

# Constant Penalty Functions to Simplify Optimization of the Choquet Integral under Constraints

Tanja Magoč

Center for Bioinformatics and Computational Biology

University of Maryland

College Park, MD

Email: t.magoc@gmail.com

**Abstract**—The Choquet integral have been used in many decision making problems as the optimization function that needs to be maximized or minimized. The complexity of these optimization problems is often increased by numerous constraints that are imposed to the solution. Several methods have been developed for optimization under constraints, one of them being penalty functions method. This method transforms a constrained optimization problem in an equivalent unconstrained optimization problem by adding penalties to the solutions that do not satisfy one of more constraints. Penalty functions method is an iterative process, which adjusts penalty functions in each iteration, often yielding a large number of iterations until a satisfying solution is found. To speed the process of the Choquet integral optimization under constraints, we propose to use constant penalty functions. This approach requires only one iteration, and is therefore much faster than the generic penalty function method. In this paper, we derive the constant penalty function that guarantees that the optimal solution found in the first iteration satisfies all constraints.

*Keywords*—the Choquet integral; optimization; constant penalty function; decision making

## I. INTRODUCTION

The Choquet integral has often been used as an optimization function in decision making such as finance [1], geographic information systems [2], evaluation of mental workload [3], and environmental structure evaluation [4]. Many real-life and scientific problems can be formulated in terms of decision making. Some decisions are relatively straightforward. However, other problems may require more strategizing such as in the case of decision under uncertainty and multi-criteria decision making. Multi-criteria decision making seeks an optimal solution among a set of alternatives that are characterized by several attributes, i.e., criteria. For example, if an individual wants to buy a car, this individual would most probably consider the price, miles per gallon, safety rating, color, and other characteristics of cars. The first difficulty faced with different criteria used to make decisions is the domain of values that each criterion could take. Some criteria take qualitative values while others take quantitative values that might be discrete or continuous. The color could be expressed using qualitative values (e.g., {red, blue, black}); the safety rating could be expressed using discrete integers representing the number of stars rating a car's safety (e.g., {1,2,3,4,5}); and the price could be expressed in terms of a continuous variable taking values within an interval (e.g., [11000,50000]).

Furthermore, a decision maker usually has preferences within each criterion in terms of the values he/she would rather choose compared to other possible values of the criterion. For example, a car buyer would like to buy the car that is the cheapest, has the highest safety rating, travels the most miles per gallon of gas, and is of the decision buyer's favorite color. However, only in miracle situations, an available alternative (e.g., a car) is characterized by the best value for each criterion. In general, the alternative that reaches the maximum value of one attribute does not reach the maximum value of another attribute (e.g., the cheapest car usually does not have five stars safety rating). Thus, a decision maker has to accept some tradeoff between criteria, and maximize the total payoff of choosing a particular alternative. The total payoff could be calculated in various ways. Maximax, maximin, weighted sum, and the Choquet integral are most commonly used methods to combine values of each criterion into a global value. While maximax, maximin, and weighted sum are attractive approaches in decision making situation where criteria are mutually independent, the Choquet integral is well suited for situations where there are dependences between criteria used in decision making. Different approaches have been employed to optimize the Choquet integral including genetic algorithm [5], branch and bound method [6], and particle swarm optimization technique [7].

In many situations, certain constraints are imposed to the possible solution, which results in a more complex optimization problem. Numerous methods have been developed to deal with constraints while optimizing a function including simplex method ([8],[9]), Lagrange multiplier approach ([10],[11]), branch and prune method [12], and penalty functions technique [13]. The method of penalty functions relies on adding the penalty (i.e., decreasing the value of a maximizing function, or increasing the value of a minimizing function) at the points or intervals that do not satisfy constraints. Typically, the penalty function is defined as a polynomial (often the first degree polynomial) function of the inputs representing the values of the criteria considered in the decision making. The coefficients of the penalty function are adjusted in each iteration of the optimization process. However, in some situations, this functions could be replaced by a constant function, which simplifies optimization process while guaranteeing that the points not satisfying the constraint will not be selected as an

optimal solution.

