

where R is $m \times m$ upper triangular and Q^T is $n \times n$ orthogonal, we have

$$x = Q_1^T R^{-T} y.$$

(Q_1 consists of the first m rows of Q .)

Using a direct error analysis approach similar to that used in Section 2 the calculated solution may be written,

$$\bar{x} = Q_1^T \{\bar{R}^{-T} y + e_2\} + e_1.$$

Expanding, we have,

$$\begin{aligned} \bar{x} &= (\bar{Q}_1^T - Q_1^T) \{\bar{R}^{-T} y + e_2\} + Q_1^T \{\bar{R}^{-T} y + e_2\} + e_1 \\ &= (\bar{Q}_1^T - Q_1^T) \{\bar{R}^{-T} y + e_2\} + Q_1^T Q_1' x - Q_1^T \bar{R}^{-T} \delta A' x + Q_1^T e_2 + e_1 \end{aligned}$$

where

$$\delta A' = \bar{R}^T Q_1' - A.$$

Also

$$Q_1^T Q_1' = Q_1^T Q_1 + \delta Q_1^T Q_1 + Q_1^T \delta Q_1' + \delta Q_1^T \delta Q_1'.$$

hence

$$\begin{aligned} \bar{x} - x &= (Q_1^T - Q_1^T) \{\bar{R}^{-T} y + e_2\} + \{\delta Q_1^T Q_1 + Q_1^T \delta Q_1' - Q_1^T \bar{R}^{-T} \delta A'\} x \\ &\quad + Q_1^T e_2 + e_1 + \delta Q_1^T \delta Q_1' x. \end{aligned} \quad (4.1)$$

We can now estimate $\|\bar{x} - x\|$ from (4.1) by applying the inequalities (1.1), (1.2), and (2.8). The dependence that interests us is that on the condition number of A , and provided $\text{eps}\chi(A)$ is small (say < 0.1), this dependence has the form,

$$\|\bar{x} - x\| \leq k \text{eps}\chi(A), \quad (4.2)$$

as the terms in higher powers of $\text{eps}\chi(A)$ can then be submerged into a term of this form. As $\chi(A) \geq 1$, the terms just proportional to eps can also be accounted for in this way.

Acknowledgments. The authors would like to thank Dr. C. Paige of McGill University for helpful discussion, and a referee for helpful and constructive suggestions.

References

1. Björck, A.: Solving Linear Least Squares Problems by Gram-Schmidt Orthogonalization. BIT 7, 1-21 (1967)
2. Björck, A.: Iterative Refinement of Linear Least Squares Solutions I. BIT 7, 257-278 (1967)
3. Golub, G. H., Wilkinson, J. H.: Iterative Refinement of Least Squares Solutions. Num. Math. 9, 139-148 (1966)
4. Wilkinson, J. H.: Error Analysis of Transformations Based on the Use of Matrices of the Form $I - 2\omega\omega^H$. In: Error in Digital Computation, Vol. 2, Editor L. B. Rall, 1965

L. S. Jennings
M. R. Osborne
Computer Centre
Australian National University
P.O. Box 4
Canberra, A.C.T. 2600
Australia

Approximations to Stochastic Programs with Complete Fixed Recourse

Peter Kall

Received April 29, 1974

Abstract. The probability distribution of the data entering a recourse problem is replaced by finite discrete distributions. It is proved that the convergence of the objective functions of the approximating problems to that one of the original problem can be achieved by choosing the discrete distributions in quite a natural way. For bounded feasible sets this implies the convergence of the optimal values. Finally some error bounds are derived.

I. Introduction

We are concerned with stochastic programs with complete fixed recourse, which may be written as

$$\min \{c'x + \psi(x) \mid x \in X\}, \quad (1)$$

where $X \subset \mathbb{R}^n$ is a closed convex polyhedral set, $c \in \mathbb{R}^n$ is a constant vector with transpose c' and

$$\psi(x) = \int_{\Omega} Q(x, A(\omega), b(\omega), q(\omega)) dP. \quad (2)$$

$A(\omega)$, $b(\omega)$ and $q(\omega)$ are an $(m \times n)$ -matrix, an m -vector and an n -vector, respectively, the elements of which are random variables defined on a probability space $(\Omega; \mathcal{F}, P)$. Finally

$$Q(x, A, b, q) = \min \{q'y \mid y \in \mathbb{R}^{\bar{n}}, Wy = b - Ax, y \geq 0\}, \quad (3)$$

where W is a constant $(m \times \bar{n})$ -matrix (fixed recourse matrix).

