2009.

[15] R. A. Polyak, "Smooth optimization methods for minimax problems," *SIAM Journal on Control and Optimization*, vol. 26, no. 6, pp. 1274–1286, 1988.

[16] F. Wang, "A hybrid algorithm for linearly constrained minimax problems," *Annals of Operations Research*, vol. 206, pp. 501–525, 2013.

[17] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, MPS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 2009.

[18] S. Ahmed and A. Shapiro, "The sample average approximation method for stochastic programs with integer recourse," *SIAM Journal on Optimization*, vol. 12, pp. 479–502, 2002.

[19] H. Mohammadi Bidhandi and J. Patrick, "Accelerated sample average approximation method for two-stage stochastic programming with binary first-stage variables," *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 41, pp. 582–595, 2017.

[20] X. Chen, R. J. Wets, and Y. Zhang, "Stochastic variational inequalities: residual minimization smoothing sample average approximations," *SIAM Journal on Optimization*, vol. 22, no. 2, pp. 649–673, 2012.

[21] S. He, P. Zhang, X. Hu, and R. Hu, "A sample average approximation method based on a D-gap function for stochastic variational inequality problems," *Journal of Industrial and Management Optimization*, vol. 10, no. 3, pp. 977–987, 2014.

[22] S. He, M. Wei, and H. Tong, "A smooth penalty-based sample average approximation method for stochastic complementarity problems," *Journal of Computational and Applied Mathematics*, vol. 287, pp. 20–31, 2015.

[23] G.-H. Lin, X. Chen, and M. Fukushima, "New restricted NCP functions and their applications to stochastic NCP and stochastic MPEC," *Optimization. A Journal of Mathematical Programming and Operations Research*, vol. 56, no. 5-6, pp. 641–953, 2007.

[24] M. Wang and M. M. Ali, "Stochastic nonlinear complementarity problems: stochastic programming reformulation and penalty-based approximation method," *Journal of Optimization Theory and Applications*, vol. 144, no. 3, pp. 597–614, 2010.

[25] M. Wang, G. Lin, Y. Gao, and M. Ali, "Sample average approximation method for a class of stochastic variational inequality problems," *Journal of Systems Science & Complexity*, vol. 24, no. 6, pp. 1143–1153, 2011.

[26] H. Xu and F. Meng, "Convergence analysis of sample average approximation methods for a class of stochastic mathematical programs with equality constraints," *Mathematics of Operations Research*, vol. 32, no. 3, pp. 648–668, 2007.

[27] H. Xu, "Sample average approximation methods for a class of stochastic variational inequality problems," *Asia-Pacific Journal of Operational Research*, vol. 27, no. 1, pp. 103–119, 2010.

[28] J. Zhang, L. Zhang, and S. Lin, "A class of smoothing SAA methods for a stochastic mathematical program with complementarity constraints," *Journal of Mathematical Analysis and Applications*, vol. 387, no. 1, pp. 201–220, 2012.

[29] J. Zhang, S.-x. He, and Q. Wang, "A SAA nonlinear regularization method for a stochastic extended vertical linear complementarity problem," *Applied Mathematics and Computation*, vol. 232, pp. 888–897, 2014.

[30] J. Zhang, Y.-Q. Zhang, and L.-W. Zhang, "A sample average approximation regularization method for a stochastic mathematical program with general vertical complementarity constraints," *Journal of Computational and Applied Mathematics*, vol. 280, pp. 202–216, 2015.