Bounding in Multi-Stage Stochastic Programming Problems

Olga Fiedler*

and

András Prékopa†

31 May 1995

Abstract

We assume that the random variables corresponding to the subsequent periods are all discrete with finite supports. We also assume that the problem is written in the λ-representation form. Starting from the last period, and proceeding in the backward direction, we create first a dual feasible basis which provides us with a lower bound. Next, a primal feasible basis is created which gives us an upper bound. These steps are repeated several times. If the bounds are not satisfactorily, then a few dual steps are performed. The proposed bounding technique is very simple in the case of a multiperiod simple recourse problem. In this case, the dual steps are executed effectively by careful selections of the incomimg and outgoing vectors.

*Freie Universität Berlin, FB Mathematik u. Informatik, Arnimalle 2-6, 14195 Berlin, Germany

†RUTCOR, Rutgers Center for Operations Research, New Brunswick, New Jersey 08903-5062, USA

0. Introduction

A simple *Bounding Procedure* to obtain fast bounds for the optimum value of the Simple Recourse Stochastic Programming Problem, where the violation of the random constraints are penalized by piecewise linear, convex functions, was proposed Prekopa 1990 (see [8]). Later, Prekopa and Li used this bounding technique in a PERT optimization problem (see [9]) and then generalized it for the case, where the penalty-function is a *multivariate* polyhedral convex function (see [10]). The main idea is the following: assuming that the random variables are all discrete with finite supports and that the problem is written in the λ-representation form, one can easily construct an initial dual feasible basis and compute the corresponding *lower* bound for the optimum value; then, a primal feasible basis is constructed which gives an *upper* bound.

In this paper the *Bounding Procedure* will be generalized for *Multi-Stage* Stochastic Programming Problems. If we formulate the multi-stage problem using the ideas from dynamic programming, we can observe that the matrix of the equality constraints (written in the λ-representation form in each stage), has the same block-structure as those in [8] and [10]. Now, starting from the last stage, and proceeding in the backward direction, one can create a dual feasible basis, combining the bases of all stages. Then, starting from the first stage and proceeding in the forward direction, one constructs a primal feasible basis. This provides us with fast lower and upper bounds for the optimum value. The procedure is repeated several times. If the bounds are not satisfactorily close, then a few primal or dual steps are performed.

Bounding procedures are very important in single as well as multi-stage stochastic programming problems. In the solving algorithms for single stage problems, such as the dual type method (DTM) of Prékopa (see [8]) and Improved DTM of Fiedler, Prékopa and Fábián (see [4]), one can use them to get fast information about the optimum value or to gain a good starting position to solve the problem. Even more important are bounding procedures in multi-stage problems.

In the last decade several authors, such as Birge, Gassman, Higle, Sen, Louveaux, Ruszczynski and Wets, developed various solution techniques for multi-stage stochastic programming problems (see e.g. [1], [5], [6], [2], [3], [7]). These methods are capable to solve problems at most 4-5 stages, because the size of the stochastic programming problem exponentially increases with the number of stages. This is called the "curse of dimensionality" in dynamic programming. Thus, bounding procedures may be the only tools to obtain information about the optimum value in case of a large number of stages, where the solution of the problem cannot be carried out.

1. Formulation of the problem

We start from the following underlying deterministic (N+1)-period linear programming problem with *staircase* structure:

$$
\begin{array}{llccccccccl}
 & \min \{ & q_0^T x^0 & + & q_1^T x^1 & + & q_2^T x^2 + & \dots & + & q_N^T x^N \} & \\
 & \text{s.t.} & A\, x^0 & & & & & & & & = b \\
 & & T^1 x^0 & + & W^1 x^1 & & & & & & = \xi^1 \\
(1.1) & & & & T^2 x^1 & + & W^2 x^2 & & & & = \xi^2 \\
 & & & & & & \ddots & & & & \vdots \\
 & & & & & & & T^N x^{N-1} + & & W^N x^N & = \xi^N \\
 & & x^0 \geq 0\,, & & x^1 \geq 0\,, & & x^2 \geq 0\,, & \dots, & & x_N \geq 0\,; &
\end{array}
$$

here $\xi^t \in R^{n_t}$, $W^t \in R^{n_t}$ x R^{m_t}, $t = 1, \dots, N$, $b \in R^{n_0}$, $x^0 \in R^{m_0}$ and the other data are of suitable sizes, compatible with the formulation (1.1).

