

求解具有等式约束的随机优化问题的二阶微分方程方法

吕琪楠, 孙菊贺, 王 莉, 谢红俭

沈阳航空航天大学理学院, 辽宁 沈阳

收稿日期: 2023年9月26日; 录用日期: 2023年10月21日; 发布日期: 2023年10月30日

摘 要

针对具有等式约束的随机优化问题, 本文首先考虑基于SAA方法获取原始问题目标函数的SAA问题, 并利用惩罚函数法将其转化为一般无约束凸优化问题。接下来建立二阶微分方程算法对所得到的问题进行求解。本文还证明了当算法的参数满足一定条件时, 二阶微分方程算法具有良好的收敛性, 其收敛率为 $o\left(\frac{1}{t^2}\right)$, 最后给出具体的数值算例验证算法的可行性。

关键词

随机优化问题, 二阶微分方程算法, 凸优化问题, SAA方法

Second Order Differential Equation Algorithm for Solving Stochastic Optimization Problems with Equality Constraints

Qinan Lv, Jue Sun, Li Wang, Hongjian Xie

School of Science, Shenyang University of Aeronautics and Astronautics, Shenyang Liaoning

Received: Sep. 26th, 2023; accepted: Oct. 21st, 2023; published: Oct. 30th, 2023

Abstract

For the stochastic optimization problem with equality constraint, this paper first considers the SAA problem of obtaining the objective function of the original problem based on the SAA method, and

文章引用: 吕琪楠, 孙菊贺, 王莉, 谢红俭. 求解具有等式约束的随机优化问题的二阶微分方程方法[J]. 应用数学进展, 2023, 12(10): 4529-4537. DOI: 10.12677/aam.2023.1210444

uses the penalty function method to transform it into a general unconstrained convex optimization problem. Next, a second-order differential equation algorithm is established to solve the obtained problem. This paper also proves that when the parameters of the algorithm meet certain conditions, the second-order differential equation algorithm has good convergence, and its convergence rate is $o\left(\frac{1}{t^2}\right)$, and finally a specific numerical example is given to verify the feasibility of the algorithm.

Keywords

Stochastic Optimization Problem, Second-Order Differential Equation System, Convex Optimization Problem, SAA Method

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着经济发展,许多实际问题往往存在许多不确定性因素,例如物理不确定性,经济不确定性,统计不确定性等[1][2]。因此,许多学者在研究优化问题时引入了随机变量,得到随机优化问题。随机优化问题作为确定性优化问题的延伸,在机械、经济、人工智能等领域都有很多应用。

随机优化问题首先被 G. Dantzig 提出,在文献[3]中,他提出了带补偿的两阶段随机优化问题,随着经济学和人类需求的进步,越来越多的学者开始关注随机优化问题。同时,针对各种形式的随机优化问题,许多学者给出了不同的算法[2][3][4][5][6]。

针对 $\min_{x \in C} \{f(x) = E[F(x, \xi)]\}$ 形式的问题,Guo 等人在文献[7]中提出了随机无梯度的 Frank-wolf 方法,主要解释了目标函数中的随机函数可以证明算法在海德尔连续性条件下的收敛性,无需满足利普希茨的连续性。此外,Guo 等人还指出,在算法的收敛分析中,梯度方差是均匀有界的假设是没有必要的。

Curtis 等在文献[8]中提出了一种基于梯度下降算法方法的随机平均梯度(SAG)计算方法,并证明了一定条件下的线性收敛性。Schmidt 等在文献[9]中提出了一种基于一阶 SGD 方法的后续方法。Chen 等人在文献[10]中将不精确的牛顿方法与随机分解方法相结合,提出了一种解决两阶段随机问题的随机牛顿方法。但是对于特定结构之外的问题,计算起来比较麻烦。

然而,解决上述随机优化问题的最常用的方法是文献[11]中的样本均值近似(SAA)和随机近似(SA)。文献[12]研究了 SAA 方法并验证了其收敛性。虽然有很多学者[13][14]致力于改进 SA 方法,但解决某些问题更麻烦,因为它涉及投影问题。

2. 预备知识

对于非光滑函数,本文选择用一串光滑函数来代替非光滑函数,下面给出光滑函数的定义。

定义 2.1 Φ 为 $R \rightarrow R$ 的一个凸函数,称 $\tilde{\Phi}: R \times [0, \infty] \rightarrow R$ 为 Φ 的光滑函数。如果 $\tilde{\Phi}(z, \mu)$ 满足以下条件:

