

On complexity of Shmoys - Swamy class of two-stage linear stochastic programming problems

Arkadi Nemirovski* Alexander Shapiro †

July 28, 2006

Abstract

We consider a class of two-stage linear stochastic programming problems, introduced by Shmoys and Swamy [4], motivated by a relaxation of a stochastic set cover problem. We show that the sample size required to solve this problem by the sample average approximation (SAA) method with a relative accuracy $\kappa > 0$ and confidence $1 - \alpha$ is polynomial in κ , $\log(\alpha^{-1})$, dimensions of the problem and a certain parameter λ . This implies that such problems can be solved in time polynomial in these parameters, the result obtained in [4] by a different method.

*Georgia Institute of Technology, Atlanta, Georgia 30332, USA, nemirovs@isye.gatech.edu

†Georgia Institute of Technology, Atlanta, Georgia 30332, USA, ashapiro@isye.gatech.edu. Research of this author was partly supported by the NSF grant DMS-0510324.

1 Introduction

Consider the problem:

$$\begin{aligned} \text{Min}_{x \geq 0, r \geq 0} \quad & \sum_{S \in \mathfrak{S}} w_S^I x_S + \sum_{S \in \mathfrak{S}, A \in \mathfrak{A}} p_A w_S^{II} r_{A,S} \\ \text{s.t.} \quad & \sum_{S: e \in S} x_S + \sum_{S: e \in S} r_{A,S} \geq 1, \quad e \in A, \quad A \in \mathfrak{A}. \end{aligned} \quad (1)$$

Here U is a finite set of cardinality n , $\mathfrak{S} = \{S_1, \dots, S_m\}$ is a family of m subsets of U , $\mathfrak{A} = 2^{\cup_{i=1}^m S_i}$ is the family of subsets of the set $\cup_{i=1}^m S_i$, $x = (x_S)_{S \in \mathfrak{S}}$, $r = (r_{A,S})_{S \in \mathfrak{S}, A \in \mathfrak{A}}$, w_S^I and w_S^{II} are nonnegative weights and p_A are nonnegative numbers such that $\sum_{A \in \mathfrak{A}} p_A = 1$. We assume that $w_S^I > 0$ for all $S \in \mathfrak{S}$.

Problem (1) can be viewed as a two-stage linear stochastic programming problem with scenarios $A \in \mathfrak{A}$, taken with respective probabilities p_A . It can be written in the following equivalent form

$$\text{Min}_{x \geq 0} \sum_{S \in \mathfrak{S}} w_S^I x_S + \sum_{A \in \mathfrak{A}} p_A F_A(x), \quad (2)$$

with $F_A(x)$ being the optimal value of the second stage problem:

$$\begin{aligned} \text{Min}_{r \geq 0} \quad & \sum_{S \in \mathfrak{S}} w_S^{II} r_{A,S} \\ \text{s.t.} \quad & \sum_{S: e \in S} r_{A,S} \geq 1 - \sum_{S: e \in S} x_S, \quad e \in A. \end{aligned} \quad (3)$$

It is not difficult to see that if $(\bar{x}_S)_{S \in \mathfrak{S}}$ is an optimal solution of problem (2), then $\bar{x}_S \in [0, 1]$, $S \in \mathfrak{S}$. Therefore we can add the constraint $x \in [0, 1]^m$ to the first stage problem (2). Also if $w_S^I > w_S^{II}$ for some $S \in \mathfrak{S}$, then $\bar{x}_S = 0$ for every optimal solution of the first stage problem. Therefore we can assume without loss of generality that $w_S^I \leq w_S^{II}$ for all $S \in \mathfrak{S}$. We denote by Opt the optimal value of problem (1).

Problem (1) is motivated by a relaxation of a set cover problem. The deterministic (weighted) set-cover problem can be formulated as follows: “given a universe U of n elements and a family \mathfrak{S} of m subsets of U , we want to choose a minimum-weight collection of sets so that each element e is contained in some chosen set”. Relaxation of the involved integrality constraints and introduction of a probabilistic structure of scenarios $A \in \mathfrak{A}$ leads to the above problem (1) (see Shmoys and Swamy [4] and references therein). The main result of [4] is to show that it is possible to solve problem (1) to a relative optimality $\kappa > 0$ with a given confidence $1 - \alpha$ in time polynomial in n, m, κ and $\lambda := \max_{S \in \mathfrak{S}} \frac{w_S^{II}}{w_S^I}$.

