IJCRM

International Journal of Contemporary Research In Multidisciplinary

Research Article

Singular Value Decomposition and Spectral Methods for Low-Rank Learning in Modern Artificial Intelligence

Tathagata Satapathy

Assistant Professor, Department of Mathematics, NIIS Institute of Engineering and Technology Bhubaneswar, Odisha, India

Corresponding Author: * Tathagata Satapathy

DOI: https://doi.org/10.5281/zenodo.20522517

Abstract

Efficient mathematical modelling and optimisation are essential for modern computational learning systems, as Artificial Intelligence and Machine Learning systems are increasingly dependent on high-dimensional data representations and large-scale neural architectures. Singular Value Decomposition (SVD) and spectral methods are mathematical tools in AI, and have become powerful techniques for dimensionality reduction, latent feature extraction, and low-rank approximation. In this paper, we discuss the mathematical basis of SVD from the perspective of vector spaces, orthogonality, matrix transformations, eigenvalue analysis and spectral decomposition theory and analyse their importance in modern Artificial Intelligence systems. We study the role of low-rank spectral structures in applications such as principal component analysis, recommendation systems, image compression, natural language processing and neural representation learning. Special attention is devoted to the computational challenges posed by Large Language Models. Keywords—Singular Value Decomposition; Spectral Methods; Low-Rank Learning; Artificial Intelligence; Large Language Models.

Manuscript Information

- **ISSN No:** 2583-7397
- **Received:** 08-04-2026
- **Accepted:** 31-05-2026
- **Published:** 03-06-2026
- **IJCRM:**5(3); 2026: 561-567

- **Plagiarism Checked:** Yes
- **Peer Review Process:** Yes

How to Cite this Article

Satapathy T. Singular Value Decomposition and Spectral Methods for Low-Rank Learning in Modern Artificial Intelligence. Int J Contemp Res Multidiscip. 2026;5(3):561-567.

Access this Article Online

www.multiarticlesjournal.com

KEYWORDS: Singular Value Decomposition, Spectral Methods, Low-Rank Learning, Artificial Intelligence, Large Language Models.

INTRODUCTION

Artificial Intelligence (AI) and Machine Learning (ML) have emerged as computational paradigms that can handle large high-dimensional data sets in areas such as Natural Language Processing, computer vision, recommendation systems and autonomous learning architectures. Modern intelligent systems, and in particular Large Language Models (LLMs), are characterised by very large parameter spaces, as evidenced by high-rank matrices and complicated neural transformations. These architectures show impressive learning capability, but the increasing dimensionality raises considerable problems concerning computational complexity, memory usage, parameter redundancy and scalable optimisation.

In modern AI systems, linear algebra and spectral theory are fundamental tools for analysing hidden structural relationships in high-dimensional data spaces. Among various spectral methods, Singular Value Decomposition (SVD) has become one of the most important mathematical tools for dimensionality reduction, latent feature extraction, noise suppression and low-rank approximation. Decomposition of a matrix into orthogonal spectral components,

$$A = U\Sigma V^T$$

SVD reveals dominant singular structures that keep important information while reducing unnecessary computational complexity. Low-rank spectral learning has become increasingly important in recent years, as transformer architectures and Large Language Models have been growing at a rapid pace. Many modern neural systems exhibit a significant spectral redundancy in which meaningful learning information is confined to lower-dimensional subspaces. Therefore, techniques for low-rank approximation have gained increasing importance for efficient parameter adaptation and scalable learning.

More recent methods, such as Low-Rank Adaptation (LoRA), show that updates to neural parameters can be well approximated by low-rank spectral structures, rather than full-rank optimization . These recent developments have motivated the present work to study the mathematical foundations of Singular Value Decomposition and spectral methods for low-rank learning in Artificial Intelligence systems.

This paper further develops the spectral interpretation of matrix decomposition, orthogonality, eigen analysis, and low-rank approximation, and investigates the applications of these techniques in di-mensionality reduction, neural representation learning, transformer optimization, and Large Language Models. Moreover, a conceptual spectral low-rank optimization framework is proposed to study the potential role of dominant singular structures in computationally efficient and scalable AI architectures.

