The L-shaped Method for Large-scale Mixed-integer Waste Management Decision Making Problems

Jakub Kůdela[a,*], Pavel Popela[a], Radovan Šomplák[b], Martin Málek[a], Adam Rychtář[a], Dušan Hrabec[c]

[a]Institute of Mathematics, Faculty of Mechanical Engineering, Brno University of Technology – VUT Brno, Technická 2896/2, 616 69 Brno, Czech Republic
[b]Institute of Process Engineering, Faculty of Mechanical Engineering, Brno University of Technology – VUT Brno, Technická 2896/2, 616 69 Brno, Czech Republic
[c]Institute of Mathematics, Faculty of Applied Informatics, Tomas Bata University in Zlín, Nad Stráněmi 4511, 760 05 Zlín, Czech Republic
jakub.kudela89@gmail.com

It is without a doubt that deciding upon strategic issues requires us to somehow anticipate and consider possible variations of the future. Unfortunately, when it comes to the actual modelling, the sheer size of the problems that accurately describe the uncertainty is often extremely hard to work with. This paper aims to describe a possible way of dealing with the issue of large-scale mixed integer models (in term of the number of possible future scenarios it can handle) for the studied waste management decision making problem. The algorithm is based on the idea of decomposing the overall problem alongside the different scenarios and solving these smaller problems instead. The use of the algorithm is demonstrated on a strategic waste management problem of choosing the optimal sites to build new incineration plants, while minimizing the expected cost of waste transport and processing. The uncertainty was modelled by 5,000 scenarios and the problem was solved to high accuracy using relatively modest means (in terms of computational power and needed software).

1. Introduction

In the last few years, there have been several papers devoted to the study of optimal strategic decision making under uncertain conditions and to the economic and environmental impact of these decisions. The strategic aspect of the models comes from the situation when some decisions (very often in terms of integer or binary variables) have to be made ahead of time. Poplewski and Jezowski (2009) approached the issue of uncertainty by constructing a so called flexible network. When dealing with inventory management, You and Grossmann (2011) derived a piecewise linear approximation for a MINLP problem and modelled the uncertainty with two scenarios. In supply chain management, Šomplák et al. (2013) presented another type of mixed integer models, where uncertainty in data plays a crucial role. Similar problem, but with a multi-period formulation was constructed by Rodriguez et al. (2013). Nemet et al. (2013) handled the problem of multi-period plant production, using three different price projections to model the randomness followed with an even more complex nonlinear mixed integer model as described in Nemet et al. (2015). The number of considered projections rose to five. A different approach to solving these mixed integer problems is by using nature inspired metaheuristics as the ones in Viktorin et al. (2016).

A more risk-averse approach can be found in Rad and Rashtchain (2016), where a similar MILP problem is considered but the number of possible scenarios was almost 400. The shift towards increasing the number of scenarios included in the computations is quite palpable and understandable. It is unwise to expect a good approximation of the impact our decision will have (in terms of the value of the chosen objective) if the number of possibilities included in the computation is not capable of capturing the "nooks and crannies" (most notably the tail properties) of the underlying distributions.

The way of approaching problems with large number of scenarios used in this paper is based upon the one presented in a seminal book on stochastic programming by Birge and Louveaux (1997). The methodology of

resolving if the chosen number of scenarios is enough to accurately describe the problem is derived from King and Wallace (2012).

2. Two-stage stochastic programming models

In the last two decades, the methodology proposed by stochastic programming earned its place among the most widely used approaches in decision making and operations research problems. The particular model of interest in this passage is the two-stage linear stochastic model. It describes a very common decision-making situation, when the strategic decision, that must be made, depends on uncertain parameters. After the values of random parameters become known, some other decisions (recourse decisions) can be made.

Since some parameters are random, questions arise about how to appropriately formulate our objective and the constraints in our model. The recommended way of formulating the objective is using a risk functional, such as expected value or the more risk averse CVaR (Conditional Value at Risk). The constraints must either hold almost surely, or at least with a probability proscribed beforehand.

