The Theory of Algorithm Convergence for Multi-source Economic Data Fusion

Yuhan Ni

Affiliation: Meishi International School, Chengdu, Sichuan, 610041, China
Email: 3916478306@qq.com

Abstract

This paper focuses on the theory of algorithm convergence in the field of multi-source economic data fusion. In today's digital economy era, multi-source economic data is widely present and of great value. The convergence of data fusion algorithms is crucial for ensuring the accuracy and stability of the fusion results. Starting from the characteristics of multi-source economic data, this paper analyzes the key significance of algorithm convergence in data fusion, explores the factors influencing algorithm convergence, expounds the basic framework of algorithm convergence theory, and looks forward to future research directions, aiming to provide theoretical support for the design and application of multi-source economic data fusion algorithms.

Keywords: Multi-source economic data; Data fusion algorithm; Convergence theory.

1. Introduction

Driven by the digital wave, the economic sector has generated a vast amount of data, which comes from a wide range of sources, including government statistics departments, financial institutions, enterprise operation systems, and Internet platforms. Multi-source economic data contains rich information and can comprehensively reflect all aspects of economic operation [1], such as market dynamics, consumer behavior, and industrial development trends.

However, data heterogeneity leads to three core challenges: First, format conflicts (such as the mixed use of CSV and JSON) increase the cost of data cleaning by more than 40% [2]; Secondly, semantic ambiguity (such as the differences in the definition of "inflation" under different statistical standards) leads to deviations in the fusion results; Thirdly, the dynamic time-varying characteristics (such as stock prices being updated every second) require the algorithm to have real-time processing capabilities [3]. Against this backdrop, algorithm convergence has become a key indicator determining the reliability of the fusion system - if the iterative process fails to reach a stable solution within a limited time, it will lead to serious consequences such as lagging monetary policy adjustments or misjudgment of market risks [4]. Therefore, constructing a convergence theoretical framework that ADAPTS to the characteristics of economic data has significant theoretical value and practical significance.

Algorithm convergence is one of the core issues of multi-source economic data fusion algorithms [5]. Convergence refers to the fact that during an algorithm's iterative process, as the number of iterations increases, it can gradually approach a stable state or an optimal solution. In data fusion, convergence directly affects the accuracy and stability of the fusion results [6]. If the algorithm does not converge or converges too slowly, it may lead to inaccurate fusion results, failing to reflect changes in economic phenomena in a timely manner, thereby affecting the scientificity and effectiveness of economic decisions. Therefore, in-depth research on the algorithm convergence theory of multi-source economic data fusion has significant theoretical and practical significance.

2. Characteristics of Multi-source Economic Data and Their Impact on Algorithm Convergence

2.1 Characteristics of Multi-source Economic Data

Multi-source economic data has a wide range of sources and covers different types and formats of data [7]. For instance, the data released by government statistics departments are usually structured tabular data, including macroeconomic indicators, industry statistics, etc. Economic data on Internet platforms, on the other hand, may be unstructured text data, such as news reports and user comments. In addition, there are semi-structured data, such as financial data in XML format. This

diversity poses challenges to the design of data fusion algorithms, requiring the algorithms to be capable of handling different types of data and effectively integrating them.

Economic data is often influenced by multiple factors and has considerable uncertainty. For instance, changes in market demand, adjustments in policies and regulations, and unexpected events such as natural disasters can all lead to fluctuations in economic data. In addition, errors may also exist during the data collection process, such as measurement errors and statistical errors. Uncertainty makes it difficult to accurately determine the true value of economic data, increasing the difficulty of data fusion. When dealing with uncertain data, algorithms need to have a certain degree of robustness, be able to resist the uncertainty of the data to a certain extent, and ensure the reliability of the fusion results.

Economic phenomena are dynamic and changing, and economic data also have strong timeliness. Timely acquisition and analysis of economic data are crucial for grasping the economic situation and making correct economic decisions. Therefore, the multi-source economic data fusion algorithm needs to have high efficiency, be able to complete the data fusion processing in a relatively short time, and output the fusion results in a timely manner. If the convergence speed of the algorithm is too slow, it may lead to a lag in the fusion result and fail to meet the requirements of practical applications.

