A Lagrangean Duality based Branch and Bound for Solving Linear Stochastic Programs with Decision Dependent Uncertainty

Vikas Goel*and Ignacio E. Grossmann†
Department of Chemical Engineering, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213

Abstract

We address a class of planning problems where the optimization decisions influence the time of information discovery for a subset of the uncertain parameters. The standard stochastic programming approach cannot be used for these problems. We present a hybrid mixed-integer disjunctive programming formulation and a Lagrangean duality based branch and bound algorithm for these problems and illustrate the advantages of this approach using examples for a manufacturing problem.

Keywords: Design and operation, endogenous uncertainty, stochastic programming, Lagrangean dual.

1 Introduction

Handling uncertainty in design and planning of process systems has received significant attention over the last few years. Stochastic programming offers a general and systematic way of incorporating uncertainty in deterministic optimization models. Most previous work in stochastic programming (Sahinidis (2004)) deals with problems with *exogenous uncertainty* (Jonsbraten (1998); e.g. demands), where the optimization decisions cannot influence the stochastic process. In contrast, previous work on problems with *endogenous uncertainty* (e.g. uncertainty in process yields, properties of oil and gas fields), where the optimization decisions influence the underlying stochastic process, is limited to a few papers only. Goel and Grossmann (2004b) considered the problem of optimal design and planning for oil and gas field infrastructure. In this problem, the uncertainty in the size and quality of a field is resolved only when a well platform is installed at the field. The authors showed that a major complication in these problems is that the structure of the scenario tree depends on the design decisions and hence that the optimization model must include conditional equations through the use of disjunctions.

In this paper, we consider stochastic problems that have both exogenous and endogenous uncertainties. The endogenous uncertainty considered is such that optimization decisions determine when uncertainty is

*E-mail: vgoel@andrew.cmu.edu
†To whom all correspondence should be addressed. Tel.: (412) 268-3642. Fax: (412) 268-7139. E-mail: grossmann@cmu.edu.

resolved. The problem is described in section 2 while the proposed stochastic program and the branch and bound algorithm are presented in sections 3 and 4, respectively. Results for our approach are presented in section 5.

2 Problem description

We consider problems with a discretized time horizon, $\mathcal{T} = \{1, 2, \ldots, T\}$, and a finite set of "sources" of endogenous uncertainty, $\mathcal{I} = \{1, 2, \ldots, I\}$. Variables $b_{i,t}$, y_t and x_t have to be optimized over the time horizon. $b_{i,t}$ are 0-1 variables that represent design decisions and determine whether the endogenous uncertainty in source i is resolved in period t or not. Variables y_t and x_t could represent expansions in capacity and sales, respectively, in period t. These variables are vectors that could have both continuous and discrete components. ξ_t represents the *exogenous* uncertain parameter associated with time period $t \in \mathcal{T}$, while θ_i is the *endogenous* uncertain parameter associated with source $i \in \mathcal{I}$. The uncertainty in ξ_t will be resolved automatically in time period t while the resolution of uncertainty in θ_i depends on decisions $b_{i,t}$. The sequence of events in each time period is as follows. Decisions y_t and $b_{i,t}$ are implemented at the beginning of time period t. This is followed by resolution of uncertainty in the exogenous parameters ξ_t and in the endogenous parameter θ_i for source i if $b_{i,t} = 1$ and $b_{i,\tau} = 0$ for all $\tau < t$. Decisions x_t are implemented at the end of the time period.

We assume a discrete set of possible realizations, Ξ, for the vector $\xi = (\xi_1, \xi_2, \ldots, \xi_T)$ and a discrete set of possible realizations, Θ_i, for θ_i. The set of scenarios corresponds to $\Xi \times (\times_{i \in \mathcal{I}} \Theta_i)$, *i.e.*, for any realization of the vector of exogenous parameters, $\xi = (\xi_1, \xi_2, \ldots, \xi_T)$, the set of scenarios includes scenarios corresponding to all possible combinations of realizations for the endogenous parameters. Individual scenarios are indexed as $s \in \mathcal{S}$, where $\mathcal{S} = \{1, 2, \ldots, S\}$ represents the set of indices corresponding to all the scenarios.

