Solving Discrete Optimization Problems when Element Costs are Random

By

Diptesh Ghosh*
Shubhabrata Das

July 2000

Please address all correspondence to:

Prof. Shubhabrata Das
Assistant Professor
Indian Institute of Management
Bannerghatta Road
Bangalore – 560 076
India
Fax: (080) 6584050
Email: shubho@iimb.ernet.in

*Faculty of Economic Sciences, University of Groningen, The Netherlands. On leave from Indian Institute of Management, Lucknow, India.

Copies of the Working Papers may be obtained from the FPM & Research Office

Solving Discrete Optimization Problems when Element Costs are Random

Diptesh Ghosh*

Faculty of Economic Sciences, University of Groningen, The Netherlands.

Shubhabrata Das

QM&IS Group, Indian Institute of Management, Bangalore, India.

Abstract

In a general class of discrete optimization problems with min-sum objective function, some of the elements may have random costs associated with them. In such a situation, the notion of optimality needs to be suitably modified. We define an optimal solution to be a feasible solution with the minimum risk. It is shown that the knowledge of the means of these random costs is enough to reduce such a problem into one with no random costs.

1 Introduction

In discrete optimization problems (DOPs), it is more of a norm than exception that the costs of some of the problem elements are not fixed. The practical solution in most of such cases is to assume some "good" approximation of the data and solve the problem. Once an optimal solution is obtained, post-optimality analysis techniques like sensitivity analysis is typically used to gain insight into the robustness of the solution obtained. In many situations, however, the decision maker has a fairly good idea about the distribution of these random elements. In this work, we try to find out how information about the distribution of the random valued data can be used to aid decision making. In general,

*On leave from Indian Institute of Management, Lucknow, India.

such a study can be considered as a part of stochastic integer programming, but we consider discrete optimization problems as a special case of general integer programming and try to obtain results that are elegant.

We consider a discrete optimization problem Π as a collection of problem instances $\pi = (G, \mathbb{S}, z)$, where G is the ground set consisting of n elements, with each element $e \in G$ having an associated cost c_e. The set $\mathbb{S}$, ($\subseteq 2^{|G|}$), is usually not described explicitly, but rather by a set of rules that each $S \in \mathbb{S}$ must satisfy; thus, $\mathbb{S}$ is collection of all feasible solutions. The function $z : \mathbb{S} \rightarrow \Re$ is referred to as the objective function (or the cost function). In this paper, we limit ourselves with *min-sum* objective functions, i.e. cases where $z(S) = \sum_{e \in S} c_e$. Such a generic framework covers a wide variety of discrete optimization problems as shown in the following examples.

Example 1 (Minimum Spanning Tree Problem) *In this problem we are given an undirected graph* $\mathbb{G} = (\mathbb{V}, \mathbb{E})$, *where each edge* $e \in \mathbb{E}$ *has a cost associated with it, and we have to find a minimum cost spanning tree in the graph. In our notation, an element refers to an edge in the graph,* G *is the set* $\mathbb{E}$, c_e *is the edge-length of edge* e *for all* $e \in \mathbb{E}$, $\mathbb{S}$ *is the set of all spanning trees in* $\mathbb{G}$, *and* $z(S) = \sum_{e \in S} c_e$ *for all* $S \in \mathbb{S}$.

Example 2 (0/1 Knapsack Problem) *In this problem we are given a set of* n *elements* $\mathbb{E} = \{e_1, \ldots, e_n\}$, *each element* e_j *having an associated profit* p_j *and an associated weight* w_j, *and a capacity* B, *and we are to determine the combination of elements that would maximize the profit and not exceed the capacity. In our notation, an element refers to an element in* $\mathbb{E}$ *(*G *is the set* $\mathbb{E}$*),* $c_{e_j} = p_j$ *for all* $e \in \mathbb{E}$, $\mathbb{S}$ *is the set of all* $S \subseteq \mathbb{E}$ *such that* $\sum_{e_j \in S} k_j \leq B$, *and* $z(S) = \sum_{e_j \in S} p_j$ *for all* $S \in \mathbb{S}$.