1) 对任意的固定的 $\mu > 0$, $\Phi(\cdot, \mu)$ 在 R 上是连续可微的,并且对任意固定的 $z \in R$, $\tilde{\Phi}(z, \mu)$ 在 $(0, \infty)$ 上是连续可微的。

- 2) $\lim_{h \rightarrow z, \mu \downarrow 0} \tilde{\Phi}(h, \mu) = \Phi(z), \forall z \in R$ 。
 3) $\tilde{\Phi}(z, \mu)$ 是在 R 上的凸函数, 对于任何固定的 $\mu > 0$ 。
 4) 存在一个正的常数 κ , 使得,

$$|\nabla_{\mu} \tilde{\Phi}(z, \mu) \nabla_{\mu} \tilde{\Phi}(z, \mu)| \leq \kappa, \forall \mu \in (0, \infty), z \in R \quad (2.1)$$

5) 对于任意的 $\mu \in (0, \infty)$, 存在一个常数 $L > 0$ $\nabla_z \tilde{\Phi}(\cdot, \mu)$ 在 R 上是 Lipschitz 连续的, 并且有 Lipschitz 常数 $L\mu^{-1}$

6) 对于任意的 $h \in C$, $\nabla_h \tilde{\Phi}(h, \mu)$ 在 $(0, \infty)$ 上是连续的, 根据定义 2.1, 可得

$$|\tilde{\Phi}(z, \mu_2) - \tilde{\Phi}(z, \mu_1)| \leq \kappa |\mu_1 - \mu_2|, \forall z \in R, \mu_1, \mu_2 \in (0, \infty)$$

并且

$$|\tilde{\Phi}(z, \mu) - \Phi(z)| \leq \kappa \mu, \forall z \in R, \mu \in (0, \infty) \quad (2.2)$$

3. 随机优化问题

本文主要研究以下形式的随机优化问题:

$$\begin{aligned} \min \quad & E_p[f(x, \xi)] \\ \text{s.t.} \quad & g(x, \xi) = 0 \\ & x \in C \end{aligned} \quad (3.1)$$

其中 ξ 为定义在概率空间 (Ω, F, P) 上的随机变量, $E_p[f(x, \xi)] = \int_{\Omega} f(x, \xi) dP(\xi)$ 表示随机变量 ξ 的数学期望。问题(3.1)中的随机函数 $f(x, \xi): R^n \times \Omega \rightarrow R$ 关于 x 连续并且 $g(x, \xi): R^n \times \Omega \rightarrow R^m$ 关于 x 连续。

对于以下形式的非光滑凸优化问题:

$$\min_{x \in H} f(x)$$

其中 $f(x)$ 是一个非光滑凸函数, H 为希尔伯特空间。Qu 和 Bianciti 在文献[15]中提出了一种二阶微分方程方法, 形式如下:

$$\ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \nabla_x \tilde{f}(x(t), \mu(t)) = 0$$

对于问题(3.1), 本文利用样本均值近似法(SAA)得到目标函数的 SAA 问题, 则问题(3.1)可转化为

$$\begin{aligned} \min \quad & \frac{1}{N} \sum_{i=1}^N f(x, \xi_i) \\ \text{s.t.} \quad & g(x, \xi) = 0 \\ & x \in C \end{aligned} \quad (3.2)$$

利用罚函数方法将问题(3.2)转换为下面无约束优化形式:

$$\min_{x \in C} \Phi(z) = \frac{1}{N} \sum_{i=1}^N f(x, \xi_i) + \frac{1}{2} \|g(x, \xi)\|^2$$

进而, 采用二阶微分方程方法求解上述优化问题, 其形式如下:

$$\ddot{z}(t) + \frac{\alpha}{t} \dot{z}(t) + \nabla_z \tilde{\Phi}(z(t), \mu(t)) = 0 \quad (3.3)$$

其中参数 $\alpha > 0$, $\tilde{\Phi}(z(t))$ 是 $\Phi(z(t))$ 的光滑函数, $\tilde{f}(z(t))$ 是 $f(z(t))$ 的光滑函数, $\mu: [0, \infty] \rightarrow [0, \infty]$ 是

