Recourse problem of the 2-stage robust location transportation problem

V. Gabrel, C. Murat, N. Remli, M. Lacroix [1]

Lamsade
Université Paris-Dauphine
Place du Maréchal De Lattre de Tassigny, 75775 Paris Cedex 16, France

Abstract

In this paper, we are interested in the recourse problem of the 2-stage robust location transportation problem. We propose a solution process using a mixed-integer formulation with an appropriate tight bound.

Keywords: Location transportation problem, robust optimization, mixed-integer linear programming.

1 Introduction

Robust optimization is a recent methodology for handling problems affected by uncertain data, and where no probability distribution is available. In robust optimization two decisional contexts are considered for taking decision under uncertainty. The first one is the *single-stage* context where the decision-maker has to select a solution before knowing the realization (values) of the uncertain parameters. Generally, the single-stage approaches provide the worst

[1] Email: {gabrel,murat,remli,lacroix}@lamsade.dauphine.fr

case solutions (Soyster [9]) that are very conservative and far from optimality in real-world applications. The second approach concerns the *multi-stage* context (or dynamic decision-making) where the information is revealed in stages. Indeed, a part of the decisions must be taken before actual realization of the uncertain parameters, and another part, called *recourse decisions* is taken when the information is known. The multi-stage approach was firstly introduced by Ben-Tal et. al. [2], and initial focus was on two-stage decision making on linear programs with uncertain feasible set. Note that the formulations obtained following this approach are generally untractable.

In this paper, we are interested in a robust version of the location transportation problem with an uncertain demand using a 2-stage formulation. Recently, Atamturk and Zhang [1] used a two-stage robust optimization in network flow and design problem to obtain a good approximation of the robust solutions. Furthermore, Thiele et. al. [10] describe a two-stage robust approach to address general linear programs affected by uncertain right hand side. The robust formulation they obtained is a convex (not linear) program, and they propose a cutting plane algorithm to exactly solve the problem. Indeed, at each iteration, they have to solve an NP-hard recourse problem on an exact way, which is time-expensive. Here, we go further in the analysis of the recourse problem of the location transportation problem, in particular we define a tight bound for the mixed-integer reformulation.

The paper is organized as follows: in Section 2, the nominal location transportation problem is introduced and its corresponding 2-stage robust formulation. A mixed integer program is then proposed in Section 3 to solve the quadratic recourse problem with a tight bound. Finally, in Section 4, the results of numerical experiments are discussed.

2 Robust location transportation problem

We consider the following location transportation problem: a commodity has to be transported from each of m potential sources, to each of n destinations. The sources capacities are $C_i,\ i = 1, \dots, m$ and the demands at the destinations are $\beta_j,\ j = 1, \dots, n$. To guarantee feasibility, we assume that the total sum of the capacities at the sources is greater than or equal to the sum of the demands at the destinations. The fixed and variable costs of supplying from source $i = 1, \dots, m$ are f_i and d_i, respectively. The cost of transporting one unit of the commodity from source i to destination j is μ_{ij}. The goal is to determine which sources to open (r_i), the supply level y_i and the amounts t_{ij} to be transported such that the total cost is minimized. The mathematical formulation of the location transportation problem is the following linear program, (T):

$$(T)\begin{cases} \min \sum_{i=1}^m d_i y_i + \sum_{i=1}^m f_i r_i + \sum_{i=1}^m \sum_{j=1}^n \mu_{ij} t_{ij} & \\ \text{s.t. } \sum_{j=1}^n t_{ij} \le y_i & i = 1 \dots m \\ \quad \sum_{i=1}^m t_{ij} \ge \beta_j & j = 1 \dots n \\ \quad y_i \le C_i r_i & i = 1 \dots m \\ \quad y_i,\ t_{ij} \ge 0,\ r_i \in \{0,1\} & i = 1 \dots m,\ j = 1 \dots n \end{cases}$$

In case of uncertainty on the demands, we model each demand β_j by the interval $[\overline{\beta}_j - \hat{\beta}_j, \overline{\beta}_j + \hat{\beta}_j]$, where $\overline{\beta}_j$ represents the nominal value of β_j, and $\hat{\beta}_j \ge 0$ its maximum deviation. We denote (T^β) the location transportation problem for a given $\beta \in [\overline{\beta} - \hat{\beta}, \overline{\beta} + \hat{\beta}]$, with a nonempty feasible set. Finally, we denote $Z(T^\beta)$ the optimal value (bounded value) of (T^β) for a given β.

