Control and Cybernetics

vol. **25** (1996) No. 6

Optimality conditions and sensitivity analysis for combinatorial optimization problems[1]

by

Marek Libura

Systems Research Institute, Polish Academy of Sciences,
01-447 Warszawa, Newelska 6, Poland

Abstract: Sensitivity analysis is an important phase of the solution procedure in which an optimal solution of the problem has been already found and additional calculations are performed in order to investigate how the optimal solution depends on changes in the problem data. The paper describes the main questions of sensitivity analysis which are specific for combinatorial optimization problems. Most of them are related to the problem of finding so called stability regions, defined as subsets of data for which some specified solution remains optimal. Methods, which can be used to compute the stability regions or its subsets, are presented in the context of optimality conditions on which they are based. As illustrations of various approaches, the sensitivity analysis results obtained for the problem of finding the minimum weight base of matroid, the binary knapsack problem and the symmetric traveling salesman problem are given.

1. Main problems in sensitivity analysis

Let $S = \{e_1, \ldots, e_n\}$ be a given set and $\mathcal{F} \subseteq 2^S$, where 2^S is the set of subsets of S. For any $e \in S$, $c(e) \in \mathbf{R}$, defines a *weight* of the element e. If $X \in \mathcal{F}$ then

$$C(X) = \sum_{e \in X} c(e) \tag{1}$$

is a *weight* of the set X.

The (linear) *combinatorial optimization problem* consists in finding a subset $X \in \mathcal{F}$ such that $C(X) \leq C(Y)$ for any $Y \in \mathcal{F}$. We will use the following standard notation

$$(P) \qquad C(X^o) = \min_{X \in \mathcal{F}} C(X).$$

[1]A preliminary version of this paper was presented on 9th Polish-Italian and 5th Polish-Finnish Symposium on System Analysis and Decision Support in Economics and Technology, Radziejowice, Poland, Oct.25-29, 1993

The set X^o is called an *optimal solution* of (P). Any element of $\mathcal{F}$ is called a *feasible solution* of the problem (P). A large variety of discrete optimization problems can be formulated in the above form.

In the *sensitivity analysis* for the problem (P) we assume that the vector of weights $c = (c(e_1), \ldots, c(e_n))^T$ as well as the family of feasible sets can vary. Assume that there is a specified subset $\mathcal{C} \subseteq \mathbf{R}^n$ of possible vectors c. Similarly, assume that the family of feasible sets depends on a parameter $u \in \mathcal{U}$, where $\mathcal{U}$ is a given set of possible parameters. We will write that $\mathcal{F} = \mathcal{F}(u)$. Moreover, there is a specified set

$$\mathcal{P} \subseteq \mathcal{C} \times \mathcal{U},$$

which will be called a *set of possible problem data.*

Any pair $p = (c, u) \in \mathcal{P}$ defines an instance of the problem (P). Denote by $\Omega(p)$ the set of optimal solutions of the problem (P) with data p.

Consider an optimal solution X^o of the problem (P) obtained for the data $p^o = (c^o, u^o)$.

The main object investigated in the sensitivity analysis is so called *stability region* of the solution X^o defined in the following way:

$$P(X^o) = \{p \in \mathcal{P} : X^o \in \Omega(p)\}.$$

The set $P(X^o)$ describes all allowed variations of problem data for which the solution X^o remains optimal.

Usually, it is very difficult to describe the stability region for a given solution of the problem (P). Sometimes it is difficult even to determine such properties of $P(X^o)$ as connectedness or convexity. Frequently we try to describe only subsets of the stability region imposing various requirements that some elements in problem data be fixed. This is equivalent to defining in an appropriate way the set $\mathcal{P}$ of possible problem data. Such a sensitivity analysis can be more 'tractable' from the computational point of view and still useful in practice. If we assume for example that $\mathcal{U} = \{u^o\}$ and $\mathcal{P} = \mathbf{R}^n \times \{u^o\}$, then $P(X^o)$ is a polyhedral convex cone (see e.g. Libura, 1977) which in some cases can be completely described. In the extreme case we can assume that only one coefficient in the problem data can vary, and then the sensitivity analysis consists in finding so called *tolerances* of this coefficient, i.e., maximum changes of its value which do not forfeit the optimality of X^o (see e.g. Libura, 1993).

From the practical point of view the possibility of finding any subset of the stability region is of interest, because such a subset describes the problem data for which the solution considered is still valid.

A very convenient type of such a subset is so called *stability ball* defined as a ball with a center in p^o such that its part belonging to $\mathcal{P}$ is fully contained in $P(X^o)$. Unfortunately, in many cases the maximum radius of this ball, called the *stability radius* of solution X^o, is equal to zero. It always happens if the optimal solution X^o is not unique. Various approaches have been considered in this situation. One of them (see Leontev, 1979) consists in defining a stability region with respect to the entire set of optimal solutions. Such a region, denoted by $P_\Omega(p^o)$, is the set of such data for which no additional solution enters the optimal set, i.e., $P_\Omega(p^o) = \{p \in \mathcal{P} : \Omega(p) \subseteq \Omega(p^o)\}$. The stability ball and the stability radius are now defined in a similar way as for the region $P(X^o)$. In this case the stability radius is always positive.