Consider the so-called sample average approximation (SAA) of problem (2). That is, a random sample A_1, \dots, A_N of independent realizations of scenarios is generated and problem (2) is approximated by

$$\text{Min}_{x \geq 0} \sum_{S \in \mathfrak{S}} w_S^I x_S + \frac{1}{N} \sum_{i=1}^N F_{A_i}(x). \quad (4)$$

The two-stage problem (1) is linear and has a finite number of scenarios. Therefore, w.p.1 for N large enough any optimal solution of the SAA problem (4) is an optimal solution of the true problem (2) and, moreover, probability of that event approaches one exponentially fast with increase of the sample size N (cf., [1]). This is a qualitative type result. In the next section we discuss estimates of the sample size N which, for given $\varepsilon > 0$ and $\alpha \in (0, 1)$, guarantee with probability at least $1 - \alpha$ that any optimal solution of the SAA problem (4) is an ε -optimal solution of the “true” problem (2). In section 3 we discuss such estimates for the relative accuracy $\varepsilon = \kappa \text{Opt}$, for a given $\kappa > 0$.

2 Sample size estimates

Let us observe that $F_A(x) = Q_A(\chi)$, where $\chi = (\chi_e)_{e \in A}$, with $\chi_e := 1 - \sum_{S: e \in S} x_S$, and $Q_A(\chi)$ is the optimal value of the problem

$$\begin{aligned} \text{Min}_{r \geq 0} \quad & \sum_{S \in \mathfrak{S}} w_S^H r_{A,S} \\ \text{s.t.} \quad & \sum_{S: e \in S} r_{A,S} \geq \chi_e, \quad e \in A. \end{aligned} \quad (5)$$

The function $Q_A(\cdot)$ is convex (and piecewise linear) and its subdifferential at a point χ is given by the set of optimal solutions of the dual problem

$$\begin{aligned} \text{Max}_{\mu \geq 0} \quad & \sum_{e \in A} \chi_e \mu_e \\ \text{s.t.} \quad & w_S^H \geq \sum_{e \in S \cap A} \mu_e, \quad S \in \mathfrak{S}_A, \end{aligned} \quad (6)$$

where $\mathfrak{S}_A := \{S \in \mathfrak{S} : S \cap A \neq \emptyset\}$.

Now consider two vectors $x = (x_S)_{S \in \mathfrak{S}}$ and $x' = (x'_S)_{S \in \mathfrak{S}}$. Then

$$\begin{aligned} F_A(x') - F_A(x) &= Q_A(\chi') - Q_A(\chi) = \sum_{e \in A} \mu'_e (\chi'_e - \chi_e) \\ &= \sum_{e \in A} \mu'_e \left(\sum_{S: e \in S} x_S - \sum_{S: e \in S} x'_S \right) = \sum_{S \in \mathfrak{S}_A} \sum_{e \in S \cap A} \mu'_e (x_S - x'_S), \end{aligned}$$

where $\mu' = (\mu'_e)_{e \in A}$ is a subgradient of $Q_A(\cdot)$ at some point of the segment joining points χ and χ' . Consequently

$$|F_A(x') - F_A(x)| \leq \sum_{S \in \mathfrak{S}_A} \sum_{e \in S \cap A} \mu'_e |x'_S - x_S| \leq \sum_{S \in \mathfrak{S}_A} \sum_{e \in S \cap A} \mu'_e \|x' - x\|_\infty,$$

where $\|\cdot\|_\infty$ denotes the max-norm. Then, by the feasibility constraint of problem (6), it follows that

$$|F_A(x') - F_A(x)| \leq \sum_{S \in \mathfrak{S}_A} w_S^H \|x' - x\|_\infty \leq \sum_{S \in \mathfrak{S}} w_S^H \|x' - x\|_\infty, \quad \forall A \in \mathfrak{A}. \quad (7)$$

That is, with respect to the max-norm, the function $F_A(\cdot)$ is Lipschitz continuous with Lipschitz constant less than or equal to $\sum_{S \in \mathfrak{S}} w_S^I$. It follows (cf., [2, 3]) that the sample size

$$N \geq O(1) \left(\frac{\sum_{S \in \mathfrak{S}} w_S^I}{\varepsilon} \right)^2 \left[m \log \left(\frac{\sum_{S \in \mathfrak{S}} w_S^I}{\varepsilon} \right) + \log \left(\frac{1}{\alpha} \right) \right] \quad (8)$$

guarantees that an $\varepsilon/2$ -optimal solution of the corresponding SAA problem solves the true problem (2) with accuracy $\varepsilon > 0$ and probability at least $1 - \alpha$.