Following the approach suggested by [1], [6] and [10], which is a natural adaptation of the original Bertsimas and Sim approach (see [3]), we define a parameter Γ, called *the budget of uncertainty* representing the range of uncertain demands that can deviate from their nominal values. We have Γ a real number belongs to $[0, n]$. For $\Gamma = 0$, every right hand side is equal to its nominal value, while $\Gamma = n$ leads to consider the problem with the worst demands.

We are interested in solving a robust version of the problem (T^β) with a 2-stage formulation. Indeed, the problem is to determine the minimum cost of choosing the facility i, $i = 1, \dots, m$ to be opened (with the r_i variables), and the supply level y_i, such that the worst demand is satisfied with a minimum cost. In this case, r_i and y_i variables are decided before the realization of the uncertainty (first stage decisions), while the t_{ij} variables represent the recourse variables to decide after the demands are revealed (second stage decisions). The robust problem is the following

$$T_{Rob}(\Gamma)\begin{cases} \min\limits_{\substack{y_i \le C_i r_i \\ y_i \ge 0,\ r_i \in \{0,1\}}} \sum_{i=1}^m d_i y_i + \sum_{i=1}^m f_i r_i + \max\limits_{\beta \in U} \min\limits_{\substack{\sum_{j=1}^n t_{ij} \le y_i,\ i=1\dots m \\ \sum_{i=1}^m t_{ij} \ge \beta_j,\ j=1\dots n \\ t_{ij} \ge 0,\ i=1\dots m,\ j=1\dots n}} \sum_{i=1}^m \sum_{j=1}^n \mu_{ij} t_{ij} \end{cases}$$

where $U = \{\beta \in \mathbb{R}^n : \beta_j = \overline{\beta}_j + z_j\hat{\beta}_j,\ j = 1, \ldots, n,\ z \in Z\}$ and $Z = \{z \in \mathbb{R}^n : \sum_{j=1}^{n} |z_j| \leq \Gamma,\ -1 \leq z_j \leq 1,\ \ j = 1 \ldots n\}$.

The problem $T_{Rob}(\Gamma)$ is a convex optimization problem that can be solved using *Kelley's algorithm* (see [8], [10]) that optimizes iteratively the master problem and the recourse problem by generating cuts. In this work, we focus on the recourse problem, namely

$$Q(y,\Gamma) \begin{cases} \max\limits_{\substack{\sum_{j=1}^{n} |z_j| \leq \Gamma \\ -1 \leq z_j \leq 1,\ j=1,\ldots,n}} \quad \min\limits_{\substack{\sum_{j=1}^{n} t_{ij} \leq y_i,\ i=1\ldots m \\ \sum_{i=1}^{m} t_{ij} \geq \overline{\beta}_j + \hat{\beta}_j z_j,\ j=1\ldots n \\ t_{ij} \geq 0,\ i=1\ldots m,\ j=1\ldots n}} \quad \sum_{i=1}^{m}\sum_{j=1}^{n} \mu_{ij} t_{ij} \end{cases}$$

At optimality $Z(Q(y,\Gamma))$ represents the worst transportation cost value for a fixed capacity level y, and Γ deviations. Furthermore, we assume that $Q(y,\Gamma)$ has a nonempty feasible set.

