Mathematical Programming 83 (1998) 451–464

L-shaped decomposition of two-stage stochastic programs with integer recourse

Claus C. Carøe *, Jørgen Tind

Department of Operations Research, University of Copenhagen, Universitetsparken 5, DK-2100, Copenhagen Ø, Denmark

Received 11 September 1995; revised manuscript received 17 February 1997

Abstract

We consider two-stage stochastic programming problems with integer recourse. The L-shaped method of stochastic linear programming is generalized to these problems by using generalized Benders decomposition. Nonlinear feasibility and optimality cuts are determined via general duality theory and can be generated when the second stage problem is solved by standard techniques. Finite convergence of the method is established when Gomory's fractional cutting plane algorithm or a branch-and-bound algorithm is applied. © 1998 The Mathematical Programming Society, Inc. Published by Elsevier Science B.V.

Keywords: Stochastic programming; Integer programming; Benders decomposition; General duality theory

1. Introduction

A stochastic programming problem arises when parameters of a deterministic mathematical programming problem are replaced by random variables. A common way of modelling uncertainty in mathematical programs is via a two-stage stochastic program where a long-term anticipatory decision must be made prior to full information about random parameters of the problem and short-term decisions are available as recourse actions once the uncertainty has been revealed. The aim is to determine a here-and-now decision which minimizes the total expected costs associated with both the long-term and the short-term decisions.

Stochastic programming problems are well known for being challenging both from theoretical and computational points of view. Stochastic programs with integer recourse are problems for which the recourse decision is required to be integer. Adding integrality restrictions to the constraints that define the second stage problem

* Corresponding author. Fax: +45 35320678; e-mail: caroe@math.ku.dk.

PII S0025-5610(98)00007-0

significantly increases the complexity of the problem and only few results have been obtained for this type of problem. The reason for this is that given any first stage decision and outcome of the random variable, the resulting decision problem is an integer programming problem which is NP-hard, and together with a large number of outcomes of the random variables this makes the problem very difficult to solve. For an overview of stochastic integer programming we refer to [1,2]. In [3], a branch-and-cut algorithm called the integer L-shaped method is developed for problems with binary first stage variables and arbitrary second stage variables. In this paper we propose a general framework for L-shaped decomposition of stochastic programs with integer recourse. The method is based on general Benders (resource) decomposition and general duality theory, see [4,5], and the integer L-shaped algorithm of Laporte and Louveaux [3] appears as a special case of this framework. A main issue in stochastic programming is whether first stage decisions give rise to infeasible second stage problems. This is also handled by the method.

The paper is divided in sections as follows. In Section 2 we give a mathematical description of the problem and collect some results from integer programming duality which are needed in subsequent sections. In Section 3 we state the L-shaped decomposition method and Section 4 deals with the case where cutting plane procedures are applied to solve the second stage problem. Use of branch-and-bound techniques is discussed in Section 5 and Section 6 contains some final remarks.

2. Two-stage stochastic programs with integer recourse

We consider the following problem:

$$\begin{array}{ll} \min & cx + E_\xi \min\{qy \mid T(\xi)x + Wy \geqslant h(\xi),\ y \in \mathbb{Z}_+^{n_2}\} \\ \text{s.t.} & Ax \geqslant b,\ x \in \mathbb{R}_+^{n_1}, \end{array} \tag{2.1}$$

where c is a known n_1-vector, q a known n_2-vector, b a known m_1-vector and A and W are known matrices of sizes $m_1 \times n_1$ and $m_2 \times n_2$, respectively. Furthermore, ξ is a random variable having support $\Xi \subset \mathbb{R}^k$ and for each ξ the variables $T(\xi)$ and $h(\xi)$ have conformable sizes. Transposes have been eliminated for simplicity. We suppose throughout that W and q are rational. The part of the objective and the constraints that are only related to the first stage decision variable x constitute here a linear programming problem. This formulation has been selected for simplicity of exposition. In particular the first stage decision variable could be restricted to be integers as well. For later use we define $X := \{x \in \mathbb{R}_+^{n_1} \mid Ax \geqslant b\}$.

The challenge of (2.1) lies in the multivariate integration of a function, which is only given implicitly as the value function of a parametric integer programming problem. In many integer programming problems with an underlying combinatorial structure the uncertainty must necessarily be of discrete nature. A continuous probability distribution on the random variable ξ will cause severe difficulties when evaluating the integral in (2.1). Moreover, in [6] it is shown that if ξ follows a continuous distribution, solutions to (2.1) can under mild conditions be approximated within any given accuracy by discrete distributions. This motivates the following assumption:

(A1) The random variable ξ has a discrete distribution with finite support, say $\Xi = \{\xi^1, \dots, \xi^r\}$ and $P(\xi = \xi^j) = p^j$.

Problem (2.1) is then equivalent to the following mixed-integer program, where the constraints have a dual blockangular structure or L-shaped form,

$$\begin{aligned} \min \quad & cx + \sum_{j=1}^{r} p^j q y^j \\ \text{s.t.} \quad & Ax \geqslant b, \quad T(\xi^j)x + Wy^j \geqslant h(\xi^j), \; j = 1, \dots, r, \quad x \in \mathbb{R}_+^{n_1}, \; y^j \in \mathbb{Z}_+^{n_2}. \end{aligned} \tag{2.2}$$

The decision variables y^j and constraints are very large in number in this formulation but since we are primarily interested in the optimal first stage decision, we rewrite problem (2.2) in terms of the first stage variables only,

$$\min\{cx + Q(x) \mid x \in X\}, \tag{2.3}$$

where the *expected recourse* function Q is defined by

$$Q(x) := E_\xi \Phi(h(\xi) - T(\xi)x) = \sum_{j=1}^{r} p^j \Phi(h(\xi^j) - T(\xi^j)x)$$

and Φ is the value function of the second stage problem,

$$\Phi(d) = \min\{qy \mid Wy \geqslant d, \; y \in \mathbb{Z}_+^{n_2}\}, \quad d \in \mathbb{R}^{m_2}. \tag{2.4}$$

In order to have (2.3) well defined we make the following standard assumption:

(A2) There exists a $u \in \mathbb{R}_+^{m_2}$ such that $uW \leqslant q$.

