A Fast Parameters Estimation for Nonlinear Multi-regressions Based on Choquet Integral with Quantum-Behaved Particle Swarm Optimization

You-Min Jau[1], Chia-Ju Wu[2], Jin-Tsong Jeng[3]

1. Department of Electrical Engineering, National Yunlin University of Science & Technology, Taiwan
2. Department of Electrical Engineering, National Yunlin University of Science & Technology, Taiwan
3. Department of Computer Science & Information Engineering, National Formosa University, Taiwan

E-mail: g9510809@yuntech.edu.tw WUCJ@yuntech.edu.tw tsong@nfu.edu.tw

Abstract: In general, the inherent interaction among attributes must be considered circumspectly in the study of data mining and information fusion, a nonlinear model with nonlinear multi-regression model based on Choquet integral is suitable to deal with these problems. However, this nonlinear model is an over-determined systems and it is difficult to find the analytic solution. Hence, many researchers had proposed many algorithms; namely, genetic algorithm, neural network, particle swarm optimization, quantum-behaved particle swarm optimization (QPSO), to estimate parameters of this nonlinear model. In this study, an effective QPSO (EQPSO) algorithm which is used to estimate parameters of nonlinear multi-regression model was proposed. That is, the proposed EQPSO algorithm applied the concept of the genetic algorithm to the QPSO algorithm that can improve the convergent speed with premature phenomenon, stagnation and reducing the inference of the creative coefficient β in the QPSO algorithm. From the simulation results, the proposed EQPSO has the parameter estimation more precise and faster convergence speed than the genetic algorithm for the nonlinear multi-regression base on the Choquet integral.

Keywords: Choquet integral, Nonlinear multi-regression**,** QPSO, Mutation, Stagnation.

Ⅰ. Introduction

In data mining and information fusion, the most common aggregation tools are the weighted average method and the linear regression. These methods are all linear and must assumption that there is no interaction among attributes. However, in many practical problems, the inherent interaction among attributes must be considered circumspectly. Hence, a nonlinear multi-regression based on Choquet integral (NMRCI) with respect to non-additive measures has been proposed [1,3,4,11,15]. Liu et al [15] proposed Choquet integral's regression that derived from fuzzy support and deal with interaction among attributes based on the correlation in statistics. Wang et al [3-4] proposed a weighted NMRCI to deal with the data with some categorical attributes. It is more complex to estimate parameters of the nonlinear multi-regression. Hence, the non-weighted nonlinear multi-regression model [1] to deal with numerical attributes with the same dimension is adopted in this paper.

At present, there are some soft computing techniques [2-4] which can be used for determining regression parameters of the NMRCI. In this study, the modify particle swarm optimization with quantum behavior (QPSO) is proposed for the NMRCI. The QPSO [5] ensures the congregation of the particle swarm without losing the randomness. Particles can appear any position of the whole space which is searched in a certain probability. Because of the swift convergence speed of QPSO, when one particle finds a local optima state, the others will quickly converge to it. If particles cannot find any better state, the QPSO will take on the premature phenomenon. In order to depart from the local optima, some improvements have been proposed [6-9]. These improved algorithms try to change the state of the particles by the mutation mechanism and to get out of the local optima. But it is difficult to be realized, when the explicit distances between the best and the local optima state is difficult to determine. In this paper, a QPSO algorithm with elitism of genetic algorithm (GA), named EQPSO is proposed. From the results of experiment show that the proposed algorithm can improve the convergent speed and the global searching ability. This paper is organized as follows. In section Ⅱ, the nonlinear multi-regression model base on Choquet integral is introduced. In section Ⅲ PSO and QPSO theory will be presented. In Section Ⅳ, The EQPSO algorithm will be introduced in detail. Simulation results and comparisons are described in Section Ⅴ and the paper is concluded in Section Ⅵ.

