L-SHAPED METHOD FOR TWO STAGE PROBLEMS OF STOCHASTIC CONVEX PROGRAMMING

John R. Birge

Hengyong Tang

Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, MI 48109-2117

Technical Report 93-22

August 1993

L-shaped Method for Two Stage Problems of Stochastic Convex Programming

John R. Birge
Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor MI 48109 U.S.A.

Hengyong Tang
Mathematics Department Liaoning University
Shenyang Liaoning 110036 P.R. China

August 18, 1993

Abstract

In this paper we extend directly L-shaped method of Van Slyke and Wets to solving two stage problems of stochastic convex programming. The implement of the algorithm is simple, so less computation work is needed. The algorithm has certain convergence. The result of computation shows that the algorithm is effective.

Keywords: L-shaped method, Quadratic programming, Convex programming, Stochastic programming, Two stage problem with recourse.

1 Introduction

Van Slyke and Wets [1] have proposed an algorithm for solving L-shaped linear programs which is applicable to linear optimal control problems and two stage problems of stochastic linear programming. J. Birge has extended the two stage L-shaped method of Van Slyke and Wets to solving multistage problems of stochastic linear programming [2]. F. V. Louveaux has proposed an algorithm for solving piecewise convex programs which is used to solve two stage stochastic programs [3] [4]. The method of F. V. Louveaux is not direct extension of the method of Van Slyke and Wets, and is not applicable to linear stochastic programming, but has quadratic finite convergence. The algorithm of F. V. Louveaux is complex and has difficulty for solving general convex stochastic programming.

Two stage problems of stochastic convex programming are difficult for solving. For solving they are transformed into problems of linear programming by linearization. So the size of the problems is greatly increased. It is not feasible for the problems of large size.

We have directly extended L-shaped method of Van Slyke and Wets to solving two stage problems of stochastic convex programs. The method is simple and has certain convergence.

In section 2, we formulate mathematical model to be discussed. In section 3 and 4, we discuss the basis of the algorithm. The algorithm is given in section 5. The convergence of the algorithm is discussed in section 6. In section 7, the result of computation is given. In section 8, we do brief summary.

2 Problem formulation

We consider following stochastic convex programming

$$\begin{array}{lrcl} \min & f_1(x) & & \\ \text{s.t} & A_1x & = & b \\ & Bx & = & \xi(\omega) \\ & x & \geq & 0 \end{array} \tag{1}$$

where x is n–dimensional decision vector, $\xi(\omega)$ is m_1–dimensional random vector defined on the discrete probability space $(\Omega, \mathcal{F}, P)$, A_1, b, B are matrices of size $m \times n, m \times 1, m_1 \times n$ respectively. $f_1(x)$ is convex function.

Two stage problem with recourse of problem (1) is

$$\begin{array}{lrcl} \min & f_1(x) & + & Q(x) \\ \text{s.t} & A_1x & = & b \\ & x & \geq & 0 \end{array} \tag{2}$$

where $Q(x) = EQ(x, \omega)$,

$$\begin{array}{rcllcl} Q(x,\omega) & = & \min & f_2(y, & \omega) & \\ & & \text{s.t} & A_2y & = & \xi(\omega) - Bx \\ & & & y & \geq & 0 \end{array} \tag{3}$$

where y_2 is n_1–dimensional variable vector, A_2 is $m_1 \times n_1$ matrix, $f_2(y, \omega)$ is convex function on y for each $\omega \in \Omega$.

To be convenient we

$$\begin{array}{ll} K_1 = \{x| & A_1x = b, x \geq 0\}, \\ K_2 = \{x| & \text{for all } \omega \in \Omega, \text{ there exists } y, \\ & \text{such that } A_2y = \xi(\omega) - Bx, y \geq 0\}. \end{array}$$

Then feasible set of problem (2) is $X = K_1 \bigcap K_2$.

3 Feasibility cuts

The method of getting feasibility cuts is the same as one for stochastic linear programming [1].

Solve linear programming

$$\begin{aligned} z \;=\; \min \quad & e^T y^+ + e^T y^- \\ \text{s.t} \quad & A_2 y + I y^+ - I y^- = \xi(\omega) - B\overline{x} \\ & y \geq 0, y^+ \geq 0, y^- \geq 0 \end{aligned} \tag{4}$$

where $e = (1, \ldots, 1)^T$, I is an unit matrix.