We make the following assumptions:

$$A.1) \{z \mid \exists y \in \mathbb{R}^{\bar{n}}: y \geq 0, Wy = z\} = \mathbb{R}^m.$$

$$A.2) \forall \omega \in \Omega: \{u \mid u \in \mathbb{R}^m, W'u \leq q(\omega)\} \neq \emptyset.$$

A.3) The elements of $A(\omega)$, $b(\omega)$, $q(\omega)$ are square integrable with respect to P .

Assumption A.1) assures the so-called complete fixed recourse, which means that every violation of the "desired" constraints $b(\omega) - A(\omega)x = 0$, which is caused by choosing $x \in X$ before knowing the realization $\omega \in \Omega$, can be compensated by the recourse program (second stage or emergency program) stated in (3). Necessary and sufficient conditions for A.1) were derived in [2].

Assumption A.2) then guarantees that—according to the duality theorem of linear programming— $Q(x, A(\omega), b(\omega), q(\omega))$ is always finite. If the practical meaning of $q_i(\omega)$ is "penalty costs per unit of y_i ", i.e. if $q(\omega) \geq 0$ for all $\omega \in \Omega$,

then A.2) is obviously satisfied. Otherwise A.2) defines a convex polyhedral cone K_W , which has to contain $q(\omega)$ for all $\omega \in \Omega$. Assumption A.3) asserts the existence of $\psi(x)$ according to (2).

We refer to the following known results ([2], [5]):

Theorem 1. For $q \in K_W$, a) $Q(x, A, b, q)$ is continuous;

b) $Q(x, A, b, q)$ is convex in (A, b) for fixed (x, q) and concave in q for fixed (x, A, b) ;

c) $Q(x, A, b, q)$ is convex in x for fixed (A, b, q) .

Theorem 2. a) $\psi(x): \mathbb{R}^n \rightarrow \mathbb{R}^1$ is finite and convex;

b) $\psi(x)$ is Lipschitz continuous;

c) if the probability distribution of $(A(\omega), b(\omega), q(\omega))$ is given by a density function, then $\psi(x)$ is continuously differentiable.

Theorem 3. For a finite discrete probability distribution, i.e. $\omega = \omega_i$ with probability $p_i > 0$, $i = 1, \dots, \kappa$, and $\sum_{i=1}^{\kappa} p_i = 1$, problem (1) becomes a linear program with (dual) decomposition structure, namely:

$$\min \left[c'x + \sum_{i=1}^{\kappa} p_i q'(\omega_i) y^{(i)} \right] \tag{4}$$

subject to

$$\begin{aligned} A(\omega_i)x + W y^{(i)} &= b(\omega_i), \quad i = 1, \dots, \kappa \\ x \in X; \quad y^{(i)} &\geq 0, \quad i = 1, \dots, \kappa. \end{aligned}$$

Theorem 2 seems to suggest the application of some usual optimization method, if a continuous type distribution is given. However, the repeated evaluation of $\psi(x)$ or its gradient involves the repeated numerical quadrature of multiple integrals, where the integration sets are polyhedral sets in $\mathbb{R}^{m \times n + m + \bar{n}}$ depending on x , which seems to be a rather difficult task due to the present state of numerical quadrature.

On the other hand Theorem 3 suggests to approximate the given probability measure P by finite discrete measures P_v or, equivalently, to approximate $(A(\omega), b(\omega), q(\omega))$ by (vector valued) integrable simple functions $(A_v(\omega), b_v(\omega), q_v(\omega))$. As a matter of fact, this approach was already proposed in [7] for a special simple recourse problem, namely for $W = (I, -I)$, where I is the $(m \times m)$ -identity matrix, and for constant $A(\omega), b(\omega)$. And in spite of the fact, that these approximating problems of type (4) become really large scale linear programs, just the recent publications [1, 4, 6] indicate, that one can take advantage of some special properties of these problems.

Our aim is to show, that this approximation is meaningful, i.e. that the convergence of the approximating objective functions to the original one always can be achieved by a rather natural choice of the integrable simple functions, and to derive error bounds.