Because of the sense of the constraints of $Q(y,\Gamma)$, the optimal values of the z_j variables will never be negative, and necessarily belong to $[0,1]$. Moreover, by strong duality theorem, one can replace the minimization problem by its dual (since the problem is always feasible), resulting

$$Q(y,\Gamma) \begin{cases} \max -\sum_{i=1}^{m} y_i u_i + \sum_{j=1}^{n} \overline{\beta}_j v_j + \sum_{j=1}^{n} \hat{\beta}_j v_j z_j & \\ \quad \text{s.t. } v_j - u_i \leq \mu_{ij} & i = 1 \ldots m,\quad j = 1 \ldots n \\ \qquad \sum_{j=1}^{n} z_j \leq \Gamma & \\ \qquad 0 \leq z_j \leq 1 & j = 1 \ldots n \\ \qquad u_i,\quad v_j \geq 0 & i = 1 \ldots m,\quad j = 1 \ldots n \end{cases}$$

where $u_i,\ v_j$ are the dual variables.

The obtained program has a quadratic shape with $(m + 2n)$ variables and $(nm + n + 1)$ constraints. More precisely, it is a bilinear program subject to linear constraints, which is a class of convex maximization problems proven NP-hard (see [4]). From a complexity viewpoint, the resulting problem is not solvable in polynomial time. Instead of solving it on a direct way, we will reformulate $Q(y,\Gamma)$ as a mixed integer program. We present this formulation in Section 3.

3 Mixed-integer program reformulation

In the current formulation of $Q(y, \Gamma)$, Γ is a real number varying between 0 and n. Nevertheless, one can assume Γ to be integer, representing the number of the constraints for which $\beta_j \neq \overline{\beta}_j$. In this case, proposition 3.1 is required to give a MIP formulation of the problem $Q(y, \Gamma)$.

Proposition 3.1 *If Γ is an integer number then there exists an optimal solution (u^*, v^*, z^*) of $Q(y, \Gamma)$ such that $z_j^* \in \{0, 1\}$, $j = 1, \ldots, n$.*

From Proposition 3.1 and assuming that $\Gamma \in \mathbb{N}$ ($\Gamma \leq n$), we deduce that, at optimality either β_j is equal to its nominal value $\overline{\beta}_j$, or its worst value $\overline{\beta}_j + \hat{\beta}_j$. Furthermore, because of binary variables z_j we are able to linearize the problem $Q(y, \Gamma)$ by replacing each product $v_j z_j$ in the objective function with a new variable ω_j and adding constraints that enforce ω_j to be equal to v_j if $z_j = 1$, and zero otherwise (see [7]). The problem becomes a mixed integer program

$$Q(y,\Gamma)\begin{cases} \max -\sum_{i=1}^{m} y_i u_i + \sum_{j=1}^{n} \overline{\beta}_j v_j + \sum_{j=1}^{n} \hat{\beta}_j \omega_j & \\ \text{s.t. } v_j - u_i \leq \mu_{ij} & i = 1 \ldots m, \quad j = 1 \ldots n \\ \sum_{j=1}^{n} z_j \leq \Gamma & \\ \omega_j \leq v_j & j = 1 \ldots n \\ \omega_j \leq M z_j & j = 1 \ldots n \\ u_i, \; v_j, \; \omega_j \geq 0, \quad z_j \in \{0, 1\} & j = 1 \ldots n, \quad i = 1 \ldots m \end{cases}$$

where M is a sufficiently large constant. For reducing the integrality gap, M needs to be as small as possible. We give the following tight bound for M :

$$M_j = v_j^*(n)$$

where $v_j^*(n)$, $j = 1, \ldots, n$ is the optimal solution value of v variables in $Q(y, n)$. For the proof, see [5]. In the next Section, we are interested in numerical experiments, performed on the transportation problem in order to compare the tight bound previously defined with an arbitrarily large M.