Ⅱ. Nonlinear multi-regression model base on the Choquet integral

Let the data consist of k recorders (or observations) of input $X = \{x_1, x_2, \cdots x_n\}$ and output Y, denoted by $f_{j1}, f_{j2}, \cdots f_{jn}, y_j$ and has a form as:

x_1	x_2	…	x_n	y
f_{11}	f_{12}	…	f_{1n}	y_1
f_{21}	f_{22}	…	f_{2n}	y_2
⋮	⋮		⋮	⋮
f_{k1}	f_{k2}	…	f_{kn}	y_k

where k is the number of observation and should take much large than n at 5 times of 2^n. Let X be a finite set of predictive attributes and Y be the objective attribute. In many real-world problems, the inherent interaction

among predictive attributes must be considered circumspectly and the kind of interaction is essentially different from the correlation in statistics. Hence, the traditional linear multi-regression model will fail in these practical applications. In order to effectively describe the inherent interaction among predictive attributes, Wang and Klir [10,11] proposed a non-additive set function μ (named general measure) defined on the power set of predictive attributes *X*, i.e., $\mu : P(X) \to [0,1]$, which satisfy the conditions $\mu(\phi) = 0$ and $\mu(X) = 1$. When it deals with data mining and information fusion where data and information posses inherent interaction, it must adopt nonlinear integral with respect to the general measure. In general, the nonlinear integral has Sugeno integral, Wang integral, Choquet integral, etc [11]. The Choquet integral is the most frequent form and is defined as follows:

$$\int f d\mu = \int_{-\infty}^{0} [\mu(F_\alpha) - \mu(X)] d\alpha + \int_{0}^{\infty} \mu(F_\alpha) d\alpha , \qquad (1)$$

where *f* is nonnegative function, and F_α is the α-cut set of function *f*. Because of *X* is finite, the Choquet integral can be express as:

$$\int f d\mu = \sum_{i=1}^{n} \left(f(x_i^*) - f(x_{i-1}^*) \right) \cdot \mu(\{x_i^*, x_{i+1}^*, \cdots, x_n^*\}) , \qquad (2)$$

where $\{x_i^*, x_{i+1}^*, \cdots, x_n^*\}$ is a permutation of $\{x_1, x_2, \cdots, x_n\}$ and $f(x_1^*) \le f(x_2^*) \le \cdots \le f(x_n^*)$, $f(x_0^*) = 0$. Thus, *f* defined on *X* and the relation between predictive attributes *X* and objective attribute *Y* can be express as a nonlinear multi-regression model as follows:

$$y = c + q \cdot \int f d\mu + \varepsilon , \qquad (3)$$

where *c*, *q* is a constant, $\int f d\mu$ is the Choquet integral of function *f* with respect to general measure μ, and $\varepsilon \sim N(0, \sigma^2)$ is a normally distributed random perturbation with zero mean and variance σ^2. In Eq. (3), the constant *c*, *q* and μ are called regression parameters. These regression parameters can be determined by minimizing the squared error [11]

$$e^2 = \sum_{j=1}^{k} \left(y_j - c - q \int f_j d\mu \right)^2 . \qquad (4)$$

However, it is an over-determined system and difficult to find an analytic solution of *c*, *q*, and μ. Hence, these parameters, *c* and *q*, can be estimated by the least square method.

Ⅲ. Particle Swarm Optimization (PSO) and Quantum Behavior PSO (QPSO)

From Eq. (3), to finds out a set of parameters satisfying the specific criteria is a heavy burden. Hence, a proper PSO algorithm can help to finish this work and it possesses an excellent ability of global search. In PSO, each particle keeps trajectory of its own position and velocity in the problem space. At the each iteration, the new positions and velocities of the particles are updated by the following two equations:

$$v_i(t+1) = v_i(t) + c_1 \cdot r_1 (p_i^{loc} - \varphi_i) + c_2 \cdot r_2 (p^{gol} - \varphi_i) , \qquad (5)$$

$$\varphi_i(t+1) = \varphi_i(t) + v_i(t+1) , \qquad (6)$$

where $\varphi_i = [\varphi_{i,1}, \varphi_{i,2} \cdots \varphi_{i,n}]$ and $v_i = [v_{i,1}, v_{i,2} \cdots v_{i,n}]$ are the position and the velocity of the i_{th} particle in the n-dimensional search space, respectively. $p_i^{loc} = [p_{i,1}^{loc}, p_{i,2}^{loc} \cdots p_{i,n}^{loc}]$ and $p^{gol} = [p_1^{gol}, p_2^{gol} \cdots p_n^{gol}]$ are the best position of the i_{th} particle and global best position found so far. *i*=1,2,⋯*M*, *M* is the number of particle population, c_1 is called cognitive parameter, c_2 is called social parameter and r_1, r_2 are random numbers between [0,1]. Clerc and Kennedy [12] proved that if the upper bound of $c_j * r_j$ is properly selected, the particle's position φ_i will converge to the center of potential field $pf^{cnt} = [pf_1^{cnt}, pf_2^{cnt} \cdots pf_n^{cnt}]$ (called equilibrium point), where pf_i^{cnt}, $i = 1, 2 \cdots, n$ is the coordinate of pf^{cnt} in i_{th} dimension, and is obtained by:

$$pf^{cnt}(i) = \frac{\left(r_1 \cdot p_i^{loc} + r_2 \cdot p^{gol} \right)}{\left(r_1 + r_2 \right)} , i = 1, 2, \cdots, M . \qquad (7)$$