Example 3 (Symmetric Traveling Salesperson Problem) *In this problem we are given an undirected graph* $\mathbb{G} = (\mathbb{V}, \mathbb{E})$, *where each edge* $e \in \mathbb{E}$ *has a cost associated with it, and we have to find a minimum cost Hamiltonian cycle in the graph. In our notation, an element refers to an edge in the graph,* G *is the set* $\mathbb{E}$, c_e *is the edge-length of edge* e *for all* $e \in \mathbb{E}$, $\mathbb{S}$ *is the set of all Hamiltonian cycles in* $\mathbb{G}$, *and* $z(S) = \sum_{e \in S} c_e$ *for all* $S \in \mathbb{S}$.

We are concerned with the situation where the costs associated with certain elements are random variables.

We first formalize our problem through the following definitions and set up. Since several of the elements in the problems we consider do not have constant values, we will classify the problem elements using the following notation:

Definition 4 *An element* $e \in G$ *in* $\pi = (G, \mathbb{S}, z)$ *is called* fixed*(*random*) if* c_e *is constant (random valued).*

Definition 5 *Given any fixed set of values for* c_e*'s, the* loss *associated with a solution* $S \in \mathbb{S}$ *is defined by*

$$L(S) = [z(S) - Z^*],$$

where Z^* *is the minimum possible value of the objective function for given values of* c_e*'s (and hence is a function of these* c_e*'s).*

Obviously, with some of the c_e's being random, the loss of any feasible solution S is also a random variable. In practice it would not be desirable to compute a new course of action with every alteration of the c_e's, especially if we deal with $\mathcal{NP}$-hard problems. So we need to find a solution which would be good overall. With this in mind, we define the risk associated with a solution in the following manner:

Definition 6 *The* risk *associated with a solution* $S \in \mathbb{S}$ *is given by*

$$R(S) = \mathbb{E}[z(S) - Z^*]$$

where the expectation is taken w.r.t. the z-values of the random edges.

We define the optimization problem for DOPs with random elements as the problem of finding a feasible solution with minimum risk. Notice that if all the elements of the instance are fixed, the minimum risk solution corresponds to the traditional concept of an optimal solution, i.e., one that minimizes the objective function value.

In the next section we analyse the simplest case of DOPs with one random element. We show that knowledge of the mean of the distribution function for this element is sufficient to obtain an optimal solution in this situation. We generalize the result for DOPs with an arbitrary number of random elements in Section 3. We conclude the paper with Section 4, where we summarize our results and propose directions for further research.

2 DOPs with Single Random Element

Let us assume that we have a DOP with a single random element $e \in G$ First we study the objective value (Z^*) of an optimal solution (in the least cost sense) as a function of c_e. We then use this function to obtain decision rules to calculate optimal tours. Let X represent a random variable denoting the cost of the element e, $H(x)$ denote $P(X \leq x)$ — the distribution function of c_e, and μ denote its mean value.

We can partition $\mathbb{S}$ into $\mathbb{S}_e$ consisting of all solution containing the element e and $\mathbb{S}^e$ of all solution *not* containing e. Let S_e be a least cost solution in $\mathbb{S}_e$ and S^e be a least cost solution in $\mathbb{S}^e$. Notice that S_e and S^e need not be unique, and that they remain least cost solutions in their respective partitions regardless of the value of c_e. Also notice that for low values of c_e, $z(S_e) < z(S^e)$. Let θ denote the value of c_e for which $z(S_e) = z(S^e)$. Using θ we can describe $Z^*(c_e)$ as follows.

Lemma 7 *$Z^*(c_e)$ is a continuous function with a slope of 1 when $c_e < \theta$ and a slope of 0 when $c_e > \theta$.*

Proof: For low values of c_e, $z(S_e) < z(S^e)$. When c_e increases, the cost of all solutions in $\mathbb{S}_e$ increase while the cost of all solution in $\mathbb{S}^e$ remain the same. So S_e remains optimal until c_e increases to become larger than θ. If c_e increases further, $z(S_e) > z(S^e)$, and S^e becomes a new optimal solution. Clearly, no further increase in c_e will make S^e suboptimal.