Indeed, this assumption implies $\Phi(d) > -\infty$ for all $d \in \mathbb{R}^{m_2}$ and hence $Q(x) > -\infty$. Assumption (A2) is independent of ξ since W and q are fixed. However, no assumptions on primal feasibility will be made in this paper. Thus it might happen that $Q(x) = +\infty$ for some $x \in X$. Particular attention should therefore be devoted to first stage decisions that satisfy the induced constraint $x \in K$, where $K = \{x \in \mathbb{R}_+^{n_1} \mid Q(x) < \infty\}$. One is therefore lead to study the properties of the value function Φ. The function Φ is nondecreasing and subadditive on its domain of definition, i.e., $\Phi(d^1) + \Phi(d^2) \geqslant \Phi(d^1 + d^2)$ for $d^i \in \mathbb{R}^{m_2}$ with $\Phi(d^i) < \infty$, $i = 1, 2$, see [7]. For a full characterization we shall consider the following two classes of Chvátal and Gomory functions, respectively. For a function F the integer round-up $\lceil F \rceil$ is defined by $\lceil F \rceil(d) = \lceil F(d) \rceil$, where $\lceil F(d) \rceil$ is the smallest integer which is larger than $F(d)$.

Definition 2.1. The class $\mathscr{C}^m$ of m-dimensional *Chvátal-functions* is the smallest class $\mathscr{H}$ of functions satisfying (i) $F \in \mathscr{H}$ if $F(d) = \lambda d$ and $\lambda \in \mathbb{Q}^m$, (ii) $F, G \in \mathscr{H}$ and α, β nonnegative rationals implies $\alpha F + \beta G \in \mathscr{H}$, (iii) $F \in \mathscr{H}$ implies $\lceil F \rceil \in \mathscr{H}$.

The class $\mathscr{C}$ of all Chvátal functions is $\mathscr{C} = \cup_{m \in \mathbb{N}} \mathscr{C}^m$.

The class $\mathscr{G}^m$ of m-dimensional *Gomory-functions* is the smallest class $\mathscr{H}$ of functions satisfying (i)–(iii) and (iv) $F, G \in \mathscr{H}$ implies $\max\{F, G\} \in \mathscr{H}$.

The class $\mathscr{G}$ of all Gomory-functions is $\mathscr{G} = \cup_{m\in\mathbb{N}}\mathscr{G}^m$.

By definition, a function $F : \mathbb{R}^{m_2} \to \mathbb{R}$ is a Chvátal-function if it can be written as $F(d) = u\lceil M_t\lceil M_{t-1}\cdots\lceil M_2\lceil M_1 d\rceil\rceil\cdots\rceil\rceil$, where $M_t, \ldots, M_1$ are nonnegative rational matrices of conformable sizes and u is a nonnegative rational vector. It is easily proven that every Gomory-function is the maximum of a finite number of Chvátal-functions.

Theorem 2.2 (Blair and Jeroslow [8]). *Let $W = (w_1, \ldots, w_{n_2})$ be a rational matrix and q a rational vector. Assuming* (A2) *there exists Gomory-functions $F, G : \mathbb{Q}^{m_2} \to \mathbb{R}$ with $G(w_j) \leqslant 0, F(w_j) \leqslant q_j$, $j = 1, \ldots, n_2$, such that for each $d \in \mathbb{R}^{m_2}$,*

(i) *$\{y \in \mathbb{Z}^{n_2}_+ \mid Wy \geqslant d\} \neq \emptyset$ if and only if $G(d) \leqslant 0$.*

(ii) *If $G(d) \leqslant 0$ then $F(d) = \min\{qy \mid Wy \geqslant d,\ y \in \mathbb{Z}^{n_2}_+\}$.*

In other words, feasibility of the second stage problem is determined by the *consistency tester* G and when consistent the value function is a Gomory function. Note that except for the rounding-up operation, this is in exact analogy with linear programming. The value function is on its domain of definition the maximum of a finite number of Chvátal functions. The Chvátal functions constituting the value function can be thought of as the "slopes" or "facets" of the value function. The expected recourse function is thus nonconvex and discontinuous. By Fatous Lemma it is lower semicontinuous even for arbitrary probability distributions.

