

# A note on robust 0 – 1 optimization with uncertain cost coefficients

Mustafa Ç. Pınar<sup>1</sup>

Department of Industrial Engineering, Bilkent University, 06533 Ankara, Turkey  
(e-mail mustafap@bilkent.edu.tr)

Received: September 2003 / Revised version: March 2004

**Abstract.** Based on the recent approach of Bertsimas and Sim (2004, 2003) to robust optimization in the presence of data uncertainty, we prove an easily computable and simple bound on the probability that the robust solution gives an objective function value worse than the robust objective function value, under the assumption that only cost coefficients are subject to uncertainty. We exploit the binary nature of the optimization problem in proving our results. A discussion on the cost of ignoring uncertainty is also included.

**Key words:** Robust optimization, 0-1 optimization, uncertain cost coefficients

**MSC classification:** 90C10, 90C15

---

## 1 Introduction

In two recent papers Bertsimas and Sim (2004, 2003), introduced a novel approach to robust optimization where they consider a robustness paradigm which provides a restricted protection against uncertainty in the coefficients while providing a probabilistic guarantee against perturbations outside the domain of protection of the robust problem. The distinguishing feature of their approach is the preservation of the computational complexity status of the original problem. In other words, unlike another recent approach due to Ben-Tal and Nemirovski (1998), the robust counterpart is no harder than the original problem of departure, be it a 0 – 1 linear program or a network-structured linear program.

The contribution of the present note is to give an alternative probabilistic bound in the special case where uncertainty is confined into the cost coefficients for 0 – 1 linear programming. Although our bound is potentially weaker than the Bertsimas-Sim bound, it is easier to compute, and conveys accurate information on the robust-

ness properties of the solution. The Bertsimas-Sim probabilistic bounds are shown in Bertsimas and Sim (2004) to be sharp, i.e., it is possible to display a simple bimodal distribution where the bounds are attained. However, they are computationally very demanding, and do not exploit the additional information provided by the 0 – 1 nature of the optimization problem nor the presence of uncertainty in the objective function only. Such problems where uncertainty is in the objective function constitute a sufficiently important subclass of uncertain optimization problems, such as the shortest path problem, the minimum spanning tree problem, the weighted non-bipartite matching and so on. In this perspective, our bounds complement the results of Bertsimas and Sim (2004) by offering probabilistic bounds that are very simple to compute, and that utilize the binary variable structure and uncertain cost nature of the optimization problems.

## 2 The Bertsimas-Sim approach

Consider the linear 0 – 1 program

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & x \in X \end{aligned}$$

where  $X = \{x \in \{0, 1\}^n : Ax \geq b\}$ . We assume that the problem is subject to uncertainty in the objective function where the coefficients  $c_j$  take values in the interval  $[\bar{c}_j, \bar{c}_j + d_j]$ , independently of one another. The biggest damage nature/adversary can inflict, assuming  $x$  has been fixed a priori is given by

$$\sum_{j=1}^n d_j x_j.$$

But, Bertsimas and Sim argue that most likely, nature/adversary will not push all  $c_j$ 's to their upper bounds simultaneously. Therefore, they consider hedging against a restricted adversary: assume only a subset of the  $c_j$ 's will be pushed to their upper bounds. Choose a  $\Gamma$  between 1 and  $n$ . And consider the maximum restricted damage:

$$\max_{\{S \subseteq \{1, \dots, n\}, |S| = \Gamma\}} \sum_{j \in S} d_j x_j$$

Therefore, one is led to solve the robust counterpart

$$\min_{x \in X} c^T x + \max_{\{S \subseteq \{1, \dots, n\}, |S| = \Gamma\}} \sum_{j \in S} d_j x_j$$

$\Gamma$  will control the conservatism in the solution, c.f. Theorem 1 of Bertsimas and Sim (2004) reproduced below. The inner problem is simplified as follows.

$$\begin{aligned}
\max \quad & \sum_{j=1}^n x_j d_j s_j \\
\text{s.t.} \quad & \sum_{j=1}^n s_j = \Gamma \\
& s_j \in \{0, 1\}, j = 1, \dots, n.
\end{aligned}$$