Mathematical Preliminaries

Modern Artificial Intelligence systems operate in high dimensional vector spaces, and matrix form is often used to describe data representations, neural activations, embeddings and transformer parameters. Thus the mathematical background to understand computational learning structures and optimization mechanisms is linear algebra and spectral analysis. For a vector space $V \subseteq \mathbb{R}^n$, a vector may be represented as

$$\mathbf{v} = (x_1, x_2, \ldots, x_n)$$

where each coordinate is a feature component in the representation space. In large scale AI systems the dimensionality of such spaces becomes extremely large often creating redundancy and computational inefficiency. The inner product is a measure of the similarity between vectors

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i$$

which geometrically defines angular alignment and structural correlation between representations. Orthogonality is crucial for spectral learning because orthogonal vectors correspond to independent directions of information that satisfy

$$\mathbf{u} \cdot \mathbf{v} = 0$$

Orthogonal structures are fundamental in dimensionality reduction, latent feature extraction and neural representation learning, due to their property of preserving geometric consistency while minimizing redundancy. Matrices can be considered as linear transformations of vector spaces. For a matrix transform,

$$Ax = b$$

the matrix A transforms the vector x into the new representation b. Repeated matrix transformations in transformer architectures and neural networks result in hierarchical feature representations in high dimensional embedding spaces. Eigen structures are also introduced by spectral analysis by the relation

$$Av = \lambda v$$

Where v is an eigenvector and is the eigen value. Eigenvectors specify directions of invariant structures, eigenvalues specify their spectral significance.

These concepts form the algebraic backbone of matrix factorization and low rank approximation for us. Numerous large scale AI applications have low dimensional dominant spectral structure in a high rank parameter space, spectral decomposition methods are very important in effective representation, compression and scalable optimization.

Singular Value Decomposition

Singular Value Decomposition (SVD) is one of the most important spectral factorization techniques and can reveal hidden low rank structures in high dimensional matrices. It has become very important in Artificial Intelligence, Machine Learning and large-scale neural optimization, due to its ability to separate dominant information from redundancy and noise. For a matrix

$$A \in \mathbb{R}^{m \times n}$$

the Singular Value Decomposition is defined as

$$A = U\Sigma V^T$$

where U and V are orthogonal matrices and Σ is a diagonal matrix containing singular values satisfying

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$$

The orthogonality conditions are given by

$$U^T U = I$$

and

$$V^T V = I$$

while the singular vectors are obtained from the spectral relations

$$A^T A\, v_i = \sigma_i^2 v_i$$

and

$$AA^T u_i = \sigma_i^2 u_i$$

An equivalent spectral representation of SVD is

$$\sum_{i=1}^{r} \sigma_i u_i v_i^T$$

representing the matrix as a weighted sum of orthogonal rank-one components. Geometrically, SVD breaks down a matrix transformation into a series of orthogonal rotations and spectral scalings. The larger singular values correspond to the main directions of information in the space of data. The smaller singular values often correspond to redundant or noisy information. As a result, a truncated SVD gives an efficient low-rank approximation:

$$A_k = U_k \Sigma_k V_k^T$$

Where only the leading k singular structures are kept. In this way, the dimensional complexity is greatly reduced, yet the key spectral features of the original matrix are retained.

Artificial Intelligence nowadays uses SVD as a major workhorse for dimensionality reduction, latent feature extraction, recommendation systems, image compression, Natural Language Processing, and transformer optimization. Low-rank spectral methods that are inspired by SVD have, in fact, lately acquired major prominence in Large Language Models, where very large parameter matrices pose huge computational and memory challenges.

Low-Rank Approximation and Spectral Compression

The parameter matrices in modern Artificial Intelligence systems are often very large and directly optimizing them results in high computational complexity and memory overhead. However, many of the high-dimensional matrices have a dominant low-rank spectral structure with the relevant infor-mation concentrated in a smaller subspace. This observation is the mathematical basis of low-rank approximation and spectral compression. For matrices factorization,

$$A = U\Sigma V^T$$

the singular values satisfy

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$$

where the larger singular values correspond to the dominant directions of the spectral information. The dominant singular components are retained to obtain a truncated rank-k approximation:

$$A_k = U_k \Sigma_k V_k^T$$

where:

- U_k contains the first k left singular vectors,
- V_k contains the first k right singular vectors,
- and Σ_k contains the largest singular values.

