Fuzzy Optimization and Decision Making, 4, 119–130, 2005.

Learning Weights in the Generalized OWA Operators

GLEB BELIAKOV gleb@deakin.edu.au
School of Information Technology, Deakin University, 221 Burwood Hwy, Burwood, 3125, Australia

Abstract. This paper discusses identification of parameters of generalized ordered weighted averaging (GOWA) operators from empirical data. Similarly to ordinary OWA operators, GOWA are characterized by a vector of weights, as well as the power to which the arguments are raised. We develop optimization techniques which allow one to fit such operators to the observed data. We also generalize these methods for functional defined GOWA and generalized Choquet integral based aggregation operators.

Keywords: aggregation operators, ordered weighted averaging, Choquet integral, fuzzy sets

1. Introduction

In a recent article in this journal R. Yager introduced a generalization of Ordered Weighted Averaging (OWA) operators, called GOWA (Yager (2004)). He studied many properties of these operators and several special cases, including generalized means, Hurwicz operator, min and max, ordered weighted geometric operator, and so on. Further, Yager proposed a generalization of the discrete Choquet integral, which is also commonly used as an aggregation operator (Denneberg (1994), Yager and Kacprzyk (1997)). At the end of the paper he poses the problem of learning the weights of GOWA from empirical data, as in (Filev and Yager (1998)), which now also includes learning the parameter λ characterizing GOWA.

This paper addresses the issue of learning the parameters of GOWA, as well as a more general instance of learning the fuzzy measure characterizing generalized Choquet aggregation operator. We provide two formulations of the problem of learning, one which is reduced to linear programming, and the second one which is reduced to a quadratic programming problem. We use similar techniques as those of (Beliakov (2002, 2003b), Beliakov et al (2004b)), designed for ordinary OWA operators and Choquet integrals.

The next section reviews GOWA operators and generalized Choquet aggregation as proposed in (Yager (2004)). In Section 3 we pose the problem of learning weights from the data, and in Section 4 we discuss various techniques to identify the parameters of GOWA operators. We consider functional defined GOWA operators in Section 5, and generalized Choquet integral based operators in Section 6. Finally we further generalize GOWA in Section 7, and show how the earlier algorithms for generalized means can be adapted for this case. The presented methods were incorporated into the AOTool software package, www.deakin.edu.au/~gleb/ao-tool.html, which provides many methods of identifying aggregation operators from empirical data, subject to user defined restrictions.

2. Review of GOWA Operators

In (Yager (2004)) Yager proposed a generalization of Ordered Weighted Averaging operators, called GOWA,

$$F(x_1,\ldots,x_n)=\left(\sum_{i=1}^{n} w_i x_{(i)}^{\lambda}\right)^{1/\lambda}, \tag{1}$$

where $x_{(i)}$ denotes the ith largest element of $x=(x_1,\ldots,x_n)^t$, $\lambda\in R$, and w is the vector of non-negative weights, which add to one. The case of $\lambda=1$ corresponds to ordinary OWA operators. We consider a special case when the arguments are drawn from the unit interval $I=[0,1]$, and $F: I^n\to I$.

As it was shown on multiple occasions (Yager (1988), Yager (1993), special cases of OWA operator include max, min and simple average, which correspond to the following vectors of weights: $w=(1,0,0,\ldots,0)$, $w=(0,\ldots,0,1)$ and $w=(\frac{1}{n},\ldots,\frac{1}{n})$. For GOWA these vectors of weights result in max, min and generalized mean operators.

Another special case occurs when $w=(\alpha,0,\ldots,0,1-\alpha)$. It corresponds to Hurwicz type aggregation. For instance, when $\lambda=-1$ we obtain a sort of harmonic mean operator

$$F(x_1,\ldots,x_n)=\frac{x_{(1)}x_{(n)}}{\alpha x_{(n)}+(1-\alpha)x_{(1)}}.$$

The attitudinal character of the GOWA operator (or the measure of orness) is defined in (Yager (2004)) as

$$AC(W;\lambda)=\left(\sum_{i=1}^{n} w_i\left(\frac{n-i}{n-1}\right)^{\lambda}\right)^{1/\lambda}=\frac{1}{n-1}\left(\sum_{i=1}^{n} w_i(n-i)^{\lambda}\right)^{1/\lambda}. \tag{2}$$

The values close to 0 correspond to min-like aggregation and values close to 1 correspond to max-like aggregation.

