Research Article

A Nonlinear Lagrange Algorithm for Stochastic Minimax Problems Based on Sample Average Approximation Method

Suxiang He, Yunyun Nie, and Xiaopeng Wang

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

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

An implementable nonlinear Lagrange algorithm for stochastic minimax problems is presented based on sample average approximation method in this paper, in which the second step minimizes a nonlinear Lagrange function with sample average approximation functions of original functions and the sample average approximation of the Lagrange multiplier is adopted. Under a set of mild assumptions, it is proven that the sequences of solution and multiplier obtained by the proposed algorithm converge to the Kuhn-Tucker pair of the original problem with probability one as the sample size increases. At last, the numerical experiments for five test examples are performed and the numerical results indicate that the algorithm is promising.

1. Introduction

Consider the stochastic minimax problems of the form

$$\min_{x \in \Re^n}\text{-}\max_{i \in I} \mathbb{E}\left[f_i(x, \omega) \right], \tag{1}$$

where $I = \{1, \ldots, m\}$, ω is a random vector supported on the probability space $(\Omega, \mathscr{F}, \mathscr{P})$, $f_i : \Re^n \times \Omega \to \Re(i \in I)$, $\mathbb{E}$ denotes expectation with respect to the distribution of $\omega \in \Omega$, and $\mathbb{E}[f_i(x, \omega)]$ $(i \in I)$ is well defined. Problem (1) has drawn much attention in recent years, which arises in various situations such as inventory theory, robust optimization, and engineering filed; for example, see [1–5].

A nonlinear Lagrange function for problem (1) can be established based on Zhang and Tang [6]; that is,

$$G(x, u, t) = t \ln\left(\sum_{i \in I} u_i e^{(1/t)\mathbb{E}[f_i(x,\omega)]} \right), \tag{2}$$

where u is Lagrange multiplier and $u \in \Gamma(u) = \{u \in \Re^m \mid \sum_{i \in I} u_i = 1, u_i \geq 0\}$, and $t > 0$ is a controlling parameter. The good properties of function (2) were investigated in [6] and the convergence analysis of the corresponding nonlinear Lagrange algorithm was presented in [7]. Although function (2) overcomes the nondifferentiability of the objective function $F(x) = \max_{i \in I} \mathbb{E}[f_i(x, \omega)]$ in problem (1), the exact

numerical evaluation of the expected value $\mathbb{E}[f_i(x, \omega)]$ $(i \in I)$ in (2) is very difficult because either distribution of random vector ω is unknown or it is too complex to compute the multidimensional integral.

The sample average approximation (in short, SAA) method [8–15] is a well-behaved approach for bypassing this difficulty. The idea of SAA method is to generate a random sample $\omega^1, \ldots, \omega^N$ of the random variable $\omega \in \Omega$ with sample size N and approximate the involved expected value function $\mathbb{E}[f_i(x, \omega)]$ $(i \in I)$ by the corresponding sample average function $(1/N) \sum_{j=1}^{N} f_i(x, \omega^j)$ $(i \in I)$. Inspired by the SAA method, we present the SAA function of $G(x, u, t)$ as follows:

$$\widehat{G}_N(x, u, t) = t \ln\left(\sum_{i \in I} u_i e^{(1/t)\widehat{g}_{N,i}(x)} \right), \tag{3}$$

where $\widehat{g}_{N,i}(x) = (1/N) \sum_{j=1}^{N} f_i(x, \omega^j)$. Furthermore, we will propose an implementable nonlinear Lagrange algorithm based on SAA function (3), in which function (3) is minimized and the Lagrange multiplier is updated by its SAA form. Under some mild assumptions on problem (1), we will show that the sequences of solution and multiplier generated by the SAA method-based nonlinear Lagrange algorithm converge to the Kuhn-Tucker pair of the original problem with probability one as the sample size N increases.

The remainder of this paper is organized as follows. Preliminaries are given in Section 2. The SAA method-based nonlinear Lagrange algorithm and convergence analysis are established in Section 3. Section 4 reports the numerical results by using the proposed algorithm to solve five test examples. Finally, conclusions are drawn in Section 5.

2. Preliminaries

This section serves as a preparation for the convergence analysis of the proposed SAA method-based nonlinear Lagrange algorithm. The assumptions on problem (1) are provided firstly. Furthermore, some results that are essential to our discussion are listed. At last, we recall the nonlinear Lagrange algorithm in [7].

Let (x^*, u^*) denote the Kuhn-Tucker pair of problem (1). Let $\delta > 0$ be small enough and define $S(x^*, \delta) = \{x \in \Re^n \mid \|x^* - x\| \le \delta\}$. The Lagrange function for problem (1) is defined by $L(x, u) = \sum_{i \in I} u_i \mathbb{E}[f_i(x, \omega)]$. Set $B(x^*) = \{i \mid \mathbb{E}[f_i(x^*, \omega)] = F(x^*), i \in I\}$ and $f(x, \omega) = (f_1(x, \omega), \ldots, f_m(x, \omega))^T$. We list the following assumptions on problem (1), which will be used in the subsequent theoretical analysis.

(A1) $\mathbb{E}[f(x, \omega)]$ is twice continuously differentiable on $\Re^n$.

(A2) There exists a nonnegative measurable function $\phi(\omega)$ such that $\mathbb{E}[\phi(\omega)]$ is finite and for every $x \in S(x^*, \delta)$ the inequality

$$\|f(x, \omega)\| < \phi(\omega) \tag{4}$$

holds with probability one.

(A3) The random sample $\omega^1, \ldots, \omega^N$ is independent and identically distributed.