The optimality of truncated SVD is established through the Eckart–Young theorem:

$$\|A - A_k\|_F = \min_{\text{rank}(B) \leq k} \|A - B\|_F$$

where the Frobenius norm is defined by

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$$

Thus, truncated SVD provides the best rank-k approximation while preserving maximum spectral information. The retained spectral energy ratio is given by

$$E_k = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2}$$

which measures how much information is preserved after low-rank compression. In many practical AI systems, a small number of dominant singular values preserve most of the essential structural information. Low-rank spectral compression has become increasingly important in modern Machine Learning and transformer architectures. In image compression and latent semantic analysis, dominant singular structures capture the principal information patterns while removing redundancy and noise. More recently, Large Language Models have demonstrated substantial parameter redundancy within extremely highrank neural weight matrices. Consequently, spectral lowrank approximation has emerged as an important strategy for scalable parameter-efficient optimization. If W denotes a neural parameter matrix, a low-rank spectral update may be expressed as

$$W' = W + U_k \Sigma_k V_k^T$$

where only the dominant spectral components are used for parameter adaptation This significantly reduces the memory footprint and the complexity of the optimization process, while keeping the necessary information on the learning process. Thus, the mathematical basis of modern scalable AI optimization frameworks such as transformer compression, parameter-efficient fine-tuning, and low-rank adaptation methods such as LoRA is low-rank approximation and spectral compression.

Applications in Artificial Intelligence

Spectral methods and low-rank matrix approximations have established themselves as fundamental tools in modern Artificial Intelligence due to their ability to extract dominant latent structures from high dimensional data. Singular Value Decomposition and related spectral methods are widely used for dimensionality reduction, feature learning, semantic analysis and scalable neural optimization. In Principal Component Analysis (PCA), covariance matrices are spectrally decomposed to find dominant directions of variance in the data space . If X is a data matrix with zero mean the covariance matrix is

$$C = \frac{1}{n} X^T X$$

and the principal components are obtained through the eigenvalue relation

$$Cv_i = \lambda_i v_i$$

where the main eigenvectors capture the most variance of the data. Therefore, PCA reduces the complexity of dimensions while keeping the important information structures. Spectral decomposition is often used in latent semantic analysis and embedding compression in Natural Language Processing (NLP). Given a term-document matrix A,

$$A = U \Sigma V^T$$

enables the representation of implicit semantic relations in lower dimensional latent subspaces. Domi-nant singular vectors encode contextual semantic structure, but suppress redundancy and sparse noise. Low-rank matrix factorization is also used to model the latent user-item interaction in recommendation systems. If R is a sparse recommendation matrix, spectral factorization approximates

$$R \approx PQ^T$$

where P and Q are the low-rank latent feature matrices. This allows scalable prediction of the missing interactions while significantly reducing the computational complexity. Spectral compression has also important applications in image processing and neural representation learning. Truncated singular structures are noise resistant, need less storage and maintain most of the visual information. More recently, transformer architectures and Large Language Models have increasingly relied on low-rank spectral methods for efficient parameter learning. Attention and embedding matrices often contain a lot of redundancy and so spectral approximation can be used for efficient compression and scalable optimization of modern neural systems.

Spectral learning frameworks have thus evolved from classical matrix analysis tools to central compu-tational mechanisms for scalable Artificial Intelligence and high-dimensional representation learning.

Advanced Spectral Learning Frameworks

The recent developments in Artificial Intelligence have made low-rank spectral optimization an impor-tant topic for large-scale neural architectures. Modern transformer systems and Large Language Models (LLMs) are composed of highly high-dimensional parameter matrices whose optimization in full-rank requires extensive computation resources and memory consumption. To overcome these restrictions, parameter-efficient spectral learning methods, e.g., Low-Rank Adaptation (LoRA), approximate neural parameter updates with low-rank matrix structures. For a neural weight matrix W, LoRA replaces the full parameter optimization by a low-rank update of the form

$$W' = W + BA$$

Where

$$B \in \mathbb{R}^{d \times r}, \quad A \in \mathbb{R}^{r \times k}, \quad r \ll \min(d, k)$$

Therefore, the parameter adaptation is constrained to a low dimensional spectral subspace, which reduces significantly the trainable parameters while maintaining learning efficiency.

