7E Singular Value Decomposition

Singular Values

We will need the following result in this section.

7.64 properties of T^*T

Suppose $T \in \mathcal{L}(V, W)$. Then

- (a) T^*T is a positive operator on V;
- (b) $\operatorname{null} T^*T = \operatorname{null} T$;
- (c) range T^*T = range T^* ;
- (d) $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$.

Proof

(a) We have

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

Thus T^*T is self-adjoint.

If $v \in V$, then

$$\left\langle \left(T^*T\right)v,v\right\rangle = \left\langle T^*(Tv),v\right\rangle = \left\langle Tv,Tv\right\rangle = \|Tv\|^2 \geq 0.$$

Thus T^*T is a positive operator.

(b) First suppose $v \in \text{null } T^*T$. Then

$$||Tv||^2 = \langle Tv, Tv \rangle = \langle T^*Tv, v \rangle = \langle 0, v \rangle = 0.$$

Thus Tv = 0, proving that null $T^*T \subseteq \text{null } T$.

The inclusion in the other direction is clear, because if $v \in V$ and Tv = 0, then $T^*Tv = 0$.

Thus null $T^*T = \text{null } T$, completing the proof of (b).

(c) We already know from (a) that T^*T is self-adjoint. Thus

range
$$T^*T = (\operatorname{null} T^*T)^{\perp} = (\operatorname{null} T)^{\perp} = \operatorname{range} T^*,$$

where the first and last equalities come from 7.6 and the second equality comes from (b).

(d) To verify the first equation in (d), note that

$$\dim \operatorname{range} T = \dim (\operatorname{null} T^*)^{\perp} = \dim W - \dim \operatorname{null} T^* = \dim \operatorname{range} T^*,$$

where the first equality comes from 7.6(d), the second equality comes from 6.51, and the last equality comes from the fundamental theorem of linear maps (3.21).

The equality dim range $T^* = \dim \operatorname{range} T^*T$ follows from (c).

The eigenvalues of an operator tell us something about the behavior of the operator. Another collection of numbers, called the singular values, is also useful. Eigenspaces and the notation E (used in the examples) were defined in 5.52.

7.65 definition: singular values

Suppose $T \in \mathcal{L}(V, W)$. 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 the dimension of the corresponding eigenspace of T^*T .

7.66 example: singular values of an operator on \mathbf{F}^4

Define $T\in\mathcal{L}\big(\mathbf{F}^4\big)$ by $T(z_1,z_2,z_3,z_4)=(0,3z_1,2z_2,-3z_4).$ A calculation shows that

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

as you should verify. Thus the standard basis of \mathbf{F}^4 diagonalizes T^*T , and we see that the eigenvalues of T^*T are 9, 4, and 0. Also, the dimensions of the eigenspaces corresponding to the eigenvalues are

$$\dim E(9, T^*T) = 2$$
 and $\dim E(4, T^*T) = 1$ and $\dim E(0, T^*T) = 1$.

Taking nonnegative square roots of these eigenvalues of T^*T and using dimension information from above, we conclude that the singular values of T are 3, 3, 2, 0.

The only eigenvalues of T are -3 and 0. Thus in this case, the collection of eigenvalues did not pick up the number 2 that appears in the definition (and hence the behavior) of T, but the list of singular values does include 2.

7.67 example: singular values of a linear map from F^4 to F^3

Suppose $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ has matrix (with respect to the standard bases)

$$\left(\begin{array}{cccc} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{array}\right).$$

You can verify that the matrix of T^*T is

$$\left(\begin{array}{cccc} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{array}\right)$$

and that the eigenvalues of the operator T^*T are 25, 2, 0, with dim $E(25, T^*T) = 1$, dim $E(2, T^*T) = 1$, and dim $E(0, T^*T) = 2$. Thus the singular values of T are $5, \sqrt{2}, 0, 0$.

See Exercise 2 for a characterization of the positive singular values.

7.68 role of positive singular values

Suppose that $T \in \mathcal{L}(V, W)$. Then

- (a) T is injective \iff 0 is not a singular value of T;
- (b) the number of positive singular values of T equals dim range T;
- (c) T is surjective \iff number of positive singular values of T equals dim W.