For scenarios $s, s' \in \mathcal{S}$, $\mathcal{D}(s, s') = \{i | i \in \mathcal{I}, \theta_i^s \neq \theta_i^{s'}\}$ represents the set of sources of endogenous uncertainty that distinguish scenarios s, s' while $|\mathcal{D}(s, s')|$ represents the cardinality of this set. $\mathbf{t}(s, s') = \max_t \{t | t \in \mathcal{T}, \xi_\tau^s = \xi_\tau^{s'} \ \forall \tau \in \mathcal{T}, \tau \leq t\}$ is the latest time period t such that scenarios s, s' are indistinguishable based on exogenous uncertainty resolved up till and including period t. Note that if $\xi_1^s \neq \xi_1^{s'}$ then we choose $\mathbf{t}(s, s') = 0$. $\mathcal{L}^0 = \{(s, s') | s, s' \in \mathcal{S}, s < s', |\mathcal{D}(s, s')| = 0\}$ represents the set of pairs (s, s') such that scenarios s, s' are identical in realizations for all endogenous parameters. Similarly, $\mathcal{L}^{1+} = \{(s, s') | s, s' \in \mathcal{S}, s < s', |\mathcal{D}(s, s')| \geq 1\}$, $\mathcal{L}_T^1 = \{(s, s') | (s, s') \in \mathcal{L}^1, \mathbf{t}(s, s') = T\}$.

3 Model

The declarative form of stochastic programs for this class of problems is given by the following hybrid mixed-integer disjunctive programming model, $(P1)$.

$$(P1) \quad \phi = \min \quad \sum_{s \in \mathcal{S}} p^s \sum_{t \in \mathcal{T}} \left({}^w c_t^s w_t^s + {}^x c_t^s x_t^s + {}^y c_t^s y_t^s + \sum_{i \in \mathcal{I}} {}^b c_{i,t}^s b_{i,t}^s \right) \tag{1}$$

$$\text{s.t.} \quad \sum_{\tau\in\mathcal{T},\tau\le t}\left({}^{w}A^{s}_{\tau,t}w^{s}_{\tau} + {}^{x}A^{s}_{\tau,t}x^{s}_{\tau} + {}^{y}A^{s}_{\tau,t}y^{s}_{\tau} + \sum_{i\in\mathcal{I}} {}^{b}A^{s}_{i,\tau,t}b^{s}_{i,\tau}\right) \le a^{s}_{t} \quad \forall s\in\mathcal{S}, t\in\mathcal{T} \tag{2}$$

$$b^{s}_{i,1} = b^{s'}_{i,1} \quad \forall s,s'\in\mathcal{S}, s<s', i\in\mathcal{I} \tag{3a}$$

$$y^{s}_{1} = y^{s'}_{1} \quad \forall s,s'\in\mathcal{S}, s<s' \tag{3b}$$

$$x^{s}_{t} = x^{s'}_{t} \quad \forall (s,s')\in\mathcal{L}^{0}, t\in\mathcal{T}, t\le \mathbf{t}(s,s') \tag{4a}$$

$$b^{s}_{i,t+1} = b^{s'}_{i,t+1} \quad \forall (s,s')\in\mathcal{L}^{0}, t\in\mathcal{T}, t\le \mathbf{t}(s,s'), i\in\mathcal{I} \tag{4b}$$

$$y^{s}_{t+1} = y^{s'}_{t+1} \quad \forall (s,s')\in\mathcal{L}^{0}, t\in\mathcal{T}, t\le \mathbf{t}(s,s') \tag{4c}$$

