Singular value decomposition

In linear algebra, the **singular value decomposition** (**SVD**) is a factorization of a real or complex matrix. It is the generalization of the eigendecomposition of a positive semidefinite normal matrix (for example, a symmetric matrix with non-negative eigenvalues) to any $m \times n$ matrix via an extension of the polar decomposition. It has many useful applications in signal processing and statistics.

Formally, the singular value decomposition of an $m \times n$ real or complex matrix $\mathbf{M}$ is a factorization of the form $\mathbf{U\Sigma V^*}$, where $\mathbf{U}$ is an $m \times m$ real or complex unitary matrix, $\mathbf{\Sigma}$ is an $m \times n$ rectangular diagonal matrix with non-negative real numbers on the diagonal, and $\mathbf{V}$ is an $n \times n$ real or complex unitary matrix. The diagonal entries σ_i of $\mathbf{\Sigma}$ are known as the **singular values** of $\mathbf{M}$. The columns of $\mathbf{U}$ and the columns of $\mathbf{V}$ are called the **left-singular vectors** and **right-singular vectors** of $\mathbf{M}$, respectively.

The singular value decomposition can be computed using the following observations:

- The left-singular vectors of $\mathbf{M}$ are a set of orthonormal eigenvectors of $\mathbf{MM^*}$.
- The right-singular vectors of $\mathbf{M}$ are a set of orthonormal eigenvectors of $\mathbf{M^*M}$.
- The non-negative singular values of $\mathbf{M}$ (found on the diagonal entries of $\mathbf{\Sigma}$) are the square roots of the non-negative eigenvalues of both $\mathbf{M^*M}$ and $\mathbf{MM^*}$.

Applications that employ the SVD include computing the pseudoinverse, least squares fitting of data, multivariable control, matrix approximation, and determining the rank, range, and null space of a matrix.

Visualization of the SVD of a 2D, real shearing matrix $\mathbf{M}$. First, we see the unit disc in blue together with the two canonical unit vectors. We then see the actions of $\mathbf{M}$, which distorts the disk to an ellipse. The SVD decomposes $\mathbf{M}$ into three simple transformations: an initial rotation $\mathbf{V^*}$, a scaling $\mathbf{\Sigma}$ along the coordinate axes, and a final rotation $\mathbf{U}$. The lengths σ_1 and σ_2 of the semi-axes of the ellipse are the singular values of $\mathbf{M}$, namely $\mathbf{\Sigma}_{1,1}$ and $\mathbf{\Sigma}_{2,2}$.

Contents

Visualisation of the matrix multiplications in singular value decomposition

Statement of the theorem

Suppose $\mathbf{M}$ is an $m \times n$ matrix whose entries come from the field K, which is either the field of real numbers or the field of complex numbers. Then the singular value decomposition of $\mathbf{M}$ exists, and is a factorization of the form

$$\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$$

where

- $\mathbf{U}$ is an $m \times m$ unitary matrix over K (if $K = \mathbb{R}$, unitary matrices are orthogonal matrices),
- $\boldsymbol{\Sigma}$ is a diagonal $m \times n$ matrix with non-negative real numbers on the diagonal,
- $\mathbf{V}$ is an $n \times n$ unitary matrix over K, and $\mathbf{V}^*$ is the conjugate transpose of $\mathbf{V}$.

The diagonal entries σ_i of $\boldsymbol{\Sigma}$ are known as the **singular values** of $\mathbf{M}$. A common convention is to list the singular values in descending order. In this case, the diagonal matrix, $\boldsymbol{\Sigma}$, is uniquely determined by $\mathbf{M}$ (though not the matrices $\mathbf{U}$ and $\mathbf{V}$ if $\mathbf{M}$ is not square, see below).

Intuitive interpretations

Rotation, scaling

In the special, yet common case when $\mathbf{M}$ is an $m \times m$ real square matrix with positive determinant: $\mathbf{U}$, $\mathbf{V}^*$, and $\boldsymbol{\Sigma}$ are real $m \times m$ matrices as well. $\boldsymbol{\Sigma}$ can be regarded as a scaling matrix, and $\mathbf{U}$, $\mathbf{V}^*$ can be viewed as rotation matrices. Thus, the expression $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$ can be intuitively interpreted as a composition of three geometrical transformations: a rotation or reflection, a scaling, and another rotation or reflection. For instance, the figure explains how a shear matrix can be described as such a sequence.

Using the polar decomposition theorem, we can also consider $\mathbf{M} = \mathbf{RP}$ as the composition of a stretch (positive definite matrix $\mathbf{P} = \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^*$) with eigenvalue scale factors σ_i along the orthogonal eigenvectors $\mathbf{V}_i$ of $\mathbf{P}$, followed by a single rotation (unitary matrix $\mathbf{R} = \mathbf{UV}^*$). If the rotation is done first, $\mathbf{M} = \mathbf{P'R}$, then $\mathbf{R}$ is the same and $\mathbf{P'} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^*$ has the same eigenvalues, but is stretched along different (post-rotated) directions. This shows that the SVD is a generalization of the eigenvalue decomposition of pure stretches in orthogonal directions (symmetric matrix $\mathbf{P}$) to arbitrary matrices ($\mathbf{M} = \mathbf{RP}$) which both stretch and rotate.

Singular values as semiaxis of an ellipse or ellipsoid

As shown in the figure, the singular values can be interpreted as the magnitude of the semiaxis of an ellipse in 2D. This concept can be generalized to n-dimensional Euclidean space, with the singular values of any $n \times n$ square matrix being viewed as the magnitude of the semiaxis of an n-dimensional ellipsoid. Similarly, the singular values of any $m \times n$ matrix can be viewed as the magnitude of the semiaxis of an n-dimensional ellipsoid in m-dimensional space, for example as an ellipse in a (tilted) 2D plane in a 3D space. Singular values encode magnitude of the semiaxis, while singular vectors encode direction. See below for further details.

The columns of *U* and *V* are orthonormal bases

Since $\mathbf{U}$ and $\mathbf{V}^*$ are unitary, the columns of each of them form a set of orthonormal vectors, which can be regarded as basis vectors. The matrix $\mathbf{M}$ maps the basis vector $\mathbf{V}_i$ to the stretched unit vector $\sigma_i \mathbf{U}_i$ (see below for further details). By the definition of a unitary matrix, the same is true for their conjugate transposes $\mathbf{U}^*$ and $\mathbf{V}$, except the geometric interpretation of the singular values as stretches is lost. In short, the columns of $\mathbf{U}$, $\mathbf{U}^*$, $\mathbf{V}$, and $\mathbf{V}^*$ are orthonormal bases. When the $\mathbf{M}$ is a normal matrix, $\mathbf{U}$ and $\mathbf{V}^*$ reduce to the unitary used to diagonalize $\mathbf{M}$. However, when $\mathbf{M}$ is not normal but still diagonalizable, its eigendecomposition and singular value decomposition are distinct.

Geometric meaning
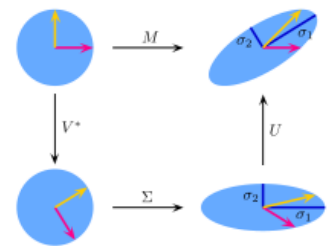
$$M = U \cdot \Sigma \cdot V^*$$

Upper left: The unit disc with the two canonical unit vectors.

Upper right: Unit disc transformed with M and singular values σ_1 and σ_2 indicated.

Lower left: The action of $\mathbf{V}^*$ on the unit disc. This is just a rotation.

Lower right: The action of $\boldsymbol{\Sigma}\mathbf{V}^*$ on the unit disc. Σ scales in vertically and horizontally.

In this special case, the singular values are φ and $1/\varphi$ where φ is the golden ratio. $\mathbf{V}^*$ is a (counter clockwise) rotation by an angle alpha where alpha satisfies $\tan(\alpha) = -\varphi$. $\mathbf{U}$ is a rotation by an angle beta with $\tan(\beta) = \varphi - 1$

Because $\mathbf{U}$ and $\mathbf{V}$ are unitary, we know that the columns $\mathbf{U}_1$, ..., $\mathbf{U}_m$ of $\mathbf{U}$ yield an orthonormal basis of K^m and the columns $\mathbf{V}_1$, ..., $\mathbf{V}_n$ of $\mathbf{V}$ yield an orthonormal basis of K^n (with respect to the standard scalar products on these spaces).

The linear transformation

$$\begin{cases} T : K^n \to K^m \\ x \mapsto \mathbf{M}x \end{cases}$$

has a particularly simple description with respect to these orthonormal bases: we have

$$T(\mathbf{V}_i) = \sigma_i \mathbf{U}_i, \qquad i = 1, \ldots, \min(m, n),$$

where σ_i is the i-th diagonal entry of $\boldsymbol{\Sigma}$, and $T(\mathbf{V}_i) = 0$ for $i > \min(m,n)$.

