Characterization and Analysis of Multistage Stochastic Programming Problems with Type II Endogenous Uncertainty to Develop Universally Applicable Bounding Approaches

Yasuhiro Shoji\*, Selen Cremaschi

Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA

yzs0108@auburn.edu, szc0113@auburn.edu

Abstract

Optimization challenges featuring endogenous uncertainties that unfold sequentially based on decisions are prevalent in the process industry. While multistage stochastic programming (MSSP) serves as a viable solution method, addressing real-world-scale problems poses computational challenges. This study examines and characterizes MSSP problems involving type II endogenous uncertainty, exploring the suitability of two heuristics: the absolute expected value solution (AEEV) and the generalized knapsack-problem-based decomposition algorithm (GKDA), developed by our research group. We applied the heuristics to eight MSSP problems, which revealed that both AEEV and GKDA may generate no feasible solutions when the MSSP problem lacks complete recourse. If the heuristics generate solutions, AEEV and GKDA provide 0 to 3 % and 0.5 to 45 % optimality gaps, respectively, with 0.003 to 0.3 % computational times compared to solving the MSSP problem.

**Keywords**: multistage stochastic program, endogenous uncertainty, heuristics

* 1. Introduction

Optimization challenges within the process industry frequently involve endogenous uncertainty (Goel & Grossmann, 2004; Tarhan & Grossmann, 2008; Solak et al., 2010). Goel & Grossmann (2006) classified endogenous uncertainty into two sub-classes: type I and type II, in which the probability distribution and realization time depend on the decisions, respectively. The constraints to enforce scenario indistinguishability before realization in type II endogenous uncertainty are referred to as non-anticipativity constraints (NACs). The exponential increase in the number of NACs with the growing number of scenarios leads to computational complexities when addressing real-world size problems (Apap & Grossmann, 2017). This challenge has prompted the development of various heuristic and decomposition solution approaches.

To tackle these issues, our research group has introduced two heuristics, AEEV (Zeng & Cremaschi, 2019) and GKDA (Zeng & Cremaschi, 2018), designed to generate feasible solutions for MSSP problems involving type II endogenous uncertainty. The heuristics decompose MSSP problems into deterministic sub-problems using expected values of uncertain parameters. They iterate constructing and solving sub-problems, judging scenario realizations, and determining recourse actions based on realized information until the end of the planning horizon. Universal applicability to any MSSP problem with type II endogenous uncertainty is among the anticipated properties of these heuristics besides the fast computational speed and providing tight bounds. However, the applicability of the heuristics depends on the characteristics of MSSP problems. We conducted an analysis of eight published problems using our heuristics. The paper compares the optimality gaps and computational times of successful instances where AEEV and GKDA provided a feasible solution. Then, it discusses the key properties limiting the applicability.

* 1. The framework of MSSP with type II endogenous uncertainty

MSSP problems involving type II endogenous uncertainty incorporate scenarios and a discretized time horizon. In these scenarios, here-and-now decisions progressively differentiate among potential outcomes over time. Wait-and-see decisions allow taking corrective action after the here-and-now decisions reveal the outcomes. The NACs ensure that the variables must take identical values when scenarios are indistinguishable. A general formulation of MSSP with endogenous uncertainties is shown in Equations (1) – (9). Note that the model is an extension of the formulation in Zeng & Cremaschi (2019).

|  |  |
| --- | --- |
|  | (1) |
| s.t. |  |
|  | (2) |
|  | (3) |
|  | (4) |
|  | (5) |
|  | (6) |
|  | (7) |
|  | (8) |
|  | (9) |

The objective function (Eq. 1) comprises the contribution and the probability of scenario , where and denote predetermined deterministic and uncertain parameters, respectively. The model involves four variables, namely , , and , with uncertainty resource , time horizon , and scenario . The two here-and-now decision variables, binary decision variables and continuous or integer decision variables , are forced to take identical values by both the initial NACs (Eqs. 4 and 5) and the conditional NACs (Eq. 6) until scenario pairs and become distinguishable. Conditional NACs constrain the wait-and-see decision variables (Eq. 7). Equations 2 and 3 are scenario-specific inequality and equality constraints, such as demand and capacity constraints, and material balances. The conditional NACs (Eqs. 6 and 7) are enforced using binary variable , determined based on here-and-now decisions via the expression (Eq. 8), where takes a value of one if scenarios and are indistinguishable and takes a value of zero otherwise.