$$\begin{bmatrix} & Z^{s,s'}_{t} & & \\ x^{s}_{t} & = & x^{s'}_{t} & \\ b^{s}_{i,t+1} & = & b^{s'}_{i,t+1}\ \forall i\in\mathcal{I} & \text{if } t\le T-1 \\ y^{s}_{t+1} & = & y^{s'}_{t+1} & \text{if } t\le T-1 \end{bmatrix} \vee \left[\neg Z^{s,s'}_{t}\right] \forall (s,s')\in\mathcal{L}^{1+}, t\in\mathcal{T}, t\le \mathbf{t}(s,s') \tag{5}$$

$$Z^{s,s'}_{t} \Leftrightarrow \bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{t}\left(\neg b^{s}_{i,\tau}\right)\right] \quad \forall (s,s')\in\mathcal{L}^{1+}, t\in\mathcal{T}, t\le \mathbf{t}(s,s') \tag{6}$$

$$Z^{s,s'}_{t} \Leftrightarrow \bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{t}\left(\neg b^{s'}_{i,\tau}\right)\right] \quad \forall (s,s')\in\mathcal{L}^{1+}, t\in\mathcal{T}, t\le \mathbf{t}(s,s') \tag{7}$$

$$w^{s}_{t}\in\mathcal{W}^{s}_{t}, x^{s}_{t}\in\mathcal{X}^{s}_{t},\ y^{s}_{t}\in\mathcal{Y}^{s}_{t},\ b^{s}_{i,t}\in\{0,1\} \quad \forall s\in\mathcal{S}, t\in\mathcal{T}, i\in\mathcal{I}$$

$$Z^{s,s'}_{t}\in\{True, False\} \quad \forall (s,s')\in\mathcal{L}^{1+}, t\in\mathcal{T}, t\le \mathbf{t}(s,s')$$

In $(P1)$, variables $b^s_{i,t}$, x^s_t and y^s_t represent the decisions to be made in time period t of scenario s. $b^s_{i,t}$ are binary variables while x^s_t and y^s_t are vectors that may have both integer and continuous components. Vector w^s_t represents other variables associated with (t,s). $\mathcal{W}^s_t$, $\mathcal{X}^s_t$ and $\mathcal{Y}^s_t$ represent the bounds and integrality restrictions on variables w^s_t, x^s_t and y^s_t. Equation (1) represents the objective of minimizing the expectation of an economic criterion. Inequality (2) represents constraints that govern decisions in individual scenarios.

Decisions for different scenarios are linked by non-anticipativity constraints, (3)-(7). The non-anticipativity rule requires that if scenarios s and s' are indistinguishable at some time, then decisions in s,s' should be the same at that time. Based on the sequence of events described in section 2, uncertainty is resolved in time period t after the implementation of decisions y^s_t and $b^s_{i,t}$. Thus, if scenarios s,s' are indistinguishable after resolution of uncertainty in time period t, then decisions $x^{(\cdot)}_t$, $b^{(\cdot)}_{i,t+1}$ and $y^{(\cdot)}_{t+1}$ should be the same for scenarios s,s'. Since all scenarios are indistinguishable before decisions $b^s_{i,1}$ and y^s_1 are implemented, therefore these decisions have to be the same for all scenarios (3). The condition $s<s'$ avoids duplication of constraints (3) for the same pair of scenarios s,s'.

The equations in (4) represent non-anticipativity constraints linking scenarios s, s' if $(s,s')\in\mathcal{L}^0$. (5)-(7) are non-anticipativity constraints linking scenarios s,s' if $(s,s')\in\mathcal{L}^{1+}$. In this case, the indistinguishability of scenarios s,s' in time period t depends on both, endogenous and exogenous uncertainty resolved in the past. Boolean variable $Z^{s,s'}_t$ represents whether scenarios s,s' are indistinguishable after the resolution of uncertainty in time period t. Disjunction (5) imposes non-anticipativity constraints on variables $x^{(\cdot)}_t, y^{(\cdot)}_{t+1}, b^{(\cdot)}_{i,t+1}$ for scenarios s,s' only if $Z^{s,s'}_t$ is $True$. Logic constraints (6) and (7) relate the indistinguishability of scenarios s,s' in time period t with decisions $b^s_{i,\tau}$ and $b^{s'}_{i,\tau}$ respectively.