The geometric content of the SVD theorem can thus be summarized as follows: for every linear map $T : K^n \to K^m$ one can find orthonormal bases of K^n and K^m such that T maps the i-th basis vector of K^n to a non-negative multiple of the i-th basis vector of K^m, and sends the left-over basis vectors to zero. With respect to these bases, the map T is therefore represented by a diagonal matrix with non-negative real diagonal entries.

To get a more visual flavour of singular values and SVD factorization — at least when working on real vector spaces — consider the sphere S of radius one in $\mathbf{R}^n$. The linear map T maps this sphere onto an ellipsoid in $\mathbf{R}^m$. Non-zero singular values are simply the lengths of the semi-axes of this ellipsoid. Especially when $n = m$, and all the singular values are distinct and non-zero, the SVD of the linear map T can be easily analysed as a succession of three consecutive moves: consider the ellipsoid $T(S)$ and specifically its axes; then consider the directions in $\mathbf{R}^n$ sent by T onto these axes. These directions happen to be mutually orthogonal. Apply first an isometry $\mathbf{V}^*$ sending these directions to the coordinate axes of $\mathbf{R}^n$. On a second move, apply an endomorphism $\mathbf{D}$ diagonalized along the coordinate axes and stretching or shrinking in each direction, using the semi-axes lengths of $T(S)$ as stretching coefficients. The composition $\mathbf{D} \circ \mathbf{V}^*$ then sends the unit-sphere onto an ellipsoid isometric to $T(S)$. To define the third and last move $\mathbf{U}$, apply an isometry to this ellipsoid so as to carry it over $T(S)$. As can be easily checked, the composition $\mathbf{U} \circ \mathbf{D} \circ \mathbf{V}^*$ coincides with T.

Example

Consider the 4×5 matrix

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix}$$

A singular value decomposition of this matrix is given by $\mathbf{U\Sigma V}^*$

$$\mathbf{U} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{V}^* = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ -\sqrt{0.8} & 0 & 0 & 0 & \sqrt{0.2} \end{bmatrix}$$

The scaling matrix $\boldsymbol{\Sigma}$ is zero outside of the diagonal (grey italics) and one diagonal element is zero (red bold). Furthermore, because the matrices $\mathbf{U}$ and $\mathbf{V}^*$ are unitary, multiplying by their respective conjugate transposes yields identity matrices, as shown below. In this case, because $\mathbf{U}$ and $\mathbf{V}^*$ are real valued, each is an orthogonal matrix.

$$\mathbf{UU^*} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{I_4}$$

$$\mathbf{VV^*} = \begin{bmatrix} 0 & 0 & \sqrt{0.2} & 0 & -\sqrt{0.8} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \sqrt{0.8} & 0 & \sqrt{0.2} \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ -\sqrt{0.8} & 0 & 0 & 0 & \sqrt{0.2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{I_5}$$

This particular singular value decomposition is not unique. Choosing V such that

$$\mathbf{V^*} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ \sqrt{0.4} & 0 & 0 & \sqrt{0.5} & -\sqrt{0.1} \\ -\sqrt{0.4} & 0 & 0 & \sqrt{0.5} & \sqrt{0.1} \end{bmatrix}$$

is also a valid singular value decomposition.

SVD and spectral decomposition

Singular values, singular vectors, and their relation to the SVD

A non-negative real number σ is a **singular value** for $\mathbf{M}$ if and only if there exist unit-length vectors $\vec{u}$ in K^m and $\vec{v}$ in K^n such that

$$\mathbf{M}\vec{v} = \sigma\vec{u} \ \text{ and } \mathbf{M^*}\vec{u} = \sigma\vec{v}$$

The vectors $\vec{u}$ and $\vec{v}$ are called **left-singular** and **right-singular vectors** for σ, respectively.

In any singular value decomposition

$$\mathbf{M} = \mathbf{U\Sigma V^*}$$

the diagonal entries of $\mathbf{\Sigma}$ are equal to the singular values of $\mathbf{M}$. The first $p = \min(m, n)$ columns of $\mathbf{U}$ and $\mathbf{V}$ are, respectively, left- and right-singular vectors for the corresponding singular values. Consequently, the above theorem implies that:

- An $m \times n$ matrix $\mathbf{M}$ has at most p distinct singular values.
- It is always possible to find a unitary basis $\mathbf{U}$ for K^m with a subset of basis vectors spanning the left-singular vectors of each singular value of $\mathbf{M}$.
- It is always possible to find a unitary basis $\mathbf{V}$ for K^n with a subset of basis vectors spanning the right-singular vectors of each singular value of $\mathbf{M}$.

A singular value for which we can find two left (or right) singular vectors that are linearly independent is called *degenerate*. If $\vec{u_1}$ and $\vec{u_2}$ are two left-singular vectors which both correspond to the singular value σ, then any normalized linear combination of the two vectors is also a left-singular vector corresponding to the singular value σ. The similar statement is true for right-singular vectors. The number of independent left and right-singular vectors coincides, and these singular vectors appear in the same columns of $\mathbf{U}$ and $\mathbf{V}$ corresponding to diagonal elements of $\mathbf{\Sigma}$ all with the same value σ.

As an exception, the left and right-singular vectors of singular value 0 comprise all unit vectors in the kernel and cokernel, respectively, of $\mathbf{M}$, which by the rank–nullity theorem cannot be the same dimension if $m \neq n$. Even if all singular values are nonzero, if $m > n$ then the cokernel is nontrivial, in which case $\mathbf{U}$ is padded with $m - n$ orthogonal vectors from the cokernel. Conversely, if $m < n$, then $\mathbf{V}$ is padded by $n - m$ orthogonal vectors from the kernel. However, if the singular value of 0 exists, the extra columns of $\mathbf{U}$ or $\mathbf{V}$ already appear as left or right-singular vectors.

Non-degenerate singular values always have unique left- and right-singular vectors, up to multiplication by a unit-phase factor $e^{\mathbf{i}\varphi}$ (for the real case up to a sign). Consequently, if all singular values of a square matrix $\mathbf{M}$ are non-degenerate and non-zero, then its singular value decomposition is unique, up to multiplication of a column of $\mathbf{U}$ by a unit-phase factor and simultaneous multiplication of the corresponding column of $\mathbf{V}$ by the same unit-phase factor. In general, the SVD is unique up to arbitrary unitary transformations applied uniformly to the column vectors of both $\mathbf{U}$ and $\mathbf{V}$ spanning the subspaces of each singular value, and up to arbitrary unitary transformations on vectors of $\mathbf{U}$ and $\mathbf{V}$ spanning the kernel and cokernel, respectively, of $\mathbf{M}$.

Relation to eigenvalue decomposition

The singular value decomposition is very general in the sense that it can be applied to any $m \times n$ matrix whereas eigenvalue decomposition can only be applied to diagonalizable matrices. Nevertheless, the two decompositions are related.

Given an SVD of $\mathbf{M}$, as described above, the following two relations hold:

$$\mathbf{M}^*\mathbf{M} = \mathbf{V}\boldsymbol{\Sigma}^*\mathbf{U}^* \ \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^* = \mathbf{V}(\boldsymbol{\Sigma}^*\boldsymbol{\Sigma})\mathbf{V}^*$$
$$\mathbf{M}\mathbf{M}^* = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^* \ \mathbf{V}\boldsymbol{\Sigma}^*\mathbf{U}^* = \mathbf{U}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*)\mathbf{U}^*$$

The right-hand sides of these relations describe the eigenvalue decompositions of the left-hand sides. Consequently:

- The columns of $\mathbf{V}$ (right-singular vectors) are eigenvectors of $\mathbf{M}^*\mathbf{M}$.
- The columns of $\mathbf{U}$ (left-singular vectors) are eigenvectors of $\mathbf{M}\mathbf{M}^*$.
- The non-zero elements of $\boldsymbol{\Sigma}$ (non-zero singular values) are the square roots of the non-zero eigenvalues of $\mathbf{M}^*\mathbf{M}$ or $\mathbf{M}\mathbf{M}^*$.

In the special case that $\mathbf{M}$ is a normal matrix, which by definition must be square, the spectral theorem says that it can be unitarily diagonalized using a basis of eigenvectors, so that it can be written $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ for a unitary matrix $\mathbf{U}$ and a diagonal matrix $\mathbf{D}$. When $\mathbf{M}$ is also positive semi-definite, the decomposition $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ is also a singular value decomposition. Otherwise, it can be recast as an SVD by moving the phase of each σ_i to either its corresponding $\mathbf{V}_i$ or $\mathbf{U}_i$. The natural connection of the SVD to non-normal matrices is through the polar decomposition theorem: $\mathbf{M} = \mathbf{S}\mathbf{R}$, where $\mathbf{S} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^*$ is positive semidefinite and normal, and $\mathbf{R} = \mathbf{U}\mathbf{V}^*$ is unitary.