Let σ be optimal multiplier vector. If $z = 0$, $\overline{x}$ is a feasible point of problem (2), otherwise we get feasibility cut

$$\sigma^T Bx \geq \sigma^T \xi(\omega) \tag{5}$$

4 Optimality cuts

For given $\omega_i \in \Omega$, we

$$\begin{aligned} K_{\omega_i} = \{x| \quad & \text{there exists } y, \text{ such that} \\ & A_2 y = \xi(\omega_i) - Bx, y \geq 0\}. \end{aligned}$$

Proposition 1 *For any $\omega_i \in \Omega$ and on all $x \in K_{\omega_i}$, $Q(x, \omega_i)$ is either a finite convex function in x or $Q(x, \omega_i)$ is identically $-\infty$.*

proof. For all $x \in K_{\omega_i}$, the following convex program is feasible

$$\begin{aligned} Q(x, \omega_i) \;=\; \min \quad & f_2(y, \omega_i) \\ \text{s.t} \quad & A_2 y = \xi(\omega_i) - Bx \\ & y \geq 0 \end{aligned} \tag{6}$$

Its dual problem is

$$\begin{aligned} \max \quad & \theta(u, v, x, \omega_i) \\ \text{s.t} \quad & u \geq 0 \end{aligned} \tag{7}$$

where

$$\theta(u, v, x, \omega_i) = \inf\{f_2(y, \omega_i) - u^T e + v^T(A_2 y - \xi(\omega_i) + Bx) | y \in R^{m_1}\}$$

If $Q(\overline{x}, \omega_i) = -\infty$, for some $\overline{x} \in K_{\omega_i}$, then $\theta(u, v, \overline{x}, \omega_i) = -\infty$ for each $u \geq 0$ [5]. Thus for all $x \in K_{\omega_i}$, we have $\theta(u, v, x, \omega_i) = -\infty$ for each $u \geq 0$, and

$$Q(x, \omega_i) = \max\{\theta(u, v, x, \omega_i) | u \geq 0\} = -\infty.$$

It remains to show that if $Q(x, \omega_i)$ is finite on the convex set K_{ω_i}, then it is convex.

$\forall\ x_1, x_2 \in K_{\omega_i}$, $x_\lambda = \lambda x_1 + (1-\lambda)x_2, \lambda \in [0,1]$, let y_1, y_2 and y_λ be optimal solution to (6) when x equals x_1, x_2 and x_λ respectively. Since $\lambda y_1 + (1-\lambda)y_2$ is a feasible solution to (6) when $x = \lambda x_1 + (1-\lambda)x_2$ (but not necessarily optimal, e.g in general $y_\lambda \neq \lambda y_1 + (1-\lambda)y_2$) and $f_2(y, \omega_i)$ is a convex function for fixed ω_i, we have

$$\begin{aligned} Q(x_\lambda, \omega_i) &= f_2(y_\lambda, \omega_i) \\ &\leq f_2(\lambda y_1 + (1-\lambda)y_2, \omega_i) \\ &\leq \lambda f_2(y_1, \omega_i) + (1-\lambda) f_2(y_2, \omega_i) \\ &= \lambda Q(x_1, \omega_i) + (1-\lambda) Q(x_2, \omega_i). \end{aligned}$$

This completes the proof of the proposition.

Proposition 2 *Let $\overline{y}_{\omega_i}$ be an optimal solution of following problem*

$$\begin{aligned} Q(\overline{x}, \omega_i) = \min\ & f_2(y,\ \omega_i) \\ s.t\quad & A_2 y = \xi(\omega_i) - B\overline{x} \\ & y \geq 0 \end{aligned} \tag{8}$$

$(\overline{u}_{\omega_i}, \overline{v}_{\omega_i})$ is corresponding optimal Lagrangian multipliers, then the linear function

$$f_2(\overline{y}_{\omega_i}, \omega_i) - \overline{u}_{\omega_i}^T e + \overline{v}_{\omega_i}^T A_2 \overline{y}_{\omega_i} - \overline{v}_{\omega_i}^T \xi(\omega_i) + \overline{v}_{\omega_i}^T B x$$

is a support of $Q(x, \xi(\omega_i))$.

