

# Using 2-Additive Measures in Nonlinear Multiregressions

L. Scott Bock and Zhenyuan Wang

**Abstract**—When a nonlinear integral with respect to a signed fuzzy measure is used in multiregression, people face a serious problem that, comparing to the number of variables (attributes), there are exponentially many unknown parameters in the model. However, in many real-world problems, the higher-order interactions among the variables can be omitted, and then only consider the second-order one with an acceptable small error in the result. Thus, a 2-additive measure based on the Möbius transformation and its inverse can be used to replace the signed fuzzy measure. In such a way, the complexity of the computation will be significantly reduced.

**Index Terms**—2-additive, genetic algorithm, möbius transformation, nonlinear multiregression,

## I. INTRODUCTION

A nonlinear multiregression model based on the generalized Choquet integral with respect to signed fuzzy measures has already been created[1]. However, the complexity of the algorithm is very high[2]. In a situation where interaction amongst 3 attributes or more is inconsequential we can use a 2-additive measure[3]. In order to reduce the number of unknown variables we use the Möbius transformation

$$\nu(A) = \sum_{B \subset A} (-1)^{|A-B|} \mu(B)$$

and its inverse

$$\mu(A) = \sum_{B \subset A} \nu(B)$$

setting  $\nu(A) = 0$  where  $|A| \geq 3$

Using a  $\lambda$ -fuzzy measure  $\mu$  we can see how  $\nu$  gives us the pure interaction of the attributes.

Let  $A = \{x_1\}$  and  $B = \{x_2\}$ ,

$$\text{then } \mu(A \cup B) = \mu(A) + \mu(B) + \lambda\mu(A)\mu(B),$$

$$\nu(A) = \mu(A), \quad \nu(B) = \mu(B), \quad \text{and}$$

$$\begin{aligned} \nu(A \cup B) &= -\mu(A) - \mu(B) + \mu(A \cup B) \\ &= -\mu(A) - \mu(B) + \mu(A) + \mu(B) + \lambda\mu(A)\mu(B) \\ &= \lambda\mu(A)\mu(B) \end{aligned}$$

In the standard multiregression model[1], if we have  $n$  predictive attributes, there are  $2n + 2^n$  unknown variables that we must search for. By omitting the higher-order interactions with a Möbius transformation, we can reduce the number of unknown variables to  $2n + n(n+1)/2 + 1$ .

## II. THE MODEL

Let  $x_1, x_2, \dots, x_n$  be predictive attributes and  $y$  be the objective attribute. Denote  $X = \{x_1, x_2, \dots, x_n\}$ . The data consists of  $l$  observations of  $x_1, x_2, \dots, x_n$  and  $y$ , and have a form as

$$\begin{array}{cccccc} x_1 & x_2 & \dots & x_n & y & \\ \hline f_{11} & f_{12} & \dots & f_{1n} & y_1 & \\ f_{21} & f_{22} & \dots & f_{2n} & y_2 & \\ & & & & \vdots & \\ f_{l1} & f_{l2} & \dots & f_{ln} & y_l & \end{array}$$

Where row  $f_{j1} \ f_{j2} \ \dots \ f_{jn} \ y_j$  is the  $j$ -th observation of attributes  $x_1, x_2, \dots, x_n$  and  $y$ ,  $j=1, 2, \dots, l$ . Positive integer  $l$  is called the size of the data, and should be much larger than  $n$ , usually, at least 5 times of  $2^n$ . Each observation of  $x_1, x_2, \dots, x_n$  can be regarded as a function  $f: X \rightarrow (-\infty, \infty)$ . Thus, the  $j$ -th observation of  $x_1, x_2, \dots, x_n$  is denoted by  $f_j$ , and we write  $f_{ji} = f_j(x_i)$ ,  $i=1, 2, \dots, n$ ,  $j=1, 2, \dots, l$ .

The interaction among attributes toward the objective attribute is described by a set function  $\mu$  defined on the power set of  $X$  satisfying the condition of vanishing at the empty set, i.e.,  $P(X) \rightarrow (-\infty, \infty)$  with  $\mu(\emptyset) = 0$ . Set function  $\mu$  is called

Manuscript received February 4, 2006.