The idea in the L-shaped method for stochastic linear programming [9] is to use dual information of the second stage program to represent the recourse function, which is known to be a polyhedral function. In the integer case the recourse function is nonconvex and cannot be supported by hyperplanes. Due to the integrality gap "nonlinear dual variables" must be considered, which in Theorem 2.2 take the form of Chvátal functions. This is formalized in the setting of general duality theory. The presentation here will be adapted to our setting and we refer to [4] for a general survey. A more thorough discussion of duality in integer programming is given in Wolsey [10]. We denote the extended real line by $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ and consider the set $\overline{\mathscr{F}}$ of all functions $F : \mathbb{R}^{m_2} \to \overline{\mathbb{R}}$ that satisfy $F(0) = 0$ and are nondecreasing, i.e., $F(d^1) \leqslant F(d^2)$ whenever $d^1_i \leqslant d^2_i$, $i = 1, \ldots, m_2$. The set $\overline{\mathscr{F}}$ is the set of *dual price functions*. Consider also a subset $\mathscr{F}$ of $\overline{\mathscr{F}}$ together with the maximization problem

$$\begin{aligned} w = \ & \max_F \quad F(d) \\ & \text{s.t.} \quad F(Wy) \leqslant qy \ \text{ for all } y \in \mathbb{Z}^{n_2}_+, \\ & \qquad\ \ F \in \mathscr{F}. \end{aligned} \tag{2.5}$$

Problem (2.5) is the ($\mathscr{F}$-)dual of (2.4). By construction it follows that $w \leqslant \Phi(d)$, since y feasible in (2.4) and F dual feasible implies $F(d) \leqslant F(Wy) \leqslant qy$. Moreover, we have the following version of the Duality Theorem and Farkas' Lemma [4,11]:

Theorem 2.3. *If the function class $\mathscr{F}$ is suitably large then* (2.4) *is infeasible if and only if there exists $\hat{G} \in \mathscr{F}$ with $\hat{G}(Wy) \leqslant 0$ for all $y \in \mathbb{Z}_+^{n_2}$ and $\hat{G}(d) > 0$. The function $\hat{G}$ is then called a dual ray. If* (2.4) *is feasible, then $\hat{y}$ is optimal in* (2.4) *if and only if there exists $\hat{F} \in \mathscr{F}$ feasible in* (2.5) *such that $q\hat{y} = \hat{F}(d)$.*

The meaning of *suitably large* is that the duality gap between (2.4) and (2.5) should be closed. This is always the case if $\Phi \in \mathscr{F}$ because then Φ is feasible in (2.5). Thus Theorem 2.3 is of practical value only if it is possible to employ price functions that are more simple than the value function itself. One such class of functions is given by the Chvátal functions, cf. Theorem 2.2. Also note that due to the nonconvexity of the primal problem, we will have to work with nonlinear price functions. The selection of a particular class of functions is determined by the algorithmic context. Below we shall be concerned with the choice of $\mathscr{F}$ when (2.4) is solved by cutting plane technique or branch-and-bound, respectively.

3. Generalized L-shaped decomposition

We now give a procedure for solving stochastic programming problems with integer recourse. The idea is to rewrite (2.3) in the following way:

$$\min\{cx + \theta \mid \theta \geqslant Q(x),\ x \in X\} \tag{3.1}$$

and then represent the constraint $\theta \geqslant Q(x)$ by means of dual price functions. For each outcome $\xi^j \in \Xi$ we have a second stage problem,

$$\min\{qy \mid Wy \geqslant h(\xi^j) - T(\xi^j)x,\ y \in \mathbb{Z}_+^{n_2}\}, \tag{3.2}$$

and an associated dual problem,

$$\max_F\{F(h(\xi^j) - T(\xi^j)x) \mid F(Wy) \leqslant qy \text{ for all } y \in \mathbb{Z}_+^{n_2}, F \in \mathscr{F}\}. \tag{3.3}$$

We suppose that the algorithm applied to solve (3.2) simultaneously generates an optimal solution of (3.3) when both problems have optimal solutions. This assumption holds for all standard algorithms. Note that dual feasibility is independent of the realization of the random variable.

Definition 3.1. Let $g : \mathbb{R}^{m_2} \to \mathbb{R}$. The inequality $g(x) \leqslant 0$ is said to be a *feasibility cut* for Q at x^* if (i) $x^* \in X \setminus K$, (ii) $g(x) \leqslant 0$ for all $x \in X \cap K$, (iii) $g(x^*) > 0$.

If for some ξ^j problem (3.2) is infeasible at x^*, we add a "valid inequality" $g(x) \leqslant 0$, which is satisfied by any feasible x but violated by x^*. To detect infeasibility and generate a feasibility cut we perform the following variant of a Phase I procedure:

$$\begin{array}{ll} \text{Minimize} & v = et \\ \text{subject to} & Wy + It \geqslant h(\xi^j) - T(\xi^j)x^*, \\ & y \in \mathbb{Z}_+^{n_2}, \quad t \in \mathbb{Z}_+^{m_2}, \end{array} \tag{3.4}$$

where e is the m_2-dimensional row vector $(1, \ldots, 1)$.

Proposition 3.2. *If the optimal value of problem* (3.4) *is positive, then the inequality* $\hat{G}(h(\xi^j) - T(\xi^j)x) \leqslant 0$ *is a feasibility cut for* Q at x^*, *where* $\hat{G}$ *is an optimal dual solution of* (3.4).

Proof. The optimal value is always nonnegative and finite and x^* belongs to $X \cap K$ if and only if the optimal value is zero. Otherwise a dual price function $\hat{G}$ has been determined which satisfies $\hat{G}(h(\xi^j) - T(\xi^j)x^*) > 0$ and $\hat{G}(Wy + It) \leqslant et$ for all $y \in \mathbb{Z}_+^{n_2}$ and all $t \in \mathbb{Z}_+^{m_2}$. Letting $t = 0$ we obtain the desired property. Moreover, if $x \in X \cap K$ then $\hat{G}(h(\xi^j) - T(\xi^j)x) \leqslant 0$, since $\hat{G}$ is a feasible (but not necessarily optimal) solution of the problem

$$\max_G \{G(h(\xi^j) - T(\xi^j)x) \mid G(Wy + It) \leqslant et \ \forall (y, t) \in \mathbb{Z}_+^{n_2} \times \mathbb{Z}_+^{m_2}, \ G \in \mathscr{F}\}.$$

The function $\hat{G}$ thus has the desired properties. □

Remark 3.3. The requirement that t is integral can be replaced by $t \geqslant 0$, in which case (3.4) becomes a mixed integer programming problem. The result remains valid, but is then based on the duality theory for mixed integer programming.