* 1. Heuristics
     1. Absolute Expected Value Solution Approach (AEEV)

A feasible solution was produced by AEEV (Zeng & Cremaschi, 2019) through the transformation of an MSSP problem into two distinct types of sub-problems: deterministic expected value sub-problems (s) and recourse deterministic expected value sub-problems (s). s and s are responsible for determining here-and-now decisions and the corresponding recourse actions. AEEV begins with the construction of a at . The endogenous uncertain parameters are replaced with their expected values of all outcomes for the at . AEEV solves the and stores the values of here-and-now decision variables of while discarding the rest of the solution. Based on the recommended here-and-now decisions and realized outcomes, AEEV judges which scenarios can be differentiated. If realization occurs and the MSSP formulation has wait-and-see decision variables, s are generated and solved. AEEV only stores the values of wait-and-see decision variables of *t* = 1 and repeats constructing and solving s and s at each time. Equations 10 - 16 are general formulation.

|  |  |
| --- | --- |
|  | (10) |
| s.t. |  |
|  | (11) |
|  | (12) |
|  | (13) |
|  | (14) |
|  | (15) |
|  | (16) |

The objective function (Eq. 10) provides the optimum objective value of under the expected values of uncertain parameter associated with sub-problem . The expected values are updated in each sub-problem as scenarios are distinguished. Since AEEV decomposes scenarios, it removes NACs and indicator variable (Eqs. 4 to 8) and only has scenario-specific constraints (Eqs. 11 and 12). AEEV fixes all decisions made before , using fixed variables , , , and (Eqs. 13 to 16). General formulation is similar to the formulation. Equations (10) to (12), (15), and (16) are identical to those of the formulation, while Equations (13) and (14) are replaced with (17) and (18). The fixes here-and-now decisions (, ) when because here-and-now decisions must be the same after realization at , while wait-and-see decisions can be different.

|  |  |
| --- | --- |
|  | (17) |
|  | (18) |

* + 1. Generalized Knapsack-Problem based Decomposition Algorithm (GKDA)

Zeng et al. (2018) transformed an MSSP problem into multiple sub-problems called knapsack sub-problems (KSPs) through the decomposition of the scenario and time indices, resulting in the generation of a feasible solution. GKDA converts the remaining indexes into the combination of them as an item set for KSPs and enumerates all admissible items into eligible item lists. GKDA assigns each item in the eligible item list with a computed item value, and both the eligible item list and the item values undergo updates with each KSP based on the prior KSP solutions. The solutions to KSPs determine both the here-and-now and wait-and-see decisions. At , GKDA generates the eligible item list and item values, then constructs and solves a KSP with the expected uncertain parameters over all scenarios. GKDA stores the values of here-and-now decision variables and judges distinguishable scenarios according to the realized outcomes. If realization occurs and the MSSP formulation has wait-and-see decision variables, GKDA generates and solves new KSPs with the expected uncertain parameters to find the recourse actions. After , GKDA persistently updates the item list and corresponding item values, construct KSPs, and figures them out until the end of the predetermined time. Equations 19 - 22 are the general KSP formulation.

|  |  |
| --- | --- |
|  | (19) |
| s.t. |  |
|  | (20) |
|  | (21) |
|  | (22) |