2.2 The Influence of the Characteristics of Multi-source Economic Data on Algorithm Convergence

2.2.1 The Impact of Diversity on Convergence

The diversity of multi-source economic data makes it necessary for algorithms to consider the characteristics and fusion rules of different types of data when processing them. Different types of data may require different fusion methods. For instance, structured data can be fused using methods based on mathematical models, while unstructured data needs to be processed with technologies such as natural language processing for feature extraction and fusion. This diversity increases the complexity of the algorithm and may lead to the algorithm being difficult to converge quickly during the iterative process. In addition, semantic differences among different types of data may also affect the convergence of the algorithm. It is necessary for the algorithm to be able to effectively solve semantic conflict problems and achieve accurate data fusion.

2.2.2 The Impact of Uncertainty on Convergence

The uncertainty of economic data makes it difficult for algorithms to determine the optimal fusion result during the iterative process. Because the true value of the data is uncertain, the algorithm may only be able to obtain one approximate solution in each iteration, and there is a certain error between this approximate solution and the true value. As the number of iterations increases, errors may gradually accumulate, causing the algorithm to fail to converge to the true solution. Therefore, the algorithm needs to have a certain degree of fault tolerance, and be able to gradually reduce errors and improve the accuracy of the fusion results through reasonable strategies when there is uncertainty in the data.

2.2.3 The Impact of Timeliness on Convergence

The timeliness requirement demands that the algorithm have a relatively fast convergence speed. In practical applications, economic decisions often need to be made within a relatively short period of time, so data fusion algorithms must complete iterations and output results within a limited time. If the convergence speed of the algorithm is too slow, even if an accurate fusion result can be obtained eventually, it may lose its practical application value due to time delay. Therefore, when designing multi-source economic data fusion algorithms, it is necessary to fully consider the convergence speed of the algorithm and adopt effective acceleration strategies, such as parallel computing and adaptive adjustment of iterative parameters, to improve the efficiency of the algorithm.

3. The Basic Framework of the Convergence Theory of Multi-source Economic Data Fusion Algorithms

3.1 Definition and Measurement Indicators of Algorithm Convergence

3.1.1 Definition of Convergence

In the fusion of multi-source economic data, algorithm convergence refers to the fact that as the number of iterations increases, the output fusion result approaches a stable state or the optimal solution. This stable state or optimal solution is either a pre-set target value or an optimal estimate under certain criteria. For example, in data fusion based on the least square method, its convergence is manifested as the error between the fusion result and the true value gradually decreases to approach zero as the number of iterations increases.

3.1.2 Measurement Indicators

The main indicators for measuring the convergence of an algorithm include convergence speed, convergence accuracy and convergence domain. Convergence speed refers to the number of iterations or the time required for an algorithm to reach a convergent state. The faster the speed, the higher the efficiency. Convergence accuracy refers to the degree of closeness between the fusion result and the true value when the algorithm converges. The higher the accuracy, the more accurate the result. The convergence domain is the initial range of values within which an algorithm can converge. The larger the range, the wider its applicability.

3.2 Analysis of Factors Affecting Algorithm Convergence

3.2.1 Structure and Parameters of the Algorithm Itself

The structure and parameters of the algorithm have a significant impact on convergence. The convergence characteristics of different algorithm structures vary. Some have global convergence and can converge to the global optimal solution under any initial value. Some only have local convergence and converge only when the initial value approaches the optimal solution. The setting of algorithm parameters can also affect the convergence speed and accuracy. For instance, in the gradient descent algorithm, a learning rate that is too high may cause the algorithm to diverect, while one that is too low will result in a slow convergence speed.

3.2.2 Quality and Characteristics of Data

Multi-source economic data has characteristics such as diversity, uncertainty and timeliness, which directly affect the convergence of algorithms. The higher the data quality, the better the accuracy and completeness, the less interference the algorithm is subject to during fusion, and the better the convergence. Conversely, if the data contains a large amount of noise, missing values or outliers, the algorithm needs more iterations to eliminate the interference, which affects the convergence speed and accuracy. In addition, data dimensions, distribution characteristics, etc. will also affect the performance of the algorithm.