3 Sample size estimates for relative near optimality

In this section we discuss an estimate of the sample size which is required in order to solve the true problem with *relative* accuracy $\varepsilon = \kappa \text{Opt}$, for some $\kappa > 0$. Recall that $\lambda := \max_{S \in \mathfrak{S}} \frac{w_S^I}{w_S}$.

Note that because of the assumption that $w_S^I \leq w_S^II$, $S \in \mathfrak{S}$, we have that $\lambda \geq 1$. We can formulate now the main result of this paper.

Theorem 1 *Given constants $\kappa > 0$ and $\alpha \in (0, 1)$, and for $\varepsilon := \kappa \text{Opt}$ and the sample size satisfying*

$$N \geq O(1) m^2 n^4 \lambda^6 \kappa^{-2} \left[m \log (mn \lambda \kappa^{-1}) + \log(\alpha^{-1}) \right], \quad (9)$$

we have that with probability at least $1 - \alpha$ any optimal solution of the SAA problem (4) is an ε -optimal solution of the true problem (2).

The above estimate (9), of the sample size, is polynomial in parameter λ , dimension parameters m and n , relative accuracy parameter κ and $\log(\alpha^{-1})$, where $1 - \alpha$ can be viewed as a confidence (probabilistic) level of the corresponding statement. This implies that with confidence $1 - \alpha$, the considered two-stage linear stochastic programming problem can be solved in time polynomial in m, n, κ, λ , and $\log(\alpha^{-1})$. This result was derived in Shmoys and Swamy [4] by analyzing a specific (ellipsoid type) algorithm.

In the remainder of this section we give a proof of the above theorem. We proceed in a few steps. Define $\rho := \frac{1}{2n\lambda}$,

$$p_e := \text{Prob}\{A \in \mathfrak{A} : e \in A\} = \sum_{A: e \in A} p_A \quad \text{and} \quad U' := \{e \in U : p_e \geq \rho/2\},$$

$$w_e := \min_{S: e \in S} w_S^II, \quad \bar{w} := \max_{e \in U'} w_e \quad \text{and} \quad \mathfrak{S}' := \{S \in \mathfrak{S} : w_S^II \leq \rho^{-1} \bar{w}\}.$$

Note that since $\lambda \geq 1$, we have that $\rho \leq 1/(2n)$, and that $\sum_{e \in U} p_e \geq \sum_{A \in \mathfrak{A}} p_A = 1$. It follows that the set U' is nonempty. Note also that

$$\{e \in U : p_e \geq \rho\} \subset U' \subset \{e \in U : p_e \geq \rho/4\}. \quad (10)$$

Lemma 1 *Let (\bar{x}, \bar{r}) be an optimal solution of the problem (1). Then: (i) $\text{Opt} \geq \lambda^{-1} \rho \bar{w}/4$, and (ii) $\bar{x}_S = 0$ for every $S \notin \mathfrak{S}'$.*

Proof. Let us prove (i). Let $e' \in U'$ be such that $w_{e'} = \bar{w}$, i.e., $e' \in \arg \max_{e \in U'} w_e$. Note that, by the definition of w_e , we have that $w_S^I \geq \bar{w}$ for any $S \in \mathfrak{S}$ such that $e' \in S$, and recall that $\lambda \geq 1$. Then we can write

