

Mathematical Foundations of Eigenvalue Decomposition, Singular Value Decomposition (SVD) and Principle Component Analysis (PCA)

Name: Alec Follmer

Date: 11/19/2025

Introduction

Many modern areas of chemistry that we work on from structural biology, spectroscopy, ultrafast dynamics, to molecular simulations produce large numerical datasets with correlated structure. Often underlying these datasets is the assumption that the measured signals arise from a small number of chemically meaningful processes (electronic transitions, structural modes, reaction intermediates) embedded in a much larger space of possible measurements. Linear algebra provides the mathematical framework that identifies the directions of correlated variation, allowing us to separate signal from noise and extract physically interpretable components.

The power of techniques such as eigenvalue decomposition, singular value decomposition (SVD), and principal component analysis (PCA) comes from their ability to reveal low-dimensional structure within high-dimensional measurements. These methods do not require explicit knowledge of the underlying chemistry or kinetics. Instead, they rely on the assumptions that: (1) chemical processes produce coherent, correlated changes in the data; (2) noise is incoherent and therefore distributed across many directions; and (3) linear superposition is a valid approximation for spectroscopic or structural signals. Under these assumptions, the dominant eigenvectors or singular vectors extract the chemically meaningful subspace.

These decompositions are valuable because they offer model-independent insight. Eigenvalue decomposition identifies the intrinsic “directions” along which a linear operator acts; SVD generalizes this to any rectangular dataset; and PCA provides a statistically optimal low-dimensional representation via variance maximization. Together, these tools allow us to identify components, quantify their contributions, and build mechanistic hypotheses grounded in the linear algebraic structure of the data.

1 Eigenvalue Decomposition

For a square matrix $A \in \mathbb{R}^{n \times n}$, an eigenvalue λ and eigenvector x satisfy

$$Ax = \lambda x.$$

If A is diagonalizable, then

$$A = V\Lambda V^{-1},$$

where:

- $V = [x_1 \ x_2 \ \cdots \ x_n]$ contains eigenvectors,
- $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ contains eigenvalues.

Eigenvalue decomposition identifies the “natural” coordinate system of the operator A : directions in which the transformation reduces to simple scaling. This structure will later inform the mathematics of SVD and PCA.

$$A\vec{x} = \lambda\vec{x}$$

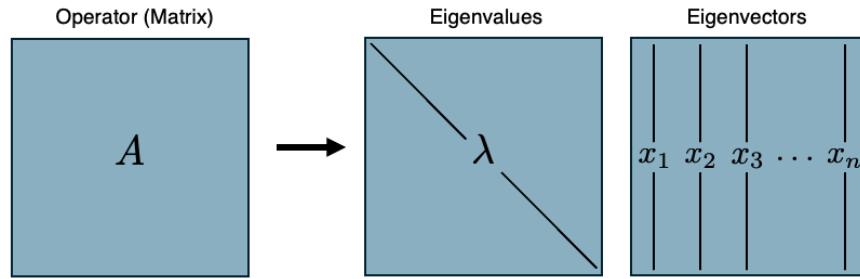


Figure 1: Conceptual illustration of eigenvalue decomposition for a square operator A . The transformation $A\vec{x} = \lambda\vec{x}$ identifies special directions \vec{x}_i (eigenvectors) that are simply scaled by the corresponding eigenvalues λ_i . Diagonalizing A via its eigenvalues and eigenvectors reveals the natural coordinate system in which the action of the operator is decoupled.