Each particle moves around and careens toward pf^{cnt} with declining its kinetic energy to zero, like a retuning satellite orbiting the earth. Inspired by analysis of convergence of the classical PSO [5], an individual particle can be seen as moving in a Delta Potential Well which center is equilibrium point pf^{cnt} in search space, and the quantum mode of a particle is depicted by state of energy (or, called wave function) $\Psi(\vec{x}, t)$ which follows the time-dependent Schrödinger equation:

$$i\hbar \partial / \partial t \, \Psi(\vec{x}, t) = -\hbar / 2m \, \nabla^2 \Psi(\vec{x}, t) + V(\vec{x}) \Psi(\vec{x}, t) . \qquad (8)$$

Sun et al [13] had showed the answer of Schrödinger equation for this model with time-independent in one-dimensional space as:

$$\varphi_{i+1} = pf^{cnt}(i) \pm L / 2 \cdot \ln(1/u) , \qquad (9)$$

where u is random number uniformly distribute on [0,1]. *L* is called "Creativity" or "Imagination" of particle. In order to avoid premature of the algorithm, Mainstream Thought Point is employed to evaluate parameter *L*. The Mainstream Thought Point or called Mean Best Position (*mbest*) and *L* are given as:

$$mbest = \left[\sum_{i=1}^{m} \frac{\varphi_{i,1}}{M}, \sum_{i=1}^{m} \frac{\varphi_{i,2}}{M}, \cdots, \sum_{i=1}^{m} \frac{\varphi_{i,n}}{M} \right] , \qquad (10)$$

$$L = 2 \cdot \beta \left| mbest - \varphi_i \right| , \qquad (11)$$

where β is creative coefficient whose value will control the convergence speed of individual particle and the performance of the algorithm.

Ⅳ. An Efficient QPSO Algorithm

Huang et al [14] proposed an algorithm of improved QPSO (called IQPSO). Its basic concept is

sharing the public information and variation by preserving the fore *N* elites at each generation. Though, this algorithm possesses some advances batter than the traditional PSO and QPSO and is more suitable for a complex nonlinear system, preventing premature, and improving the convergence speed, etc. But it still represents stagnating phenomenon in searching the global optimal solution, and is strongly influenced by the creative coefficient β and preserving number *N*. In order to improve this drawback, a modify algorithm EQPSO that combined the QPSO algorithm with the concept of GA to improve the IQPSO algorithm is proposed. In EQPSO, an index of stagnation (named *matu*) to monitor the convergence of particle's state is set. If the best value of p^{gol} is not updated over then ten generations that means the procedure getting into local minimum. In order to depart from this situation as soon as possible, EQPSO adopts the similar policy of elitism. That is, this procedure preserves the p^{gol} and sorts these collected parameters in each iteration. Once *matu*, the index of stagnation, reaches a specified value, the new particle swarm of next generation will be replaced by these collected parameters. The proposed EQPSO algorithm is as follow:

Step 1: Randomly initialize the particle swarm φ_i , $i = 1, 2, \ldots, M$ and evaluate their fitness value.

Step 2: Initialize p_i^{loc} with the best fitness value of i_{th} particle and p^{gol} with the best one of p_i^{loc}.

Step 3: Calculate the Mean Best Position ($mbest$) by Eq. (10).

Step 4: Calculate the center of potential field pf^{cnt} by Eq. (7).

Step 5: Randomly select the Eq. (12) with equal probability to update φ_i .

$$\varphi_i(t+1) = pf^{cnt}(i) \pm \beta \cdot \left\| \left(mbest - \varphi_i(t) \right) \right\| \cdot \ln(1/u), \quad (12)$$

Step 6: Evaluate the fitness value of $\varphi_i(t+1)$.

Step 7: By value calculated in step 6, sort $\varphi_i(t+1)$.

Step 8: Check whether the maximum iteration is reached or the termination criteria is satisfied. If yes, then go to step 14. If no, carry on next step.

Step 9: Check whether p_i^{loc} and p^{gol} should be updated. If p^{gol} is updated, then set matu = 0 and go back to step 3. If p^{gol} is not updated, then increase matu = 1 and go on next step.