$$\begin{aligned}
\text{Opt} &= \sum_{S \in \mathfrak{S}} w_S^I \bar{x}_S + \sum_{S \in \mathfrak{S}, A \in \mathfrak{A}} p_A w_S^I \bar{r}_{A,S} \\
&\geq \sum_{S: e' \in S} w_S^I \bar{x}_S + \sum_{A: e' \in A} p_A \sum_{S: e' \in S} w_S^I \bar{r}_{A,S} \\
&\geq \sum_{S: e' \in S} \lambda^{-1} \bar{w} \bar{x}_S + \sum_{A: e' \in A} p_A \sum_{S: e' \in S} \bar{w} \bar{r}_{A,S} \\
&\geq \sum_{A: e' \in A} p_A \left(\sum_{S: e' \in S} \lambda^{-1} \bar{w} \bar{x}_S + \sum_{S: e' \in S} \bar{w} \bar{r}_{A,S} \right) \\
&\geq \lambda^{-1} \bar{w} \sum_{A: e' \in A} p_A \sum_{S: e' \in S} (\bar{x}_S + \bar{r}_{A,S}).
\end{aligned}$$

Moreover, $\sum_{A: e' \in A} p_A = p_{e'}$ and $\sum_{S: e' \in S} (\bar{x}_S + \bar{r}_{A,S}) \geq 1$ by feasibility of (\bar{x}, \bar{r}) . It follows that

$$\text{Opt} \geq \lambda^{-1} \bar{w} p_{e'} \geq \lambda^{-1} \bar{w} \rho / 4,$$

where the last inequality follows from the second inclusion of (10). This proves (i).

Let us prove (ii). Consider $\Omega := \sum_{S \in \mathfrak{S} \setminus \mathfrak{S}'} w_S^I \bar{x}_S$. It suffices to show that $\Omega = 0$. Consider the following modification of the optimal solution (\bar{x}, \bar{r}) .

- (1) Let us reduce all \bar{x}_S , $S \in \mathfrak{S} \setminus \mathfrak{S}'$, to zero keeping remaining \bar{x}_S , $S \in \mathfrak{S}'$, unchanged. This will result in getting a new first-stage solution, denoted x^* , and reducing the associated cost of the first stage by Ω .
- (2) The corresponding solution (x^*, \bar{r}) may be infeasible for the second-stage problem (3). We correct this by replacing $\bar{r}_{A,S}$ with $r_{A,S}^* := \bar{r}_{A,S} + \delta_{A,S}$, where the corrections $\delta_{A,S}$ are defined as follows. Given $A \in \mathfrak{A}$, we start by setting all $\delta_{A,S}$ to zero. Then we look in some order, one after another, at the elements $e \in A$. For current $e \in A$, we pick among the sets $S \in \mathfrak{S}$ containing e , one (be denoted S_e) with the smallest possible $w_{S_e}^I$, so that $w_{S_e}^I = w_e$. Consequently we increase the value of δ_{A,S_e} from 0 to $\gamma_e := \sum_{S \in \mathfrak{S}': e \in S} \bar{x}_S$.

Let us observe that we end up with a feasible solution (x^*, r^*) of the second-stage problem (3). Indeed, for every $A \in \mathfrak{A}$ and $e \in A$ we have

$$\begin{aligned}
\sum_{S: e \in S} (r_{A,S}^* + x_S^*) &= \sum_{S \in \mathfrak{S}': e \in S} (r_{A,S}^* + x_S^*) + \sum_{S \notin \mathfrak{S}': e \in S} (r_{A,S}^* + x_S^*) \\
&\geq \sum_{S \in \mathfrak{S}': e \in S} (\bar{r}_{A,S} + \bar{x}_S) + \sum_{S \notin \mathfrak{S}': e \in S} (\bar{r}_{A,S} + \delta_{A,S}) \\
&\geq \sum_{S \in \mathfrak{S}': e \in S} (\bar{r}_{A,S} + \bar{x}_S) + \sum_{S \notin \mathfrak{S}': e \in S} \bar{r}_{A,S} + \delta_{A,S_e} \\
&= \sum_{S \in \mathfrak{S}': e \in S} (\bar{r}_{A,S} + \bar{x}_S) + \sum_{S \notin \mathfrak{S}': e \in S} \bar{r}_{A,S} + \sum_{S \in \mathfrak{S}': e \in S} \bar{x}_S \\
&= \sum_{S \in \mathfrak{S}: e \in S} (\bar{r}_{A,S} + \bar{x}_S) \geq 1.
\end{aligned}$$