4 Branch and bound algorithm

Goel and Grossmann (2004a) showed that (5)-(7) need to be applied only for $(s, s') \in \mathcal{L}_1^T, t \in \mathcal{T}$, which greatly reduces the dimensionality of the model[1]. Hence, we will use the reduced domain for (s, s', t) for (5)-(7). Model $(P1)$ is coupled in scenarios through the non-anticipativity constraints. We propose a branch and bound algorithm that generates lower bounds by solving a Lagrangean dual problem obtained by relaxing these constraints.

$$(P1_{LR}) \quad \phi_{RLR}({}^b\lambda, {}^x\lambda, {}^y\lambda, {}^z\lambda) =$$

$$\min \sum_{s\in\mathcal{S}} p^s \sum_{t\in\mathcal{T}} \left({}^w c_t^s w_t^s + {}^x c_t^s x_t^s + {}^y c_t^s y_t^s + \sum_{i\in\mathcal{I}} {}^b c_{i,t}^s b_{i,t}^s \right) + \sum_{\substack{s,s'\in\mathcal{S}\\ s<s'}} \left[\sum_{i\in\mathcal{I}} {}^b\lambda_{i,0}^{s,s'} \left(b_{i,1}^s - b_{i,1}^{s'} \right) + {}^y\lambda_0^{s,s'} \left(y_1^s - y_1^{s'} \right) \right]$$

$$+ \sum_{(s,s')\in\mathcal{L}^0} \sum_{t=1}^{\mathbf{t}(s,s')} \left(\sum_{i\in\mathcal{I}} {}^b\lambda_{i,t}^{s,s'} \left(b_{i,t+1}^s - b_{i,t+1}^{s'} \right) + {}^y\lambda_t^{s,s'} \left(y_{t+1}^s - y_{t+1}^{s'} \right) + {}^x\lambda_t^{s,s'} \left(x_t^s - x_t^{s'} \right) \right)$$

$$+ \sum_{(s,s')\in\mathcal{L}_T^1} \sum_{t\in\mathcal{T}} {}^z\lambda_t^{s,s'} \left(z_t^{s,s'} - z_t^{s',s} \right)$$

s.t. (2)

$$1 - \sum_{\substack{\tau\in\mathcal{T}\\ \tau\le t}} b_{i,\tau}^s \le z_t^{s,s'} \le 1 - b_{i,\tau}^s \qquad \forall (s,s') \in \mathcal{L}_T^1, t\in\mathcal{T}, \{i\} = \mathcal{D}(s,s'), \tau\in\mathcal{T}, \tau\le t \tag{8}$$

$$1 - \sum_{\substack{\tau\in\mathcal{T}\\ \tau\le t}} b_{i,\tau}^{s'} \le z_t^{s',s} \le 1 - b_{i,\tau}^{s'} \qquad \forall (s,s') \in \mathcal{L}_T^1, t\in\mathcal{T}, \{i\} = \mathcal{D}(s,s'), \tau\in\mathcal{T}, \tau\le t \tag{9}$$

$$w_t^s \in \mathcal{W}_t^s,\; x_t^s \in \mathcal{X}_t^s,\; y_t^s \in \mathcal{Y}_t^s,\; b_{i,t}^s \in \{0,1\} \qquad \forall s\in\mathcal{S}, t\in\mathcal{T}, i\in\mathcal{I}$$

$$0 \le z_t^{s,s'}, z_t^{s',s} \le 1 \qquad \forall (s,s') \in \mathcal{L}_T^1, t\in\mathcal{T}$$