L. Scott Bock is with Department of Mathematics, University of Nebraska at Omaha, Omaha, NE 68182 USA (email: lsbock@mail.unomaha.edu)

Zhenyuan Wang is with Department of Mathematics, University of Nebraska at Omaha, Omaha, NE 68182 USA (email: zhenyuanwang@mail.unomaha.edu)

a signed fuzzy measure.

The generalized Choquet integral of a function  $f: X \rightarrow (-\infty, \infty)$ , with respect to a signed fuzzy measure,  $\mu$ , is defined by

$$\int f d\mu = \int_{-\infty}^0 [\mu(F_\alpha) - \mu(X)] d\alpha - \int_0^{\infty} \mu(F_\alpha) d\alpha$$

when not both terms on the right side are infinite, where set  $F_\alpha = \{x \mid f(x) \geq \alpha\}$ , and is called the  $\alpha$ -cut set of function  $f$ .

The non-linear multiregression model is

$$y = c + \int (a + bf) d\mu + N(0, \sigma^2)$$

where  $c$  is a constant, both  $a$  and  $b$  are real-valued functions defined on  $X$ ,  $f$  is an observation of  $x_1, x_2, \dots, x_n$ ,  $\mu$  is a signed fuzzy measure, and  $N(0, \sigma^2)$  is a normally distributed random perturbation with expectation 0 and variance  $\sigma^2$ . Functions  $a$  and  $b$  can be expressed as vectors, i.e.,  $a = \{a_1, a_2, \dots, a_n\}$  and  $b = \{b_1, b_2, \dots, b_n\}$ . They should satisfy the following constraints:

$$a_i \geq 0 \text{ for } i = 1, 2, \dots, n, \text{ with } \min_{1 \leq i \leq n} a_i = 0;$$

$$-1 \leq b_i \leq 1 \text{ for } i = 1, 2, \dots, n \text{ with } \max_{1 \leq i \leq n} |b_i| = 1$$

In order to reduce the number of unknown variables we use the Möbius transformation and its inverse.

In this multiregression model, the regression coefficients are constant  $c$ , all elements of vectors  $a$  and  $b$ , and  $v(A)$  for every  $A \in P(X) \setminus \phi$  such that  $|A| \leq 2$ . These regression coefficients may be determined by using a genetic algorithm. Finally,  $\sigma^2$  may be estimated by the regression residual.

### III. THE ALGORITHM

(1) Input  $n$ ,  $l$ , and the data.

(2) Set the value for each of the following:

$\lambda$ : The bit length of each gene... Default is 10.

$p$ : The population size. It should be a large positive even integer. Its default is 200.

$\alpha$  and  $\beta$ : The probabilities used in a random switch to control the choice of genetic operators for producing offspring from selected parents. They should satisfy the condition that  $\alpha \geq 0$ ,  $\beta \geq 0$ , and  $\alpha + \beta \leq 1$ . Their defaults are 0.2 and 0.6 respectively.

$A$  and  $B$ : The maximum number of mutations and crossover points respectively, which are selected by a genetic switch. Their defaults are  $4n$  and  $2n$ .

$\varepsilon$  and  $\delta$ : Small positive numbers used in the stopping controller. Their defaults are  $10^{-6}$  and  $10^{-9}$  respectively.

$w$ : The limit number of generations that have not made significant progression. Its default is 30.

$$(3) \text{ Calculate } \sigma_y^2 = \frac{1}{l} \sum_{j=1}^l (y_j - \bar{y})^2$$

$$\text{where } \bar{y} = \frac{1}{l} \sum_{j=1}^l y_j$$

(4) Randomly create the initial population that consists of  $p$  chromosomes. Each chromosome consists of  $2n + n(n+1)/2 + 1$  genes denoted by  $g_1, g_2, \dots, g_{2n+n(n+1)/2+1}$ . The first  $n$  of them represent vector  $a$ , the next  $n$  represent vector  $b$ , the next  $n(n+1)/2$  represent the  $v$ 's, and the last of which represent  $c$ . Each gene consists of  $\lambda$  bits and represents a number in  $[0, 1)$ . Initialize counter WT by 0, and SE by  $\sigma_y^2$ .