Thus while related, the eigenvalue decomposition and SVD differ except for positive semi-definite normal matrices $\mathbf{M}$: the eigenvalue decomposition is $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1}$ where $\mathbf{U}$ is not necessarily unitary and $\mathbf{D}$ is not necessarily positive semi-definite, while the SVD is $\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$ where $\boldsymbol{\Sigma}$ is diagonal and positive semi-definite, and $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices that are not necessarily related except through the matrix $\mathbf{M}$. While only non-defective square matrices have an eigenvalue decomposition, any $m \times n$ matrix has a SVD.

Applications of the SVD

Pseudoinverse

The singular value decomposition can be used for computing the pseudoinverse of a matrix. Indeed, the pseudoinverse of the matrix $\mathbf{M}$ with singular value decomposition $\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$ is

$$\mathbf{M}^+ = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^*$$

where $\boldsymbol{\Sigma}^+$ is the pseudoinverse of $\boldsymbol{\Sigma}$, which is formed by replacing every non-zero diagonal entry by its reciprocal and transposing the resulting matrix. The pseudoinverse is one way to solve linear least squares problems.

Solving homogeneous linear equations

A set of homogeneous linear equations can be written as $\mathbf{A}\mathbf{x} = \mathbf{0}$ for a matrix $\mathbf{A}$ and vector $\mathbf{x}$. A typical situation is that $\mathbf{A}$ is known and a non-zero $\mathbf{x}$ is to be determined which satisfies the equation. Such an $\mathbf{x}$ belongs to $\mathbf{A}$'s null space and is sometimes called a (right) null vector of $\mathbf{A}$. The vector $\mathbf{x}$ can be characterized as a right-singular vector corresponding to a singular value of $\mathbf{A}$ that is zero. This observation means that if $\mathbf{A}$ is a square matrix and has no vanishing singular value, the equation has no non-zero $\mathbf{x}$ as a solution. It also means that if there are several vanishing singular values, any linear combination of the corresponding right-singular vectors is a valid solution. Analogously to the definition of a (right) null vector, a non-zero $\mathbf{x}$ satisfying $\mathbf{x}^*\mathbf{A} = \mathbf{0}$, with $\mathbf{x}^*$ denoting the conjugate transpose of $\mathbf{x}$, is called a left null vector of $\mathbf{A}$.

Total least squares minimization

A total least squares problem refers to determining the vector $\mathbf{x}$ which minimizes the 2-norm of a vector $\mathbf{A}\mathbf{x}$ under the constraint $\|\mathbf{x}\| = 1$. The solution turns out to be the right-singular vector of $\mathbf{A}$ corresponding to the smallest singular value.

Range, null space and rank

Another application of the SVD is that it provides an explicit representation of the range and null space of a matrix $\mathbf{M}$. The right-singular vectors corresponding to vanishing singular values of $\mathbf{M}$ span the null space of $\mathbf{M}$ and the left-singular vectors corresponding to the non-zero singular values of $\mathbf{M}$ span the range of $\mathbf{M}$. E.g., in the above example the null space is spanned by the last two columns of $\mathbf{V}$ and the range is spanned by the first three columns of $\mathbf{U}$.

As a consequence, the rank of $\mathbf{M}$ equals the number of non-zero singular values which is the same as the number of non-zero diagonal elements in $\boldsymbol{\Sigma}$. In numerical linear algebra the singular values can be used to determine the *effective rank* of a matrix, as rounding error may lead to small but non-zero singular values in a rank deficient matrix. Singular values beyond a significant gap are assumed to be numerically equivalent to zero.

Low-rank matrix approximation

Some practical applications need to solve the problem of approximating a matrix $\mathbf{M}$ with another matrix $\tilde{\mathbf{M}}$, said truncated, which has a specific rank r. In the case that the approximation is based on minimizing the Frobenius norm of the difference between $\mathbf{M}$ and $\tilde{\mathbf{M}}$ under the constraint that $\mathbf{rank}(\tilde{\mathbf{M}}) = r$ it turns out that the solution is given by the SVD of $\mathbf{M}$, namely

$$\tilde{\mathbf{M}} = \mathbf{U}\tilde{\boldsymbol{\Sigma}}\mathbf{V}^*$$

where $\tilde{\Sigma}$ is the same matrix as Σ except that it contains only the r largest singular values (the other singular values are replaced by zero). This is known as the **Eckart–Young theorem**, as it was proved by those two authors in 1936 (although it was later found to have been known to earlier authors; see Stewart 1993).

Separable models

The SVD can be thought of as decomposing a matrix into a weighted, ordered sum of separable matrices. By separable, we mean that a matrix $\mathbf{A}$ can be written as an outer product of two vectors $\mathbf{A} = \mathbf{u} \otimes \mathbf{v}$, or, in coordinates, $A_{ij} = u_i v_j$. Specifically, the matrix $\mathbf{M}$ can be decomposed as:

$$\mathbf{M} = \sum_i \mathbf{A}_i = \sum_i \sigma_i \mathbf{U}_i \otimes \mathbf{V}_i$$

Here $\mathbf{U}_i$ and $\mathbf{V}_i$ are the i-th columns of the corresponding SVD matrices, σ_i are the ordered singular values, and each $\mathbf{A}_i$ is separable. The SVD can be used to find the decomposition of an image processing filter into separable horizontal and vertical filters. Note that the number of non-zero σ_i is exactly the rank of the matrix.

Separable models often arise in biological systems, and the SVD factorization is useful to analyze such systems. For example, some visual area V1 simple cells' receptive fields can be well described[1] by a Gabor filter in the space domain multiplied by a modulation function in the time domain. Thus, given a linear filter evaluated through, for example, reverse correlation, one can rearrange the two spatial dimensions into one dimension, thus yielding a two-dimensional filter (space, time) which can be decomposed through SVD. The first column of $\mathbf{U}$ in the SVD factorization is then a Gabor while the first column of $\mathbf{V}$ represents the time modulation (or vice versa). One may then define an index of separability,

$$\alpha = \frac{\sigma_1^2}{\sum_i \sigma_i^2},$$

which is the fraction of the power in the matrix M which is accounted for by the first separable matrix in the decomposition.[2]

Nearest orthogonal matrix

It is possible to use the SVD of a square matrix $\mathbf{A}$ to determine the orthogonal matrix $\mathbf{O}$ closest to $\mathbf{A}$. The closeness of fit is measured by the Frobenius norm of $\mathbf{O} - \mathbf{A}$. The solution is the product $\mathbf{U}\mathbf{V}^*$.[3] This intuitively makes sense because an orthogonal matrix would have the decomposition $\mathbf{U}\mathbf{I}\mathbf{V}^*$ where $\mathbf{I}$ is the identity matrix, so that if $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^*$ then the product $\mathbf{A} = \mathbf{U}\mathbf{V}^*$ amounts to replacing the singular values with ones. Equivalently, the solution is the unitary matrix $\mathbf{R} = \mathbf{U}\mathbf{V}^*$ of the Polar Decomposition $\mathbf{M} = \mathbf{R}\mathbf{P} = \mathbf{P}'\mathbf{R}$ in either order of stretch and rotation, as described above.

A similar problem, with interesting applications in shape analysis, is the orthogonal Procrustes problem, which consists of finding an orthogonal matrix $\mathbf{O}$ which most closely maps $\mathbf{A}$ to $\mathbf{B}$. Specifically,

$$\mathbf{O} = \underset{\Omega}{\operatorname{argmin}} \|\mathbf{A}\Omega - \mathbf{B}\|_F \quad \text{subject to} \quad \Omega^\mathsf{T}\Omega = \mathbf{I}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

This problem is equivalent to finding the nearest orthogonal matrix to a given matrix $\mathbf{M} = \mathbf{A}^\mathsf{T}\mathbf{B}$.

The Kabsch algorithm

The Kabsch algorithm (called Wahba's problem in other fields) uses SVD to compute the optimal rotation (with respect to least-squares minimization) that will align a set of points with a corresponding set of points. It is used, among other applications, to compare the structures of molecules.

Signal processing

The SVD and pseudoinverse have been successfully applied to signal processing,[4] image processing[5] and big data (e.g., in genomic signal processing).[6][7][8][9]

Other examples

The SVD is also applied extensively to the study of linear inverse problems, and is useful in the analysis of regularization methods such as that of Tikhonov. It is widely used in statistics where it is related to principal component analysis and to Correspondence analysis, and in signal processing and pattern recognition. It is also used in output-only modal analysis, where the non-scaled mode shapes can be determined from the singular vectors. Yet another usage is latent semantic indexing in natural language text processing.

The SVD also plays a crucial role in the field of quantum information, in a form often referred to as the Schmidt decomposition. Through it, states of two quantum systems are naturally decomposed, providing a necessary and sufficient condition for them to be entangled: if the rank of the Σ matrix is larger than one.

One application of SVD to rather large matrices is in numerical weather prediction, where Lanczos methods are used to estimate the most linearly quickly growing few perturbations to the central numerical weather prediction over a given initial forward time period; i.e., the singular vectors corresponding to the largest singular values of the linearized propagator for the global weather over that time interval. The output singular vectors in this case are entire weather systems. These perturbations are then run through the full nonlinear model to generate an ensemble forecast, giving a handle on some of the uncertainty that should be allowed for around the current central prediction.

