Approximation Algorithms for 2-stage and Multi-stage Stochastic Optimization

David B. Shmoys[1], Chaitanya Swamy[2]

[1] Dept. of Computer Science, Cornell University, Ithaca, NY 14853, USA.
shmoys@cs.cornell.edu
[2] Chaitanya Swamy, Center for the Mathematics of Information, Caltech
Pasadena, CA 91125, USA. cswamy@ist.caltech.edu

Abstract. Stochastic optimization problems provide a means to model uncertainty in the input data where the uncertainty is modeled by a probability distribution over the possible realizations of the data. We consider the well-studied paradigm of stochastic recourse models, in which the realized input is revealed through a series of stages and one can take decisions in each stage in response to the new information learned. We obtain the first approximation algorithms for a variety of 2-stage and k-stage stochastic linear and integer optimization problems where the underlying random data is given by a "black box" and no restrictions are placed on the recourse costs: one can merely sample data from this distribution, but no direct information about the distributions is given. Our contributions are twofold. First, we give a *fully polynomial approximation scheme* for solving a broad class of 2-stage and k-stage linear programs, where k is not part of the input, that is, we show that using only sampling access to the underlying distribution, one can, for any $\epsilon > 0$, compute a solution of cost guaranteed to be within a $(1+\epsilon)$ factor of the optimum, in time polynomial in $\frac{1}{\epsilon}$ and the size of the input. To the best of our knowledge, this is the first such result that shows that (a class) of multi-stage stochastic programs can be solved to near-optimality in polynomial time. Second, we give a rounding approach for stochastic integer programs that shows that approximation algorithms for a deterministic analogue yields, with a small constant-factor loss, provably near-optimal solutions for the stochastic generalization. Thus we obtain approximation algorithms for several stochastic problems, including the stochastic versions of the set cover, vertex cover, facility location, multicut (on trees) and multicommodity flow problems.

Keywords. Algorithms, approximation algorithms, randomized algorithms, stochastic optimization, convex optimization.

1 Introduction

Stochastic optimization problems attempt to handle uncertainty in the input data by modeling the uncertainty by a probability distribution over possible realizations of the actual data called scenarios. We shall consider a broad class

Dagstuhl Seminar Proceedings 05031
Algorithms for Optimization with Incomplete Information
http://drops.dagstuhl.de/opus/volltexte/2005/72

of these problems, called *multi-stage stochastic programming problems with recourse*, where the uncertainty evolves through a series of stages and one take decisions in each stage in response to the new information learned. Multi-stage stochastic programming is an area that has received a great deal of attention within the Operations Research community, both in terms of asymptotic convergence results, as well as computational work in a wide variety of application domains. These problems are often computationally quite difficult, both from a practical perspective, as well from the point of view of computational complexity theory with even extremely specialized (sub)problems being $\#P$-complete [4].

Our Results We obtain the first approximation algorithms for a variety of multi-stage stochastic linear and integer optimization problems without placing any restrictions on the underlying probability distribution or on the cost structure of the input. Our results are obtained in the "black-box" model where one is only provided with sampling access to the distribution, but no direct information about the distributions is given. For multi-stage problems, we require a procedure that can generate, given a series of outcomes for the initial stages, a sample of the input according to the conditional distribution (given those outcomes). The results mentioned here are taken from two papers [13,16] and a note [17].

Our results have two principal components. First, for a broad class of 2-stage and k-stage stochastic linear programs (LPs), where k is not part of the input, we devise an algorithm, that given any $\epsilon > 0$, computes a solution of objective function value within $(1+\epsilon)$ of the optimum, in time polynomial in $\frac{1}{\epsilon}$, the input size, and a parameter λ that is the ratio of the cost of the same action in successive stages, which is a lower bound on sample complexity in the black-box model. The class of LPs considered is rich enough to capture the fractional versions of a variety of combinatorial optimization problems, such as, multicommodity flow problems, covering problems, facility location problems, connectivity problems. The algorithm in [13] for 2-stage programs, is based on reformulating the stochastic linear program, which has both an exponential number of variables and an exponential number of constraints, as a compact convex program, and adapting the ellipsoid method from convex optimization to solve the resulting program to near optimality. In doing so, a significant difficulty that we must overcome is that even evaluating the objective function of this convex program at a given point may be quite difficult and provably hard. This algorithm uses a suitably defined approximate subgradient, which we show is computable in polynomial time using samples from the distribution, to generate cutting planes for use in the ellipsoid method.