2 Full Singular Value Decomposition

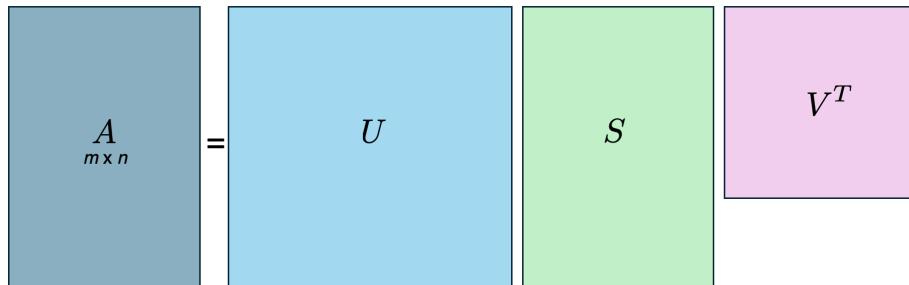


Figure 2: Schematic representation of the full singular value decomposition (SVD). Any $m \times n$ matrix A can be factored into the product $A = USV^T$, where U and V are orthonormal matrices containing the left and right singular vectors, respectively, and S contains the singular values. This decomposition expresses the linear transformation encoded in A as a rotation, scaling, and a second rotation.

Let $A \in \mathbb{R}^{m \times n}$ be *any* (possibly rectangular) matrix. The full SVD expresses A as

$$A = USV^T,$$

where:

- $U \in \mathbb{R}^{m \times m}$ is orthonormal,
- $V \in \mathbb{R}^{n \times n}$ is orthonormal,
- $S \in \mathbb{R}^{m \times n}$ is diagonal (rectangular) with entries $s_1 \geq s_2 \geq \dots \geq 0$.

The matrices AA^T and A^TA connect full SVD to eigenvalue theory:

$$AA^T = USS^TU^T, \quad A^TA = VS^TSV^T.$$

Thus:

$$A^TA\mathbf{v}_i = s_i^2\mathbf{v}_i,$$

meaning singular values are the square roots of the eigenvalues of A^TA .

Coordinates and Meaning

- U : orthonormal basis in the m -dimensional measurement space (e.g., “spectral space”).
- S : strengths of each component.
- V^T : orthonormal basis in the n -dimensional condition space (e.g., “time” or “experimental coordinate” space).

Singular Value Decomposition (SVD) is especially powerful for time-resolved spectroscopic experiments—such as the energy-vs-time maps characteristic of TR-XES, TR-XAS, or ultrafast optical spectroscopy because it provides a mathematically rigorous way to separate spectral signatures from temporal dynamics. In these experiments, each measurement is a 2D dataset where one axis corresponds to energy (or wavelength) and the other to delay time. The raw data inevitably contain a mixture of meaningful signal, overlapping spectral contributions from multiple transient species, and substantial noise. SVD decomposes this matrix into orthogonal spectral and temporal components, allowing us to quickly assess the dimensionality of the dataset, identify dominant time-dependent processes, and determine whether clean kinetic behavior is present before further modeling. Practically, this means SVD acts as an immediate diagnostic tool: it reveals whether we have captured real dynamical evolution, whether additional species or states participate in the reaction, and whether the data quality supports deeper kinetic or spectroscopic analysis.

3 From Full SVD to Thin SVD

Although full SVD is mathematically complete, experimental data almost always satisfy $m \gg n$, meaning only the first n singular values can be nonzero.

Thus we can truncate:

$$U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_n] \in \mathbb{R}^{m \times n},$$

$$S_n = \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

This yields the **thin SVD**:

$$A = U_n S_n V^T.$$

This is the form used in the presentation:

$$A_{m \times n} = \underbrace{U_{m \times n}}_{\text{spectral modes}} \quad \underbrace{S_{n \times n}}_{\text{singular values}} \quad \underbrace{V^T_{n \times n}}_{\text{temporal / experimental modes}}.$$

