Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 23 (2011) 155 – 160

Procedia Engineering

www.elsevier.com/locate/procedia

Research on Genetic Algorithm and Data Information based on Combined Framework for Nonlinear Functions Optimization

Zhigang Ji, Zhenyu Li, Zhiqiang Ji

Hebei University of Engineering, Handan，China

Abstract

In recent years, piecewise linear change has become an attractive tools, used for all kinds of complicated nonlinear system. Piecewise linear individual function to provide the platform segmental affine nonlinear system contains a large amount of counter approximate nonlinear function value. Even if section of linearization method widely used the best approximation of the nonlinear function of continuous time a minimum number of piecewise functions did not mention liveried with appropriate literature. This paper presents a method of optimization based on clustering evolution get optimal piecewise linear approximation of a class of nonlinear function. The technology is based on the balance between the approximate precision and simplified, and improves the approximate Linear A minimum number of department. The technology has been successfully applied in some common nonlinear function.

Keywords: Optimization, Clustering, Genetic Algorithm.

1. INTRODUCTION

Including piecewise linear change of the nonlinear function of said by a group of linear function. Linear function is defined in a different intervals fixed exchange, the separation [1]. Although, a lot of work on nonlinear function happened the piecewise linear change, and their application in the analysis of real-time nonlinear dynamic system, and not have been too piecewise linear methods for obtaining the optimal nonlinear function. Section of the linear zed a important factor is a lot of subsection departments should consider nonlinear function, represent certain region of interest (tied up and down the state variable). This is usually based on the balance between the will of decision function approximation and minimizing the number field of piecewise etc. A lot of fields will be able to capture more in within the scope of the nonlinear characteristics and realize with the potential to improve the conservative, but the resulting approximation model will be highly complex. However, a low in the simplified model also can't represent the industry nonlinear dynamics in different operating ranges. On this basis, this paper analyzes

doi:10.1016/j.proeng.2011.11.2482

the optimal balance is divided into non-linear function approaching the United States has been made based on clustering methods used. The use of previous works of piecewise linear cluster technology mainly focused on the conventional K refers to [2] and fuzzy clustering [3]. However, these technology, based on local optimization suffer trapped the shortcomings of the local optimal point. At the same time, the final solution depends heavily on the initial parameters. In this respect, an evolutionary algorithm (genetic algorithm) as the foundation of the method is use of cluster analysis. Genetic algorithm (GA) minimize the risk of converging local optimal value, get the whole search space because at the same time processing. This method is based on the combination of genetic algorithm framework, data clustering and curve fitting, has been successfully applied in the collection of nonlinear function.

2. Piecewise Linear Approximation

Widely used in subsection linear approximation of the nonlinear system. The main reason for its use a wide range of nonlinear system modeling and analysis is to build the effectiveness of linear system theory as the tool, the analysis in every region about linear. A typical approximation of a continuous. For obtaining the piecewise equivalent of $f(x_k)$ in figure 1, the first step involves partitioning the independent variable into a set of N sectors as

$$\left[x_{k1_1} \quad x_{k1_2}\right],\left[x_{k2_1} \quad x_{k2_2}\right],\left[x_{k3_1} \quad x_{k3_2}\right],\ldots,\left[x_{kN_1} \quad x_{kN_2}\right] \tag{1}$$

Over each sector, the continuous single-valued nonlinear function $f(x_k)$ is approximated by the following linear functions as

$$f(x_k)=\alpha(x_k)x_k+\beta(x_k) \tag{2}$$

where $\left[\alpha(x_k)\beta(x_k)\right]=\left[\alpha_{ki}\beta_{ki}\right]$ iff $x_{ki_1}\le x_k\le x_{ki_2}$ for [1,2,…,N]

However, the spacing of the position is not a simple process. For a given number of intervals, not all possible position will lead to the same approximation. Therefore, it is an ideal, to obtain the optimal location of the interval. For a nonlinear system with much nonlinear function, it will lead to a simple expression, and the whole system. Here the best sex refers to the best position, such as nitrogen interval all possible location of the interval, let the best approximate nonlinear function. This paper aims to obtain the optimal location of the hybrid method based on interval curve fitting clustering and genetic algorithm.

3. Genetic Algorithm Based Clustering

Genetic algorithm is a kind of probability search algorithm iteratively transforms, a set of mathematical objects (usually fixed length binary string), each with a certain predefined fitness value (accepted), entered a new population offspring object [4]. This algorithm is based on the theory of Darwin's natural selection principle. The searching space of the parameters of the code string form (called chromosomes). Such a set is called a string population. At first, a random population is created; reflect different points of the search space. About a fitness function each string, represents the level of fitness or degree of acceptance of the string. Based on the principle of state of existence, a few strings (those who choose to have high fitness value) and each is assigned a number of copies to the mating pool. In a certain generation, copy, crossover and mutation operator average increase fitness value line and fitness best string to a certain pre-decided level.

