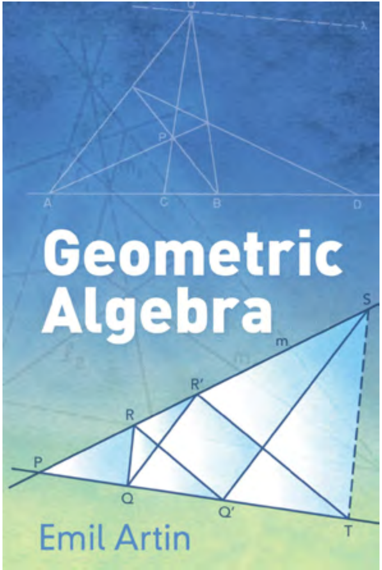
The Singular Value Decomposition

Sheldon Axler

27 January 2023 and 3 February 2023

It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out.

–Emil Artin, 1957



notation and definitions

$\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$.

V and W are finite-dimensional inner product spaces over $\mathbf{F}$.

$n = \dim V$.

$T\colon V \to W$ is a linear map.

$T^*\colon W \to V$ is the linear map defined by

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

for all $v \in V$ and all $w \in W$.

T^* is called the *adjoint* of T.

An operator is called *self-adjoint* if it equals its adjoint.

$T^*T\colon V \to V$ is self-adjoint because

$$(T^*T)^* = T^*(T^*)^* = T^*T.$$

Suppose $\lambda \in \mathbf{F}$ is an eigenvalue of T^*T with eigenvector $v \in V$. Then

$$\begin{aligned}\lambda\|v\|^2 &= \langle \lambda v, v\rangle\\ &= \langle T^*Tv, v\rangle\\ &= \langle Tv, Tv\rangle\\ &\geq 0.\end{aligned}$$

Thus $\lambda \geq 0$. Hence all eigenvalues of T^*T are nonnegative numbers.

The *multiplicity* of the eigenvalue λ of T^*T is $\dim \operatorname{null}(T^*T - \lambda I)$.

singular values

definition: *singular values*

The *singular values* of T are the nonnegative square roots of the eigenvalues of T^*T, listed in decreasing order, each included as many times as its multiplicity.

Example: Define $T\colon \mathbf{F}^4 \to \mathbf{F}^4$ by

$$T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4).$$

Then

$$T^*T(z_1, z_2, z_3, z_4) = (9z_1, 4z_2, 0, 9z_4).$$

Thus the eigenvalues of T^*T are 9, 4, and 0, with multiplicities 2, 1, and 1.

Hence the singular values of T are $3, 3, 2, 0$.

Example: Suppose $T\colon \mathbf{F}^4 \to \mathbf{F}^3$ has matrix (with respect to standard bases)

$$\begin{pmatrix} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

Then T^*T has matrix

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix}.$$

The eigenvalues of T^*T are 25, 2, 0, with multiplicities 1, 1, and 2. Thus the singular values of T are $5, \sqrt{2}, 0, 0$.

singular values

V and W are finite-dimensional inner product spaces over $\mathbf{F}$.

$n = \dim V$.

$T\colon V \to W$ is a linear map.

definition: *singular values*

The *singular values* of T are the nonnegative square roots of the eigenvalues of T^*T, listed in decreasing order, each included as many times as its multiplicity.

T has n singular values.

basic results on singular values

(a) T is injective $\iff$ 0 is not a singular value of T.

(b) The number of positive singular values of T equals $\dim \operatorname{range} T$.

(c) T is surjective $\iff$ number of positive singular values of T equals $\dim W$.

(d) $T = 0$ $\iff$ all singular values of T are 0.

(e) T is an isometry ($\|Tv\| = \|v\|$ for all $v \in V$) $\iff$ all singular values of T are 1.

Erhard Schmidt result from 1907

characterization of positive singular values

Suppose $s > 0$. Then s is a singular value of T if and only if there exist nonzero vectors $v \in V$ and $w \in W$ such that

$$Tv = sw \quad \text{and} \quad T^*w = sv.$$

The earliest appearance I have found of the singular value decomposition in a linear algebra textbook is in Gil Strang's 1976 textbook *Linear Algebra and Its Applications*, which includes the following:

"These simple matrices are much more valuable than they look, because of a new way to factor the matrix A. It is called the singular value decomposition, and it is not nearly as famous as it should be."

comparison of eigenvalues and singular values

list of eigenvalues	list of singular values
context: vector spaces	context: inner product spaces
defined only for linear maps from a vector space to itself	defined for linear maps from an inner product space to a possibly different inner product space
can be arbitrary real numbers (if $\mathbf{F} = \mathbf{R}$) or complex numbers (if $\mathbf{F} = \mathbf{C}$)	are nonnegative numbers
can be the empty list if $\mathbf{F} = \mathbf{R}$	length of list equals dimension of domain
includes $0 \iff$ operator is not invertible	includes $0 \iff$ linear map is not injective
no standard order, especially if $\mathbf{F} = \mathbf{C}$	always listed in decreasing order

singular value decomposition

Suppose the positive singular values of T are $s_1, \dots, s_m$. Then there exist orthonormal lists $e_1, \dots, e_m$ in V and $f_1, \dots, f_m$ in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$.