Columns of A are denoted

$$A = [a_1 \ a_2 \ \cdots \ a_n],$$

and the decomposition gives

$$a_j = \sum_{i=1}^n s_i \mathbf{u}_i v_{ij}.$$

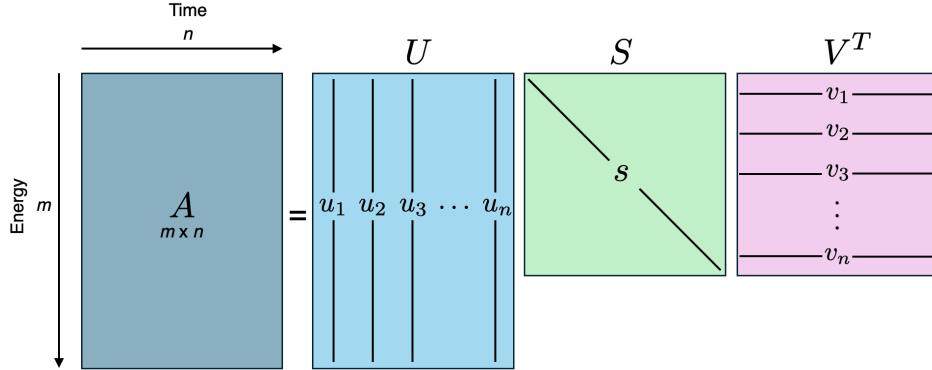


Figure 3: Thin singular value decomposition (SVD) for a data matrix $A \in \mathbb{R}^{m \times n}$. Here A is decomposed as $A = U_n S_n V^T$, where the columns of U_n (u_1, u_2, \dots, u_n) represent orthonormal basis functions in the measurement space, S_n contains the nonzero singular values, and the rows of V^T (v_1, v_2, \dots, v_n) encode the amplitudes or time dependences of each component. This form reflects the dimensionally reduced version of the full SVD.

4 Low-Rank Approximation

The low-rank approximation that naturally arises from SVD is particularly important for chemical systems because the number of chemically meaningful species involved in a reaction is almost always far smaller than the number of measured time points or energy channels. Even though a dataset may contain hundreds or thousands of sampled energies and dozens of time delays, the underlying sys-

tem might involve only two or three intermediates. SVD exploits this structure by concentrating the chemically meaningful information into the first few singular components, while higher-order components tend to correspond to noise or minor experimental artifacts. This provides not only a powerful form of denoising, but also a principled way to reduce model complexity: kinetic modeling, spectral decomposition, and global fitting can all be performed in a reduced space where each component corresponds approximately to a true chemical species or a distinct dynamical process. In practice, this makes SVD indispensable for interpreting large spectroscopic datasets because it bridges the gap between high-dimensional experimental measurements and the low-dimensional mechanistic space in which the chemistry actually occurs.

SVD provides the expansion:

$$A = \sum_{i=1}^r s_i \mathbf{u}_i \mathbf{v}_i^T,$$

where $r = \text{rank}(A)$.

The best rank- k approximation (in the least-squares sense) is

$$A_k = \sum_{i=1}^k s_i \mathbf{u}_i \mathbf{v}_i^T.$$

This is fundamental for denoising and analyzing spectroscopic datasets: large s_i correspond to coherent signal; small s_i usually reflect noise.

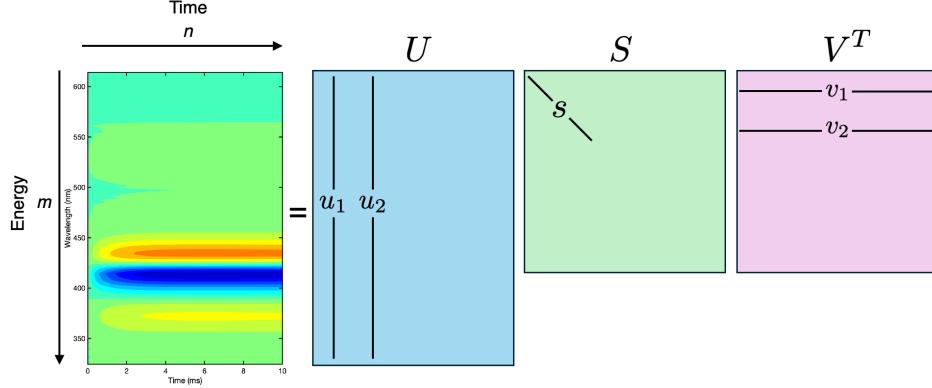


Figure 4: Application of the thin SVD to a two-dimensional stopped-flow spectroscopic dataset $A_{m \times n}$ (e.g., wavelength vs. pump–probe delay). The matrix is decomposed into spectral basis vectors u_1, u_2, \dots , singular values s_i , and temporal components v_i . The left singular vectors describe orthogonal spectral signatures, while the right singular vectors capture their corresponding time evolutions, providing a compact representation of the dataset.