We obtain that: (x^*, r^*) is a feasible solution of the second-stage problem (3), the first-stage cost of this solution is less by Ω than the optimal cost \mathbf{Opt} (associated with solution (\bar{x}, \bar{r})), the second-stage cost associated with (x^*, r^*) is greater than the corresponding quantity for the solution (\bar{x}, \bar{r}) by at most $\Theta := \sum_{A \in \mathfrak{A}} p_A \sum_{S \in \mathfrak{S}} w_S^I \delta_{A,S}$. Note since the total cost associated with (x^*, r^*) should be greater than or equal to the optimal cost associated with (\bar{x}, \bar{r}) , and hence we have that $\Theta \geq \Omega$. We also have

$$\begin{aligned} \Theta &= \sum_{A \in \mathfrak{A}} p_A \sum_{e \in A} w_{S_e}^I \sum_{S \notin \mathfrak{S}': e \in S} \bar{x}_S = \sum_{A \in \mathfrak{A}} p_A \sum_{e \in A} w_e \sum_{S \notin \mathfrak{S}': e \in S} \bar{x}_S \\ &= \sum_{e \in U} w_e \underbrace{\left(\sum_{A: e \in A} p_A \right)}_{p_e} \underbrace{\left(\sum_{S \notin \mathfrak{S}': e \in S} \bar{x}_S \right)}_{\gamma_e} = \sum_{e \in U'} w_e p_e \gamma_e + \sum_{e \notin U'} w_e p_e \gamma_e. \end{aligned} \quad (11)$$

Now for $e \in U'$ we have

$$w_e p_e \gamma_e \leq w_e \sum_{S \notin \mathfrak{S}'} \bar{x}_S \leq \frac{w_e}{\rho^{-1} \bar{w}} \sum_{S \notin \mathfrak{S}'} w_S^I \bar{x}_S$$

(recall that $w_S^I \geq \rho^{-1} \bar{w}$ for $S \notin \mathfrak{S}'$). Also recalling that $w_S^I \leq \lambda w_S^I$ and that for $e \in U'$, $w_e \leq \bar{w} = \max_{e' \in U'} w_{e'}$, we conclude that:

$$e \in U' \text{ implies } w_e p_e \gamma_e \leq \rho \lambda \sum_{S \notin \mathfrak{S}'} w_S^I \bar{x}_S = \rho \lambda \Omega. \quad (12)$$

For $e \notin U'$ and $S \in \mathfrak{S}$ such that $e \in S$ we have that $w_S^I \geq \lambda^{-1} w_S^I \geq \lambda^{-1} w_e$, i.e., $\lambda w_e^{-1} w_S^I \geq 1$, and hence

$$\gamma_e = \sum_{S \notin \mathfrak{S}': e \in S} \bar{x}_S \leq \lambda w_e^{-1} \sum_{S \notin \mathfrak{S}': e \in S} w_S^I \bar{x}_S \leq \sum_{S \notin \mathfrak{S}': e \in S} \bar{x}_S \leq \lambda w_e^{-1} \sum_{S \notin \mathfrak{S}'} w_S^I \bar{x}_S \leq \lambda w_e^{-1} \Omega.$$

It follows that

$$e \notin U' \text{ implies } w_e p_e \gamma_e \leq \lambda \Omega p_e \leq \rho \lambda \Omega, \quad (13)$$

where the last inequality holds since $e \notin U'$ combined with (10).

Combining (11) with (12) and (13) we obtain $\Theta \leq n \rho \lambda \Omega$. Since $\Theta \geq \Omega$, it follows that $\Omega \leq n \rho \lambda \Omega$. Since $\Omega \geq 0$ and by the definition of ρ we have that $n \rho \lambda = 1/2 < 1$, it follows that $\Omega = 0$. This completes the proof. ■

Proof of theorem 1. Note that the SAA problem (4) is of the same form as the true problem (2) and can be obtained from (2) by replacing the true distribution $(p_A)_{A \in \mathfrak{A}}$ with its sample estimate. Consider the sample average estimate

$$\hat{p}_e := N^{-1} \text{card}\{i : e \in A_i, i = 1, \dots, N\}$$

of p_e , and respective sample analogues

$$\hat{U}' := \{e \in U : \hat{p}_e \geq 3\rho/4\}, \quad \hat{w} := \max_{e \in \hat{U}'} w_e \text{ and } \hat{\mathfrak{S}}' := \{S \in \mathfrak{S} : w_S^I \leq \rho^{-1} \hat{w}\}.$$

Let \hat{x} be an optimal solution of the first stage of the corresponding SAA problem. By part (ii) of Lemma 1 we have that $\hat{x}_S = 0$ for every $S \notin \widehat{\mathfrak{S}}'$.