We assume that the RHS-vectors ξ^t , $t = 1, \dots, N$ are discrete random variables, each with a finite number of possible values ξ^t_k and with corresponding *path probabilities* p^t_k (cf. [5]). Furthermore, we suppose that the technology matrix in stage t is equal to the recourse matrix or its negative in the previous stage $t-1$ for $t = 2, \dots, N,$, i.e.

$$T^t = \pm W^{t-1} \quad (t = 2, \dots, N).$$

One formulates the multi-period stochastic programming problem, based on problem (1.1), in the form of a single large scale linear programming problem as follows:

$$
\begin{aligned}
\min \{ \quad & q_0^T x^0 \ + \sum_{k=1}^{M_1} p_k^1 q_1^T x_k^1 \ + \sum_{k=1}^{M_2} p_k^2 q_2^T x_k^2 \ + \ldots \quad + \sum_{k=1}^{M_N} p_k^N q_N^T x_k^N \ \} \\
\text{s.t.} \quad & A\, x^0 \ = \ b \\
& T_k^1 x^0 \ + \ W_k^1 x_k^1 \ = \ \xi_k^1, \quad k = 1, \ldots, M_1 \ , \\
& \pm W_{\alpha(k)}^{t-1} x_{\alpha(k)}^{t-1} \ + \ W_k^t x_k^t \ = \ \xi_k^t, \quad k = 1, \ldots, M_t \ , \ t = 2, \ldots, N \\
& x^0 \geq 0 \ , \ x_k^t \geq 0 \ , \quad k = 1, \ldots, \sum_{t=1}^{N} M_t \ , \ \ t = 1, \ldots, N \ ,
\end{aligned}
\tag{1.2}
$$

where W_k^t is the recourse matrix for node k at stage t, $\alpha(k)$ is the *immediate* ancestor of node k at stage t, M_t is the number of possible values at stage t, $t = 1, \ldots, N$.

2. The basic idea

We reformulate the problem by the use of the λ-representation and the construct dual and primal feasible bases. First a dual feasible basis (DFB) is constructed which provides us with a *lower bound*. Then, we construct a primal feasible basis (PFB) and obtain an *upper bound*. To create the DFB for (1.2) we adopt the idea for the construction of the initial DFB in the DTM, developed by Prékopa (see [8]) for the simple recourse problems, and then extended by Prékopa and Lee (see [10]) for linearly constrained optimization problems with convex polyhedral objective functions. The λ-representation gives rise to a specially block-constrained LP for which the DFB can easily be constructed based on the *fundamental property* of DFBs for such problems (see Th. 2.1 in [8]).

We start with the application of λ-linearization method. For each decision variable x_k^t, $k = 1, \ldots, \sum_{t=1}^{N} M_t$, $t = 1, \ldots, N$, we introduce the following functions:

$$
\begin{aligned}
f_k^t(z_k^t) = \ & \min_{x_k^t} \ p_k^t q_t^T x_k^t \\
\text{s.t.} \ \ & W_k^t x_k^t = z_k^t, \quad x_k^t \geq 0, \ z_k^t \in R^{n_k^t}.
\end{aligned}
\tag{2.1}
$$

Since the optimum value of a linear program, that depends on its RHS parameters $z_k^t \in R^{n_k^t}$, is a convex polyhedral function, there exists a subdivision of the space $R^{n_k^t}$ into the convex polyhedra (simplices) S_{kh}^t, with pairwise disjoint interiors, such that f_k^t is linear on each S_{kh}^t and continuous and convex on their union $\bigcup_h S_{kh}^t$. For fixed k and t let $z_{k1}^t, \ldots, z_{kH_k^t}^t$

designate the set of all vertices of all S^t_{kh}. Then, the λ-representation of the function $f^t_k(\cdot)$ is the following:

$$
\text{(2.2)} \qquad
\begin{aligned}
f^t_k(z^t_k) = \min\nolimits_\lambda \quad & \sum_{h=1}^{H^t_k} f^t_k(z^t_{kh})\lambda^t_{kh} \\
\text{s.t.} \quad & \textstyle\sum_h \; z^t_{kh}\lambda^t_{kh} = z^t_k \\
& \textstyle\sum_h \; \lambda^t_{kh} = 1, \quad \lambda^t_{kh} \geq 0,
\end{aligned}
$$

where the summation over h in the last two rows is the same as in the first one. Let us specialize z^t_k from (1.2) as follows:

$$
z^t_k = \xi^t_k \pm W^{t-1}_{\alpha(k)} x^{t-1}_{\alpha(k)}, \quad k = 1, \ldots, \sum_{t=1}^{N} M_t \;, \quad t = 1, \ldots, N,
$$

where $\alpha(k)$ is the immediate ancestor of node k at stage t and $W^0_{\alpha(k)} := T^1_k$. If we substitute the functions $f^t_k(z^t_k)$ by their λ-representations from (2.2), we can reformulate the multi-stage stochastic programming problem (1.2) in the following manner:

(2.3)

$$
\begin{array}{llllllr}
\min\limits_{x^0,\lambda} & \{q_0^T x^0 + & \displaystyle\sum_{k=1}^{M_1}\sum_{h=1}^{H^1_k} f^1_{kh}\lambda^1_{kh} + & \displaystyle\sum_{k=1}^{M_2}\sum_{h=1}^{H^2_k} f^2_{kh}\lambda^2_{kh} + & \cdots & + \displaystyle\sum_{k=1}^{M_N}\sum_{h=1}^{H^N_k} f^N_{kh}\lambda^N_{kh}\} & \\
& Ax^0 & & & & & = b \\
& T^1_k x^0 + & \displaystyle\sum_{h=1}^{H^1_k} z^1_{kh}\lambda^1_{kh} \quad (k = 1, \ldots, M_1) & & & & = \xi^1_k \\
& & \displaystyle\sum_{h=1}^{H^1_k} \lambda^1_{kh} & & & & = 1 \\
& & \displaystyle\sum_{h=1}^{H^1_k} z^1_{\alpha(k)h}\lambda^1_{\alpha(k)h} + & \displaystyle\sum_{h=1}^{H^2_k} z^2_{kh}\lambda^2_{kh} \quad (k = 1, \ldots, M_2) & & & = \xi^2_k \\
& & & \textstyle\sum_{h=1}^{H^2_k} \; \lambda^2_{kh} & & & = 1 \\
& & & & \ddots & & \vdots \\
& & & & \displaystyle\sum_{h=1}^{H^{N-1}_k} z^{N-1}_{\alpha(k)h}\lambda^{N-1}_{\alpha(k)h} + & \textstyle\sum_{h=1}^{H^N_k} \; z^N_{kh}\lambda^N_{kh} & = \xi^N_k \\
& & & & & \displaystyle\sum_{h=1}^{H^N_k} \lambda^N_{kh} & = 1 \\
& & & & & (k = 1, \ldots, M_N) &
\end{array}
$$

The matrix of problem (2.3) has a special block-structure which is illustrated in Table 1 in case of three stages, i.e., $N = 2$, and the number of scenarios at the 3rd stage is equal to the number of descendants at the 2nd stage, i.e., $M_1 = M_2 = M$.

It is easy to see that in our special situation, where $T^t = \pm W^{t-1}$ all scenarios $\{\xi_k^t,\ k = 1, \dots, M_t\}$ at stage t ($t = 2, \dots, N$) belong to *at most* M_1 different spaces $R^{n_j}, j = 1, \dots, M_1$, where M_1 is the number of all possible values of the random RHS-vector ξ^1. Let $\beta^1(k,t)$ denotes the 2nd-stage ancestor of node (k,t). Then, if $\xi^1_{\beta^1(k,t)} = \xi^1_j \in R^{n_j},\ \ j \in \{1, \dots, M_1\}$, then $\xi_k^t \in R^{n_j}$, too.

Let us introduce the following notations:

$$\text{(2.4)} \qquad \left.\begin{aligned} &K^t := \{1, \dots, M_t\} \\ &I_j^t := \{k \mid k \in K^t,\ \xi_k^t \in R^{n_j}\}\ ,\ j = 1, \dots, M_1 \\ &K^t = \bigcup_{j=1}^{M_1} I_j^t \end{aligned}\right\} \quad t = 1, \dots, N.$$

Table 1

Matrix of equality constraints in case of 3-stage problem:

q_o	$f^1_{11} \dots f^1_{1H^1_1}$	$\dots$	$f^1_{M1} \dots f^1_{MH^1_M}$	$f^2_{11} \dots f^2_{1H^2_1}$	$\dots$	$f^2_{M1} \dots f^2_{MH^2_M}$	
A							b
T^1_1	$z^1_{11} \dots z^1_{1H^1_1}$						ξ^1_1
0	$1 \dots 1$						1
$\vdots$		$\ddots$					$\vdots$
T^1_M			$z^1_{M1} \dots z^1_{MH^1_M}$				ξ^1_M
0			$1 \dots 1$				1
	$z^1_{11} \dots z^1_{1H^1_1}$			$z^2_{11} \dots z^2_{1H^2_1}$			ξ^2_1
	$0 \dots 0$			$1 \dots 1$			1
		$\ddots$			$\ddots$		$\vdots$
			$z^1_{M1} \dots z^1_{MH^1_M}$			$z^2_{M1} \dots z^2_{MH^2_M}$	ξ^2_M
			$0 \dots 0$			$1 \dots 1$	1