SVD has also been applied to reduced order modelling. The aim of reduced order modelling is to reduce the number of degrees of freedom in a complex system which is to be modelled. SVD was coupled with radial basis functions to interpolate solutions to three-dimensional unsteady flow problems.[10]

Interestingly, SVD has been used to improve gravitational waveform modeling by the ground based gravitational wave interferometer aLIGO.[11] SVD can help to increase the accuracy and speed of waveform generation to support gravitational waves searches and update two different waveform models.

Singular value decomposition is used in recommender systems to predict people's item ratings.[12] Distributed algorithms have been developed for the purpose of calculating the SVD on clusters of commodity machines.[13]

Another code implementation of the Netflix Recommendation Algorithm SVD (the third optimal algorithm in the competition conducted by Netflix to find the best collaborative filtering techniques for predicting user ratings for films based on previous reviews) in platform Apache Spark is available in the following GitHub repository [14] implemented by Alexandros Ioannidis. The original SVD algorithm,[15] which in this case is executed in parallel encourages users of the GroupLens website, by consulting proposals for monitoring new films tailored to the needs of each user.

Low-rank SVD has been applied for hotspot detection from spatiotemporal data with application to disease outbreak detection .[16] A combination of SVD and higher-order SVD also has been applied for real time event detection from complex data streams (multivariate data with space and time dimensions) in Disease surveillance.[17]

Existence proofs

An eigenvalue λ of a matrix $\mathbf{M}$ is characterized by the algebraic relation $\mathbf{M}u = \lambda u$. When $\mathbf{M}$ is Hermitian, a variational characterization is also available. Let $\mathbf{M}$ be a real $n \times n$ symmetric matrix. Define

$$\begin{cases} f : \mathbb{R}^n \to \mathbb{R} \\ f(x) = x^{\mathsf{T}}\mathbf{M}x \end{cases}$$

By the extreme value theorem, this continuous function attains a maximum at some u when restricted to the closed unit sphere $\{\|x\| \le 1\}$. By the Lagrange multipliers theorem, u necessarily satisfies

$$\nabla f = \nabla x^{\mathsf{T}}\mathbf{M}x - \lambda \cdot \nabla x^{\mathsf{T}}x = 2(\mathbf{M} - \lambda\mathbf{I})x = 0$$

where the nabla symbol, ∇, is the del operator (differentiation with respect to x).

A short calculation shows the above leads to $\mathbf{M}u = \lambda u$ (symmetry of $\mathbf{M}$ is needed here). Therefore, λ is the largest eigenvalue of $\mathbf{M}$. The same calculation performed on the orthogonal complement of u gives the next largest eigenvalue and so on. The complex Hermitian case is similar; there $f(x) = x^* M x$ is a real-valued function of $2n$ real variables.

Singular values are similar in that they can be described algebraically or from variational principles. Although, unlike the eigenvalue case, Hermiticity, or symmetry, of $\mathbf{M}$ is no longer required.

This section gives these two arguments for existence of singular value decomposition.

Based on the spectral theorem

Let $\mathbf{M}$ be an $m \times n$ complex matrix. Since $\mathbf{M}^*\mathbf{M}$ is positive semi-definite and Hermitian, by the spectral theorem, there exists an $n \times n$ unitary matrix $\mathbf{V}$ such that

$$\mathbf{V}^*\mathbf{M}^*\mathbf{M}\mathbf{V} = \bar{\mathbf{D}} = \begin{bmatrix} \mathbf{D} & 0 \\ 0 & 0 \end{bmatrix},$$

where $\mathbf{D}$ is diagonal and positive definite, of dimension $\ell \times \ell$, with ℓ the number of non-zero eigenvalues of $\mathbf{M}^*\mathbf{M}$ (which can be shown to verify $\ell \le \min(n, m)$). Note that $\mathbf{V}$ is here by definition a matrix whose i-th column is the i-th eigenvector of $\mathbf{M}^*\mathbf{M}$, corresponding to the eigenvalue $\bar{\mathbf{D}}_{ii}$. Moreover, the j-th column of $\mathbf{V}$, for $j > \ell$, is an eigenvector of $\mathbf{M}^*\mathbf{M}$ with eigenvalue $\bar{\mathbf{D}}_{jj} = 0$. This can be expressed by writing $\mathbf{V}$ as $\mathbf{V} = [\mathbf{V}_1 \quad \mathbf{V}_2]$, where the columns of $\mathbf{V}_1$ and $\mathbf{V}_2$ therefore contain the eigenvectors of $\mathbf{M}^*\mathbf{M}$ corresponding to non-zero and zero eigenvalues, respectively. Using this rewriting of $\mathbf{V}$, the equation becomes:

$$\begin{bmatrix} \mathbf{V}_1^* \\ \mathbf{V}_2^* \end{bmatrix} \mathbf{M}^*\mathbf{M}[\mathbf{V}_1 \quad \mathbf{V}_2] = \begin{bmatrix} \mathbf{V}_1^*\mathbf{M}^*\mathbf{M}\mathbf{V}_1 & \mathbf{V}_1^*\mathbf{M}^*\mathbf{M}\mathbf{V}_2 \\ \mathbf{V}_2^*\mathbf{M}^*\mathbf{M}\mathbf{V}_1 & \mathbf{V}_2^*\mathbf{M}^*\mathbf{M}\mathbf{V}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{D} & 0 \\ 0 & 0 \end{bmatrix}.$$

This implies that

$$\mathbf{V}_1^*\mathbf{M}^*\mathbf{M}\mathbf{V}_1 = \mathbf{D}, \qquad \mathbf{V}_2^*\mathbf{M}^*\mathbf{M}\mathbf{V}_2 = 0.$$

Moreover, the second equation implies $\mathbf{M}\mathbf{V}_2 = 0$.[18] Finally, the unitarity of $\mathbf{V}$ translates, in terms of $\mathbf{V}_1$ and $\mathbf{V}_2$, into the following conditions:

$$\begin{aligned} \mathbf{V}_1^*\mathbf{V}_1 &= \mathbf{I}_1, \\ \mathbf{V}_2^*\mathbf{V}_2 &= \mathbf{I}_2, \\ \mathbf{V}_1\mathbf{V}_1^* + \mathbf{V}_2\mathbf{V}_2^* &= \mathbf{I}_{12}, \end{aligned}$$

where the subscripts on the identity matrices are used to remark that they are of different dimensions.

Let us now define

$$\mathbf{U}_1 = \mathbf{M}\mathbf{V}_1\mathbf{D}^{-\frac{1}{2}}.$$

Then,

$$\mathbf{U}_1\mathbf{D}^{\frac{1}{2}}\mathbf{V}_1^* = \mathbf{M}\mathbf{V}_1\mathbf{D}^{-\frac{1}{2}}\mathbf{D}^{\frac{1}{2}}\mathbf{V}_1^* = \mathbf{M}(\mathbf{I} - \mathbf{V}_2\mathbf{V}_2^*) = \mathbf{M} - (\mathbf{M}\mathbf{V}_2)\mathbf{V}_2^* = \mathbf{M},$$

since $\mathbf{M}\mathbf{V}_2 = \mathbf{0}$. This can be also seen as immediate consequence of the fact that $\mathbf{M}\mathbf{V}_1\mathbf{V}_1^* = \mathbf{M}$. Note how this is equivalent to the observation that, if $\{v_i\}_{i=1}^l$ is the set of eigenvectors of $\mathbf{M}^*\mathbf{M}$ corresponding to non-vanishing eigenvalues, then $\{\mathbf{M}v_i\}_{i=1}^l$ is a set of orthogonal vectors, and $\{\lambda^{-1/2}\mathbf{M}v_i\}_{i=1}^l$ a (generally not complete) set of *orthonormal* vectors. This matches with the matrix formalism used above denoting with $\mathbf{V}_1$ the matrix whose columns are $\{v_i\}_{i=1}^l$, with $\mathbf{V}_2$ the matrix whose columns are the eigenvectors of $\mathbf{M}^*\mathbf{M}$ which vanishing eigenvalue, and $\mathbf{U}_1$ the matrix whose columns are the vectors $\{\lambda^{-1/2}\mathbf{M}v_i\}_{i=1}^l$.

We see that this is almost the desired result, except that $\mathbf{U}_1$ and $\mathbf{V}_1$ are in general not unitary, since they might not be square. However, we do know that the number of rows of $\mathbf{U}_1$ is no smaller than the number of columns, since the dimensions of $\mathbf{D}$ is no greater than m and n. Also, since

$$\mathbf{U}_1^*\mathbf{U}_1 = \mathbf{D}^{-\frac{1}{2}}\mathbf{V}_1^*\mathbf{M}^*\mathbf{M}\mathbf{V}_1\mathbf{D}^{-\frac{1}{2}} = \mathbf{D}^{-\frac{1}{2}}\mathbf{D}\mathbf{D}^{-\frac{1}{2}} = \mathbf{I}_1$$

the columns in $\mathbf{U}_1$ are orthonormal and can be extended to an orthonormal basis. This means that we can choose $\mathbf{U}_2$ such that $\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix}$ is unitary.