The most common way of dealing with random parameters is sampling or scenario generation – Monte Carlo methods or importance sampling being the favourable techniques. The obvious problem these methods bring is that in order to get satisfactory results (especially when one is using a quantile-based functional or probabilistic constraints), the number of scenarios needed grows to unmanageable highs. An approach of deciding an appropriate number of scenarios was proposed in King and Wallace (2012).

The model is going to be a one with expected value as its risk functional and constraints that must hold almost surely. This means that after the generation of scenarios, the second-stage constraints (the ones with random parameters and recourse decisions) must hold for all scenarios. The problem then takes the following general decomposable form:

$$\min_{x, y_s} \quad c^\mathsf{T} x + \sum_{s \in S} p_s q_s^\mathsf{T} y_s \tag{1}$$

$$\text{s.t.} \quad Ax = b, \tag{2}$$

$$T_s x + W_s y_s = h_s, \qquad \forall s \in S, \tag{3}$$

$$x, y_s \geq 0, \qquad \forall s \in S. \tag{4}$$

The objective Eq(1) to minimize is the cost that corresponds to taking the first-stage decision x (with associated cost vector c) plus the expected cost from recourse second-stage decisions y_s (with associated cost vector q_s) that depend on the particular realization of the random parameters in the scenario $s \in S$, where p_s is the probability of the scenario. The constraint Eq(2) restricts the first-stage decision (using a matrix A and a vector of right hand sides b), the constraints Eq(3) ties together the first-stage and the second-stage decisions (by matrices T_s, W_s and a vector of right hand sides h_s) and Eq(4) requires the decision variables to be positive. Eq(3) and Eq(4) are prescribed in the almost surely fashion (i.e., must hold for all scenarios). The sizes of the matrices and vectors are such that the products and equalities are coherent. The resulting problem is an "ordinary" linear program, but as the number of scenarios grows, even this in present terms rather simple problem quickly grows outside the realm of our computational powers.

3. The L-shaped method

This section provides a description of a powerful method for solving large scale two-stage stochastic linear problems of the type Eq(1) - Eq(4), i.e. exactly the ones that require a large number of scenarios (how large a number is this method able to handle is going to be the subject of the case study).

The method is based on developing an approximation (Θ) of the expected value, using Lagrangian dual methodology and simplex multipliers. This should not be an issue since most current solvers will provide these values automatically. The computational superiority of the algorithm stems from its scenarios decomposition – for each scenario, the second-stage problem is solved separately, computing the multipliers and updating the approximation of the expected value. Another strength comes from possible parallelization – for multi-core processors, the tasks of computing optimal solutions for each scenario can be divided between the individual cores, providing additional speedup. The overall procedure can be summarized as follows:

$$\min_{x, \Theta} \quad c^\mathsf{T} x + \Theta \tag{5}$$

$$\text{s.t.} \quad Ax = b, \tag{6}$$

$$D_l x \geq d_l, \qquad l = 1, \dots, r_1, \tag{7}$$

$$E_l x + \Theta \geq e_l, \quad l = 1, \dots, r_2, \tag{8}$$

$$\Theta \in \mathbb{R}, \; x \geq 0. \tag{9}$$

Step 1.   Set $r_1 = r_2 = \nu = 0$ and ε sufficiently small.

Step 2.   Set $\nu = \nu + 1$. Solve linear program (called a master problem) Eq(5)-Eq(9).
Let (Θ^ν, x^ν) be an optimal solution. If no constraint Eq(8) is present, $\Theta^\nu = -\infty$ and is not considered in the computation.

Step 3.   For $\forall s \in S$ solve the program

$$\min_{y, v^+, v^-} \; w = e^T v^+ + e^T v^- \tag{10}$$

$$\text{s.t.} \quad W_s y + I v^+ - I v^- = h_s - T_s x^\nu, \tag{11}$$

$$y, v^+, v^- \geq 0, \tag{12}$$

where $e^T = (1, \dots, 1)$, until, for some s, the optimal value $w > 0$. In this case, let σ^ν be the associated simplex multipliers and define

$$D_{r_1+1} = (\sigma^\nu)^T T_s, \qquad d_{r_1+1} = (\sigma^\nu)^T h_s, \tag{13}$$

to generate a constraint (called a feasibility cut) of type Eq(7). Set $r_1 = r_1 + 1$ and return to Step 2. If for all $s \in S, w = 0$, go to Step 4.