The objective function (Eq. 19) calculates the optimum objective value of KSPs under the expected uncertain parameter associated with KSP . The values of expected uncertain parameters are updated as scenarios are distinguished. As a result of GKDA’s decomposition of scenarios and times in the MSSP formulation, the resulting expression retains solely scenario-specific constraints (Eq. 20 and 21) in the formulation with the set of items . Except for sequencing constraints, these scenario-specific constraints work as KSP weight constraints. The sequencing constraints are removed from the formulation, and the eligible item lists substitute for them to abide by the rules the removed sequencing constraints define. Equations 23 and 24 are added to the formulation to fix here-and-now decisions to determine the wait-and-see decisions.

|  |  |
| --- | --- |
|  | (23) |
|  | (24) |

The main difference between AEEV and GKDA is that GKDA lacks time information while AEEV retains it. The item values help to find better feasible bounds by compensating for the information loss. Our group developed three approaches to estimate the item values of KSPs (Zeng et al., 2018). *The straightforward approach* transfers the objective function coefficients of the general MSSP formulation as the item values of KSPs. *The maximum potential gain approach (MPGA)* applies to maximization problems where the decision maker cannot earn revenue, gains, or returns until a specific stage or outcome in the planning horizon. *The capital recovery factor (CRF) approach* applies to problems that involve capital investment decisions. The approach annualizes the capital investment made only once in the planning horizon.

* 1. Case study results and discussion

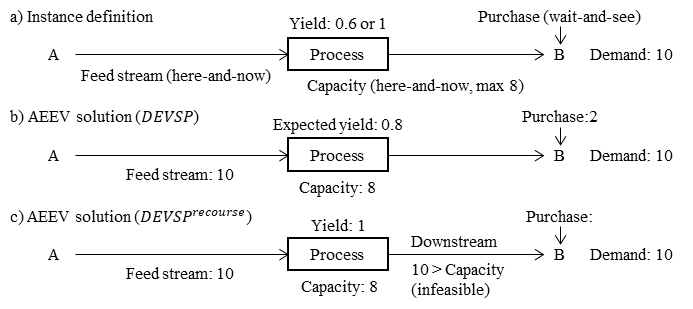
We analyze eight published MSSP problems with type II endogenous uncertainty by applying our heuristics. Table 1 compares the optimality gaps and computational times of successful instances obtained by AEEV and GKDA. We chose Pyomo version 6.4.0 and CPLEX version 20.10 to formulate and solve all instances, employing 48 processors on a node from Auburn University Easley Cluster. The definition of the optimality gap is the relative error between the MSSP and our heuristic solutions.

**Table 1.** Optimality gaps and computational times of successful instances



In most instances, AEEV provides tighter bounds than GKDA, while the computational time is shorter for GKDA than AEEV. The result implies that the optimality gap tends to be tight while the computational time tends to be long as the approaches retain more information in their formulation. In simpler terms, AEEV provides tighter bounds with longer computational time as it decomposes solely the scenario indexes and preserves the time information in the MSSP formulation, compared to GKDA which removes both scenario and time information by decomposing their indexes.

The analysis revealed two challenges for the general applicability of the heuristics. The first challenge is that AEEV and GKDA may fail to generate a feasible solution if the MSSP model does not have complete recourse, which ensures no outcome produces infeasible results. The example in Figure 1 is used to illustrate this challenge.



**Figure 1.** A process network illustrating the infeasibility of a sub-problem after realization.

The process in Figure 1 produces B from A with uncertain yield (0.6 or 1). The here-and-now decisions are the feed stream's flow rate and the installed process's capacity, and one wait-and-see decision is the amount of B purchased. Suppose the demand for B is 10, and the maximum capacity expansion is 8. AEEV first solves (Figure 1b) to determine here-and-now decisions with the expected yield (0.8). The installation of the process realizes actual yield, 0.6 or 1, and AEEV solves with fixed here-and-now decisions, i.e., 10 for the feed stream's flow rate and 8 for the capacity. Focusing on the outcome where the yield is 1 (Figure 1c), it is evident that the downstream flow rate is infeasible since the flow rate (10) exceeds the capacity. GKDA generates infeasible KSPs for the same reason. Thus, the updated uncertain parameter after realization does not always satisfy the constraints satisfied before realization without complete recourse. Table 2 complies with the studied MSSP problems, which reveals that most models do not have complete recourse. The second challenge arises from the lack of time information in GKDA. The item value estimation approaches to compensate for time information loss apply to most optimization problems (Table 2) except for the demand-side response planning problem. This problem has a time lag between decisions and implementations, which cannot be captured within the current GKDA.

