

On Solution Quality in Stochastic Programming

Güzin Bayraksan · David P. Morton

guzinb@mail.utexas.edu · morton@mail.utexas.edu

Graduate Program in Operations Research

The University of Texas at Austin, Austin, TX 78712 USA

January 2005

Abstract

Determining whether a solution is of high quality (optimal or near optimal) is a fundamental question in optimization theory and algorithms. We develop a Monte Carlo sampling-based procedure for assessing solution quality in stochastic programs. Quality is defined via the optimality gap and our procedure's output is a confidence interval on this gap. We present a result that justifies a single-replication procedure that only requires solving one optimization problem. This is in contrast to an existing multiple-replications procedure. Even though the single replication procedure is computationally significantly less demanding, the resulting confidence interval might have low coverage probability for small sample sizes for some problems. We point to variants of this procedure that require two replications instead of one, use ϵ -optimal solutions, and perform better empirically.

We consider a stochastic optimization problem of the form

$$z^* = \min_{x \in X} E f(x, \tilde{\xi}), \quad (\text{SP})$$

where f is a real-valued function that determines the cost of operating with decision x under a realization of the random vector $\tilde{\xi}$, whose distribution is assumed known. We make the following assumptions:

(A1) $f(\cdot, \tilde{\xi})$ is continuous on X , w.p.1,

(A2) $E \sup_{x \in X} f^2(x, \tilde{\xi}) < \infty$,

(A3) $X \neq \emptyset$ and is compact.

(SP) is often computationally intractable and an intuitive approach is to resort to sampling and approximate the problem with

$$z_n^* = \min_{x \in X} \frac{1}{n} \sum_{i=1}^n f(x, \tilde{\xi}^i). \quad (\text{SP}_n)$$

We assume $\tilde{\xi}^1, \tilde{\xi}^2, \dots, \tilde{\xi}^n$ are independent and identically distributed (i.i.d.) as $\tilde{\xi}$. Let x^* denote an optimal solution to (SP) with optimal cost z^* . Similarly, let x_n^* and z_n^* denote an optimal solution and the optimal cost of (SP_n). Given a candidate solution $\hat{x} \in X$, we define its quality by its optimality gap, $\mu_{\hat{x}} = E f(\hat{x}, \tilde{\xi}) - z^*$. There are two problems associated with computing this quantity. First, z^* is not known and a lower bound (since we are dealing with a minimization problem) on z^* needs to be computed. A second difficulty is that for a given \hat{x} , it is not always possible to compute $E f(\hat{x}, \tilde{\xi})$ exactly. So, we will estimate $\mu_{\hat{x}}$ by

$$G_n(\hat{x}) = \frac{1}{n} \sum_{i=1}^n f(\hat{x}, \tilde{\xi}^i) - \min_{x \in X} \frac{1}{n} \sum_{i=1}^n f(x, \tilde{\xi}^i). \quad (1)$$

The second term on the right-hand side of (1) is a lower bound in expectation on z^* and was independently introduced by Mak et al. [4] and Norkin et. al. [5]. The former used it in a multiple-replications procedure for assessing solution quality while the latter used it in a branch-and-bound methodology for global stochastic optimization. Other algorithmic work that uses Monte Carlo simulation-based bounds and multiple replications includes [1, 3].

Here, we show how a single observation of (1) can be used to make a valid statistical inference on the quality of a candidate solution. This is in contrast to [4], where multiple replications (e.g., 20 or 30) are used. For a feasible solution, $x \in X$, let $\bar{f}_n(x) = \frac{1}{n} \sum_{i=1}^n f(x, \tilde{\xi}^i)$, $\sigma_{\hat{x}}^2(x) = \text{var}[f(\hat{x}, \tilde{\xi}) - f(x, \tilde{\xi})]$ and $s_n^2(x) = \frac{1}{n-1} \sum_{i=1}^n [(f(\hat{x}, \tilde{\xi}^i) - f(x, \tilde{\xi}^i)) - (\bar{f}_n(\hat{x}) - \bar{f}_n(x))]^2$. Note that $G_n(\hat{x})$ given in equation (1) can be written as $\bar{f}_n(\hat{x}) - z_n^*$, with the understanding that the same n observations $\tilde{\xi}^1, \tilde{\xi}^2, \dots, \tilde{\xi}^n$ are used in $\bar{f}_n(\hat{x})$ and z_n^* . Below we state the single replication procedure (SRP).

SRP:

Input: Desired value of $0 < \alpha < 1$, sample size n and a candidate solution $\hat{x} \in X$.

Output: $(1 - \alpha)$ -level confidence interval on $\mu_{\hat{x}}$.