Relax the binary requirements to

$$0 \leq s_j \leq 1, j = 1, \dots, n$$

It is easy to see that the relaxed problem always has a binary optimal solution (see Theorem 1 of Bertsimas and Sim 2003). Now, one uses LP duality to obtain the following robust counterpart

$$\min_{x \in X, y} \bar{c}^T x + \Gamma y + \sum_{j=1}^n \max(0, d_j x_j - y). \quad (1)$$

One can linearize the max terms by introducing extra nonnegative variables. This is a problem of the same nature as the problem of departure. In other words, if it was an LP to begin with, it stays an LP. If it was a 0 – 1 LP it remains so. In the 0 – 1 case, further simplification results into the solution of at most  $n$  problems of identical structure to obtain the robust solution. More precisely, the robust problem can be recast as

$$\min_{x \in X, y} \bar{c}^T x + \Gamma y + \sum_{j=1}^n \max(0, d_j - y)x_j. \quad (2)$$

Furthermore, it can be shown that the optimal  $y$  is either zero or equal to one of the perturbation bounds  $d_j$  (Theorem 3 of Bertsimas and Sim 2003). Note that with this approach the objective function value  $\bar{c}^T x^*$ , where  $x^*$  is an optimal solution of the robust problem is smaller than or equal to  $t^*$ , the robust optimal value, for all possible values of  $c$  if at most  $\Gamma$  coefficients of  $c$  are perturbed. If more than  $\Gamma$  coefficients vary, the result will most likely be still not that dramatic as the following result of Bertsimas and Sim shows.

Assuming the  $c_j$ 's are iid random variables, symmetrically distributed, and rewriting the uncertain problem as

$$\begin{aligned}
\min \quad & t \\
\text{s.t.} \quad & c^T x \leq t \\
& x \in X
\end{aligned}$$

we can reiterate a theorem of Bertsimas and Sim (2004).

**Theorem 1.** *Let  $x^*$  denote a robust solution for some  $\Gamma$ . Then*

$$\text{Prob} \left( \sum_{j=1}^n c_j x_j^* > t^* \right) \leq \frac{1}{2^n} \left\{ (1 - \mu) \sum_{l=\lfloor v \rfloor}^n \binom{n}{l} + \mu \sum_{l=\lfloor v \rfloor + 1}^n \binom{n}{l} \right\}$$

where  $v = \frac{\Gamma + n}{2}$  and  $\mu = v - \lfloor v \rfloor$ .

They also show an approximate bound which is computationally simpler than the above; see Theorem 3(c) of Bertsimas and Sim (2004). Note that the above theorem does not exploit the fact that we are dealing with a 0 – 1 program. Furthermore, the analysis in Bertsimas and Sim (2004) concentrates on uncertainty in only the constraint matrix coefficients (observing that the objective function could be treated as a constraint). However, we believe that the class of 0 – 1 problems with uncertain cost coefficients deserve separate attention due to its importance.

### 3 The bound

Against this background, now, we can give the main results of this note. We assume for simplicity that all nominal values of cost coefficients  $\bar{c}_j$  are positive. If a nominal value is zero, our model of uncertainty (3) says that this coefficient is not uncertain. It is certainly a reasonable assumption that a coefficient which is known to be zero will remain so. On the other hand, we can modify the derivation of the robust problem accordingly, to include only those coefficients that are uncertain in (1). We refrain from this not to encumber notation any further.

**Theorem 2.** *Assume that the cost coefficients randomly take values in the following fashion: For positive  $\epsilon$  and  $\bar{c}_j > 0$ ,  $j = 1, \dots, n$ , let*

$$c_j = \bar{c}_j(1 + \epsilon \xi_j), \quad (3)$$

where  $\xi_j$  are iid, symmetrically distributed in  $[-1, 1]$ , and assume w.l.o.g.  $\sum_{j=1}^n \epsilon^2 \bar{c}_j^2 = 1$ . Let  $x^{rob}$ ,  $y^{rob}$  denote a robust solution for some  $\Gamma < n$ , and  $t^{rob}$  the corresponding optimal value. If  $y^{rob} > 0$  then, we have

$$\Pr\{c^T x^{rob} > t^{rob}\} \leq e^{-\frac{\Gamma^2 \epsilon^2 (\min_j \bar{c}_j)^2}{2}}. \quad (4)$$