Notice that GOWA exhibit some similarity to the generalized quasi-linear means (Dyckhoff and Pedrycz (1984)), but are symmetrized by using a permutation of the arguments.

Yager (Yager (2004)) poses the problem of identification of the weights w and the parameter λ from empirically collected data. Since GOWA is a special class of functions, they require special regression techniques to fit them to the data. The remainder of this paper addresses this issue in detail.

3. Problem of Weights Identification

GOWA operators certainly bring much flexibility into modelling aggregation process in the decision making. We are now interested in specifying particular operators (i.e., particular vector of weights and parameter λ) for concrete situations. It is often not enough to specify just some general properties the aggregation operator must satisfy to adequately model a given aggregation process, as these properties define an infinite family of operators. We need to use the requirement of empirical fit (Zimmermann (1996)), that is to fit an operator with certain properties to some sort of empirical data. The data can be collected in an experiment, by questioning experts in the field (Zimmermann (1996), Beliakov and Warren (2001), Sicilia et al (2003)), or by conducting a mental experiment: what would be the aggregated value if the argument x has these specific values?

The problem of identification of aggregation operators has been studied in (Beliakov (2002, 2003b)). Consider the data set consisting of K $(n+1)$-tuples $\{(x^k, y^k)\}_{k=1}^K$, where $x^k \in I^n$ are observed arguments of F and $y^k \in I$ are observed aggregated values. The goal is to identify the vector of weights and the parameter λ, such that GOWA operator $F(x)$ in (1) fits all the data best,

$$F(x^k) = y^k, \quad k = 1, \ldots, K.$$

Not all the data can be fit exactly due to observation inaccuracies, or perhaps inadequacy of GOWA model in this specific situation, hence we require the above system of equations to be solved in the approximate sense as described below.

Consider first the case of a fixed λ. We are interested in using fast and proven linear regression techniques. Let us linearize the data. Take the linearized data set $\{(z^k, (y^k)^\lambda)\}_{k=1}^K$, where the components $z_i^k = x_{(i)}^k$. Then we find the weights by minimizing the least squares criterion

$$\min_w \left(\sum_{k=1}^K \left[\sum_{i=1}^n w_i (z_i^k)^\lambda - (y^k)^\lambda \right]^2 \right)^{1/2}, \tag{3}$$

subject to the restrictions $0 \le w_i \le 1, i = 1, \ldots, n, \sum w_i = 1$.

Alternatively, we can minimise the absolute difference between the predicted and observed values, the method frequently used in robust regression:

$$\min_w \sum_{k=1}^K \left| \sum_{i=1}^n w_i (z_i^k)^\lambda - (y^k)^\lambda \right|, \tag{4}$$

subject to the same restrictions. In this case approximation process is less sensitive to outliers. In addition, one may specify one further constraint on the orness measure

$$\sum_{i=1}^n w_i (n-i)^\lambda = ((n-1)AC)^\lambda,$$

where $AC \in I$ is the desired value of the orness measure. Solution to problems (3) and (4) is discussed in the next section.