Consider the formulation of the Lagrangean dual for the problem at the root node. Model $(P1_{LR})$ is obtained from $(P1)$ by relaxing (5), replacing (3)-(4) by penalty terms in the objective with Lagrange mulitpliers ${}^b\lambda_{i,t}^{s,s'}$, ${}^x\lambda_t^{s,s'}$ and ${}^y\lambda_t^{s,s'}$, and reformulating (6)-(7) as linear constraints (8)-(9) in terms of algebraic variables $z_t^{(\cdot,\cdot)}$. Also, we introduce dummy variables $z_t^{s',s}$ for $(s,s') \in \mathcal{L}_1^T$ so that (8) links $z_t^{s',s}$ to decisions $b_{(\cdot,\cdot)}^s$ while (9) links $z_t^{s,s'}$ to decisions $b_{(\cdot,\cdot)}^{s'}$. In addition, the equality $z_t^{s',s}$ between $z_t^{s,s'}$ is dualized with Lagrange multipliers ${}^z\lambda_t^{s,s'}$. Note that $z_t^{s,s'}, z_t^{s',s}$ are declared as continuous variables but will take 0-1 values automatically.

$(P1_{LR})$ is an MILP model and clearly a relaxation of $(P1)$ for any ${}^b\lambda, {}^x\lambda, {}^y\lambda, {}^z\lambda$. Further, $(P1_{LR})$ can be decomposed into one MILP sub-problem for each scenario. The following Larangean dual problem (Guignard and Kim (1987), Nemhauser and Wolsey (1988)) gives a lower bound to ϕ.

$$\phi_{LD} = \max_{{}^b\lambda, {}^x\lambda, {}^y\lambda, {}^z\lambda} \phi_{LR}\left({}^b\lambda, {}^x\lambda, {}^y\lambda, {}^z\lambda\right)$$

We generate lower bounds at each node of the branch and bound tree by solving one such Lagrangean dual problem. Upper bounds at each node are obtained by applying problem-specific heuristics to the solution of the Lagrangean dual in order to generate feasible solutions to the problem.

[1] Note that $\left\{(s,s',t) | (s,s') \in \mathcal{L}_T^1, t\in\mathcal{T}\right\} \subseteq \left\{(s,s',t) | (s,s') \in \mathcal{L}^{1+}, t\in\mathcal{T}, t \le \mathbf{t}(s,s')\right\}$.

Table 1: Computational results for Sizes problem

Problem Specifications						Proposed Branch and Bound After t CPU seconds				LP based Branch and Bound After $10 \cdot t$ CPU seconds		
I	T	S	Binary variables	Continuous variables	Constraints	t	Nodes	Best sol. found	% gap	Nodes	Best sol. found	% gap
5	5	16	400	4,977	13,927	502	9	120,026	0.010	26,490	120,026	0.025
4	5	27	540	6,103	22,160	430	4	112,608	0.010	6,899	112,621	0.157
6	6	32	1,152	18,081	52,581	11,546	15	144,054	0.065	79,844	144,154	0.332
7	7	64	3,136	60,993	181,939	13,507	3	245,930	0.038	71,151	246,025	0.237