Step 10: Check whether matu reaches 10 times *M*. If yes, let $\varphi(t+1) = p^{gol}$ and go back to step 3. If no, keep $\varphi(t+1)$ unchanged and go back to step 3.

Step 11: Check whether p_i^{loc} and p^{gol} should be updated and output results.

Ⅴ. Results and Comparisons

The simulations were conducted in the Matlab environment with Intel Core 2 Duo CPU P8400 and 4GB Ram. It has been successfully run for a number of examples.

Example 1: Let the regression parameters *q*=2.5, *c*=6, the dimension of predictive attribute is 4, $X = \{x_1, x_2, x_3, x_4\}$ and data size is 50 show in Table 1. The EQPSO and the GA algorithm are independently used with stop criteria $er \le 1e^{-3}$ and $er \le 1.5e^{-2}$, respectively. The average results are obtained as shown in Table 2. Besides, we also compare with PSO, QPSO, and IQPSO. In this example, the stop criteria $er \le 1e^{-2}$ usually cannot be achieved in running GA algorithm [1]. Hence, the stop criteria is released as $er \le 1.5e^{-2}$ for running GA algorithm. From Table 2, the proposed algorithm is superior to the GA, PSO, QPSO and IQPSO for the parameters of estimation in the NMRCI.

Ⅵ. Conclusions

The NMRCI can describe the multi-input single-output systems or multi-input multi-output systems and has been widely developed under the non-additive measures. Many researchers had proposed some algorithms such as GA, neural network, PSO and QPSO to estimate parameters of the NMRCI. In the study, we successfully combine the mutation concept of GA into the QPSO algorithm. From the results of simulation, the proposed algorithm is superior to the GA, PSO, QPSO and IQPSO for the parameters of estimation in the NMRCI.

Acknowledgments

This work was supported by the National Science Council of Taiwan, R. O. C. under Grand NSC98-2221-E-224-047.

References

[1] Wang Z, Leung KS, Wong ML, Fang J, Xu K (2000), Nonlinear nonnegative multi-regressions based on Choquet integrals. International Journal of Approximate Reasoning, Vol. 25, No. 2, pp. 71-87.

[2] Wang Z, Wang J (1999), A genetic algorithm for determining non additive set functions in information fusion. Fuzzy Sets System, Vol. 102, No. 3, pp. 463–469.

[3] Xu K, Wang Z, Wong ML, Leung KS (2001), Discover dependency pattern among attributes by using a new type of nonlinear multi-regression. International Journal of Intelligent Systems, Vol. 16, pp. 949-962.

[4] Wang Z, Guo H (2003), A new genetic algorithm for nonlinear multi-regressions based on generalized Choquet integrals. Proc. of Fuzz/IEEE, pp. 819-821.

[5] Sun J, Feng B, Xu WB (2004), Particle swarm optimization with particles having quantum behavior.

IEEE Congress on Evolutionary Computation, USA, pp. 325-331.

[6] Lin X, Feng B, Sun J (2008), Chaos quantum-behaved particle swarm optimization algorithm. Computer Engineering and Design, Vol. 29, No. 10, pp. 2610-2612.

[7] Ge HW, J WH (2007), Application of quantum-behaved PSO algorithm with mutation operator in system parameters identification. Computer Engineering and Applications, Vol. 43, No. 29, pp. 222-224.

[8] Wang Z, Feng B and Sun J (2008), Quantum-behaved particle swarm optimization with dimension mutation operator. Computer Engineering and Design, Vol. 29, No.6, pp. 1478-1481.

[9] Zhong W, Qian F, Xing J (2008), An improved theta-PSO algorithm with crossover and mutation. Proc. of the World Congress on Intelligent Control and Automation, pp. 5308-5312.

[10] Wang Z, Klir GJ (1992), Fuzzy Measure Theory, Plenum Press, New York.

[11] Wang Z, Klir GJ (2008), Generalized Measure Theory, Springer, New York.

[12] Clerc M and Kennedy J (2002), The particle swarm: explosion, stability, and convergence in a multi-dimensional complex space. IEEE Transactions on Evolutionary Computation, Vol. 6, No. 1, pp.68-73.

[13] Sun J, Feng B, Xu WB (2004), A global search strategy of quantum-behaved particle swarm optimization. Proc. of IEEE Conference on Cybernetics and Intelligent Systems, pp. 111-116.