Proof The linear map T is injective if and only if null $T = \{0\}$, which happens if and only if null $T^*T = \{0\}$ [by 7.64(b)], which happens if and only if 0 is not an eigenvalue of T^*T , which happens if and only if 0 is not a singular value of T, completing the proof of (a).

The spectral theorem applied to T^*T shows that dim range T^*T equals the number of positive eigenvalues of T^*T (counting repetitions). Thus 7.64(d) implies that dim range T equals the number of positive singular values of T, proving (b).

Use (b) and 2.39 to show that (c) holds.

The table below compares eigenvalues with 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 $F = R$) or complex numbers (if $F = C$)	are nonnegative numbers
can be the empty list if $F = R$	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 $F = C$	always listed in decreasing order

The next result nicely characterizes isometries in terms of singular values.

7.69 isometries characterized by having all singular values equal 1

Suppose that $S \in \mathcal{L}(V, W)$. Then

S is an isometry \iff all singular values of S equal 1.

Proof We have

$$S$$
 is an isometry $\iff S^*S = I$
 \iff all eigenvalues of S^*S equal 1
 \iff all singular values of S equal 1,

where the first equivalence comes from 7.49 and the second equivalence comes from the spectral theorem (7.29 or 7.31) applied to the self-adjoint operator S^*S .

SVD for Linear Maps and for Matrices

The next result shows that every linear map from V to W has a remarkably clean description in terms of its singular values and orthonormal lists in V and W. In the next section we will see several important applications of the singular value decomposition (often called the SVD).

The singular value decomposition is useful in computational linear algebra because good techniques exist for approximating eigenvalues and eigenvectors of positive operators such as T*T, whose eigenvalues and eigenvectors lead to the singular value decomposition.

7.70 singular value decomposition

Suppose $T \in \mathcal{L}(V, W)$ and 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

7.71
$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$.

Proof Let $s_1, ..., s_n$ denote the singular values of T (thus $n = \dim V$). Because T^*T is a positive operator [see 7.64(a)], the spectral theorem implies that there exists an orthonormal basis $e_1, ..., e_n$ of V with

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

for each k = 1, ..., n.

For each k = 1, ..., m, let

$$f_k = \frac{Te_k}{s_k}.$$

If $j, k \in \{1, \dots, m\}$, then

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

Thus f_1, \dots, f_m is an orthonormal list in W.

If $k \in \{1, ..., n\}$ and k > m, then $s_k = 0$ and hence $T^*Te_k = 0$ (by 7.72), which implies that $Te_k = 0$ [by 7.64(b)].

Suppose $v \in V$. Then

$$Tv = T(\langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n)$$

$$= \langle v, e_1 \rangle Te_1 + \dots + \langle v, e_m \rangle Te_m$$

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

where the last index in the first line switched from n to m in the second line because $Te_k = 0$ if k > m (as noted in the paragraph above) and the third line follows from 7.73. The equation above is our desired result.

274

Suppose $T \in \mathcal{L}(V,W)$, the positive singular values of T are s_1,\ldots,s_m , and e_1,\ldots,e_m and f_1,\ldots,f_m are as in the singular value decomposition 7.70. The orthonormal list e_1,\ldots,e_m can be extended to an orthonormal basis $e_1,\ldots,e_{\dim V}$ of V and the orthonormal list f_1,\ldots,f_m can be extended to an orthonormal basis $f_1,\ldots,f_{\dim W}$ of V. The formula 7.71 shows that

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

Thus the matrix of T with respect to the orthonormal bases $(e_1, \dots, e_{\dim V})$ and $(f_1, \dots, f_{\dim W})$ has the simple form

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

If dim $V = \dim W$ (as happens, for example, if W = V), then the matrix described in the paragraph above is a diagonal matrix. If we extend the definition of diagonal matrix as follows to apply to matrices that are not necessarily square, then we have proved the wonderful result that every linear map from V to W has a diagonal matrix with respect to appropriate orthonormal bases.

7.74 definition: diagonal matrix

An M-by-N matrix A is called a *diagonal matrix* if all entries of the matrix are 0 except possibly $A_{k,k}$ for $k=1,\ldots,\min\{M,N\}$.