Let us now consider the case of unknown parameter λ which also needs to be fitted to the data. We need to minimize expression (3) (or (4)) with respect to both w and λ. We can represent this as a bi-level optimization problem

$$\min_{w,\lambda} \Phi(w,\lambda) = \min_{\lambda} \min_{w} \left(\sum_{k=1}^{K} \left[\sum_{i=1}^{n} w_i (z_i^k)^\lambda - (y^k)^\lambda \right]^2 \right)^{1/2}, \tag{5}$$

or

$$\min_{w,\lambda} \Phi(w,\lambda) = \min_{\lambda} \min_{w} \sum_{k=1}^{K} \left| \sum_{i=1}^{n} w_i (z_i^k)^\lambda - (y^k)^\lambda \right|, \tag{6}$$

subject to the same set of linear constraints on w, and λ unrestricted. At the outer step we perform optimization with respect to one nonlinear variable λ, and at the inner step, we solve (3) or (4) with a fixed λ using efficient methods described in the next section. The need to solve (3) or (4) a large number of times as a sub-problem of the minimization with respect to λ is a significant factor when choosing the method of solution, which has to be robust and extremely fast.

Let us examine the outer problem of minimization with respect to λ. Firstly, we need to consider special cases

$\lambda \to \infty$, which translates into $F(x) = \max_{j:w_j>0} x_{(j)}$,

$\lambda \to -\infty$, which translates into $F(x) = \min_{j:w_j>0} x_{(j)}$,

$\lambda \to 0$, which translates into $F(x) = \prod_{j=1}^{n} x_{(j)}^{w_j}$, called Ordered Weighted Geometric (OWG) operator (Chiclana et al (2000), Xu and Da (2002)).

Note that the first two cases do not correspond to the ordinary max and min operators, as the weights are also taken into account. For instance, by letting $w = (0, 0, \ldots, 1)$, we obtain min and not max in the first case. These special cases have to be implemented explicitly in any software that uses GOWA, rather than using the general formula (1), because of numerical instabilities when λ approaches any of these critical values.

Another important issue is that the function to be minimized in (5) or (6) is not necessarily convex with respect to λ, and as such may possess multiple local minima. This is despite the fact that for any *fixed* vector of weights w, $\Phi(w,\lambda)$ will have a unique minimum with respect to λ, see the discussion of the 'main property' in (Dyckhoff and Pedrycz (1984)). This does not imply the uniqueness of the local minimum with respect to both variables λ and w. This phenomenon is well known, and can be illustrated on the example of fuzzy c-means functional used in clustering (Bezdek (1981)), which is convex with respect to two subsets of variables, but not convex with respect to all variables. A consequence of this is the existence of potentially large number of local minima of $\Phi(w,\lambda)$. It is therefore incorrect to use a simple local optimization tool when solving the outer problem with respect to λ, like Newton's or conjugate gradient method, as it may converge to a locally, but not globally optimal solution. A global optimization method is required.

For univariate case there are many global optimization methods (Horst et al (2000)). For instance Piyavsky–Shubert method (Pijavski (1972)), Cutting angle (Rubinov (2000), Beliakov (2003a)) or even a simple grid search can all be used.

4. Optimization Techniques to Identify the Parameters

We start with the case of a fixed λ, as it is important for solving the inner sub-problem in (5) or (6). Problem (3) is a quadratic programming problem, which can be solved using standard general algorithms. However, several specialized algorithms for this instance of the problem are available (Lawson and Hanson (1995)). (3) is known as the linear least squares with equality and inequality constraints problem LSEI (Haskell and Hanson (1981), Hanson and Haskell (1982)). The method in (Hanson and Haskell (1982)) uses the active set and penalty function approach to deal with both kinds of constraints. It is formulated as the system of equations and inequalities,

$$\text{Solve } \mathbf{Aw} \approx \mathbf{b}, \mathbf{Cw} \geq \mathbf{d}, \mathbf{Ew} = \mathbf{e}, \tag{7}$$

where the first system is solved in the least squares sense (by orthogonal factorization). It is therefore appropriate to represent (3) in the form (7) for algorithmic purposes, and then apply the LSEI algorithm from (Hanson and Haskell 1982). Note that the elements of the matrix **A** are given by $A_{ki} = (z_i^k)^\lambda$, and the elements of **b** are $b_k = (y^k)^\lambda$.