Another approach consists in introducing so called stability index (see Libura, van der Poort, Sierksma, van der Veen, 1995), which is a pair composed of the set of elements for which weights must not change to preserve the optimality of X^o, and the stability radius defined for the subset of remaining elements.

Similar definitions can be introduced when we consider approximate solutions instead of the optimal one. Then, for example, the *ϵ-stability region* of $X \in \mathcal{F}$ is defined for $\epsilon \geq 0$ as a set of problem data for which $C(X) \leq (1+\epsilon)C(Y)$ for any $Y \in \mathcal{F}$. In a similar way the ϵ-stability ball and the ϵ-stability radius are introduced.

The mentioned problems of the sensitivity analysis are closely connected to so called parametric analysis. The parametric problem for (P) consists in covering the set $\mathcal{P}$ of possible problem data by stability regions of solutions of (P). Also this problem can be in practice solved only partially for particular subsets of $\mathcal{P}$.

2. Sensitivity analysis and optimality conditions

All methods used to find stability regions or their subsets can be related to optimality conditions on which they are based. This section presents main sources of optimality conditions in combinatorial optimization and presents the results obtained for some standard discrete optimization problems as illustrations for various approaches derived from different optimality conditions.

An ideal situation would be if for a given problem (P) we could find for some index set Q a family of functions $J_q : \mathcal{F} \times \mathcal{C} \times \mathcal{U} \to \mathbf{R}$, $q \in Q$, such that the following conditions held:

X^o is an optimal solution of the problem (P) *with data* $p = (c, u)$
if and only if $J_q(X^o, c, u) \leq 0$ *for* $q \in Q$.

If the above necessary and sufficient conditions hold, then the stability region is obviously defined in the following way:

$$P(X^o) = \{(c,u) \in \mathcal{P} : J_q(X^o, c, u) \leq 0,\ q \in Q\}. \tag{2}$$

Unfortunately, the situations in which we are able to formulate such a necessary and sufficient optimality conditions are very seldom in combinatorial optimization. Nevertheless, we face such a situation for some important problems, and then a complete sensitivity analysis is possible. The following example describes such a case. (Descriptions of examples end with □.)

EXAMPLE 2.1

Let $M = (S, \mathcal{I})$ be a matroid on S (see e.g. Welsh, 1976). Assume that $\mathcal{F} \subseteq \mathcal{I}$ is a family of bases of M, and $c(e)$, $e \in S$, denotes the weight of matroid element e. Then (P) becomes the well known and important problem of finding a minimum weight base of the matroid M. Specifying in an appropriate way the matroid and the weights of its elements we can obtain various combinatorial optimization problems (see Lawler, 1976). If for example M is so called *cycle matroid* of a given undirected graph G with a set of edges S, and $c(e)$ denotes the length of edge e, then (P) becomes a problem of finding the minimum weight forest in G.

For the problem of finding the minimum weight base of matroid we are able to formulate necessary and sufficient optimality conditions. These conditions are expressed through the family of so called *fundamental circuits* or the family of *fundamental cutsets* (see Libura, 1991).

Let $C(e_j, X^o)$ denote the fundamental circuit in the base X^o of the matroid M defined by the element $e_j \in S \setminus X^o$ (see e.g. Welsh, 1976), and assume that $\mathcal{C} = \mathbf{R}^n$. Then the stability region of the minimum weight base X^o can be completely described in the following way:

$$P(X^o) = \{c \in \mathbf{R}^n : c(e_i) \leq c(e_j),\ e_j \in S \setminus X^o,\ e_i \in C(e_j, X^o)\}.$$

□

2.1. 'Trivial' optimality conditions

Necessary and sufficient optimality conditions are available rather seldom in discrete optimization. Nevertheless, for any optimization problem we can state 'trivial' optimality conditions which follow immediately from the definition of optimal solution. Using the notation introduced in Section 1, they can be formulated in the following way :

$X^o \in \Omega(p)$ *for* $p = (c, u)$, *if and only if*

$$X^o \in \mathcal{F}(u) \tag{3}$$

$$C(X^o) \leq C(X) \ \textit{for any}\ X \in \mathcal{F}(u). \tag{4}$$

One should not expect that the above optimality conditions can be directly applied to describe the stability region of X^o, unless the problem (P) is very simple and 'highly structured'. Nevertheless, such conditions can be exploited in two manners :

- they can be simplified in particular situations;
- they can be used in some qualitative analyses.

The following examples illustrate both possibilities.

EXAMPLE 2.2

This example shows how a subset of the stability region can be described by reformulating the 'trivial' optimality conditions.