*Proof.* The robust 0 – 1 problem in this case is

$$\begin{aligned} \min \quad & y\Gamma + \bar{c}^T x + \mathbf{1}^t z \\ \text{s.t.} \quad & \epsilon \bar{c}_i x_i - y \leq z_i, \quad i = 1, \dots, n \\ & x \in X \\ & z \geq 0 \end{aligned}$$

where  $\mathbf{1} = (1, \dots, 1)$  is the  $n$ -vector of ones. Denote an optimal vector  $(x^{rob}, y^{rob}, z^{rob})$ . Clearly, by nonnegativity we have

$$\mathbf{1}^T z^{rob} \geq 0. \quad (5)$$

Now, we can bound the probability of the event  $\{c^T x^{rob} > t^{rob}\}$ . We have

$$\begin{aligned} \Pr\{c^T x^{rob} > t^{rob}\} &= \Pr\left\{\sum_{i=1}^n \bar{c}_i(1 + \epsilon \xi_i)x_i^{rob} > \Gamma y^{rob} + \bar{c}^T x^{rob} + \mathbf{1}^T z^{rob}\right\} \\ &\leq \Pr\left\{\bar{c}^T x^{rob} + \epsilon \sum_{i=1}^n \bar{c}_i \xi_i x_i^{rob} > \Gamma y^{rob} + \bar{c}^T x^{rob}\right\} \\ &= \Pr\left\{\epsilon \sum_{i=1}^n \bar{c}_i \xi_i x_i^{rob} > \Gamma y^{rob}\right\} \\ &\leq \exp\left(-\Gamma^2 (y^{rob})^2\right) \prod_{i=1}^n \exp\left(\Gamma^2 (y^{rob})^2 \frac{\epsilon^2 \bar{c}_i^2 (x_i^{rob})^2}{2}\right) \end{aligned}$$

where the first inequality follows from (5), and the second inequality follows from a well-known result from probability theory as  $\Gamma y^{rob}$  is positive by hypothesis; see the proof of Proposition 1 of Ben-Tal and Nemirovski (2000). Now, the result follows since  $\sum_{i=1}^n \epsilon^2 \bar{c}_i^2 = 1$ ,  $x_i$ 's are binary valued, and  $y^{rob} \geq \epsilon \min_i \bar{c}_i$  by Theorem 3 of Bertsimas and Sim (2003) and by hypothesis of strict positivity.  $\square$

*Remark 1.* Notice that we can always rescale the coefficients  $\bar{c}_i$  so as to satisfy  $\sum_{i=1}^n \epsilon^2 \bar{c}_i^2 = 1$ .

*Remark 2.* The assumption of strict positivity of  $y^{rob}$  is a reasonable assumption to make for the following simple reason. If  $y^{rob}$  is zero, then the parameter  $\Gamma$  plays no role in the robust counterpart and we are dealing with the following robust counterpart

$$\min_{x \in X} \bar{c}^T x + \sum_{j=1}^n \max(0, d_j) x_j.$$

which is obtained from (2) by setting  $y$  to zero. Since in our model of uncertainty of Theorem 2 all perturbations  $d_j$  are positive then the robust problem is simply

$$\min_{x \in X} \bar{c}^T x + \sum_{j=1}^n d_j x_j,$$

which consists of hedging against the worst possible contingency. Obviously, this situation is uninteresting probabilistically since in this case the random objective function value will always be less than or equal to the robust optimal value.

If we allow equality in the event  $\{c^T x^{rob} > t^{rob}\}$ , i.e., if we want to bound the probability of the event  $\{c^T x^{rob} \geq t^{rob}\}$  we can derive a result similar to Theorem 2 under the condition that the random variables  $\xi_i$  have the following distribution:  $\Pr\{\xi_i = 1\} = \Pr\{\xi_i = -1\} = \frac{1}{2}$ .