The problem (4) is easily converted to a linear programming problem. By splitting a real number u into positive and negative parts $u = u_+ - u_-$, we can write $|u| = u_+ + u_-$. Then the objective function in (4) becomes

$$\min_{u,w} \sum_{k=1}^{K} (u_+^k + u_-^k),$$

and together with the set of linear constraints

$$u_+^k - u_-^k = \sum_{i=1}^{n} w_i (z_i^k)^\lambda - (y^k)^\lambda, \quad k = 1, \ldots, K,$$

and $0 \leq w_i \leq 1$, $\sum w_i = 1$, $u_+^k, u_-^k \geq 0$, we obtain an LP problem, see (Watson (2000)) for details.

There are special versions of the simplex algorithm designed for this instance of LP (Barrodale and Roberts (1980)). The problem is written in the form of (7), but now the system of approximate equations is solved in the smallest absolute deviation sense, and the algorithm from (Barrodale and Roberts (1980)) is applied.

To solve the general problem of minimizing $\Phi(w, \lambda)$ in (5) or (6) with respect to both variables, we apply a global optimization algorithm, such as Piyavsky–Shubert method, at the outer level, and for each value of λ we compute the value of $\Phi(w, \lambda)$

by solving (3) or (4) as described above. The minimizer of $\Phi(w, \lambda)$ yields the optimal weights and optimal parameter λ.

5. Functional Defined GOWA Operators

Yager discusses various methods for generating OWA and GOWA weights via a monotone function $f: I \to I$, for which $f(0) = 0$, $f(1) = 1$, called basic unit monotonic (BUM) function (Yager (1996, 2004)). Using these functions we can generate the OWA weights as $w_i = f(\frac{i}{n}) - f(\frac{i-1}{n})$. This method allows one to define not just one, but a whole family of $2, 3, \ldots,$-variate OWA operators, as the vector of weights of any dimension can be generated from f. This is explored in (Beliakov et al (2004b)), where such families (called generalized aggregation operators (Calvo et al (2002)) are identified.

An interesting problem arises: can we identify a BUM function f rather than the individual weights from the data? If so, this brings an opportunity to identify from a particular data set not only one GOWA operator which fits best this data set, but the whole family of GOWA operators (of any dimension $n \geq 2$). For OWA operators, this problem was resolved positively in (Beliakov et al (2004b)).

Let us approximate $f(t)$ with a monotone linear regression spline (i.e., a piecewise linear continuous monotone function) (Beliakov (2000))

$$S(t) = \sum_{j=1}^{J} c_j B_j(t), \tag{8}$$

where $B_j(t)$ are some basis functions (like B-splines, or their linear combinations, as in (Beliakov (2000)), and $c \in R^J$ is a vector of spline coefficients that need to be identified from the data. Monotonicity of the spline S translates into a simple condition of non-negativity of spline coefficients c with suitably chosen basis functions (Beliakov (2000)).

Then we have

$$\begin{aligned}
F(x) &= \left(\sum_{i=1}^{n} \left[f\left(\frac{i}{n}\right) - f\left(\frac{i-1}{n}\right) \right] x_{(i)}^{\lambda} \right)^{1/\lambda} \\
&= \left(\sum_{i=1}^{n} \left[\sum_{j=1}^{J} c_j B_j\left(\frac{i}{n}\right) - \sum_{j=1}^{J} c_j B_j\left(\frac{i-1}{n}\right) \right] x_{(i)}^{\lambda} \right)^{1/\lambda} \\
&= \left(\sum_{i=1}^{n} \left[\sum_{j=1}^{J} c_j \left(B_j\left(\frac{i}{n}\right) - B_j(\frac{i-1}{n}) \right) \right] x_{(i)}^{\lambda} \right)^{1/\lambda} \\
&= \left(\sum_{i=1}^{n} \left[\sum_{j=1}^{J} c_j G_j(i) \right] x_{(i)}^{\lambda} \right)^{1/\lambda} = \left(\sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} x_{(i)}^{\lambda} G_j(i) \right] \right)^{1/\lambda},
\end{aligned} \tag{9}$$