In this paper, we first present the basics of the Choquet integral and its use in multi-criteria decision making (section 2). In the section 3, we present a few common constraints to an optimizing function such as the Choquet integral. The common method of penalty functions to deal with constraints is presented in the section 4. Finally, in the section 5, we describe how a constant function could be used as a penalty function in optimizing the Choquet integral. We conclude the paper with future work in extending the proposed method to other optimizing functions.

## II. THE CHOQUET INTEGRAL AS THE OPTIMIZATION FUNCTION IN MULTI-CRITERIA DECISION MAKING

A multi-criteria decision making (MCDM) problem seeks the optimal choice among a (usually finite) set of alternatives. It can formally be defined as a triple  $(X, I, (\succeq_i)_{i \in I})$  where

- $X \subset X_1 \times \dots \times X_n$  is the set of alternatives with each set  $X_i$  representing a set of values of the criterion  $i$ .
- $I$  is the (finite) set of criteria (or attributes).
- $\forall i \in I$ ,  $\succeq_i$  is a preference relation (a weak order) over  $X_i$ .

The next task is to “combine” the partial preference relations  $\succeq_i$  of an alternative into a global value for the alternative such that the final order of the alternatives is in the agreement with the decision maker’s partial preferences. The natural way to construct a global preference is by using a utility function for each criterion to reflect partial preferences of a decision-maker, and then combine these monodimensional utilities into a global utility function using an aggregation operator. The utility functions

$$u_i : X_i \rightarrow \mathbb{R} \quad (1)$$

such that

$$\forall x_i, y_i \in X_i, u_i(x_i) \geq u_i(y_i) \text{ if and only if } x_i \succeq_i y_i \quad (2)$$

map the values of all attributes onto a common scale (often the interval  $[0, 1]$ ). The existence of monodimensional utility functions is guaranteed under relatively loose hypotheses by the work presented in [14].

Numerous aggregation operators could be used to combine monodimensional utilities into a single number that represents the value of an alternative. Two simple approaches that correspond to optimistic and pessimistic behavior of the decision maker are maximax and maximin strategies, respectively, assuming that the goal of a decision-maker is to maximize the utility. The maximax method compares the utilities of all criteria of an alternative and chooses the highest utility value,  $\max_i u_i(x_i)$ , to represent the global utility of the alternative  $\mathbf{x} = (x_1, \dots, x_n)$ . This approach reflects an optimistic behavior of the decision-maker since he/she is concerned only with the attribute that has the highest utility for the given alternative. The maximax method tries to maximize the best criterion:

$$\max_{\mathbf{x} \in X} (\max_{i \in \{1, \dots, n\}} u_i(x_i)). \quad (3)$$

On the contrary, the maximin method reflects a pessimistic behavior of the decision-maker as the decision-maker is concerned only with the criterion that could result in the worst value. This method compares the utilities of all attributes of an alternative and chooses the lowest utility value,  $\min_i u_i(x_i)$ , to represent the global utility of the alternative  $\mathbf{x}$ . The decision-maker tries to maximize the value of the worst case scenario:

$$\max_{\mathbf{x} \in X} (\min_{i \in \{1, \dots, n\}} u_i(x_i)). \quad (4)$$

To allow for a position between these extremes when making a decision, a simple combination of maximax and maximin approaches is achieved by a weighted aggregation operator, named Hurwicz criterion [15],

$$\max_{\mathbf{x} \in X} (\alpha \max_i u_i(x_i) + (1 - \alpha) \min_i u_i(x_i)), \quad (5)$$

where  $\alpha \in [0, 1]$  is the weight given by the decision maker. This approach simplifies to the optimistic case if  $\alpha = 1$  and to the pessimistic case if  $\alpha = 0$ .

