Journal of Computational Mathematics
Vol.43, No.2, 2025, 315–344.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2208-m2022-0035

A STOCHASTIC AUGMENTED LAGRANGIAN METHOD FOR STOCHASTIC CONVEX PROGRAMMING*

Jiani Wang
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
Email: wjiani@lsec.cc.ac.cn
Liwei Zhang[1)]
School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China
Email: lwzhang@dlut.edu.cn

Abstract

In this paper, we analyze the convergence properties of a stochastic augmented Lagrangian method for solving stochastic convex programming problems with inequality constraints. Approximation models for stochastic convex programming problems are constructed from stochastic observations of real objective and constraint functions. Based on relations between solutions of the primal problem and solutions of the dual problem, it is proved that the convergence of the algorithm from the perspective of the dual problem. Without assumptions on how these random models are generated, when estimates are merely sufficiently accurate to the real objective and constraint functions with high enough, but fixed, probability, the method converges globally to the optimal solution almost surely. In addition, sufficiently accurate random models are given under different noise assumptions. We also report numerical results that show the good performance of the algorithm for different convex programming problems with several random models.

Mathematics subject classification: 49N15, 90C15, 90C25.
Key words: Stochastic convex optimization, Stochastic approximation, Augmented Lagrangian method, Duality theory.

1. Introduction

In this paper, we consider the following stochastic convex optimization problem:

$$\begin{aligned} \min_{x\in X_0} \ & f(x) = \mathbb{E}[F(x,\xi)] \\ \text{s.t.} \ \ & g_i(x) = \mathbb{E}[G_i(x,\xi)] \le 0, \quad i=1,\ldots,p. \end{aligned} \tag{1.1}$$

Here $X_0 \subset \mathbb{R}^n$ is a nonempty bounded closed convex set, $\xi : \Omega \to \Xi$ is a random vector defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $F : O\times\Xi \to \mathbb{R}, G_i : O\times\Xi \to \mathbb{R}, i{=}1,\ldots,p$, where $O{\supset}X_0$ is an open convex set and Ξ is a measurable space. Without loss of generality, we assume that expectations $\mathbb{E}[F(x,\xi)]$ and $\mathbb{E}[G_i(x,\xi)]$ are well defined and finite valued for every $x \in O$ and the expected value function $f(\cdot)$ and $g_i(\cdot)$ are continuous and convex on O. Any algorithm for solving problem (1.1) has to be faced with the difficulty that the full evaluations of expectations $\mathbb{E}[F(x,\xi)]$ and $\mathbb{E}[G_i(x,\xi)]$ are either impossible or expensive in practice. There are two types of methods to resolve this problem: the sample average approximation (SAA) method and the

* Received February 9, 2022 / Revised version received June 6, 2022 / Accepted August 31, 2022 / Published online January 19, 2024 /
1) Corresponding author

stochastic approximation (SA). The SAA method usually solves the random approximation model through sample averaging estimators of random variables. Let $\xi_1, \cdots, \xi_N$ be an i.i.d. sample of realizations of random vector ξ of size N and the average sample approximation model is defined as

$$\begin{aligned} &\min_{x\in X_0} \ \frac{1}{N}\sum_{m=1}^{N} F(x,\xi_m) \\ &\text{s.t.} \ \ \frac{1}{N}\sum_{m=1}^{N} G_i(x,\xi_m) \leq 0, \quad i=1,\ldots,p. \end{aligned} \tag{1.2}$$

Usually, the convergence of SAA depends on the special choice of parameters and the expensive iteration cost, like [18, 26, 28]. However, so far no study has applied sample averaging methods to the case of biased noise for stochastic convex optimization problem (1.1).

On the other hand, the stochastic approximation, in most studies (for example, [17, 31]), is to generate stochastic oracles $\mathbb{F}^{t_k} : \mathbb{R}^n \times \Xi \to \mathbb{R}$ and $\mathbb{G}_i^{s_k} : \mathbb{R}^n \times \Xi \to \mathbb{R}$ of the stochastic function values of f and g_i. More specifically, the random approximation model of (1.1) is defined as

$$\begin{aligned} &\min_{x\in X_0} \ \mathbb{F}^{t_k}(x) \\ &\text{s.t.} \ \ \mathbb{G}_i^{s_k}(x) \leq 0, \quad i=1,\ldots,p, \end{aligned} \tag{1.3}$$

where $\mathbb{F}^{t_k}$ and $\mathbb{G}_i^{s_k}$ are function which are constructed by one or mini-batches of random samples t_k, s_k of stochastic function [5, 29]. For each $k, \mathbb{F}^{t_k}$ and $\mathbb{G}_i^{s_k}$ are continuous on $x \in O$. Obviously, the iterate $x_{k+1} = x_{k+1}(\xi_{[k]})$ can be seen as a function of the history $\xi_{[k]} := (\xi_1, \cdots, \xi_k)$ of the generated random process. The above two random models are considered as the noisy computable version of the real optimization problem (1.1) and the convergence of both methods relies on zero-mean noise with bounded variance (or even with decreasing variance), so estimators in these random models need to be carefully chosen [1, 30]. To the best of our knowledge, no study so far has mentioned the convergence of the stochastic convex programming with inequality constraints under the above two methods, for the case of biased noise. Regardless of the random approximation model (1.2) or (1.3), we propose a stochastic augmented Lagrange method and prove that when the random models are merely sufficiently close to the real optimization problems with high enough, but fixed, probability, the sequence generated by the stochastic algorithm converges to the optimal solution almost surely. In this paper, we consider a general random approximation model of (1.1) as follows:

$$\begin{aligned} &\min_{x\in X_0} \ f^k(x) \\ &\text{s.t.} \ \ G^k(x) \leq 0. \end{aligned} \tag{1.4}$$

For each k, f^k and $G^k =: (g_1^k, \cdots, g_p^k)$ are stochastic approximations of f and g_i and continuous on $x \in O$. The augmented Lagrangian function of problem (1.4) is defined by

$$\mathcal{L}_r^k(x,\lambda) = f^k(x) + \frac{1}{2r}\Big[\big\|\Pi_{\mathbb{R}_+^p}\big(\lambda + rG^k(x)\big)\big\|^2 - \|\lambda\|^2\Big], \quad \forall\,(x,\lambda) \in \mathbb{R}^n \times \mathbb{R}^p, \tag{1.5}$$

where $\Pi_{\mathbb{R}_+^p}(y)$ represents the projection of y onto $\mathbb{R}_+^p$ for any $y \in \mathbb{R}^p$. In the following we denote $[y]_+ := \Pi_{\mathbb{R}_+^p}(y)$. The stochastic augmented Lagrangian method for solving (1.1) with the help of the random model (1.4) can be described as Algorithm 1.1.

The augmented Lagrangian method for solving the optimization problem with constraints can be traced back to the pioneering paper by Rockafellar [23]. Since the augmented Lagrangian