Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2019, 6(3):88-93**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Application of Two-Population Evolutionary Index Model

DING Hua, XIA Hui-ming, ZHU Yi-wen

Nanjing Normal University Taizhou College,Taizhou 225300,China
Email:dingh@nnutc.edu.cn

Abstract In this article, we first introduce the principle, characteristics and process of the evolution strategy algorithm, then we propose a new double-populations evolutionary exponent model. What's more, we analyze the algorithm's objective function, state generation function, evolutionary curve fitting model and ending rule for learning. Finally, we use Matlab to realize this algorithm. Comparing the simulation result of numerical fitting with Genetic Algorithm and Linear Regression, we find that the double-populations evolutionary exponent model can obtain optimal fitting better for the result is more accurate and the convergence speed is faster.

Keywords Double-populations; Evolutionary exponent model; Curve fitting; nonlinear curve model; Evolution strategy

1. Introduction

In the process of solving practical problems and doing scientific experiments, we often need to study the functional relationship between some variables in order to understand the inherent laws and essential attributes of things. The unknown relationship between these variables is usually implicit in a set of discrete data obtained from observation and experiment. The key to solving practical problems is to find out the relatively accurate functional relationship between variables based on observation and experiment data. The method of curve fitting takes the influence of random observation errors on observation data into account. And the overall error that it needs to seek is the smallest, which can better reflect the approximate function relation of observation data.

Evolution is common in nature. However, it is a valuable research issue that we put evolution into computers and use computers to solve problems about evolution. The evolution strategy was first put forward by two German scholars in 1960s.It uses the optimal form of evolutionary theory and the inheritance and variation of genetic information from generation to generation to acquire features that do not exist in previous generations. It evolves according to the theory of survival of the fittest. General evolutionary strategies use real numbers instead of binary coding, which can solve many problems that are composed of real numbers in real life. The intensity of variation plays a decisive role in evolutionary strategies. In the evolutionary strategy, two genetic information can be passed on to offspring, one is the mean which record all locations, the other is the intensity of variation which record the mean. In recent years, the theory of evolutionary strategies has been deeply studied. Many important achievements have been made in the convergence, algorithm construction and application of evolutionary strategies. In this paper, we apply evolutionary strategies algorithm to curve fitting and construct an evolutionary index model. We mainly use the incidence of infectious diseases in document [1] to carry out numerical simulation experiments, make efficiency analysis and comparison. Compared with several well-known algorithms, the simulation result has higher degree of accuracy, smaller error and faster convergence speed. We hope that these analyses will contribute to the design and application of new algorithms.