Proof Let $s_1, \dots, s_n$ denote the singular values of T. By the spectral theorem, there exists an orthonormal basis $e_1, \dots, e_n$ of V with

$$T^* T e_k = s_k^2 e_k$$

for $k = 1, \dots, n$.

For $k = 1, \dots, m$, let $f_k = \frac{Te_k}{s_k}$. Then

$$\begin{aligned} \langle f_j, f_k \rangle &= \frac{1}{s_j s_k} \langle Te_j, Te_k \rangle = \frac{1}{s_j s_k} \langle e_j, T^* T e_k \rangle \\ &= \frac{s_k}{s_j} \langle e_j, e_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases} \end{aligned}$$

Thus $f_1, \dots, f_m$ is an orthonormal list.

If $1 \leq k \leq m$, then $Te_k = s_k f_k$.
If $m < k \leq n$, then $Te_k = 0$.

Suppose $v \in V$. Then

$$\begin{aligned} Tv &= T\bigl(\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n \bigr) \\ &= \langle v, e_1 \rangle Te_1 + \cdots + \langle v, e_m \rangle Te_m \\ &= s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m. \quad \blacksquare \end{aligned}$$

singular value decomposition

Suppose the positive singular values of T are $s_1, \ldots, s_m$. Then there exist orthonormal lists $e_1, \ldots, e_m$ in V and $f_1, \ldots, f_m$ in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$.

With notation as above,

$$Te_k = s_k f_k$$

for $k = 1, \ldots, m$. Also,

$$\operatorname{range} T = \operatorname{span}(f_1, \ldots, f_m).$$

Extend $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ to orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_{\dim W}$ of V and W.

Then

$$Te_k = \begin{cases} s_k f_k & \text{if } 1 \leq k \leq m, \\ 0 & \text{if } m < k \leq n. \end{cases}$$

The entry in row j, column k of the matrix of T with respect to these bases is

$$\mathcal{M}\big(T, (e_1, \ldots, e_n), (f_1, \ldots, f_{\dim W})\big)_{j,k} = \begin{cases} s_k & \text{if } 1 \leq j = k \leq m, \\ 0 & \text{otherwise.} \end{cases}$$

Thus T has a "diagonal" matrix with respect to these bases.

comparison of spectral theorem and singular value decomposition

spectral theorem	**singular value decomposition**
describes only self-adjoint operators (when $\mathbf{F} = \mathbf{R}$) or normal operators (when $\mathbf{F} = \mathbf{C}$)	describes arbitrary linear maps from an inner product space to a possibly different inner product space
produces a single orthonormal basis	produces two orthonormal lists, one for domain space and one for range space, that are not necessarily the same even when range space equals domain space
different proofs depending upon whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$	same proof works regardless of whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$

adjoint via SVD

SVD of adjoint

Suppose $s_1, \ldots, s_m$ are the positive singular values of T. Suppose $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ are orthonormal lists in V and W such that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for every $v \in V$. Then

$$T^*w = s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m$$

for every $w \in W$.

Proof

If $v \in V$ and $w \in W$ then

$$\begin{aligned}
\langle Tv, w\rangle \\
&= \big\langle s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m, w\big\rangle \\
&= s_1\langle v, e_1\rangle\langle f_1, w\rangle + \cdots + s_m\langle v, e_m\rangle\langle f_m, w\rangle \\
&= \big\langle v, s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m\big\rangle.
\end{aligned}$$

This implies that

$$T^*w = s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m,$$

as desired. ∎

inverse via SVD

SVD of inverse

Suppose $s_1, \ldots, s_m$ are the positive singular values of T. Suppose $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ are orthonormal lists in V and W such that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for every $v \in V$. Then

$$T^*w = s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m$$

for every $w \in W$.
If T in invertible, then

$$T^{-1}w = \frac{\langle w, f_1\rangle}{s_1}e_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m}e_m$$

for every $w \in W$.

Proof

Suppose T in invertible and $w \in W$. Let

$$v = \frac{\langle w, f_1\rangle}{s_1}e_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m}e_m.$$

Apply T to both sides, getting

$$\begin{aligned} Tv &= \frac{\langle w, f_1\rangle}{s_1}Te_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m}Te_m \\ &= \langle w, f_1\rangle f_1 + \cdots + \langle w, f_m\rangle f_m \\ &= w, \end{aligned}$$

where the last line holds because $f_1, \ldots, f_m$ is an orthonormal basis of $\operatorname{range} T$. The equation above shows that $v = T^{-1}w$. ∎

SVD for matrices

matrix version of SVD

Suppose A is an M-by-n matrix with rank $m \geq 1$. Then there exist an M-by-m matrix B with orthonormal columns, an m-by-m diagonal matrix D with positive entries on the diagonal, and an n-by-m matrix C with orthonormal columns such that

$$A = BDC^*.$$

Proof Let $T\colon \mathbf{F}^n \to \mathbf{F}^M$ be the linear map whose matrix equals A. Let $s_1, \ldots, s_m$ be the positive singular values of T. By SVD, there exist orthonormal lists $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ in $\mathbf{F}^n$ and $\mathbf{F}^M$ such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for every $v \in \mathbf{F}^n$.

