AN ALTERNATING METHOD FOR STOCHASTIC LINEAR PROGRAMMING WITH SIMPLE RECOURSE

by

Liqun Qi

Computer Sciences Technical Report #515

October 1983

An Alternating Method for Stochastic Linear Programming with Simple Recourse*

Liqun Qi

Computer Sciences Department
University of Wisconsin-Madison
Madison, Wisconsin 53706

Abstract

Stochastic linear programming with simple recourse arises naturally in economic problems and other applications. One way to solve it is to discretize the distribution functions of the random demands. This will considerably increase the number of variables and will involve discretization errors. Instead of doing this, we describe a method which alternates between solving some n-dimensional linear subprograms and some m-dimensional convex subprograms, where n is the dimension of the decision vector and m is the dimension of the random demand vector. In many cases, m is relatively small. This method converges in finitely many steps.

*Sponsored by the National Science Foundation under Grant No. MCS-8200632.

1. Introduction

The standard form of stochastic linear programs with simple recourse is as follows

$$\begin{aligned} \min \quad & cx + E(Q(x,\underset{\sim}{\xi})) \\ \text{s.t.} \quad & Ax = b \qquad\qquad (SLP) \\ & x \geq 0 \end{aligned}$$

with

$$\begin{aligned} Q(x,\underset{\sim}{\xi}) = \min \quad & q^+y^+ + q^-y^- \\ \text{s.t.} \quad & y^+ - y^- = \underset{\sim}{\xi} - Tx \qquad\qquad (RP) \\ & y^+ \geq 0,\ y^- \geq 0 \end{aligned}$$

where $x, c \in \mathbb{R}^n$, $y^+, y^-, q^+, q^- \in \mathbb{R}^m$, $b \in \mathbb{R}^k$, $A \in L(\mathbb{R}^n, \mathbb{R}^k)$, $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $\underset{\sim}{\xi}$ is an m-dimensional random vector with known distribution F, $\bar{\underset{\sim}{\xi}} < \infty$, $q = q^+ + q^- \geq 0$.

One of the economic interpretations of this model is as follows: x is the decision vector, which is nonnegative. $Ax = b$ is the resource constraint on x. T is the technology matrix which yields the linear transformation of the activities in finished products $w = Tx$. $\underset{\sim}{\xi}$ represents demand. The goal is to minimize the total expected cost which consists of two parts: the production cost cx and the penalty cost. For the ith product, $i=1,\ldots,m$, the output w_i is to be compared with the observed value ξ_i of the stochastic demand $\underset{\sim}{\xi}_i$. Any discrepancies between ξ_i and w_i, $i=1,\ldots,m$, are penalized as follows:

$$\text{if } \xi_i < w_i \text{ then the penalty is } q_i^-(w_i-\xi_i),$$

$$\text{if } \xi_i \geq w_i \text{ then the penalty is } q_i^+(\xi_i-w_i).$$

The first eventuality corresponds to excess product, the second to shortage.

The equivalent deterministic problem can be written as

$$\begin{aligned} \min_{x,w} \quad & cx + \phi(w) \\ \text{s.t.} \quad & Ax = b \\ & Tx = w \\ & x \geq 0 \end{aligned} \qquad \text{(EDP)}$$

where $w \in \mathbb{R}^m$ represents the output as mentioned above, $\phi(w) = \sum_{i=1}^{m} \phi_i(w_i)$ is separable convex and continuous,

$$\phi_i(w_i) = q_i^+ \int_{\xi_i < w_i} (w_i - \xi_i) dF_i(\xi_i) + q_i^- \int_{\xi_i > w_i} (\xi_i - w_i) dF_i(\xi_i),$$

F_i is the marginal distribution of $\underset{\sim}{\xi}_i$, $i=1,\ldots,m$, [6] [12] [13].

Suppose that $m + k$ is small comparing with n. This assumption is practical. For example, in the stochastic transportation problem [1] [8] [14] [15] [16] [17], $n = m \times k$.