Consider the binary knapsack problem (see e.g. Garfinkel, Nemhauser, 1972)

$$\begin{aligned} \min \sum_{i=1}^{n} c_i x_i & \\ \sum_{i=1}^{n} a_i x_i &\geq a_0 \\ x_i &= 0 \text{ or } 1, \quad i = 1, \ldots, n. \end{aligned} \tag{5}$$

To state (5) as the problem (P) assume that for a given $S = \{e_1, \ldots, e_n\}$ and for $X \subseteq S$ the vector $x = (x_1, \ldots, x_n)^T \in \{0,1\}^n$ is a characteristic vector of X, i.e., $x_i = 1$ if and only if $e_i \in X$. Moreover, $X \in \mathcal{F}$ if and only if $a^T x \geq a_0$, where $a = (a_1, \ldots, a_n)^T \in \mathbf{R}^n_+$, $a_0 \in \mathbf{R}_+$. Assume that $\mathcal{C} = \{c^o\}$ and let x^o be a characteristic vector of the optimal solution X^o of the problem (5) with $c = c^o$. Then, as it was shown in Libura (1977),

$$P(X^o) = \{c^o\} \times \{(a_0, a_1, \ldots, a_n)^T \in \mathbf{R}^{n+1}_+ :$$

$$\sum_{i=1}^{n} a_i x_i^o \geq a_0 \tag{6}$$

$$\sum_{i \in Q} a_i < a_0 \quad \text{for } Q \in \mathcal{Q}\}, \tag{7}$$

where $\mathcal{Q}$ is a family of maximal subsets $Q \subseteq \{1, \ldots, n\}$ such that the inequality $\sum_{i \in Q} c_i < \sum_{i=1}^{n} c_i x_i^o$ holds.

In the description of $P(X^o)$ the inequality (6) corresponds to the inequality (3), which implies the feasibility of the solution X^o. The inequalities (7) are derived from the conditions (4). □

EXAMPLE 2.3

This example illustrates the possibility of simple sensitivity analysis based on a formula which is derived directly from the 'trivial' optimality conditions.

Assume that $\mathcal{P} = \mathbf{R}^n \times \{u^o\}$ and consider a Chebyshev metric in $\mathbf{R}^n$. Let $\rho(X^o, p^o)$ denote the so called *stability radius* of the optimal solution X^o, defined as the maximum radius of the ball with center in p^o and entirely contained in the stability region.

The 'trivial' optimality conditions lead directly to the following formula (see Libura, 1993):

$$\rho(X^o, p^o) = \min_{X \in \mathcal{F} \setminus \{X^o\}} \frac{C^o(X) - C^o(X^o)}{|X| + |X^o| - 2|X^o \cap X|}, \tag{8}$$

where $|X|$ denotes cardinality of the set X and $C^o(X)$ is the weight of X for the problem data p^o.

The formula (8) can hardly be regarded as an efficient way of calculating the stability radius, but it can provide a simple lower bound for this radius.

Assume for example that we know that the problem (P) has a single optimal solution X^o for the data p^o. (Observe that from (8) it follows immediately that if $|\Omega(p)| > 1$ then $\rho(X^o, p^o) = 0$.) Moreover, assume that all coefficients of the vector c^o are rational numbers, i.e., $c_i^o = l_i^o / m_i^o$, $i = \{1, \ldots, n\}$. Then from (8) it is easy to see that $\rho(X^o, p^o) \geq 1/(2Mn)$, where M is the smallest common multiple of m_i^o, $i = 1, \ldots, n$. □

2.2. Sufficient optimality conditions based on simple relaxations and restrictions of (P)

It was mentioned above that the necessary and sufficient optimality conditions are very seldom available in combinatorial optimization. On the other hand, sometimes we are able to formulate various *sufficient* optimality conditions. These conditions play a very important role in the sensitivity analysis due to the fact that they allow to describe some proper subsets of the stability region. The most immediate sufficient optimality conditions follow directly from simple relaxations and restrictions of the problem (P).

Let (P) be an original discrete optimization problem

$$\text{(P)} \qquad \min_{X \in \mathcal{F}(u)} C(X).$$

The problem

$$\text{(R)} \qquad \min_{X \in \mathcal{F}'} C'(X)$$

is called a *relaxation* of (P) if the following conditions are satisfied: $\mathcal{F}(u) \subseteq \mathcal{F}'$ and $C'(X) \leq C(X)$ for all $X \in \mathcal{F}(u)$.

The problem

$$\text{(Q)} \qquad \min_{X \in \mathcal{F}''} C''(X)$$

is called a *restriction* of (P) if : $\mathcal{F}'' \subseteq \mathcal{F}(u)$ and $C''(X) \geq C(X)$ for all $X \in \mathcal{F}''$.

The following sufficient optimality conditions (see e.g. Geoffrion, Nauss, 1977) follow directly from the definitions of relaxations and restrictions.

Sufficient optimality condition based on relaxation of (P):

If X^* is an optimal solution for the problem (R) *and $X^* \in \mathcal{F}(u)$, $C(X^*) = C'(X^*)$, then X^* is also an optimal solution of the problem* (P).

Sufficient optimality condition based on restriction of (P):

If $X^o \in \Omega(p)$ is feasible for the problem (Q) *and $C(X^o) = C''(X^o)$, then X^o is an optimal solution of* (Q).