For $\mathbf{V}_1$ we already have $\mathbf{V}_2$ to make it unitary. Now, define

$$\mathbf{\Sigma} = \begin{bmatrix} \begin{bmatrix} \mathbf{D}^{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix} \\ 0 \end{bmatrix}$$

where extra zero rows are added **or removed** to make the number of zero rows equal the number of columns of $\mathbf{U}_2$, and hence the overall dimensions of $\mathbf{\Sigma}$ equal to $m \times n$. Then

$$\begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} D^{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix} \\ 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix}^* = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{D}^{\frac{1}{2}}\mathbf{V}_1^* \\ 0 \end{bmatrix} = \mathbf{U}_1\mathbf{D}^{\frac{1}{2}}\mathbf{V}_1^* = \mathbf{M}$$

which is the desired result:

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$$

Notice the argument could begin with diagonalizing $\mathbf{M}\mathbf{M}^*$ rather than $\mathbf{M}^*\mathbf{M}$ (This shows directly that $\mathbf{M}\mathbf{M}^*$ and $\mathbf{M}^*\mathbf{M}$ have the same non-zero eigenvalues).

Based on variational characterization

The singular values can also be characterized as the maxima of $\mathbf{u}^\top\mathbf{M}\mathbf{v}$, considered as a function of $\mathbf{u}$ and $\mathbf{v}$, over particular subspaces. The singular vectors are the values of $\mathbf{u}$ and $\mathbf{v}$ where these maxima are attained.

Let $\mathbf{M}$ denote an $m \times n$ matrix with real entries. Let S^{k-1} be the unit $(k-1)$-sphere in $\mathbb{R}^k$, and define $\sigma(\mathbf{u}, \mathbf{v}) = \mathbf{u}^\top\mathbf{M}\mathbf{v}$, $\quad \mathbf{u} \in S^{m-1}, \mathbf{v} \in S^{n-1}$.

Consider the function σ restricted to $S^{m-1} \times S^{n-1}$. Since both S^{m-1} and S^{n-1} are compact sets, their product is also compact. Furthermore, since σ is continuous, it attains a largest value for at least one pair of vectors $\mathbf{u} \in S^{m-1}$ and $\mathbf{v} \in S^{n-1}$. This largest value is denoted σ_1 and the corresponding vectors are denoted $\mathbf{u}_1$ and $\mathbf{v}_1$. Since σ_1 is the largest value of $\sigma(\mathbf{u}, \mathbf{v})$ it must be non-negative. If it were negative, changing the sign of either $\mathbf{u}_1$ or $\mathbf{v}_1$ would make it positive and therefore larger.

Statement. $\mathbf{u}_1$, $\mathbf{v}_1$ are left and right-singular vectors of $\mathbf{M}$ with corresponding singular value σ_1.

Proof. Similar to the eigenvalues case, by assumption the two vectors satisfy the Lagrange multiplier equation:

$$\nabla\sigma = \nabla\mathbf{u}^\top\mathbf{M}\mathbf{v} - \lambda_1 \cdot \nabla\mathbf{u}^\top\mathbf{u} - \lambda_2 \cdot \nabla\mathbf{v}^\top\mathbf{v}$$

After some algebra, this becomes

$$\mathbf{M}\mathbf{v}_1 = 2\lambda_1\mathbf{u}_1 + 0$$
$$\mathbf{M}^\top\mathbf{u}_1 = 0 + 2\lambda_2\mathbf{v}_1$$

Multiplying the first equation from left by $\mathbf{u}_1^\top$ and the second equation from left by $\mathbf{v}_1^\top$ and taking $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ into account gives

$$\sigma_1 = 2\lambda_1 = 2\lambda_2.$$

Plugging this into the pair of equations above, we have

$$\mathbf{M}\mathbf{v_1} = \sigma_1 \mathbf{u_1}$$
$$\mathbf{M}^\mathsf{T}\mathbf{u_1} = \sigma_1 \mathbf{v_1}$$

This proves the statement.

More singular vectors and singular values can be found by maximizing $\sigma(\mathbf{u}, \mathbf{v})$ over normalized $\mathbf{u}$, $\mathbf{v}$ which are orthogonal to $\mathbf{u}_1$ and $\mathbf{v}_1$, respectively.

The passage from real to complex is similar to the eigenvalue case.

Calculating the SVD

Numerical approach

The SVD of a matrix $\mathbf{M}$ is typically computed by a two-step procedure. In the first step, the matrix is reduced to a bidiagonal matrix. This takes $O(mn^2)$ floating-point operations (flop), assuming that $m \geq n$. The second step is to compute the SVD of the bidiagonal matrix. This step can only be done with an iterative method (as with eigenvalue algorithms). However, in practice it suffices to compute the SVD up to a certain precision, like the machine epsilon. If this precision is considered constant, then the second step takes $O(n)$ iterations, each costing $O(n)$ flops. Thus, the first step is more expensive, and the overall cost is $O(mn^2)$ flops (Trefethen & Bau III 1997, Lecture 31).

The first step can be done using Householder reflections for a cost of $4mn^2 - 4n^3/3$ flops, assuming that only the singular values are needed and not the singular vectors. If m is much larger than n then it is advantageous to first reduce the matrix M to a triangular matrix with the QR decomposition and then use Householder reflections to further reduce the matrix to bidiagonal form; the combined cost is $2mn^2 + 2n^3$ flops (Trefethen & Bau III 1997, Lecture 31).

The second step can be done by a variant of the QR algorithm for the computation of eigenvalues, which was first described by Golub & Kahan (1965). The LAPACK subroutine DBDSQR[19] implements this iterative method, with some modifications to cover the case where the singular values are very small (Demmel & Kahan 1990). Together with a first step using Householder reflections and, if appropriate, QR decomposition, this forms the DGESVD[20] routine for the computation of the singular value decomposition.

The same algorithm is implemented in the GNU Scientific Library (GSL). The GSL also offers an alternative method that uses a one-sided Jacobi orthogonalization in step 2 (GSL Team 2007). This method computes the SVD of the bidiagonal matrix by solving a sequence of 2 × 2 SVD problems, similar to how the Jacobi eigenvalue algorithm solves a sequence of 2 × 2 eigenvalue methods (Golub & Van Loan 1996, §8.6.3). Yet another method for step 2 uses the idea of divide-and-conquer eigenvalue algorithms (Trefethen & Bau III 1997, Lecture 31).

There is an alternative way that does not explicitly use the eigenvalue decomposition.[21] Usually the singular value problem of a matrix $\mathbf{M}$ is converted into an equivalent symmetric eigenvalue problem such as $\mathbf{M}\,\mathbf{M}^*$, $\mathbf{M}^*\mathbf{M}$, or

$$\begin{pmatrix} \mathbf{O} & \mathbf{M} \\ \mathbf{M}^* & \mathbf{O} \end{pmatrix}.$$

The approaches that use eigenvalue decompositions are based on the QR algorithm, which is well-developed to be stable and fast. Note that the singular values are real and right- and left- singular vectors are not required to form similarity transformations. One can iteratively alternate between the QR decomposition and the LQ decomposition to find the real diagonal Hermitian matrices. The QR decomposition gives $\mathbf{M} \Rightarrow \mathbf{Q}\,\mathbf{R}$ and the LQ decomposition of $\mathbf{R}$ gives $\mathbf{R} \Rightarrow \mathbf{L}\,\mathbf{P}^*$. Thus, at every iteration, we have $\mathbf{M} \Rightarrow \mathbf{Q}\,\mathbf{L}\,\mathbf{P}^*$, update $\mathbf{M} \Leftarrow \mathbf{L}$ and repeat the orthogonalizations. Eventually, this iteration between QR decomposition and LQ decomposition produces left- and right- unitary singular matrices. This approach cannot readily be accelerated, as the QR algorithm can with spectral shifts or deflation. This is because the shift method is not easily defined without using similarity transformations. However, this iterative approach is very simple to implement, so is a good choice when speed does not matter. This method also provides insight into how purely orthogonal/unitary transformations can obtain the SVD.