The table below compares the spectral theorem (7.29 and 7.31) with the singular value decomposition (7.70).

spectral theorem	singular value decomposition
	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 on whether $F = R$ or $F = C$	same proof works regardless of whether $F = R$ or $F = C$

The singular value decomposition gives us a new way to understand the adjoint and the inverse of a linear map. Specifically, the next result shows that given a singular value decomposition of a linear map $T \in \mathcal{L}(V,W)$, we can obtain the adjoint of T simply by interchanging the roles of the e's and the f's (see 7.77). Similarly, we can obtain the pseudoinverse T^{\dagger} (see 6.68) of T by interchanging the roles of the e's and the f's and replacing each positive singular value s_k of T with $1/s_k$ (see 7.78).

Recall that the pseudoinverse T^{\dagger} in 7.78 below equals the inverse T^{-1} if T is invertible [see 6.69(a)].

7.75 singular value decomposition of adjoint and pseudoinverse

Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are s_1, \dots, s_m . Suppose e_1, \dots, e_m and f_1, \dots, f_m are orthonormal lists in V and W such that

7.76
$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$. Then

7.77
$$T^*w = s_1 \langle w, f_1 \rangle e_1 + \dots + s_m \langle w, f_m \rangle e_m$$

and

7.78
$$T^{\dagger}w = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \dots + \frac{\langle w, f_m \rangle}{s_m} e_m$$

for every $w \in W$.

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

$$\begin{split} \langle Tv,w\rangle &= \left\langle s_1\langle v,e_1\rangle f_1 + \dots + s_m\langle v,e_m\rangle f_m,w\right\rangle \\ &= s_1\langle v,e_1\rangle\langle f_1,w\rangle + \dots + s_m\langle v,e_m\rangle\langle f_m,w\rangle \\ &= \left\langle v,s_1\langle w,f_1\rangle e_1 + \dots + s_m\langle w,f_m\rangle e_m\right\rangle. \end{split}$$

This implies that

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

proving 7.77.

To prove 7.78, suppose $w \in W$. Let

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

Apply T to both sides of the equation above, getting

$$Tv = \frac{\langle w, f_1 \rangle}{s_1} Te_1 + \dots + \frac{\langle w, f_m \rangle}{s_m} Te_m$$

$$= \langle w, f_1 \rangle f_1 + \dots + \langle w, f_m \rangle f_m$$

$$= P_{\text{range } T} w,$$

where the second line holds because 7.76 implies that $Te_k = s_k f_k$ if k = 1, ..., m, and the last line above holds because 7.76 implies that $f_1, ..., f_m$ spans range T and thus is an orthonormal basis of range T [and hence 6.57(i) applies]. The equation above, the observation that $v \in (\text{null } T)^{\perp}$ [see Exercise 8(b)], and the definition of $T^{\dagger}w$ (see 6.68) show that $v = T^{\dagger}w$, proving 7.78.

7.79 example: finding a singular value decomposition

Define $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ by $T(x_1, x_2, x_3, x_4) = (-5x_4, 0, x_1 + x_2)$. We want to find a singular value decomposition of T. The matrix of T (with respect to the standard bases) is

$$\left(\begin{array}{cccc} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{array}\right).$$

Thus, as discussed in Example 7.67, the matrix of T^*T is

$$\left(\begin{array}{cccc} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{array}\right),$$

and the positive eigenvalues of T^*T are 25, 2, with $\dim E(25, T^*T) = 1$ and $\dim E(2, T^*T) = 1$. Hence the positive singular values of T are $5, \sqrt{2}$.

Thus to find a singular value decomposition of T, we must find an orthonormal list e_1, e_2 in \mathbf{F}^4 and an orthonormal list f_1, f_2 in \mathbf{F}^3 such that

$$Tv = 5\langle v, e_1 \rangle f_1 + \sqrt{2}\langle v, e_2 \rangle f_2$$

for all $v \in \mathbf{F}^4$.

An orthonormal basis of $E(25, T^*T)$ is the vector (0, 0, 0, 1); an orthonormal basis of $E(2, T^*T)$ is the vector $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0)$. Thus, following the proof of 7.70, we take

$$e_1 = (0, 0, 0, 1)$$
 and $e_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)$

and

$$f_1 = \frac{Te_1}{5} = (-1, 0, 0)$$
 and $f_2 = \frac{Te_2}{\sqrt{2}} = (0, 0, 1)$.