(A4) (x^*, u^*) satisfies the K-T condition. That is,

$$\nabla_x L(x^*, u^*) = 0, \quad u^* \in \Gamma(u),$$
$$u_i^*(F(x^*) - \mathbb{E}[f_i(x, \omega)]) = 0, \quad i \in I. \tag{5}$$

(A5) Strict complementary condition holds; that is, $u_i^* > 0$ for $i \in B(x^*)$.

(A6) Linear independent constraint qualification holds. That is,

$$\left\{ \left(\frac{1}{\nabla \mathbb{E}[f_i(x^*, \omega)]} \right) \mid i \in B(x^*) \right\} \tag{6}$$

is a set of linear independent vectors.

(A7) For all $y \in \Re^n$ satisfying $\nabla \mathbb{E}[f_i(x^*, \omega)]^T y = 0, i \in B(x^*)$ ($y \ne 0$), it holds that

$$y^T \nabla_x^2 L(x^*, u^*) y \ge \lambda \|y\|^2, \tag{7}$$

where $\lambda > 0$ is a constant.

Definition 1 (see [11]). For nonempty sets A and B in $\Re^n$, one denotes by $\text{dist}(x, A) = \inf_{x' \in A} \|x - x'\|$ the distance from $x \in \Re^n$ to A and by $\mathbb{D}(A, B) = \sup_{x \in A} \text{dist}(x, B)$ the deviation of the set A from the set B.

Lemma 2 (Heine-Cantor theorem; see [16]). *If $h(x) : M \to N$ is continuous function and M is compact, then $h(x)$ is uniformly continuous, where M and N are two metric spaces.*

Note. An important special case is that every continuous function from a closed interval to the real numbers is uniformly continuous.

Lemma 3. *Define $W_N^j(x), W^j(x) : \Re^n \to \Re$ for $j = 1, \ldots, p$. Suppose that $W_N^j(x)$ converges to $W^j(x)$ with probability one uniformly on $\Re^n$ for $j = 1, \ldots, p$. Then $\sum_{j=1}^p W_N^j(x)$ converges to $\sum_{j=1}^p W^j(x)$ with probability one uniformly on $\Re^n$.*

Proof. From the given condition, for $j = 1, \ldots, p$, one has that, for any $\epsilon > 0$, there exists $N^j \in N_+$ such that when $N > N^j$, $|W_N^j(x) - W^j(x)| < \epsilon/p$ holds with probability one for any $x \in \Re^n$.

Let $N_0 = \max_{1 \le j \le p} \{N^j\}$. Thus we have that, for any $\epsilon > 0$, when $N > N_0$, for any $x \in \Re^n$, the following holds with probability one:

$$\left| \sum_{j=1}^p W_N^j(x) - \sum_{j=1}^p W^j(x) \right| \le \sum_{j=1}^p \left| W_N^j(x) - W^j(x) \right|$$
$$< p \cdot \frac{\epsilon}{p} = \epsilon, \tag{8}$$

which means that Lemma 3 is true. □

Algorithm 4 is from [7].

Algorithm 4. We have the following.

Step 1. Choose $t \in (0, \hat{t})$, where $\hat{t} \in (0, 1)$, $u^{(0)} \in \Gamma(u)$, and $\epsilon \in (0, 1)$ small enough and set $k = 0$.

Step 2. Solve

$$\min_{x \in \Re^n} G(x, u^{(k)}, t) \tag{9}$$

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

Step 3. If $|\sum_{i \in I} u_i^{(k)}(\mathbb{E}[f_i(x^{(k)}, \omega)] - F(x^{(k)}))| < \epsilon$, then stop. Otherwise go to Step 4.

Step 4. Update the Lagrange multiplier $u^{(k)}$ by

$$u_i^{(k+1)} = \frac{u_i^{(k)} e^{(1/t)\mathbb{E}[f_i(x^{(k)})]}}{\sum_{j \in I} u_j^{(k)} e^{(1/t)\mathbb{E}[f_j(x^{(k)})]}}, \quad i = 1, \ldots, m. \tag{10}$$

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

3. The SAA Method-Based Nonlinear Lagrange Algorithm and Its Convergence

In view of the numerical computation difficulty in Algorithm 4 and motivated by the SAA method, we

provide the following implementable nonlinear Lagrange algorithm based on the SAA method firstly. Furthermore we establish the convergence analysis of the SAA method-based algorithm under assumptions (A1)–(A7) in this section.

Implementable SAA method-based **Algorithm 5** is presented as follows.

Algorithm 5. We have the following.

Step 1. Choose $t \in (0, \widehat{t})$, where $\widehat{t} \in (0, 1)$, $\epsilon \in (0, 1)$ small enough, $\widehat{u}_N^{(0)} \in \Gamma(u)$, and N is large enough. Set $k = 0$.

Step 2. Solve

$$\min_{x \in \mathfrak{R}^n} \widehat{G}_N\left(x, \widehat{u}_N^{(k)}, t\right) \qquad (11)$$

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

Step 3. If $|\sum_{i \in I} \widehat{u}_{N,i}^{(k)}(\widehat{g}_{N,i}(\widehat{x}_N^{(k)}) - F(\widehat{x}_N^{(k)}))| < \epsilon$, then stop. Otherwise go to Step 4.

Step 4. Update the Lagrange multiplier $\widehat{u}_N^{(k)}$ by

$$\widehat{u}_{N,i}^{(k+1)} = \frac{\widehat{u}_{N,i}^{(k)} e^{(1/t)\widehat{g}_{N,i}(\widehat{x}_N^{(k)})}}{\sum_{j \in I} \widehat{u}_{N,j}^{(k)} e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(k)})}}, \quad i = 1, \dots, m. \qquad (12)$$

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

Taking into account the local convergence analysis of Algorithm 4 given in [7], next we will study the convergence of the sequence pair $(\widehat{x}_N^{(k)}, \widehat{u}_N^{(k)})$ obtained by **Algorithm 5** on $S(x^*, \delta) \times S(u^*, \delta)$. Let $v^{(k)}$ and $S^{(k)}$ denote the optimal value and the set of optimal solutions of $\min_{x \in S(x^*, \delta)} G(x, u^{(k)}, t)$ and $\widehat{v}_N^{(k)}$ and $\widehat{S}_N^{(k)}$ denote the optimal value and the set of optimal solutions of $\min_{x \in S(x^*, \delta)} \widehat{G}_N(x, \widehat{u}_N^{(k)}, t)$, respectively. Set $g_i(x) = \mathbb{E}[f_i(x, \omega)]$.