2. Evolutionary strategy algorithm

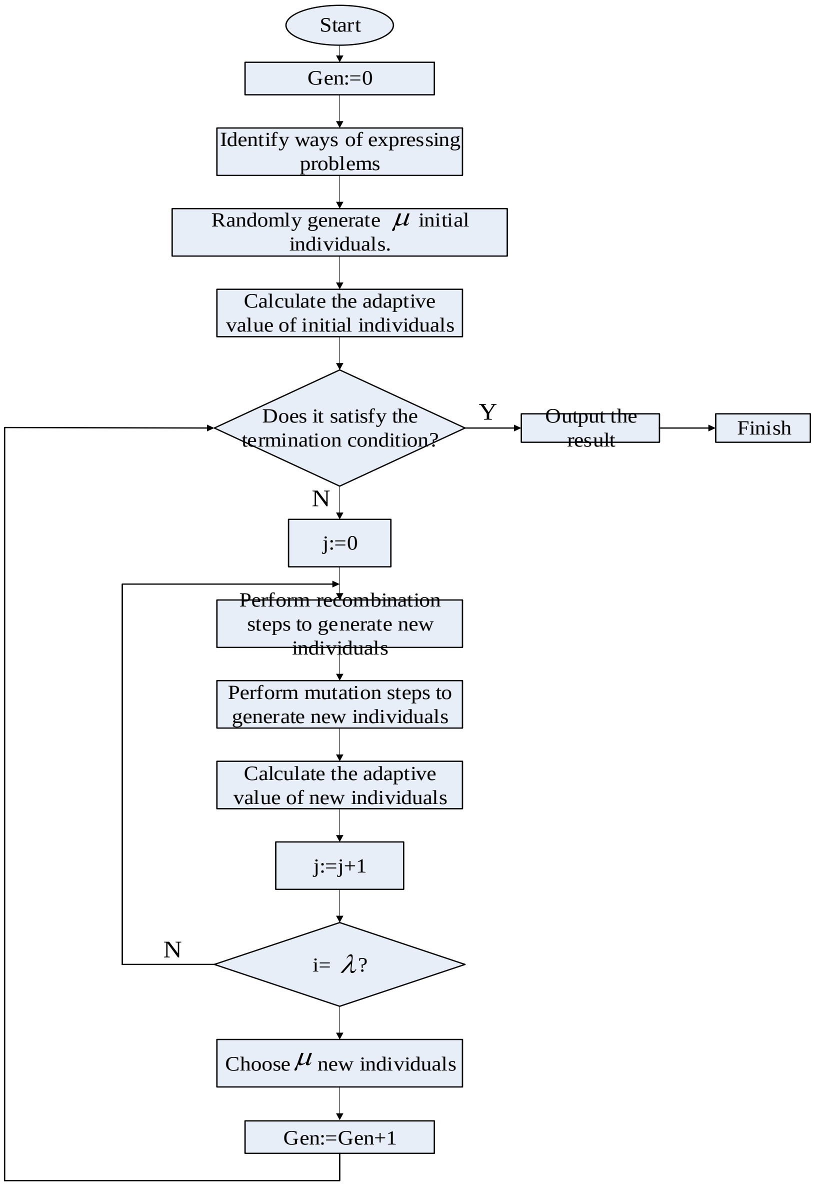

Figure 1: The Workflow of the Evolution Strategy

2.1 The principle of the algorithm

Randomly generate an initial population suitable for the given problem environment, that is, search space. Each individual in the population is a real number encoding. Then we calculate the adaptive value of each individual. According to Darwin's principle of evolution, we choose genetic operators (recombination, mutation, etc.) to iterate and optimize over the population, until an optimal solution or approximate optimal solution is found on a generation.

2.2 Main features of the algorithm

(1) The evolutionary strategy selects μ individuals from λ new individuals or from $(\mu+\lambda)$ individuals to form new population, and the choice is determined.

(2) The evolution strategy directly uses the feasible solution of the problem as the form of individual expression. We don't need to encode the individual and consider the influence of random disturbance on individuals anymore. These make it more convenient for application of the evolution strategy.
(3) The evolutionary strategy takes the optimization problem in the n dimensional real number space as the main processing object.

2.3 Steps of the algorithm
(1) Identify ways of expressing problems;
(2) Generate initial population randomly and calculate the adaptive value;
(3) According to evolution strategy, new population is generated by the following operations:
1) Recombination: exchange target variables and random factors of two parent individuals to generate new individuals;
2) Mutation: add random variables to reorganized individuals to generate new individuals;
3) Calculation: calculate the adaptive value of new individuals;
4) Selection: choose the best individuals to form the next generation on the basis of the selection strategy.
(4)Repeat steps (3) until the end condition is reached. Then select best individuals as the result of the evolutionary strategy.
The workflow of the evolution strategy is shown in Figure 1.The Gen in the figure represents the generation of evolution. In the zeroth generation, according to the way the problem is expressed as binary or triple, μ initial individuals are randomly generated and their adaptive value is calculated. Then, we perform recombination and mutation steps to generate new individuals. The j in the figure is used to count the number of new statistical individuals. λ times it performs recombination and mutation , λ new individuals are produced. Afterwards, we calculate the adaptive value of new individuals and select μ best individuals from λ new individuals or from $(\mu+\lambda)$ individuals to form new population. In this way, we complete the evolution of one generation. Then repeat this evolutionary process until the end condition is satisfied.