[14] Huang Z, Wang Y, Yang C, Wu C (2009), A new improved quantum-behaved particle swarm optimization model. Proc. of the IEEE Conference on Industrial Electronics and Applications, pp. 1560–1564.

[15] Liu HC, Tu YC, Chen CC, Weng WS (2008), The Choquet integral with respect to λ measure based on γ support. Proce. of the International Conferences on Machine Learning and Cybernetics, Kuming, China.

Table 1: Predictive and object attributes for Example 1.

$f(x_1)$	$f(x_2)$	$f(x_3)$	$f(x_4)$	y
0.5	0.9	0.7	1.1	7.9750
0.5	1.3	0.3	0.2	7.0250
1.3	1.5	1.2	0.3	8.2375
1.7	0.7	0.5	0.3	7.7250
0.7	1.3	0.1	1.3	7.6750
1.5	0.8	1.5	1	9.0125
⋮	⋮	⋮	⋮	⋮
1.4	1	0.8	0.3	7.8750
0.5	0.8	1.5	1.7	9.2250
0.3	1.3	1.6	1	8.8250
1.2	1.5	1.6	1.7	9.8500
0.9	1	0.7	0.4	7.6500

Table 2: The average results of example 1, where MSE is minimum mean square error, iters is iterations.

	μ_{calc} GA	μ_{calc} PSO	μ_{calc} QPSO	μ_{calc} IQPSO	μ_{calc} EQPSO	μ_{org}		y_{calc} GA	y_{calc} PSO	y_{calc} QPSO	y_{calc} IQPSO	y_{calc} EQPSO	y_{org}
$\{x_1\}$	0.197	0.105	0.224	0.170	0.199	0.200	data(1)	7.963	8.030	8.023	8.028	7.975	7.975
$\{x_2\}$	0.103	0.036	0.065	0.083	0.099	0.100	data(2)	7.019	7.116	7.053	7.028	7.024	7.025
$\{x_1,x_2\}$	0.338	0.353	0.394	0.322	0.350	0.350	data(3)	8.274	8.069	8.210	8.263	8.240	8.238
$\{x_3\}$	0.384	0.333	0.263	0.238	0.398	0.400	data(4)	7.716	7.620	7.843	7.672	7.725	7.725
$\{x_1,x_3\}$	0.460	0.395	0.593	0.478	0.449	0.450	data(5)	7.657	7.777	7.650	7.759	7.676	7.675
$\{x_2,x_3\}$	0.498	0.348	0.287	0.509	0.499	0.500	data(6)	9.025	9.030	9.207	9.035	9.012	9.013
$\{x_1,x_2,x_3\}$	0.617	0.474	0.576	0.611	0.601	0.600	data(7)	7.904	8.004	8.021	7.885	7.926	7.925
$\{x_4\}$	0.282	0.177	0.280	0.349	0.298	0.300	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\{x_1,x_4\}$	0.458	0.459	0.419	0.378	0.449	0.450	data(44)	8.363	8.297	8.393	8.381	8.362	8.363
$\{x_2,x_4\}$	0.340	0.323	0.306	0.352	0.350	0.350	data(45)	7.037	6.986	6.979	7.027	7.024	7.025
$\{x_3,x_4\}$	0.602	0.577	0.589	0.634	0.600	0.600	data(46)	7.884	7.799	7.935	7.871	7.876	7.875
$\{x_1,x_2,x_4\}$	0.703	0.746	0.692	0.788	0.699	0.700	data(47)	9.235	9.235	9.248	9.398	9.224	9.225
$\{x_1,x_3,x_4\}$	0.893	0.990	0.951	0.887	0.900	0.900	data(48)	8.833	8.673	8.711	8.735	8.824	8.825
$\{x_2,x_3,x_4\}$	0.811	0.761	0.878	0.812	0.800	0.800	data(49)	9.864	9.764	9.864	9.867	9.851	9.850
X	1.000	1.000	1.000	1.000	1.000	1.000	data(50)	7.653	7.724	7.700	7.671	7.651	7.650

	Stop criteria	q	c	MSE	Elapse time (sec)
GA	0.015	2.51626	5.98430	0.014801	28.062
PSO	3000 iters	2.25588	6.33291	0.0957334	44.013
QPSO	3000 iters	2.39674	6.12287	0.0959836	43.785
IQPSO	3000 iters	2.43673	6.07210	0.0882585	86.308
EQPSO	0.001	2.50015	6.00107	0.0009860	8.847