Theorem 6. *If assumptions (A1)–(A3) hold and $\widehat{u}_{N,i}^{(k)}$ converges to $u_i^{(k)}$ ($i = 1, \dots, m$) with probability one for some k, then the following statements hold:*

(i) *$\widehat{G}_N(x, \widehat{u}_N^{(k)}, t)$ converges to $G(x, u^{(k)}, t)$ with probability one uniformly on $S(x^*, \delta)$;*

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

Proof. (i) Let $\varphi(a, b) = b e^{a/t}$, where $a, b \in \mathfrak{R}$. Then one has

$$\widehat{G}_N\left(x, \widehat{u}_N^{(k)}, t\right) = t \ln \sum_{i \in I} \varphi\left(\widehat{g}_{N,i}(x), \widehat{u}_{N,i}^{(k)}\right). \qquad (13)$$

Now we prove that $\varphi(\widehat{g}_{N,i}(x), \widehat{u}_{N,i}^{(k)})$ converges to $\varphi(g_i(x), u_i^{(k)})$ with probability one uniformly on $S(x^*, \delta)$. Considering the definition of $\varphi(a, b)$, we have

$$\left| \varphi\left(\widehat{g}_{N,i}(x), \widehat{u}_{N,i}^{(k)}\right) - \varphi\left(g_i(x), u_i^{(k)}\right) \right|$$
$$\leq \left| \varphi\left(\widehat{g}_{N,i}(x), \widehat{u}_{N,i}^{(k)}\right) - \varphi\left(\widehat{g}_{N,i}(x), u_i^{(k)}\right) \right|$$
$$\quad + \left| \varphi\left(\widehat{g}_{N,i}(x), u_i^{(k)}\right) - \varphi\left(g_i(x), u_i^{(k)}\right) \right|$$
$$= \left| \widehat{u}_{N,i}^{(k)} e^{(1/t)\widehat{g}_{N,i}(x)} - u_i^{(k)} e^{(1/t)\widehat{g}_{N,i}(x)} \right|$$
$$\quad + \left| u_i^{(k)} e^{(1/t)\widehat{g}_{N,i}(x)} - u_i^{(k)} e^{(1/t)g_i(x)} \right|$$
$$= e^{(1/t)\widehat{g}_{N,i}(x)} \left| \widehat{u}_{N,i}^{(k)} - u_i^{(k)} \right| + u_i^{(k)} \left| e^{(1/t)\widehat{g}_{N,i}(x)} - e^{(1/t)g_i(x)} \right|. \qquad (14)$$

It follows from assumption (A1) and Theorem 7.48 in [11] that both $\widehat{g}_{N,i}(x)$ and $g_i(x)$ ($i = 1, \dots, m$) are continuous at x on $S(x^*, \delta)$. Consequently, for any $x \in S(x^*, \delta)$, there exist constants a_1 and a_2 such that $\widehat{g}_{N,i}(x) \in [a_1, a_2]$ and $g_i(x) \in [a_1, a_2]$ for $i = 1, \dots, m$. Since $\varphi(a, 1)$ is continuous at a on $\mathfrak{R}$ and from Lemma 2, we have that $\varphi(a, 1)$ is uniformly continuous on $[a_1, a_2]$; that is, for any $\epsilon > 0$ and $x, y \in [a_1, a_2]$, there exists $\widehat{\delta} > 0$ such that, for $|x - y| < \widehat{\delta}$, it holds that

$$\left| \varphi(x, 1) - \varphi(y, 1) \right| < \epsilon. \qquad (15)$$

Furthermore, from Theorem 7.48 in [11] we know that $\widehat{g}_{N,i}(x)$ converges to $g_i(x)$ with probability one uniformly on $S(x^*, \delta)$, which means that, for the above given $\widehat{\delta}$, there exists $N_0 \in N_+$ such that when $N > N_0$, the inequality

$$\left| \widehat{g}_{N,i}(x) - g_i(x) \right| < \widehat{\delta} \qquad (16)$$

holds with probability one for any $x \in S(x^*, \delta)$. In view of formula (15) and formula (16), one draws the conclusion that, for any $\epsilon > 0$, there exists $N_0 \in N_+$ such that when $N > N_0$, it holds that

$$\left| \varphi\left(\widehat{g}_{N,i}(x), 1\right) - \varphi\left(g_i(x), 1\right) \right| < \epsilon \qquad (17)$$

with probability one for any $x \in S(x^*, \delta)$. That is, $e^{(1/t)\widehat{g}_{N,i}(x)}$ converges to $e^{(1/t)g_i(x)}$ with probability one uniformly on $S(x^*, \delta)$.

In view of $\widehat{u}_{N,i}^{(k)}$ converging to $u_i^{(k)}$ with probability one, $e^{(1/t)\widehat{g}_{N,i}(x)}$ being bounded on $S(x^*, \delta)$ with probability one, $u_i^{(k)} \in \Gamma(u)$, and formula (14), we obtain that $\varphi(\widehat{g}_{N,i}(x), \widehat{u}_{N,i}^{(k)})$ converges to $\varphi(g_i(x), u_i^{(k)})$ with probability one uniformly on $S(x^*, \delta)$ as $N \to \infty$. Moreover, one gets that $\sum_{i \in I} \varphi(\widehat{g}_{N,i}(x), \widehat{u}_{N,i}^{(k)})$ converges to $\sum_{i \in I} \varphi(g_i(x), u_i^{(k)})$ with probability one uniformly on $S(x^*, \delta)$ as $N \to \infty$ from Lemma 3.

