

Innovative Application and Optimization Strategy of Algebraic Method in Big Data Analysis

Shasha Liu

College of Science and Technology of China Three Gorges University, Yichang, Hubei, 443002, China

339079811@qq.com

Keywords: Algebraic method; Big data analysis; Optimization strategy

Abstract: This article aims to explore the innovative application of algebraic method in big data analysis and its optimization strategy. This article expounds the background of the era of big data and its influence on social economy, scientific research and other fields, and emphasizes the importance of algebraic method as a mathematical tool in data analysis and its potential innovative value. Based on this, this article deeply studies the application of algebraic method in data preprocessing, data mining and big data visualization, and puts forward a series of optimization strategies for algebraic method, including algorithm optimization, model optimization and system integration and optimization. The research shows that algebraic methods have shown great application potential in big data analysis. Through innovative application and optimization strategy, the practicability and expansibility of algebraic method in big data analysis can be improved. Algebraic method improves the efficiency of data processing and provides us with more accurate and comprehensive data analysis results. Its application prospect is broad, and it will be combined with more cutting-edge technologies in the future to make greater contributions to the development of data science.

1. Introduction

With the rapid development of information technology, we are in an era of data explosion, that is, the era of big data [1]. The arrival of this era has benefited from the remarkable improvement of computer storage capacity, the popularization of Internet technology and the wide application of various sensors and intelligent devices [2]. Big data is large in scale, diverse in types and extremely fast in production. It covers structured data, semi-structured data and unstructured data. For example: texts, images and videos on social media, and sensor data in the Internet of Things[3].

The arrival of the era of big data has had a profound impact on social economy and scientific research [4]. In the economic field, big data provides enterprises with unprecedented market insight, helping enterprises to conduct market research, customer portrait, risk assessment and decision-making more accurately [5]. In scientific research, big data provides massive observation data and experimental data for astronomy, biology, medicine and other fields, which promotes the acceleration and deepening of scientific discovery [6].

Algebraic method is a basic subject of mathematics. It plays an important role in data analysis [7]. Algebraic method can accurately describe the relationship between data by constructing mathematical model, and provide a powerful tool for data processing, analysis and interpretation [8]. In big data analysis, algebraic methods can help us extract useful information from the data, and also reveal the hidden laws and patterns behind the data. Its potential innovation value lies in that through the introduction of algebraic methods, we can develop more efficient and accurate data analysis algorithms and open up new paths for the application of big data.

The purpose of this article is to explore the application of algebraic method in big data analysis, analyze its advantages, challenges and prospects, and put forward corresponding solutions and innovative ideas. The research significance lies in providing new theoretical support and technical means for big data analysis and promoting the development and application of big data analysis technology through in-depth research on the combination of algebraic methods and big data

analysis.

2. Overview of algebraic methods fundamentals

Algebraic method is an important branch of mathematics. It studies the operational relations and structural properties between mathematical objects. The basic concepts of algebraic methods include sets, elements, operations, relations, etc., while the principles involve the properties and theorems of algebraic structures such as groups, rings and fields [9]. According to the different research objects and application fields, algebraic methods can be divided into linear algebra, abstract algebra and other branches, as shown in Table 1.

Table 1 Overview of Algebraic Method Branches

Branch Name	Description	Main Research Objects
Linear Algebra	Studies vector spaces, linear transformations, etc.	Vectors, matrices, linear equations
Abstract Algebra	Studies algebraic structures such as groups, rings, fields, and their properties	Group elements, ring elements, field elements
Universal Algebra	Studies the commonalities and structures of general algebraic systems	Algebraic systems, algebraic homomorphisms
Homological Algebra	Uses homology groups to study the properties of algebraic objects	Modules, homology groups, exact sequences

Linear algebra is an algebraic method to study mathematical objects such as vector space, linear transformation and matrix. It is widely used in data processing, such as linear regression analysis and principal component analysis. Abstract algebra focuses more on the general properties and classification of algebraic structures, such as group theory and ring theory. These theories have important applications in cryptography, coding theory and other fields. Algebraic method has obvious advantages in processing data. First of all, algebraic methods are accurate and can accurately describe the relationships and laws between data. Secondly, algebraic method is systematic, which can establish a complete theoretical system and methodology and provide strong support for data analysis. Algebraic method is also universal and can be applied to different types of data and problem analysis.