These simple approaches are very tempting to use for quick decisions. However, they focus only on one or two criteria and ignore the impact of other characteristics of alternatives. This approach often does not suit a given situation in which a decision has to be made. Thus, we usually need to consider more complex aggregation operators that take into consideration all attributes. The simplest and most natural of them is a weighted sum approach, in which the decision-maker is asked to provide weights,  $w_i$ , that reflect the importance of each criterion. Thus, the global utility of the alternative  $\mathbf{x} = (x_1, \dots, x_n) \in X$  is given by

$$u(\mathbf{x}) = \sum_{i=1}^n w_i u_i(x_i). \quad (6)$$

The best alternative is the one that maximizes this value. Even though this approach is attractive due to its low complexity, it can be shown that using an additive aggregation operator, such as weighted sum, is equivalent to assuming that all the attributes are independent [16]. In practice, this is usually not realistic and therefore, we need to turn to non-additive approaches, that is to aggregation operators that are not linear combinations of partial preferences.

Before approaching non-additive methods, we give the definition of a non-additive measure, a tool for building non-additive aggregation operators such as the Choquet integral. For practical purposes, we restrict ourselves to measures defined on a finite set  $I$ .

**Definition 1. [Non-additive (fuzzy) measure]:** Let  $I$  be the set of attributes and  $\mathcal{P}(I)$  the power set of  $I$ . A set function  $\mu : \mathcal{P}(I) \rightarrow [0, 1]$  is called a non-additive measure (or a fuzzy measure) if it satisfies the following three axioms:

- (1)  $\mu(\emptyset) = 0$  : the empty set contains no information.
- (2)  $\mu(I) = 1$  : the maximal set contains all the information.
- (3)  $\mu(B) \leq \mu(C)$  if  $B, C \subset \mathcal{P}(I)$  and  $B \subset C$ : a new criterion added cannot make the importance of a coalition (a set of criteria) diminish.

A non-additive integral, such as the Choquet integral [17], is a type of a general averaging operator that can model the behavior of a decision maker. The decision-maker provides a set of values of importance, this set being the values of the non-additive measure on which the non-additive integral is computed from. Formally, the Choquet integral in finite case is defined as follows:

**Definition 2. [Choquet integral]:** Let  $\mu$  be a fuzzy measure on  $(I, \mathcal{P}(I))$  and an application  $f : I \rightarrow \mathbb{R}^+$ . The Choquet integral of  $f$  w.r.t.  $\mu$  is defined by:

$$(C) \int_I f d\mu = \sum_{i=1}^n (f(\sigma(i)) - f(\sigma(i-1)))\mu(A_{(i)}), \quad (7)$$

where  $\sigma$  is a permutation of the indices in order to have  $f(\sigma(1)) \leq \dots \leq f(\sigma(n))$ ,  $A_{(i)}$  is  $A_{(i)} = \{\sigma(i), \dots, \sigma(n)\}$ , and  $f(\sigma(0)) = 0$ , by convention.

It can be shown that many aggregation operators can be represented by Choquet integrals with respect to some fuzzy measure. Although the Choquet integral is well suited for quantitative measurements, it has a major drawback. The decision maker needs to input a value of importance of each subset of attributes, that is total of  $2^n$  values, which is intractable due to its exponential complexity. However, we can overcome intractability by using 2-additive measure to limit the complexity to a  $O(n^2)$  (as shown in [18]) and still get accurate results.

Before giving the definition of a 2-additive measure, we need to define notion of interaction indices of orders 1 and 2. The importance of a criterion (or the interaction index of degree 1) is best described as the value this attribute brings to each coalition it does not belong to. It is given by the Shapley value [19]:

**Definition 3. [Shapley value]:** Let  $\mu$  be a non-additive measure over  $I$ . The Shapley value of index  $i$  is defined by:

$$v(i) = \sum_{B \subset I \setminus \{i\}} \gamma_I(B) [\mu(B \cup \{i\}) - \mu(B)] \quad (8)$$

with

$$\gamma_I(B) = \frac{(|I| - |B| - 1)! \cdot |B|!}{|I|!} \quad (9)$$

and  $|B|$  denoting the cardinal of  $B$ .