Definition 3.4. Let $f : \mathbb{R}^{m_2} \to \mathbb{R}$. The inequality $f(x) \leqslant \theta$ is said to be an *optimality cut* for Q at x^* if (i) $x^* \in X \cap K$, (ii) $f(x) \leqslant Q(x)$ $\forall x \in X$, (iii) $f(x^*) = Q(x^*)$.

Proposition 3.5. *Suppose* $\hat{F}^j$, $j = 1, \ldots, r$ *are optimal dual price functions obtained by solving* (3.3) *with* $x = x^*$ *for each* $\xi^j \in \Xi$. *Then an optimality cut for* Q *at* x^* *is given by the inequality* $\theta \geqslant \sum_{j=1}^r p^j \hat{F}^j(h(\xi^j) - T(\xi^j)x)$.

Proof. For every $j \in \{1, \ldots, r\}$ optimal solutions $\hat{y}^j$ and $\hat{F}^j$ of (3.2) and (3.3) corresponding to $x = x^*$ satisfy $q\hat{y}^j = \hat{F}^j(h(\xi^j) - T(\xi^j)x^*)$ as well as $F(h(\xi^j) - T(\xi^j)x^*) \leqslant \hat{F}^j(h(\xi^j) - T(\xi^j)x^*) = q\hat{y}^j \leqslant qy$ for all feasible $y \in \mathbb{Z}_+^{n_2}$ and all feasible $F \in \mathscr{F}$. Hence for $x \in X \cap K$ and corresponding optimal dual price functions $\tilde{F}^j$ we have

$$Q(x) = \sum_{j=1}^r p^j \tilde{F}^j(h(\xi^j) - T(\xi^j)x) \geqslant \sum_{j=1}^r p^j \hat{F}^j(h(\xi^j) - T(\xi^j)x)$$

with equality for $x = x^*$. □

The procedure can be summarized as follows: At each iteration of the method we consider a *current problem*, which is a relaxation of (2.3). For each outcome ξ^j let $s(j)$

be the number of feasibility cuts added sofar and let t be the number of optimality cuts added sofar. The current problem is:

$$\begin{array}{ll} \min & cx+\theta \\ \text{s.t.} & 0 \geqslant \hat{G}_{k_j}(h(\xi^j)-T(\xi^j)x), \ \ k_j=1,\dots,s(j), \ \ j=1,\dots,r, \\ & \theta \geqslant \sum_{j=1}^{r} p^j \hat{F}_k^j(h(\xi^j)-T(\xi^j)x), \ \ k=1,\dots,t, \\ & x\in X. \end{array} \tag{3.5}$$

In multilevel planning (3.5) is referred to as the relaxed master problem or the supremal subproblem, whereas the second stage problem(s) (3.2) is referred to as the infimal subproblem(s). The LP-relaxation of (3.2) can provide linear feasibility cuts and optimality cuts for the recourse functions, in which case the current problem will have linear constraints too. In general, these cuts will not be tight.

Algorithm 1

Step 0: Set $n := t := 0$ and $s(j) := 0$ for all $j = 1,\dots,r$ and $\bar{z}^1 = \infty$ (or any upper bound on the value of (2.3)).

Step 1: Set $n := n+1$. Solve the current problem. If the current problem is infeasible, then (2.3) is infeasible; stop. Otherwise let (x^n, θ^n) be an optimal solution if it exists; if the current problem is unbounded, then let (x^n, θ^n) be a feasible solution with $cx^n + \theta^n < \bar{z}^n$.

Step 2: If $cx^n + \theta^n = \bar{z}^n$ then (x^n, θ^n) is optimal; stop. Otherwise solve (3.2) and (3.3) for all $\xi^j \in \Xi$ with $x = x^n$.

Step 3: (i) Suppose the second stage problem is infeasible for some $\xi^{j_1},\dots,\xi^{j_v}$. Then add the feasibility cuts $0 \geqslant \hat{G}^{j_i}(h(\xi^{j_i}) - T(\xi^{j_i})x)$, $i = 1,\dots,v$, to the current problem, where $\hat{G}^{j_i}$ is the dual ray obtained from the Phase I problem (3.4) with $x = x^n$ and $\xi = \xi^{j_i}$. Let $s(j_i) := s(j_i)+1$, $i = 1,\dots,v$, and $\bar{z}^{n+1} = \bar{z}^n$. Return to Step 1. (ii) Suppose the second stage problem for each $\xi^j \in \Xi$ has a finite optimum. Denote the corresponding optimal dual price functions by $\hat{F}^j$, $j = 1,\dots,r$. Add the optimality cut $\theta \geqslant \sum_{j=1}^{r} \hat{F}^j(h(\xi^j) - T(\xi^j)x)$ to the current problem and update $\bar{z}^n$ as

$$\bar{z}^{n+1} = \min\left\{\bar{z}^n,\ cx^n + \sum_{j=1}^{r} p^j \hat{F}^j(h(\xi^j) - T(\xi^j)x^n)\right\}.$$

Let $l := l+1$ and return to step 1.