Analytic result of 2 × 2 SVD

The singular values of a 2 × 2 matrix can be found analytically. Let the matrix be $\mathbf{M} = z_0\mathbf{I} + z_1\sigma_1 + z_2\sigma_2 + z_3\sigma_3$

where $z_i \in \mathbb{C}$ are complex numbers that parameterize the matrix, $\mathbf{I}$ is the identity matrix, and σ_i denote the Pauli matrices. Then its two singular values are given by

$$\begin{aligned} \sigma_\pm &= \sqrt{|z_0|^2 + |z_1|^2 + |z_2|^2 + |z_3|^2 \pm \sqrt{(|z_0|^2 + |z_1|^2 + |z_2|^2 + |z_3|^2)^2 - |z_0^2 - z_1^2 - z_2^2 - z_3^2|^2}} \\ &= \sqrt{|z_0|^2 + |z_1|^2 + |z_2|^2 + |z_3|^2 \pm 2\sqrt{(\operatorname{Re} z_0 z_1^*)^2 + (\operatorname{Re} z_0 z_2^*)^2 + (\operatorname{Re} z_0 z_3^*)^2 + (\operatorname{Im} z_1 z_2^*)^2 + (\operatorname{Im} z_2 z_3^*)^2 + (\operatorname{Im} z_3 z_1^*)^2}} \end{aligned}$$

Reduced SVDs

In applications it is quite unusual for the full SVD, including a full unitary decomposition of the null-space of the matrix, to be required. Instead, it is often sufficient (as well as faster, and more economical for storage) to compute a reduced version of the SVD. The following can be distinguished for an $m \times n$ matrix M of rank r:

Thin SVD

$$\mathbf{M} = \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{V}^*$$

Only the *n* column vectors of *U* corresponding to the row vectors of *V** are calculated. The remaining column vectors of *U* are not calculated. This is significantly quicker and more economical than the full SVD if $n \ll m$. The matrix U_n is thus $m \times n$, Σ_n is $n \times n$ diagonal, and *V* is $n \times n$.

The first stage in the calculation of a thin SVD will usually be a QR decomposition of *M*, which can make for a significantly quicker calculation if $n \ll m$.

Compact SVD

$$\mathbf{M} = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^*$$

Only the *r* column vectors of *U* and *r* row vectors of *V** corresponding to the non-zero singular values Σ_r are calculated. The remaining vectors of *U* and *V** are not calculated. This is quicker and more economical than the thin SVD if $r \ll n$. The matrix U_r is thus $m \times r$, Σ_r is $r \times r$ diagonal, and V_r^* is $r \times n$.

Truncated SVD

$$\tilde{\mathbf{M}} = \mathbf{U}_t \mathbf{\Sigma}_t \mathbf{V}_t^*$$

Only the *t* column vectors of *U* and *t* row vectors of *V** corresponding to the *t* largest singular values Σ_t are calculated. The rest of the matrix is discarded. This can be much quicker and more economical than the compact SVD if $t \ll r$. The matrix U_t is thus $m \times t$, Σ_t is $t \times t$ diagonal, and V_t^* is $t \times n$.

Of course the truncated SVD is no longer an exact decomposition of the original matrix *M*, but as discussed above, the approximate matrix $\tilde{\mathbf{M}}$ is in a very useful sense the closest approximation to *M* that can be achieved by a matrix of rank *t*.

Norms

Ky Fan norms

The sum of the *k* largest singular values of *M* is a matrix norm, the Ky Fan *k*-norm of *M*.[22]

The first of the Ky Fan norms, the Ky Fan 1-norm, is the same as the operator norm of *M* as a linear operator with respect to the Euclidean norms of K^m and K^n. In other words, the Ky Fan 1-norm is the operator norm induced by the standard ℓ^2 Euclidean inner product. For this reason, it is also called the operator 2-norm. One can easily verify the relationship between the Ky Fan 1-norm and singular values. It is true in general, for a bounded operator *M* on (possibly infinite-dimensional) Hilbert spaces

$$\|\mathbf{M}\| = \|\mathbf{M}^*\mathbf{M}\|^{\frac{1}{2}}$$

But, in the matrix case, $(M^* M)^{\frac{1}{2}}$ is a normal matrix, so $\|M^* M\|^{\frac{1}{2}}$ is the largest eigenvalue of $(M^* M)^{\frac{1}{2}}$, i.e. the largest singular value of *M*.

The last of the Ky Fan norms, the sum of all singular values, is the trace norm (also known as the 'nuclear norm'), defined by $\|M\| = \mathrm{Tr}[(M^* M)^{\frac{1}{2}}]$ (the eigenvalues of $M^* M$ are the squares of the singular values).

Hilbert–Schmidt norm

The singular values are related to another norm on the space of operators. Consider the Hilbert–Schmidt inner product on the $n \times n$ matrices, defined by

$$\langle \mathbf{M}, \mathbf{N} \rangle = \mathrm{trace}(\mathbf{N}^* \mathbf{M}).$$

So the induced norm is

$$\|\mathbf{M}\| = \sqrt{\langle \mathbf{M}, \mathbf{M} \rangle} = \sqrt{\mathrm{trace}(\mathbf{M}^* \mathbf{M})}.$$

Since the trace is invariant under unitary equivalence, this shows

$$\|\mathbf{M}\| = \sqrt{\sum_i \sigma_i^2}$$

where σ_i are the singular values of $\mathbf{M}$. This is called the **Frobenius norm**, **Schatten 2-norm**, or **Hilbert–Schmidt norm** of $\mathbf{M}$. Direct calculation shows that the Frobenius norm of $\mathbf{M} = (m_{ij})$ coincides with:

$$\sqrt{\sum_{ij} |m_{ij}|^2}.$$

In addition, the Frobenius norm and the trace norm (the nuclear norm) are special cases of the Schatten norm.

Variations and generalizations

Mode-*k* representation

$M = USV^\mathsf{T}$ can be represented using mode-*k* multiplication of matrix S applying $\times_1 U$ then $\times_2 V$ on the result; that is $M = S \times_1 U \times_2 V$.[23]

Tensor SVD

Two types of tensor decompositions exist, which generalise the SVD to multi-way arrays. One of them decomposes a tensor into a sum of rank-1 tensors, which is called a tensor rank decomposition. The second type of decomposition computes the orthonormal subspaces associated with the different factors appearing in the tensor product of vector spaces in which the tensor lives. This decomposition is referred to in the literature as the higher-order SVD (HOSVD) or Tucker3/TuckerM. In addition, multilinear principal component analysis in multilinear subspace learning involves the same mathematical operations as Tucker decomposition, being used in a different context of dimensionality reduction.

Scale-invariant SVD

The SVD singular values of a matrix *A* are unique and are invariant with respect to left and/or right unitary transformations of *A*. In other words, the singular values of *UAV*, for unitary *U* and *V*, are equal to the singular values of *A*. This is an important property for applications in which it is necessary to preserve Euclidean distances, and invariance with respect to rotations.

The Scale-Invariant SVD, or SI-SVD,[24] is analogous to the conventional SVD except that its uniquely-determined singular values are invariant with respect to diagonal transformations of *A*. In other words, the singular values of *DAE*, for nonsingular diagonal matrices *D* and *E*, are equal to the singular values of *A*. This is an important property for applications for which invariance to the choice of units on variables (e.g., metric versus imperial units) is needed.

HOSVD of functions – numerical reconstruction – TP model transformation

TP model transformation numerically reconstruct the HOSVD of functions. For further details please visit:

- HOSVD-based canonical form of TP functions and qLPV models
- Tensor product model transformation
- TP model transformation in control theory

Bounded operators on Hilbert spaces

The factorization $\mathbf{M} = \mathbf{U\Sigma V}^*$ can be extended to a bounded operator *M* on a separable Hilbert space *H*. Namely, for any bounded operator *M*, there exist a partial isometry *U*, a unitary *V*, a measure space (X, μ), and a non-negative measurable *f* such that

$$\mathbf{M} = \mathbf{U} T_f \mathbf{V}^*$$

where T_f is the multiplication by *f* on $L^2(X, \mu)$.

This can be shown by mimicking the linear algebraic argument for the matricial case above. $VT_f V^*$ is the unique positive square root of *M*M*, as given by the Borel functional calculus for self adjoint operators. The reason why *U* need not be unitary is because, unlike the finite-dimensional case, given an isometry U_1 with nontrivial kernel, a suitable U_2 may not be found such that

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$$

is a unitary operator.

As for matrices, the singular value factorization is equivalent to the polar decomposition for operators: we can simply write

$$\mathbf{M} = \mathbf{UV}^* \cdot \mathbf{V} T_f \mathbf{V}^*$$

and notice that *U V** is still a partial isometry while $VT_f V^*$ is positive.

Singular values and compact operators

The notion of singular values and left/right-singular vectors can be extended to compact operator on Hilbert space as they have a discrete spectrum. If *T* is compact, every non-zero λ in its spectrum is an eigenvalue. Furthermore, a compact self adjoint operator can be diagonalized by its eigenvectors. If $\mathbf{M}$ is compact, so is $\mathbf{M}^*\mathbf{M}$. Applying the diagonalization result, the unitary image of its positive square root T_f has a set of orthonormal eigenvectors $\{e_i\}$ corresponding to strictly positive eigenvalues $\{\sigma_i\}$. For any $\psi \in H$,

$$\mathbf{M}\psi = \mathbf{U} T_f \mathbf{V}^* \psi = \sum_i \langle \mathbf{U} T_f \mathbf{V}^* \psi, \mathbf{U} e_i \rangle \mathbf{U} e_i = \sum_i \sigma_i \langle \psi, \mathbf{V} e_i \rangle \mathbf{U} e_i$$

where the series converges in the norm topology on H. Notice how this resembles the expression from the finite-dimensional case. σ_i are called the singular values of $\mathbf{M}$. $\{\mathbf{U}e_i\}$ (resp. $\{\mathbf{V}e_i\}$) can be considered the left-singular (resp. right-singular) vectors of $\mathbf{M}$.