While the Shapley value gives the importance of a single attribute to the entire set of attributes, the interaction index of degree 2 represents the interaction between two attributes, and is defined by ([20],[21]):

**Definition 4. [Interaction index of degree 2]:** Let  $\mu$  be a non-additive measure over  $I$ . The interaction index between  $i$

and  $j$  is defined by:

$$I(i, j) = \sum_{B \subset I \setminus \{i, j\}} (\xi_I(B) \cdot (\mu(B \cup \{i, j\}) - \mu(B \cup \{i\}) - \mu(B \cup \{j\}) + \mu(B)))$$

with

$$\xi_I(B) = \frac{(|I| - |B| - 2)! \cdot |B|!}{(|I| - 1)!}. \quad (10)$$

The interaction indices belong to the interval  $[-1, +1]$  and

- $I(i, j) > 0$  if the attributes  $i$  and  $j$  are complementary;
- $I(i, j) < 0$  if the attributes  $i$  and  $j$  are redundant;
- $I(i, j) = 0$  if the attributes  $i$  and  $j$  are independent.

Even though we can define interaction indices of any order, defining the importance of attributes and the interaction indices between two attributes is generally enough in multi-criteria decision making problems. Thus, 2-additive measures constitute a feasible and accurate tool in this setting. The formal definition of a 2-additive measure follows [20]:

**Definition 5. [2-additive measure]:** A non-additive measure  $\mu$  is called 2-additive if all its interaction indices of order equal to or larger than 3 are null and at least one interaction index of degree two is not null.

It has been shown [22] that the Shapley values and the interaction indices of order two offer us an elegant way to represent a Choquet integral. Therefore, in a decision-making problem, we can ask the decision maker to give the Shapley values,  $I_i$ , and the interaction indices,  $I_{ij}$ , and then use the Choquet integral with respect to a 2-additive measure,  $\mu$ , to obtain the aggregation operator:

$$(C) \int f d\mu = \sum_{I_{ij} > 0} (f(i) \wedge f(j)) I_{ij} + \sum_{I_{ij} < 0} (f(i) \vee f(j)) |I_{ij}| + \sum_{i=1}^n f(i) (I_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}|).$$

This form of the Choquet integral is a practical approach to many multi-criteria decision making situations ([1],[2],[3],[4]).

### III. EXAMPLES OF CONSTRAINTS

Numerous constraints could be imposed on the input variables (i.e., variables representing the values of criteria in multi-criteria decision making). A typical constraint in many optimization problems is that (some) variables could not take negative values,  $x_i \geq 0$ . As we have seen in the previous section, the values of variables in a multi-criteria decision making are often mapped to the interval  $[0,1]$ ; thus the non-negativity constraint is imposed to all variables. Moreover, in this particular situation, the maximum value of each variable is

1, thus another constraint should be imposed to each variable,  $x_i \leq 1$ .

However, not every constraint should be imposed to all variables. For example, when buying a car, we might want to buy a car only if it is relatively safe, so we might decide to consider only cars with a safety rating of at least three stars. Using utility functions, the rating of 3 stars would map to the value of 0.5. Thus, the constraint that we need to consider is  $x_{\text{safety}} \geq 0.5$ .

All of the above examples of constraints are hard constraints, so they must always be satisfied. On the other hand, there might be soft constraints, which we would like to be satisfied, but it is not necessary that they are satisfied. For example, we might not want to buy a car that costs more than \$20,000, thus yielding the constraint  $x_{\text{price}} \leq 20,000$ . However, if there is a great deal available for a car that is a bit more expensive, we might be willing to somewhat sacrifice this soft constraint and buy a car that doesn't satisfy this constraint.