The dual vector corresponding to any basis $\hat{B}$ for problem (2.3) will be partitioned as

$$(2.5) \qquad \begin{cases} y & \\ v_k^1 & k = 1, \ldots, M_1 \\ w_k^1 & \\ \vdots & \\ v_k^N & k = 1, \ldots, M_N, \\ w_k^N & \end{cases}$$

where $y \in R^{n_0}$, $(v_k^1, w_k^1) \in R^{n_k} \times R^1$, $k \in K^1$, $(v_k^t, w_k^t) \in R^{n_j} \times R^1$, $k \in K^t$, and $v_k^t \in R^{n_j}$ for $k \in I_j^t$, i.e. if the 2nd-stage ancestor of the scenario k at stage t $\xi^1_{\beta^1(k,t)}$ belongs to the space R^{n_j}, $j \in K^1$, $t = 2, \ldots, N$.

3. An Algorithm

Step 0:

Set **it** $:=$ **1**. For all $t = 1, \ldots, N$, $j = 1, \ldots, M_1$ and for each block $k \in K^t$ at stage t such that $k \in I_j^t$ choose any (n_j+1) vectors of z_{kh}^t with subscripts $\{(k, h_1), \ldots, (k, h_{n_j+1})\}$ so that the selected vectors constitute the set of the vertices of one of the (subdividing) $R^{n_j}-$dimensional simplices S_{kh}^t.

Step 1:

Set $\mathbf{t} := \mathbf{N}$ and $\bar{f}_{kh}^N := f_{kh}^N$ for all $h = 1, \ldots, H_k^N$, $k = 1, \ldots, M_N$.

Go to Step 3.

Step 2:

Compute the *modified* costs-coefficients at stage t (corresponding to the vectors selected in Step 1 as follows:

$$(2.6) \qquad \bar{f}_{kh_i}^t = f_{kh_i}^t - v_k^{t+1} z_{kh_i}^t \ , i = 1, \ldots, n_j + 1, \ k = 1, \ldots, M_t,$$

Step 3:

Solve the system of linear equations:

$$(2.7)\qquad \begin{array}{l} (z^t_{kh_i})^T \; v^t_k + w^t_k = \bar{f}^t_{kh_i}, \quad i \in \{1, \ldots, n_j + 1\}, \\ v^t_k \in R^{n_j}, \quad w^t_k \in R^1, \quad k \in I^t_j \quad (\bigcup_{j=1}^{M_1} I^t_j = K^t) \; . \end{array}$$

Step 4:

If $\mathbf{t} = \mathbf{1}$, then go to Step 5. Otherwise, set $\mathbf{t} = \mathbf{t} - \mathbf{1}$, and go to Step 2.

Step 5:

Solve the following linear programming problem by any method which produces a pair of primal and dual solutions but not necessarily an optimal basis:

$$(2.8)\qquad \begin{array}{ll} \min_{x_1} \{ (\; q_0^T - & \sum_{j=1}^{M_1} (v^1_k)^T T^1_k) x^0 \} \\ \text{s.t.} & A x^0 = b \; , \; x^0 \geq 0. \end{array}$$

Step 6:

Compute a lower bound $\underline{V}$ for the optimum value V of problem (2.3):

$$(2.9)\qquad \underline{V} := b^T y + \sum_{t=1}^{N} \sum_{k=1}^{M_t} \; [(v^t_k)^T \; \xi_t + \mathbf{1}^T w^t_k] \; \leq \; V \; ,$$

where y is the optimal dual vector for the problem (2.8), $v^t_k, \; w^t_k \; (k \in K^t, \; t = 1, \ldots, N)$ are the solutions of (2.7) and the symbol $\mathbf{1}$ denotes a vector with all components equal to 1.

Step 7:

If $\mathbf{t} > \mathbf{N}$, then go to Step 8.