5 PCA via the Covariance Matrix

Let A be mean-centered such that each row has zero mean. The covariance matrix is

$$C = \frac{1}{n-1} AA^T.$$

Eigenvalue decomposition of C gives

$$C\mathbf{u}_i = \lambda_i \mathbf{u}_i.$$

If $A = USV^T$ is the thin SVD, then

$$C = U \left(\frac{S^2}{n-1} \right) U^T.$$

Thus:

- PCA eigenvectors = \mathbf{u}_i (left singular vectors),
- PCA variances = $s_i^2/(n-1)$,
- PCA scores = SV^T .

This shows that PCA is *exactly* SVD applied to a mean-centered dataset, with no additional assumptions.

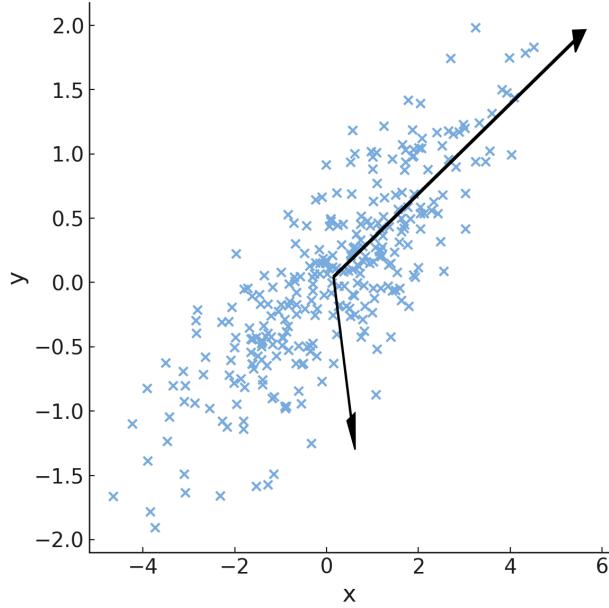


Figure 5: Illustration of principal component analysis (PCA) for a two-dimensional dataset. The centered data points (light blue) exhibit correlated variation. The black arrows represent the first and second principal components, obtained from the singular value decomposition of the data matrix. The first principal component aligns with the direction of maximum variance, while the second spans the orthogonal subspace capturing the remaining variance.

6 Summary

Eigenvalue decomposition, singular value decomposition, and principal component analysis form a unified mathematical framework for understanding structure in high-dimensional chemical data. Eigenvalue decomposition provides the foundational idea: that a linear transformation can often be understood through a set of special directions that it simply stretches or compresses. Singular value decomposition extends this principle to any rectangular matrix, allowing even complex experimental datasets to be expressed as orthogonal components whose contributions are ranked by their singular values. By moving from the full SVD to its thin form, we isolate precisely the subspace where meaningful chemical variation occurs while discarding directions dominated by noise.

PCA emerges naturally from this framework. By examining the covariance of the data or equivalently, by applying SVD to a mean-centered matrix, we can identify the dominant axes of variation that best capture correlated chemical changes. These principal components separate spectral signatures from their kinetic or experimental trajectories, making it possible to decode underlying processes even when they are not directly observable.

In practice, these tools offer a powerful way to interpret the multidimensional measurements common in modern spectroscopy, crystallography, and dynamics studies. They sharpen our ability to see patterns, distinguish signal from noise, and recover mechanistic insight from large datasets. Together, they provide a coherent strategy for reducing complexity while preserving the essential physics and chemistry that drive molecular behavior.