Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 4, pp. 2260–2265, DOI 10.12694/scpe.v25i4.2961

RESEARCH ON SURVEYING AND MAPPING DATA PROCESSING BASED ON NONLINEAR MATHEMATICAL MODELS AND DEEP LEARNING OPTIMIZATION

HUIXIAN WEI* AND XIANGHUI WEI†

Abstract. In order to deeply understand the mapping data processing of nonlinear mathematical model optimization, this paper uses nonlinear model optimization theory to process mapping data. When the precision of parameter approximation is high, the calculation results of each algorithm are the same, and the iteration times of the fastest descent method and simplex method are significantly increased compared with the other three algorithms. Therefore $x_{01} = 5.42$ in the parameter is kept unchanged near the truth value, and the convergence range of Newton method, the fastest descent method, conjugate gradient method and simplex method away from the truth value of parameter $x_{02} = -0.25$ is investigated. When the approximate value of undetermined parameters is high, the results of Newton algorithm and simplex algorithm are consistent, but the Newton iterative algorithm has faster convergence speed and higher computational efficiency than the simplex algorithm. Lower value when the undetermined parameter approximation precision, namely the undetermined parameter approximation and its true value is far off, may complete failure type Newton iteration algorithm, the simplex algorithm can be a supplement of the Newton method for the most part, the simplex method significantly reduces the requirements for initial value of parameters, data calculation efficiency has improved significantly.

Key words: Nonlinear, Mathematical model optimization, Surveying data processing

1. Introduction. With the development of data processing technology and ability unceasing enhancement, the computer in many industrial control occasions, there is a kind of such variables: they are closely related to the quality of the product, should be strictly controlled, but because of economic or technology (such as online measuring instrument is expensive, or not work) in a poor working environment, are hard to measure online, analytical values can only be obtained through off-line laboratory analysis [1]. However, offline laboratory analysis often has the problem of long time lag, which cannot meet the requirements of online real-time control and optimization operation [2].

In recent years, surveying and mapping techniques have made significant advances in order to solve the estimation and control problems of such variables. The basic point of surveying and mapping technology is to select the secondary variables that are closely related to the primary variables and easy to measure, such as temperature, pressure and flow, etc., according to some optimal criteria. Use computer software to estimate the dominant variables. In the control system that takes the estimated value of surveying sensor as the feedback signal, the controller and estimator are separated, thus bringing great convenience to the design of both the controller and (Soft Sensor) [3]. In addition to "measuring" the leading variables, the soft sensor can also estimate the process parameters of some reaction process characteristics, overcoming the problems of offline analysis lag and large sampling interval. At present, mapping technology has become one of the key research directions in the field of engineering control [4].Linear principal component analysis (LPCA) is mainly aimed at steady-state data with high dimension, noise and collinearity between variables. However, many systems in the real society are often not stable systems, such as a financial decision-making system and many real-time industrial process systems. In addition to the above characteristics, the data generated during the operation of these systems often have time series relationship. At this time, the traditional principal component analysis is used to extract the features of these systems, but the principal component can not reflect the changing characteristics of the data. Aiming at this kind of problem, the theory of dynamic principal component analysis

*Shijiazhuang Institute of Railway Technology, Shijiazhuang, Hebei 050061, China (HuixianWei@126.com)
† Shijiazhuang Institute of Railway Technology, Shijiazhuang, Hebei 050061, China (Corresponding author, XianghuiWei9@163.com)

method is to obtain the augmented matrix by using the sampling data of the past time to the current analysis data matrix. Based on the augmented matrix, the traditional principal component analysis is carried out, and the method to determine the number of the augmented sequence vector is further proposed. Linear principal component analysis is generally used to extract linear relationships of systems or variables, which is a linear dimension reduction technique in essence. Principal component analysis has been successfully applied in many fields, such as chemical industry, finance, biology and so on, and has achieved good research results. But for the actual industrial process, its nature is nonlinear. At this point, there will be some problems if linear principal component analysis is used to extract the features of nonlinear processes.

