Application research based on matrix eigenvalues and eigenvectors

Fangfei Zhu

The International Division of No. 2 High School of East China Normal University, Shanghai, China

18221512187@163.com

Abstract. Current common challenges such as high-dimensional data processing and steady-state analysis of complex systems have become increasingly prominent. Eigenvalues and eigenvectors, leveraging their unique mathematical properties, play an irreplaceable role in fields such as data mining and system modeling, serving as a crucial bridge connecting theoretical mathematics with practical applications. Through literature review, this study investigates the application of matrix eigenvalues and eigenvectors in Principal Component Analysis (PCA) and Markov chain steady-state analysis. The results demonstrate that matrix eigenvalues and eigenvectors exhibit significant universality and effectiveness in cross-domain applications. Validated in scenarios including PCA and Markov chain steady-state analysis, they help address key issues including high-dimensional data dimensionality reduction, system steady-state prediction, and information prioritization, thereby providing mathematical support for technological optimization. Simultaneously, they can reveal intrinsic system patterns, reflecting a deep analytical capability for system structures. Future research may focus on optimizing algorithms for solving sparse matrix eigenvalues and exploring integration with deep learning and graph neural networks to expand their application boundaries in large-scale complex systems.

Keywords: Principal Component Analysis, Markov chain, data dimensionality reduction, matrix eigenvalues, eigenvectors

1. Introduction

Eigenvalues and eigenvectors, as core tools in linear algebra, reveal the intrinsic structure of matrices and demonstrate universal value across fields such as data science, dynamical systems analysis, and network science. The theory of eigenvalues and eigenvectors provides a fundamental framework for understanding linear transformations, with its significance extending beyond theoretical construction into a wide range of practical applications [1]. For instance, in Principal Component Analysis (PCA), eigenvalue decomposition is employed for high-dimensional data dimensionality reduction and feature extraction, a method that has become a standard tool in statistics and machine learning [2]. Conversely, eigenvalues are utilized in Markov chains to analyze steady-state behavior, offering crucial theoretical support for stochastic process modeling [3]. These applications underscore the broad interdisciplinary impact of eigenvalue based methods.

This study examines the theoretical properties, application scenarios, advantages, and limitations of matrix eigenvalues and eigenvectors. Theoretically, it clarifies core properties, the application conditions of the Spectral Theorem, and compares the suitability of manual versus numerical solution methods. Practically, it focuses on data processing, system analysis, and information retrieval, analyzing their roles in PCA and Markov chain steady-state analysis. This research holds dual theoretical and practical value: theoretically, it contributes to refining the applied theoretical framework of eigenvalues and eigenvectors, illustrating the transition from theoretical principles to practical applications, thereby providing a reference for subsequent research that balances theoretical depth with practical orientation, while also laying groundwork for future theoretical expansion. Practically, the focused scenarios of PCA and Markov chain steady-state analysis address core demands in interdisciplinary research; the findings can respectively aid data processing in data science, quantitative modeling in economic analysis, and technical optimization in information retrieval, offering clear practical guidance [4].

Regarding the research background, this study adopts a combined methodology of systematic literature review and case analysis. By systematically reviewing the theoretical foundations of eigenvalues and eigenvectors, such as the Spectral Theorem and matrix diagonalization conditions [5], and analyzing typical application cases, it aims to reveal their practical utility across different domains. The selection of PCA and Markov chains as research subjects is primarily based on two reasons: firstly, PCA is a representative application of eigenvalue decomposition in dimensionality reduction, possessing a broad practical foundation and theoretical depth [2]; secondly, the steady-state analysis of Markov chains relies heavily on eigenvalue theory, representing a classical problem in probability modeling and system convergence research [3]. Together, they exemplify the central role of eigenvalue methods in statistics and probability theory.

2. Matrix decomposition

2.1. Key properties

(1) Eigenspace:

All eigenvectors corresponding to the same eigenvalue λ, together with the zero vector, form a linear subspace referred to as the eigenspace of eigenvalue λ, denoted as

$$E_\lambda = \{v | A_v = \lambda_v\} \tag{1}$$

(2) Characteristic Polynomial:

The equation $A_v = \lambda_v$ is equivalent to $(A - \lambda I)v = 0$ (where I is the identity matrix). This homogeneous system of linear equations has a non-zero solution v if and only if the determinant of the coefficient matrix $(A - \lambda I)$ is zero, i.e.:

$$det(A - \lambda I)v = 0 \tag{2}$$

This is an nth-degree polynomial equation in λ, known as the characteristic polynomial of matrix A, whose roots are the eigenvalues of A.