3 Objective function and control condition
3.1 Objective function
We assume that y_i is the measured value of point x_i $(i=1,2,\cdots,n)$, $y_i^{'}$ is the value calculated by the fitting function at point x_i .Then the sum of squares of errors on n points is $g=\sum_{i=1}^{n}(y_i-y_i^{'})^2$ .The i here represent the number of data involved in fitting. Obviously, the smaller the value of function, the better.

3.2 State producing function
In this paper, the initial population is randomly generated. Three operators, namely recombination, mutation and selection, are used to train and learn in the process of evolution. The new individuals are constructed according to the formula $\sigma_i^{'}=\sigma_i\cdot\exp(r^{'}\cdot N(0,1)+r\cdot N_i(0,1))$ and $x_i^{'}=x_i+\sigma_i\cdot N_i(0,1)$ or formula $\sigma_i^{'}(j)=\sigma_i^{'}(j)\cdot\exp(\tau_2\cdot N(0,1)+\tau_1\cdot N_j(0,1))$ and $x_i^{'}(j)=x_i(j)+\sigma_i^{'}(j)\cdot\eta_j$ .In the formula, $x_i^{'}$ and x_i represent offspring individuals and parent individuals respectively, while $\sigma_i^{'}$ and σ_i represent standard deviation of offspring and parent respectively. The i here represent the generation of evolutionary.

3.3 The model of fitting curve
We number the year as an independent variable and record it as t. At the same time, we take the total incidence rate as dependent variable and record it as y .The nonlinear curve fitted by the evolutionary index model is in

the form of $y = 10^{a+bt}$. The main purpose of this paper is to obtain the optimal solution of parameter satisfying the termination condition by training and learning parameters a and b.

3.4 End criterion for learning

The end criterion satisfying the given precision is adopted in the cycle process. In the process of each internal loop, we record the current adaptive value as f. When the adaptive value satisfies the given accuracy, the program stops and the optimal value is recorded as $y_i^{'}$.The expression of function of adaptive value is $f = \frac{1}{1+g}, g = \sum_{i=1}^{n}(y_i - y_i^{'})^2$ in it, and the i here is the total number of data involved in fitting.

4. Examples of numerical algorithms

Based on the above model, we can fit the total incidence of infectious diseases in Yangxin County from 1976 to 1989 cited in document [1].

Table 1: Raw data on the incidence of infectious diseases

Time	1976	1977	1978	1979	1980	1981	1982
t	1	2	3	4	5	6	7
Total Incidence	1016.1	84.89	553.4	339.7	268.5	414.7	416.3
Time	1983	1984	1985	1986	1987	1988	1989
t	8	9	10	11	12	13	14
Total Incidence	317.5	242.2	141.2	80.4	115.9	70.3	22.9

4.1 Main flow of the algorithm

Step 1: Randomly generate two populations with initial parameters of a and b, which are defined as population 1 and population 2 respectively.

Step 2: Calculate the adaptive value of two populations. If the termination condition is satisfied, the optimal parameter value will be selected from the two populations. Otherwise, step 3 is executed.

Step 3: Do operations of recombination and mutation on two populations to generate new individuals.

Population 1: Use discrete recombination and Gauss mutation. The specific formula is as follows:

$$\sigma'_i = \sigma_i \cdot \exp(\tau_2 \cdot N(0,1) + \tau_1 \cdot N_i(0,1))$$

$$x'_i = x_i + \sigma_i \cdot N_i(0,1),$$

Here, the value of τ_1 and τ_2 is 1. $N(0,1)$ and $N_i(0,1)$ are the random numbers which obey the standard normal distribution. $N_i(0,1)$ is a random number that meets the standard normal distribution generated by the i component.