Step 4.   For $\forall s \in S$ solve the program

$$\min_y \; q_s^T y \tag{14}$$

$$\text{s.t.} \quad W_s y = h_s - T_s x^\nu, \tag{15}$$

$$y \geq 0. \tag{16}$$

Let λ_s be the simplex multipliers associated with the optimal solution of problem s of type Eq(14) - Eq(16). Define

$$E_{r_2+1} = \sum_{s \in S} \lambda_s T_s, \qquad e_{r_2+1} = \sum_{s \in S} \lambda_s h_s. \tag{17}$$

Let $w^\nu = e_{r_2+1} - E_{r_2+1} x^\nu$. If $\Theta^\nu + \varepsilon \geq w^\nu$, stop; x^ν is an ε-optimal first-stage solution. Otherwise, set $r_2 = r_2 + 1$, add a so-called optimality cut to Eq(8), and return to Step 2.

Several extensions to this algorithm have been proposed, a contemporary computational analysis of the different variations can be found in Zverovich et al. (2012). A convex programming enhancement of the method, using an oracle with on-demand accuracy, was proposed in Wolf et al. (2014).

4. Case study

As one of our main areas of applied research is in waste management decision making problems, the case study is going to investigate the use of the aforementioned algorithm in such a setting. The goal is to decide on the location of new incinerator plants – this is a highly-debated topic, since the changes in legislature in our country restrict the use of landfills so other means of dealing with waste are anticipated. There are already four of these plants working in the Czech Republic, but if they were to process all the produced waste, their capacity would be highly insufficient.

The problem in question is by no means small, there are 206 micro-regions producing waste and 30 possible locations for new incinerator plants (the four already functioning are not included in the 30 proposed). The incidence matrix describing possible flows (either between micro-regions or between a micro-region and a plant) has 2,072 columns.

The objective function to be minimized comprises the transportation costs, processing costs and costs associated with situations when the build incinerator is not running at full capacity (a building cost of the facility could be used instead, but the operation/capacity costs are easier to work with).

There are two sources of randomness considered. The first one is the production of waste in different micro-regions – based on expert forecasts an interval on which the distribution of waste production is uniform was constructed for each micro-region. The second random parameter is the processing cost in the incinerator plants – for each such plant different triangular distribution is constructed (based, once again, on the collected data). The costs that correspond to incinerator plants not running at full capacity are estimated as 10 % of the processing costs (but could be modelled as separate random parameters). The transportation costs on each route and capacities of the proposed incinerators are considered fixed.

4.1 Model formulation

The situation described above is summarized in the following network model:

$$\min_{\Delta, x^s, y^s, u^s} \sum_{s \in S} p^s (\sum_{i \in I} f_i^s y_i^s + \sum_{i \in I} g_i^s u_i^s + \sum_{j \in J} c_j x_j^s) \tag{18}$$

$$\text{s.t.} \quad \sum_{j \in J} a_{ij} x_j^s - y_i^s = -q_i^s, \quad \forall i \in I, \forall s \in S, \tag{19}$$

$$y_i^s + u_i^s = \Delta_i o_i, \qquad \forall i \in I, \forall s \in S, \tag{20}$$

$$x^s, y^s, u^s \geq 0, \qquad \forall s \in S, \tag{21}$$

$$\Delta_i \in \{0,1\}, \qquad \forall i \in I. \tag{22}$$

The index i describes the nodes (micro-regions and plants), the index j the arcs (columns of the incidence matrix) and the index s the different scenarios for random parameters. The decision variables are Δ for the integer first-stage decision on building the facility and the second-stage positive variables are x^s, y^s, u^s the amount of waste transported on different arcs, the amount processed by different facilities and the facility capacities left unused, respectively. The parameters f_i^s, g_i^s, c_j are the processing costs, under-capacity costs and transportation costs, respectively. The Eq(19) describes the balances in each node – the amounts entering and leaving the node on different arcs (distinguished by the incidence matrix value a_{ij}), the amounts processed and produced (q_i^s) must be in balance. The Eq(20) guarantees, that processing takes place only in facilities, that have been selected, and that if the full capacity (o_i) is not met, the penalty is applied. An additional condition is placed securing that the four incinerator plants that are already processing waste will always be selected.