On the one hand, since linear principal elements only reflect the linear characteristics of the system, in order to reflect as much information as possible, the number of principal elements will increase, thus losing the role of principal component analysis dimension reduction. On the other hand, after the nonlinear system is analyzed by linear principal component analysis method, the residual matrix often contains nonlinear information, rather than invalid information such as noise as the linear principal component thought. Therefore, it is often inappropriate to use linear principal component analysis to analyze nonlinear systems. Nonlinear principal component analysis is an extension of linear principal component analysis, which can more effectively complete the extraction of information with nonlinear relationship [5].

2. Literature review. Nonlinear principal component analysis is a hot and difficult research topic in recent years. In view of this research problem, Mirzaei, F. et al. proposed a mapping model based on DPCA-RBF network for industrial data featuring nonlinear, noisy, collinearity and dynamic strength. Firstly, the dynamic PCA method was used to preprocess the industrial modeling data to obtain the principal components, and then the RBF network model of the principal components and key quality variables was established. This model effectively reduces the number of model variables, removes noise and dynamic information, and reduces the number of parameters of RBF network training, thus improving the accuracy of the model [6].

Su, Y. T. et al. proposed a nonlinear regression model based on master curve for industrial process data with high dimension, data coupling and strong nonlinearity. This model draws on the basic idea of PLS, and takes into account the correlation between independent variables and dependent variables while extracting implicit variable information by using the master curve. In the space of hidden variables, polynomial function is used to fit the nonlinear relation between hidden variables. In the case study, pure function data and real-time operation data of vinyl chloride distillation tower were used to verify the model [7].

Cai, L. et al. made a pioneering study on the application of principal component analysis method in the field of fault diagnosis and monitoring of chemical process, and determined whether there were faults or anomalies by testing the statistics of data. With the development of large-scale database systems, algorithms based on data analysis have been greatly developed [8].

On the basis of the current study, this article for the study of surveying and mapping data processing based on nonlinear mathematical model of optimization, optimization theory and typical optimization algorithm for nonlinear model characteristic analysis, substitution analysis model of nonlinear optimization theory to deal with the data of surveying and mapping, when parameter approximation precision, the algorithm of calculation result is the same.

The number of iteration of the steepest descent method and simplex method increases obviously compared with the other three algorithms. When the parameter approximation value is taken, Newton method, quasi-Newton method, Gauss-Newton method, trust region method and other solutions do not converge. In this case, the norm of parameter error at this point indicates that the iterative algorithm based on Newton algorithm is highly dependent on the parameter approximation value. Therefore, the convergence range of Newton method, fastest descent method, conjugate gradient method and simplex method away from the parameter truth value is investigated by keeping the parameters near the truth value unchanged. When the accurate approximation value of parameters cannot be obtained, the efficiency of surveying and mapping data processing results can be effectively improved [9].

3. Methods.

3.1. Nonlinear model optimization theory and typical optimization algorithm characteristics.

(1) Nonlinear model optimization theory. The unconstrained nonlinear model is optimized, as shown in Equation 3.1.

$$min f_{x\epsilon R^N}(x) \tag{3.1}$$

where $f(x) = f(x_1, x_2, ..., x_n)$ is the n-element nonlinear real-valued function defined in R^n, $x = (x_1, x_2, ..., x_n)^T$.

Nonlinear model optimization iterative algorithm, that is, step size factor α_k is determined through some search method, as shown in Equation 3.2.

$$f(x_k + \alpha_k d_k) < f(x_k) \tag{3.2}$$

That is, move the objective function $f(x)$ in a specified direction until Equation 3.1 is satisfied. Different displacement (different selection of search direction d_k and step factor α_k) results in different iterative algorithms. In order to ensure the convergence of the algorithm, the search direction is required to be the downward direction. The direct optimization algorithm of nonlinear model represented by simplex algorithm only needs to calculate the value of the function without calculating the derivative of the objective function, which is also the most effective method to search for the minimum value. It is suitable for situations where the expression of the objective function is very complex and it is difficult to calculate the derivative [10].

(2) Characteristic analysis of typical optimization algorithms for nonlinear models. The fastest descent method is one of the simplest and oldest methods for solving nonlinear least squares problems. It is the basis of other unconstrained algorithms. It is based on the first approximation of the function and uses the negative gradient direction $d_k = -\nabla f(x)$ as the search direction. Because the steepest descent method in the direction of the adjacent two iterations perpendicular, convergence path appear jagged, thus began a few steps of step length is longer, the approximate solution and the change of the objective function value is bigger, but when close to convergence, step size is small, the approximate solution and the change of the objective function values are also small, thus greatly affect the convergence rate of the steepest descent method.