Finally, a constraint does not have to be imposed to only a single variable. It could be imposed to a combination of variables, such as a sum of two or more variables. Let us consider the problem of selecting an optimal financial investment portfolio where the variables represent the portion of money invested in each financial investment assets. A multi-criteria decision making problem is designed to determine the best distribution of money among  $n$  assets so that the largest amount of money will be earned in a year [2]. Certain amount of money needs to be invested in each of  $n$  assets in order to reduce risk of losing all money ([23],[24]) (we assume here that the  $n$  assets were previously selected from thousands of investment assets that exist in the world, so that  $n$  is a realistic number for the size of an investment portfolio). Thus, all the variables should be positive (rather than just non-negative),  $x_i > 0$ , and strictly smaller than 1,  $x_i < 1$ , since not all money could be invested in the same asset. Moreover, not more than all money could be invested in the portfolio (we are making simplifying assumptions that an investor can not borrow money from one institution to invest it in another asset), thus the sum of all money should be no more than 1 (i.e., 100% of money),  $\sum_{i=1}^n x_i \leq 1$ . This is a typical example of a hard constraint when the value of one variable depends on the values of other variables.

Soft constraints could also be imposed to the variables that depend on each other. For example, we might require that at least 50% of all money is invested in the financial assets,  $\sum_{i=1}^n x_i \geq 0.5$ . However, in some situations (e.g., in economic recessions, gain of investing money could be very low compared to the risk of investing money), we might accept to invest smaller amount of money if no good investment option is available.

Even though the presented examples are all represented by linear inequalities, a more complex constraints could be imposed to the solution of an optimization problem.

#### IV. PENALTY FUNCTIONS

Several methods have been developed to deal with constraints imposed to the solution of an optimization problem. The method of penalty functions ([11],[13]) transforms a constrained optimization problem into an unconstrained optimization problem, which is accomplished by adding penalties in a minimization problem (or subtracting penalties in a maximization problem) to the solutions that do not satisfy one or more constraints. A penalty function is assigned to each constraint and is adjusted through iterative process of the algorithm. Each penalty function is designed so that its value is equal to 0 if the constraint is satisfied, and greater than 0 if the constraint is not satisfied. Let  $c_i(\mathbf{x})$  be the  $i^{\text{th}}$  constraint. Then the penalty function is defined as:

$$\begin{aligned} p_i(\mathbf{x}) &= 0 \text{ if } c_i(\mathbf{x}) \text{ is satisfied.} \\ p_i(\mathbf{x}) &> 0 \text{ if } c_i(\mathbf{x}) \text{ is not satisfied.} \end{aligned} \quad (11)$$

Let us consider a constrained optimization problem:

$$\begin{aligned} &\text{maximize } f(\mathbf{x}) \\ &\text{subject to the following constraints:} \\ &g_i(\mathbf{x}) \geq 0, i = 1, \dots, m; \\ &h_j(\mathbf{x}) = 0, j = 1, \dots, s. \end{aligned}$$

This problem is transformed into maximizing the function  $f'(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^m r_i(\mathbf{x}) - \sum_{j=1}^s q_j(\mathbf{x})$ ,

where  $r_i$  and  $q_j$  are penalty functions. The goal of penalty functions is to decrease the values of the objective function at the elements of the search space that do not satisfy all the constraints. However, finding the maximum of the function  $f'(\mathbf{x})$  does not guarantee that the solution is found since the maximum of this function still might not satisfy all the constraints. In order to find a solution that maximizes the objective function and satisfies all the constraints, the penalty functions are changed and the objective function is evaluated again. The process is iterated until an optimal solution that satisfies all the constraints is reached.

The method could be simply described by the following algorithm:

```

PenaltyFunctions
{
    for each constraint
    {
        initialize penalty function
    }
    while (solution not found)
    {
        evaluate objective function
        for each constraint
        {
            add penalties to objective function
        }
        increase penalty functions
    }
}

```

Note that each minimization problem,  $\min f(\mathbf{x})$ , could be transformed in the maximization problem  $\max (-f(\mathbf{x}))$ . Thus, in this paper, without loss of generality, we assume that we have a maximization problem.