Although the L-shaped method as described in the previous section dealt with the linear case and real first-stage decision variables, the same procedure can be applied even if in the first-stage (master) problem is mixed-integer (our case) or convex. To clarify, the individual parts of the method will be shown in detail. The first-stage problem (in Step 2.):

$$\min_{\Delta, \Theta} \quad \Theta \tag{23}$$

$$\text{s.t.} \quad D\Delta \geq d, \tag{24}$$

$$E\Delta + \Theta \geq e, \tag{25}$$

$$\Theta \in \mathbb{R}, \Delta_i \in \{0,1\}, \qquad \forall i \in I. \tag{26}$$

The second-stage problems (in Step 4.):

$$\min_{x,y,u} \quad \sum_{i \in I} f_i^s y_i + \sum_{i \in I} g_i^s u_i + \sum_{j \in J} c_j x_j \tag{27}$$

$$\text{s.t.} \quad \sum_{j \in J} a_{ij} x_j - y_i = -q_i^s, \qquad \forall i \in I, \tag{28}$$

$$y_i + u_i = \Delta_i^y o_i, \qquad \forall i \in I, \tag{29}$$

$$x, y, u \geq 0, \tag{30}$$

where the first-stage decision Δ_i^y is considered fixed. The cutting hyperplanes in Eq(24) and Eq(25) are computed as:

$$D_{r_1+1} = (\sigma^v)^\top \begin{pmatrix} 0 \\ -\text{diag}(o) \end{pmatrix}, \qquad d_{r_1+1} = (\sigma^v)^\top \begin{pmatrix} -q^s \\ 0 \end{pmatrix}, \tag{31}$$

$$E_{r_2+1} = \sum_{s \in S} (\lambda_s)^\top \begin{pmatrix} 0 \\ -\text{diag}(o) \end{pmatrix}, \qquad e_{r_2+1} = \sum_{s \in S} (\lambda_s)^\top \begin{pmatrix} -q^s \\ 0 \end{pmatrix}, \tag{32}$$

where the operation diag transforms a vector into a diagonal matrix and the notation 0 is overloaded to mean both a zero vector and a zero matrix, depending on the context.

4.2 Results of the computation

As indicated earlier, the aim is to use a high enough number of scenarios to get an accurate approximation of the problem. There were 5,000 scenarios used to obtain the optimal first-stage decision on facility construction. This results in the considered problem having over 13 M second-stage variables. Furthermore, two different cases were studied – one when the production of waste is uncorrelated among the individual micro-regions and one with positive correlation (close to 1). To further support the use of stochastic programming, an additional model using only one scenario (having only 2,602 second-stage variables) with average values for each parameter is constructed and consequently tested on all the scenarios (i.e., the average value scenario was

used to decide on the facility location and based on this decision found the best recourse in terms of the second-stage variables for all the scenarios).

The results are best described using the Figure 1 and Figure 2, where the distribution of optimal values for individual scenarios is depicted as a histogram. In Figure 1a) and Figure 2a) is the distribution using 5,000 scenarios for the strategic decision and in Figure 1b) and Figure 2b) is the result when only the average values were used (where the value 0 describes a situation when the first-stage decision resulted an infeasible problem for the second-stage). The number of scenarios, where the solution from an easier problem is used and as a result has an infeasible second-stage, is understandably high – this model has no way of adjusting to varying parameters and its usage is meant to illustrate the dangers of this approach.

For the 5,000 scenario models, out of the 30 possible locations, 18 and 19 were selected and the average objective function value (i.e., the value being minimized in our model) was 2.716×10^9 CZK (approx. 100.5 M EUR) and 2.754×10^9 CZK (101.9 M EUR) for the uncorrelated and correlated case, respectively. A natural question to ask is whether these numbers are an accurate approximation of the expected value for each of the cases. To test it, the methodology proposed in King and Wallace (2012) was used and 10,000 new scenarios were generated. The strategic first-stage decision from the 5,000 cases was used, for all the 10,000 scenarios, the optimal second-stage decisions were found and the average was computed. The difference between the optimal values was 0.013 %. This level of accuracy can be considered quite satisfactory.