There are simplex-type methods to solve (EDP) by discretizing ϕ [3] [11] (for more references, see [10]). Usually, this will considerably increase the number of variables and will involve discretization errors. To avoid these two drawbacks, we describe here a method which alternates between solving some n-dimensional linear subprograms by fixing w and some m-dimensional convex subprograms by restricting $n-m-k$ activities to zero levels. In Section 2, we define such subprograms and discuss their solvability. In Section 3, we give the algorithm. In Section 4, we prove the convergence of our method.

2. Subprograms

Definition 2.1 For any fixed $w \in \mathbb{R}^m$, the following linear program is called a linear subprogram of (EDP) and denote by $L(w)$:

$$
\begin{aligned}
\min_x \quad & cx + \phi(w) \\
\text{s.t.} \quad & Ax = b \\
& Tx = w \\
& x \geq 0 .
\end{aligned}
\qquad L(w)
$$

Without confusion, we also denote its optimal value by $L(w)$. ∎

The constant term $\phi(w)$ in the objective function does not affect the solution. However, we keep it there for comparison with (EDP).

Theorem 2.2 Suppose (EDP) is solvable. Then $L(w)$ is solvable if and only if it is feasible.

Proof "Only if" is obvious. For the "if" part, it suffices to know that $L(w)$ is bounded, then we know that $L(w)$ is solvable according to the duality theorem of linear programming. However, $L(w)$ cannot be unbounded, since otherwise (EDP) would be unbounded. This proves the theorem. ∎

If (EDP) is solvable, then for certain w^*, $L(w^*)$ will yield optimal solutions of (EDP). Such w^* is called a certainty equivalent of (EDP) [12]. If we can find a certainty equivalent, then we have almost solved (EDP). Our method will provide a way to seek such a certainty equivalent.

Denote the support vector of x by supp (x).

$$
\langle \ell = \text{supp}\ (x) \rangle := \begin{cases} \ell_i = 1 & \text{if} \quad x_i \neq 0 \\ \ell_i = 0 & \text{if} \quad x_i = 0 . \end{cases}
$$

Definition 2.3 Let M be the set of all n-dimensional vectors with $m+k$ components being "1" and others being "0". Suppose that $\ell \in M$. The following convex program is called a convex subprogram of (EDP) and is denoted by $C(\ell)$:

$$\begin{aligned} \min_{x,w} \quad & cx + \phi(w) \\ \text{s.t.} \quad & Ax = b \\ & Tx = w \\ & \text{supp}(x) \le \ell \\ & x \ge 0 \end{aligned} \qquad C(\ell)$$

Without confusion, we also denote its optimal value by $C(\ell)$. ■

The following theorem is proved in [12]. Since it is important to our discussion, we still make a proof here.

Theorem 2.4 Suppose (EDP) is solvable. Then there exists an $\ell^* \in M$ such that $C(\ell^*)$ yields an optimal solution of (EDP).

Proof Suppose (x^*, w^*) is an optimal solution of (EDP). Consider $L(w^*)$. $L(w^*)$ is solvable by Theorem 2.2. Suppose x^0 is a basic optimal solution of $L(w^*)$. Then there exists $\ell^* \in M$ such that $\text{supp}(x^0) \le \ell^*$. Then $C(\ell^*)$ yields an optimal solution (x^0, w^*) of (EDP). ■

We call such ℓ^* an optimal support of (EDP). If we find an optimal support of (EDP), we have also almost solved (EDP). The solvability of $C(\ell)$, however, is not so simple even if it is feasible and (EDP) is solvable. In fact, the solvability of (EDP) only implies the boundedness of $C(\ell)$.