In n dimensions Euclidean space division is the clustering process involving a given n data-points into groups (), based on some of the same different weights. /For a given population, clustering algorithm in each data point distribution of the population $Z=\left\{(x_k,y_k):k=1,2,\ldots,M\right\}$ to a certain class depending on the similarity/dissimilarity measure. The most widely used k-means algorithm clustering

problems to solve the Euclidean distance, through the optimization clustering center and data points. Update clustering center generation in order to minimize the sum total of the Euclidean distance all data points. The major disadvantage of mean algorithm, it provides the optimal K clustering subprime depends on the initial cluster choice. Many peak system, it can converge to the optimal value, is not. It is because of these factors, the approximate methods to solve the problem of the optimization of the potential. In the present work, to overcome the limitation of the traditional algorithm, in for the global optimal solution method, this method is used.

Genetic algorithm is proposed to search ability can be used, looking for the right clustering center, in the feature space, producing a cluster of similarity measure is to optimize the [5]. Use is clustering measure the sum total of the Euclidean distance, points from their respective clustering center. Mathematically, the clustering metric M for the K clusters $C_1, C_2, C_3, \ldots, C_K$ is given by

$$M\left(C_1, C_2, C_3, \ldots, C_K\right) = \sum_{i=1}^{K} \sum_{x_j \in C_j} \left\| x_j - z_j \right\| \tag{3}$$

The task of the GA is to search for the appropriate cluster centers $Z_1, Z_2, \ldots, Z_K$, such that the clustering metric M is minimized. The basic steps of GA described previously, which are also followed in the GA-clustering algorithm, are summarized below.

String representation and initialization

Each string is as vector in the K clustering center representative. Center will be initialized to randomly selected data points within the search space.

Fitness computation

On the basis of the cluster, determine the Euclidean distance between the centers, data points. This is followed by clustering center update the data points in the average of the composition. Then, fitness value calculation is on the basis of the similarity measure (3) software.

Selection

In the selection process, personal string of copy based on sports fitness value and sent to the mating pool. Operation and implementation for use wheel is proportional to the number of copies of each string in the fitness value.

Crossover

In the single point cross implementation in this work, two strings to choose from random mating pool with a certain cross predefined crossover probability a randomly selected point, produce two new string. The new generation has certain features cross string.

Mutation

Depending on whether a randomly generated number is larger than a predefined mutation probability or not each bit in the string obtained after crossover is altered (changing 0 to 1 and 1 to 0).

Termination criterion

Just as in each generation of the members of the fitness function value than ever before, the survival of the fittest. If a small improvement is in some generation and then stop, the entire operations algorithm or the execution to a predetermined the most offspring. Elite strategy, which can better protect the superior string in the present work, the worst string replacement of traditional a particular generation and the best string in past lives. This ensures that the superior clustering center preservation in the initial offspring.