proof. The dual problem of the problem (8) is

$$\begin{aligned} \max\quad & \theta(u, v, \overline{x}, \omega_i) \\ \text{s.t}\quad & u \geq 0 \end{aligned} \tag{9}$$

where

$$\theta(u, v, \overline{x}, \omega_i) = \inf\{f_2(y, \omega_i) - u^T e + v^T (A_2 y - \xi(\omega_i) + B\overline{x}) | y \in R^{m_1}\}. \tag{10}$$

Since $(\overline{u}_{\omega_i}, \overline{v}_{\omega_i})$ is optimal solution of (8), by the duality theory we have that

$$f_2(\overline{y}_{\omega_i}, \omega_i) - \overline{u}_{\omega_i}^T e + \overline{v}_{\omega_i}^T A_2 \overline{y}_{\omega_i} - \overline{v}_{\omega_i}^T \xi(\omega_i) + \overline{v}_{\omega_i}^T B\overline{x} = Q(\overline{x}, \omega_i).$$

For all $x \in K_{\omega_i}$, $(\overline{u}_{\omega_i}, \overline{v}_{\omega_i})$ is a feasible solution of all problem (7), but is not necessarily an optimal solution. Thus again by duality theory we have

$$\begin{aligned} Q(x, \omega_i) &= \max\{\theta(u, v, x, \omega_i) | u \geq 0\} \\ &\geq \theta(\overline{u}_{\omega_i}, \overline{v}_{\omega_i}, x, \omega_i) \\ &= \inf\{f_2(y, \omega_i) - \overline{u}_{\omega_i}^T e + \overline{v}_{\omega_i}^T A_2 y - \overline{v}_{\omega_i}^T \xi(\omega_i) + \overline{v}_{\omega_i}^T B x | y \in R^{m_1}\} \\ &= f_2(\overline{y}_{\omega_i}) - \overline{u}_{\omega_i}^T e + \overline{v}_{\omega_i}^T A_2 \overline{y}_{\omega_i} - \overline{v}_{\omega_i}^T \xi(\omega_i) + \overline{v}_{\omega_i}^T B x. \end{aligned} \tag{11}$$

last equality is according to

$$\begin{aligned} &\arg \quad \min_y f_2(y,\omega_i) - \overline{u}_{\omega_i}^T e + \overline{v}_{\omega_i}^T A_2 y - \overline{v}_{\omega_i}^T \xi(\omega_i) + \overline{v}_{\omega_i}^T Bx \\ = &\arg \quad \min_y f_2(y,\omega_i) - \overline{u}_{\omega_i}^T e + \overline{v}_{\omega_i}^T A_2 y - \overline{v}_{\omega_i}^T \xi(\omega_i) + \overline{v}_{\omega_i}^T B\overline{x}. \end{aligned}$$

This completes the proof of the proposition.

Proposition 1 and proposition 2 are the basis of getting optimality cuts (see following algorithm).

5 Algorithm

Algorithm 1 *Step 1. Solve the problem*

$$\begin{array}{llcllr} \min & f_1(x)+\theta & & & & \\ s.t & A_1x & = & b & & (12.1) \\ & \sigma_k^T Bx & \geq & \sigma_k^T \xi_k & k=1,\ldots,s & (12.2) \\ & -v_k^T Bx+\quad \theta & \geq & \rho_k & k=1,\ldots,t & (12.3) \\ & \qquad\qquad x & \geq & 0 & & \end{array} \tag{12}$$

Initially s and t are zero. If no constraints of the form (12.3) *are present, θ is set equal to $-\infty$, is ignored in the computation. Let x_l, θ_l be an optimal solution of* (12).

Step 2. Solve the problems

$$\begin{array}{rcll} z_i & = & \min & e^Ty^+ + e^Ty^- \\ & & s.t & A_2y + Iy^+ - Iy^- = \xi(\omega_i) - Bx_l \quad i=1,\ldots,N_1 \\ & & & y\geq 0, y^+\geq 0, y^-\geq 0 \end{array} \tag{13}$$

If for some $z_r > 0$, $\xi_r = \xi(\omega_r)$ and optimal multiplier σ_r are used to generate a feasibility cut of the form (12.2) *go to step* 1, *otherwise go to step* 3.