Suppose $\ell \in M$. Then there are $m+k$ positive integers $i_s, s=1,\dots,m+k$ such that $1 \le i_1 < i_2 < \dots < i_{m+k} \le n$ and

$$\ell_h = \begin{cases} 1 & \text{if } h = i_s \text{ for some } s, \\ 0 & \text{otherwise}. \end{cases}$$

Let x_e and c_e be the $m+k$ dimensional vectors consisting of the i_s-th components, $s=1,\dots,m+k$, of x and c. Let A_e and T_e be the sub-matrices of A and T, consisting of the i_s-th columns, $s=1,\dots,m+k$, of A and T. Then $C(\ell)$ is equivalent to

$$\begin{aligned} \min_{x_e, w} \ & c_e x_e + \phi(w) \\ \text{s.t. } & A_e x_e = b \\ & T_e x_e = w \\ & x_e \ge 0 . \end{aligned} \tag{2.1}$$

Write $B = \begin{pmatrix} A_e \\ T_e \end{pmatrix}$. This is an $(m+k) \times (m+k)$ square matrix. Suppose $G = B^{-1}$ exists. Then (2.1) is equivalent to

$$\begin{aligned} \min_{w} \ & c_e G \begin{pmatrix} b \\ w \end{pmatrix} + \phi(w) \\ \text{s.t. } & G \begin{pmatrix} b \\ w \end{pmatrix} \ge 0 \end{aligned} \tag{2.2}$$

Let $c_e G = (g,h)$, where $g \in \mathbb{R}^k$, $h \in \mathbb{R}^m$. Then (2.2) is equivalent to

$$\begin{aligned} \min_{w} \ & \sum_{j=1}^{m} [\phi_i(w_i) + h_i w_i] + gb \\ \text{s.t. } & G \begin{pmatrix} b \\ w \end{pmatrix} \ge 0 . \end{aligned} \tag{2.3}$$

Another equivalent form of (2.1) is simply

$$\begin{aligned} \min_{x_e} \quad & c_e x_e + \phi(T_e x_e) \\ \text{s.t.} \quad & A_e x_e = b \\ & x_e \geq 0 . \end{aligned} \tag{2.4}$$

Remark 2.5 (2.3) is an m-variable convex program with linear constraints and a separable convex objective function. There are many methods to solve it [2] [4] [5] [7] [9]. We shall not discuss this here. ∎

The solvability of $C(\ell)$ can be assured by some conditions on the random vector $\tilde{\xi}$.

Theorem 2.6 Suppose $q \geq 0$ and Ω, the support of the random vector $\tilde{\xi}$, is compact. Then $C(\ell)$ is solvable for any $\ell \in M$ if it is feasible and (EDP) is solvable.

Proof If Ω is compact, then Ω_i, the support of $\tilde{\xi}_i$, $i=1,\ldots,m$, are also compact. This also implies that $\bar{\tilde{\xi}}$ exists and thus ϕ is finite. Since (EDP) is solvable, (2.1), i.e., $C(\ell)$ is bounded. Then there exists feasible point sequence $\{(x_e^r, w^r) | r=0,1,2,\ldots\}$ such that

$$\lim_{r\to\infty} [c_e x_e^r + \phi(w^r)] = \text{Inf } \{C(\ell)\} .$$

If $\{(x_e^r, w^r) | r=0,1,2,\ldots\}$ has a limiting point, then this limiting point will be an optimal solution of $C(\ell)$ since the feasible set of $C(\ell)$ is closed and the objective function of $C(\ell)$ is continuous. Suppose it has no limiting point. Then it has a limiting direction (x_e^c, w^c). Since the feasible set of $C(\ell)$ is closed convex and the objective function of $C(\ell)$ is continuous,

$$\lim_{\lambda\to\infty} [c_e(x_e^0 + \lambda x_e^c) + \phi(w^0 + \lambda w^c)] = \text{Inf } \{C(\ell)\} .$$

Now the only case for (2.1) failing to be solvable is that

$$c_e(x_e^0 + \lambda x_e^c) + \phi(w^0 + \lambda w^c) > \text{Inf } \{C(\ell)\} > -\infty , \quad \forall\lambda \geq 0 .$$

But this is impossible since $\phi(w^0 + \lambda w^c)$ is linear for λ sufficiently large (see 10.2 and 12.4 of [12]). ∎

3. The Alternating Algorithm

Algorithm 3.1 Starting from any $\ell^0 \in M$ or starting from any $w^0 \in \mathbb{R}^m$, do the following two procedures alternatively until the method stops in step 2. This (x^j, w^j) is an optimal solution.