The next lemma shows that int($\mathbb{S}_e$) and int($\mathbb{S}^e$) are disjoint although $\mathbb{S}_e$ and $\mathbb{S}^e$ are not.

Lemma 8 *The following statements are true*
(a) When $c_e < \theta$ every optimal solution contains e.
(b) When $\theta < c_e$ no optimal solution contains e.
(c) At $c_e = \theta$, both S_e and S^e are optimal.

Proof: (a) Assume to the contrary that there exists an optimal solution S_o not containing e. Let $T : c_e \to c_e + \delta$, $0 < \delta < \theta - c_e$. Since $e \notin S_o$, $z(S_o)$ remains unchanged after the transformation, i.e. the optimal objective value does not increase after T. But this contradicts Lemma 7.

(b) Can be proved using arguments similar to those in (a).

(c) Follows from the fact that Z^* is continuous (Lemma 7).

From Lemma 8 and the discussion preceeding Lemma 7, we get the following result.

Lemma 9 *For any value of c_e, at least one of S_e and S^e is an optimal optimal solution.*

We are now in a position to prove the key theorem for this section.

Theorem 10 *If c_e has a finite mean μ, then the optimal solution is the least cost solution when $c_e = \mu$.*

Proof: From Lemma 9, we know that either S_e or S^e is optimal. The expected loss for S_e, $\mathbb{E}L(S_e) = \int_\theta^\infty (x - \theta) dH(x)$ while that of S^e is $\mathbb{E}L(S^e) = \int_{-\infty}^\theta (\theta - x) dH(x)$.

Now S_e is optimal if $\mathbb{E}L(S^e) - \mathbb{E}L(S^e) \geq 0$, otherwise S^e is optimal. But

$$\begin{aligned} & \mathbb{E}L(S^e) - \mathbb{E}L(S^e) \\ = & \int_{-\infty}^{\theta} (\theta - x) dH(x) - \int_{\theta}^{\infty} (x - \theta) dH(x) \\ = & \theta \int_{-\infty}^{\infty} dH(x) - \int_{-\infty}^{\infty} x dH(x) \\ = & \theta - \mu \end{aligned} \quad (1)$$

which means that S_e will be an optimal solution if and only if $\mu \leq \theta$, the only interval where S_e is a least cost solution. The theorem follows.

Remark 11 *It is easy to see that, if* X *has a finite support (*$[a,b]$*), then*

$$R(S_e) = (b-\theta) - \int_{\theta}^{b} H(x)dx \tag{2}$$

$$R(S^e) = \int_{a}^{\theta} H(x)dx \tag{3}$$

Thus,

$$R(S_e) \leq R(S^e) \iff (b-\theta) \leq \int_{a}^{b} H(x)dx,$$

and the right hand side of the above equation reduces to $b-\mu$*, using integration by parts. This presents an alternative (equivalent) proof of the Theorem 10 for this special case.*

Theorem 10 implies that knowledge of the mean of the distribution function is enough to compute an optimal soution for a DOP with one random element. In the next section we generalize this result to DOPs with more than one random elements.

3 DOPs with general Number of Random Elements

In this section we consider the case where k of the elements are random. Accordingly, we partition G into $G_R = \{e_1, \ldots, e_k\}$ of random elements, and $G_F = \{e_{k+1}, \ldots, e_n\}$ of fixed elements. Let $X_1, \ldots, X_k$ be the random variables denoting the values of $c_{e_1} \ldots, c_{e_k}$ and $H(x_1, \ldots, x_n)$ denote $Pr(X_1 \leq x_1, \ldots, X_k \leq x_k)$. We represent the objective function value of any solution S as

$$z(S) = F(S) + \sum_{i: e_i \in S \cap G_R} X_i \tag{4}$$

where $F(S) = \sum_{e \in S \cap G_F} c_e$ is the fixed component of the cost $z(S)$.

Let $K_1 \ldots, K_{2^k}$ be the 2^k subsets of $K = \{1, \ldots, k\}$. For $i = 1, \ldots, 2^k$, let

$$\mathbb{S}_i = \{S : S \in \mathbb{S};\ e_j \in S \quad \forall j \in K_i;\ e_j \notin S \quad \forall j \in K \setminus K_i,\} \tag{5}$$

constitute a partition of $\mathbb{S}$.