(5) Decode each chromosome to get vectors  $a$  and  $b$  and the values for the  $v$ 's and  $c$  thusly:

$$a_i = \frac{g_i - m(g)}{(1 - g_i)(1 - m(g))},$$

$$b_i = \frac{2g_{n+i} - 1}{M(g)},$$

$$v_j = \tan\left(\frac{\pi}{2}(2g_{2n+j} - 1)\right)$$

$$c = \tan\left(\frac{\pi}{2}(2g_{2n+n(n+1)/2+1} - 1)\right)$$

for  $i = 1, 2, \dots, n$ , and  $j = 1, 2, \dots, n(n+1)/2$

where  $m(g) = \min_{1 \leq k \leq n} g_k$  and

$$M(g) = \max_{1 \leq k \leq n} |2g_{n+k} - 1|$$

(6) Create  $\bar{v}_i$  for  $i = 1, 2, \dots, 2^n - 1$  thusly:

for each  $i$  write it as a binary string  $S_i = k_{i,1}k_{i,2} \dots k_{i,2^n-1}$

set  $j = 1$

for each  $i = 1, 2, \dots, 2^n - 1$

if  $S_i$  contains more than 2, 1's  $\bar{v}_i = 0$ ,

else  $\bar{v}_i = v_j$  and  $j \in j+1$ .

(7) Create  $\mu$ 's using the Inverse Möbius function:

$$\mu_i = \sum_{k_{i,j}=1} \bar{v}_j$$

where

$$\mu_i = \mu(\{x_j | k_{ij} = 1\})$$

(8) The residual error of the  $r$ -th chromosome in the current population is denoted by  $\bar{\sigma}_r^2$ . Let  $m(\sigma^2) = \min_{1 \leq r \leq p} \bar{\sigma}_r^2$ . Set

$R = \{r | m(\sigma^2) = \bar{\sigma}_r^2\}$  and save  $a$ ,  $b$ ,  $c$ , and  $\mu$ 's of  $r$ -th chromosome for all  $r \in R$ . Calculate  $\bar{\sigma}_r^2$  thusly:

$$\bar{\sigma}_r^2 = \sum_{j=1}^l (\bar{y}_j - \bar{y})^2$$

where  $\bar{y}_j = \int f_j d\mu$

and is calculated as

$$\bar{y}_j = \sum_{i=1}^{2^n-1} \{[0 \vee (\min_{k_{ii}=1} (a_i + b_i f_{ii}) - \max_{k_{ii}=0} (a_i + b_i f_{ii}))]\mu_i$$

(9) If  $m(\delta^2) \leq \varepsilon \sigma_y^2$ , then go to (16), otherwise, take the next step.

(10) If  $SE - m(\sigma_y^2) \leq \delta \sigma_y^2$ , then  $WT + 1 = WT$ , otherwise  $0 = WT$

(11) If  $WT > w$ , then go to (16); otherwise, take the next step

(12) Sort the chromosomes in the population by  $\bar{\sigma}_r^2$  in increasing order and renumber  $1, 2, \dots, p$  then set

$$p_r = \frac{p+1-r}{2} \text{ for } r = 1, 2, \dots, p.$$

(13) According to the probability distribution  $p_r$  (by using a random switch) select two different chromosomes in the

current population as the parents. Randomly select one operator among the mutation with probability  $\alpha$  and with the number of bits mutated chosen by a random switch from 1 to  $A$ , the crossover with probability  $\beta$  and with the number of crossover points chosen by a random switch from 1 to  $B$ , and the two-point realignment with probability  $1 - \alpha - \beta$  to produce two new chromosomes as the offspring.

(14) Repeat step (13) for  $p/2$  times totally to get  $p$  new chromosomes. Find  $\bar{\sigma}_r^2$  for each new chromosome.