1. Sample i.i.d. observations $\tilde{\xi}^1, \tilde{\xi}^2, \dots, \tilde{\xi}^n$ from the distribution of $\tilde{\xi}$.
2. Solve (SP_n) to obtain x_n^* .
3. Calculate $G_n(\hat{x})$ as given in (1) and

$$s_n^2(x_n^*) = \frac{1}{n-1} \sum_{i=1}^n \left[(f(\hat{x}, \tilde{\xi}^i) - f(x_n^*, \tilde{\xi}^i)) - (\bar{f}_n(\hat{x}) - \bar{f}_n(x_n^*)) \right]^2.$$

4. Output one-sided CI on $\mu_{\hat{x}}$,

$$\left[0, G_n(\hat{x}) + \frac{t_{n-1, \alpha} s_n(x_n^*)}{\sqrt{n}} \right]. \quad (2)$$

Next we give a proposition that establishes consistency of the estimators. Let X^* denote the set of optimal solutions to (SP) and let $x_{\min}^* \in \arg \min_{x \in X^*} \text{var}[f(\hat{x}, \tilde{\xi}) - f(x, \tilde{\xi})]$. In other words, x_{\min}^* is a solution with minimum variance of $f(\hat{x}, \tilde{\xi}) - f(x, \tilde{\xi})$, $\sigma_{\hat{x}}^2(x_{\min}^*)$, among all the optimal solutions.

Proposition 1 *Assume (A1)-(A3), $\hat{x} \in X$, and that $\tilde{\xi}^1, \tilde{\xi}^2, \dots, \tilde{\xi}^n$ are i.i.d. as $\tilde{\xi}$. Then,*

- (i) $z_n^* \rightarrow z^*$, w.p.1,
- (ii) all limit points of $\{x_n^*\}$ lie in X^* , w.p.1,
- (iii) $\liminf_{n \rightarrow \infty} s_n^2(x_n^*) \geq \sigma_{\hat{x}}^2(x_{\min}^*)$, w.p.1.

We note that (i) and (ii) follow immediately from Theorem A1 of [6, p.69]. A proof of (iii) can be found in [2]. When (SP) has multiple optimum solutions, we cannot expect $\{x_n^*\}$ to have a unique limit point. However, by part (ii) of Proposition 1, all its limit points belong to the set of optimum solutions, X^* . Similarly, $\{s_n^2(x_n^*)\}$ may not have a unique limit. That is why “lim inf” appears in part (iii) of Proposition 1 instead of a “lim.” We next present the main result regarding the validity of the SRP.

Theorem 2 *Assume (A1)-(A3), $\hat{x} \in X$, and that $\tilde{\xi}^1, \tilde{\xi}^2, \dots, \tilde{\xi}^n$ are i.i.d. as $\tilde{\xi}$. Given $0 < \alpha < 1$, for the SRP,*

$$\liminf_{n \rightarrow \infty} P \left(\mu_{\hat{x}} \leq G_n(\hat{x}) + \frac{t_{n-1, \alpha} s_n(x_n^*)}{\sqrt{n}} \right) \geq 1 - \alpha. \quad (3)$$

Again, see [2] for the proof. Theorem 2 justifies construction of the approximate $(1-\alpha)$ -level one-sided confidence interval for $\mu_{\hat{x}} = Ef(\hat{x}, \tilde{\xi}) - z^*$, given in (2) *without* requiring $G_n(\hat{x}) = \bar{f}_n(\hat{x}) - z_n^*$ to be asymptotically normal. The intuitive reason for this is that minimization of the sample mean in z_n^* , while making asymptotic analysis of this random variable more difficult, projects the normal distribution so that the resulting confidence interval is conservative.

Even though our single replication procedure is computationally significantly less demanding, solving a single minimization problem might also create some problems. For instance, in step 2 of the procedure, if the minimization problem used to calculate the gap estimate yields a solution x_n^* that is equal to \hat{x} , then both the gap estimate $G_n(\hat{x})$ and the variance estimate $s_n(x_n^*)$ are 0 and consequently the CI on the optimality gap given in (2) has width 0. This can happen even though the candidate solution \hat{x} is far from optimal, for small sample sizes. Of course, Proposition 1 eliminates this possibility as the sample size grows large, but this still indicates that care must be exercised in employing the single-replication procedure. So, in [2] we introduce two-replication procedures and discuss other techniques such as using ϵ -optimal solutions in order to “safeguard” the procedure and achieve good empirical coverage results.

Acknowledgements

This research is partially supported by the National Science Foundation under Grant DMI-0217927.

References

- [1] S. Ahmed and A. Shapiro. The sample average approximation method for stochastic programs with integer recourse. Technical report, Optimization Online, <http://www.optimization-online.org/>, 2002.
- [2] G. Bayraksan and D.P. Morton. Assessing solution quality in stochastic programs. Technical report, Graduate Program in Operations Research, The University of Texas at Austin, 2004.
- [3] A.J. Kleywegt, A. Shapiro, and T. Homem-de-Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, pages 479–502, 2001.
- [4] W.K. Mak, D.P. Morton, and R.K. Wood. Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, 24:47–56, 1999.
- [5] V.I. Norkin, G.Ch. Pflug, and A. Ruszczyński. A branch and bound method for stochastic global optimization. *Mathematical Programming*, 83:425–450, 1998.
- [6] R.Y. Rubinstein and A. Shapiro. *Discrete Event Systems: Sensitivity and Stochastic Optimization by the Score Function Method*. John Wiley & Sons, Chichester, 1993.