Singular Values, Randomized Decomposition has also emerged as a significant computational approach for largescale spectral approximation. Randomized projections provide approximations of the dominant singular structures at a much lower computational cost than direct computation of the full matrix decompositions. for a random projection matrix , Ω,

$$Y = A\Omega$$

captures the dominant spectral subspace of A, allowing for scalable approximation for massive datasets and transformer architectures. Tensor decomposition methods generalize low-rank spectral learning to higher-order neural representations. Factorized components of a tensorX can be used to approximate:

$$X \approx \sum_{i=1}^{r} a_i \otimes b_i \otimes c_i$$

where ⊗ denotes the tensor product. Such decompositions are useful for reducing parameter redundancy in multi-dimensional neural systems. More recently, spectral learning frameworks have gained traction for transformer optimisation, attention compression, and scalable Large Language Models. The dominant singular structures in neural parameter matrices allow spectral low-rank approximation to efficiently use memory, adapt faster, and reduce optimization complexity. Thus, modern spectral learning methods offer an important mathematical foundation for scalable and computationally efficient Artificial Intelligence systems.

Proposed Spectral Low-Rank Optimization Framework

Modern LLMs have very high dimension transformer parameter matrices and optimizing them directly is computationally intensive and memory consuming. Full-rank optimization enhances the representational ability; however, existing neural systems are often highly spectrally redundant, i.e., the majority of learning information is accumulated in the lower-dimensional latent subspaces.

Consider a transformer weight matrix:

$$W \in \mathrm{R} m \times n$$

Standard full-rank fine-tuning performs parameter adaptation through

$$W' = W + \Delta W$$

where

$$\Delta W \in \mathrm{R}^{m\times n}$$

Thus, the total trainable parameter complexity becomes

$$P_{\text{full}} = mn$$

which becomes computationally expensive for large transformer architectures. To reduce optimization cost, Low-Rank Adaptation (LoRA) approximates the parameter update using low-rank matrices:

$$\Delta W = BA$$

were

$$B \in \mathrm{R}^{m\times r}, \qquad A \in \mathrm{R}^{r\times n}, \qquad r \ll \min(m, n)$$

The trainable parameter complexity is therefore reduced to

$$P_{\text{LoRA}} = r(m+n)$$

Although LoRA significantly reduces optimization dimensionality, the factorization does not explicitly preserve dominant spectral structures of the original neural parameter space. Consequently, spectral information preservation is not mathematically guaranteed. Motivated by this limitation, the proposed framework introduces an explicitly spectral low-rank optimization strategy based on Singular Value Decomposition. For the matrix W,

$$W = U\Sigma V^T$$

where:

U and V are orthogonal singular vector matrices,

and

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$$

contains ordered singular values satisfying

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$$

The matrix may be expanded spectrally as

$$W = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

where each term represents a rank-one spectral component weighted by its singular value.

Since dominant singular values preserve the principal latent structure of the parameter space, the proposed framework retains only the leading k singular structures:

$$W_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

with

$$k \ll r$$

The discarded spectral error becomes

$$W - W_k = \sum_{i=k+1}^{r} \sigma_i u_i v_i^T$$

Using the Frobenius norm,

$$\|W - W_k\|_F^2 = \sum_{i=k+1}^{r} \sigma_i^2$$

Thus, the approximation error depends only on the discarded singular values. Now consider any arbitrary rank-k matrix B. By the Eckart–Young theorem,

$$\|W - W_k\|_F = \min_{\text{rank}(B)=k} \|W - B\|_F$$

Hence, W_k provides the optimal rank-k approximation among all possible rank-k matrices.

This result establishes the central mathematical advantage of the proposed framework. Unlike conventional low-rank adaptation methods that employ arbitrary factorisation, the proposed framework explicitly preserves dominant singular structures associated with maximum spectral information en-ergy. The retained spectral energy ratio is given by

$$E_k = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2}$$

If

$E_k \rightarrow 1$

then, despite major dimensional compression, the structural information of most of the original neural parameter matrix is retained. So the adaptive spectral optimization update is then written as

$$W' = W + \alpha U_k \Sigma_k V^T$$

where:

- $U_k \Sigma_k V^T_k$ preserves dominant spectral directions,
- k controls the compression rank,
- and α denotes the spectral adaptation coefficient.