Theorem 3.6. *The L-shaped algorithm does not cycle and termination in Step* 2 *implies that a global optimum has been reached.*

Proof. We first show that (x^n, θ^n) is infeasible at iteration $n+1$. If a feasibility cut $0 \geqslant \hat{G}^j(h(\xi^j) - T(\xi^j)x)$, is generated from x^n then x^n is cut off since $\hat{G}^j(h(\xi^j) - T(\xi^j)x^n) > 0$ by construction. If termination in Step 2 is implied then the upper bound $\bar{z}^n$ has been reached. Otherwise $\theta^n < \sum_{j=1}^{r} p^j\hat{F}^j(h(\xi^j) - T(\xi^j)x^n)$ and hence an optimality cut $\theta \geqslant \sum_{j=1}^{r} p^j \hat{F}^j(h(\xi^j) - T(\xi^j)x)$ is generated which cuts off (x^n, θ^n). To prove optimality we note that $cx^n + \theta^n$ is always an upper bound on the value z of (3.1). □

Remark 3.7. Following the lines of [12] we can use a multicut approach for (2.3). The objective of (3.5) is then replaced by $cx + \sum_{j=1}^{r} p^j \theta^j$ and the corresponding optimality cuts take the form $\theta^j \geqslant \hat{F}^j(h(\xi^j) - T(\xi^j)x)$, $j = 1, \ldots, r$, where the functions $\hat{F}^j$ are obtained as above.

Remark 3.8. As noted in [5], the performance of this method depends heavily on the class of dual price functions applied. At one extreme we can consider the value function itself or some class of functions which contains the value function and at another extreme we can consider point estimates, viz. functions of the type

$$F(d) = \begin{cases} \Phi(d) & \text{for } d \geqslant h(\xi) - T(\xi)x^*, \\ l(d) & \text{otherwise,} \end{cases}$$

where $l(d)$ is any lower bound on $\Phi(d)$. This is the type of price functions employed by Laporte and Louveaux in their branch-and-cut procedure for integer recourse problems. This yields a finite procedure if X is finite, but the number of constraints needed is in general exponential in n_1. With binary first stage variables linear optimality cuts can be obtained in this way, see [3].

Definition 3.9 (Laporte and Louveaux [3]). A set of feasibility cuts is said to be *valid* if there exists some finite number s, such that $x \in X \cap K$ if and only if $0 \geqslant g_k(x)$, $k = 1, \ldots, s$. A set of t optimality cuts is said to be *valid* if for all $x \in X \cap K$ it holds that $(x, \theta) \in \{(x, \theta) \mid \theta \geqslant g_l(x),\ l = 1, \ldots, t\}$ implies $\theta \geqslant Q(x)$.

Obviously, if a valid set of feasibility cuts and a valid set of optimality cuts is generated in the L-shaped method, then the procedure converges in a finite number of steps. Note that finiteness here only means that the relaxed master problem and the subproblems have to be solved a finite number of times. We have not addressed the question of how to solve the current problem. The form of the constraints in the current problem will depend on the class of price functions employed and thereby the algorithm applied to solve (3.2).

A precise description of constraints of the master problem is available when the rank of the polyhedron $\operatorname{conv}\{y \in \mathbb{Z}_+^{n_2} \mid Wy \geqslant h(\xi) - T(\xi)x\}$ is known (see [7] for a definition of the rank of a polyhedron).

Example 3.10. Let $G = (V, E)$ be a graph with vertices V and edges E. Let n be the number of vertices and denote by $E(i)$ the set of edges incident to vertex $i \in V$. The first stage problem is a knapsack problem, where a set of vertices has to be selected such that total weight is maximized without exceeding the capacity b $(b < n)$.

$$\begin{aligned} \max\quad & \sum_{i \in V} c_i x_i \\ \text{s.t.}\quad & \sum_{i \in V} x_i \leqslant b, \quad x_i \in \{0, 1\}. \end{aligned}$$

The resulting subgraph is given by the set of vertices i having $x_i = 1$ and the set of edges incident to these vertices. In the second stage a b-matching of maximum weight on this subgraph has to be determined. The number of edges that can be assigned to vertex i on the subgraph is ξ_i, where ξ_i is a random variable with nonnegative, integer values $\xi_i^1, \ldots, \xi_i^r$. The second stage problem hence is

$$\begin{aligned} \max\quad & \sum_{e \in E} q_e y_e \\ \text{s.t.}\quad & \sum_{e \in E(i)} y_e \leqslant \xi_i x_i, \quad i \in V, \\ & y_e \geqslant 0 \text{ and integer}. \end{aligned} \tag{3.6}$$

The optimal solution of the subadditive dual of problem (3.6) is of the form

$$F(d) = \sum_{i \in V} u_i d_i + \sum_{U \subseteq V} \left\lfloor \frac{1}{2} \sum_{i \in U} d_i \right\rfloor v_U,$$

where $u_i, v_U \geqslant 0$ for all $i \in V$ and all $U \subseteq V$. The optimality cuts have the form

$$\theta \leqslant \sum_{j=1}^{r} p^j \left(\sum_{i \in V} u_i^j \xi_i^j x_i + \sum_{U \subseteq V} \left\lfloor \frac{1}{2} \sum_{i \in U} \xi_i^j x_i \right\rfloor v_U^j \right),$$

where the multipliers u_i^j, $i \in V$ and v_U^j, $U \subseteq V$ and $\sum_{i \in U} \xi_i^j x_i$ odd, are determined by solving for each outcome ξ_i^j, $i \in V$, $j = 1, \ldots, r$ the problem (cf. [10])

$$\begin{aligned} \min\quad & \sum_{i \in V} u_i \xi_i x_i + \sum_{U \subseteq V} \left\lfloor \frac{1}{2} \sum_{i \in U} \xi_i x_i \right\rfloor v_U \\ \text{s.t.}\quad & u_i + u_j + \sum_{E(U) \ni e} v_U \geqslant q_e, \quad \text{for all } e = (i, j) \in E \\ & u_i \geqslant 0, \quad v_U \geqslant 0 \text{ for all } i \in V, \ U \subseteq V. \end{aligned}$$

