What Is the Singular Value Decomposition?

Nicholas J. Higham[*]

October 13, 2020

A singular value decomposition (SVD) of a matrix $A \in \mathbb{R}^{m \times n}$ is a factorization

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

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal, $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m \times n}$, where $p = \min(m, n)$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$.

Partition $U = [u_1, \ldots, u_m]$ and $V = [v_1, \ldots, v_n]$. The σ_i are called the *singular values* of A and the u_i and v_i are the left and right *singular vectors*. We have $Av_i = \sigma_i u_i$, $i = 1\colon p$. The matrix Σ is unique but U and V are not. The form of Σ is

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline & 0 & \end{bmatrix} \text{ for } m \geq n, \quad \Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_m & \end{array} \right] \text{ for } m \leq n$$

Here is an example, in which the entries of A have been specially chosen to give simple forms for the elements of the factors:

$$A = \begin{bmatrix} 0 & \frac{4}{3} \\ -1 & -\frac{5}{3} \\ -2 & -\frac{2}{3} \end{bmatrix} = \underbrace{\frac{1}{3} \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{V^T}.$$

The power of the SVD is that it reveals a great deal of useful information about norms, rank, and subspaces of a matrix and it enables many problems to be reduced to a trivial form.

Since U and V are nonsingular, $\operatorname{rank}(A) = \operatorname{rank}(\Sigma) = r$, where $r \leq p$ is the number of nonzero singular values. Since the 2-norm and Frobenius norm are invariant under orthogonal transformations, $\|A\| = \|\Sigma\|$ for both norms, giving

$$\|A\|_2 = \sigma_1, \quad \|A\|_F = \left(\sum_{i=1}^{r} \sigma_i^2 \right)^{1/2},$$

and hence $\|A\|_2 \leq \|A\|_F \leq r^{1/2} \|A\|_2$. The range space and null space of A are given in terms of the columns of U and V by

$$\operatorname{null}(A) = \operatorname{span}\{v_{r+1}, \ldots, v_n\},$$
$$\operatorname{range}(A) = \operatorname{span}\{u_1, u_2, \ldots, u_r\}.$$

---

[*]Department of Mathematics, University of Manchester, Manchester, M13 9PL, UK (nick.higham@manchester.ac.uk).

We can write the SVD as

$$A = \begin{bmatrix} u_1, u_2 \ldots, u_r \end{bmatrix} \operatorname{diag}(\sigma_1, \ldots, \sigma_r) \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{bmatrix} = \sum_{i=1}^{r} \sigma_i u_i v_i^T, \qquad (*)$$

which expresses A as a sum of r rank-1 matrices, the ith of which has 2-norm σ_i. The famous Eckart–Young theorem (1936) says that

$$\min_{\operatorname{rank}(B)=k} \|A - B\|_q = \begin{cases} \sigma_{k+1}, & q = 2, \\ \left(\sum_{i=k+1}^{r} \sigma_i^2 \right)^{1/2}, & q = F, \end{cases}$$

and that the minimum is attained at

$$A_k = U D_k V^T, \quad D_k = \operatorname{diag}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0).$$

In other words, truncating the this sum $(*)$ after $k < r$ terms gives the best rank-k approximation to A in both the 2-norm and the Frobenius norm. In particular, this result implies that when A has full rank the distance from A to the nearest rank-deficient matrix is σ_r.

Relations with Symmetric Eigenvalue Problem

The SVD is not directly related to the eigenvalues and eigenvectors of A. However, for $m \geq n$, $A = U\Sigma V^T$ implies

$$A^T A = V \operatorname{diag}(\sigma_1^2, \ldots, \sigma_n^2) V^T, \quad AA^T = U \operatorname{diag}(\sigma_1^2, \ldots, \sigma_n^2, \underbrace{0, \ldots, 0}_{m-n}) U^T,$$

so the singular values of A are the square roots of the eigenvalues of the symmetric positive semidefinite matrices $A^T A$ and AA^T (modulo $m - n$ zeros in the latter case), and the singular vectors are eigenvectors. Moreover, the eigenvalues of the $(m+n) \times (m+n)$ matrix

$$C = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$$

are plus and minus the singular values of A, together with $|m - n|$ additional zeros if $m \neq n$, and the eigenvectors of C and the singular vectors of A are also related.