(3) Trace and Determinant:

The sum of all eigenvalues equals the trace of matrix A (i.e., the sum of the main diagonal elements, $tr(A)$), and the product of all eigenvalues equals the determinant of matrix A, $det(A)$.

$$\sum_{i=1}^{n} \lambda_i = tr(A) and \prod_{i=1}^{n} \lambda_i = det(A) \tag{3}$$

(4) Similarity Diagonalization:

If an $n \times n$ square matrix A has n linearly independent eigenvectors, then it can be diagonalized. That is, there exists an invertible matrix P (composed of the linearly independent eigenvectors of A) and a diagonal matrix Λ (whose diagonal elements are the eigenvalues of A), such that:

$$A = P\Lambda P^{-1} \tag{4}$$

This process is called similarity diagonalization, which can greatly simplify matrix power operations (such as $A^K = P\Lambda^k P^{-1}$) and function operations (such as $e^A = Pe^\Lambda P^{-1}$).

2.2. Computation methods

The computation of eigenvalues and eigenvectors can be divided into two categories: exact computation and numerical computation.

Exact Computation (suitable for low-dimensional matrices):

Step 1: Solve for Eigenvalues.

Compute the characteristic polynomial $det(A - \lambda I) = 0$, and find all roots of this polynomial. These roots are all the eigenvalues λ_i.

Step 2: Solve for Eigenvectors.

Substitute each obtained eigenvalue λ_i, into the equation $(A - \lambda_i I)v = 0$, and solve for the fundamental solution set of this homogeneous linear system. The non-zero solutions are the eigenvectors corresponding to λ_i.

In practical applications, iterative numerical methods are predominantly used. A comparison of the power method and the QR algorithm is shown in Table 1.

Table 1. Comparison of power method and QR algorithm

Property/Dimension	Power Method	QR Algorithm
What it computes	One dominant eigenvalue (the largest one)	All eigenvalues
Speed	Slow (linear convergence)	Fast (quadratic/cubic convergence)
Computational Cost	Small	Large
Advantages	Simple, memory-efficient, suitable for very large sparse matrices	Comprehensive functionality, fast, numerically stable
Disadvantages	Limited functionality, slow convergence	High computational cost, complex implementation
Applicable Scenarios	Big data scenarios like PageRank, where only the key eigenvalue is needed	Small to medium-sized matrices, requiring all eigenvalues

2.3. Related theorems

Spectral Theorem: This is a fundamental theorem in the theory of symmetric matrices. It states that any real symmetric matrix can be orthogonally diagonalized. That is, there exists an orthogonal matrix $Q(Q^T = Q^{-1})$ and a diagonal matrix Λ such that $A = Q\Lambda Q^T$.

Perron-Frobenius Theorem: This theorem applies to matrices with non-negative elements (non-negative matrices), particularly to positive matrices or irreducible matrices. It states that such matrices possess a unique, largest-in-magnitude positive real eigenvalue (called the Perron root), and the components of its corresponding eigenvector are all positive. This theorem serves as the theoretical foundation for steady-state analysis of Markov chains and the PageRank algorithm in graph theory.

3. Applications of eigenvalues and eigenvectors

3.1. Principal Component Analysis

Principal Component Analysis (PCA) is an unsupervised linear dimensionality reduction technique. Its core objective is to project high-dimensional data into a lower-dimensional space while preserving the maximum amount of information from the original data.

The mathematical essence of PCA is the eigenvalue decomposition of the covariance matrix. The steps are as follows:

Data Standardization and Covariance Matrix Calculation: Assume there are m data samples, each of n dimensions. First, center the data (so that the mean is 0), then compute its $n \times n$ covariance matrix C. The covariance matrix quantifies the correlation strength between different dimensions of the data.

Eigenvalue Decomposition: Perform eigenvalue decomposition on the covariance matrix C. According to the Spectral Theorem, C is a real symmetric matrix. Therefore, there exists an orthogonal matrix Q and a diagonal matrix Λ such that $C = Q\Lambda Q^T$. Here, the diagonal elements of Λ, denoted as $\lambda_1, \lambda_2, \ldots \lambda_n$ are the eigenvalues of C, and they satisfy $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$. Each column of Q is the corresponding unit eigenvector $v_1, v_2, \ldots, v_n$.

Principal Component Extraction and Dimensionality Reduction: The magnitude of an eigenvalue λ_i represents the variance of the data projection onto the direction of its corresponding eigenvector v_i. A larger variance indicates that more original data information is retained in that direction. Therefore, we select the eigenvectors corresponding to the top k largest eigenvalues (i.e., the principal components) to form an $n \times k$ projection matrix W.