## V. A CONSTANT PENALTY FUNCTION TO OPTIMIZE THE CHOQUET INTEGRAL WITH RESPECT TO A 2-ADDITIVE MEASURE

Even though the penalty functions method could be very effective in solving constrained optimization problems, it suffers a few drawbacks, one of them being the complexity of selecting a good starting penalty function. It is not clear how to initialize a penalty function. The most common approach used to define a penalty function is to square the subrange of the objective function that does not satisfy the constraint. The next question is how to adjust the penalty functions after each iteration of the optimization process. Usually, the adjusted penalty function is simply a multiply of the original penalty function. However, this method might lead towards numerous iterations until the optimal solution is found, which yields another drawback, namely how to define the end of the iterative optimization process.

While the general form of the penalty function method is designed so that it could be applied in any constrained optimization problem regardless of the type of the optimization function, it sometimes takes a large amount of time to obtain the optimal solution. Instead, the penalty function could be defined by a much simpler form such as a constant function. This form is well suited for optimization functions that are bounded on the entire interval that input variables could take. Note that it is enough that the optimization function be bounded from above in an maximization problem, and from below in a minimization problem. One such an optimization function is the Choquet integral with respect to a 2-additive measure constructed by the utility method, which maps the values of all criteria to the interval  $[0,1]$  (as described in the section 2).

In this situation, the penalty applied should be a value that is greater or equal to any value that the Choquet integral with respect to a 2-additive measure could possibly take. The reason for this penalty is that if the constraint is not satisfied at a particular point in the search space, we do not want to consider this point as a possible solution to the problem since the problem becomes impractical if a (hard) constraint is not satisfied. Thus applying the penalty whose value is larger than the maximum value the Choquet integral could theoretically take will ensure that the value of the objective function is a negative value and therefore will certainly not be selected as a maximum value.

To determine the value of the penalty, let us consider the formula of the Choquet integral with respect to a 2-additive

measure:

$$(C) \int f d\mu = \sum_{I_{ij} > 0} (f(i) \wedge f(j)) I_{ij} + \sum_{I_{ij} < 0} (f(i) \vee f(j)) |I_{ij}| + \sum_{i=1}^n f(i) (I_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}|).$$

If we consider the first two sums, we conclude that each interaction index of order 2 is covered by at most one of the sums (the interaction indices of 0 are not covered at all). Each element of each summation is a minimum or a maximum of two utility values multiplied by the absolute value of the corresponding interaction index of order 2. Each utility value and the absolute value of each interaction index belong to the interval  $[0,1]$ . Using interval computation arithmetics (see e.g., [25]), we conclude that each element of any of the first two summation is at most equal to one. The maximum number of elements in two summations is the number of interaction indices of order 2, which is  $\frac{n \cdot (n-1)}{2}$ , where  $n$  is the number of criteria considered. Further, the third summation is the multiplication of a Shapley value and a utility value decreased by a certain amount. Thus, to find its maximum value, we can disregard the subtracted part, and only find the maximum value of the utility value multiplied by a Shapley value. As both of these values fall into the interval  $[0,1]$ , the maximum value of each element of the third summation is one, and thus, the value of the summation can not exceed  $n$ . The maximum value that the Choquet integral can theoretically take is therefore the sum of the maximum value the first two sums could take,  $\frac{n \cdot (n-1)}{2}$ , and the maximum value the third sum could take,  $n$ :

$$\frac{n \cdot (n-1)}{2} + n = \frac{n^2 + n}{2} < \frac{n^2 + n^2}{2} = n^2. \quad (12)$$

Thus, the penalty that should be applied for a violation of each hard constraint should be equal to at least  $n^2$  to ensure that even if just one constraint is not satisfied at a particular point, the penalty will decrease the value of the optimization function at this point below zero, and therefore this point will surely not yield the maximum value of the optimization function.