The crucial point in applying these optimality conditions is an appropriate choice of relaxations and restrictions. Namely, the data in modified problems (R) and (Q) must be closely connected to the data of the original problem (P). Usually this correspondence is quite natural, and one assumes some functional dependence between the data of related pairs of problems. If $\mathcal{P}_R$ is the set of possible problem data of the relaxation (R) then we must be able to define a mapping $r : \mathcal{P}_R \to \mathcal{P}$ such that the problem (R) with data q is a relaxation of the problem (P) with any data $p \in r(q)$.

A general scheme of exploiting the optimality conditions based on relaxations to construct a subset of the stability region $P(X^o)$ is the following:

- We try to find a subset $R(X^o) \subseteq \mathcal{P}_R$ such that: (i) X^o is optimal in (R) with the data belonging to $R(X^o)$, (ii) X^o is feasible in (P) for any data obtained as a mapping $r(R(X^o))$ and the optimal values of problems (P) and (R) are equal;
- Then, from the optimality conditions based on the relaxation (R), it follows that the subset $r(R(X^o)) \cap \mathcal{P}$ is contained in the stability region $P(X^o)$.

Observe that such a scheme requires that we be able to perform the sensitivity analysis for the relaxation (R). The main assumption which decides about the efficiency of such an approach is that the relaxation can be much simpler than the original problem, and that the sensitivity analysis for (R) will be much simpler as well.

The use of particular relaxations in sensitivity analysis will be discussed more deeply in Section 2.5. in the context of duality in combinatorial optimization.

The use of restrictions in the sensitivity analysis consists in direct application of the optimality conditions. Now we need an appropriate mapping $q : \mathcal{P} \to \mathcal{P}_Q$, where $\mathcal{P}_Q$ is the set of data of the problem (Q). The general scheme is as follows:

- We try to find a subset $Q(X^o) \subseteq \mathcal{P}$ such that for any $p \in Q(X^o)$ the following conditions hold: (i) the problem (Q) with the data $q(p)$ is a restriction of (P); (ii) X^o is a feasible solution of (Q) and the optimal value of (Q) is equal to the optimal value of the original problem.
- Then from the optimality conditions for restrictions it follows immediately that $P_Q \subseteq P(X^o)$.

2.3. Optimality conditions based on linear programming description of the problem (P)

This section concerns a relaxation of the problem (P) connected with the linear programming description of (P).

Assume that the family $\mathcal{F}$ of feasible sets is fixed and all data changes concern the vector c of weights, i.e., $\mathcal{P} = \mathcal{C} \times \{u^o\}$, $\mathcal{C} = \mathbf{R}^n$.

The problem (P) can be equivalently stated in the form

$$\min c^T \xi(X) \qquad (9)$$
$$\xi(X) \in \xi(\mathcal{F}).$$

where $\xi(X) \in \{0,1\}^n$ denotes a characteristic vector of X, and $\xi(\mathcal{F})$ is the set of characteristic vectors of elements of the family $\mathcal{F}$.
Consider the following relaxation of (9):

$$\min c^T x \qquad (10)$$
$$x \in F$$

where $F \subseteq \mathbf{R}^n$ is the convex hull of the set $\xi(\mathcal{F})$. The set F is a polyhedral convex set, which follows from the fact that the set S is finite. If we have the description of F in the form of the system of linear equations or inequalities, then the problem (P) can be solved as a linear programming problem. This fact is well known in the combinatorial optimization and is used in various algorithms (see e.g. Schrijver, 1986). Usually the problem of finding a linear programming description of F is a very difficult task. From the point of view of the sensitivity analysis it would be enough to have a description of F only in the neighborhood of the point $\xi(X^o)$. Namely, let $D(X^o)$ denote the smallest convex cone such that $F \subseteq D(X^o) + \xi(X^o)$. Then it can be shown (see e.g. Libura, 1977) that

$$P(X^o) = G(X^o) \times \{u^o\}, \qquad (11)$$

where

$$G(X^o) = \{c \in \mathbf{R}^n : c^T y \leq 0, \;\; y \in -D(X^o)\}. \qquad (12)$$

The formula (12) is a consequence of optimality conditions known from the theory of linear programming. Thus we have:

X^o is an optimal solution of (P) *for a given $u = u^o$ if and only if the vector c belongs to the polar cone of $-D(X^o)$* .

Also a complete description of the cone $D(X^o)$ is usually too difficult in practice. Nevertheless, it is enough to have some cone D with the property

$$D(X^o) \subseteq D$$

beeing an approximation of the cone $D(X^o)$. Observe that then, instead of the above *necessary and sufficient* condition, we have the following *sufficient* optimality condition:

If c belongs to the polar cone of $-D$ then X^o is optimal in the problem (P).

As the polar cone of D is contained in the polar cone of $D(X^o)$, from this condition we obtain the subset of the stability region $P(X^o)$.