4. Cutting plane techniques

We now explain how the generalized L-shaped method works when the second stage problem is solved by cutting plane techniques. In this section $\mathscr{F}$ will denote the class of functions $F : \mathbb{R}^{m_2} \to \mathbb{R}$ that are nondecreasing, subadditive and satisfy $F(0) = 0$. Using these properties of the price functions, the dual problem (3.3) is equivalent to the following problem, see [7], where the constraints are independent of the primal variable:

$$\begin{aligned} \max\quad & F(h(\xi^j) - T(\xi^j)x) \\ \text{s.t.}\quad & F(w_j) \leqslant q_j, \quad j = 1, \ldots, n_2, \quad F \in \mathscr{F}. \end{aligned} \tag{4.1}$$

The subadditive dual of (3.4) can similarly be written

$$\begin{array}{ll} \max & G(h(\xi^j) - T(\xi^j)x) \\ \text{s.t.} & G(w_j) \leqslant 0,\ G(e_i) \leqslant 1,\ j = 1, \ldots, n_2,\ i = 1, \ldots, m_2, \\ & G \in \mathscr{F}, \end{array} \tag{4.2}$$

where e_i is the ith unit vector of $\mathbb{R}^{m_2}$. The constraints $G(e_i) \leqslant 1$, $i = 1, \ldots, m_2$, ensure that the problem is bounded.

In a standard cutting plane procedure valid inequalities are successively generated and added to the constraints defining the feasibility set. The resulting problem is then solved by linear programming without integrality restrictions. This procedure is repeated until the current LP-solution is integral. The cuts can be written in the form

$$\sum_{j=1}^{n_2} F^{(l)}(w_j)y_j \geqslant F^{(l)}(q), \quad l = 1, \ldots, \tau, \tag{4.3}$$

where $F^{(l)} \in \mathscr{F}$, $l = 1, \ldots, \tau$. See e.g., [7] or [13]. At termination of the cutting plane procedure dual variables $(u_1, \ldots, u_{m_2}, u_{m_2+1}, \ldots, u_{m_2+\tau}) \geqslant 0$ are obtained from the LP-solution. We then define a function $F : \mathbb{R}^{m_2} \to \mathbb{R}$ by

$$F(d) := \sum_{i=1}^{m_2} u_i d_i + \sum_{i=1}^{\tau} u_{m_2+i} F^{(i)}(d). \tag{4.4}$$

By construction it follows that $F \in \mathscr{F}$ and by LP-duality we see that F is a feasible solution of (4.1). Moreover, it is an optimal solution of (4.1) and the duality gap between (4.1) and (3.2) is closed. For details see [10,13].

As an example consider the case where the cutting planes are generated using Gomory's Fractional Cutting Plane Algorithm (Gomory FCPA). Gomory cuts have long been disfavoured due to poor computational experience in the sixties and early seventies of implementations. However, recent experiments with Gomory cuts shows that it can be a very efficient way of solving integer programs. This success is largely due to the greatly improved LP-codes now available, together with a branch-and-cut framework. The Gomory cut generated from a source row of the final Simplex tableau of the LP-relaxation of the second stage problem is (see, e.g., [7]) $\sum_{j=1}^{n_2} F(w_j)y_j \geqslant F(q)$, where $F(d) = \lceil \sum_{i=1}^{m_2} \lambda_i d_i \rceil$ and $\lambda_i = \lceil \mu v_{pi} \rceil - \mu v_{pi}$. Here μ is some parameter and v_{pi} are the elements of the current inverse basis matrix corresponding to the source row. The successive cuts are of the form $\sum_{j=1}^{n_2} F^{(l)}(w_j)y_j \geqslant F^{(l)}(q)$ where

$$F^{(l)}(d) = \left\lceil \sum_{i=1}^{m_2} \lambda_i^{(l-1)} d_i + \sum_{i=1}^{l-1} \lambda_{m_2+i}^{(l-1)} F^{(i)}(d) \right\rceil, \quad l = 1, \ldots, \tau, \tag{4.5}$$

$F^{(0)} \equiv 0$ and the vector $\lambda^{(l-1)} \geqslant 0$ is obtained from the coefficients of the slack variables in the $(l-1)$th iteration. We see that $F^{(l)} \in \mathscr{F}$, $l = 1, \ldots, t$, and that they are Chvátal functions too. Introducing auxiliary variables and representing the round-up operations in (4.5) by integrality requirements, the current problem (3.5) can be rewritten as a mixed-integer linear program. However, the number of auxiliary variables will equal the total number of nested round-up operations in the price functions.