1. From ℓ^j, solve $C(\ell^j)$ to get an optimal solution w^j.
2. From w^j, solve $L(w^j)$ to get a basic optimal solution x^{j+1}. Pick $\ell^{j+1} \in M$ such that

$$\text{supp}\ (x^{j+1}) \leq \ell^{j+1}. \tag{3.1}$$

If there is more than one basic optimal solution of $L(w^j)$, we should choose x^{j+1} such that there exists $\ell^{j+1} \in M$ satisfying (3.1) and

$$C(\ell^{j+1}) < C(\ell^j). \tag{3.2}$$

If no such ℓ^{j+1} can be found, or $\ell^{j+1} = \ell^j$, stop. ∎

Remark 3.2 To start this algorithm, we can pick any $x^0 \in \{x | Ax - b,\ x \geq 0\}$ and let $w^0 = Tx^0$. However, a good starting point should be a good estimate of the certainty equivalent w^*. According to our model, we can take $w^0 = \bar{\xi}$. ∎

Remark 3.3 If $L(w^j)$ is nondegenerate, we can simply take

$$\ell^{j+1} = \text{supp}\ (x^{j+1}). \quad ∎$$

Remark 3.4 This algorithm yields a sequence $\ell^0, \ell^1, \ldots, \ell^j$ = an optimal support. We will prove in Section 5 that $\ell^r \neq \ell^s$ for $r \neq s$. In this sense, it looks like a pivoting method in M. However, ℓ^{r+1} is not necessarily a "neighbor" vector of ℓ^r. They may be different in more than two components. ∎

4. The Convergence Theorem

Theorem 4.1 (Convergence Theorem) Suppose that $q \geq 0$, that $K = \{x | Ax = b, x \geq 0\} \neq \phi$, that (EDP) is solvable and that the support of the random variable $\tilde{\xi}$ is compact. Then Algorithm 3.1 is well-defined and stops in finitely many steps if it starts from a feasible point w^0. Furthermore, (3.2) holds for every j in this case.

Proof In Algorithm 3.1, $C(\ell^j)$ and $L(w^j)$ are always feasible if the algorithm begins from a feasible point. Therefore, by Theorems 2.2, 2.6 and the hypotheses of this theorem, we know that $C(\ell^j)$ and $L(w^j)$ are always solvable. In fact, it now suffices to prove that (3.2) holds for each ℓ^j which is not an optimal support. Then we get the conclusion since M is finite. Suppose $C(\ell^j)$ has an optimal solution $(\bar{x}_e, w^j)$. Let $\bar{x}$ be the n-dimensional vector whose components consist of $\bar{x}_e$ and 0 correspondingly. If $\bar{x}$ is not an optimal solution of $L(w^j)$, then there exists an $\ell^{j+1} \in M$ with

$$C(\ell^{j+1}) \leq L(w^j) < C(\ell^j).$$

Thus, suppose $\bar{x}$ is an optimal solution of $L(w^j)$. We shall prove that there is an optimal solution of $L(w^j)$ such that the associated ℓ^{j+1} satisfies (3.2). Suppose that w^* is a certainty equivalent. Let x^* be an optimal solution of $L(w^*)$. Since ℓ^j is not an optimal support,

$$cx^* + \phi(w^*) < c\bar{x} + \phi(w^j)$$

Let

$$(x^\lambda, w^\lambda) = \lambda(\bar{x}, w^j) + (1 - \lambda)(x^*, w^*), \quad 0 < \lambda < 1$$

According to the convexity of ϕ, we have

$$cx^{\lambda} + \phi(w^{\lambda}) < c\bar{x} + \phi(w^{j}), \quad \forall 0 < \lambda < 1 .$$

Suppose $\bar{x}^{\lambda}$ is a basic optimal solution of $L(w^{\lambda})$. Then

$$c\bar{x}^{\lambda} + \phi(w^{\lambda}) \leq cx^{\lambda} + \phi(w^{\lambda}) < c\bar{x} + \phi(w^{j}) .$$

Now pick $\ell^{\lambda} \in M$ corresponding to the optimal basis of $\bar{x}^{\lambda}$. We have

$$C(\ell^{\lambda}) \leq c\bar{x}^{\lambda} + \phi(w^{\lambda}) < c\bar{x} + \phi(w^{j}) = C(\ell^{j}) .$$