However, when considering soft constraints, we would like that they are satisfied, but would accept a solution that yield an optimal value even if the soft constraint is not satisfied. In this situation, we could consider a lower penalty function for soft constraints than for hard constraints. Moreover, since we would allow an optimal solution to not satisfy soft constraints, we do not want to decrease the value of the optimization function below zero if only one soft constraint is not satisfied. In some situations, we might allow for multiple soft constraints to be unsatisfied. If we allow  $i$  soft constraints to be unsatisfied, the constant penalty function should be defined as

$$\frac{n^2}{i-1}. \quad (13)$$

The proposed penalty function will assure that if  $i$  soft constraints are unsatisfied, the optimization function still might

yield a positive value; however, if the  $(i + 1)^{\text{th}}$  soft constraint is not satisfied, the optimization function will yield a non-positive value.

Furthermore, if not all soft constraints have the same importance, we might define penalty functions differently for each soft constraints. For example, if not satisfying a soft constraint  $C1$  is less important than not satisfying a soft constraint  $C2$ , then the penalty function for the constraint  $C1$  should be smaller than the penalty function for the constraint  $C2$ .

Finally, in some situations, we might prefer a solution that barely missed to satisfy a soft constraint over a solution that is far from satisfying a soft constraint. In these situations, we apply a penalty function to the objective function. However, the constant penalty function is increased by the distance from satisfying the constraint. The distance is usually calculated as the Euclidean distance between the proposed solution and the constraints.

## VI. CONCLUSION

Numerous situation is life require decision making of some sort. Some situations are so simple that we do not even realize that we are making a decision. However, many situations require more strategies to be employed when making a decision. Different approaches have been taken to select an optimal alternative among available options, one of them being the Choquet integral with respect to a 2-additive measure, which is especially well suited for situations where criteria considered in making decision are mutually dependent.

Optimization of the Choquet integral is not a trivial process itself. The complexity of this process is increased even more by imposing constraints to possible solutions. The constraints could range from simple linear equalities or inequalities applied to one variable at a time to non-linear equalities or inequalities that combine multiple variables. Moreover, constraints could be hard (i.e., constraints that must be satisfied) or soft (i.e., constraints that we would like to be satisfied, but would accept that some of them not be satisfied if that yields a much better optimal solution than if these constraints were satisfied).

Numerous approaches have been used to deal with constraints, one of them being the penalty functions method, which adds penalty to the value of the optimization function if a constraint is not satisfied by subtracting (adding in minimization problem) the value of the penalty function from the maximizing function. The penalty functions method could result in a large number of iterations due to difficulty in adjusting the penalty from one iteration to another. To simplify this method, a constant penalty function could be applied when maximizing the Choquet integral since the Choquet integral is bounded from above. In the case of hard constraints, the penalty function should take the value  $n^2$ , where  $n$  is the number of criteria used in the decision making problem to ensure that each point that does not satisfy even one hard constraint will yield a non-positive value, and therefore will not be selected as an optimal value. In the case of soft constraints,

the constant penalty function should be defined in proportion to the number of soft constraints allowed, and could also be increased based on the distance from satisfying the constraint to give a better chance of being selected as the optimum to points that are closer to satisfying the constraints than the points that are further away from satisfying the constraints.

In this paper, we presented how constant penalty functions could be used in the optimization of the Choquet integral with respect to a 2-additive measure. However, a similar approach could be applied to any other function that is bounded from above in the maximization problem, or bounded below in a minimization problem.