where $G_j(i) = B_j(\frac{i}{n}) - B_j(\frac{i-1}{n})$. After linearization of the data set, we obtain a system of K equations with respect to unknown c

$$\sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} (z_i^k)^\lambda G_j(i) \right] = (y^k)^\lambda, \quad k = 1, \dots, K,$$

which we resolve in the least squares (or least absolute deviation) sense, by minimizing

$$\min_{c,\lambda} \left(\sum_{k=1}^{K} \left[\sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} (z_i^k)^\lambda G_j(i) \right] - (y^k)^\lambda \right]^2 \right)^{1/2}, \tag{10}$$

or

$$\min_{c,\lambda} \sum_{k=1}^{K} \left| \sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} (z_i^k)^\lambda G_j(i) \right] - (y^k)^\lambda \right|, \tag{11}$$

subject to non-negativity of c_j. For BUM functions continuous on $[0, 1]$ we add boundary conditions

$$S(0) = \sum_{j=1}^{J} c_j B_j(0) = 0, \quad S(1) = \sum_{j=1}^{J} c_j B_j(1) = 1,$$

whereas for BUM functions continuous on $]0, 1[$, we add

$$S(0) = \sum_{j=1}^{J} c_j B_j(0) \geq 0, \quad S(1) = \sum_{j=1}^{J} c_j B_j(1) \leq 1.$$

After examining both problems, we notice that for a fixed λ (10) is a quadratic programming problem LSEI, and (11) can be converted to a linear programming problem as discussed above. In both cases we write each problem in the form (7), where the elements of the matrix $\mathbf{A}$ are given by $A_{kj} = \sum_{i=1}^{n} (z_i^k)^\lambda G_j(i)$, and the other two matrices correspond to linear constraints on c given above (non-negativity and boundary conditions).

For variable λ we write (10), (11) as bi-level optimization problems, where at the outer level we perform optimization with respect to λ (using global optimization), and at the inner level we solve LSEI or LP problem.

6. Generalized Choquet Aggregation

Choquet integral is frequently used as an aggregation tool (Grabisch et al (1995), Benvenuti and Mesiar (2000), Calvo et al (2002)). The Choquet integral based aggregation operator is defined as

$$C_v(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_{(i)}[v(H_i) - v(H_{i+1})], \tag{12}$$

where $v : 2^X \to I$ is a fuzzy measure on the set $X = \{X_1, X_2, \ldots, X_n\}$, which is a monotonic (i.e. $v(S) \leq v(T)$ whenever $S \subseteq T$) set function satisfying $v(\emptyset) = 0, v(X) = 1$, $H_i = \{X_{(1)}, X_{(2)}, \ldots, X_{(i)}\}$, $X_{(i)}$ denotes the ith largest variable (rather than its value), and $H_{n+1} = \emptyset$ by convention. Equation (12) can also be written as (Grabisch (1997)), (Grabisch (2000)), p. 110,

$$C_v(x_1, \ldots, x_n) = \sum_{i=1}^{n} \left[x_{(i)} - x_{(i-1)}\right] v(H_i). \tag{13}$$

In this notation, C_v is a *linear* function of the coefficients of the fuzzy measure $v(H_i)$.

In multicriteria decision making, Choquet aggregation explicitly models the importance of not only individual criteria, but of their subsets, as well as various interactions between the criteria. In the context of learning aggregation operators from data, identification of Choquet aggregation operator is equivalent to identification of the fuzzy measure $v(T)$, described by 2^n coefficients. This problem was addressed in (Grabisch et al (1995), Sicilia et al (2003), Beliakov et al (2004b)).