Next we prove that $\widehat{G}_N(x, \widehat{u}_N^{(k)}, t)$ converges to $G(x, u^{(k)}, t)$ with probability one uniformly on $S(x^*, \delta)$. Let $\Phi_N(x) = \sum_{i \in I} \varphi(\widehat{g}_{N,i}(x), \widehat{u}_{N,i}^{(k)})$ and $\Phi(x) = \sum_{i \in I} \varphi(g_i(x), u_i^{(k)})$. Hence, one has

$$\left| \widehat{G}_N\left(x, \widehat{u}_N^{(k)}, t\right) - G\left(x, u^{(k)}, t\right) \right| = t \left| \ln \Phi_N(x) - \ln \Phi(x) \right|. \qquad (18)$$

From the above discussion, we know that $\Phi_N(x), \Phi(x) \in [e^{a_1/t}, e^{a_2/t}]$ for any $x \in S(x^*, \delta)$. Since $\phi(y) = \ln y$ is continuous at y on $\mathfrak{R}_{++}$ and by Lemma 2, we obtain that $\phi(y)$ is uniformly continuous on the interval $[e^{a_1/t}, e^{a_2/t}]$. That is, for any $\epsilon_0 > 0$ and $x', y' \in [e^{a_1/t}, e^{a_2/t}]$, there exists $\overline{\delta} > 0$ such that, for $|x' - y'| < \overline{\delta}$, it holds that

$$\left| \ln x' - \ln y' \right| < \epsilon_0. \tag{19}$$

Moreover, for the given $\overline{\delta}$, there exists $N_0 \in N_+$ such that, for $N > N_0$, it holds that

$$\left| \Phi_N(x) - \Phi(x) \right| < \overline{\delta} \tag{20}$$

with probability one for any $x \in S(x^*, \delta)$. From formulas (19) and (20), it follows that, for any $\epsilon_0 > 0$, there exists $N_0 \in N_+$ such that, for $N > N_0$, the following inequality holds:

$$\left| \ln \Phi_N(x) - \ln \Phi(x) \right| < \epsilon_0 \tag{21}$$

with probability one for any $x \in S(x^*, \delta)$. Combined with formula (18), statement (i) is true.

(ii) From statement (i) and Theorem 5.3 in [11], statement (ii) is obtained. The proof of Theorem 6 is completed.

□

Theorem 7. *If assumptions (A1)–(A3) hold, letting $\widehat{u}_{N,i}^{(0)} = u_i^{(0)}$, $i = 1, \ldots, m$, then, for any $k \geq 1$, the following statements hold:*

(i) $\widehat{u}_{N,i}^{(k)}$ *converges to $u_i^{(k)}$ with probability one for $i = 1, \ldots, m$;*

(ii) $\widehat{G}_N(x, \widehat{u}_N^{(k)}, t)$ *converges to $G(x, u^{(k)}, t)$ with probability one uniformly on $S(x^*, \delta)$;*

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

Proof. (i) We use the mathematical induction method to show that statement (i) is true below.

(a) Let $k = 1$; then for $i = 1, \ldots, m$ we have

$$\widehat{u}_{N,i}^{(1)} = \frac{\widehat{u}_{N,i}^{(0)} e^{(1/t)\widehat{g}_{N,i}(\widehat{x}_N^{(0)})}}{\sum_{j \in I} \widehat{u}_{N,j}^{(0)} e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(0)})}},$$

$$u_i^{(1)} = \frac{u_i^{(0)} e^{(1/t)g_i(x^{(0)})}}{\sum_{j \in I} u_j^{(0)} e^{(1/t)g_j(x^{(0)})}}. \tag{22}$$

Considering $\widehat{u}_{N,i}^{(0)} = u_i^{(0)}$, we have that $\widehat{x}_N^{(0)}$ converges to $x^{(0)}$ with probability one from Theorem 6. For $j = 1, \ldots, m$, it holds that

$$\left| \widehat{g}_{N,j}(\widehat{x}_N^{(0)}) - g_j(x^{(0)}) \right|$$
$$\leq \left| \widehat{g}_{N,j}(\widehat{x}_N^{(0)}) - g_j(\widehat{x}_N^{(0)}) \right| \tag{23}$$
$$+ \left| g_j(\widehat{x}_N^{(0)}) - g_j(x^{(0)}) \right|.$$

Noting that the first term in the right-hand side of formula (23) converges to 0 with probability one by Theorem 7.48 in [11] and the second term converges to 0 with probability one for $g_j(x)$ being continuous, we obtain that $\widehat{g}_{N,j}(\widehat{x}_N^{(0)})$ converges to $g_j(x^{(0)})$ with probability one. Moreover, since $m(y) = e^{y/t}$ is continuous at y on $\mathfrak{R}$, one gets that $e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(0)})}$ converges to $e^{(1/t)g_j(x^{(0)})}$ with probability one. Then it follows from the properties of convergent sequence that $\widehat{u}_{N,i}^{(1)}$ converges to $u_i^{(1)}$ with probability one for $i = 1, \ldots, m$.

(b) When $k = p$, we assume that $\widehat{u}_{N,i}^{(p)}$ converges to $u_i^{(p)}$ with probability one for $i = 1, \ldots, m$. Then, when $k = p + 1$, next we prove that $\widehat{u}_{N,i}^{(p+1)}$ converges to $u_i^{(p+1)}$ with probability one for $i = 1, \ldots, m$.