一个连续可微的递减函数，满足 $\lim_{t \rightarrow \infty} \mu(t) = 0$ 。

对于式(3.3)中的函数 $\mu(\cdot)$ ，本文做出如下假设

$$\int_{t_0}^{\infty} t \mu(t) dt < \infty \tag{3.4}$$

带回定义 1.1，则对于任意固定的 $t_0 > 0$ ，式(3.4)可以推出

$$\int_{t_0}^{\infty} \frac{1}{t} \mu(t) dt < \infty$$

4. 收敛定理

定理 1: 假设 $z: [t_0, \infty) \rightarrow R^n$ 是式(3.3)的轨迹，并假设 $\arg \min \Phi(z)$ 非空。

1) 假设 $\alpha \geq 3$ ，那么

$$\Phi(z(t)) - \min \Phi = O\left(\frac{1}{t^2}\right).$$

2) 假设 $\alpha > 3$ ，那么

$$\begin{aligned} \int_{t_0}^{\infty} t (\Phi(z(t)) - \min \Phi) dt &< \infty \\ \int_{t_0}^{\infty} t \|\dot{z}(t)\|^2 dt &< \infty \\ \Phi(z(t)) - \min \Phi &= o\left(\frac{1}{t^2}\right) \end{aligned} \tag{4.1}$$

证明: 1) 令 $z^* \in \arg \min \Phi$ ，并考虑如下函数

$$\beta(t) = t^2 (\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) + \frac{1}{2} \|(\alpha - 1)(z(t) - z^*) + t\dot{z}(t)\|^2$$

根据式(2.2)，可以得出

$$\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t) \geq \Phi(z(t)) - \Phi(z^*) \geq 0 \tag{4.2}$$

因此推导出 $\beta(t) \geq 0$ 。由(3.3)，得

$$\begin{aligned} \frac{d}{dt} \beta(t) &= 2t (\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) \\ &\quad + t^2 (\langle \nabla_z \tilde{\Phi}(z(t), \mu(t)), \dot{z}(t) \rangle + \nabla_{\mu} \tilde{\Phi}(z(t), \mu(t)) \dot{\mu}(t) - \nabla_{\mu} \tilde{\Phi}(z^*, \mu(t)) \dot{\mu}(t) + 2\kappa\dot{\mu}(t)) \\ &\quad + \langle (\alpha - 1)(z(t) - z^*) + t\dot{z}(t), \alpha\dot{z}(t) + t\ddot{z}(t) \rangle \\ &= 2t (\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) - (\alpha - 1)t \langle z(t) - z^*, \nabla_z \tilde{\Phi}(z(t), \mu(t)) \rangle \\ &\quad + t^2 (\nabla_{\mu} \tilde{\Phi}(z(t), \mu(t)) \dot{\mu}(t) - \nabla_{\mu} \tilde{\Phi}(z^*, \mu(t)) \dot{\mu}(t) + 2\kappa\dot{\mu}(t)) \end{aligned} \tag{4.3}$$

通过 $\dot{\mu}(t) \leq 0$ 以及式(2.2)推导出

$$\nabla_{\mu} \tilde{\Phi}(z(t), \mu(t)) \dot{\mu}(t) - \nabla_{\mu} \tilde{\Phi}(z^*, \mu(t)) \dot{\mu}(t) \leq -2\kappa\dot{\mu}(t) \tag{4.4}$$

因为对于任意固定的 μ ， $\tilde{\Phi}(z, \mu)$ 相对于 z 都是凸函数，故有

$$\tilde{\Phi}(z^*, \mu(t)) - \tilde{\Phi}(z(t), \mu(t)) \geq \langle \nabla_z \tilde{\Phi}(z(t), \mu(t)), z^* - z(t) \rangle \tag{4.5}$$

若 $\alpha \geq 3$ ，则将式(4.4)和式(4.3)代入到式(4.2)中，则可以得到

$$\begin{aligned} \frac{d}{dt}\beta(t) &\leq -(\alpha-3)t(\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) + 2(\alpha-1)\kappa t\mu(t) \\ &\leq 2(\alpha-1)\kappa t\mu(t) \end{aligned} \quad (4.6)$$