## REFERENCES

- [1] T. Magoc, "New algorithms for optimal portfolio selection," Ph.D. dissertation, University of Texas at El Paso, 2009.
- [2] T. Magoc, A. Kassin, and R. Romero, "A line of sight algorithm using fuzzy measures," *Proceedings of the North American Fuzzy Information Processing Society Annual Meeting*, 2010.
- [3] M. Grabisch, H. Prade, E. Raufaste, P. and Terrier, "Application of the Choquet integral to subjective mental workload evaluation," *Automated Systems Based on Human Skill and Knowledge*, vol. 9, no. 1, 2008.
- [4] Y. Nakamori and N. Iwamoto, "Analysis of environmental evaluation structure by the Choquet integral model," *International Joint Conference of the 4th IEEE International Conference on Fuzzy Systems and the 2nd International Fuzzy Engineering Symposium*, 1995.
- [5] I. Saad, S. Hammadi, M. Benrejeb, and P. Borne, "Choquet integral for criteria aggregation in the flexible job-shop scheduling problems," *Mathematics and Computers in Simulation*, vol. 76, no. 5-6, 2008.
- [6] L. Galand, P. Perny, P., and O. Spanjaard, "A branch and bound algorithm for Choquet optimization in multicriteria problems," *Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems*, vol. 634, no. 4, pp. 355-365, 2010.
- [7] T. Xiao, N. Zhang, and D.-M. Huang, "Solution to the model of nonlinear multiregression based on generalized Choquet integral and its application in estimating forest volume," *Proceedings of the 2010 International Conference on Computer Application and System Modeling*, 2010.
- [8] G. B. Dantzig, "Maximization of linear function subject to linear inequalities," in T. C. Koopmans (ed.), *Activity Analysis of Production and Allocation*, John Wiley & Sons, New York, 1951, pp. 339-347.
- [9] D. Gale, "Linear programming and the simplex method," *Notices of the AMC*, 2007, vol. 54, no. 3, pp. 364-369.
- [10] G. Arfken, "Lagrange multipliers," in *Mathematical methods for physicists*, 3<sup>rd</sup> ed. Orlando, Florida: Academic Press, 1985, pp. 945-950.
- [11] G. Thomas and N. Sonilal, "Use of penalty function or Lagrange multipliers for constrained deconvolution," pp. 1559-1562.
- [12] A. Goldsztejn, "A branch and prune algorithm for the approximation of non-linear AE-solution set," *Proceedings of the 2006 ACM Symposium on Applied Computing*, 2006.
- [13] A. E. Smith and D. W. Coit, "Penalty functions," *handbook of evolutionary computation*, joint publication of Exford University Press and Institute of Physics Publishing, 1995.
- [14] D. Krantz, R. Luce, P. Suppes, and A. Tverski, *Foundations of measurement*. Academic Press, 1971.
- [15] J.-Y. Jaffray and M. and Jeleva, "Information processing under imprecise risk with the Hurwicz criterion," *Proceedings of International Symposium on Imprecise Probability: Theories and Applications*, Prague, Czech Republic, 2007.
- [16] F. Modave and M. Grabisch, "Preferential independence and the Choquet integral," 8<sup>th</sup> *International Conference on the Foundations and Applications of Decision Under Risk and Uncertainty*, Mons, Belgium, 1997.
- [17] G. Choquet, "Theory of capacities," *Annales de l'Institut Fourier*, no. 5, 1953.
- [18] M. Ceberio and F. Modave, "Interval-based multicriteria decision making," in *Modern Information Processing: from Theory to Applications*, B. Bouchon-Meunier, G. Coletti, and R. R. Yager (Editors), Elsevier ed., 2006.
- [19] L. S. Shapley, "A value for  $n$ -person games," in *Contributions to the Theory of Games*, H. W. Kuhn and A. W. Tucker (Editors), no. 2, pp. 307-317, Princeton University Press, 1953.

- [20] D. Denneberg and M. Grabisch, "Shapley value and interaction index," *Mathematics of Interaction Index*, 1996.
- [21] M. Grabisch and M. Roubens, "Application of the Choquet integral in multicriteria decision making," in *Fuzzy Measures and Integrals: Theory and Applications*, M. Garbisch, T. Murofushi, and M. Sugeno (Editors), Physica Verlag, 2000.
- [22] M. Grabisch, "The interaction and Mobius representation of fuzzy measures on finite spaces,  $k$ -additive measures: a survey," in *Fuzzy measures and integrals: Theory and applications*, M. Grabisch, T. Murofushi, and M. Sugeno (Editors), Physica Verlag, 2000.
- [23] J. Estrada, *Finance in a nutshell: A no-nonsense companion to the tools and techniques of finance*. Prentice Hall, 2005.
- [24] D. Kansas, *The Wall Street Journal: Complete money & investing guide book*. Three Rivers Press, New York, 2005.
- [25] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*. Springer-Verlag, London, 2001.