II. Convergence

Let $(A_v(\omega), b_v(\omega), q_v(\omega))$, $v = 1, 2, \dots$, be a sequence of integrable simple functions such that

$$(A_v(\omega), b_v(\omega), q_v(\omega)) \xrightarrow{v \rightarrow \infty} (A(\omega), b(\omega), q(\omega)) \tag{5}$$

pointwise on Ω ,

$$\|A_v(\omega)\|_{\infty} \leq \|A(\omega)\|_{\infty}, \quad \|b_v(\omega)\|_{\infty} \leq \|b(\omega)\|_{\infty}, \quad \|q_v(\omega)\|_{\infty} \leq \|q(\omega)\|_{\infty} \tag{6}$$

for all $\omega \in \Omega$ ($\|\dots\|_{\infty}$ means for vectors the maximum norm and for matrices the corresponding matrix norm),

assumption A.2) is also valid for $q_v(\omega)$, $v = 1, 2, \dots$. (7)

Taking one of the usual partitions of $\mathbb{R}^{m \times n + m + \bar{n}}$, the elements of which are halfopen intervals contained in one of the orthants, and choosing the norm minimal vertex of every interval with its probability, one can obviously achieve the requirements (5) and (6). But to assure (7), we have to be careful. If for example

$$W = (1, -1)$$

and $q(\omega)$ has the range $R(q) = \{(\xi, \eta) | \xi \geq -2, \eta \geq 2, \eta \leq 5\}$, then A.1) and A.2) are satisfied for the original problem. Now let $M = \{(\xi, \eta) | -4 < \xi \leq -2, 1 \leq \eta < 3\}$ be an interval of some partition. Then $q^{-1}[M] \neq \emptyset$, such that M could have a positive probability. Choosing on M the norm minimal vertex $v = \{(\xi, \eta) | \xi = -2, \eta = 1\}$ as value of $q_v(\omega)$ does not satisfy A.2), since $W'u \leq q_v(\omega)$ yields $-1 \leq u \leq -2$. But if we choose the norm minimal element of the intersection of M and

$$\begin{aligned} K_W &= \{(\xi, \eta) | \exists u: -\eta \leq u \leq \xi\} \\ &= \{(\xi, \eta) | \xi + \eta \geq u\}, \end{aligned}$$

i.e. $q_v(\omega) = (-2, -2)$, then A.2) is satisfied. In general, the analogous way (choosing the norm minimal element of the intersection of every interval and K_W) yields a sequence satisfying (7) too.

For $\psi_v(x) = \int_{\Omega} Q(x, A_v(\omega), b_v(\omega), q_v(\omega)) dP$ we can prove

Theorem 4. Provided (5), (6) and (7), $\lim_{v \rightarrow \infty} \psi_v(x) = \psi(x)$.

Proof. $Q(x, A_v(\omega), b_v(\omega), q_v(\omega))$ is finite on Ω due to (7). From (5) and Th. 1 a) follows

$$Q(x, A_v(\omega), b_v(\omega), q_v(\omega)) \xrightarrow{v \rightarrow \infty} Q(x, A(\omega), b(\omega), q(\omega)) \tag{8}$$

pointwise on Ω .

For any fixed x, ω and v we have

$$Q(x, A_v(\omega), b_v(\omega), q_v(\omega)) = \tilde{q}'_{v,i}(\omega) B_i^{-1} [b_v(\omega) - A_v(\omega)x],$$

where B_i is an optimal feasible basis out of W and $\tilde{q}_{v,i}(\omega)$ consists of the corresponding components of $q_v(\omega)$. With the Euclidean norm $\|\dots\|$ follows from (6)

$$\begin{aligned} |Q(x, A_v(\omega), b_v(\omega), q_v(\omega))| &\leq \|\tilde{q}_{v,i}(\omega)\| \|B_i^{-1}\| \|b_v(\omega) - A_v(\omega)x\| \\ &\leq \beta \|q_v(\omega)\| \{\|b_v(\omega)\| + \|A_v(\omega)x\|\} \\ &\leq \beta \sqrt{\bar{n} \cdot \bar{m}} \|q_v(\omega)\|_\infty \{\|b_v(\omega)\|_\infty + \|A_v(\omega)\|_\infty \cdot \|x\|_\infty\} \\ &\leq \beta \sqrt{\bar{n} \cdot \bar{m}} \|q(\omega)\|_\infty \{\|b(\omega)\|_\infty + \|A(\omega)\|_\infty \cdot \|x\|_\infty\} \\ &= C(x, \omega), \end{aligned} \tag{9}$$

where $\beta = \max_i \|B_i^{-1}\|$.