3.2.3 Selection of Initial Values

For some iterative algorithms, the selection of initial values has a significant impact on convergence. An appropriate initial value can enable the algorithm to converge to the optimal solution more quickly, while a poor initial value may lead to slow convergence or even divergence. In practical applications, due to the complexity and uncertainty of economic data, it is difficult to determine the optimal initial value. Therefore, strategies such as random initialization and initialization based on prior knowledge need to be adopted to select appropriate initial values.

3.3 Main Contents of Algorithm Convergence Theory

3.3.1 Method for Proving Convergence

Proving the convergence of algorithms is one of the core contents. Common methods include Lyapunov stability theory, monotonic boundedness principle, and compression mapping principle, etc. The Lyapunov stability theory proves convergence by constructing Lyapunov functions and judging the stability of algorithms based on the properties of the functions. The monotonically bounded principle utilizes the monotonicity and boundedness of the iterative sequence of an algorithm to prove convergence. The principle of compressed mapping is applicable to algorithms that satisfy the conditions of compressed mapping, and the convergence is proved by demonstrating that the algorithm is a compressed mapping.

3.3.2 Convergence Rate Analysis Method

Analyzing the convergence speed of an algorithm helps understand its efficiency and provides a basis for optimization. Common methods include the Big O notation and asymptotic analysis, etc. The Big O notation describes the time complexity or space complexity of an algorithm and evaluates the convergence speed by analyzing the running time or space occupation under different input scales. Asymptotic analysis studies the convergence behavior of algorithms when the number of iterations tends to infinity, and evaluates the convergence speed by analyzing indicators such as the convergence order of the iteration sequence.

3.3.3 Convergence Improvement Strategies

To enhance the convergence of the algorithm, various improvement strategies can be adopted. For example, adaptively adjust parameters, dynamically adjust parameters according to the iterative process of the algorithm, and improve the convergence speed and accuracy; By adopting parallel computing, the iterative process is decomposed into multiple sub-tasks and executed in parallel on multiple processors to shorten the running time. A hybrid algorithm is designed by combining the advantages of multiple algorithms to improve the performance and convergence of the algorithm.

4. Application and Challenges of Convergence Theory for Multi-Source Economic Data Fusion Algorithms

4.1 Application Fields

4.1.1 Macroeconomic Forecast

Macroeconomic forecasting is one of the important tasks in the economic field. Accurate macroeconomic forecasting can provide an important basis for the government to formulate economic policies and for enterprises to make strategic decisions. The multi-source economic data fusion algorithm can integrate economic data from multiple channels such as government statistics departments, financial institutions, and enterprises, extract valuable information through fusion analysis, and improve the accuracy of macroeconomic forecasting. The convergence performance of the algorithm can ensure the stability and reliability of the prediction results, enabling them to promptly reflect changes in the economic situation.

4.1.2 Financial Market Analysis

The financial market is a complex and volatile system, influenced by multiple factors. The multi-source economic data fusion algorithm can integrate multi-source data such as stock market data, bond market data, and macroeconomic data to analyze the operational rules and trends of the financial market. The convergence performance of the algorithm is sufficient to ensure the accuracy of the analysis results and provide investors with scientific investment decision-making suggestions. For instance, by integrating and analyzing data such as the price trends and trading volumes of different stocks, the future trends of stocks can be predicted, helping investors formulate reasonable investment strategies.

4.1.3 Enterprise Decision Support

During the operation process, enterprises need to handle a large amount of economic data, such as sales data, cost data, market data, etc. The multi-source economic data fusion algorithm can integrate multi-source data both inside and outside an enterprise, providing comprehensive market information and competitive situation analysis for the enterprise. The convergence performance of the algorithm can ensure the timeliness and accuracy of the analysis results, helping enterprises formulate scientific development strategies and marketing strategies. For instance, by integrating and analyzing market demand data, competitor data and the company's own data, enterprises can understand the changing trends of market demand, adjust their product structure and marketing strategies, and enhance their market competitiveness.

4.2 Challenges Faced

4.2.1 Data Security and Privacy Protection

In the process of integrating multi-source economic data, a large amount of sensitive economic data is involved, such as the financial data of enterprises and the consumption data of individuals. The security and privacy protection of these data are of vital importance. However, current data fusion algorithms often focus on the fusion processing and analysis of data, with relatively less consideration given to data security and privacy protection. How to effectively protect the security and privacy of data while ensuring the effect of data fusion is one of the important challenges faced by the application of the convergence theory of multi-source economic data fusion algorithms.