Gomory has shown that upon further specification of the Gomory FCPA, lexicographic dual feasibility and a known lower bound on the sequence of optimal values of the LP-relaxation, finite convergence is guaranteed. As Theorem 2 reveals, the class of dual price functions can indeed be tightened to the class $\mathscr{C}^{m_2}$ of m_2-dimensional Chvátal-functions. The dual price functions generated by the Gomory FCPA will not necessarily correspond to the Chvátal functions constituting the value functions, but finiteness can still be guaranteed. The following assumption is necessary to bound the number of Gomory cuts uniformly in x and ξ:

(A3) Suppose that for some suitably large $a \in \mathbb{Z}_+$ we have

$$\left\{ y \in \mathbb{Z}_+^{n_2} \;\middle|\; Wy \geqslant h(\xi^j) - T(\xi^j)x \right\} \subseteq \left\{ y \in \mathbb{R}_+^{n_2} \;\middle|\; \sum_{j=1}^{n_2} y_j \leqslant a \right\}$$

for all $\xi^j \in \Xi$ and all $x \in X$.

In particular assumption (A3) implies that the recourse function is bounded from below.

Theorem 4.1. *Suppose assumption* (A3) *is fulfilled. Then valid sets of feasibility cuts and optimality cuts are generated in finitely many step and the L-shaped method converges finitely.*

Proof. We show that only a finite number of dual price functions can be generated from the second stage problems. This implies finiteness since eventually the value function will be generated. Consider first the generation of optimality cuts. In the first step a Gomory cut $\sum_{j=1}^{n_2} F^1(w_j)y_j \geqslant F^1(q)$ is generated from a source row and added to the LP-relaxation of (3.2). A dual price function is then obtained by solving the resulting dual LP-problem. Only a finite number of dual price functions can be obtained by this procedure, because we can restrict attention (cf. (4.4)) to basic solutions of the system

$$\left\{ u \in \mathbb{R}_+^{m_2+1} \;\middle|\; \sum_{i=1}^{m_2} u_i w_{ij} + u_{m_2+1} F^1(w_j) \leqslant q_j \right\},$$

which are independent of the right-hand side of the second stage problem (3.2). In each step of the FCPA, only a finite number of dual price functions can be obtained, since the number of source rows and the number of basic feasible solutions of the dual LP-problem both are finite. By a result from integer programming (see e.g., Theorem II.4.3.8 in [7]), the maximum number of cuts needed by the Gomory FCPA to solve the second stage problem to optimality or detect infeasibility is independent of the right-hand side. Thus the number of optimality cuts is finite.

Consider now the generation of feasibility cuts. Feasibility cuts are generated by solving the problem

$$\min \left\{ et \;\mid\; Wy + It \geqslant h(\xi^j) - T(\xi^j)x,\; (y,t) \in \mathbb{Z}_+^{n_2} \times \mathbb{Z}_+^{m_2} \right\} \tag{4.6}$$

for right-hand sides $h(\xi^j) - T(\xi^j)x$, $j = 1, \ldots, r$. If we denote the value of (4.6) by $\varphi(x, \xi^j)$ we see that feasibility cuts are simply optimality cuts for the functions $\varphi(\cdot, \xi^j)$. The argument above applies and the number of feasibility cuts is also finite. □

5. Branch-and-bound

In this section we discuss the generalized L-shaped method when branch-and-bound is employed for solving the second stage problem. A branch-and-bound algorithm generates a branching tree during the solution process. At each node of the tree integer upper and lower bounds (possibly $-\infty$ or $+\infty$) are stated for the integer variables and the LP-relaxation of the resulting problem is considered. The goal is to solve the second stage problem (3.2) completely by generating a tree such that the optimum of (3.2) is found as an integer solution to an LP-problem at one of the nodes. The class of price functions appropriate for this type of algorithm was shown in [10] to be price functions of the form

$$F(d) := \min_{i=1,\ldots,P} \{u^i d + b^i\}, \quad u^i = (u^i_1, \ldots, u^i_{m_2}) \geqslant 0$$

for some finite $P \in \mathbb{N}$, i.e., the class of all nondecreasing, concave, polyhedral functions. These dual price functions are generated as follows. Let $i = 1, \ldots, P$ index the terminal nodes of a current tree and let $d \in \mathbb{R}^{m_2}$ be some right-hand side. Denoting upper and lower bounds for the ith node by l^i and k^i, the problem corresponding to the ith node is:

$$\begin{aligned} \min \quad & qy \\ \text{s.t.} \quad & Wy \geqslant d, \\ & k^i \leqslant y \leqslant l^i. \end{aligned} \tag{5.1}$$

Introducing dual variables $\underline{u}$ and $\overline{u}$ for the lower and upper bounds, the dual problem of (5.1) is

$$\begin{aligned} \max \quad & ud + \underline{u}k^i - \overline{u}l^i \\ \text{s.t.} \quad & uW + \underline{u} - \overline{u} \leqslant q, \\ & u, \underline{u}, \overline{u} \geqslant 0. \end{aligned} \tag{5.2}$$

If (5.2) is unbounded a dual ray $(r^i, \underline{r}^i, \overline{r}^i) \geqslant 0$ satisfying $r^i W + \underline{r}^i - \overline{r}^i \leqslant 0$ and $r^i d + \underline{r}^i k^i - \overline{r}^i l^i > 0$ is obtained from the usual Phase I problem and we put $(u^i, \underline{u}^i, \overline{u}^i) := (u, \underline{u}, \overline{u}) + \lambda(r^i, \underline{r}^i, \overline{r}^i)$, where $(u, \underline{u}, \overline{u})$ is a feasible solution of some previously solved subproblem and $\lambda \geqslant 0$. Otherwise, (5.2) has an optimal solution $(u^i, \underline{u}^i, \overline{u}^i)$. In both cases we define the function $f_i(d) := u^i d + \underline{u}^i k^i - \overline{u}^i l^i = u^i d + b^i$. Then $f_i(d)$ equals the optimal value of (5.1) and (5.2) if finite; otherwise $f_i(d)$ can be made arbitrary large by letting λ tend to $+\infty$. Performing these operations for all terminal nodes we then define the price function F by

$$F(d) := \min_{i=1,\ldots,P} f_i(d) = \min_{i=1,\ldots,P} \{u^i d + b^i\}. \tag{5.3}$$

If y is integer then $F(Wy) \leqslant \min_{i=1,\dots,P}\{u^i Wy + \underline{u}^i k^i - \overline{u}^i l^i\} \leqslant qy$ by LP-duality and this implies that F is dual feasible in (3.3). Moreover, if the minimum is taken over all end nodes of the branching tree at termination, then F is an optimal solution of (3.3).