4. Genetic Algorithm Based Clustering for Piecewise Linearization

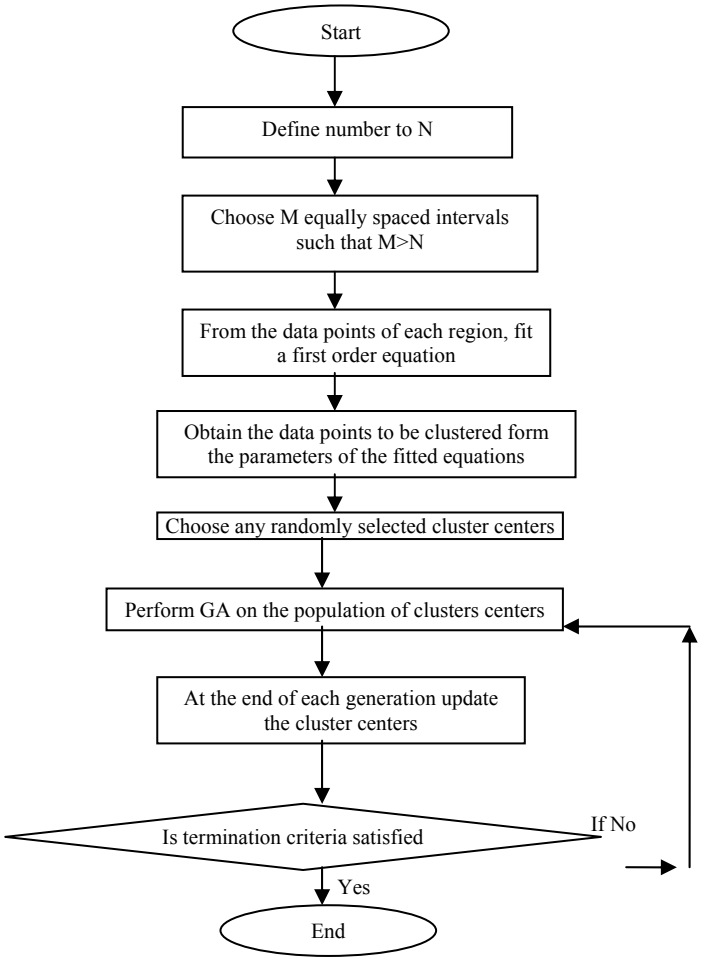

Figure1: obtaining the piecewise linear approximation

In the present problems, clustering decided the interval of specific countries into approximation, different fell on the crowd (intervals) qualitative compare their section approximation parameters, i.e the values of α and β . The idea behind the use of clustering in the present case is that, there are certain intervals of x_k, where there is a very marginal difference in the value of linear parameters α and β. Hence, replacing those intervals by a single value of α and β will lead to the decrease in the number of space, so the complexity of the performance. The reductionism is in the number of because the interval of the cluster will lead to a drop in the proportion of the complexity of the approximate model. As a nonlinear system contains many counter approximate nonlinear function value by a linearised model, reduce the complexity of the model is more and more obvious increase and the degree of nonlinear function and the country. The clustering process is initiated by first converting the range of x_k in a set of M equally spaced regions. The number of M more in number to take the optimal expected and then number of clusters N

(area). And in every area of the interval, the first order equation trefoil, based on the curve fitting region as previously program. Fitting the parameters of the equation is as follows. α and β constitutes the population for the clustering. Hence, the problem at hand is to perform a two dimensional clustering of M data points into N. The flowchart for performing the clustering is shown in Figure 2. In the next section, the proposed approach has been tested for common scaler valued nonlinear functions.

5. Results and Discussion

In this period of genetic algorithm in combination with the clustering method has been applied to get a single block linearised approximate nonlinear function value. The determination of efficiency is proposed based on approximation precision as follows.

Figure 2 shows the approximate function clustering of piecewise linear parameters using genetic algorithm. With the increase of the number of the cluster, improve the approximation observation.

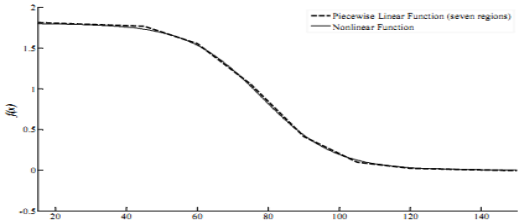

Figure 2: Piecewise linear approximation of monotonically decreasing nonlinear functions.

Clustering is the region to be equal interval, with a small interval field a steep change in non-linear characteristics is seen. But, it is due to gather in large interval department to get in areas around the characteristics of the linear or almost the same. As mentioned previously, define the scope of the interests of the state variables as a group, the spacing of the department first. The GA based clustering is carried out by obtaining the linear parameters (α and β) for each of the sectors by curve fitting.

Based on an outstanding man approach to keep healthy is string with higher consideration. A population had better choose to the string as it in the next generation. The location of the initial selected equally spaced parameters and their corresponding clusters is shown in Figure 3. A further increase in the number of clusters would result in new cluster centers around the data points, which are sparsely located i.e. points with less α and high β.

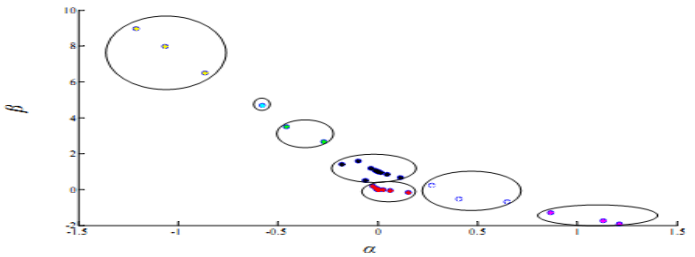

Figure 3: Clusters obtained for equally spaced intervals for function in figure 4.

6. Conclusion

Piecewise linear nonlinear function counter value framework of genetic algorithm and use combination data clustering. This method has a set of common test of nonlinear function. The results show that the

method in the AD hoc way using the linear interval.

References

[1] Storace, M. and Feo, Oscar, D., 'Piecewise-Linear Approximation of Nonlinear Dynamical Systems'. IEEE Transactions On Circuit Systems - 1, Regular Papers, Vol. 51., No. 4, (2004), pp. 830-842.
[2] Ferrari-Trecate, G, Muselli, M., Liberati, D. and Morari, M., 'A Clustering Technique for Identification of Piecewise Affine Systems', Automatica, Vol. 39, (2003), pp. 205 – 217.
[3] Ghosh, S. and Maka, S., 'A Fuzzy Clustering Based Approach for the Piecewise Affine Approximation of a Class of Nonlinear Systems', Communications in Nonlinear Science and Numerical Simulation, Vol. 15 (9), (2010), pp. 2235-2244.
[4] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, MA, 1989.
[5] Maulik, U. and Bandyopadhyay, S., 'Genetic algorithm-based clustering technique'. Pattern Recognition 33 (2000), pp. 1455-1465.