Yager proposes in (Yager (2004)) a generalized Choquet aggregation operator

$$C_{v,\lambda}(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} x_{(i)}^{\lambda}[v(H_i) - v(H_{i+1})]\right)^{1/\lambda}. \tag{14}$$

It is not difficult to see that this implies

$$C_{v,\lambda}(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n}[x_{(i)}^{\lambda} - x_{(i-1)}^{\lambda}]v(H_i)\right)^{1/\lambda}. \tag{15}$$

The sum in the brackets is again a linear function of the fuzzy measure coefficients. Thus we can apply the methods of fuzzy measure identification studied in (Grabisch et al (1995), Sicilia et al (2003), Beliakov et al (2004a, b)), with one distinction that the data are linearized (i.e., taking $\{z_i^k, (y^k)^{\lambda}\}$). The problem becomes a quadratic or linear programming problem for a fixed λ, and it is solved as a bi-level optimization problem if λ also has to be identified from the data.

In the same fashion as in (Grabisch et al (1995), Beliakov et al (2004a, b)), we can add further conditions on the fuzzy measure, such as k-additivity (Grabisch (1997)), sub- or super-additivity, substitutivity of certain variables, and so on, which all translate into linear restrictions on the values of v. Furthermore, following (Beliakov et al (2004b)) we can study symmetric k-additive generated fuzzy measures (whose coefficients are defined by a generating function similar to BUM), after linearizing the data.

7. Further Generalization

We mentioned that the GOWA operators resemble generalized quasi-linear means (Dyckhoff and Pedrycz (1984)), but involve permutation of the arguments. Quasi-linear means are a special case of generalized quasi-arithmetic means (Aczel (1969)) defined as

$$F(x_1, \dots, x_n) = g^{-1}\left(\sum_{i=1}^{n} w_i g(x_i)\right),$$

where g is a continuous strictly monotone function. In this special case we took $g(x) = x^\lambda$.

It makes sense to further generalize GOWA by replacing x^λ with a general continuous strictly monotone function $g(x)$. We obtain

$$F(x_1, \dots, x_n) = g^{-1}\left(\sum_{i=1}^{n} w_i g(x_{(i)})\right). \tag{16}$$

Many of the results concerning GOWA in (Yager (2004)) can be directly extended for the case of operators in the form (16), especially the ones concerning the behavior of the operator for various vectors of weights. The orness measure in this general case is written as

$$AC(w) = g^{-1}\left(\sum_{i=1}^{n} w_i g\left(\frac{n-i}{n-1}\right)\right).$$

Operators (16) offer even more flexibility in modelling aggregation process. Let us consider the task of identifying the vector of weights w and the function g simultaneously. For this approximate function g with a monotone linear spline (8). Our goal is to determine the coefficients of the spline, which must be positive to enforce strict monotonicity of g, as well as the weights w, which are non-negative and add to one. Let us write (16) as

$$g(F(x_1, \dots, x_n)) = \sum_{i=1}^{n} w_i g(x_{(i)}).$$

By using spline representation we have

$$\sum_{j=1}^{J} c_j B_j(F(x_1, \dots, x_n)) = \sum_{i=1}^{n} w_i \sum_{j=1}^{J} c_j B_j(x_{(i)}) = \sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} w_i B_j(x_{(i)})\right],$$

which we re-write as

$$\sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} w_i B_j(x_{(i)}) - B_j(F(x_1, \dots, x_n)) \right] = 0$$

By using our dataset, we obtain a system of K equations, needed to be solved in the least squares (or least absolute deviation) sense

$$\sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} w_i B_j(z_i^k) - B_j(y^k) \right] = 0, \quad k = 1, \dots, K.$$

Thus we minimize

$$\min_{c,w} \left(\sum_{k=1}^{K} \left[\sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} w_i B_j(z_i^k) - B_j(y^k) \right] \right]^2 \right)^{1/2}, \tag{17}$$

or

$$\min_{c,w} \sum_{k=1}^{K} \left| \sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} w_i B_j(z_i^k) - B_j(y^k) \right] \right|, \tag{18}$$

subject to the mentioned linear restrictions on the values of c, w.

Similarly to the case of generalized means, studied in (Beliakov (2003b)), for a fixed c (i.e., fixed $S(x)$) we have either a quadratic or linear programming problem to find w, and for a fixed w, we have a quadratic or linear programming problem to find c. However if we consider both c, w as variables, we obtain a difficult global optimization problem, similar to the one that arises in fuzzy c-means clustering (Bezdek (1981)).