Let $k = p + 1$; then for $i = 1, \ldots, m$ one has

$$\widehat{u}_{N,i}^{(p+1)} = \frac{u_i^{(p)} e^{(1/t)\widehat{g}_{N,i}(\widehat{x}_N^{(p)})}}{\sum_{j \in I} u_j^{(p)} e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})}},$$

$$u_i^{(p+1)} = \frac{u_i^{(p)} e^{(1/t)g_i(x^{(p)})}}{\sum_{j \in I} u_j^{(p)} e^{(1/t)g_j(x^{(p)})}}. \tag{24}$$

From Theorem 6, we know that $\widehat{x}_N^{(p)}$ converges to $x^{(p)}$ with probability one as $N \to \infty$. By a similar proof process to that in (a), we have that $e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})}$ converges to $e^{(1/t)g_j(x^{(p)})}$ with probability one for $j = 1, \ldots, m$. For $j = 1, \ldots, m$, one has that

$$\left| \widehat{u}_{N,j}^{(p)} e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})} - u_j^{(p)} e^{(1/t)g_j(x^{(p)})} \right|$$

$$\leq \left| \widehat{u}_{N,j}^{(p)} e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})} - u_j^{(p)} e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})} \right|$$

$$+ \left| u_j^{(p)} e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})} - u_j^{(p)} e^{(1/t)g_j(x^{(p)})} \right| \tag{25}$$

$$= e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})} \left| \widehat{u}_{N,j}^{(p)} - u_j^{(p)} \right|$$

$$+ u_j^{(p)} \left| e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})} - e^{(1/t)g_j(x^{(p)})} \right|.$$

Noting that the first term of (25) tends to 0 with probability one as $N \to \infty$ for $\widehat{u}_{N,j}^{(p)}$ converging to $u_j^{(p)}$ with probability one and $e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})}$ being bounded on $S(x^*, \delta)$ with probability one and the second term tends to 0 with probability one as $N \to \infty$ for $u_j^{(p)} \in (0, 1)$, we obtain that $\widehat{u}_{N,j}^{(p)} e^{(1/t)\widehat{g}_{N,j}(\widehat{x}_N^{(p)})}$ converges to $u_j^{(p)} e^{(1/t)g_j(x^{(p)})}$ with probability one. Then it follows from properties of convergent sequence that $\widehat{u}_{N,i}^{(p+1)}$ converges to $u_i^{(p)}$ with probability one for $i = 1, \ldots, m$.

According to (a) and (b), we have that statement (i) holds.

(ii) From statement (i) and Theorem 6, we obtain that statement (ii) is true.

(iii) From statement (ii) and Theorem 5.3 in [11], one has that statement (iii) holds.

□

The above theorem shows that the sample average approximation Lagrange multiplier $\widehat{u}_N^{(k)}$ converges to its counterpart $u^{(k)}$ with probability one, and the optimal value and optimal solutions of the subproblem $\min_{x \in S(x^*, \delta)} \widehat{G}_N(x, \widehat{u}_N^{(k)}, t)$ converge to their counterparts of the subproblem $\min_{x \in S(x^*, \delta)} G(x, u^{(k)}, t)$ with probability one under some mild conditions. Next we will analyze the convergence of Algorithm 5 under some mild conditions.

Theorem 8. *If assumptions (A1)–(A7) hold, letting $\widehat{u}_{N,i}^{(0)} = u_i^{(0)}$ ($i = 1, \ldots, m$), then there exist $\delta > 0$ and $\widehat{t} \in (0, 1)$ such that, for any $(u^{(0)}, t) \in S(u^*, \delta) \times (0, \widehat{t})$, it holds that the sequence pair $(\widehat{x}_N^{(k)}, \widehat{u}_N^{(k)})$ converge to the K-T pair (x^*, u^*) with probability one.*

Proof. Under assumptions (A1) and (A4)–(A7), from Theorem 3.1 in [7], we have that there exist $\delta > 0$ and $\widehat{t} \in (0, 1)$ such that, for any $k \geq 1$ and $(u^{(0)}, t) \in S(u^*, \delta) \times (0, \widehat{t})$, the following inequality holds:

$$\max \left\{ \left\| x^{(k)} - x^* \right\|, \left\| u^{(k)} - u^* \right\| \right\} \leq c^k t^k \left\| u^{(0)} - u^* \right\|, \quad (26)$$

where $c > 0$ is a constant, which implies that the pair $(x^{(k)}, u^{(k)})$ tend to the K-T pair (x^*, u^*) of the original problem (1) as $k \to \infty$.

Since assumptions (A1)–(A3) hold and $\widehat{u}_{N,i}^{(0)} = u_i^{(0)}$ ($i = 1, \ldots, m$), it follows by Theorem 7 that the pair $(\widehat{x}_N^{(k)}, \widehat{u}_N^{(k)})$ converge to the pair $(x^{(k)}, u^{(k)})$ with probability one as $N \to \infty$.

Furthermore, since

$$\begin{aligned} \left\| \widehat{x}_N^{(k)} - x^* \right\| &\leq \left\| \widehat{x}_N^{(k)} - x^{(k)} \right\| + \left\| x^{(k)} - x^* \right\|, \\ \left\| \widehat{u}_N^{(k)} - u^* \right\| &\leq \left\| \widehat{u}_N^{(k)} - u^{(k)} \right\| + \left\| u^{(k)} - u^* \right\|, \end{aligned} \quad (27)$$

the conclusion is obtained. □

Remark 9. Theorem 8 shows that, under some mild assumptions, the sequence pair $(\widehat{x}_N^{(k)}, \widehat{u}_N^{(k)})$ generated by Algorithm 5 locally tend to the K-T pair (x^*, u^*) of the original problem (1) with probability one as $N \to \infty$ and $k \to \infty$ when the controlling parameter t is less than the threshold $\widehat{t}$.