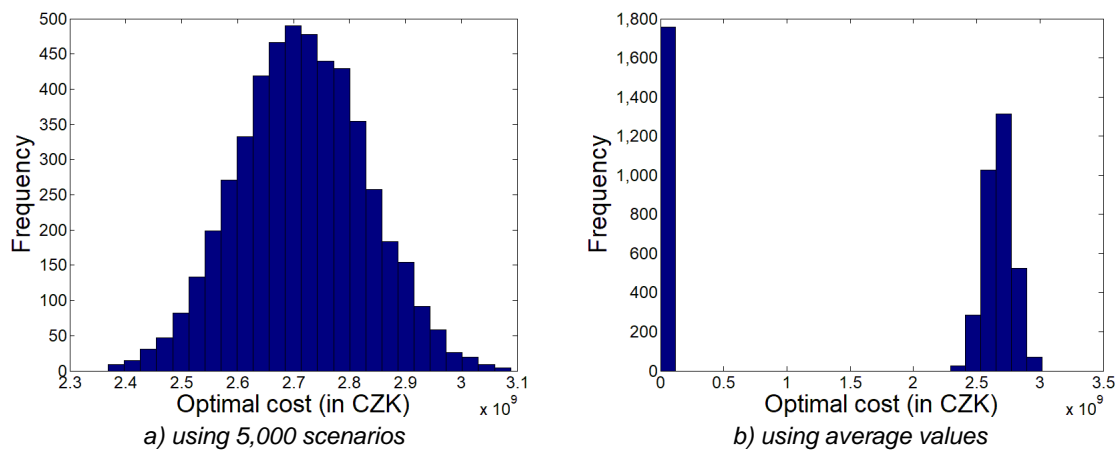*a) using 5,000 scenarios*   *b) using average values*

Figure 1: Distribution of optimal objective function values for different scenarios (a) using 5,000 scenarios and (b) using average values, the uncorrelated case
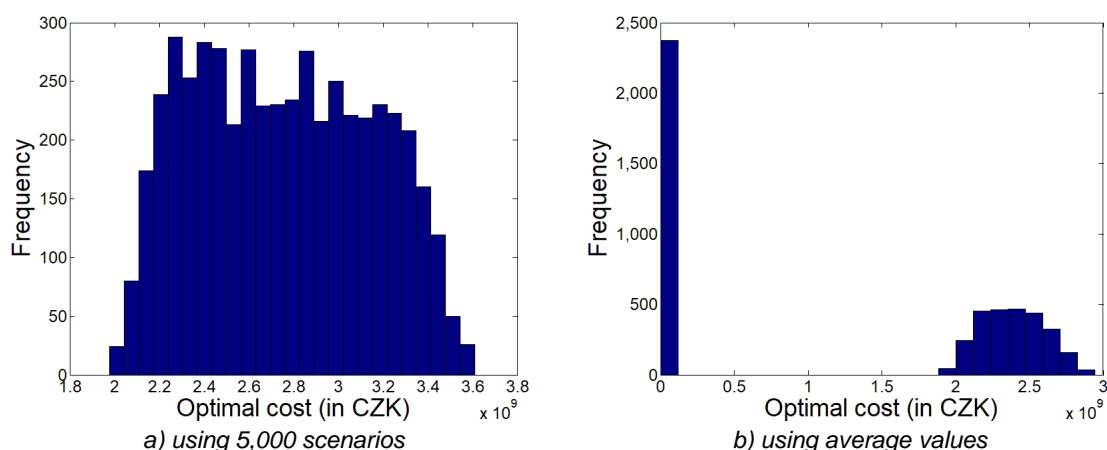

a) using 5,000 scenarios   *b) using average values*

Figure 2: Distribution of optimal objective function values for different scenarios (a) using 5,000 scenarios and (b) using average values, the correlated case

The computations were carried out on an ordinary machine (3.2 GHz i5-4460 CPU, 16 GB RAM) using basic optimization library embedded in MATLAB (linprog and intlinprog functions). The computation of each of the

cases took about 50 min (without any parallelization). Using more powerful solvers, better machines and taking advantage of the possible parallelization can drastically reduce the computational time as shown in Zverovich et al. (2012).