Moreover, by Hoeffding's inequality we have for a given $e \in U$, that

$$\text{Prob}\{\hat{p}_e < p_e - \rho/4\} \leq e^{-2(\rho/4)^2 N},$$

and hence probability that there is $e \in U$ such that $\hat{p}_e < p_e - \rho/4$ is less than or equal to $ne^{-\rho^2 N/8}$. Therefore, with probability at least $1 - \alpha$, the sample size N satisfying

$$N \geq 8\rho^{-2} \log(n/\alpha) = 32n^2\lambda^2 \log(n/\alpha) \quad (14)$$

guarantees that $\hat{p}_e \geq p_e - \rho/4$ for all $e \in U$. It follows that for sample size N satisfying (14) we have with probability at least $1 - \alpha$ that \widehat{U}' is a subset of U' , and hence $\hat{w} \leq \bar{w}$ and $\widehat{\mathfrak{S}}'$ is a subset of \mathfrak{S}' . We obtain that for sample size N satisfying (14), $\hat{x}_S = 0$ for every $S \notin \mathfrak{S}'$ with probability at least $1 - \alpha$. Consequently, for sample size N satisfying (14) and confidence at least $1 - \alpha$, we can replace the set \mathfrak{S} of the true and SAA problems by the set \mathfrak{S}' . That is, the true problem (2) is replaced by the problem:

$$\text{Min}_{0 \leq x'_S \leq 1} \sum_{S \in \mathfrak{S}'} w_S^I x'_S + \sum_A p_A f_A(x'), \quad (15)$$

where $x' := (x'_S)_{S \in \mathfrak{S}'}$ and $f_A(x')$ is the optimal value of the second stage problem:

$$\begin{aligned} & \text{Min}_{r \geq 0} \sum_{S \in \mathfrak{S}} w_S^I r_{A,S} \\ & \text{s.t.} \quad \sum_{S \in \mathfrak{S}': e \in S} (r_{A,S} + x'_S) + \sum_{S \notin \mathfrak{S}': e \in S} r_{A,S} \geq 1, \quad e \in A. \end{aligned} \quad (16)$$

In a way similar to derivations of section 2 it is possible to show that for any x', x'' it follows that

$$|f_A(x'') - f_A(x')| \leq \sum_{S \in \mathfrak{S}'} w_S^I \|x'' - x'\|_\infty \leq m\rho^{-1}\bar{w}\|x'' - x'\|_\infty,$$

where the last inequality follows from the definition of \mathfrak{S}' . Consequently, an estimate of the required sample size can be written as

$$N \geq O(1) \left(\frac{m\rho^{-1}\bar{w}}{\kappa \text{Opt}} \right)^2 \left[m \log \left(\frac{m\rho^{-1}\bar{w}}{\kappa \text{Opt}} \right)^2 + \log \left(\frac{1}{\alpha} \right) \right], \quad (17)$$

which together with the estimate of Opt given in Lemma 1 gives the required estimate (9). ■

References

- [1] Shapiro, A., and Homem-de-Mello, T., On rate of convergence of Monte Carlo approximations of stochastic programs, *SIAM J. Optimization*, 11, 70-86 (2000).

- [2] Shapiro, A., Monte Carlo sampling methods, in: Ruszczyński, A. and Shapiro, A., (Eds.), *Stochastic Programming*, Handbook in OR & MS, Vol. 10, North-Holland Publishing Company, Amsterdam, 2003.
- [3] Shapiro, A. and Nemirovski, A., On complexity of stochastic programming problems, in: *Continuous Optimization: Current Trends and Applications*, pp. 111-144, V. Jeyakumar and A.M. Rubinov (Eds.), Springer, 2005.
- [4] Shmoys, D.B. and Swamy, C., Stochastic optimization is (almost) as easy as deterministic optimization, *Proceedings of FOCS*, 228-237 (2004).