Newton's method is a classical unconstrained algorithm, which uses the first derivative (gradient) and second derivative (Hesse matrix) at the iterative point X to perform quadratic function approximation to the objective function, and then takes the minimum point of the quadratic function as the new iterative point, and repeats this process until the approximate minimum point satisfying the accuracy requirement is obtained.

Conjugate gradient method is an optimization algorithm based on conjugate direction method. Conjugate direction is in solving n positive definite quadratic objective function is minimum point to produce a set of conjugate direction as gradient direction, under the condition of exact line search algorithm at most iterative step n minimum point can be obtained, because the general objective function approximation of near the minimum point in quadratic function, so you can imagine an algorithm for quadratic function is more effective. It is expected to have a good effect on general functions. The conjugate gradient method is to generate the conjugate direction of Hesse matrix of convex quadratic function $f(x)$ in each iteration step without using the fastest descending direction at the current point. Each conjugate vector depends on the negative gradient at the iteration point, so as to establish the minimum point of not $f(x)$[11].

Trust region method first given the trust region radius as the upper bound of the length of the displacement, and in the current iteration point as the center, to the upper bound for the radius of trust region area, by solving the region's "letter Lai Yu subproblems" (objective function of the quadratic approximation model) of optimal point to determine the displacement of "candidate", if the candidate displacement can make the objective function values are fully drop, The candidate displacement is accepted as the new displacement, and the radius of trust region is maintained and expanded to continue the new iteration. Otherwise, it indicates that the approximation between the quadratic model and the objective function is not ideal, and the radius of the trust region needs to be reduced, and then the new candidate displacement is obtained by solving the sub-problem in the new trust region. The process is repeated until the iteration termination condition is satisfied.

Iterative algorithms need to calculate the derivative of the objective function, the solution directly without calculating the value of the objective function derivative calculation function and when the objective function

Table 4.1: Correlation observations and their truth values

i	1	2	3	4	5
The true value	4.202384	3.258924	2.527006	1.959459	1.519394
Observations	4.20	3.25	2.52	1.95	1.51

Table 4.2: Least-squares data processing results of nonlinear models under different calculation methods

The serial number	Equalization method	computing result	Error norm	iterations
1	steepest descent method	5.421940625 -0.255618484	0.00219887	121
2	Newton method	5.422744582 -0.255672087	0.00291896	6
3	conjugate gradient method	5.422159251 -0.255670512	0.00238897	20
4	Trust domain method	5.422748251 -0.255670512	0.00291233	6
5	simplex method	5.422748251 -0.255674147	0.00292317	64

or analytical expression is complex, calculating derivative difficulty, represented by the simplex algorithm of the simplex method is one of the most effective method of searching the minimum of the objective function of analyticity did not demand, wide applicable [12].

4. Results and analysis.

4.1. Analysis of surveying and mapping data case by nonlinear model optimization theory. It is known that the nonlinear model is $L_i = x_1 e^{ix2}$, and the truth values of x_1 and x_2 are expressed as vectors, that is, the five truth values of $X = (5.420136187 - 0.25436189)^T$, L_i and the corresponding five independent observations with the same accuracy are listed in Table 4.1.

When the approximate value of the parameter is $X_0 = (x_{01}, x_{02})^T = (5.4, -0.3)^T$, the error norm of the parameter is $||X - X_0||_2 = 0.0499$. On this basis, combined with MATLAB software programming technology, the data processing results of the above typical optimization algorithm are obtained, as shown in Table 4.2.

As can be seen from the calculation results in Table 4.2, when the parameter approximation value is of high accuracy, the calculation results of all algorithms are the same, and the number of iterations of the fastest descent method and simplex method increases significantly compared with the other three algorithms [13].