5. Conclusions

As it is with most of the advancements in mathematical disciplines, their adoption to the areas of applied research, where they can be utilized, takes some time. The goal of this paper was to describe a decomposition-based method that allows for much higher levels of description and accuracy in optimization models where uncertainty plays a critical part. It is hard to overestimate the impact a good approximation of the uncertainty and the subsequent proper use of this approximation has on the results of optimization models. A general algorithmic procedure, that is able to deal with such situations, was described and the particular case of deciding upon the construction of incineration plant was studied and discussed. Although the technique for solving the two-stage models is well known (at least in the stochastic programming community), it is still underused in applications dealing with stochastic optimization models. The paper presents the impact the possibility of dealing with large models has on the "robustness" of the strategic decision one makes, which is best summarized by Figure 1 and Figure 2. The authors of this paper are not aware of any such contribution in the area of waste management and hope this paper will encourage researcher dealing with decision making under uncertainty to utilize the aforementioned method. Further enhancement both in the model of the strategic waste management decision problem and in the decomposition algorithm itself are possible and will be the subject of our future work. Moreover, the constant rise of computational power will allow us to consider and compute problems of even larger magnitude than the one presented.

References

Birge J.R., Louveaux F., 1997, Introduction to stochastic programming, New York, USA: Springer. ISBN: 978-0-3879-8217-5.

King A.J., Wallace S.W., 2012, Modelling with stochastic programming, New York, NY, USA: Springer. ISBN: 978-0-387-87817-1.

Nemet A., Klemeš J.J., Kravanja Z., 2013, Life Span Production Plant Optimisation under Varying Economic Conditions. Chemical Engineering Transactions, 35, pp.103-108, doi:10.3303/CET1335017.

Nemet A., Klemeš J.J., Kravanja Z., 2015, Designing a Total Site for an entire lifetime under fluctuating utility prices. Computers and Chemical Engineering, 72, 159-182, doi:10.1016/j.compchemeng.2014.07.004.

Poplewski G., Jezowski J., 2009, Optimisation based approach for designing flexible water usage network. Chemical Engineering Transactions, 18, pp.409-414, doi:10.3303/CET0918066.

Rad A., Rashtchain D., 2016, A Risk-based MILP Approach for Optimal Placement of Flammable Gas Detectors. Chemical Engineering Transactions, 53, pp.145-150, doi:10.3303/CET1653025.

Rodriguez M.A., Vecchietti A.R., Grossmann I.E., Harjunkoski I., 2013, Supply chain design and inventory management optimization in the motors industry. Chemical Engineering Transactions, 32, pp.1171-1176, doi:10.3303/CET1332196.

Šomplák R., Procházka V., Pavlas M., Popela P., 2013, The logistic model for decision making in waste management. Chemical Engineering Transactions, 35, pp.817-822, doi:10.3303/CET1335136.

Viktorin A., Hrabec D., Pluháček M., 2016, Multi-chaotic differential evolution for vehicle routing problem with profits. Proceedings - 30th European Conference on Modelling and Simulation, ECMS 2016, pp.245-251, ISBN: 978-099324402-5.

Wolf C., Fabian C.I., Koberstein A., Suhl L., 2014, Applying oracles of on-demand accuracy in two-stage stochastic programming: A computational study. European Journal of Operational Research 239(2), pp.437-448, doi:10.1016/j.ejor.2014.05.010.

You F., Grossmann I.E., 2011, Stochastic inventory management for tactical process planning under uncertainties: MINLP models and algorithms. AIChE Journal, 57(5), pp.1250-1277, doi:10.1002/aic.12338.

Zverovich V., Fabián C.I., Ellison E.F.D., Mitra G., 2012, A computational study of a solver system for processing two-stage stochastic LPs with enhanced Benders decomposition. Mathematical Programming Computation, 4(3), pp.211-238, doi:10.1007/s12532-012-0038-z.