Step 3. Solve the problems

$$\begin{array}{llcl} \min & f_2(y, \ \omega_i) & & \\ s.t & \quad A_2y & = & \xi(\omega_i) - Bx_l \quad i=1,\ldots,N_2 \\ & \qquad y & \geq & 0 \end{array} \tag{14}$$

Let y_{ω_i} be optimal solution, $(u_{\omega_i}, v_{\omega_i})$ be corresponding optimal Lagrangian multipliers. compute

$$\rho(\omega_i) = f_2(y_{\omega_i},\omega_i) - u_{\omega_i}^T e + v_{\omega_i}^T A_2 y_{\omega_i} - v_{\omega_i}^T \xi(\omega_i),$$

$$\rho_l = \sum_{k=1}^{N_2} p_i \rho(\omega_i), \qquad v_l = \sum_{k=1}^{N_2} p_i v_{\omega_i}.$$

where p_i is probability of ω_i.

If $\rho_l + v_l^T B x_l \le \theta_l$, terminate.

Otherwise use ρ_l, v_l to generate a new optimality cut of the form (12.3) *and odd it to problem* (12), *go to step* 1.

6 Convergence of the algorithm

Theorem 1 *If for some (x_l, θ_l), such that*

$$\rho_l + v_l^T B x_l = \theta_l,$$

then x_l is an optimal solution to problem (2).

proof. Problem (2) is equivalent to

$$\begin{array}{ll} \min & f_1(x) + \theta \\ \text{s.t} & Q(x) \le \theta \\ & x \in K \end{array} \tag{15}$$

Let

$$\begin{array}{ll} \overline{K} = \{(x,\theta) | & Q(x) \le \theta, x \in K\}, \\ \overline{\overline{K}} = \{(x,\theta) | & A_1 x = b, x \ge 0 \\ & \sigma_k^T B x \ge \sigma T_k \xi_k, k = 1, \dots, s \\ & -v_k^T B x + \theta \ge \rho_k, k = 1, \dots, t\}, \end{array}$$

then $\overline{K} \subset \overline{\overline{K}}$.

According to proposition 2 we have

$$\theta_l = \rho_l + v_l^T B x_l = Q(x_l),$$

thus $(x_l, \theta_l) \in \overline{K}$.

Since (x_l, θ_l) is an optimal solution to problem (12) and $\overline{K} \subset \overline{\overline{K}}$, then (x_l, θ_l) is an optimal solution to problem (15) and x_l is an optimal solution to problem (2). This completes the proof of the theorem.

Theorem 2 *Suppose the algorithm generates an infinite sequence of (x_l, θ_l), $(\overline{x}, \overline{\theta_l})$ is the limit point of an arbitrary subsequence (x_{l_i}, θ_{l_i}), and*

$$\lim_{i \to \infty} \rho_{l_i} + v_{l_i}^T B x_{l_i} - \theta_{l_i} = 0$$

then $\overline{x}$ is an optimal solution to problem (2).

Proof. Since K is closed convex set [6] and $Q(x)$ is a convex function, then $\overline{K}$ is a closed convex set too. thus $\overline{x} \in \overline{K}$.

According to algorithm and proposition 2 we have

$$\theta_{l_i} < Q(x_{l_i}) = \rho_{l_i}^T B x_{l_i}.$$

Since

$$\lim_{i \to \infty} \rho_{l_i} + v_{l_i}^T B x_{l_i} - \theta_{l_i} = 0,$$

we have

$$\overline{\theta} = Q(\overline{x}).$$

Thus $(\overline{x}, \overline{\theta})$ is a feasible point to problem (15). Let (x^*, θ^*) be an optimal solution to problem (15), then

$$f_1(x_{l_i}) + \theta_{l_i} \leq f_1(x^*) + \theta^*.$$

By convexity of $f_1(x)$, $f_1(x)$ is a continuous function [7], thus

$$f_1(\overline{x}) + \overline{\theta} \leq f_1(x^*) + \theta^*.$$

Hence $(\overline{x}, \overline{\theta})$ is an optimal solution to problem (15), and $\overline{x}$ is an optimal solution to problem (2). This completes the proof of theorem.

7 Computational experience

To see the feasibility and effect of the algorithm we have computed some sample problems of stochastic convex programming.