其中最后一个不等式可使用式(4.1)。

通过式(3.4), 有 $\left[\frac{d}{dt}\beta(t)\right]_+$ 属于 $L_1(t_0, \infty)$, 令 $\rho(t) := \beta(t) - \int_{t_0}^t \left[\frac{d}{dt}\beta(t)\right]_+ dt$, 则由 $\beta(t)$ 和 $\left[\frac{d}{dt}\beta(t)\right]_+ \in L_1(t_0, \infty)$ 的有界性可以发现 $\rho(\cdot)$ 也是有界的。又因为 $\frac{d}{dt}\rho(t) = \frac{d}{dt}\beta(t) - \left[\frac{d}{dt}\beta(t)\right]_+ \leq 0$, 故 $\lim_{t \rightarrow \infty} \rho(t)$ 是存在的。因此可得

$$\lim_{t \rightarrow \infty} \beta(t) = \lim_{t \rightarrow \infty} w(t) + \int_{t_0}^{\infty} [\beta(s)]_+ ds < \infty \quad (4.7)$$

由此可知 $\beta(\cdot)$ 在 $[t_0, \infty)$ 上是有界的, 且 $t^2(\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t))$ 在 $[t_0, \infty)$ 上是有界的, 故存在一个常数 C , 使得

$$t^2(\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) \leq C, \forall t \geq t_0.$$

因此, 根据式(4.1)有

$$\Phi(z(t)) - \Phi(z^*) \leq \frac{C}{t^2}$$

即

$$\Phi(z(t)) - \Phi(z^*) = O\left(\frac{1}{t^2}\right)$$

2) 假设 $\alpha > 3$, 对式(4.6)做 t_0 到 t 的积分, 得到

$$\begin{aligned} \int_{t_0}^t s(\tilde{\Phi}(z(s), \mu(s)) - \tilde{\Phi}(z^*, \mu(s)) + 2\kappa\mu(s)) ds &\leq \frac{1}{\alpha-3}(\beta(t_0) - \beta(t)) + \frac{2(\alpha-1)\kappa}{\alpha-3} \int_{t_0}^t s\mu(s) ds \\ &\leq \frac{1}{\alpha-3}\beta(t_0) + \frac{2(\alpha-1)\kappa}{\alpha-3} \int_{t_0}^t s\mu(s) ds \end{aligned}$$

在式(3.4)下, 有

$$\int_{t_0}^{\infty} s(\tilde{\Phi}(z(s), \mu(s)) - \tilde{\Phi}(z^*, \mu(s)) + 2\kappa\mu(s)) ds < \infty \quad (4.8)$$

通过式(4.2), 可以得到

$$\int_{t_0}^{\infty} t(\Phi(z(t)) - \min \Phi) dt < \infty$$

接下来证明 $\Phi(z(t)) - \min \Phi = o\left(\frac{1}{t^2}\right)$

取算法(3.3)的标量乘以 $t^2\dot{x}(t)$, 则有

$$\begin{aligned} &\frac{t^2}{2} \frac{d}{dt} \|\dot{z}(t)\|^2 + \alpha t \|\dot{z}(t)\|^2 + t^2 \frac{d}{dt} (\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) \\ &= t^2 (\nabla_{\mu} \tilde{\Phi}(z(t), \mu(t)) \dot{\mu}(t) - \nabla_{\mu} \tilde{\Phi}(z^*, \mu(t)) \dot{\mu}(t) + 2\kappa \dot{\mu}(t)) \end{aligned}$$

对上述方程做 t_0 到 t 的积分, 则有

$$\begin{aligned} & \frac{t^2}{2} \|\dot{z}(t)\|^2 - \frac{t_0^2}{2} \|\dot{z}(t_0)\|^2 + (\alpha - 1) \int_{t_0}^t s \|\dot{z}(s)\|^2 ds + t^2 (\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) \\ & - t_0^2 (\tilde{\Phi}(z(t_0), \mu(t_0)) - \tilde{\Phi}(z^*, \mu(t_0)) + 2\kappa\mu(t_0)) - 2 \int_{t_0}^t s (\tilde{\Phi}(z(s), \mu(s)) - \tilde{\Phi}(z^*, \mu(s)) + 2\kappa\mu(s)) ds \\ & = \int_{t_0}^t s^2 \nabla_{\mu} \tilde{\Phi}(z(s), \mu(s)) \dot{\mu}(s) - \nabla_{\mu} \tilde{\Phi}(z^*, \mu(s)) \end{aligned}$$

结合上述关系式以及式(4.2)和(4.4)，由于 $\frac{t^2}{2} \|\dot{z}(t)\|^2 \geq 0$ ，则有

$$(\alpha - 1) \int_{t_0}^t s \|\dot{z}(s)\|^2 ds \leq \tilde{C} + 2 \int_{t_0}^t s (\tilde{\Phi}(z(s), \mu(s)) - \tilde{\Phi}(z^*, \mu(s)) + 2\kappa\mu(s)) ds$$