**Theorem 3.** *Assume that the cost coefficients randomly take values under the uncertainty model (3) where  $\xi_j$  are iid, symmetrically distributed in  $[-1, 1]$  as  $\Pr\{\xi_i = 1\} = \Pr\{\xi_i = -1\} = \frac{1}{2}$ . Let  $x^{rob}, y^{rob}$  denote a robust solution for some  $\Gamma < n$ , and  $t^{rob}$  the corresponding optimal value. Furthermore, let  $\mathcal{A} = \{i : x_i^{rob} = 1\}$ . If  $y^{rob} > 0$  and  $|\mathcal{A}| \geq 1$  then, we have*

$$\Pr\{c^T x^{rob} \geq t^{rob}\} \leq |\mathcal{A}| e^{-\Gamma^2 \frac{(\min_{i \in \mathcal{A}} \bar{c}_i)^2}{2(\min_{i \in \mathcal{A}} \bar{c}_i)^2}}. \quad (6)$$

*Proof.* Using arguments similar to those in the proof of Theorem 2 we have

$$\begin{aligned} \Pr\{c^T x^{rob} \geq t^{rob}\} &\leq \Pr\left\{\sum_{i=1}^n \bar{c}_i \in \xi_i x_i^{rob} \geq \Gamma y^{rob}\right\} \\ &= \Pr\left\{\sum_{i \in \mathcal{A}} \bar{c}_i \in \xi_i \geq \Gamma y^{rob}\right\} \\ &= \Pr\left\{\sum_{i \in \mathcal{A}} \left(\bar{c}_i \in \xi_i - \frac{\Gamma y^{rob}}{|\mathcal{A}|}\right) \geq 0\right\} \\ &\leq \Pr\left\{\max_{i \in \mathcal{A}} \left(\bar{c}_i \in \xi_i - \frac{\Gamma y^{rob}}{|\mathcal{A}|}\right) \geq 0\right\} \\ &\leq \sum_{i \in \mathcal{A}} \Pr\left\{\left(\bar{c}_i \in \xi_i - \frac{\Gamma y^{rob}}{|\mathcal{A}|}\right) \geq 0\right\}. \end{aligned}$$

Now, from the proof of Lemma A.3 in Ben-Tal et al. (2002), for any continuous random variable  $v$  one has for any  $\rho \geq 0$

$$E(e^{\rho v}) \geq \Pr\{v \geq 0\}.$$

Therefore, we have

$$\Pr\left\{\bar{c}_i \in \xi_i - \frac{\Gamma y^{rob}}{|\mathcal{A}|} \geq 0\right\} \leq E\left(e^{\rho\left(\bar{c}_i \in \xi_i - \frac{\Gamma y^{rob}}{|\mathcal{A}|}\right)}\right).$$

Now, since

$$E(e^{\rho \bar{c}_i \in \xi_i}) = \cosh(\rho \bar{c}_i \epsilon) \leq e^{\frac{1}{2} \rho^2 \epsilon^2 \bar{c}_i^2}$$

we obtain

$$\Pr\left\{\bar{c}_i \in \xi_i - \frac{\Gamma y^{rob}}{|\mathcal{A}|} \geq 0\right\} \leq e^{\frac{1}{2} \rho^2 \bar{c}_i^2 \epsilon^2 - \frac{\Gamma y^{rob}}{|\mathcal{A}|}}.$$

The right side is minimized by choosing  $\rho = \frac{\Gamma y^{rob}}{|\mathcal{A}| \epsilon^2 \bar{c}_i^2}$ . Thus we get

$$\Pr \left\{ \bar{c}_i \in \xi_i - \frac{\Gamma y^{rob}}{|\mathcal{A}|} \geq 0 \right\} \leq e^{-\frac{1}{2} \frac{\Gamma^2 (y^{rob})^2}{|\mathcal{A}|^2 \epsilon^2 \bar{c}_i^2}}.$$

Summing up over all indices in  $\mathcal{A}$  we arrive at

$$\sum_{i \in \mathcal{A}} \Pr \left\{ \bar{c}_i \in \xi_i - \frac{\Gamma y^{rob}}{|\mathcal{A}|} \geq 0 \right\} \leq \sum_{i \in \mathcal{A}} e^{-\frac{1}{2} \frac{\Gamma^2 (y^{rob})^2}{|\mathcal{A}|^2 \epsilon^2 \bar{c}_i^2}}.$$

Now the proof is completed using the hypotheses that  $|\mathcal{A}| \geq 1$  and that  $y^{rob} > 0$  and arguments similar to the last paragraph of the proof of Theorem 2.  $\square$

*Remark 3.* Notice that we dispensed with the rescaling of coefficients  $\bar{c}_i$  (so as to satisfy  $\sum_{i=1}^n \epsilon^2 \bar{c}_i^2 = 1$ ).