We consider two stage problem with recourse of stochastic convex quadratic programming

$$\begin{array}{lrcl} \min & \frac{1}{2}x^T H_1 x & + & d_1^T + Q(x) \\ \text{s.t} & A_1 x & = & b \\ & x & \geq & 0 \end{array} \tag{16}$$

where $Q(x) = EQ(x, \omega)$,

$$\begin{array}{rclrcl} Q(x,\omega) & = & \min & \frac{1}{2}y^T H_2 y & + & d_2^T \\ & & \text{s.t} & A_2 y & = & \xi(\omega) - Bx \\ & & & y & \geq & 0 \end{array} \tag{17}$$

where A_1, A_2, B, b, d_1, d_2 are matrices of size $m \times n, m_1 \times n_1, m_1 \times n, m \times 1, n \times 1, n_1 \times 1$ respectively. H_1 is $n \times n$ positive semidefinite matrix, H_2 is positive definite matrix. $\xi(\omega)$ is m_1-dimensional random vector defined on the discrete probability space $(\Omega, \mathcal{F}, P)$, and has N scenarios.

We have computed 3 sample problems which have different size using IBM RS/6000 workstation. We use OSL (Optimization Subroutine Library) to compute the subproblems of linear programming and convex quadratic programming. The code is implemented in **Fortran 77** language.

	n	m	n_1	m_1	N
example1	8	5	6	4	16
example2	36	35	28	23	32
example3	60	53	70	40	64

In all examples, problem (15) has equality constraints and inequality constraints. When $\rho_l + v_l^T B x_l - \theta_l \leq 0.001$, we terminate computing. Theoretically, $\rho_l + v_l^T B x_l - \theta_l \geq 0$, but sometimes it probably is a negative number which absolute value is very small due to computation error. The result of the computation is following:

example	NFE	NOP	NIE	CPU	ITI	HST	LST	RVT
1	3	3	6	470	78.3	935	847	-0.609280×10^{-5}
2	2	3	5	918	183.6	2582	1221	0.412843×10^{-6}
3	0	10	10	6049	604.9	4737	1723	0.634673×10^{-3}

Where

NFE: Number of feasibility cuts.

NOP: Number of optimality cuts.

NIE: Number of iterations.

CPU: Average CPU time (microseconds).

ITI: Average time of each iteration (microseconds).

HST: High storage.

LST: Low storage.

RVT: $\rho_l + v_l^T B x_l - \theta_l$.

We have together computed N problems (13) (linear programming) and N problems (14) (convex quadratic programming). So the work of the computation is decreased and the units of storage are saved.

If A_2, d_2, H_2 are random, computation can also implemented. The result of computation shows that the algorithm is effective for stochastic convex quadratic programming.

8 Summary

This paper extends L-shaped method of Van Slyke and Wets to solving stochastic convex programming with recourse. Algorithm is a kind of decomposition method, so it possesses good character of decomposition method. The implement of the algorithm is simple and it has certain convergence. The result of the computing convex quadratic programming shows that algorithm is effective. A number of sample problems of general stochastic convex programming have to be computed to see practical efficiency of the algorithm for general stochastic convex programming.

Acknowledgment

The author acknowledge the lot of help given by Mr. Dereck Holmes. He has also supplied sample problems in the paper. Mr. Samer Takriti has also given us help when we do computation.

References

[1] R. Van Slyke and R. J. Wets, "L-shaped linear program with application to optimal control and stochastic linear programming," SIAM Journal and Applied Mathematics 17, 638-663,1969.

[2] John R. Birge, "Decomposition and partitioning methods for multistage stochastic linear programs," Operations Research, 33, 989-1007, 1985.

[3] Francois V. Louveaux, "Piecewise convex programs," Mathematical Programming, 15, 53-62, 1978.

[4] Francois V. Louveaux, "A solution method for multistage stochastic programs with recourse with application to an energy investment problem," Operations Research, 28, 889-902, 1980.

[5] M. S. Bazaraa and C. M. Shetty, Nonlinear Programming, Theory and Algorithm, John Wiley and Sons, New York, 1979.

[6] P. Kall, Stochastic Linear Programming, Springer, Berlin, 1974.

[7] R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.

[8] Richard P. O'neill, "Nested decomposition of multistage convex programs," SIAM J. Control nad Optimization, Vol 14, No 3, 409-418, 1976.