There are differences and connections between algebraic methods and other data analysis methods [10]. Statistical methods focus on analyzing data through the principles of probability theory and mathematical statistics, emphasizing the randomness and uncertainty of data. Machine learning method learns rules and patterns from data by constructing models and algorithms, and emphasizes the ability of data prediction and classification. Algebraic method pays more attention to the structure and operational relationship of data, and emphasizes the accurate description and systematic analysis of data. In practical application, these methods are generally integrated and complementary to each other, and together form a complete system of data analysis.

3. Innovative application of algebraic method in big data analysis

3.1. The application of algebraic method in data preprocessing

In big data analysis, data preprocessing is a crucial step. In the face of massive, complex and noisy data, algebraic methods provide us with powerful tools. Through algebraic means such as matrix operation and vector space transformation, we can efficiently clean data, remove or correct abnormal values and errors in data. Algebraic methods can also help us to transform data. Such as standardization and normalization, which enables data from different sources and different scales to be compared and analyzed in the same dimension.

Using the dimension reduction technology of algebraic method, we can greatly reduce the dimension of data while retaining the main information of data, and improve the efficiency and accuracy of subsequent data analysis. Therefore, the application of algebraic method in data preprocessing improves the quality of data processing and lays a solid foundation for subsequent

data mining and visualization.

3.2. Innovative application of algebraic method in data mining

Data mining is the core of big data analysis, which aims to extract valuable information and knowledge from massive data. Algebraic method shows its unique innovation in this field, as shown in Table 2. These innovative applications enrich the technical means of data mining and improve the accuracy and efficiency of data mining.

Table 2 Innovative Applications of Algebraic Methods in Data Mining

Algebraic Method Category	Innovative Application Points	Contributions to Data Mining
Linear Algebra	Principal Component Analysis (PCA), Singular Value Decomposition (SVD)	Dimensionality reduction, feature extraction, improving mining accuracy
Abstract Algebra	Application of Group Theory in Clustering Analysis	Discovering intrinsic structures in data, optimizing clustering results
Matrix Algebra	Matrix Decomposition, Non-negative Matrix Factorization (NMF)	Decomposing complex data, revealing hidden patterns
Algebraic Graph Theory	Application of Graph Theory Algorithms in Social Network Analysis	Analyzing network structures, mining social relationships

3.3. Exploration of algebraic method in big data visualization

Visualization of big data is a process of presenting complex data in an intuitive and easy-to-understand way. It is very important for data understanding and decision-making. Algebraic methods are also widely explored and applied in this field. Through algebraic methods, we can construct a low-dimensional representation of high-dimensional data, such as using multidimensional scaling analysis (MDS) and other technologies to map high-dimensional data into two-dimensional or three-dimensional space, so that the structure and distribution of data can be more clearly seen. At the same time, algebraic method can also help us to design more exquisite visual charts and interactive interfaces, such as optimizing the layout and color coding of charts by matrix operation, and improving the information density and readability of charts.

4. Optimization strategy of algebraic method in big data analysis

(1) Algorithm optimization strategy

In big data analysis, the algorithm optimization of algebraic method is the key to improve the efficiency and accuracy of analysis. Aiming at the computational bottleneck in algebraic algorithm, this article puts forward a variety of optimization strategies. First of all, by parallelizing the algorithm, using multi-core processors and distributed computing resources, large-scale matrix operations, vector operations and other tasks are decomposed into multiple subtasks for parallel execution, which can significantly improve the computing speed. Secondly, we can use sparse matrix technology to optimize the storage and calculation methods and reduce unnecessary calculation and storage expenses according to the characteristics of a large number of zero elements or nearly zero elements in data. In addition, we can also introduce preprocessing technology to speed up the execution of algebraic operations. These algorithm optimization strategies improve the computational efficiency of algebraic methods, and also make algebraic methods better adapt to the real-time analysis requirements in big data environment.

(2) Model optimization strategy

The application of algebraic methods in big data analysis generally depends on the construction of appropriate mathematical models. In order to improve the performance and generalization ability of the model, this article considers adopting a series of model optimization strategies. On the one hand, by fine tuning the model parameters (using grid search, random search and other strategies), the optimal solution is found in the parameter space, so that the model can better fit the data characteristics. On the other hand, this article thinks that regularization techniques (L1

regularization, L2 regularization, etc.) can be introduced to prevent the model from over-fitting and improve the generalization performance of the model. In addition, we can also explore the method of model fusion. For example, multiple algebraic models are combined to make use of their respective advantages to improve the overall analysis effect. These model optimization strategies can enable algebraic methods to build more accurate and stable models in big data analysis and provide strong support for decision-making.