其中

$$\tilde{C} = \frac{t_0^2}{2} \|\dot{z}(t_0)\|^2 + t_0^2 (\tilde{\Phi}(z(t_0), \mu(t_0)) - \tilde{\Phi}(z^*, \mu(t_0)) + 2\kappa\mu(t_0))$$

由式(4.8)和 $\alpha > 3$ 可得式(4.1)。

接下来考虑

$$E(t) := \frac{1}{2} \|\dot{z}(t)\|^2 + \tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)$$

由于他在 $[t_0, \infty)$ 上是非负的，根据式(4.1)和(4.8)，

$$\int_{t_0}^{\infty} tE(t) dt < \infty \tag{4.9}$$

对 $t^2E(t)$ 进行微分，则有

$$\begin{aligned} \frac{d}{dt}(t^2E(t)) &= 2tE(t) + t^2 \frac{dE(t)}{dt} \\ &= 2t \left(\frac{1}{2} \|\dot{z}(t)\|^2 + \tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t) \right) \\ &\quad + t^2 \left(\langle \dot{z}(t), \ddot{z}(t) \rangle + \langle \nabla_z \tilde{\Phi}(z(t), \mu(t)), \dot{z}(t) \rangle \right) \\ &\quad + \nabla_{\mu} \tilde{\Phi}(z(t), \mu(t)) \dot{\mu}(t) - \left(\nabla_{\mu} \tilde{\Phi}(z^*, \mu(t)) \dot{\mu}(t) + 2\kappa\dot{\mu}(t) \right) \\ &\quad + t^2 \left(\langle \dot{z}(t), \ddot{z}(t) \rangle + \langle \nabla_z \tilde{\Phi}(z(t), \mu(t)), \dot{z}(t) \rangle \right) \\ &\quad + \nabla_{\mu} \tilde{\Phi}(z(t), \mu(t)) \dot{\mu}(t) - \left(\nabla_{\mu} \tilde{\Phi}(z^*, \mu(t)) \dot{\mu}(t) + 2\kappa\dot{\mu}(t) \right) \end{aligned} \tag{4.10}$$

在(3.3)和式(4.4)的基础上可推出

$$\begin{aligned} \frac{d}{dt}(t^2E(t)) &\leq 2t (\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) + t(1 - \alpha) \|\dot{z}(t)\| \\ &\leq 2t (\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t)) \end{aligned}$$

结合式(4.8)和(4.10)发现 $\left[\frac{d}{dt}(t^2E(t)) \right]_+$ 属于 $L_1(t_0, \infty)$ ，因此类比式(4.7)的推导过程可知 $\lim_{t \rightarrow \infty} t^2E(t)$ 存

在，通过式(4.9)可知 $\int_{t_0}^{\infty} tE(t) dt = \int_{t_0}^{\infty} \frac{1}{t} (t^2E(t)) dt$ ，又由于 $\int_{t_0}^{\infty} \frac{1}{t} dt = \infty$ 和 $\lim_{t \rightarrow \infty} t^2E(t)$ 的存在性可知

$$\lim_{t \rightarrow \infty} t^2E(t) = 0$$

在 E 的定义下有

$$0 \leq \lim_{t \rightarrow \infty} t^2 \left(\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t) \right) \leq \lim_{t \rightarrow \infty} t^2 E(t) = 0$$

并且

$$\begin{aligned} 0 &\leq \lim_{t \rightarrow \infty} t^2 \left(\Phi(z(t)) - \Phi(z^*) \right) \\ &\leq \lim_{t \rightarrow \infty} t^2 \left(\tilde{\Phi}(z(t), \mu(t)) - \tilde{\Phi}(z^*, \mu(t)) + 2\kappa\mu(t) \right) = 0 \end{aligned}$$

即

$$\Phi(z(t)) - \min \Phi = o\left(\frac{1}{t^2}\right)$$

5. 数值模拟

在数值模拟部分, 本文将二阶微分方程算法用于求解以下两个问题, 它们的目标函数是凸优化函数, 其中问题 1 是无约束随机优化问题。结果表明, 二阶微分方程算法计算的结果具有较好的收敛性, 计算程序用 MATLAB 编写。