4. Numerical Results

The numerical results for five test examples by using Algorithm 5 are presented in this section, where the five test problems are compiled based on the deterministic optimization problems in the literature [17, 18]. The numerical experiments are implemented in Matlab 7.1 runtime environment on the same computer, whose basic parameters are Intel CORE i3-2310 M@2.10 GHz and memory 2 Gb.

TABLE 1: The numerical results for Example 1.

N	$1/t$	Iter.	$\|x^* - \widehat{x}_N^{(k)}\|$	$\|\widehat{v}_N^{(k)} - v^*\|$
10^2	10	3	$1.904585e - 001$	$1.929105e - 001$
10^3	10	2	$5.976346e - 002$	$2.127908e - 002$
10^4	10	1	$3.668249e - 002$	$8.282325e - 002$
10^5	10	1	$5.525511e - 003$	$3.098093e - 003$
10^6	10	1	$1.319634e - 003$	$1.279392e - 003$
10^7	10	1	$5.493409e - 004$	$9.275895e - 004$

In the experiments, the sample $\omega^1, \ldots, \omega^N$ with sample size N is generated by *unifrnd* in Matlab 7.1. For each problem, we choose $N = 10^2$, $N = 10^3$, $N = 10^4$, $N = 10^5$, $N = 10^6$, and $N = 10^7$, respectively, to make comparison. The initial value $u_N^{(0)} = (1/m, \ldots, 1/m)^T$ for each example. Unconstrained minimization problem in Step 2 of Algorithm 5 is solved by BFGS quasi-Newton method combined with Wolf nonexact linear search rule, and the control precision is 10^{-6} in this step. The stopping criterion in Step 3 is

$$\left| \sum_{i \in I} \widehat{u}_{N,i}^{(k)} \left(\widehat{g}_{N,i} \left(\widehat{x}_N^{(k)} \right) - F \left(\widehat{x}_N^{(k)} \right) \right) \right| \leq \varepsilon, \quad (28)$$

where $\varepsilon = 10^{-8}$.

The obtained numerical results are reported in Tables 1–5, in which N, $1/t$, iter., $\|x^* - \widehat{x}_N^{(k)}\|$, and $\|\widehat{v}_N^{(k)} - v^*\|$ represent the sample size, the value of controlling parameter, the number of iterations, the error between the solution sequence $\widehat{x}_N^{(k)}$ by Algorithm 5 and the optimal solution x^* of problem (1), and the error between the optimal value $\widehat{v}_N^{(k)}$ by Algorithm 5 and the optimal value v^* of problem (1), respectively.

Example 1 (Hald-Madson [17]). Consider the unconstrained min-max stochastic problem (1), in which ω is uniformly distributed on $\Omega = [0, 1]$ and $f_i : \mathfrak{R}^2 \times \Omega \to \mathfrak{R}(i = 1, \ldots, 4)$ are given by

$$\begin{aligned} f_1(x, \omega) &= 10 \left(2\omega x_2 - 3\omega^2 x_1^2 \right), \\ f_2(x, \omega) &= -f_1(x, \omega), \\ f_3(x, \omega) &= 4\omega^3 - 2\omega x_1, \\ f_4(x, \omega) &= -f_3(x, \omega). \end{aligned} \quad (29)$$

This problem has the optimal solution $x^* = (1, 1)^T$ and the optimal value $v^* = 0$. The numerical results for this example obtained by Algorithm 5 are shown in Table 1.

Example 2 (Beale [18]). Consider the unconstrained min-max stochastic problem (1), in which ω is uniformly distributed on $\Omega = [0, 1]$ and $f_i : \mathfrak{R}^3 \times \Omega \to \mathfrak{R}(i = 1, \ldots, 5)$ are given by

TABLE 2: The numerical results for Example 2.

N	$1/t$	Iter.	$\|x^* - \hat{x}_N^{(k)}\|$	$\|\hat{v}_N^{(k)} - v^*\|$
10^2	5	2	$4.503171e - 001$	$1.050617e - 001$
10^3	5	2	$1.604489e - 001$	$1.028518e - 001$
10^4	5	2	$4.014238e - 002$	$1.801205e - 002$
10^5	5	2	$6.732682e - 003$	$4.511915e - 003$
10^6	5	4	$4.870377e - 003$	$1.690755e - 003$
10^7	5	4	$3.176635e - 004$	$1.530397e - 004$

TABLE 3: The numerical results for Example 3.

N	$1/t$	Iter.	$\|x^* - \hat{x}_N^{(k)}\|$	$\|\hat{v}_N^{(k)} - v^*\|$
10^2	11	4	$8.816640e - 002$	$2.888003e - 001$
10^3	11	4	$2.816097e - 002$	$1.355825e - 001$
10^4	11	4	$2.859081e - 003$	$1.536729e - 003$
10^5	11	3	$4.731965e - 003$	$1.960690e - 003$
10^6	11	4	$4.279962e - 004$	$8.171314e - 004$
10^7	11	4	$2.830422e - 004$	$1.131960e - 004$

$$
\begin{aligned}
f_1(x, \omega) &= 4\omega x_1^2 + 2x_2^2 + x_3^2 + 6\omega^2 x_1 x_2 \\
&\quad + 2x_1 x_3 - 8x_1 - 18\omega^2 x_2 - 8\omega x_3 + 9, \\
f_2(x, \omega) &= f_1(x, \omega) + 10\left(x_1 + 3\omega^2 x_2 + 2x_3 - 6\omega\right), \\
f_3(x, \omega) &= f_1(x, \omega) - 10x_1, \\
f_4(x, \omega) &= f_1(x, \omega) - 30\omega^2 x_2, \\
f_5(x, \omega) &= f_1(x, \omega) - 20\omega x_3.
\end{aligned}
\tag{30}
$$

This problem has the optimal solution $x^* = (4/3, 7/9, 4/9)^T$ and the optimal value $v^* = 1/9$. The numerical results for this example obtained by Algorithm 5 are shown in Table 2.