**Table 2.** Complete recourse availability in MSSP and item value estimation approach in GKDA

|  |  |  |
| --- | --- | --- |
| Optimization problem and authors | Complete recourse | Item value estimation in GKDA |
| Size (Jonsbråten et al., 1998) | No | Straightforward |
| Oil/gas-field development (Goel & Grossmann, 2004) | No | CRF |
| Process network synthesis (Tarhan & Grossmann, 2008) | No | CRF |
| Open-pit mine production scheduling (Boland et al., 2008) | No | Straightforward |
| Clinical trial planning (Colvin & Maravelias, 2008) | Yes | MPGA |
| R&D project portfolio (Solak et al., 2010) | No | MPGA |
| Vehicle routing (Khaligh & MirHassani, 2016) | No | Straightforward |
| Demand-Side Response Planning (Giannelos et al., 2018) | No | None |

* 1. Conclusions and future directions

We examined eight MSSP problems involving Type II endogenous uncertainty with AEEV and GKDA. For the cases with feasible solutions, AEEV provided tighter bounds than GKDA, while the computational time was shorter for GKDA. The analyses revealed two challenges for expanding the applicability of AEEV and GKDA. Future studies will focus on AEEV as it is more universally applicable and yields tighter bounds than GKDA. We will investigate two approaches. The first approach will add complete recourse to the MSSP problems, eliminating infeasibility without changing the optimum. The second approach will use weighted uncertain parameters instead of expected uncertain parameters, improving feasibility without changing the MSSP formulation.

References

Boland, N., Dumitrescu, I., & Froyland, G., 2008. A multistage stochastic programming approach to open pit mine production scheduling with uncertain geology. Optimization online, 1-33.

Colvin, M. & Maravelias, C. T., 2008. A stochastic programming approach for clinical trial planning in new drug development. Comput. & Chem. Eng., 32, 2626-2642.

Goel, V. & Grossmann, I. E., 2004. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Comput. & Chem. Eng., 28(8), 1409-1429.

Goel, V. & Grossmann, I. E., 2006. A class of stochastic programs with decision dependent. uncertainty Math. Program., 108(2-3), 355-394.

Jonsbråten, T.W., Wets, R. J-B., & Woodruff, D.L., 1998. A class of stochastic programs with decision dependent random elements. Ann. Oper. Res., 82, 83-106.

Khaligh, F. H. & MirHassani, S. A., 2016. A mathematical model for vehicle routing problem under endogenous uncertainty. Int. J. Prod. Res., 54:2, 579-590.

Solak, S., Clarke, J. P. B., Johnson, E. L., & Barnes, E. R., 2010. Optimization of R&D project portfolios under endogenous uncertainty. EJOR, 207(1), 420-433.

Tarhan, B. & Grossmann, I. E., 2008. A multistage stochastic programming approach with strategies for uncertainty reduction in the synthesis of process networks with uncertain yields. Comput. & Chem. Eng., 32(4-5), 766-788.

Zeng, Z., Christian, B., & Cremaschi, S., 2018. A generalized knapsack-problem based decomposition heuristic for solving multistage stochastic programs with endogenous and/or exogenous uncertainties. Ind. Eng. Chem. Res., 57(28), 9185-9199.

Zeng, Z. & Cremaschi, S., 2019. A general primal bounding framework for large-scale multistage stochastic programs under endogenous uncertainties. Chem. Eng. Res. Des., 141, 464-480.