Therefore

$$\ell^{\lambda} \neq \ell^{j}, \quad \forall 0 < \lambda < 1 .$$

Therefore, since M is finite, there exists a sequence $\{\lambda^{r} | r=1,2,\ldots\}$ such that $\lambda^{r} \to 1^{-}$, $\ell^{\lambda^{r}} \equiv \ell^{j+1}$ for some $\ell^{j+1} \in M$. Let B be the basic matrix of $\binom{A}{T}$, corresponding to ℓ^{j+1}. Then $\hat{x} = B^{-1}\binom{b}{w^{j}} = \lim_{r\to\infty} B^{-1}\binom{b}{w^{\lambda^{r}}} = \lim_{r\to\infty} x^{\lambda^{r}} \geq 0$ exists. Therefore, ℓ^{j+1} also corresponds to an optimal basic solution of $L(w^{j})$, and

$$C(\ell^{j+1}) = C(\ell^{\lambda^{r}}) < C(\ell^{i}) .$$

This proves the theorem. ∎

Corollary 4.2 Let $\bar{x}$ be defined in the above proof. Then a necessary condition for w^j to be a certainty equivalent is that $\bar{x}$ be an optimal solution of $L(w^j)$. A sufficient condition for w^j to be a certainty equivalent is that $\bar{x}$ be the unique optimal solution of $L(w^j)$. ∎

Acknowledgment

I am thankful to Professor S. M. Robinson for his suggestions.

References

[1] V. Balachandran, "Generalized transportation networks with stochastic demands: an operator theoretic approach", Network 9, pp. 169-184, 1979.

[2] M. Bazaraa and C. M. Shetty, "Nonlinear Programming Theory and Algorithms", Wiley, New York, 1979.

[3] G. B. Dantzig, "Linear Programming and Extensions", Princeton University Press, Princeton, 1963.

[4] P. E. Gill and W. Murray, "Linearly constrained problems including linear and quadratic programming", In: "The State of the Art in Numerical Analysis" (D. Jacobs, ed.), pp. 313-363, Academic Press, New York, 1977.

[5] P. E. Gill, W. Murray and M. H. Wright, "Practical optimization", Academic Press, New York, 1981.

[6] P. Kall, "Stochastic Linear Programming", Springer-Verlag, Berlin, 1976.

[7] L. Qi, "Base set strategy for solving linearly constrained convex programs", Comp. Sci. Tech. Rept. #505, University of Wisconsin-Madison, 1983.

[8] L. Qi, "Forest iteration method for stochastic transportation problem", Comp. Sci. Tech. Rept. #522, University of Wisconsin-Madison, 1983.

[9] S. M. Robinson, "Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear programming algorithms", Math. Programming 7, pp. 1-16, 1974.

[10] I. M. Stancu-Minasian and M. J. Wets, "A research bibliography in stochastic programming 1955-1975", Oper. Res. 24, pp. 1078-1179, 1976.

[11] B. Strazicky, "Some results concerning an algorithm for the discrete recourse problem", in "Stochastic Programming", (M. A. H. Dempster, ed.), pp. 263-274, 1980.

[12] R. J.-B. Wets, "Solving Stochastic Programs with Simple Recourse, I", Dept. of Math., University of Kentucky, 1974.

[13] R. J.-B. Wets, "Stochastic programs with fixed recourse: the equivalent deterministic problem", SIAM Review 16, pp. 309-339, 1974.

[14] A. C. Williams, "A stochastic transportation problem", Oper. Res. 11, pp. 759-770, 1963.

[15] D. Wilson, "On a priori bounded model for transportation problems with stochastic demand and integer solutions", AIEE Transaction 4, pp. 186-193, 1972.

[16] D. Wilson, "Tighter bounds for stochastic transportation problems", AIEE Transaction 5, pp. 180-185, 1973.

[17] D. Wilson, "A mean cost approximation for transportation problems with stochastic demands", Naval Res. Logist. Quart. 22, pp. 181-187, 1975.