Data Projection: The original data matrix X is projected onto the selected principal component directions, resulting in the reduced k-dimensional data matrix $Y = XW$.

For example, in face recognition, the original image data has an extremely high dimensionality (e.g., a 100 x 100 pixel image has 10,000 dimensions). Direct processing faces the "curse of dimensionality," which is characterized by excessive computational costs and difficulty in capturing data patterns. After applying PCA, the eigenvectors of the covariance matrix are referred to as "eigenfaces." The "eigenfaces" corresponding to the largest few eigenvalues represent the most significant and common variation patterns in face images (such as lighting conditions and overall contours). By retaining the first 100 principal components, the data can be compressed from 10,000 dimensions to 100 dimensions, while preserving the most critical identity information in the images. This significantly improves the efficiency of subsequent classification or recognition tasks. This is one of the key functions of PCA.

PCA also plays a vital role in economic research. Its primary value lies in distilling a small set of key composite indicators from a large number of correlated economic variables, thereby simplifying the analysis of complex economic systems. For instance, when constructing a comprehensive economic development index, economists often face multiple highly correlated variables such as GDP, per capita income, consumption index, and employment rate. Using PCA, a "first principal component" that explains the maximum proportion of variance in the original data can be extracted. This component can serve as a composite score measuring the economic development level of a country or region, effectively avoiding information redundancy and subjective weight assignment inherent in multi-indicator evaluations. Furthermore, in finance, PCA is widely used for market risk modeling. By analyzing a large number of asset return time series, it helps identify common latent factors driving market fluctuations, such as overall market risk and industry-specific risks, providing a quantitative basis for portfolio optimization and risk management.

In principal component analysis, eigenvalues and eigenvectors jointly form the mathematical foundation and decision-making basis for the dimensionality reduction process. Eigenvectors qualitatively determine the structural framework for information extraction by indicating the orthogonal directions of maximum variance, thus providing the optimal projection coordinate system for the high-dimensional data. Eigenvalues, by quantifying the magnitude of variance in each direction, provide an objective criterion for dimension selection. Their values directly reflect the amount of information contained in the corresponding principal components. Consequently, by calculating the cumulative percentage of variance explained, they guide the retention of the most significant data features with minimal dimensionality loss, realizing a seamless translation from mathematical principles to engineering applications.

Despite these strengths, its application has notable limitations. Firstly, PCA is inherently a linear transformation method; its eigenvectors can only capture linear correlations within the data. For datasets with complex nonlinear structures, its dimensionality reduction effectiveness decreases significantly. Secondly, eigenvalues are highly sensitive to outliers. A few extreme values can substantially alter the structure of the covariance matrix, causing the directions of the eigenvectors to shift and thus compromising the stability of the principal components. Furthermore, although dimension selection based on eigenvalue contribution rates quantifies the retained information, it cannot guarantee that the discarded minor components do not contain subtle features with specific semantic meaning (e.g., key discriminatory information for certain classification tasks). This information loss can be irreversible. Finally, performing eigenvalue decomposition on large-scale matrices requires very high computational complexity. When the data dimensionality is extremely high, the computational cost can become prohibitive.

3.2. Markov chains

A Markov chain is a stochastic process model whose core principle is that the probability of transitioning to a future state depends exclusively on the current state and is independent of all previous states. Markov chains find wide applications across various fields including artificial intelligence, probability theory, and statistics.

In the analytical application of Markov chains, the state transition law is described by the transition probability matrix P, where each element P_{ij} represents the probability of transitioning from state i to state j. Eigenvalues and eigenvectors play a crucial role in determining the long-term behavior trend of the chain and verifying the existence of a stationary distribution [6].

The transition probability matrix P of a Markov chain has a natural and critical eigenvalue of 1 (this is determined by the property that "the sum of probabilities in each row equals 1" in Markov chains. Mathematically, it can be proven that 1 is an eigenvalue for any transition probability matrix). The magnitude of the other eigenvalues (denoted as λ) directly determines the speed and possibility of the Markov chain converging to a "steady state":

If the magnitudes of all non-1 eigenvalues are less than 1 (i.e., $|\lambda| < 1$), then the Markov chain is "ergodic"—regardless of the initial state, after a sufficient number of transitions, the state distribution will gradually converge to a unique "stationary distribution" (also called "steady-state distribution") and will no longer change with further transitions.