例 5.1

$$\begin{aligned} \min f(x) &:= E_{\theta} [f(x, \theta)] \\ &:= E_{\theta} \left[\frac{1}{2} x^T (A + \text{Adiag}(\theta)) x + b^T x \right] \end{aligned}$$

其中对角矩阵 $A \in R^{n \times n}$, 向量 $b \in R^n$, 随机向量 $\theta \in R^n$, $\text{diag}(\theta)$ 是由 θ 生成的对角矩阵。 θ 是在四维向量空间 $\theta = [-\theta_0, \theta_0]^T$ 中随机均匀选取的, 并且 $\theta_0 < 1$ 。

为了研究的简便性, 本文在 $[0, 1]^n$ 中随机选取向量 b , 并在离散集 $\{1, 10^{-1}, \dots, 10^{-\xi}\}$ 中随机选取元素作为矩阵 A 的对角元素, 矩阵 A 的变化决定了该问题的最优解。将本题参数设置为 $n = 4, \theta_0 = 0.5, \xi = 3, \alpha = 20$, 其最优解为 $x^* = A^{-1}b$, 接下来用本文的二阶微分方程算法进行求解。所求解与最小值点之间距离的变化图像如图 1 所示:

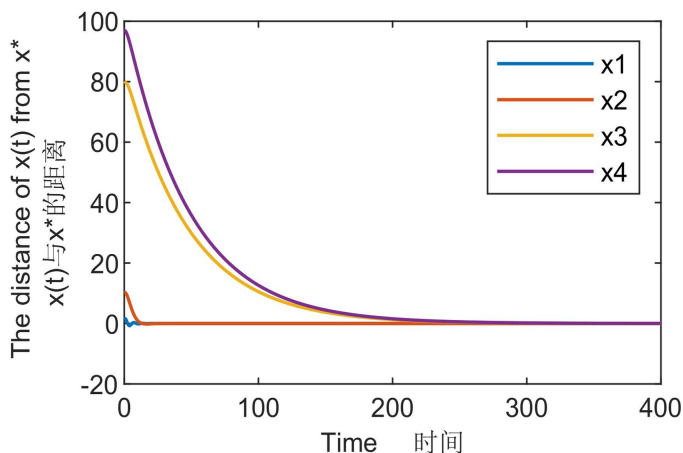


Figure 1. Image of the distance between the solved and the minimum point over time

图 1. 所求解与最小值点之间的距离随时间变化图像

例 5.2

$$\begin{aligned} \min f(x) &= E_{\theta} [f(x, \theta)] \\ &= E_{\theta} \left[\sum_{i=1}^{n/4} (1 + \theta_i) \left((x_{4i-1} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4 \right) \right] \\ \text{s.t. } g(x, \theta) &= -\theta x = 0 \end{aligned}$$

其中 $x \in C, \theta \in R^{n/4}$, 随机变量 θ 在 $[-\theta_0, \theta_0]^{n/4}$ 中均匀随机选择。

首先将目标函数转化为 SAA 问题, 接下来利用罚函数法将其转化为无约束问题, 即

$$\min_{x \in C} \Phi(z) = \frac{1}{N} \sum_{i=1}^N f(x, \theta_i) + \frac{1}{2} \|g(x, \theta)\|^2$$

接下来利用二阶微分方程算法求解, 发现例 5.2 有唯一最小值的点 $x^* = (0, 0, 0, 0)$, 并且最小值为 0。求解出的点与最小值点之间距离的变化图像如图 2 所示:

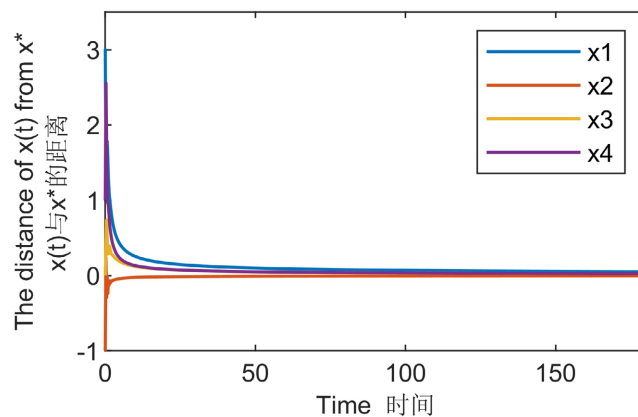


Figure 2. Image of the distance between the solved and the minimum point over time