According to A.3) and Schwarz' inequality $\int_\Omega C(x, \omega) dP$ exists. From (8), (9) and Lebesgue's bounded convergence theorem follows the proposition. q.e.d.

From Theorem 4 we get the connection between solving the approximating problems and the original one as a special application of a more general theorem proved in [3]:

Theorem 5. Assume that the set of optimal solutions of (1) is nonempty and bounded. Then

- a) $\inf \{c'x + \psi_v(x) \mid x \in X\} \xrightarrow{v \rightarrow \infty} \inf \{c'x + \psi(x) \mid x \in X\}$;
- b) if x_v solves $\min \{c'x + \psi_v(x) \mid x \in X\}$, $v = 1, 2, \dots$, then every convergent subsequence of $\{x_v \mid v = 1, 2, \dots\}$ converges to a solution of (1).

The boundedness assumption in this theorem may not be omitted in general, as was demonstrated in [3] by examples. However, in practical problems it is not very restrictive to assume that the feasible set X is bounded.

III. Error Bounds

Let $(A_v(\omega), b_v(\omega), q_v(\omega))$ be an integrable simple function such that A.2) is satisfied. We want to have an error estimate for the objective function and hence for the optimal value of the approximating problem $\min \{c'x + \psi_v(x) \mid x \in X\}$, which depends on the approximation of $(A(\omega), b(\omega), q(\omega))$ by $(A_v(\omega), b_v(\omega), q_v(\omega))$, measured by the (generalized) L_2 -norm. For any vector valued function

$$g: \Omega \rightarrow \mathbb{R}^k$$

we define

$$\varrho(g) = \sqrt{\int_\Omega \|g(\omega)\|^2 dP}, \tag{10}$$

where $\|\dots\|$ is the Euclidean norm in \mathbb{R}^k . In this connection $\varrho(A)$ means that the matrix $A(\omega)$ is handled as an $(m \cdot n)$ -vector.

Theorem 6. There are constants α, γ, δ_v such that

$$|\psi(x) - \psi_v(x)| \leq [\alpha + \gamma \|x\|] \varrho(q - q_v) + \delta_v [\varrho(b - b_v) + \|x\| \varrho(A - A_v)].$$

Proof. For every convex or concave function

$$\begin{aligned} \varphi: \mathbb{R}^l \rightarrow \mathbb{R}^1 \text{ holds} \\ |\varphi(x) - \varphi(y)| &\leq \text{Max} [|(x - y)' \nabla \varphi(x)|; |(x - y)' \nabla \varphi(y)|] \\ &\leq \text{Max} [\|\nabla \varphi(x)\| \cdot \|x - y\|; \|\nabla \varphi(y)\| \cdot \|x - y\|], \end{aligned} \tag{11}$$

where $\nabla \varphi$ is the gradient (or some subgradient) of φ . Using basic solutions, Theorem 1b) and (10), (11), we get

$$\begin{aligned} |\Delta Q_v(x, \omega)| &= |Q(x, A(\omega), b(\omega), q(\omega)) - Q(x, A_v(\omega), b_v(\omega), q_v(\omega))| \\ &\leq |Q(x, A(\omega), b(\omega), q(\omega)) - Q(x, A(\omega), b(\omega), q_v(\omega))| \\ &\quad + |Q(x, A(\omega), b(\omega), q_v(\omega)) - Q(x, A_v(\omega), b_v(\omega), q_v(\omega))| \\ &= |\tilde{q}'_i(\omega) B_i^{-1} [b(\omega) - A(\omega)x] - \tilde{q}'_{v,i}(\omega) B_i^{-1} [b(\omega) - A(\omega)x]| \\ &\quad + |\tilde{q}'_{v,i}(\omega) B_i^{-1} [b(\omega) - A(\omega)x] - \tilde{q}'_{v,k}(\omega) B_k^{-1} [b_v(\omega) - A_v(\omega)x]| \\ &\leq \text{Max}_{i \in J_1} \|B_i^{-1} [b(\omega) - A(\omega)x]\| \cdot \|q(\omega) - q_v(\omega)\| \\ &\quad + \text{Max}_{i \in J_2} \|B_i^{-1} \tilde{q}_{v,i}(\omega)\| \cdot \|b(\omega) - b_v(\omega) - [A(\omega) - A_v(\omega)]x\|, \end{aligned} \tag{12}$$