(15) Sort the  $2p$  chromosomes by  $\bar{\sigma}_r^2$  increasing. Keep the first  $p$  and discard the rest. Set  $SE = \bar{\sigma}_r^2$ . Set  $R = \{r | m(\sigma^2) = \bar{\sigma}_r^2\}$  and save  $a$ ,  $b$ ,  $c$ , and  $\mu$ 's of  $r$ -th chromosome for all  $r \in R$ . Return to step (9)

(16) Display  $a$ ,  $b$ ,  $c$ , and  $\mu$ 's of  $r$ -th chromosome for all  $r \in R$ .

(17) Stop

#### IV. TEST AND RESULTS

We ran this algorithm by setting  $n = 7$ ;  $l = 30 * 2^7$ ; and  $p = 200$ ; and choosing  $a$ ,  $b$ ,  $c$ , and  $\mu$ 's. We then created random  $f_{ji}$ 's and calculated  $y_j$  the Choquet integral with  $f_{ji}$ 's,  $a$ ,  $b$ ,  $c$ , and  $\mu$ 's. Running the algorithm took 356 minutes and 17 seconds to create 195 generations. The algorithm terminated with  $m(\sigma_y^2) = 23.364$ . Since  $\sigma_y^2 = 777.259$ , this is a 96.99% solution. The results were:

Var.	Chosen Values	Output	Var.	Chosen Values	Output
$a_1$	4.31	2.96	$a_2$	62.8	14.5
$a_3$	4.56	1.71	$a_4$	3.02	0
$a_5$	0.502	0.168	$a_6$	0	0.345
$a_7$	1.81	0.420	$b_1$	-0.970	-0.627
$b_2$	-1	-0.802	$b_3$	0.779	-0.113
$b_4$	0.684	0.287	$b_5$	0.948	0.872
$b_6$	-0.424	-1	$b_7$	-0.394	0.852
$c$	-9.57	-3.32	$\mu_1$	0.361	0.527
$\mu_2$	0.314	0.490	$\mu_3$	0.676	1.02
$\mu_4$	-0.0508	0.0957	$\mu_5$	0.311	0.623
$\mu_6$	0.264	0.586	$\mu_7$	0.625	1.11
$\mu_8$	0.229	0.584	$\mu_9$	0.590	1.11
$\mu_{10}$	0.543	1.08	$\mu_{11}$	0.904	1.60
$\mu_{12}$	0.178	0.680	$\mu_{13}$	0.539	1.21

$\mu_{14}$	0.492	1.67	$\mu_{15}$	0.854	1.70
$\mu_{16}$	0.428	0.459	$\mu_{17}$	0.789	0.986
$\mu_{18}$	0.742	0.949	$\mu_{19}$	1.10	1.48
$\mu_{20}$	0.377	0.555	$\mu_{21}$	0.738	1.08
$\mu_{22}$	0.691	1.04	$\mu_{23}$	1.05	1.57
$\mu_{24}$	0.656	1.04	$\mu_{25}$	1.02	1.57
$\mu_{26}$	0.971	1.53	$\mu_{27}$	1.33	2.06
$\mu_{28}$	0.605	1.14	$\mu_{29}$	0.967	1.67
$\mu_{30}$	0.920	1.63	$\mu_{31}$	1.28	2.16
$\mu_{32}$	0.854	0.363	$\mu_{33}$	1.21	0.891
$\mu_{34}$	1.17	0.854	$\mu_{35}$	1.53	1.38
$\mu_{36}$	0.803	0.459	$\mu_{37}$	1.16	0.986
$\mu_{38}$	1.12	0.949	$\mu_{39}$	1.48	1.48
$\mu_{40}$	1.08	0.947	$\mu_{41}$	1.44	1.47
$\mu_{42}$	1.40	1.44	$\mu_{43}$	1.76	1.96
$\mu_{44}$	1.03	1.04	$\mu_{45}$	1.39	1.57
$\mu_{46}$	1.35	1.53	$\mu_{47}$	1.71	2.06
$\mu_{48}$	1.28	0.822	$\mu_{49}$	1.64	1.35
$\mu_{50}$	1.60	1.31	$\mu_{51}$	1.96	1.84
$\mu_{52}$	1.23	0.918	$\mu_{53}$	1.59	1.45
$\mu_{54}$	1.54	1.41	$\mu_{55}$	1.91	1.94
$\mu_{56}$	1.51	1.41	$\mu_{57}$	1.87	1.93
$\mu_{58}$	1.82	1.90	$\mu_{59}$	2.19	2.42
$\mu_{60}$	1.46	1.50	$\mu_{61}$	1.82	2.03
$\mu_{62}$	1.77	1.99	$\mu_{63}$	2.13	2.52
$\mu_{64}$	0	0	$\mu_{65}$	0.361	0.527
$\mu_{66}$	0.314	0.490	$\mu_{67}$	0.676	1.02
$\mu_{68}$	-0.0508	0.0957	$\mu_{69}$	0.311	0.623
$\mu_{70}$	0.264	0.586	$\mu_{71}$	0.625	1.11
$\mu_{72}$	0.229	0.584	$\mu_{73}$	0.590	1.11
$\mu_{74}$	0.543	1.07	$\mu_{75}$	0.904	1.60
$\mu_{76}$	0.178	0.680	$\mu_{77}$	0.539	1.21
$\mu_{78}$	0.492	1.17	$\mu_{79}$	0.854	1.70
$\mu_{80}$	0.428	0.459	$\mu_{81}$	0.789	0.986
$\mu_{82}$	0.742	0.949	$\mu_{83}$	1.104	1.48