When the parameter approximation value is $X_0 = (x_{01}, x_{02})^T = (3.4, -0.8)^T$, Newton method, quasi-Newton method, Gauss-Newton method, trust region method and other solutions do not converge, and the norm of parameter error at this point is $||X - X_0||_2 = 2.0925$, indicating that the iterative algorithm based on Newton algorithm has a high dependence on parameter approximation value. Existing research shows that the nonlinear model parameters of the optimal value affected by the intensity of nonlinear function of nonlinear circle, in this case, the main factors influencing the size of the model of nonlinear strength index, in order to further parameter approximation accuracy values of different algorithms are dependent, so keep parameters in $x_{01} = 5.42$ near the true value of the same, The convergence range of Newton method, steepest descent method, conjugate gradient method and simplex method beyond the truth value of parameter $x_{02} = -0.25$ is investigated. By changing the parameter value according to $X = (x_{01}, x_{02})^T = (5.42 - 0.25 \pm k \times 0.1)^T k = 0, 1, 2, ..., n$, the convergence boundary of the corresponding algorithm and corresponding calculation results can be obtained, as shown in Table 4.3. The comparison of parameter approximations of corresponding algorithms is shown in Figure 4.1.

The calculation shows that similar data processing results can be obtained by changing the initial values of parameters in other ways. As shown in Table 4.3:

Table 4.3: Convergence boundary and corresponding calculation results under different algorithms

Algorithm	Parameter approximation	Parameter error norm	parametric solution	Parametric solution error norm
	5.42,-0.25	0.0044		
Newton method	5.42,1.45	1.7044	5.4227 -0.2557	0.0029
	5.42,-0.85	0.2956		
Steepest	5.42,-0.25	0.0044	5.4227,-0.2557	0.0029
descent	5.42,22.95	23.2044	5.4223,-0.2556	0.0026
method	5.42,-0.45	0.1956	5.4232,-0.2556	0.0034
Conjugate	5.42,-0.25	0.0044	5.4230,-0.2557	0.0032
gradient	5.42,0.15	75.5044	5.4227,-0.2557	0.0021
method	5.42,-12.95	38.6956	5.4228,-0.2557	0.0041
	5.42,-0.25	0.0044	5.4230,-0.2557	0.0029
Simplex method	5.42,74.25	75.5044	5.4227,-0.2557	0.0029
	5.42,-38.95	38.6956	5.4228,-0.2557	0.0030

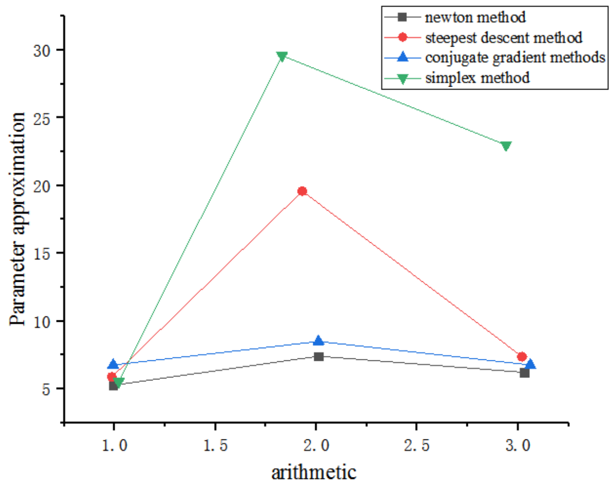

Fig. 4.1: Comparison of parameter approximations of corresponding algorithms

1. When the initial parameter value is within the parameter convergence region of the respective algorithms, Newton algorithm has the highest computational efficiency and the least number of iterations in the calculation process, but Newton algorithm has higher requirements on the accuracy of the initial parameter value, and the convergence region of the initial parameter value is the smallest compared with other types of algorithms [14].
2. Compared with the iterative calculation method, the simplex method has the lowest requirement on the accuracy of initial parameter value, although the calculation efficiency is relatively lower. Compared with the Newton algorithm, the simplex method significantly reduces the requirement on the accuracy of initial parameter value of the iterative algorithm. For general surveying and mapping data processing problems, when high precision initial values of undetermined parameters cannot be obtained, simplex method will undoubtedly be a good supplement to Newtonian algorithm [15].