4 Numerical experiments

Several series of tests were performed for various values of the parameters of the transportation problem, namely the number of sources, the number of demands, the amounts available at each source, the nominal and the highest demands at each destination and the transportation costs. To be closer to the reality, we choose to set the number of demands greater than the number of sources. All other numbers are randomly generated as follows: for all $j = 1, \ldots, n$, the nominal demand $\overline{\beta}_j$ belongs to $[10, 50]$, and the deviation $\hat{\beta}_j = p_j \overline{\beta}_j$, such that p_j represents the percentage of maximum increase of each demand j. We take p_j in $[0.1, 0.5]$, which ensures $\hat{\beta}_j$ to be strictly positive. The amounts y_i at each source $i = 1, \ldots, m$ are obtained by an equal distribution of the sum of the maximum demands. Finally, the costs are in $[1, 50]$.

The problem $Q(y, \Gamma)$ was solved with CPLEX 11.2. For each (n, m), ten instances have been generated. Table 1 shows results of average running time and percentage of solved instances, for each one of the two bounds previously mentioned (see Section 3), such that the computation was stopped after one hour.

Table 1
Running time results

		Running time (s)		% solved instances	
$n \times m$	Γ	M	$v^*(n)$	M	$v^*(n)$
	25%	2.63	0.76	100	100
250×10	50%	1178.14	14.05	50	100
	75%	215.24	0.97	90	100

The results described in Table 1 show that the computing time obtained by setting M to the bound $v^*(n)$ is significantly lower than the arbitrarily bound. Moreover, we remark that the running time increases for the value of Γ between $n/2$ and n whatever the bound is (see figure 1.a). Figure 1.b illustrates the evolution of the objective value versus Γ for a sample $m = 100$ and $n = 250$. The curve obtained is an increasing concave function, where $Z(Q(y, \Gamma))$ increases quickly for small values of Γ and slowly for high values. This is due to the model itself, since whenever Γ increases, the most influent uncertain parameters will be chosen.

Additional experiments have been performed on the uncertain transportation problem using the $v^*(n)$ bound, in order to determine the limit size of the problem that can be solved within one hour of CPU time. In Figure 2.a when $n = 500$, we observe that the running time grows as the number of sources m increases. Indeed, for $m = 10$ the problem takes few seconds to be solved. An

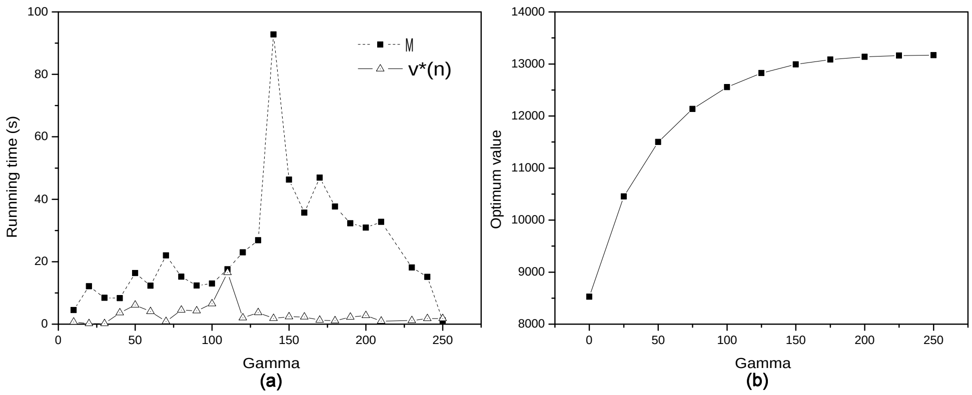

Fig. 1. A sample m=100, n=250: **a.** Running time vs Γ. **b.** Objective value vs Γ.

average of 20 minutes is needed for instances with $m = 80$ and $\Gamma = 60\%$ of total deviation, and one hour for those with $m \in [300, 500]$ and $\Gamma = 50\%$. When $n = 1000$ uncertain demands (Figure 2.b), all instances containing $m = 10$ sources are solved within one hour, whatever is Γ between 10% and 100%. For $100 \leq m \leq 500$ the solver is not able to reach the optimal within this time for $\Gamma = 50\%$, and for $m \geq 600$ there are memory issues with the solver.