Then, as expected, we see that e_1,e_2 is an orthonormal list in ${\bf F}^4$ and f_1,f_2 is an orthonormal list in ${\bf F}^3$ and

$$Tv = 5\langle v, e_1 \rangle f_1 + \sqrt{2}\langle v, e_2 \rangle f_2$$

for all $v \in \mathbf{F}^4$. Thus we have found a singular value decomposition of T.

The next result translates the singular value decomposition from the context of linear maps to the context of matrices. Specifically, the following result gives a factorization of an arbitrary matrix as the product of three nice matrices. The proof gives an explicit construction of these three matrices in terms of the singular value decomposition.

In the next result, the phrase "orthonormal columns" should be interpreted to mean that the columns are orthonormal with respect to the standard Euclidean inner product.

7.80 matrix version of SVD

Suppose A is a p-by-n matrix of rank $m \ge 1$. Then there exist a p-by-m matrix B with orthonormal columns, an m-by-m diagonal matrix D with positive numbers on the diagonal, and an n-by-m matrix C with orthonormal columns such that

$$A = BDC^*$$
.

Proof Let $T: \mathbf{F}^n \to \mathbf{F}^p$ be the linear map whose matrix with respect to the standard bases equals A. Then dim range T = m (by 3.78). Let

7.81
$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_m \langle v, e_m \rangle f_m$$

be a singular value decomposition of T. Let

 $B = \text{the } p\text{-by-}m \text{ matrix whose columns are } f_1, ..., f_m,$

 $D = \text{the } m\text{-by-}m \text{ diagonal matrix whose diagonal entries are } s_1, ..., s_m,$

 $C = \text{the } n\text{-by-}m \text{ matrix whose columns are } e_1, ..., e_m.$

Let u_1, \dots, u_m denote the standard basis of \mathbf{F}^m . If $k \in \{1, \dots, m\}$ then

$$(AC - BD) u_k = Ae_k - B(s_k u_k) = s_k f_k - s_k f_k = 0.$$

Thus AC = BD.

Multiply both sides of this last equation by C^* (the conjugate transpose of C) on the right to get

$$ACC^* = BDC^*$$
.

Note that the rows of C^* are the complex conjugates of e_1, \ldots, e_m . Thus if $k \in \{1, \ldots, m\}$, then the definition of matrix multiplication shows that $C^*e_k = u_k$; hence $CC^*e_k = e_k$. Thus $ACC^*v = Av$ for all $v \in \operatorname{span}(e_1, \ldots, e_m)$.

If $v \in (\operatorname{span}(e_1, \dots, e_m))^{\perp}$, then Av = 0 (as follows from 7.81) and $C^*v = 0$ (as follows from the definition of matrix multiplication). Hence $ACC^*v = Av$ for all $v \in (\operatorname{span}(e_1, \dots, e_m))^{\perp}$.

Because ACC^* and A agree on $\operatorname{span}(e_1, \dots, e_m)$ and on $\left(\operatorname{span}(e_1, \dots, e_m)\right)^{\perp}$, we conclude that $ACC^* = A$. Thus the displayed equation above becomes

$$A = BDC^*$$

as desired.

Note that the matrix A in the result above has pn entries. In comparison, the matrices B, D, and C above have a total of

$$m(p+m+n)$$

entries. Thus if p and n are large numbers and the rank m is considerably less than p and n, then the number of entries that must be stored on a computer to represent A is considerably less than pn.

Exercises 7E

- 1 Suppose $T \in \mathcal{L}(V, W)$. Show that T = 0 if and only if all singular values of T are 0.
- 2 Suppose $T \in \mathcal{L}(V, W)$ and s > 0. Prove that s is a singular value of T if and only if there exist nonzero vectors $v \in V$ and $w \in W$ such that

$$Tv = sw$$
 and $T^*w = sv$.

The vectors v, w satisfying both equations above are called a **Schmidt pair**. Erhard Schmidt introduced the concept of singular values in 1907.