Otherwise, for all $t = 1, \ldots, N, \; j = 1, \ldots, M_1$ and for each block $k \in K^t$ at stage t, such that $k \in I^t_j$, choose any $(n_j + 1)$ vectors of z^t_{kh} with the subscripts $(k, \tilde{h}_1), \ldots, (k, \tilde{h}_{n_j+1})$, so that the equations

$$(2.10)\qquad \left\{ \begin{array}{l} \sum_{i=1}^{n_j+1} \; z^t_{k\tilde{h}_i} \lambda^t_{k\tilde{h}_i} = \xi^t_k \pm W^{t-1}_k \; , \\ \sum_{i=1}^{n_j+1} \; \lambda^t_{k\tilde{h}_i} = 1, \quad \lambda^t_{k\tilde{h}_i} \geq 0, \quad k = 1, \ldots, M_t \end{array} \right.$$

are satisfied. Note, that the selected vectors $z_{k\tilde{h}_i}, i = 1, \ldots, n_j + 1$, constitute the set of the vertices of one of the subdividing R^{n_j} simplices S^t_{kh}; this simplex can differ from the simplex identified in Step 1 for the corresponding stage t.

Set $\mathbf{t} := \mathbf{t} + \mathbf{1}$, and go to Step 7.

Step 8:

Compute an upper bound $\bar{V}$ for the optimum value V of the problem (2.3):

$$\bar{V} := q_0^T \, x^0_{opt} + \sum_{t=1}^{N} \sum_{k=1}^{M_t} \sum_{i=1}^{n_j+1} f^t_{k\tilde{h}_i} \, \lambda^t_{k\tilde{h}_i} \, , \tag{2.11}$$

where x^0_{opt} is the primal optimal solution of problem (2.8) and $\lambda^t_{k\tilde{h}_i}$ are defined in Step 7.

□

Remark 2.1

If we combine now the vectors selected in Step 0 with the vectors in the *working basis* B we obtain a *dual* feasible basis for the problem (2.3). The dual feasibility of $\hat{B}$ can easily be shown using similar considerations as those in Lemma 2.1 in [8] and Remark 1.1 in [4]. On the other hand, if we combine the vectors selected in Step 7 with the vectors in the working basis B, we obtain a *primal* feasible basis for the problem (2.3).

□

Remark 2.2

The described *bounding procedure* can be iterated as follows:

Step 9: Compute

$$\Delta V := \bar{V} - \underline{V}.$$

If ΔV is sufficiently small, then STOP.

Otherwise, set $\mathbf{it} := \mathbf{it} + \mathbf{1}$, replace the subscripts selected in Step 0 for the subscripts defined in Step 7, i.e. set $(k, h_i) = (k, \tilde{h}_i)$, and go to Step 2.

However, there is *no garantee* that the new bounds will be better than the previous ones. Therefore, recommend to repeat the Bounding Procedure subsequently several times and accept those bounds for which the duality gap ΔV is the smallest. If ΔV is unsatifactorily large, perform a few dual steps.

$\square$

References

[1] **Birge J.R.** (1985). Decomposition and partitioning methods for multi-stage stochastic linear programs, *Operations Research*, 33, 989–1007

[2] **Birge J.R., Louveuax F.V.** (1988). A multicut algorithm for two-stage linear programs, *European Journal of OR* 34, 484–392

[3] **Birge J.R., Wets J.B.** (1986). Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse, *Mathematical Programming* 27, 54–102

[4] **Fiedler O., Prékopa A., Fábián C.I.** (1995). On a dual method for a specially structured linear programming problem, *RUTCOR Research Report* (to appear)

[5] **Gassman H.** (1990). MSLiP: A computer code for the multi-stage stochastic linear programming problem, *Mathematical Programming* 47, 407–423

[6] **Higle J., Sen. S.** (1991). Stochastic decomposition: an algorithm for two-stage linear programs with recourse, Mathematics of OR 16, 650-669

[7] **Ruszczýnski A.** (1993). Parallel decomposition of multistage stochastic programming problems, *Mathematical Programming* 58, 201–228

[8] **Prékopa A.** (1990). On a dual method... Dual method for the solution of a one-stage stochastic programming problem with random RHS obeying a discrete probability distribution, *ZOR-Methods and Models of Operations Research* 34, 441-461.

[9] **Prékopa A., Li W.** (1992). On an optimization problem concerning the stochastic PERT problem, *RUTCOR Research Report* 18–92

[10] **Prékopa A., Li W.** (1995). Solution of and bounding in a linearly constrained optimization problem with convex, polyhedral objective function, *Mathematical Programming* (to appear)