The branching step partitions the feasible space by branching on dualized equality constraints and on relaxed disjunctions. Our strategy for branching on dualized equality constraints is similar to that used by Caroe and Schultz (1999). Branching on equality constraints linking variables $b_{i,t}^{(\cdot)}$ (or the binary components of variables $x_t^{(\cdot)}$ or $y_t^{(\cdot)}$) across scenarios s, s' is based on the standard dichotomy branching strategy. When branching on constraint $x_{l,t}^s = x_{l,t}^{s'}$, where $x_{l,t}^{(\cdot)}$ is a continuous component of $x_t^{(\cdot)}$, the feasible space is partitioned about the mean value of these variables in the solution of the Lagrangean dual. The same strategy is used for branching on equality constraints on continuous components of variables $y_t^{(\cdot)}$. When branching on a relaxed disjunction, the feasible region is bifurcated into regions where $Z_t^{s,s'} = Z_t^{s',s} = True$ and $Z_t^{s,s'} = Z_t^{s',s} = False$, respectively. The set of dualized equality constraints on the up-branch ($Z_t^{s,s'} = Z_t^{s',s} = True$) is augmented by the set of equality constraints inside the disjunction corresponding to (s, s', t). Note that the solution of the relaxed Lagrangean dual may be such that $\hat{Z}_t^{s,s'} = \hat{Z}_t^{s',s} = True$. Thus, introducing the restriction $Z_t^{s,s'} = Z_t^{s',s} = True$ may not alter the solution of the Lagrangean dual. Thus, the first branch is further bifurcated to eliminate infeasibility in one of the violated equality constraints.

It is important to note that logic inferencing (Hooker (2000)) on Boolean and discrete variables can significantly impact the quality of the lower bounds. Therefore, it is important to execute this step before the Lagrangean dual is solved. Also, if some components of variables x_t and y_t are continuous, then the algorithm should be stopped after the l_∞-diameter of the feasible sets of the sub-problems has fallen below a certain threshold to guarantee finite convergence.

5 Results

Table 1 shows results for four instances of the sizes problem (Jonsbraten et al. (1998)). In this problem, an item has to be produced in I different sizes. Demands for each size have to be satisfied in each of T time periods. The uncertainty lies in the unit production costs for each size (endogenous), which is resolved as soon as the size is produced for the first time, and the demands for each size in each time period (exogenous). Decisions $b_{i,t}$ represent whether size i is produced in period t or not, while y_t and x_t represent the amounts produced and those sold to satisfied demand, respectively. It is clear from Table 1 that the proposed algorithm generates better feasible solutions and smaller optimality gaps compared to a standard LP based branch and bound algorithm (ILOG CPLEX 9.0) in one order of magnitude less time. All results were obtained on a Pentium-IV, 2.4 GHz Linux machine.

The proposed algorithm is currently being tested on the gas field problem (Goel and Grossmann (2004b)). This method can also be applied to optimal design of process networks with uncertainty in yields (Goel and Grossmann (2004a)).

Acknowledgments

Financial support from the ExxonMobil Upstream Research Company is gratefully acknowledged. The authors are specially thankful to Amr El-Bakry and Eric Mulkay from ExxonMobil Upstream Research Company for providing guidance in the implementation of the branch and bound algorithm.

References

Caroe, C. C., Schultz, R., 1999. Dual decomposition in stochastic integer programming. Oper. Res. Lett. 24, 37–45.

Goel, V., Grossmann, I. E., 2004a. A class of stochastic programs with decision dependent uncertainty. Submitted for publication.

Goel, V., Grossmann, I. E., 2004b. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Computers and Chemical Engineering 28 (8), 1409–1429.

Guignard, M., Kim, S., 1987. Lagrangean decomposition - a model yielding stronger lagrangean bounds. Mathematical Programming 39 (2), 215–228.

Hooker, J., 2000. Logic-based methods for optimization: combining optimization and constraint satisfaction. New York : John Wiley and Sons.

Jonsbraten, T. W., 1998. Optimization models for petroleum field exploitation. Ph.D. thesis, Norwegian School of Economics and Business Administration.

Jonsbraten, T. W., Wets, R. J. B., Woodruff, D. L., 1998. A class of stochastic programs with decision dependent random elements. Annals of Operations Research 82, 83–106.

Nemhauser, G. L., Wolsey, L. A., 1988. Integer and combinatorial optimization. Wiley, New York, english Series: Wiley-Interscience series in discrete mathematics and optimization.

Sahinidis, N. V., 2004. Optimization under uncertainty: State-of-the-art and opportunities. Comput. Chem. Eng. 28 (6-7), 971–983.