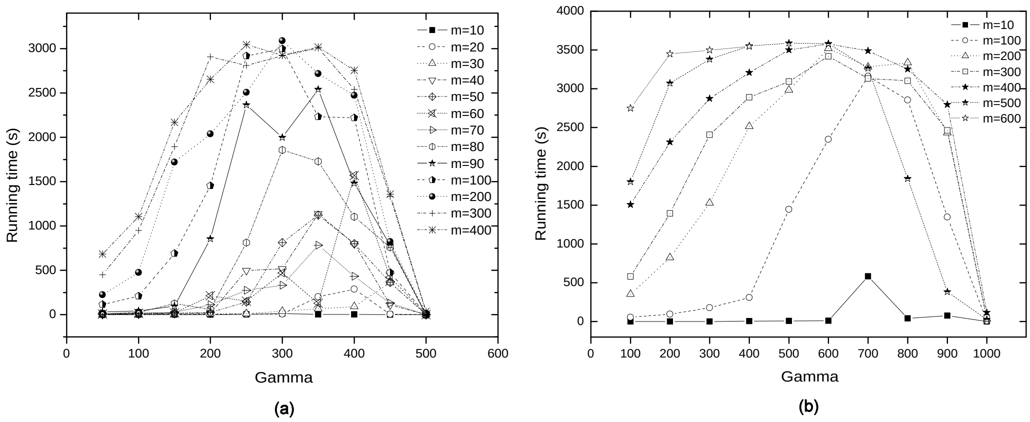

Fig. 2. Running time (s) vs Γ : **a.** Tests n=500. **b.** Tests n=1000.

5 Conclusion

The aim of this paper is to solve the recourse problem of the robust 2-stage location transportation problem. Previously, the 2-stage formulation has already been considered in [1] and [10]. Nevertheless, the limit size of solved instances with Kelley's algorithm, was performed for about 30 uncertain parameters. Here, we present the first (to our knowledge) extensive computation analysis on a particular recourse problem (namely, the location transportation problem), which is the most difficult part of the 2-stage robust optimization. Indeed, the tight bound we propose allows us to solve big size instances. Furthermore, this work seems to be promising to solve big size problems of the general 2-stage robust location transportation problem. This will be the aim of future research.

References

[1] Atamtürk, A. and M. Zhang, *Two-Stage Robust Network Flow and Design Under Demand Uncertainty*, Operations Research, **55**, 4 (2007), 662 – 673.

[2] Ben-Tal A., Goryashko, A., Guslitzer, E. and Nimerovski, A., *Adjustable robust solutions of uncertain linear programs*, Math. Program. Ser. **99** (2004), 351–376.

[3] Bertsimas, D. and Sim, M., *The Price of Robustness*, Oper. Res. **52** (2004), 35–53.

[4] Floudas, C. and Parlados, P.M., "State of the Art in Global Optimization, Global Optimization of separable concave functions under Linear Constraints with Totally Unimodular Matices", (1995).

[5] Gabrel, V. and Murat, C. and Remli, N. and Lacroix, M., *A new bound for solving the recourse problem of the 2-stage robust location transportation problem*,

[6] Gabrel, V. and Murat, C., *Duality and robustness in linear programming*, to appear in Journal of the Operational Research Society ,

[7] Glover, F. and Woolsey, E., *Converting the 0-1 polynomial programming problem to a 0-1 linear program*, Oper. Res., **22** (1974), 180–182.

[8] Kelley, J. E., *The Cutting-Plane Method for Solving Convex Programs*, Society for Industrial and Applied Mathematics, **8**,4 (1960), 703–712.

[9] Soyster, A. L., *Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming*, Operations research, **21**,5 (1973), 1154–1157.

[10] Thiele, A. and Terry, T. and Epelman, M., *Robust linear optimization with recourse*, Technical report (2009).