- 3 Give an example of $T \in \mathcal{L}(\mathbb{C}^2)$ such that 0 is the only eigenvalue of T and the singular values of T are 5, 0.
- **4** Suppose that $T \in \mathcal{L}(V, W)$, s_1 is the largest singular value of T, and s_n is the smallest singular value of T. Prove that

$${||Tv|| : v \in V \text{ and } ||v|| = 1} = [s_n, s_1].$$

- 5 Suppose $T \in \mathcal{L}(\mathbb{C}^2)$ is defined by T(x,y) = (-4y,x). Find the singular values of T.
- **6** Find the singular values of the differentiation operator $D \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ defined by Dp = p', where the inner product on $\mathcal{P}_2(\mathbf{R})$ is as in Example 6.34.
- 7 Suppose that $T \in \mathcal{L}(V)$ is self-adjoint or that $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, each included in this list as many times as the dimension of the corresponding eigenspace. Show that the singular values of T are $|\lambda_1|, \ldots, |\lambda_n|$, after these numbers have been sorted into decreasing order.
- **8** Suppose $T \in \mathcal{L}(V, W)$. Suppose $s_1 \ge s_2 \ge \cdots \ge s_m > 0$ and e_1, \ldots, e_m is an orthonormal list in V and f_1, \ldots, f_m is an orthonormal list in W such that

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

for every $v \in V$.

- (a) Prove that f_1, \dots, f_m is an orthonormal basis of range T.
- (b) Prove that e_1, \dots, e_m is an orthonormal basis of $(\text{null } T)^{\perp}$.
- (c) Prove that s_1, \dots, s_m are the positive singular values of T.
- (d) Prove that if $k \in \{1, ..., m\}$, then e_k is an eigenvector of T^*T with corresponding eigenvalue s_k^2 .
- (e) Prove that

$$TT^*w = s_1^{\,2}\langle w, f_1\rangle f_1 + \cdots + s_m^{\,2}\langle w, f_m\rangle f_m$$

for all $w \in W$.

- 9 Suppose $T \in \mathcal{L}(V, W)$. Show that T and T^* have the same positive singular values.
- **10** Suppose $T \in \mathcal{L}(V, W)$ has singular values s_1, \dots, s_n . Prove that if T is an invertible linear map, then T^{-1} has singular values

$$\frac{1}{s_n}, ..., \frac{1}{s_1}.$$

- Suppose that $T \in \mathcal{L}(V, W)$ and v_1, \dots, v_n is an orthonormal basis of V. Let s_1, \dots, s_n denote the singular values of T.
 - (a) Prove that $||Tv_1||^2 + \dots + ||Tv_n||^2 = s_1^2 + \dots + s_n^2$.
 - (b) Prove that if W = V and T is a positive operator, then

$$\langle Tv_1, v_1 \rangle + \dots + \langle Tv_n, v_n \rangle = s_1 + \dots + s_n.$$

See the comment after Exercise 5 in Section 7A.

- 12 (a) Give an example of a finite-dimensional vector space and an operator T on it such that the singular values of T^2 do not equal the squares of the singular values of T.
 - (b) Suppose $T \in \mathcal{L}(V)$ is normal. Prove that the singular values of T^2 equal the squares of the singular values of T.
- Suppose $T_1, T_2 \in \mathcal{L}(V)$. Prove that T_1 and T_2 have the same singular values if and only if there exist unitary operators $S_1, S_2 \in \mathcal{L}(V)$ such that $T_1 = S_1 T_2 S_2$.
- **14** Suppose $T \in \mathcal{L}(V, W)$. Let s_n denote the smallest singular value of T. Prove that $s_n \|v\| \le \|Tv\|$ for every $v \in V$.
- Suppose $T \in \mathcal{L}(V)$ and $s_1 \ge \cdots \ge s_n$ are the singular values of T. Prove that if λ is an eigenvalue of T, then $s_1 \ge |\lambda| \ge s_n$.
- 16 Suppose $T \in \mathcal{L}(V, W)$. Prove that $(T^*)^{\dagger} = (T^{\dagger})^*$.

 Compare the result in this exercise to the analogous result for invertible linear maps [see 7.5(f)].
- 17 Suppose $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if T^{\dagger} is self-adjoint.

Matrices unfold Singular values gleam like stars Order in chaos shines

—written by ChatGPT with input haiku about SVD