图 2. 所求解与最小值点之间的距离随时间变化图像

6. 结论

目前常用 SAA 方法求解有约束随机优化问题, 但其计算较为复杂。通过上述数值模拟可以发现, 本文提出的二阶微分方程方法求解例 5.1 和例 5.2 的速度较快, 精度较小, 并且所求的结果能够较为平滑的收敛到最优解。因此, 二阶微分方程方法在求解等式约束随机优化问题时是收敛的。

但本文只验证了二阶微分方程算法可以求解等式约束下的随机优化问题, 对于求解不等式约束的随机优化问题和随机约束的随机优化问题, 二阶微分方程方法是否也具有有良好的收敛性, 值得下一步的探讨和研究。

参考文献

- [1] Ferguson, A.R. and Dantzig, G.B. (1956) The Allocation of Aircraft to Routes—An Example of Linear Programming under Uncertain Demand. *Management Science*, **3**, 45-73. <https://doi.org/10.1287/mnsc.3.1.45>
- [2] Dantzig, G.B. (2004) Linear Programming under Uncertainty. *Management Science*, **50**, 1764-1769. <https://doi.org/10.1287/mnsc.1040.0261>
- [3] Kiefer, J. and Wolfowitz, J. (1952) Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, **23**, 462-466. <https://doi.org/10.1214/aoms/1177729392>

-
- [4] Spall, J.C. (1992) Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation. *IEEE Transactions on Automatic Control*, **37**, 332-341. <https://doi.org/10.1109/9.119632>
- [5] Chen, R., Menickelly, M. and Scheinberg, K. (2018) Stochastic Optimization Using a Trust-Region Method and Random Models. *Mathematical Programming*, **169**, 447-487. <https://doi.org/10.1007/s10107-017-1141-8>
- [6] Billups, S.C., Graf, P. and Larson, J. (2013) Derivative-Free Optimization of Expensive Functions with Computational Error Using Weighted Regression. *SIAM Journal on Optimization*, **23**, 27-53. <https://doi.org/10.1137/100814688>
- [7] Guo, J.H., Liu, H.L. and Xiao, X.T. (2020) A Unified Analysis of Stochastic Gradient-Free Frank-Wolfe Methods. *International Transactions in Operational Research*, **29**, 63-86. <https://doi.org/10.1111/itor.12889>
- [8] Curtis, F.E., Scheinberg, K. and Shi, R. (2017) A Stochastic Trust Region Algorithm. *Industrial and Systems Engineering*, 1-16.
- [9] Schmidt, M., Roux, N.L. and Bach, F. (2017) Minimizing Finite Sums with the Stochastic Average Gradient. *Mathematical Programming*, **162**, 83-112. <https://doi.org/10.1007/s10107-016-1030-6>
- [10] Nie, Y.Y. (2013) Convergence Theory of a SAA Method for Min-Max Stochastic Optimization Problems. *Applied Mechanics and Materials*, **303-306**, 1319-1322. <https://doi.org/10.4028/www.scientific.net/AMM.303-306.1319>
- [11] Ghadimi, S. and Lan, G. (2016) Accelerated Gradient Methods for Nonconvex Nonlinear and Stochastic Programming. *Mathematical Programming*, **156**, 59-99. <https://doi.org/10.1007/s10107-015-0871-8>
- [12] Chen, X.J., Qi, L. and Womersley, R.S. (1995) Newton's Method for Quadratic Stochastic Programs with Recourse. *Journal of Computational and Applied Mathematics*, **60**, 29-46. [https://doi.org/10.1016/0377-0427\(94\)00082-C](https://doi.org/10.1016/0377-0427(94)00082-C)
- [13] Nemirovski, A., Juditsky, A., Lan, G., *et al.* (2009) Robust Stochastic Approximation Approach to Stochastic Programming. *SIAM Journal on Optimization*, **19**, 1574-1609. <https://doi.org/10.1137/070704277>
- [14] Qu, X. and Bian, W. (2022) Fast Inertial Dynamic Algorithm with Smoothing Method for Nonsmooth Convex Optimization. *Computational Optimization and Applications*, **3**, 287-317. <https://doi.org/10.1007/s10589-022-00388-6>
- [15] Polyak, B.T. and Juditsky, A.B. (1992) Acceleration of Stochastic Approximation by Averaging. *SIAM Journal on Control and Optimization*, **30**, 838-855. <https://doi.org/10.1137/0330046>