Stability analysis for a parameterized multicriteria Boolean linear programming problem

Yury Nikulin and Marko M. Mäkelä

Department of Mathematics, University of Turku, FI-20014 Turku (e-mail: yurnik@utu.fi and makela@utu.fi).

Abstract: A multicriteria boolean programming problem with linear cost functions in which initial coefficients of the cost functions are subject to perturbations is considered. For any optimal alternative, with respect to parameterized principle of optimality "from Condorcet to Pareto", an appropriate measure of the quality is introduced. This measure corresponds to the so-called stability function defined earlier for optimal solutions of a generic multicriteria combinatorial optimization problem with Pareto and lexicographic optimality principles. Various properties of such function are studied and maximum norm of perturbations for which an optimal solution preserves its optimality is calculated.

Keywords: Condorcet optimality, Pareto set, stability, quality measure, parameterization, multicriteria optimization.

1. INTRODUCTION

The stability theory has its roots originating from the definition of a well-posed mathematical problem given in Hadamard (1903), where it was stated that mathematical models of physical phenomena should include, among others, the property of a solution to depend continuously on the data, in some reasonable topology. In optimization a question of stability of a problem arises in the case where the set of feasible solutions (alternatives) and/or the objective (cost) function depend on parameters. The presence of such parameters in optimization models is due to many reasons, for instance inaccuracy of initial data, non-adequacy of models to real processes, errors of numerical methods, errors of rounding off and other factors. Thus it appears to be important to allocate classes of problems in which small changes of the input data lead to small changes of the result. The problems with such properties are called stable. It is obvious that many optimization problems arising in practice cannot be correctly formulated, analyzed and solved without exploiting the results of the stability theory.

It is not very surprising that many researchers focus on analyzing various aspects of stability for large classes of optimization problems. For example, one can find a vast annotated bibliography for sensitivity and post-optimal analysis in integer programming and combinatorial optimization problems in Greenberg (1998).

The main object while studying stability of multicriteria optimization problems is usually a set of optimal (sometimes referred to as efficient) solutions or alternatives, i.e. the set of feasible solutions which satisfy a given optimality principle. In the case where the partial criteria of the problem have an equal importance, the Pareto optimality principle, originally proposed in Pareto (1909), is more often used. Generally, a feasible solution is said to be Pareto optimal if there is no other feasible solution such that at least one its objective value getting better does not deteriorate any other objective values.

If we relax the demand of non-worsening objectives in such a way that worsening for some objective values is allowed but the number of objectives which values are allowed to be deteriorated is restricted above by the number of objectives with better values, then we get the concept of Condorcet optimality principle, originally proposed in Condorcet (1785).

It is clear that the set of optimal solutions defined by Condorcet optimality principle is a subset of the set of optimal solutions given by the Pareto optimality principles, i.e. Pareto optimality principle gives more freedom for solutions to become optimal compared to the Condorcet optimality principle.

A frequently used tool of stability theory and post-optimal analysis is so-called stability radius of some given optimal solution. In single objective optimization, it gives an upper bound on a subset of problem parameters for which this solution remains optimal (see Greenberg (1998) and bibliography therein). There are already similar investigations in multiobjective case, where the stability radius defines an extreme level of problem parameter perturbations preserving efficiency of the given solution. For example, in Emelichev et al. (2002) one can find a large survey on sensitivity analysis of vector unconstrained integer linear programming, where the stability radius is a key object under investigation.

It is important to note that even in single objective case the stability radius does not provide us with any information about the quality of a given solution in the case when problem data are outside of the stability region. Some attempts to study the quality of the problem solution in this case are connected with concepts of stability function, which was originally proposed in Libura (1999) and Libura (2000) for scalar combinatorial optimization problems. Later, some of the results were extended for the case of multicriteria combinatorial optimization problems with Pareto and lexicographic optimality principles in Libura et al. (2006). In Nikulin (2008), the similar questions were investigated under framework of game theory, more precisely a stability function for a coalition game with bans, linear payoffs, antagonistic strategies and parameterized principle of optimality "from Nash to Pareto" was studied.

Here we give an extension of the concept of stability function under the parameterized optimality principle "from Condorcet to Pareto". The paper is structured as follows. In Section 2, we consider a multicriteria Boolean linear programming problem. The problem consists in finding the set of optimal solutions, i.e. alternatives which are optimal with respect to the chosen optimality principle. In Section 3, for a given solution we introduce an appropriate relative error as a function of the norm of data perturbations. Afterwards, we define so called stability radius as an extreme level of perturbations of problem parameters for which the stability function is equal to zero. We give analytical formula to calculate the value of stability function and corresponding radius. In Section 4, a short example from data mining theory is considered to illustrate the way how the stability function can be used as an efficient tool for post-optimal analysis. Final remarks and conclusions appear in Section 5.