where $\{B_i \mid i \in J_1\}$ and $\{B_i \mid i \in J_2\}$ are those bases out of W , which for some $x \in X$ and $\omega \in \Omega$ are primal feasible and dual feasible respectively.

From (12) and Schwarz' inequality follows with

$$\begin{aligned} z(x, \omega) &= \begin{cases} 0 & \text{if } b(\omega) - A(\omega)x = 0 \\ \frac{b(\omega) - A(\omega)x}{\|b(\omega) - A(\omega)x\|} & \text{else} \end{cases} \\ \text{and } r_i(\omega) &= \begin{cases} 0 & \text{if } q_v(\omega) = 0 \\ \frac{\tilde{q}_{v,i}(\omega)}{\|q_v(\omega)\|} & \text{else} \end{cases} \end{aligned}$$

$$\begin{aligned} |\psi(x) - \psi_v(x)| &\leq \int_\Omega |\Delta Q_v(x, \omega)| dP \\ &\leq \text{Max}_{i \in J_1} \|B_i^{-1} \cdot z(x, \omega)\| \cdot \varrho(b - A x) \cdot \varrho(q - q_v) \\ &\quad + \text{Max}_{i \in J_2} \|B_i^{-1} r_i(\omega)\| \cdot \varrho(q_v) \cdot \varrho(b - b_v + (A_v - A)x) \end{aligned} \tag{13}$$

Hence, since $\varrho(g + h) \leq \varrho(g) + \varrho(h)$, for

$$\begin{aligned} \alpha &= \text{Max}_{i \in J_1} \|B_i^{-1} z(x, \omega)\| \varrho(b) \\ \gamma &= \text{Max}_{i \in J_1} \|B_i^{-1} z(x, \omega)\| \varrho(A) \\ \delta_v &= \text{Max}_{i \in J_2} \|B_i^{-1} r_i(\omega)\| \varrho(q_v) \end{aligned}$$

from (13) follows

$$|\psi(x) - \psi_r(x)| \leq [\alpha + \gamma \|x\|] \varrho(q - q_r) + \delta_r [\varrho(b - b_r) + \|x\| \varrho(A - A_r)]. \quad \text{q.e.d.} \tag{14}$$

It must be mentioned that determining the constants α and γ leads in general to a considerable amount of work, since it implies more or less the inversion of all nonsingular $(m \times m)$ -submatrices of W . This difficulty diminishes rapidly for certain special cases. Determining δ_r (or at least an upper bound) is not difficult, since $B_i^{-1}r_i(\omega)$ is a feasible u -part of the set $\{(u, q) | W'u \leq q, \|q\| \leq 1\}$, which is bounded according to A.1).

Corollary 1. Assume simple recourse, i.e. $W = (I, -I)$. Then

$$|\psi(x) - \psi_r(x)| \leq [\varrho(b) + \|x\| \varrho(A)] \varrho(q - q_r) + \varrho(q_r) [\varrho(b - b_r) + \|x\| \varrho(A - A_r)].$$

Proof. $W = (I, -I)$ implies that for every basis B_i out of $W \|B_i^{-1}z(x, \omega)\| = \|z(x, \omega)\|$ and, by definition, $\|z(x, \omega)\| \leq 1$. And for every $i \in J_2$, $B_i^{-1}r_i(\omega)$ is a feasible u -part of $\{(u, q) | W'u \leq q, \|q\| \leq 1\}$, which for

$$q = \begin{pmatrix} q^+ \\ q^- \end{pmatrix} = \begin{pmatrix} q_1^+ \\ \vdots \\ q_m^+ \\ q_1^- \\ \vdots \\ q_m^- \end{pmatrix} \text{ may be written as}$$

$$\{(u, q^+, q^-) | -q^- \leq u \leq q^+; \|q^+\|^2 + \|q^-\|^2 \leq 1\}.$$

Obviously every feasible u satisfies $\|u\| \leq 1$. Using these observations, we get for the constants defined in the proof of Theorem 6

$$\begin{aligned} \alpha &\leq \varrho(b) \\ \gamma &\leq \varrho(A) \\ \delta_r &\leq \varrho(q_r). \quad \text{q.e.d.} \end{aligned}$$