#### 4 Cost of ignoring uncertainty

Coming from another angle, one can question the cost of neglecting the uncertainty in cost coefficients altogether in the above models of uncertainty. An interesting result in this direction is given in Oğuz (2000) with an emphasis on post-optimality or sensitivity analysis. Oğuz considers the following problem

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & x \in X \end{aligned}$$

where  $X$  is an arbitrary closed and bounded, non-empty set in  $\mathbb{R}_+^n$ . It is further assumed that the components of the vector  $c$  which may be equal to zero remain fixed at zero. In other words, this assumption is in agreement with our model of uncertainty (3). We now interpret Oğuz' result in our context. Define the nominal problem to be

$$\begin{aligned} \min \quad & \bar{c}^T x \\ \text{s.t.} \quad & x \in X \end{aligned}$$

Denote the optimal value for this problem  $z_1$  attained at some  $x^{nom}$  (not necessarily unique). Also define the random problem

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & x \in X \end{aligned}$$

where  $c$  is a random vector according to (3). Fix some realization of  $c$ , say  $\tilde{c}$ , and denote the optimal value by  $z_2$ . Furthermore let  $z_3 = \tilde{c}^T x^{nom}$ . Then Oğuz proves the following bound:

$$\frac{z_3 - z_2}{z_3} \leq \frac{2\epsilon}{1 + \epsilon}. \quad (7)$$

For small perturbations, e.g., for  $\epsilon = 0.05$ , the bound (7) says that the suboptimality of  $z_3$  is at most about 10% with respect to  $z_2$ ; i.e., we have  $z_3 \leq 1.10z_2$ . However, for  $\epsilon = 0.9$ , the bound (7) says that the suboptimality of  $z_3$  is given by the bound  $z_3 \leq 19z_2$ ! This suggests that the nominal solution may not yield an acceptable performance with relatively large disturbances.

By way of illustration of the bound in Theorem 2 in a similar situation, consider a problem with  $n = 100$  and where the  $\bar{c}$  varies from 10 to 15 in steps of 0.05. For  $\epsilon = 0.9$  we have  $(\epsilon \min_i \bar{c}_i)^2/2 = 0.003177$ . When  $\Gamma = 50$  the right hand of bound (4) gives a value of  $3.553 \times 10^{-4}$ . For  $\Gamma = 60$  we obtain  $1.078 \times 10^{-5}$ , whereas for  $\Gamma = 70$  we have  $1.734 \times 10^{-7}$ , and so on. Therefore, we can say that with high probability the value of a robust optimal solution in a random instance will be less than or equal to  $t^{rob}$ , the robust optimal value. This statement can also be interpreted as saying that the robust optimal value puts a probabilistic upper bound on the performance of a robust optimal solution in random instances.

## References

- Ben-Tal A, Nemirovski A (2000) Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming* 88: 411–424
- Ben-Tal A, Nemirovski A (1998) Robust convex optimization. *Mathematics of Operations Research* 23: 769–805
- Ben-Tal A, Nemirovski A, Roos C (2002) Robust solutions of uncertain quadratic and conic quadratic problems. *SIAM Journal on Optimization* 13: 535–560
- Bertsimas D, Sim M (2004) The price of robustness. *Operations Research* 52: 35–53
- Bertsimas D, Sim M (2003) Robust discrete optimization and network flows. *Mathematical Programming Series B* 98: 49–72
- Oğuz O (2000) Bounds on the opportunity cost of neglecting reoptimization in mathematical programming. *Management Science* 46: 1009–1012