Consequently, by applying results or algorithms for the eigensystem of a symmetric matrix to $A^T A$, AA^T, or C one obtains results or algorithms for the singular value decomposition of A.

Connections with Other Problems

The pseudoinverse of a matrix $A \in \mathbb{R}^{n \times n}$ can be expressed in terms of the SVD as

$$A^+ = V \operatorname{diag}(\sigma_1^{-1}, \ldots, \sigma_r^{-1}, 0, \ldots, 0) U^T.$$

The least squares problem $\min_x \|b - Ax\|_2$, where $A \in \mathbb{R}^{m \times n}$ with $m \geq n$ is solved by $x = A^+b$, and when A is rank-deficient this is the solution of minimum 2-norm. For $m < n$ this is an underdetermined system and $x = A^+b$ gives the minimum 2-norm solution.

We can write $A = U\Sigma V^T = UV^T \cdot V\Sigma V^T \equiv PQ$, where P is orthogonal and Q is symmetric positive semidefinite. This decomposition $A = PQ$ is the polar decomposition and $Q = (A^TA)^{1/2}$ is unique. This connection between the SVD and the polar decomposition is useful both theoretically and computationally.

Applications

The SVD is used in a very wide variety of applications—too many and varied to attempt to summarize here. We just mention two.

The SVD can be used to help identify to which letters vowels and consonants have been mapped in a substitution cipher (Moler and Morrison, 1983).

An inverse use of the SVD is to construct test matrices by forming a diagonal matrix of singular values from some distribution then pre- and post-multiplying by random orthogonal matrices. The result is matrices with known singular values and 2-norm condition number that are nevertheless random. Such "randsvd" matrices are widely used to test algorithms in numerical linear algebra.

History and Computation

The SVD was introduced independently by Beltrami in 1873 and Jordan in 1874. Golub popularized the SVD as an essential computational tool and developed the first reliable algorithms for computing it. The Golub–Reinsch algorithm, dating from the late 1960s and based on bidiagonalization and the QR algorithm, is the standard way to compute the SVD. Various alternatives are available; see the references.

References

This is a minimal set of references, which contain further useful references within.

- Jack Dongarra and Mark Gates and Azzam Haidar and Jakub Kurzak and Piotr Luszczek and Stanimire Tomov and Ichitaro Yamazaki, The Singular Value Decomposition: Anatomy of Optimizing an Algorithm for Extreme Scale, SIAM Rev. 60(4), 808–865, 2018.
- Gene Golub and Charles F. Van Loan, Matrix Computations, fourth edition, Johns Hopkins University Press, Baltimore, MD, USA, 2013.
- Roger Horn and Charles Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991. Chapter 3.
- Cleve B. Moler and Donald Morrison, Singular Value Analysis of Cryptograms, Amer. Math. Monthly 90, 78–87, 1983.
- Yuji Nakatsukasa and Nicholas J. Higham, Stable and Efficient Spectral Divide and Conquer Algorithms for the Symmetric Eigenvalue Decomposition and the SVD, SIAM J. Sci. Comput. 35(3), A1325–A1349, 2013.

Related Blog Posts

- Faster SVD via Polar Decomposition (2015)
- Professor SVD by Cleve Moler (2006)
- What Is a Random Orthogonal Matrix? (2020)
- What Is the Polar Decomposition? (2020)

This article is part of the "What Is" series, available from `https://nhigham.com/category/what-is` and in PDF form from the GitHub repository `https://github.com/higham/what-is`.