Lemma 12 *If* $S^1, S^2 \in \mathbb{S}_i$*, for some* i*, then* $z(S^1) - z(S^2)$ *is non-random.*

Proof: By construction (5), S^1 and S^2 have the same set of random elements and hence by (4) $z(S^1) - z(S^2) = F(S^1) - F(S^2)$ which is non-random.

For any fixed set of costs $(x_1, \ldots, x_k)$, let S_i denote a least cost solution within $\mathbb{S}_i$. While S_i need not be unique, by Lemma 12, it remains a least cost solution among the ones in $\mathbb{S}_i$ regardless of values of the cost variables X_i's.

The following lemma is useful for restricting our search for optimal solutions.

Lemma 13 *For any solution* $S \in \mathbb{S}$, $R(S) \geq \min_j\{R(S_j)\}$.

Proof: Since $\mathbb{S} = \cup_{i=1}^{2^k} \mathbb{S}_i$, $\exists j \ni S \in \mathbb{S}_j$. Then

$$R(S) = \mathbb{E}[z(S) - Z^*] = \mathbb{E}[z(S) - z(S_j) + z(S_j) - Z^*] = z(S) - z(S_j) + R(S_j) \geq R(S_j),$$

following Lemma 12 and choice of S_j.

An immediate implication of Lemma 13 is the fact that at least one among S_1 through S_{2^k} is an optimal solution in the minimum risk sense.

Let us introduce the sets $\{\mathcal{R}_i; 1 \leq i \leq 2^k\}$ in the k-dimensional Euclidean space $(\Re^k)$ through

$$\mathcal{R}_i = \{(x_1, \ldots x_k) : S_i \text{ is a least cost solution at } (x_1, \ldots x_k)\}. \tag{6}$$

Let us also introduce a partition of the same space through $\{P_i; 1 \leq i \leq 2^k\}$ such that

$$\begin{aligned} P_1 &= \mathcal{R}_1, \\ P_i &= \mathcal{R}_i \setminus (\cup_{j<i} P_j) \quad i = 2, \ldots, 2^k. \end{aligned}$$

Notice that for all $i = 1, \ldots, 2^k$,

$$P_i \subseteq \mathcal{R}_i \tag{7}$$

If S_i is a least cost solution at $(x_1, \ldots x_k)$, then for this set of costs, $z(S_i) \leq z(S_j)$, $\forall j = 1, \ldots, 2^k$. Now

$$\begin{aligned} z(S_i) - z(S_j) &= F(S_i) + \sum_{m \in K_i} x_m - \Big[F(S_j) + \sum_{m \in K_j} x_m\Big] \\ &= \Big[\sum_{m \in K_i \setminus K_j} x_m - \sum_{m \in K_j \setminus K_i} x_m\Big] + F(S_i) - F(S_j). \end{aligned} \tag{8}$$

Therefore an alternative characterization of $\mathcal{R}_i$ is

$$\mathcal{R}_i = \Big\{(x_1, \ldots, x_k) : \sum_{m \in K_i \setminus K_j} x_m - \sum_{m \in K_j \setminus K_i} x_m \leq F(S_j) - F(S_i), j = 1, \ldots, 2^k\Big\} \tag{9}$$

and of P_i is

$$\begin{aligned} P_1 = \Big\{(x_1, \ldots, x_k) : \quad & \sum_{m \in K_1 \setminus K_j} x_m - \sum_{m \in K_j \setminus K_1} x_m \leq F(S_j) - F(S_1), \\ & j = 1, \ldots, 2^k\Big\} \\ P_i = \Big\{(x_1, \ldots, x_k) : \quad & \sum_{m \in K_i \setminus K_j} x_m - \sum_{m \in K_j \setminus K_i} x_m \leq F(S_j) - F(S_i), \\ & j = 1, \ldots, 2^k; (x_1, \ldots, x_k) \notin (\cup_{j<i} P_j)\Big\} \end{aligned} \tag{10}$$

We are now in a position to prove the main theorem in this section.