More recently, in [17], using this notion of an approximate subgradient, we show that the *Sample Average Approximation* (SAA) method, which is the algorithm of choice in the computational stochastic programming literature, is also a polynomial time algorithm for the class of 2-stage problems considered in [13]. The SAA method is a natural approach to computing solutions in the black-box model where one first samples from the distribution some N times, and then solves a sample average problem where the actual distribution is approximated by the distribution induced by the samples. By arguing that the subgradient of the sample average objective function is an approximate subgradient of the true function, we are able to show that for any $\epsilon > 0$, the sample size required for guaranteeing a near-optimal solution is polynomially bounded in the input size, $\frac{1}{\epsilon}$, and λ, thereby obtaining a more efficient procedure for the 2-stage problem. Our proof technique is different from that of Kleywegt, Shapiro, and Homem-De-Mello [8], and yields a much better bound on the sample size than the bound in [8] for 2-stage problems.

Furthermore, using the framework of establishing closeness in subgradients, we were able to show recently [16] that the SAA method converges in polynomial time even for (our class of) multi-stage stochastic programs with an arbitrary distribution. To the best of our knowledge, in the black-box model, no bounds were known previously on the sample size required to guarantee near-optimality for multi-stage programs with arbitrary distributions, either for the SAA method, or for any other algorithm. The only work we are aware of is a recent result of Shapiro [12] which proves bounds on the sample size required by the SAA method under the strong assumption that *the distributions in the different stages are independent*.

Complementing our results on stochastic linear programming, we give a general rounding approach to convert a fractional solution to the stochastic LP to an integer solution, that shows that any LP-based approximation guarantee for the deterministic analogue (where all the data is known is advance) yields a guarantee for the stochastic generalization with only a small constant-factor loss in the guarantee. Thus, we can lift existing algorithms and guarantees for deterministic integer problems to obtain approximation algorithms for stochastic integer optimization problems. We thereby obtain approximation algorithms for several 2-stage and multi-stage problems, including the stochastic versions of the set cover, vertex cover, facility location, multicut (on trees) and multicommodity flow problems. Moreover, the performance guarantees we obtain, in several applications, improve upon previous results that were obtained in weaker models [2,10,7,5].

Related Work Although stochastic recourse problems, both linear and integer programs, have been extensively studied, relatively little is known about polynomial-time algorithms that deliver provably near-optimal solutions to the stochastic linear or integer program. We briefly review some of the work that deals with proving such worst-case results; for a more detailed account, the reader is referred to the papers [13,16], Swamy [15], Birge and Louveaux [1], and the survey by Stougie and van der Vlerk [14].

In the stochastic programming literature, it is common to distinguish between the 2-stage setting and the multi-stage setting. The SAA method for solving 2-stage stochastic programs has been well-studied in the stochastic programming literature, but relatively fewer results are known that bound the the sample size required to obtain a near-optimal solution (with high probability). Kleywegt, Shapiro, and Homem-De-Mello [8] (see also [11]) prove a bound that is polynomial in the dimension of the problem, but depends on the variance of a certain quantity (calculated using the scenario distribution) that might be exponential in the input size and the parameter λ, even for our structured class of LPs. The dependence on λ is unavoidable in the black-box model; ours is the first result to show an upper bound on the sample size that is polynomial in the input size and λ. In an effort to reconcile the contrast between our results and the bounds in [8], Nemirovski and Shapiro (personal communication) recently showed that for the 2-stage stochastic set cover problem with non-scenario-dependent costs, if one preprocesses the input to eliminate certain first-stage decisions, then the upper bound of [8] becomes polynomial in λ. Nesterov and Vial [9], and Dyer, Kannan, and Stougie [3] also give algorithms for solving stochastic linear programs; both these results require a number of samples that depends on the maximum variation in the objective function value, which in general is not polynomially bounded.

The study of 2-stage stochastic integer programs from the perspective of approximation algorithms design is a relatively new area. Prior to our results, all work in this area was restricted to models where there are limitations imposed either on the class of probability distributions, or on the cost structure of the two stages. The first approximation result appears to be due to Dye, Stougie, and Tomasgard [2], who consider a resource provisioning problem in the setting where the uncertainty in the data is limited to a polynomial number of scenarios. Subsequently, there has been a series of recent papers in the Computer Science literature that has considered the stochastic versions of various (integer) combinatorial optimization problems [10,7,5]. Whereas [10,7] work with a restricted class of distributions, Gupta et al. [5] consider the "black-box" model but impose a restriction on the recourse costs.