Example 3 (Rosen Suzuki [18]). Consider the unconstrained min-max stochastic problem (1), in which ω is uniformly distributed on $\Omega = [0, 1]$ and $f_i : \mathcal{R}^4 \times \Omega \to \mathcal{R}(i = 1, \ldots, 4)$ are given by

$$
\begin{aligned}
f_1(x, \omega) &= 3\omega^2 x_1^2 + x_2^2 + 6\omega^2 x_3^2 \\
&\quad + 2\omega x_4^2 - 5x_1 - 20\omega^3 x_2 - 23x_3 + 7x_4, \\
f_2(x, \omega) &= 11x_1^2 + 22\omega x_2^2 + 12x_3^2 \\
&\quad + 44\omega^3 x_4^2 + 5x_1 - 15x_2 - 11x_3 - 3x_4 - 160\omega, \\
f_3(x, \omega) &= 11x_1^2 + 21x_2^2 + 36\omega^2 x_3^2 + 21x_4^2 \\
&\quad - 15x_1 - 10\omega x_2 - 21x_3 - 9\omega^2 x_4 - 100, \\
f_4(x, \omega) &= 11x_1^2 + 11x_2^2 + 12x_3^2 + 3\omega^2 x_4^2 \\
&\quad + 15x_1 - 15x_2 - 21x_3 - 3x_4 - 150\omega^2.
\end{aligned}
\tag{31}
$$

This problem has the optimal solution $x^* = (0, 1, 2, -1)^T$ and the optimal value $v^* = -44$. The numerical results for this example obtained by Algorithm 5 are shown in Table 3.

Example 4 (Wong 1 [17]). Consider the unconstrained min-max stochastic problem (1), in which ω is uniformly distributed on $\Omega = [0, 1]$ and $f_i : \mathcal{R}^7 \times \Omega \to \mathcal{R}(i = 1, \ldots, 5)$ are given by

TABLE 4: The numerical results for Example 4.

N	$1/t$	Iter.	$\|x^* - \hat{x}_N^{(k)}\|$	$\|\hat{v}_N^{(k)} - v^*\|$
10^2	3.63	2	$8.258801e - 002$	$3.835420e + 000$
10^3	3.63	2	$5.495890e - 002$	$2.518867e + 000$
10^4	3.63	2	$3.114568e - 002$	$1.552539e + 000$
10^5	3.63	2	$7.241573e - 003$	$4.854735e - 002$
10^6	3.63	3	$6.411032e - 005$	$6.869434e - 004$
10^7	3.63	3	$4.752318e - 005$	$1.734132e - 004$

$$
\begin{aligned}
f(x, \omega) &= (x_1 - 10)^2 + 5(x_2 - 12)^2 \\
&\quad + 2\omega x_3^4 + 3(x_4 - 11)^2 + 10x_5^6 + 7x_6^2 \\
&\quad + 3\omega^2 x_7^4 - 4x_6 x_7 - 10x_6 - 8x_7. \\
f_1(x, \omega) &= f(x, \omega), \\
f_2(x, \omega) &= f(x, \omega) \\
&\quad - 10\left(-2x_1^2 - 9\omega^2 x_2^4 - x_3 \right.\\
&\qquad\qquad \left. -4x_4^2 - 5x_5 + 120 + 21\omega^2\right), \\
f_3(x, \omega) \\
&= f(x, \omega) \\
&\quad - 10\left(-7x_1 - 6\omega x_2 - 10x_3^2 - 2\omega x_4 + x_5 + 282\right), \\
f_4(x, \omega) &= f(x, \omega) \\
&\quad - 10\left(-23x_1 - 3\omega^2 x_2^2 - 6x_6^2 + 8x_7 + 196\right), \\
f_5(x, \omega) \\
&= f(x, \omega) \\
&\quad - 10\left(-4x_1^2 - 2\omega x_2^2 + 3x_1 x_2 - 6\omega^2 x_3^2 - 5x_6 + 11x_7\right),
\end{aligned}
\tag{32}
$$

This problem has the optimal solution $x^* = (2.3305, 1.9514, -0.47754, 4.3657, -0.62449, 1.0381, 1.5942)^T$ and the optimal value $v^* = 680.6301$. The numerical results for this example obtained by Algorithm 5 are shown in Table 4.

Example 5 (Wong 2 [18]). Consider the unconstrained minmax stochastic problem (1), in which ω is uniformly distributed on $\Omega = [0, 1]$ and $f_i : \mathfrak{R}^{10} \times \Omega \to \mathfrak{R}(i = 1, \ldots, 9)$ are given by

$$
\begin{aligned}
f(x, \omega) &= x_1^2 + 2\omega x_2^2 + x_1 x_2 \\
&\quad - 14x_1 - 48\omega^2 x_2 + (x_3 - 10)^2 + 4(x_4 - 5)^2 \\
&\quad + (x_5 - 3)^2 + 4\omega(x_6 - 1)^2 + 5x_7^2 \\
&\quad + 7(x_8 - 11)^2 + 2(x_9 - 10)^2 \\
&\quad + (x_{10} - 7)^2 + 45,
\end{aligned}
$$

$$
f_1(x, \omega) = f(x, \omega),
$$

$$
\begin{aligned}
f_2(x, \omega) = f(x, \omega) - 10 \Big(&-6\omega(x_1 - 2)^2 \\
&-4(x_2 - 3)^2 - 2x_3^2 + 7x_4 + 120 \Big),
\end{aligned}
$$