Population 2: Use golden section reorganization and Cauchy mutation. The specific formula is as follows:

$$\sigma_i^{'}(j) = \sigma_i^{'}(j) \cdot \exp(\tau_2 \cdot N(0,1) + \tau_1 \cdot N_j(0,1))$$

$$x_i^{'}(j) = x_i(j) + \sigma_i^{'}(j) \cdot \eta_j$$

Here, η and η_j are proportional parameters of a Cauchy random variable at $t = 1$. η_j is used to update each component. The value of τ_1 and τ_2 is 1. $N(0,1)$ and $N_j(0,1)$ are the random numbers which obey the standard normal distribution. $N_j(0,1)$ is a random number that meets the standard normal distribution generated by the j component.

Step 4: Calculate the adaptive value of two populations respectively.

Step 5: For population 1 and population 2, (μ, λ) and $(\mu + \lambda)$ selection strategies are adopted respectively to pick out good individuals.

Step 6: Repeat step 3 to step 5 until it satisfies the end condition. Select the best parameter individual as the result.

4.2 Fitting results

We use evolutionary index model to fit nonlinear curve, its form is $y = 10^{a+bt}$.The selection strategy (μ, λ) is used in the fitting process and parameters take $\mu = 20, \lambda = 7 * \mu = 140$.We composite algorithm routine under $Matlab7.0$ language environment. After calculation, the fitting function is $y = 10^{2.900840375 - 0.069838375t}$ and the sum of squares of errors is 432287.8384.

4.3 Comparison of the algorithm results

The comparison between the results of fitting and results of genetic algorithm and curvilinear regression method in document [1] is shown in Table 2.

Table 2: Comparison of fitting results of each algorithm

Extract parameters	**Evolutionary strategy algorithm**	**Genetic algorithm**	**Linear regression**
a	2.900840375	2.902344	3.118500
b	-0.069838375	-0.070312	-0.100500
Residual	432287.8384	434276.7	671718.5

From the table, we can see that the simulation results of the double-populations evolutionary exponent model are better than those of genetic algorithm and linear regression. This shows the effectiveness of the new model constructed.

5. Conclusion

The process of fitting non-linear curve with evolutionary two-population model proposed in this paper is to seek approximate value y_i'. Then minimize the sum of squared errors between it and raw data y_i. It is a process of dynamically estimating parameter values. Compared with genetic algorithm and linear regression method, the result is more satisfactory. But there are also some shortcomings in this algorithm about means of expression of individual and selection of mutation operators. It is worth studying how to improve the efficiency of the algorithm. We can consider combining it with other algorithms, such as simulated annealing algorithm, neural network, functional network, differential evolution algorithm and so on. This is where we will make further efforts in the future. The double-populations evolutionary exponent model proposed in this paper has strong generality and can also be used in other types of nonlinear curve fitting.

Acknowledgments

This article is supported by The National Natural Science Foundation of China (Grant No. 11601234, 61165015), the Project of Jiangsu Natural Science Foundation (Grant No. BK20160571, BK20150537), the Project of Jiangsu College Natural Science Foundation (Grant No. 15KJB110013).

References

[1]. Cai Yu-dong. (1995). Application of Genetic Algorithm in Medical Nonlinear Curve Fitting. *Journal of Biomedical Engineering*, 159-161.

[2]. Deng Su, Li, Xiao-yi. (2010). Application of Markov Chain in Prediction of spiratory Infectious Diseases. *Chinese Journal of Health Statistics*, 2010(27), 615-616.

[3]. HE Bing, CHE Linxian, LIU Chusheng. (2012). Novel bi-group differential evolutionary programming. *Computer Engineering and Applications*, 2012, 48(26).

[4]. WANG Peichong, QIAN Xu. (2015) Improved teaching learning based optimization with double populations competition. *Computer Engineering and Applications*, 2015, 51(24).

[5]. LI Yang-fan. (2018). Research on Production Line Balance Problem of L Company Based on Double Population Genetic Algorithm. *Value Engineering*, 2018, 37(33).