Some methods of solving the problem (P), belonging to the class of cutting planes methods (see e.g. Garfinkel, Nemhauser, 1972) or branch-and-cut methods (see Crowder, Johnson, Padberg, 1983), can generate the cone D in a natural way as a 'by-product' of the solution procedure.

2.4. Optimality conditions based on k-th best solutions

Consider the set of all feasible solutions of the problem (P) numbered in such a way, that

$$C(X^o) \leq C(X^1) \leq C(X^2) \leq \ldots \leq C(X^s), \qquad (13)$$

where $s = |\mathcal{F}| - 1$.

Given such a sequence we will call k-th element of sequence (13) the *k-th best solution* of the problem (P). Usually we obtain only the single solution X^o. But frequently the algorithm solving the problem (P) can be modified in such a way that it generates several consecutive elements of (13). Such a subset of solutions can be used in the sensitivity analysis for the solution X^o. The approach is based on an idea presented in Piper, Zoltners (1976).

Assume that we know first k elements of the sequence (13). Let $\mathcal{K} = \{X^1, \ldots, X^{k-1}\}$, and $\delta = C(X^{k-1}) - C(X^o)$. Moreover, let for $c' \in \mathcal{C}$

$$z(c') = \min\{C'(Y) : Y \in \mathcal{F},\ C(Y) \geq C(X^o) + \delta\}, \qquad (14)$$

where $C'(Y)$ is the weight of subset Y for the vector of weights c'.

Assume that the family of feasible solutions is fixed, i.e., $\mathcal{U} = \{u^o\}$. Then we have the following sufficient optimality condition (see Libura, 1993; Piper, Zoltners, 1976):

X^o is an optimal solution of the problem (P) *with the vector of weights c', if $C'(X^o) \leq C'(Y)$ for any $Y \in \mathcal{K}$, and $C'(X^o) \leq z$ for some $z \leq z(c')$.*

The main problem, which must be solved if we want to use the above optimality conditions, consists in an appropriate choice of the threshold value z. The best choice, leading to the strongest optimality condition, would be obtained for $z = z(c')$, but this is impractical, because to calculate $z(c')$ is as difficult as to solve the original problem (P). On the other hand, the value z cannot be too small, because then the assumptions in the optimality conditions are not satisfied. Therefore, some tight but computationally inexpensive lower bound for the optimal value of the problem (14) is needed. Such bounds can be obtained by solving various relaxations of (14). One possible relaxation consists in replacing in (14) the condition $Y \in \mathcal{F}$ by a weaker condition $Y \in 2^S$, which leads to the following knapsack problem

$$\min\{C'(Y) : Y \in 2^S,\ C(Y) \geq C(X^o) + \delta\}.$$

Observe that the optimality conditions considered give a single point (c', u^o) in the set $\mathcal{P}$ of admissible data. But it is known (see e.g. Libura, 1977), that when $\mathcal{U} = \{u^o\}$, then the projection of the stability region $P(X^o)$ on $\mathcal{C} = \mathbf{R}^n$ is a polyhedral convex cone. Thus, given the set of vectors c^i, $i \in T$, for which the solution X^o remains optimal in (P), we obtain a subset $P_T \subseteq P(X^o)$, such that

$$P_T = \text{cone}(c^i,\ i \in T) \times \{u^o\},$$

where $\text{cone}(c^i,\ i \in T) = \{c \in \mathbf{R}^n,\ c = \sum_{j \in T} \lambda_j c^j,\ \lambda_j \geq 0 \quad \text{for } j \in T\}$.

Given the set $\mathcal{K}$ we can also easily calculate the tolerances of weights for some subset of elements of S and give lower bounds for the tolerances of remaining elements.

Let $t^+(e)$ and $t^-(e)$ denote respectively the upper and the lower tolerance of weight $c(e)$ for $e \in S$, i.e., $t^+(e)$ is the maximum individual increase and $t^-(e)$ is the maximum individual decrease of the weight $c(e)$ which do not forfeit the optimality of X^o.

Let $\mathcal{K}^+(e) = \{X \in \mathcal{K} : e \in X\}$ and $\mathcal{K}^-(e) = \{X \in \mathcal{K} : \{e\} \cap X = \emptyset\}$.
Then for $e \in X^o$, $\quad t^-(e) = \infty$
and

$$t^+(e) = \min\{C(X) : X \in \mathcal{K}^+(e)\} - C(X^o) \quad \text{if } \mathcal{K}^+(e) \neq \emptyset$$
$$t^+(e) \geq \delta \quad \text{otherwise.}$$

Similarly, for $e \in S \setminus X^o$, $\quad t^+(e) = \infty$
and

$$t^-(e) = \min\{C(X) : X \in \mathcal{K}^-(e)\} - C(X^o) \quad \text{if } \mathcal{K}^-(e) \neq \emptyset$$
$$t^-(e) \geq \delta \quad \text{otherwise.}$$

This approach was used in computational experiments for the traveling salesman problem in Libura, van der Poort, Sierksma, van der Veen (1995).