For the general complete recourse problem we get out of the difficulties of determining α and γ , if $q(\omega)$ is constant, as is readily seen in Theorem 6. Moreover we can weaken assumption A.3) to integrability instead of square integrability. Defining a generalized L_1 -norm for vector valued functions as

$$\mu(g) = \int_{\Omega} \|g(\omega)\| dP, \tag{15}$$

we get

Theorem 7. For constant $q(\omega) \equiv q$ there is a constant δ such that

$$|\psi(x) - \psi_r(x)| \leq \delta [\mu(b - b_r) + \|x\| \mu(A - A_r)].$$

Proof. From (12) we see that

$$|AQ(x, \omega)| \leq \text{Max}_{i \in J_1} \|B_i^{-1}q_i\| \cdot \|b(\omega) - b_r(\omega) - [A(\omega) - A_r(\omega)]x\|$$

and therefore

$$|\psi(x) - \psi_r(x)| \leq \text{Max}_{i \in J_1} \|B_i^{-1}q_i\| [\mu(b - b_r) + \|x\| \mu(A - A_r)].$$

For $i \in J_2$, $B_i^{-1}q_i$ is feasible in $\{u | W'u \leq q\}$, which is bounded. q.e.d.

Every one of the error estimates above becomes independent of x , if $A(\omega) = A$ and $q(\omega) \equiv q$ are constant, i.e. in this case we get uniform convergence of $\{\psi_r(x)\}$ on X , if the simple functions chosen converge to the remaining random variables with respect to the appropriate norm.

IV. Final Remarks

One might object, that this type of approximation is not practicable, since the size of the approximating problems becomes very large. If for example in (A, b, q) 50 random variables with a joint probability distribution are involved, and if we discretize in such a way that for every random variable 10 realizations occur, then we should have in (4) $m \times 10^{50}$ constraints, which cannot be handled. However, in most of the practical problems there is a small number of random variables t_1, t_2, \dots, t_r , where often $r \leq 5$, and (A, b, q) depend on t_1, \dots, t_r . If this dependence looks like

$$\begin{aligned} A(t) &= A_0 + A_1 t_1 + \dots + A_r t_r \\ b(t) &= b_0 + b_1 t_1 + \dots + b_r t_r \\ q(t) &= q_0 + q_1 t_1 + \dots + q_r t_r \end{aligned}$$

then the discretization can be carried through with respect to the random vector t yielding problems of a size which can be handled, to-day.

It is obvious, that all above mentioned error estimates then can be expressed in $\varrho(t - t_{(r)})$ or $\mu(t - t_{(r)})$, where $t_{(r)}$ is a simple function.

References

1. Garstka, S. J., Rutenberg, D. P.: Computation in discrete stochastic programs with recourse. *Operations Research* **21**, 112-122 (1973)
2. Kall, P.: Qualitative Aussagen zu einigen Problemen der stochastischen Programmierung. *Z. Wahrscheinlichkeitstheorie* **6**, 246-272 (1966)
3. Kosmol, P.: Algorithmen zur konvexen Optimierung. *OR-Verfahren*, Band XVIII 176-186 (1974)
4. Strazicky, B.: On an algorithm for solution of the two-stage stochastic programming problem. *OR-Verfahren*, Band XVIII (to appear).
5. Walkup, D. W., Wets, R. J. B.: Stochastic programs with recourse. *SIAM J. Appl. Math.* **15**, 1299-1314 (1967)
6. Werner, M.: Ein Lösungsansatz für ein spezielles zweistufiges stochastisches Optimierungsproblem. *Z. Operations Research* **17**, 119-128 (1973)
7. Wets, R.: Programming under uncertainty: The complete problem. *Z. Wahrscheinlichkeitstheorie* **4**, 316-339 (1966)

Peter Kall
 Institut für Operations Research und
 mathematische Methoden der Wirtschaftswissenschaften
 der Universität Zürich
 Weinbergstr. 59
 CH-8006 Zürich
 Schweiz