5. Conclusions and prospects. In this paper, based on the nonlinear mathematical model for optimization of surveying and mapping data processing research, introduces the model of the nonlinear optimization theory and analysis of typical optimization algorithm features, analysis of nonlinear optimization theory model the surveying and mapping data instance when parameter approximation precision, the algorithm calculation result is the same, the steepest descent method and simplex method relative to the other three kinds of iteration algorithm is significantly increased. When the parameter approximation value is $X_0 = (x_{01}, x_{02})^T = (3.4, -0.8)^T$, Newton method, quasi-Newton method, Gauss-Newton method, trust region method and other solutions do not converge, and the norm of parameter error at this point is 001, indicating that the iterative algorithm based on Newton algorithm has a high degree of dependence on parameter approximation value. Therefore, keep $x_{01} = 5.42$ in the parameter unchanged near the truth value, and investigate the convergence range of Newton method, fastest descent method, conjugate gradient method and simplex method away from the truth value of parameter $x_{02} = -0.25$. The efficiency of surveying and mapping data processing can be improved effectively when the accurate approximation value of parameters cannot be obtained.

REFERENCES

[1] Zhongjian, Kang, Yanyan, Feng, Ruiying, & Liu. (2015). Research on fault location method based on nonlinear impedance model. Journal of communications, 10(11), 903-909.

[2] Du, J. (2020). Research on optimization of portrait sculpture data based on 3d image and mobile edge computing. IEEE Access, 8(99), 1-1.

[3] Peng, R. H., Hu, X. Y., & Han, B. (2016). 3d inversion of frequency-domain csem data based on gauss-newton optimization. Chinese Journal of Geophysics, 59(9), 3470-3481.

[4] Chen, C., F Liu, Lin, J., Zhu, K., & Wang, Y. (2016). An optimized air-core coil sensor with a magnetic flux compensation structure suitable to the helicopter tem system. Sensors, 16(4), 508.

[5] Dong, H., Qiang, F., Xu, Z., Quan, Q., & Zhang, R. (2015). Practical rotation angle measurement method by monocular vision. Applied Optics, 54(3), 425.

[6] Mirzaei, F., Delavar, M., Alzoubi, I., & Arrabi, B. N. (2018). Modeling and predict environmental indicators for land leveling using adaptive neuro-fuzzy inference system (anfis), and regression. International Journal of Energy Sector Management, 12(4), 484-506.

[7] Su, Y. T., Hu, S., & Bethel, J. S. (2017). Estimation of cylinder orientation in three-dimensional point cloud using angular distance-based optimization. Optical Engineering, 56(5), 053106.

[8] Cai, L., Li, P., Luo, Q., Zhai, P., & Zhang, Q. (2017). Geometry optimization of a segmented thermoelectric generator based on multi-parameter and nonlinear optimization method. Journal of Electronic Materials, 46(3), 1552-1566.

[9] Ignatkin, Y. M., Makeev, P. V., Shaidakov, V. I., & Shomov, A. I. (2019). Computational research of the main rotor steep descent modes based on the nonlinear blade vortex model. Russian Aeronautics (Iz VUZ), 62(2), 244-253.

[10] Zuo, Y., Gong, M., Zeng, J., Ma, L., & Jiao, L. (2015). Personalized recommendation based on evolutionary multi-objective optimization [research frontier]. IEEE Computational Intelligence Magazine, 10(1), 52-62.

[11] AixiSun, XueJin, & YuboChang. (2017). Research on the process optimization model of micro-clearance electrolysis-assisted laser machining based on bp neural network and ant colony. The International Journal of Advanced Manufacturing Technology, 88(9), 3485–3498.

[12] Guo, R., Chao, N., Jing-Yi, Z., Ping, W., & Zhen-Miao, Z. (2017). Life prediction and test period optimization research based on small sample reliability test of hydraulic pumps. High Technology Letters, 23(1), 63-70.

[13] Zhou, C., Bai, W., Gao, L., & Kang, S. (2016). Optimization research of rohs compliance based on product lifecycle. International Journal of Advanced Manufacturing Technology, 84(9-12), 1-9.

[14] Pelissier, F., Chenini, H., Berry, F., Landrault, A., & Derutin, J. P. (2016). Embedded multi-processor system-on-programmable chip for smart camera pose estimation using nonlinear optimization methods. Journal of Real-Time Image Processing, 12(4), 1-17.

[15] Sun, T., Hu, B., & Yang, Y. (2019). Research on highly nonlinear plateaued functions. IET Information Security, 13(5), 515-518.

Edited by: B Nagaraj M.E.
Special issue on: Deep Learning-Based Advanced Research Trends in Scalable Computing
Received: Dec 2, 2023
Accepted: Jan 23, 2024