$$
\begin{aligned}
f_3&(x, \omega) \\
&= f(x, \omega) \\
&\quad - 10 \left(-15\omega^2 x_1^2 - 8x_2 - (x_3 - 6)^2 + 4\omega x_4 + 40 \right),
\end{aligned}
$$

$$
\begin{aligned}
f_4(x, \omega) = f(x, \omega) \\
&\quad - 10 \left(-2\omega^3(x_1 - 8)^2 \right. \\
&\qquad \left. -2(x_2 - 2)^2 - 3x_5^2 + 2\omega x_6 + 30 \right),
\end{aligned}
$$

$$
\begin{aligned}
f_5(x, \omega) = f(x, \omega) \\
&\quad - 10 \left(-x_1^2 - 2(x_2 - 2)^2 + 2x_1 x_2 - 14x_5 + 6x_6 \right),
\end{aligned}
$$

$$
\begin{aligned}
f_6(x, \omega) = f(x, \omega) \\
&\quad - 10 \left(-4x_1 - 10\omega x_2 + 3x_7 - 9x_8 + 105 \right),
\end{aligned}
$$

$$
\begin{aligned}
f_7&(x, \omega) \\
&= f(x, \omega) - 10 \left(-10x_1 + 8x_2 + 17x_7 - 6\omega^2 x_8 \right),
\end{aligned}
$$

$$
\begin{aligned}
f_8(x, \omega) = f(x, \omega) \\
&\quad - 10 \left(3x_1 - 12\omega x_2 - 12(x_9 - 8)^2 + 7x_{10} \right),
\end{aligned}
$$

$$
\begin{aligned}
f_9&(x, \omega) \\
&= f(x, \omega) - 10 \left(8x_1 - 2x_2 - 5x_9 + 2x_{10} + 24\omega \right).
\end{aligned}
\tag{33}
$$

This problem has the optimal solution

$$
\begin{aligned}
x^* = (&2.171996, 2.363683, \\
&8.773926, 5.095985, 0.990655, 1.430574,
\end{aligned}
$$

TABLE 5: The numerical results for Example 5.

N	$1/t$	Iter.	$\|x^* - \widehat{x}_N^{(k)}\|$	$\|\widehat{v}_N^{(k)} - v^*\|$
10^2	1.26	5	$2.210460e - 001$	$3.206412e + 000$
10^3	1.26	5	$1.040470e - 002$	$1.698143e - 001$
10^4	1.26	6	$3.016663e - 003$	$2.977730e - 002$
10^5	1.26	6	$2.854231e - 003$	$2.331369e - 002$
10^6	1.26	6	$6.241527e - 004$	$5.987215e - 003$
10^7	1.26	6	$1.796381e - 004$	$1.225449e - 003$

$$
\begin{aligned}
&1.321644, 9.828726, \\
&8.280092, 8.375927)^T
\end{aligned}
\tag{34}
$$

and the optimal value $v^* = 24.306209$. The numerical results for this example obtained by Algorithm 5 are shown in Table 5.

From the above numerical results, the following remarks are proposed.

Remark 10. The preliminary numerical results show that Algorithm 5 is feasible and promising.

Remark 11. Compared with the numerical results for the same test example with the different sample size N, the above numerical results show that, with the sample size being chosen larger, the precision of the optimal solution and the optimal value by Algorithm 5 become higher, which coincides with the theoretical analysis in Section 3.

5. Conclusions

This paper investigates a nonlinear Lagrange algorithm for solving stochastic minimax problems based on the sample average approximation method. And the convergence theory of the proposed algorithm is established under some assumptions. Furthermore, the preliminary numerical results are reported to demonstrate the feasibility and effectiveness of the algorithm. The future works on improving the numerical experiments to obtain the solutions with higher precision and performing the numerical experiments for large-scale test examples deserve our further attention. And applying this proposed algorithm to some practical problems is also interesting.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

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

[2] J. Dupačováa, "The minimax approach to stochastic programming and an illustrative application," *Stochastic*, vol. 20, pp. 73–88, 1987.

[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] M. Riis and K. A. Andersen, "Applying the minimax criterion in stochastic recourse programs," Tech. Rep., Department of Operations Research, University of Aarhus, Aarhus, Denmark, 2002.

[5] S. Takriti and S. Ahmed, "Managing short-term electricity contracts under uncertainty: a mini- max approach," Tech. Rep., School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, Ga, USA, 2002.

[6] L. W. Zhang and H. W. Tang, "A maximum entropy algorithm with parameters for solving minimax problem," *Archives Control Science*, vol. 6, pp. 47–59, 1997.

[7] S. X. He and Y. Y. Nie, "A class of nonlinear Lagrangian algorithms for minimax problems," *Journal of Industrial and Management Optimization*, vol. 9, pp. 75–97, 2013.

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

[9] X. Chen, R. B.-J. Wets, and Y. Zhang, "Stochastic variational inequalities: residual minimization smoothing/sample average approximations," *SIAM Journal on Optimization*, vol. 22, pp. 649–673, 2012.

[10] F. Meng and H. Xu, "A regularized sample average approximation method for stochastic mathematical programs with nonsmooth equality constraints," *SIAM Journal on Optimization*, vol. 17, no. 3, pp. 891–919, 2006.

[11] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, MPS-SIAM Series on Optimization, 2009.

[12] 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.

[13] 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.

[14] 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.

[15] J. Zhang, L.-W. 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.

[16] R. Walter, *Principles of Mathematical Analysis*, McGraw-Hill, New York, NY, USA, 1953.

[17] C. Charalambous, "Nonlinear least pth optimization and nonlinear programming," *Mathematical Programming*, vol. 12, no. 1, pp. 195–225, 1977.

[18] G. Di Pillo, L. Grippo, and S. Lucidi, "A smooth method for the finite minimax problem," *Mathematical Programming*, vol. 60, no. 1–3, pp. 187–214, 1993.