Let us formulate a bi-level optimization problem

$$\min_{w} \min_{c} \left(\sum_{k=1}^{K} \left[\sum_{j=1}^{J} c_j \left[\sum_{i=1}^{n} w_i B_j(z_i^k) - B_j(y^k) \right] \right]^2 \right)^{1/2}.$$

At the outer level we have a global optimization problem which we solve using the cutting angle method (Beliakov (2003a), Rubinov (2000)), and at the inner level we have a quadratic programming problem that we solve using LSEI algorithm (Hanson and Haskell (1982)). For (18) we apply a similar approach. If the number of variables is not large (say, $n \leq 6$), cutting angle method is an efficient tool for deterministic global optimization, which locates the global minimum of (17) or (18). This approach is implemented for the case of generalized means in AOTool software, and is easily adapted for GOWA operators (16) by using permuted data values z^k instead of the original data x^k.

8. Conclusion

This paper addresses the problem of identification of parameters of GOWA operators proposed by Yager in (Yager (2004)). We developed a range of special regression tools that allow one to fit GOWA operators to empirical data. These tools rely on efficient solution of a bi-level optimization problem. We have shown that by linearizing the data, the inner problem can be converted to a linear or quadratic programming problem, and solved by standard algorithms.

Further we considered functional defined GOWA operators and generalized Choquet integral based operators, and formulated similar bi-level optimization problems for these cases. The optimization methods presented in this paper were successfully implemented, and included into the software package A0Tool (available from www.deakin.edu.au/˜gleb/aotool.html). A0Tool implements a large number of methods for identification of aggregation operators from the empirical data, including those for triangular norms, uninorms, means, OWA, general aggregation operators and Choquet integral based operators. The new methods presented here will complement this range.

References

Aczel, J. (1969). *On Applications and Theory of Functional Equations*. New York: Academic Press.

Barrodale, I. and F. Roberts. (1980). "Algorithm 552: Solution of the Constrained l_1 Linear Approximation Problem," *ACM Transactions Mathematical Software* 6, 231–235.

Beliakov, G. (2000). "Shape Preserving Approximation Using Least Squares Splines," *Approximation Theory and Applications* 16, 80–98.

Beliakov, G. (2002). "Monotone Approximation of Aggregation Operators Using Least Squares Splines," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 659–676.

Beliakov, G. (2003a). "Geometry and Combinatorics of the Cutting Angle Method," *Optimization* 52(4–5), 379–394.

Beliakov, G. (2003b), "How to Build Aggregation operators from data?". *International Journal of Intelligent Systems* 18, 903–923.

Beliakov, G., T. Calvo, E. García Barriocanal, and M. Sicilia. (2004a), 'Choquet Integral-Based Aggregation of Interface Usability Parameters: Identification of Fuzzy Measure'. In: *6th International Conference on Optimization Techniques and Applications*. Ballarat, Australia.

Beliakov, G., R. Mesiar, and L. Valášková. (2004b). "Fitting Generated Aggregation Operators to Empirical Data," *International Journal of Uncertainty, Fuzziness and Knowledge- Based Systems* 12, 219–236.

Beliakov, G. and J. Warren. (2001), "Appropriate Choice of Aggregation Operators in Fuzzy Decision Support Systems," *IEEE Transactions On Fuzzy Systems* 9, 773–784.

Benvenuti, P. and R. Mesiar. (2000), "Integrals with Respect to a General Fuzzy Measure," In M. Grabisch, T. Murofushi, and M. Sugeno (eds.) *Fuzzy Measures and Integrals. Theory and Applications*. Heidelberg: Physica-Verlag, pp. 205–232.

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*, Advanced applications in pattern recognition. New York: Plenum Press.