Much less is known about multi-stage stochastic programs, both in the linear and integer case. To the best of our knowledge, other than the work of [12] on multi-stage programs where the different stages are independent, we are not aware of any previous work that proves bounds on the sample size required in the black-box model, either using the SAA method or any other method, to solve the multi-stage linear program to near-optimality. Our bounds establish the first convergence rate results for the SAA method for (a class of) multi-stage stochastic LPs with arbitrary distributions. For multi-stage integer problems with recourse, the only prior work is due to Hayrapetyan, Swamy, and Tardos [6], who give an approximation algorithm for a k-stage version of the Steiner tree problem in the black-box model, but with a somewhat restricted cost structure.

References

1. J. R. Birge and F. V. Louveaux. *Introduction to Stochastic Programming*. Springer-Verlag, NY, 1997.
2. S. Dye, L. Stougie, and A. Tomasgard. The stochastic single resource service-provision problem. *Naval Research Logistics*, 50(8):869–887, 2003. Also appeared as "The stochastic single node service provision problem", COSOR-Memorandum 99-13, Dept. of Mathematics and Computer Science, Eindhoven, Technical University, Eindhoven, 1999.

3. M. Dyer, R. Kannan, and L. Stougie. A simple randomised algorithm for convex optimisation. SPOR-Report 2002-05, Dept. of Mathematics and Computer Science, Eindhoven Technical University, Eindhoven, 2002.
4. M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. SPOR-Report 2003-20, Dept. of Mathematics and Computer Science, Eindhoven Technical University, Eindhoven, 2003.
5. A. Gupta, M. Pál, R. Ravi, and A. Sinha. Boosted sampling: approximation algorithms for stochastic optimization. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 417–426, 2004.
6. A. Hayrapetyan, C. Swamy, and É. Tardos Network design for information networks. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 933–942, 2005.
7. N. Immorlica, D. Karger, M. Minkoff, and V. Mirrokni. On the costs and benefits of procrastination: approximation algorithms for stochastic combinatorial optimization problems. In *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 684–693, 2004.
8. A. J. Kleywegt, A. Shapiro, and T. Homem-De-Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal of Optimization*, 12:479–502, 2001.
9. Y. Nesterov and J.-Ph. Vial. Confidence level solutions for stochastic programming. CORE Discussion Papers, 2000. http://www.core.ucl.ac.be/services/psfiles/dp00/dp2000-13.pdf.
10. R. Ravi and A. Sinha. Hedging uncertainty: approximation algorithms for stochastic optimization problems. In *Proceedings of the 10th International Conference on Integer Programming and Combinatorial Optimization*, pages 101–115, 2004.
11. A. Shapiro. Monte Carlo sampling methods. In A. Ruszczynski and A. Shapiro, editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, North-Holland, Amsterdam, 2003.
12. A. Shapiro. On complexity of multistage stochastic programs. Published electronically in *Optimization Online*, 2005. http://www.optimization-online.org/DB_FILE/2005/01/1041.pdf.
13. D. B. Shmoys and C. Swamy. An approximation scheme for stochastic linear programming and its application to stochastic integer programs. Journal version submitted. http://ist.caltech.edu/~cswamy/papers/stochopt.pdf. Preliminary version appeared as "Stochastic optimization is (almost) as easy as deterministic optimization" in *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 228–237, 2004.
14. L. Stougie and M. H. van der Vlerk. Approximation in stochastic integer programming. To appear in D. S. Johnson, J. K. Lenstra, and D. B. Shmoys, editors, *Approximation and Heuristics*, Handbooks in Operations Research and Management Science, North-Holland, Amsterdam.
15. C. Swamy. *Approximation Algorithms for Clustering Problems*. Ph.D. thesis, Cornell University, Ithaca, NY, 2004.
16. C. Swamy and D. B. Shmoys. Approximation algorithms for multi-stage stochastic optimization. *Manuscript*, 2005.
17. C. Swamy and D. B. Shmoys. The sample average approximation method for 2-stage stochastic optimization. *Manuscript*, 2004.