Theorem 14 *If $X_1, \ldots, X_k$ are random variables having finite means $\mu_1, \ldots \mu_k$ respectively, then the least cost tour, corresponding to the costs of $c_{e_1}, \ldots, c_{e_k}$ fixed at $\mu_1, \ldots, \mu_k$, will be optimal in the least risk sense.*

Proof: The risk associated with the solution S_i can be written as

$$R(S_i) = \sum_{j=1}^{2^k} \int_{P_j} \{z(S_i) - z(S_j)\} dH(.) \tag{11}$$

So,

$$\begin{aligned}
& R(S_i) - R(S_j) \\
&= \sum_{m \neq i,j} \int_{P_m} \{z(S_i) - z(S_j)\} dH(.) + \int_{P_j} \{z(S_i) - z(S_j)\} dH(.) \\
&\quad - \int_{P_i} \{z(S_j) - z(S_i)\} dH(.) \\
&= \int_{\mathfrak{R}^k} \{z(S_i) - z(S_j)\} dH(.) \\
&= \int_{\mathfrak{R}^k} \Big[\sum_{m \in K_i \setminus K_j} x_m - \sum_{m \in K_j \setminus K_i} x_m - (F(S_j) - F(S_i)) \Big] dH(.), \text{ by (8)} \\
&= \sum_{m \in K_i \setminus K_j} \mu_m - \sum_{m \in K_j \setminus K_i} \mu_m - (F(S_j) - F(S_i)). \qquad (12)
\end{aligned}$$

Hence for any i,

$$R(S_i) = \min_{1 \leq j \leq 2^k} R(S_j) \Leftrightarrow R(S_i) \leq R(S_j) \forall j \Leftrightarrow (\mu_1, \ldots, \mu_k) \in \mathcal{R}_i, \text{ by (9) and (12).}$$

Theorem 14 tells us that knowledge of the means of the random elements is adequate to obtain an optimal tour for a generic DOP with a min-sum objective function.

4 Conclusions

In this paper we considered the problem of solving a general class of discrete optimization problems with min-sum objective functions and having random cost elements where the distribution of the costs of the random elements are known. We defined the risk associated with feasible solutions as their expected suboptimality values. In Section 2 we showed that if there was only one random cost element in the problem, then the optimal solution in the least risk sense is a least cost solution when the cost of the random element is fixed at its mean value. In Section 3 we generalized this result to discrete optimization problems with an arbitrary number of random cost elements. In order to do so, we partitioned the set of all feasible solutions and created a corresponding partition of the Euclidean space of possible values of the random cost elements. Computing the risks of the least cost solution in each set of the partition of solutions, we showed that a minimum risk solution can be obtained by pegging the costs of each random cost element to the mean of the corresponding distribution and computing a least cost solution for this instance.

A direct extension of this work would be to consider general discrete optimization problems with min-max objectives. Analysis of such problems are more complicated due to the fact that the objective function values of such problems depend only on the cost of a single element in a solution. One can also consider the connections between our result and the sensitivity analysis and stability analysis results for general discrete optimization problems.

References

[1] S. van Hoesel and A.P.M. Wagelmans, "On the complexity of postoptimality analysis of 0/1 programs", Discrete Applied Mathematics, 91(1–3) 251–263, (1999).

[2] W.K. Klein Haneveld and M.H. van der Vlerk, "Stochastic Programming (Lecture Notes)", Available on request, email m.h.van.der.vlerk@eco.rug.nl, (2000).

[3] M. Libura, "Sensitivity analysis for minimum Hamiltonian path and traveling salesman problems", Discrete Applied Mathematics 30(2) 197–211, (1991).

[4] E.S. van der Poort. "Aspects of sensitivity analysis for the traveling salesman problem", Ph.D. thesis, Graduate school/Research Institute, Systems Organization and Management, University of Groningen, The Netherlands, (1997).

[5] R. Ramaswamy. Sensitivity analysis in combinatorial optimization, Fellowship dissertation, Indian Institute of Management, Calcutta, India (1994).

[6] A.P.M. Wagelmans. Sensitivity analysis in combinatorial optimization, Ph.D. Thesis, Econometric Institute, Erasmus University, Rotterdam, The Netherlands, (1990).