Calvo, T., A. Kolesarova, M. Komornikova, and R. Mesiar. (2002). "Aggregation Operators: Properties, Classes and Construction Methods," In T. Calvo, G. Mayor, and R. Mesiar (eds.) Aggregation Operators. *New Trends and Applications*. Heidelberg, New York: Physica-Verlag, pp. 3–104.

Chiclana, F., F. Herrera, and E. Herrera-Viedma. (2000). "The Ordered Weighted Geometric Operator: Properties and Applications," In *8th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems*. Madrid, pp. 985–991.

Denneberg, D. (1994). *Non-Additive Measure and Integral*. Dordrecht: Kluwer.

Dyckhoff, H. and W. Pedrycz. (1984). "Generalized Means as Model of Compensative Connectives," *Fuzzy Sets and Systems* 14, 143–154.

Filev, D. and R. Yager. (1998). "On the Issue of Obtaining OWA Operator Weights," *Fuzzy Sets and Systems* 94, 157–169.

Grabisch, M. (1997). "k-order Additive Discrete Fuzzy Measures and Their Representation," *Fuzzy Sets and Systems* 92, 167–189.

Grabisch, M. (2000). "The Interaction and Moebius Representation of Fuzzy Measures on Finite Spaces, k-Additive Measures: A Survey," In M. Grabisch, T. Murofushi, and M. Sugeno (eds.) *Fuzzy Measures and Integrals. Theory and Applications*. Heidelberg: Physica-Verlag, pp. 70–93.

Grabisch, M., H. Nguyen, and E.Walker. (1995). *Fundamentals of Uncertainty Calculi, with Applications to Fuzzy Inference*. Dordrecht: Kluwer.

Hanson, R. and K. Haskell. (1982). "Algorithm 587. Two Algorithms for the Linearly Constrained Least Squares Problem," *ACM Transactions Mathematical Software* 8, 323–333.

Haskell, K. and R. Hanson. (1981). "An Algorithm for Linear Least Squares Problems with Equality and Nonnegativity Constraints," *Mathematical Programming* 21, 98–118.

Horst, R., P. Pardalos, and N. Thoai. (eds.) 2000. "*Introduction to Global Optimization*, Vol. 48 of *Nonconvex Optimization and its Applications*," 2nd Edition. Dordrecht: Kluwer Academic Publishers.

Lawson, C. and R. Hanson. (1995). *Solving Least Squares Problems*. Philadelphia: SIAM.

Pijavski, S. (1972). "An Algorithm For Finding the Absolute Extremum of a Function," *USSR Computational Mathematics and Mathematical Physics* 2, 57–67.

Rubinov, A. (2000). *Abstract Convexity and Global Optimization*, Vol. 44 of *Nonconvex optimization and its applications*. Dordrecht; Boston: Kluwer Academic Publishers.

Sicilia, M., E. Garcia Barriocanal, and T. Calvo. (2003). "An Inquiry-Based Method for Choquet integral-based Aggregation of Interface Usability Parameters," *Kybernetica* 39, 601–614.

Watson, G. (2000). "Approximation in normed linear spaces," *Journal of Computational and Applied Mathematics* 121, 1–36.

Xu, Z. and Q. Da. (2002). "The Ordered Weighted Geometric Averaging Operator," *International Journal of Intelligent Systems* 17, 709–716.

Yager, R. (1988). "On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decision Making," *IEEE Transactions on Systems, Man and Cybernetics* 18, 183–190.

Yager, R. (1993). "Families of OWA operators," *Fuzzy Sets and Systems* 59, 125–148.

Yager, R. (1996). "Quantifier Guided Aggregation Using OWA Operators," *International Journal of Intelligent Systems* 11, 49–73.

Yager, R. (2004). "Generalized OWA Aggregation operators," *Fuzzy Optimization and Decision Making* 3, 93–107.

Yager, R. and J. Kacprzyk (eds.) (1997). *The Ordered Weighted Averaging Operators. Theory and Applications*. Boston: Kluwer.

Zimmermann, H.-J. (1996). *Fuzzy Set Theory – and its Applications*. Boston: Kluwer.