2.5. Optimality conditions based on duality

Duality plays an important role in the sensitivity analysis in such branches of optimization as linear, convex and geometric programming (see Fiacco, 1983; Gal, 1979) . Also in combinatorial optimization this role is significant, although the situation here is different. The main difference consists in the fact that for a given combinatorial optimization problem one can formulate various dual problems and most of them are not 'strong' duals. This means that the optimal value of such a dual problem is not equal to this of the primal one. A consequence of this fact is that usually duality leads only to sufficient optimality conditions.

Most of dual problems considered in the combinatorial optimization can be obtained using the following simple scheme described in Libura (1984):

Assume that for a problem (P) we can define a family $(\mathrm{R}_q,\ q \in Q)$ of relaxations, where

$$(\mathrm{R}_q) \qquad \min_{X \in \mathcal{F}_q} C_q(X).$$

The parameter q, determining a particular relaxation, plays a role of dual variable.

The dual problem based on the family $(\mathrm{R}_q,\ q \in Q)$ is defined in the following way:

$$(\mathrm{D}) \qquad \sup_{q \in Q} v(\mathrm{R}_q),$$

where $v(\mathrm{R}_q)$ denotes the optimal value of the problem (R_q).

The following examples illustrate various possibilities of choice of the families of relaxations and corresponding dual problems.

EXAMPLE 2.4

Assume that $S \subseteq V \times V$ is the set of edges of an undirected graph $G = (V, S)$, and that the family $\mathcal{F}$ in the definition of problem (P) is a family of such a subsets X of graph edges, that the subgraph $G_X = (V, X)$ is a Hamiltonian cycle in G (see e.g. Lawler, Lenstra, Rinnoy Kan, Shmoys, 1985). Taking the vector of weights in (P) equal to the vector of edge lengths in G we obtain the well known traveling salesman problem. One possible relaxation of (P) is then so called minimum 1-tree problem (see e.g. Lawler, Lenstra, Rinnoy Kan, Shmoys, 1985), which is obtained when the family of feasible sets in (P) is replaced by the family of 1-trees (i.e., a family of such subsets of S, which form exactly one cycle in the graph G). A natural parametrization of such a relaxation consists in replacing the original length $c(e)$ of the edge $e = (i, j)$, $i, j \in V$, by a new length

$$c_q(i, j) = c(i, j) + q(i) + q(j), \tag{15}$$

where $q(i)$, $q(j)$ are the elements of the so called penalty vector $q = (q(1), \ldots, q(|V|))^T$, $q \in \mathbf{R}^{|V|}$. Such a modification has been used in various algorithms for the traveling salesman problem as well as in the sensitivity analysis for this problem, Libura (1991). In the described case, the set Q of dual parameters is $\mathbf{R}^{|V|}$, and the dual problem consists in finding a penalty vector q, which maximizes the weight of the minimum spanning 1-tree in the graph G with modified weights of edges. □

EXAMPLE 2.5

Usually we are faced with a primal problem in which the family $\mathcal{F}$ has some 'structure'. An important case is when this family can be defined as an intersection of two other families of subsets of S, i.e.,

$$\mathcal{F} = \mathcal{G} \cap \mathcal{X},$$

where $\mathcal{G}, \mathcal{X} \subseteq 2^S$, and $\mathcal{G}$ is defined through a system of inequalities:

$$\mathcal{G} = \{X \subseteq S : g(x) \geq b\}$$

for some $g : \mathbf{R}^n \to \mathbf{R}^s$, $b \in \mathbf{R}^s$. As before, x denotes the characteristic vector of the set X. In practice the family $\mathcal{X}$ describes some 'structural' properties of feasible solutions (e.g., requirements that solutions are paths, cycles, covers, etc.) end the family $\mathcal{G}$ defines additional properties (e.g., limits on the number of edges, etc.).

The primal problem is then stated as

$$(\mathrm{P}') \qquad \min\ C(X)$$
$$g(x) \geq b$$
$$X \in \mathcal{X}.$$

Assume that $Q = \mathcal{D} \subseteq \Phi^s_+$, where Φ^s_+ is the set of nondecreasing functions $\varphi : \mathbf{R}^s \to \mathbf{R}$. Consider a family of *l-relaxations* (see Libura, 1984), where for $q = \varphi,\ \varphi \in \Phi^s_+$, the relaxation (R_q) has the form

$$\min_{X \in \mathcal{X}} \{C(X) - \varphi(g(x)) + \varphi(b)\}.$$

The dual problem based on such a family, so called *l-dual problem*, is formulated in the following way:

$$(\mathrm{D}_l) \qquad \sup_{\varphi \in \mathcal{D}} \min_{X \in \mathcal{X}} \{C(X) - \varphi(g(x)) + \varphi(b)\}.$$

In a similar way the *s-dual problem*:

$$(\mathrm{D}_s) \qquad \sup_{\varphi \in \mathcal{D}} \min\{C(X) : \varphi(g(x)) \geq \varphi(b),\ X \in \mathcal{X}\}$$

is defined. It corresponds to the *s-relaxation*, which for $\varphi \in \Phi^s_+$ has the form:

$$\min\{C(X) : \varphi(g(x)) \geq \varphi(b),\ X \in \mathcal{X}\}.$$

If we take $\mathcal{D}$ as a set of nondecreasing affine real functions on $\mathbf{R}^s$, then we obtain the well known and widely used Lagrangean and surrogate relaxations of the problem (P′).