Compact operators on a Hilbert space are the closure of finite-rank operators in the uniform operator topology. The above series expression gives an explicit such representation. An immediate consequence of this is:

> **Theorem.** $\mathbf{M}$ is compact if and only if $\mathbf{M}^*\mathbf{M}$ is compact.

History

The singular value decomposition was originally developed by differential geometers, who wished to determine whether a real bilinear form could be made equal to another by independent orthogonal transformations of the two spaces it acts on. Eugenio Beltrami and Camille Jordan discovered independently, in 1873 and 1874 respectively, that the singular values of the bilinear forms, represented as a matrix, form a complete set of invariants for bilinear forms under orthogonal substitutions. James Joseph Sylvester also arrived at the singular value decomposition for real square matrices in 1889, apparently independently of both Beltrami and Jordan. Sylvester called the singular values the *canonical multipliers* of the matrix A. The fourth mathematician to discover the singular value decomposition independently is Autonne in 1915, who arrived at it via the polar decomposition. The first proof of the singular value decomposition for rectangular and complex matrices seems to be by Carl Eckart and Gale J. Young in 1936;[25] they saw it as a generalization of the principal axis transformation for Hermitian matrices.

In 1907, Erhard Schmidt defined an analog of singular values for integral operators (which are compact, under some weak technical assumptions); it seems he was unaware of the parallel work on singular values of finite matrices. This theory was further developed by Émile Picard in 1910, who is the first to call the numbers σ_k *singular values* (or in French, *valeurs singulières*).

Practical methods for computing the SVD date back to Kogbetliantz in 1954, 1955 and Hestenes in 1958.[26] resembling closely the Jacobi eigenvalue algorithm, which uses plane rotations or Givens rotations. However, these were replaced by the method of Gene Golub and William Kahan published in 1965,[27] which uses Householder transformations or reflections. In 1970, Golub and Christian Reinsch[28] published a variant of the Golub/Kahan algorithm that is still the one most-used today.

See also

- Canonical correlation
- Canonical form
- Correspondence analysis (CA)
- Curse of dimensionality
- Digital signal processing
- Dimensionality reduction
- Eigendecomposition of a matrix
- Empirical orthogonal functions (EOFs)
- Fourier analysis
- Generalized singular value decomposition
- Inequalities about singular values
- K-SVD
- Latent semantic analysis
- Latent semantic indexing
- Linear least squares
- List of Fourier-related transforms
- Locality-sensitive hashing
- Low-rank approximation
- Matrix decomposition
- Multilinear principal component analysis (MPCA)
- Nearest neighbor search
- Non-linear iterative partial least squares
- Polar decomposition
- Principal component analysis (PCA)
- Schmidt decomposition
- Singular value
- Time series
- Two-dimensional singular-value decomposition (2DSVD)
- von Neumann's trace inequality
- Wavelet compression

Notes

1. DeAngelis, G. C.; Ohzawa, I.; Freeman, R. D. (October 1995). "Receptive-field dynamics in the central visual pathways". *Trends Neurosci*. **18** (10): 451–8. doi:10.1016/0166-2236(95)94496-R (https://doi.org/10.1016%2F0166-2236%2895%2994496-R). PMID 8545912 (https://www.ncbi.nlm.nih.gov/pubmed/8545912).

2. Depireux, D. A.; Simon, J. Z.; Klein, D. J.; Shamma, S. A. (March 2001). "Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex". *J. Neurophysiol*. **85** (3): 1220–34. doi:10.1152/jn.2001.85.3.1220 (https://doi.org/10.1152%2Fjn.2001.85.3.1220). PMID 11247991 (https://www.ncbi.nlm.nih.gov/pubmed/11247991).

3. The Singular Value Decomposition in Symmetric (Lowdin) Orthogonalization and Data Compression (http://www.wou.edu/~beavers/Talks/Willamette1106.pdf)

4. Sahidullah, Md.; Kinnunen, Tomi (March 2016). "Local spectral variability features for speaker verification". *Digital Signal Processing*. **50**: 1–11. doi:10.1016/j.dsp.2015.10.011 (https://doi.org/10.1016%2Fj.dsp.2015.10.011).

5. Rowayda A. Sadek (2012). "SVD Based Image Processing Applications: State of The Art, Contributions and Research Challenges". *International Journal of Advanced Computer Science and Applications*. **3** (7): 26–34. arXiv:1211.7102 (https://arxiv.org/abs/1211.7102). Bibcode:2012arXiv1211.7102S (https://ui.adsabs.harvard.edu/abs/2012arXiv1211.7102S). doi:10.14569/IJACSA.2012.030703 (https://doi.org/10.14569%2FIJACSA.2012.030703).

6. O. Alter, P. O. Brown and D. Botstein (September 2000). "Singular Value Decomposition for Genome-Wide Expression Data Processing and Modeling" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC27718). *PNAS*. **97** (18): 10101–10106. Bibcode:2000PNAS...9710101A (https://ui.adsabs.harvard.edu/abs/2000PNAS...9710101A). doi:10.1073/pnas.97.18.10101 (https://doi.org/10.1073%2Fpnas.97.18.10101). PMC 27718 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC27718). PMID 10963673 (https://www.ncbi.nlm.nih.gov/pubmed/10963673).

7. O. Alter; G. H. Golub (November 2004). "Integrative Analysis of Genome-Scale Data by Using Pseudoinverse Projection Predicts Novel Correlation Between DNA Replication and RNA Transcription" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC534520). *PNAS*. **101** (47): 16577–16582. Bibcode:2004PNAS..10116577A (https://ui.adsabs.harvard.edu/abs/2004PNAS..10116577A). doi:10.1073/pnas.0406767101 (https://doi.org/10.1073%2Fpnas.0406767101). PMC 534520 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC534520). PMID 15545604 (https://www.ncbi.nlm.nih.gov/pubmed/15545604).

8. O. Alter; G. H. Golub (August 2006). "Singular Value Decomposition of Genome-Scale mRNA Lengths Distribution Reveals Asymmetry in RNA Gel Electrophoresis Band Broadening" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1524674). *PNAS*. **103** (32): 11828–11833. Bibcode:2006PNAS..10311828A (https://ui.adsabs.harvard.edu/abs/2006PNAS..10311828A). doi:10.1073/pnas.0604756103 (https://doi.org/10.1073%2Fpnas.0604756103). PMC 1524674 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1524674). PMID 16877539 (https://www.ncbi.nlm.nih.gov/pubmed/16877539).

9. Bertagnolli, N. M.; Drake, J. A.; Tennessen, J. M.; Alter, O. (November 2013). "SVD Identifies Transcript Length Distribution Functions from DNA Microarray Data and Reveals Evolutionary Forces Globally Affecting GBM Metabolism" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3839928). *PLOS ONE*. **8** (11): e78913. Bibcode:2013PLoSO...878913B (https://ui.adsabs.harvard.edu/abs/2013PLoSO...878913B). doi:10.1371/journal.pone.0078913 (https://doi.org/10.1371%2Fjournal.pone.0078913). PMC 3839928 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3839928). PMID 24282503 (https://www.ncbi.nlm.nih.gov/pubmed/24282503). Highlight (http://www.alterlab.org/research/highlights/pone.0078913_Highlight.pdf).

10. Walton, S.; Hassan, O.; Morgan, K. (2013). "Reduced order modelling for unsteady fluid flow using proper orthogonal decomposition and radial basis functions". *Applied Mathematical Modelling*. **37** (20–21): 8930–8945. doi:10.1016/j.apm.2013.04.025 (https://doi.org/10.1016%2Fj.apm.2013.04.025).

11. Setyawati, Y.; Ohme, F.; Khan, S. (2019). "Enhancing gravitational waveform model through dynamic calibration". *Physical Review D*. **99** (2): 024010. arXiv:1810.07060 (https://arxiv.org/abs/1810.07060). Bibcode:2019PhRvD..99b4010S (https://ui.adsabs.harvard.edu/abs/2019PhRvD..99b4010S). doi:10.1103/PhysRevD.99.024010 (https://doi.org/10.1103%2FPhysRevD.99.024010).

12. Sarwar, Badrul; Karypis, George; Konstan, Joseph A. & Riedl, John T. (2000). "Application of Dimensionality Reduction in Recommender System -- A Case Study" (http://files.grouplens.org/papers/webKDD00.pdf) (PDF). University of Minnesota. Retrieved 26 May 2014.