Let B be the M-by-m matrix whose columns are $f_1, \ldots, f_m$.

Let D be the m-by-m diagonal matrix whose diagonal entries are $s_1, \ldots, s_m$.

Let C be the n-by-m matrix whose columns are $e_1, \ldots, e_m$. Then

$$AC = BD.$$

Multiply both sides on the right by C^*, and use $ACC^* = A$ to get

$$A = BDC^*. \qquad \blacksquare$$

Frobenius norm

Hilbert–Schmidt norm and singular values

Let $s_1, \dots, s_n$ denote the singular values of T. Suppose $v_1, \dots, v_n$ is an orthonormal basis of V and $w_1, \dots, w_M$ is an orthonormal basis of W. Then

$$\sum_{k=1}^{n} \|Tv_k\|^2 = \sum_{k=1}^{n} \sum_{j=1}^{M} |\langle Tv_k, w_j \rangle|^2 = \sum_{k=1}^{n} s_k^2.$$

Proof

$$\sum_{k=1}^{n} \|Tv_k\|^2 = \sum_{k=1}^{n} \langle Tv, Tv \rangle = \sum_{k=1}^{n} \langle T^*Tv_k, v_k \rangle = \operatorname{trace} T^*T.$$

Thus this sum does not depend on the orthonormal basis $v_1, \dots, v_n$.

Hence we need only show that

$$\sum_{k=1}^{n} \|Tv_k\|^2 = \sum_{k=1}^{n} s_k^2$$

for ***some*** orthonormal basis $v_1, \dots, v_n$ of V. Suppose we have a SVD

$$Tv = \sum_{k=1}^{m} s_k \langle v, e_k \rangle f_k.$$

Extend $e_1, \dots, e_m$ to an orthonormal basis $e_1, \dots, e_n$ of V. Then

$$Te_k = \begin{cases} s_k f_k & \text{if } 1 \le k \le m \\ 0 & \text{if } m < k \le n. \end{cases}$$

Thus $\sum_{k=1}^{n} \|Te_k\|^2 = \sum_{k=1}^{n} s_k^2$. ∎

Hilbert matrix

Let H_n denote the n-by-n Hilbert matrix, whose entry in row j, column k is $\frac{1}{j+k-1}$.

Example:

$$H_4 = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}.$$

My computer claims that if $n = 20$, then

$$s_1 \approx 1.91 \quad \text{and} \quad s_{20} \approx 7.78 \times 10^{-29}.$$

Is H_{20} invertible? Equivalently, is $H_{20} = 0$?

Suppose $b = (b_1, \ldots, b_{20}) \in \mathbf{R}^{20}$ with $b \neq 0$. Then

$$\begin{aligned}
\langle H_{20}b, b\rangle &= \sum_{k=1}^{20}\sum_{j=1}^{20} b_k b_j \frac{1}{j+k-1} \\
&= \sum_{k=1}^{20}\sum_{j=1}^{20} b_k b_j \int_0^1 x^{k-1}x^{j-1}\,dx \\
&= \int_0^1 \left(\sum_{k=1}^{20} b_k x^{k-1}\right)\left(\sum_{j=1}^{20} b_j x^{j-1}\right) dx \\
&= \int_0^1 \left(\sum_{k=1}^{20} b_k x^{k-1}\right)^2 dx \\
&> 0.
\end{aligned}$$

Thus $H_{20}b \neq 0$; hence H_{20} is invertible.

$$\det H_{20} \approx 4.21 \times 10^{-226}$$

applications of the SVD

singular value decomposition

Suppose the positive singular values of T are $s_1, \ldots, s_m$. Then there exist orthonormal lists $e_1, \ldots, e_m$ in V and $f_1, \ldots, f_m$ in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$.

One week from today:

- norms of linear maps
- approximation of T by linear maps with lower-dimensional range
- polar decomposition
- operators applied to ellipsoids and parallelograms
- volume via singular values
- formula for pseudoinverse using SVD

reference

Sheldon Axler, *Linear Algebra Done Right*, fourth edition to be published in Springer's Undergraduate Texts in Mathematics series around December 2023.

Chapter 7 (Operators in Inner Product Spaces) contains the material on the singular value decomposition in Sections 7E and 7F. Chapter 7 is now freely and legally available on the book's website https://linear.axler.net.

This will be an Open Access book, meaning that the electronic version will be legally free to the world.