The corresponding trainable parameter complexity becomes

$$P_{\text{proposed}} = k\,(m + n + 1)$$

which is much less than the full rank optimization while preserving the dominant spectral information. Hence, the proposed framework theoretically provides:

- Best rank-k approximation,
- Preservation of the spectral information dominant,
- Reduced optimization dimension,
- Spectral redundancy reduction,
- Computationally efficient transformer adaptation.

The entire optimization process is carried out through four main stages:

1. Spectral decomposition of transformer parameter matrices
2. Extraction of dominant singular structures
3. Low-rank spectral adaptation
4. Reconstruction of optimized neural parameters

From a computational point of view, the framework attempts to reduce:

- Memory overhead,
- Number of trainable parameters,
- Optimization duplicity,
- Cost of fine-tuning at a large scale.

Thus, the proposed spectral framework provides a mathematically informed low-rank optimization strategy that bridges classical spectral decomposition theory with scalable transformer learning and modern Large Language Model optimization.

CONCLUSION

Artificial Intelligence and Large Language Models have opened up numerous possibilities but simul-taneously have increased the need for computationally efficient learning systems that can handle very high dimensional parameter spaces. This paper focused on the study of the mathematical underpinnings of using Singular Value Decomposition and spectral methods for low-rank learning and scalable neural optimization.

The article develops the spectral perspective of matrix decomposition orthogonality eigen analysis and low rank approximation, and then explores the ways in which these can be used for dimensionality reduction, latent semantic learning, recommendation systems, transformer compression and modern neural architectures. The research demonstrated the role of dominant singular structures in preserving the main information, while also eliminating redundancy and reducing computational complexity. And, it introduced a spectral low-rank optimization conceptual system for Large Language Models through dominant singular subspace adaptation. Unlike classical low-rank factorization methods, the proposed system clearly maintains the principal spectral directions with Singular Value Decomposition. Continuing from the Eckart-Young theorem and Frobenius norm, the setup is proved to give an optimally approximate rank-k with far less optimization dimensionality. The spectral formulation presented

$$W' = W + \alpha(U_k V_k^T)$$

demonstarte how dominant singular structures can be leveraged to boost computationally efficient transformer adaptation and scalable parameter-efficient learning. While the current paper largely offers a mathematical and conceptual setup, it also paves the way for multiple research directions like adaptive rank selection, dynamic spectral learning, distributed transformer optimization, attention compression, and scalable foundation model architectures. That means, spectral low-rank learning is the key mathematical tool that can dramatically enhance the efficiency, scalability, and interpretability of Artificial Intelligence systems as the most advanced neural computing environment.

REFERENCES

1. Golub GH, Van Loan CF. *Matrix Computations*. 4th ed. Baltimore (MD): Johns Hopkins University Press; 2013.
2. Bishop CM. *Pattern Recognition and Machine Learning*. New York (NY): Springer; 2006.
3. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge (MA): MIT Press; 2016.
4. Halko N, Martinsson PG, Tropp JA. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. SIAM Review. 2011;53(2):217-288.
5. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, *et al*. Attention is all you need. Advances in Neural Information Processing Systems. 2017;30:5998-6008.
6. Hu EJ, Shen Y, Wallis P, Allen-Zhu Z, Li Y, Wang S, *et al*. LoRA: low-rank adaptation of large language models. International Conference on Learning Representations. 2022.
7. Pearson K. On lines and planes of closest fit to systems of points in space. Philosophical Magazine. 1901;2(11):559-572.
8. Eckart C, Young G. The approximation of one matrix by another of lower rank. Psychometrika. 1936;1(3):211-218.
9. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521:436-444.

About the Corresponding Author

Tathagata Satapathy is an Assistant Professor of Mathematics at NIIS Institute of Engineering and Technology, Bhubaneswar, Odisha. His academic interests include mathematics education, applied mathematics, and interdisciplinary research. He is actively engaged in teaching, research, and mentoring students in higher education.