Accepted Manuscript

An improved averaged two-replication procedure with Latin hypercube sampling

Güzin Bayraksan

PII: S0167-6377(17)30690-9
DOI: https://doi.org/10.1016/j.orl.2017.12.005
Reference: OPERES 6319

To appear in: *Operations Research Letters*

Received date : 25 September 2016
Revised date : 7 August 2017
Accepted date : 12 December 2017

An Improved Averaged Two-Replication Procedure with Latin Hypercube Sampling

Güzin Bayraksan

Department of Integrated Systems Engineering, The Ohio State University, Columbus, Ohio 43210 bayraksan.1@osu.edu

Abstract

The averaged two-replication procedure assesses the quality of a candidate solution to a stochastic program by forming point and confidence interval estimators on its optimality gap. We present an improved averaged two-replication procedure that uses Latin hypercube sampling to form confidence intervals of optimality gap. This new procedure produces tighter and less variable interval widths by reducing the sampling error by $\sqrt{2}$. Despite having tighter intervals, it improves an earlier procedure's asymptotic coverage probability bound from $(1-\alpha)^2$ to $(1-\alpha)$.

Keywords: Stochastic optimization, Solution validation, Variance reduction, Latin hypercube sampling, Monte Carlo simulation

1. Introduction

Determining if a candidate solution is optimal or near-optimal is fundamental to optimization. Theoretical results on optimality conditions (e.g., Karush-Kuhn-Tucker conditions) and relaxation bounds on optimal values (e.g., via integrality or semidefinite relaxations) are indispensable tools for optimization theory and algorithms. Applying these tools to stochastic programs, however, is not always immediate. This is because obtaining *exact* function and gradient values is not possible for many stochastic programs.

Monte Carlo sampling provides an alternative technique to evaluate solution quality for stochastic programs. Unlike deterministic optimization, these sampling-based methods are statistical in nature. They provide point and confidence interval estimators of optimality gaps, statistical tests of optimality conditions and solution stability, along with (typically asymptotic) probabilistic guarantees [1, 2, 3]. This paper is concerned with sampling-based optimality gap estimators.

A widely used method to estimate optimality gaps in stochastic programming is the so-called Multiple Replications Procedure (MRP). MRP uses independent replications (or "batches') of observations to estimate optimality gaps. It was first fully analyzed by [4], and a similar bound on the optimal value was independently proposed by [5]. A main advantage of MRP is that it can be used for a wide variety of problems. It does not require convexity of functions and solution sets, continuity of the objective function, and so forth. However, a criticism of MRP is that it solves of a number of (e.g., 25–30) Sample Average Approximation (SAA) problems to evaluate the quality of a single candidate solution.

An alternative approach that solves only a single SAA problem—called the Single Replication Procedure (SRP)—was proposed by [6]. SRP has the same asymptotic probabilistic guarantee as MRP. That is, the confidence intervals formed by SRP have asymptotically at least the same desired coverage probability (e.g., 90%) as their MRP counterparts. However, SRP can underestimate the optimality gaps at small sample sizes for some problems; see, e.g., the example in Section 3.3.

To improve the small-sample behavior of SRP, the Averaged Two-Replication Procedure (A2RP) averages two independent SRP estimators [6]. A2RP provides an attractive alternative: It considerably improves upon the small-sample behavior of SRP but requires the solution of only two SAA problems. Therefore, it has satisfactory probabilistic behavior with low computational burden. A2RP has been used to evaluate solutions to a variety of problems including electric vehicle integration [7], harvesting and forestry management [8], and stochastic scheduling [9, 10].

Another method to improve the performance of Monte Carlo sampling-based estimators is to employ variance and bias reduction techniques. In the context of sampling-based estimators of optimal value and optimality gaps, randomized Quasi-Monte Carlo (QMC) [11, 12, 13], Latin Hypercube Sampling (LHS) [14, 12, 15, 13, 16], Antithetic Variates (AV) [15, 12, 16], and different batching structures [17, 18] have been investigated. AV has been found to be less effective than LHS for most problems [15, 16]. Randomized QMC can also be effective like LHS [13]. However, generating QMC observations can be computationally burdensome when the dimension of the random vector is large. Dimension-reduction strategies for QMC that aim to find the 'important' components of a random vector can themselves be computationally burdensome [11]. Overall, LHS presents a good choice for both variance and bias reduction on a variety of problems with minimal computational effort.

This paper studies the combination of two promising methods: LHS and A2RP. [16] first studied the application of LHS to A2RP and found the combination to yield a number of benefits