4.2.2 Complexity and Scalability of the Algorithm

With the continuous growth and multi-source diversification of economic data, the scale of data that data fusion algorithms need to process is getting larger and larger. The complexity and scalability of the algorithm have become the key factors influencing its application. If the complexity of the algorithm is too high, it may lead to an excessively long computing time, which cannot meet the requirements of practical applications. In addition, the scalability of the algorithm is also of vital importance, as it needs to be able to adapt to data fusion tasks of different scales and types. Therefore, how to design a multi-source economic data fusion algorithm with low complexity and high scalability is one of the important challenges currently faced.

4.2.3 Challenges of Cross-domain Data Fusion

Multi-source economic data often involve multiple fields, such as economics, statistics, computer science, etc. Data from different fields have distinct characteristics and patterns. Cross-domain data fusion requires a comprehensive consideration of knowledge and methods from multiple fields. However, at present, there are still certain obstacles in communication and cooperation among different fields, and there is a lack of unified data fusion standards and norms. How to achieve effective integration of cross-domain data and fully leverage the value of multi-source economic data is another challenge faced by the application of the convergence theory of multi-source economic data fusion algorithms.

5. Conclusion

This paper conducts an in-depth study on the algorithm convergence theory of multi-source economic data fusion. Firstly, the characteristics of multi-source economic data and their influence on the convergence of the algorithm were analyzed. It was pointed out that diversity, uncertainty and timeliness are the main characteristics of multi-source economic data, which bring challenges to the convergence of the algorithm. Then the basic framework of the convergence theory of multi-source economic data fusion algorithms was expounded, including the definition and measurement indicators of algorithm convergence, the analysis of factors affecting algorithm convergence, and the main contents of the algorithm convergence theory. Then, the application fields and challenges faced by the convergence theory of multi-source economic data fusion algorithms were discussed. The application fields include macroeconomic forecasting, financial market analysis, and enterprise decision support, etc. The main challenges faced include data security and privacy protection, the complexity and scalability of algorithms, and the difficulties of cross-domain data fusion, etc.

In the future, theoretical research on the convergence of algorithms for multi-source economic data fusion can be carried out from the following aspects: First, strengthen the research on data security and privacy protection technologies, develop more secure and reliable data fusion algorithms, and ensure the security and privacy of multi-source economic data during the fusion process; Second, further optimize the structure and parameters of the algorithm, reduce its complexity, and enhance its scalability to meet the integration requirements of large-scale multi-source economic data. Third, enhance cross-disciplinary cooperation and exchanges, establish unified data integration standards and norms, promote the integration of knowledge from different fields, and improve the effectiveness of cross-disciplinary data integration. Fourth, by integrating emerging technologies such as artificial intelligence and big data, we should explore more intelligent and efficient multi-source economic data fusion algorithms to provide stronger support for the development of the economic sector.

REFERENCES

[1] Li, M., Wang, F., Jia, X., Li, W., Li, T., & Rui, G. (2021). Multi-source data fusion for economic data analysis. Neural Computing and Applications, 33(10), 4729-4739.
[2] Sherman, R. (2014). Business intelligence guidebook: From data integration to analytics. Newnes.
[3] Zhang, L., Xie, Y., Xidao, L., & Zhang, X. (2018, May). Multi-source heterogeneous data fusion. In 2018 International conference on artificial intelligence and big data (ICAIBD) (pp. 47-51). IEEE.
[4] Bernanke, B. S. (2011). Monetary policy report to the congress. DIANE Publishing.
[5] Knight, P. A. (2008). The Sinkhorn−Knopp algorithm: convergence and applications. SIAM Journal on Matrix Analysis and Applications, 30(1), 261-275.
[6] Arora, S., & Singh, S. (2013). The firefly optimization algorithm: convergence analysis and parameter selection. International Journal of Computer Applications, 69(3).
[7] De Waal, T., van Delden, A., & Scholtus, S. (2020). Multi‐source statistics: basic situations and methods. International Statistical Review, 88(1), 203-228.