(3) System integration and optimization

In the field of big data analysis, the system integration and optimization of algebraic technology plays a vital role. In this article, algebra technology is closely integrated with big data processing framework and data warehousing system, which ensures the smooth flow of data and efficient analysis. Through the innovation of system architecture, such as introducing distributed storage mechanism and real-time stream processing architecture, the scalability and fault tolerance of the system can be significantly enhanced. These include adjusting system configuration parameters and optimizing data circulation path, aiming at reducing system response time and improving processing speed. These system integration and optimization measures enable algebra technology to be more smoothly embedded in the whole process of big data analysis, which provides strong and reliable technical support for big data analysis.

5. Conclusions

After in-depth research and practice, this article draws the following main conclusions: First, algebraic methods show great application potential in big data analysis. Through the innovative application of data preprocessing, data mining and big data visualization, algebraic method improves the efficiency and accuracy of data analysis, and also reveals the deeper laws and patterns behind the data for us. Secondly, this article puts forward a series of optimization strategies for algebraic methods, including algorithm optimization, model optimization and system integration and optimization. These strategies can effectively improve the practicability and scalability of algebraic methods in big data analysis. To sum up, the research in this article shows that algebraic method is an indispensable tool in the field of big data analysis. Its innovation and optimization strategy provide new ideas for the development of big data analysis.

Although this article has achieved the above research results, there are still some limitations and deficiencies in this study. Due to the complexity and diversity of big data, this study failed to cover all types of data and analysis scenarios. Therefore, the application effect of algebraic method in some specific scenarios needs further verification and optimization. In the future, with the continuous development and innovation of big data technology, algebraic methods will be more closely combined with other data analysis methods to form a more complete data analysis system. We can further explore the application of algebraic methods in more types and more complex data. I believe that in the future research, algebraic methods will play a more important role in the field of big data analysis and make greater contributions to the development of data science.

References

- [1] Chen Zihao, Xu Chen, Qian Weining, et al. Research Progress on Distributed Matrix Computing Systems for Big Data Analytics [J]. *Journal of Software*, 2023, 34(3): 1236-1258.
- [2] Cai Shanglin, Xie Bo. Design and Application Research of an Intelligent Fire Safety Evaluation System Based on Big Data Analytics [J]. *Fire World*, 2024, 10(13): 52-54.
- [3] Yuan Wei. Application and Effectiveness of Big Data Analytics in Network Security Incident Detection [J]. *Computer Application Digest*, 2024, 40(6): 49-51.
- [4] Wu Xiaodan, Wang Bowei. Simulation of Big Data Fuzzy Random Mining Based on Naive Bayes [J]. *Computer Simulation*, 2023, 40(11): 501-505.
- [5] He Zhuangzhuang. Data Dimensionality Reduction Methods for Big Data Analytics in Smart

Meters [J]. Microcomputer Applications, 2024, 40(4): 190-194.

[6] Yun Chenchao, Huang Yichen, Zhao Wei, et al. Research on Data Fusion Methods for Scientific and Technological Achievement Data Based on Big Data Analytics [J]. Microcomputer Applications, 2022, 38(4): 113-116.

[7] Lu Sian, Hou Guoqing. Research on Cloud Computing Resource Prediction Based on Big Data Analytics Technology [J]. Computer Simulation, 2022, 39(10): 502-505+537.

[8] Feng Yiping, Fei Wantang, Wang Zhuoyu, et al. Research on Methods for Implementing Big Data Analytics Based on Data Mining Algorithm Data Models [J]. Electronic Measurement Technology, 2020, 43(3): 5.

[9] Yu Yuezhang, Xia Tianyu, Jing Yinan, et al. An Intelligent Interactive Guide System for Big Data Analytics [J]. Computer Science, 2021, 48(09): 110-117.

[10] Wang Ying, Xiong Lu, Liu Xiao. Efficiency Analysis of Pattern Matching Algorithms Based on Big Data Processing [J]. Modern Electronic Technology, 2021, 044(009): 124-128.