An important dual problem, called *f-dual problem* (see e.g. Tind, Wolsey, 1981; Libura, 1984), is obtained with the same framework when as a relaxation of (P′) the problem

$$\min\{\varphi(b) : X \in \mathcal{X}\}$$

is used, and the family of dual parameters Q is defined for some $\mathcal{D} \subseteq \Phi^s_+$ in the following way:

$$Q = \{\varphi \in \mathcal{D} : \varphi(g(X)) \leq C(X) \ \text{ for } X \in \mathcal{X}\}.$$

The f-dual problem has the form

$$(\mathrm{D}_f) \qquad \sup \varphi(b)$$
$$\varphi(g(x)) \leq C(X) \ \text{ for } X \in \mathcal{X}$$
$$\varphi \in \mathcal{D}. \qquad \square$$

Consider a pair of the primal problem (P′) and the dual problem (D), and corresponding optimal values $v(\mathrm{P})$ and $v(\mathrm{D})$. An important parameter defined for this pair is so called *duality gap* Δ, where

$$\Delta = v(\mathrm{P}) - v(\mathrm{D}).$$

If $\Delta = 0$, then an optimality condition which can be formulated is, in fact, the same as the sufficient optimality condition for simple relaxation stated in Section 2.2. By solving the dual problem and obtaining q^*, we determine the best relaxation (R_{q^*}), which now can be used in the procedure of obtaining a subset of the stability region, described in Section 2.2.

Not always zero duality gap can be achieved for a given family of relaxations. It can be shown that one can guarantee zero duality gap with dual problems described in Example 2.6, if the set of dual parameters is chosen in a proper way, but this leads to very difficult dual problems (see Tind, Wolsey, 1981; Libura, 1993).

If $\Delta > 0$, then the situation is more complicated, because the procedure mentioned may not be used directly to determine a subset of the stability region $P(X^o)$. In this case similar procedure gives a subset of so called *ϵ-stability region* (with $\epsilon = \Delta/(v(P) - \Delta)$). This term is used for a subset of data for which the solution X^o remains ϵ-optimal. Observe that the value of accuracy ϵ is not determined *a priori* but after the dual problem is solved.

The f-dual problems allow to formulate optimality conditions for the problem (P′) which are not only sufficient but also necessary (see Tind, Wolsey, 1981; Wolsey, 1981). They have the following form:

X^o is an optimal solution of the problem (P′) if and only if there exists a function $\varphi^o \in \Phi^s_+$ such that the following conditions hold:

(i) $g(x^o) \geq b$

(ii) $X^o \in \mathcal{X}$

(iii) $\varphi^o(g(x^o)) \leq C(X^o)$

(iv) $C(X^o) = \varphi^o(b)$.

The solution of f-dual problem is usually not unique and there are many functions satisfying the above conditions for a particular solution X^o. Given any such function from (i)-(iv) we have a description of subset of the stability region $P(X^o)$. The main disadvantage of this approach is that the f-dual problem is usually very difficult to solve. A possibility of overcoming this problem was presented in Wolsey (1981), where it was observed that various algorithms for solving the problem (P′) can be appropriately modified to produce not only the solution of the primal problem but also some solution of the f-dual problem.

In some cases the dual problem with nonzero duality gap can be also successfully used to perform some limited sensitivity analysis for the original problem. The approach is similar as in the case $\Delta = 0$ and consists in transferring the results of the sensitivity analysis obtained for 'the best' relaxation (R_{q^*}) to the original problem. The following example (see Libura, 1991;1993) illustrates such a possibility.

EXAMPLE 2.6

Consider again the symmetric traveling salesman problem in $G = (V, S)$ and its dual problem based on a 1-tree relaxation (see Example 4). Let q^* be the optimal solution of the dual problem and T be the minimum spanning 1-tree in the graph modified according to (15) with $q = q^*$. Denote by $C_{q^*}(X^o)$ the length of the optimal Hamiltonian tour in a modified graph. Similarly, $C_{q^*}(T)$ is the value of the minimum 1-tree in this graph, and $\Delta = C_{q^*}(X^o) - C_{q^*}(T)$.

Let $t^+(e)$ and $t^-(e)$ denote the tolerances of the length of edge $e \in S$ with respect the optimal Hamiltonian tour X^o. Denote by $t^+(e, q^*, T)$ and $t^-(e, q^*, T)$ similarly defined tolerances of length of edge e with respect to the optimal 1-tree T in the graph modified with the penalty vector q^*. Then (see Libura, 1991) for $e \in X^o \cap T$ we have

$$t^+(e) \geq t^+(e, q^*, T) - \Delta$$

and for $e \in (S \setminus X^o) \cap (S \setminus T^o)$ we have

$$t^-(e) \geq t^-(e, q^*, T) - \Delta.$$

Observe that $t^+(e) = \infty$ for $e \in S \setminus X^o$, and $t^-(e) = \infty$ for $e \in X^o$. For remaining tolerances we have only obvious nonnegativity conditions.