If there exist non-unity eigenvalues with magnitude equal to 1, the Markov chain may exhibit "periodicity" or contain "multiple disconnected state subsets" (such as two independent state groups with no inter-group transitions). In such cases, there may be no unique stationary distribution, or the state distribution may cycle among multiple distributions (for example, the periodic transition pattern Monday→Tuesday→...→Sunday→Monday demonstrates no long-term stable distribution).

In the steady-state analysis of Markov chains, the solution method based on eigenvalues and eigenvectors demonstrates unique advantages while revealing corresponding limitations. Its core advantage lies in providing an analytical approach with theoretical guarantees for obtaining the stationary distribution.

First, this method offers high computational efficiency. For chains with numerous states, converging to the steady state through simulation iterations can be extremely time-consuming. By transforming the problem into solving for the left eigenvector of the transition matrix P^T corresponding to eigenvalue 1, efficient numerical linear algebra methods (such as the power method or QR algorithm) can be directly applied. This avoids protracted iterative processes, with particularly significant advantages when the state space is large.

Second, the method possesses a solid theoretical foundation. The Perron-Frobenius theorem provides theoretical assurance for the existence of a unique positive stationary distribution in irreducible and aperiodic Markov chains, confirming that this distribution corresponds precisely to the eigenvector of the dominant eigenvalue 1. This ensures that the solution is not merely a computational outcome but a necessary deduction from mathematical theorems, guaranteeing both accuracy and reliability [7].

However, this method also has significant limitations that cannot be overlooked. Its primary constraint lies in its strict theoretical prerequisites. The effectiveness of this method highly depends on the chain satisfying the conditions of being irreducible and aperiodic. For many practical complex systems, this is a strong assumption. If these conditions are not met, the single eigenvector corresponding to eigenvalue 1 cannot accurately characterize the system's behavior.

Secondly, the method faces practical computational challenges. Although theoretically elegant, performing a complete eigen decomposition of the transition matrix for Markov chains with extremely large state spaces remains computationally expensive or even infeasible. While iterative methods for sparse matrices can be used to find the dominant eigenvector, the high dimensionality still presents a bottleneck.

Finally, this method fails to capture transient behavior. The eigenvector approach directly provides the system's long-term limiting behavior but completely ignores the transient process of the system transitioning from its initial state to the steady state. This information about the transient phase is equally crucial for evaluating the system's short-term performance.

4. Conclusion

Firstly, eigenvalues and eigenvectors provide a universal and powerful mathematical framework for revealing the intrinsic structural characteristics of matrices. Secondly, their applications demonstrate significant effectiveness in transforming complex practical problems into solvable mathematical models. Specifically, in PCA, they automatically identify the directions of maximum variance in data, achieving optimal linear dimensionality reduction and feature extraction, thereby effectively overcoming the "curse of dimensionality." In Markov chains, they transform the steady-state solution problem into an eigenvector problem, avoiding lengthy iterative simulations and enabling direct, accurate prediction of the system's long-term behavior. This ability to translate abstract mathematics into concrete applications makes them an indispensable bridge connecting theoretical mathematics with engineering practice. However, this study also reveals inherent limitations of this theoretical approach. The primary limitation lies in its heavy reliance on linear assumptions: PCA cannot effectively handle data with nonlinear manifold structures, while the "memoryless" assumption of Markov chains struggles to characterize complex dynamic systems. Secondly, high computational costs present a practical challenge, as the time and space complexity of eigen decomposition for large-scale matrices can be prohibitive. Additionally, there are interpretability challenges: particularly in PCA, the principal components after dimensionality reduction are linear combinations of original features, often resulting in ambiguous physical meanings that require interpretation through domain knowledge. Future research should focus on overcoming these limitations, such as developing nonlinear dimensionality reduction techniques, optimizing eigenvalue algorithms for large-scale sparse matrices, and actively exploring integration with cutting-edge fields like deep learning and graph neural networks. This will expand their application boundaries to more complex and broader practical scenarios.

References

[1]   Axler, S. (2015). *Linear algebra done right* (3rd ed.). Springer International Publishing.
[2]   Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*(2065), 20150202.
[3]   Grinstead, C. M., & Snell, J. L. (1997). *Introduction to probability.* American Mathematical Society.
[4]   Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics.
[5]   Strang, G. (2016). *Introduction to linear algebra* (5th ed.). Wellesley-Cambridge Press.
[6]   Zhou, Zhihua. (2016). *Machine Learning*. Tsinghua University Press.
[7]   Gilbert, Strang. (2014). *Introduction to Linear Algebra* (5th Edition). Xi'an Jiaotong University Press.