2. PROBLEM FORMULATION

We consider a problem with $m \geq 2$ cost functions representing the problem objectives. Let $X \subseteq 2^{\{0,1\}^n} \backslash \emptyset,\ |X| \geq 2$, be a set of feasible solutions or alternatives $x := \{x_1, ..., x_n\}^T \neq (0, 0, ..., 0)^T$, where n denotes the problem size.

For each solution $x \in X$, a vector of cost functions

$$f(C, x) := (f_1(C, x), ..., f_m(C, x))^T$$

consists of individual cost functions $f_i(C, x),\ i \in I_m := \{1, 2, ..., m\}$, which are defined as linear functions, i.e.:

$$f_i(C, x) := C_i x.$$

Here C_i is i-th row of matrix $C = [c_{ij}] \in \mathbf{R}_+^{m \times n}$, where $\mathbf{R}_+^{m \times n}$ is a set of $m \times n$ matrices with all elements being positive.

Without loss of generality, we assume that $f_i(C, x)$ are minimized on the set of feasible solutions X for each $i \in I_m$.

Contrary to the single objective case where the concept of optimal solution is unique, under multicriteria framework the concept of optimality may vary and is usually based on binary relations reflecting preferability of one solutions over others. In its turn, any binary relation generates a principle of optimality (in other terminology, sometimes referred as a choice function).

For any $x,\ x' \in X$ and $C \in \mathbf{R}_+^{m \times n}$, we put

$$[x, x', C]^+ := |\{i \in I_m :\ f_i(C, x) > f_i(C, x')\}| = |\{i \in I_m :\ C_i(x - x') > 0\}|;$$

$$[x, x', C]^- := |\{i \in I_m :\ f_i(C, x) < f_i(C, x')\}| = |\{i \in I_m :\ C_i(x - x') < 0\}|;$$

$$[x, x', C]^0 := |\{i \in I_m :\ f_i(C, x) = f_i(C, x')\}| = |\{i \in I_m :\ C_i(x - x') = 0\}|.$$

Obviously,

$$[x, x', C]^+ + [x, x', C]^- + [x, x', C]^0 = m. \quad (1)$$

The binary relations $x \prec x'$ of a strict preference between two feasible solutions x and x' (x' is preferred to x) are given according to the formulae:

- **Condorcet (majority) domination relation** $x \prec_\mu x'$:

$$[x, x', C]^+ > [x, x', C]^-; \quad (2)$$

- **Pareto domination relation** $x \prec_\pi x'$:

$$[x, x', C]^+ > (m - 1)[x, x', C]^-. \quad (3)$$

Note, that

$$[x, x', C]^+ > \zeta \cdot [x, x', C]^-,$$

where $\zeta \geq m - 1$, also defines the Pareto domination relation $x \prec_\pi x'$, however $m - 1$ in (3) is the smallest integer value of ζ which may guarantee $x \prec_\pi x'$.

A solution $x^* \in X$ is called **Condorcet optimal** if

$$\mu(x^*, C) = \emptyset,$$

where

$$\mu(x^*, C) := \{x \in X :\ x^* \prec_\mu x\}.$$

We will refer to the set of all Condorcet optimal solutions as the **Condorcet set** and denote it by $M^m(C)$.

A solution $x^* \in X$ is called **Pareto optimal** if

$$\pi(x^*, C) = \emptyset,$$

where

$$\pi(x^*, C) := \{x \in X :\ x^* \prec_\pi x\}.$$

We will refer to the set of all Pareto optimal solutions as the **Pareto set** and denote it by $P^m(C)$.

The Condorcet principle of optimality realizes the well-known procedure of decision-making by the majority of votes. It is easy to understand that the binary relation $\prec_\mu$ is not always transitive, not even for $m = 3$. Indeed, let $X = \{x_1, x_2, x_3\}$ and let

	$f_1(x)$	$f_2(x)$	$f_3(x)$
x_1 :	1	2	3
x_2 :	2	3	1
x_3 :	3	1	2

Recall that $f_i(x) \to \min_{x \in X},\ i \in I_3$. Then it is clear that $x_2 \prec_\mu x_1$ and $x_3 \prec_\mu x_2$, but at the same time $x_1 \prec_\mu x_3$. The requirement of majority rule then provides that none can be selected among x_1, x_2 and x_3. Therefore, the set $M^m(C)$ may be empty. This explains the well-known Condorcet paradox of voting originally mentioned in Condorcet (1785), which was later comprehensively analyzed by Kenneth Arrow based on the axiomatic approach to the mechanism of collective decision-making in Arrow (1963). Notice also that $P^m(C)$ is always non-empty due to the finite number of feasible solutions.