$\mu_{84}$	0.377	0.555	$\mu_{85}$	0.738	1.08
$\mu_{86}$	0.691	1.0	$\mu_{87}$	1.05	1.57
$\mu_{88}$	0.656	1.04	$\mu_{89}$	1.02	1.57
$\mu_{90}$	0.971	1.53	$\mu_{91}$	1.33	2.06
$\mu_{92}$	0.605	1.14	$\mu_{93}$	0.967	1.67
$\mu_{94}$	0.920	1.63	$\mu_{95}$	1.28	2.16
$\mu_{96}$	0.854	0.363	$\mu_{97}$	1.21	0.891
$\mu_{98}$	1.17	0.854	$\mu_{99}$	1.53	1.38
$\mu_{100}$	0.803	0.459	$\mu_{101}$	1.16	0.986
$\mu_{102}$	1.12	0.949	$\mu_{103}$	1.48	1.48
$\mu_{104}$	1.08	0.947	$\mu_{105}$	1.44	1.47
$\mu_{106}$	1.40	1.44	$\mu_{107}$	1.76	1.96
$\mu_{108}$	1.03	1.04	$\mu_{109}$	1.39	1.57
$\mu_{110}$	1.35	1.53	$\mu_{111}$	1.71	2.06
$\mu_{112}$	1.28	0.822	$\mu_{113}$	1.64	1.35
$\mu_{114}$	1.60	1.31	$\mu_{115}$	1.96	1.84
$\mu_{116}$	1.23	0.918	$\mu_{117}$	1.59	1.45
$\mu_{118}$	1.54	1.41	$\mu_{119}$	1.91	1.94
$\mu_{120}$	1.51	1.41	$\mu_{121}$	1.87	1.93
$\mu_{122}$	1.82	1.90	$\mu_{123}$	2.19	2.42
$\mu_{124}$	1.46	1.50	$\mu_{125}$	1.82	2.03
$\mu_{126}$	1.77	1.99	$\mu_{127}$	2.13	2.52

## V. CONCLUSION

Using a 2-additive approach to nonlinear multiregression shows some very promising results; however, in order to create a more complete solution the inclusion of some other methods into some sort of a hybrid may make sense.

## REFERENCES

- [1] Z. Wang and H. Guo, A new genetic algorithm for nonlinear Multiregressions based on generalized Choquet integrals, *Proc. FUZZ—IEEE2003*, 819-821
- [2] Z. Wang and G. J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York, 1992.
- [3] M. Grabisch, The Interaction and Möbius representations 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 Eds.) Springer Verlag, Heidelberg (2000), 70-93