13. Bosagh Zadeh, Reza; Carlsson, Gunnar (2013). "Dimension Independent Matrix Square Using MapReduce" (http://stanford.edu/~rezab/papers/dimsum.pdf) (PDF). arXiv:1304.1467 (https://arxiv.org/abs/1304.1467). Bibcode:2013arXiv1304.1467B (https://ui.adsabs.harvard.edu/abs/2013arXiv1304.1467B). Retrieved 12 July 2014.

14. "GitHub - it21208/SVDMovie-Lens-Parallel-Apache-Spark" (https://github.com/it21208/SVDMovie-Lens-Parallel-Apache-Spark). 28 January 2019.

15. http://www.timelydevelopment.com/demos/NetflixPrize.aspx

16. Hadi Fanaee-T & João Gama (September 2014). "Eigenspace method for spatiotemporal hotspot detection". *Expert Systems*. **32** (3): 454–464. arXiv:1406.3506 (https://arxiv.org/abs/1406.3506). Bibcode:2014arXiv1406.3506F (https://ui.adsabs.harvard.edu/abs/2014arXiv1406.3506F). doi:10.1111/exsy.12088 (https://doi.org/10.1111%2Fexsy.12088).

17. Hadi Fanaee-T & João Gama (May 2015). "EigenEvent: An Algorithm for Event Detection from Complex Data Streams in Syndromic Surveillance". *Intelligent Data Analysis*. **19** (3): 597–616. arXiv:1406.3496 (https://arxiv.org/abs/1406.3496). doi:10.3233/IDA-150734 (https://doi.org/10.3233%2FIDA-150734).

18. To see this, we just have to notice that $trace(\mathbf{V_2^* M^* M V_2}) = \|\mathbf{M V_2}\|^2$, and remember that $\|A\| = 0 \iff A = 0$

19. Netlib.org (http://www.netlib.org/lapack/double/dbdsqr.f)

20. Netlib.org (http://www.netlib.org/lapack/double/dgesvd.f)

21. mathworks.co.kr/matlabcentral/fileexchange/12674-simple-svd (http://www.mathworks.co.kr/matlabcentral/fileexchange/12674-simple-svd)

22. Fan, Ky. (1951). "Maximum properties and inequalities for the eigenvalues of completely continuous operators" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1063464). *Proceedings of the National Academy of Sciences of the United States of America*. **37** (11): 760–766. Bibcode:1951PNAS...37..760F (https://ui.adsabs.harvard.edu/abs/1951PNAS...37..760F). doi:10.1073/pnas.37.11.760 (https://doi.org/10.1073%2Fpnas.37.11.760). PMC 1063464 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1063464). PMID 16578416 (https://www.ncbi.nlm.nih.gov/pubmed/16578416).

23. De Lathauwer, L.; De Moor, B.; Vandewalle, J. (1 January 2000). "A Multilinear Singular Value Decomposition". *SIAM Journal on Matrix Analysis and Applications*. **21** (4): 1253–1278. CiteSeerX 10.1.1.102.9135 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.102.9135). doi:10.1137/S0895479896305696 (https://doi.org/10.1137%2FS0895479896305696). ISSN 0895-4798 (https://www.worldcat.org/issn/0895-4798).

24. Uhlmann, Jeffrey (2018), *A Generalized Matrix Inverse that is Consistent with Respect to Diagonal Transformations* (http://faculty.missouri.edu/uhlmannj/UC-SIMAX-Final.pdf) (PDF), SIAM Journal on Matrix Analysis, 239:2, pp. 781–800

25. Eckart, C.; Young, G. (1936). "The approximation of one matrix by another of lower rank". *Psychometrika*. **1** (3): 211–8. doi:10.1007/BF02288367 (https://doi.org/10.1007%2FBF02288367).

26. Hestenes, M. R. (1958). "Inversion of Matrices by Biorthogonalization and Related Results". *Journal of the Society for Industrial and Applied Mathematics*. **6** (1): 51–90. doi:10.1137/0106005 (https://doi.org/10.1137%2F0106005). JSTOR 2098862 (https://www.jstor.org/stable/2098862). MR 0092215 (https://www.ams.org/mathscinet-getitem?mr=0092215).

27. Golub, G. H.; Kahan, W. (1965). "Calculating the singular values and pseudo-inverse of a matrix". *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*. **2** (2): 205–224. Bibcode:1965SJNA....2..205G (https://ui.adsabs.harvard.edu/abs/1965SJNA....2..205G). doi:10.1137/0702016 (https://doi.org/10.1137%2F0702016). JSTOR 2949777 (https://www.jstor.org/stable/2949777). MR 0183105 (https://www.ams.org/mathscinet-getitem?mr=0183105).

28. Golub, G. H.; Reinsch, C. (1970). "Singular value decomposition and least squares solutions". *Numerische Mathematik*. **14** (5): 403–420. doi:10.1007/BF02163027 (https://doi.org/10.1007%2FBF02163027). MR 1553974 (https://www.ams.org/mathscinet-getitem?mr=1553974).

References

- Banerjee, Sudipto; Roy, Anindya (2014), *Linear Algebra and Matrix Analysis for Statistics*, Texts in Statistical Science (1st ed.), Chapman and Hall/CRC, ISBN 978-1420095388
- Trefethen, Lloyd N.; Bau III, David (1997). *Numerical linear algebra*. Philadelphia: Society for Industrial and Applied Mathematics. ISBN 978-0-89871-361-9.

- Demmel, James; Kahan, William (1990). "Accurate singular values of bidiagonal matrices". *SIAM Journal on Scientific and Statistical Computing*. **11** (5): 873–912. CiteSeerX 10.1.1.48.3740 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.3740). doi:10.1137/0911052 (https://doi.org/10.1137%2F0911052).
- Golub, Gene H.; Kahan, William (1965). "Calculating the singular values and pseudo-inverse of a matrix". *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*. **2** (2): 205–224. Bibcode:1965SJNA....2..205G (https://ui.adsabs.harvard.edu/abs/1965SJNA....2..205G). doi:10.1137/0702016 (https://doi.org/10.1137%2F0702016). JSTOR 2949777 (https://www.jstor.org/stable/2949777).
- Golub, Gene H.; Van Loan, Charles F. (1996). *Matrix Computations* (3rd ed.). Johns Hopkins. ISBN 978-0-8018-5414-9.
- GSL Team (2007). "§14.4 Singular Value Decomposition" (https://www.gnu.org/software/gsl/manual/html_node/Singular-Value-Decomposition.html). *GNU Scientific Library. Reference Manual*.
- Halldor, Bjornsson and Venegas, Silvia A. (1997). "A manual for EOF and SVD analyses of climate data" (http://brunnur.vedur.is/pub/halldor/TEXT/eofsvd.html). McGill University, CCGCR Report No. 97-1, Montréal, Québec, 52pp.
- Hansen, P. C. (1987). "The truncated SVD as a method for regularization". *BIT*. **27** (4): 534–553. doi:10.1007/BF01937276 (https://doi.org/10.1007%2FBF01937276).
- Horn, Roger A.; Johnson, Charles R. (1985). "Section 7.3". *Matrix Analysis*. Cambridge University Press. ISBN 978-0-521-38632-6.
- Horn, Roger A.; Johnson, Charles R. (1991). "Chapter 3". *Topics in Matrix Analysis*. Cambridge University Press. ISBN 978-0-521-46713-1.
- Samet, H. (2006). *Foundations of Multidimensional and Metric Data Structures* (https://archive.org/details/foundationsofmul00same). Morgan Kaufmann. ISBN 978-0-12-369446-1.
- Strang G. (1998). "Section 6.7". *Introduction to Linear Algebra* (3rd ed.). Wellesley-Cambridge Press. ISBN 978-0-9614088-5-5.
- Stewart, G. W. (1993). "On the Early History of the Singular Value Decomposition" (http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.23.1831). *SIAM Review*. **35** (4): 551–566. CiteSeerX 10.1.1.23.1831 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.23.1831). doi:10.1137/1035134 (https://doi.org/10.1137%2F1035134). hdl:1903/566 (https://hdl.handle.net/1903%2F566). JSTOR 2132388 (https://www.jstor.org/stable/2132388).
- Wall, Michael E.; Rechtsteiner, Andreas; Rocha, Luis M. (2003). "Singular value decomposition and principal component analysis" (http://public.lanl.gov/mewall/kluwer2002.html). In D.P. Berrar; W. Dubitzky; M. Granzow (eds.). *A Practical Approach to Microarray Data Analysis*. Norwell, MA: Kluwer. pp. 91–109.
- Press, WH; Teukolsky, SA; Vetterling, WT; Flannery, BP (2007), "Section 2.6" (http://apps.nrbook.com/empanel/index.html?pg=65), *Numerical Recipes: The Art of Scientific Computing* (3rd ed.), New York: Cambridge University Press, ISBN 978-0-521-88068-8

External links

- Online SVD calculator (http://engineerjs.com/doc/ejs/engine/linalg-1/_svd.html)