The problem of calculating tolerances $t^+(e, q^*, T)$ and $t^-(e, q^*, T)$ is relatively easy in comparison with calculating the tolerances $t^+(e)$, $t^-(e)$, so in this case the sensitivity analysis for relaxation provides a computationally inexpensive lower bounds for tolerances of the original problem. The role of the dual problem here is that it leads to the smallest value of Δ and in a consequence, the best quality of these bounds.

The tolerances of edges describe only the maximum individual changes of edge lengths. But in the case of the minimum 1-tree problem (and more generally, in the case of the problem of finding the minimum weight base of matroid) it can be proved (see Libura, 1993) that the lengths of all edges belonging to T can be simultaneously increased within their upper tolerances without forfeiting the optimality of T. Alternatively, the lengths of all edges in $S \setminus T$ can be simultaneously decreased within their lower tolerances. This property can be transferred to some extent to the primal problem and as it was shown in Libura (1993), the tolerances $t^+(e, q^*, T)$, $t^-(e, q^*, T)$ can be used to describe partially simultaneous changes of lengths in the original traveling salesman problem. □

3. Final remarks

This paper describes general sources of optimality conditions in discrete optimization in the context of sensitivity analysis. It seems rather unlikely to expect a significant progress in developing general methods yielding to these conditions. On the other hand, it seems reasonable to investigate optimality conditions for particular problems and exploit their structure. Another promising area of investigations is connected to the question as to what kind of optimality conditions are generated in the process of solving the problem with a particular algorithm.

References

CROWDER, H., JOHNSON, E.L., PADBERG, M. (1983) Solving large-scale zero-one linear programming problems. *Operations Research*, **31**, 803–834.

FIACCO A.V. (1983) *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming.* Academic Press, New York, London.

GAL T. (1979) *Postoptimal Analyses, Parametric Programming, and Related Topics.* McGraw-Hill, New York.

GARFINKEL, R.S., NEMHAUSER, G.L. (1972) *Integer Programming.* Wiley, New York.

GEOFFRION, A.M., NAUSS, R. (1977) Parametric and postoptimality analysis in integer programming. *Management Science*, **23**, 453–466.

LAWLER, E.L., LENSTRA, J.K., RINNOY KAN, A.G., SHMOYS, D. (1985) *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization.* Wiley.

LAWLER, E.L. (1976) *Combinatorial Optimization : Networks and Matroids.* Holt, Rinehart and Winston, New York.

LEONTEV, V.K. (1979) Stability of discrete linear problems. (in Russian) *Problemy Kibiernietiki* **35**, 169–184.

LIBURA, M. (1977A) Sensitivity analysis in integer programming. (in Polish) *Archiwum Automatyki i Telemechaniki*, **XXII**, 299–311.

LIBURA, M. (1977B) Sensitivity analysis for integer knapsack problem. (in Polish) *Archiwum Automatyki i Telemechaniki*, **XXII**, 313–322.

LIBURA, M. (1984) Integer programming duality and its applications in sensitivity analysis and approximate algorithms. (in Polish) *Archiwum Automatyki i Telemechaniki*, **XXIX**, 1-2, 75–92.

LIBURA, M. (1991A) Sensitivity analysis for minimum Hamiltonian path and traveling salesman problems. *Discrete Applied Mathematics*, **30**, 197–211.

LIBURA, M. (1991B) Sensitivity analysis for minimum weight base of matroid. *Control and Cybernetics*, **20**, 7–24.

LIBURA, M. (1993) *Sensitivity Analysis for Solutions of Discrete Optimization Problems.* (in Polish) Synpress, Warszawa.

LIBURA, M. VAN DER POORT, E.S., SIERKSMA, G., VAN DER VEEN, J.A.A. (1995) *Sensitivity analysis and k-best solutions for the traveling salesman problem.* Working Paper WP-1-1995, IBS PAN.

PIPER, C.J., ZOLTNERS, A.A. (1976) Some easy postoptimality analysis for zero-one programming. *Management Science*, **22**, 759–765.

SCHRIJVER A. (1986) *Theory of Linear and Integer Programming.* Wiley, Chichester.

TIND, J., WOLSEY, L.A. (1981) An elementary survey of general duality theory in mathematical programming. *Mathematical Programming*, **21**, 241–261.

WARD, J.E., WENDELL, R.E. (1990) Approaches to sensitivity analysis in linear programming. *Annals of Operations Research*, **27**, 3–38.

WELSH, D.J.A. (1976) *Matroid Theory.* Academic Press, London, New York, San Francisco.

WOLSEY L.A. (1981) Integer programming duality: Price functions and sensitivity analysis. *Mathematical Programming*, **20**, 173–195.