When the dual price functions are obtained by branch-and-bound technique the current problem (3.5) becomes a disjunctive program. Consider again the multicut version. The optimality cut $\theta^j \geqslant F^j(h(\xi^j) - T(\xi^j)x)$ where F^j has the form (5.3) is then equal to the requirement that θ^j be greater than or equal at least one of the terms $u^1(h(\xi^j) - T(\xi^j)x), \dots, u^P(h(\xi^j) - T(\xi^j)x)$. A similar and somewhat stronger dual function, using also dual multipliers from nonterminal nodes, was suggested in [14] and applied there to integer programming sensitivity analysis.

Theorem 5.1. *Suppose a branch-and-bound algorithm terminates finitely when applied to problem* (3.2) *and that Assumption* (A3) *is fulfilled. Then valid sets of feasibility cuts and optimality cuts are generated in finitely many steps and the L-shaped method converges finitely.*

Proof. The support Ξ of ξ is finite so at most a finite number of feasibility cuts are generated for a given x. The L-shaped method does not cycle by Theorem 3.6 and by Assumption (A3) the number of possible branching trees is finite, hence valid sets of feasibility cuts and optimality cuts are generated in a finite number of steps. □

Remark 5.2. If first stage decision variables are restricted to integers, then Assumption (A3) is implied by the stronger assumption that $\{x \in \mathbb{Z}_+^{n_1} \mid Ax \geqslant b\}$ is finite. The number of possible branching trees is finite due to Assumption (A3) but grows exponential with n_2.

6. Concluding remarks

The contribution of this paper is to show how the classical L-shaped method can be generalized to stochastic programs with integer recourse. The procedure leads to nonlinear master problems where the constraints depend on the algorithm employed for solving integer programs. Two particular cases have been examined where the procedure has finite convergence provided the master problem can be solved by a finite method. Moreover, in the first case the master problem can be transformed into a mixed-integer program, whereas in the second case the master problem is equivalent to a disjunctive program. The generalized L-shaped procedure above stresses the fact that decomposition of stochastic programs with respect to time stages can not be separated from the gathering of dual information via appropriate dual price functions. The nonlinearity of the master problem is also reflected by the difficulty of performing sensitivity analysis in integer programming, where dual price functions enter naturally, see e.g., [14]. However, the master problem obtained from either of the two approaches is not in general computationally attractive. In order to become efficient, more structure on problem (2.1) should be imposed and attention should be limited to a restricted class of dual price functions, e.g., by Lagrangian relaxation of (part of) the second stage constraints, by considering only price functions separable in their arguments, or by considering only price functions with one level of round-up operations. In this way more tractable master problems can be obtained at the expense of optimality.

Acknowledgements

We are grateful to two anonymous referees for valuable comments and suggestions.

References

[1] R. Schultz, L. Stougie, M.H. van der Vlerk, Two-stage stochastic integer programming: A survey, Statist. Neerlandica 50 (1996) 404–416.
[2] L. Stougie, Design and Analysis of Methods for Stochastic Integer Programming, CWI Tract 37, Amsterdam, 1987.
[3] G. Laporte, F.V. Louveaux, The integer L-shaped method for stochastic integer programs with complete recourse, Oper. Res. Lett. 13 (1993) 133–142.
[4] J. Tind, L.A. Wolsey, An elementary survey of general duality theory in mathematical programming, Math. Programming 21 (1981) 241–261.
[5] L.A. Wolsey, A resource decomposition algorithm for general mathematical programs, Math. Programming Study 14 (1981) 244–257.
[6] R. Schultz, On structure and stability in stochastic programs with random technology matrix and complete integer recourse, Math. Programming 70 (1995) 73–89.
[7] G.L. Nemhauser, L.A. Wolsey, Integer and Combinatorial Optimization, Wiley-Interscience, New York, 1988.
[8] C.E. Blair, R.G. Jeroslow, The value function of an integer program, Math. Programming 23 (1982) 237–273.
[9] R.M. Van Slyke, R.J.-B. Wets, L-shaped linear programs with application to optimal control and stochastic programming, SIAM J. Appl. Math. 17 (1969) 638–663.
[10] L.A. Wolsey, Integer programming duality: Price functions and sensitivity analysis, Math. Programming 20 (1981) 173–195.
[11] J. Tind, Multilevel optimization: A common framework for decomposition in general mathematical programming, Publication no. 93/2, Department of Operations Research, University of Aarhus, July 1993.
[12] J.R. Birge, F.V. Louveaux, A multicut algorithm for two-stage stochastic linear programs, European J. Oper. Res. 34 (3) (1988) 384–392.
[13] R.G. Jeroslow, Cutting plane theory: Algebraic methods, Discrete Math. 23 (1978) 121–150.
[14] L. Schrage, L.A. Wolsey, Sensitivity analysis for branch and bound integer programming, Oper. Res. 33 (1985) 1008–1023.