Applications of the Singular Value Decompositi3on

Sheldon Axler

3 February 2023

SVD poem

Singular value decomposition, a tool of great might,
A method for breaking down data with insight.
From matrices to vectors, it’s a subject that’s so bright,
A fundamental tool for solving many problems with great light.

Eigenvectors and eigenvalues, so intriguing to behold,
A way to understand the behavior of a matrix to unfold,
With SVD, we can analyze and comprehend,
A tool for understanding and solving complex systems to the end.

–written by ChatGPT with input “poem about singular value decomposition”

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

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

$n = \dim V$.

$T: V \to W$ 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.

Let $s_1 \geq \cdots \geq s_m$ denote the positive singular values of T.

singular value decomposition

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 $v \in V$.

norms of linear maps

Standard definition of $\|T\|$:

$$\|T\| = \sup\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\}.$$

Pedagogical problem:

- Students may not be familiar with $\sup$.
- Replacing $\sup$ with $\max$ may not work because students may not know that a continuous function on a compact set has a maximum.

Suppose $e_1, \dots, e_m$ and $f_1, \dots, f_m$ 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 all $v \in V$.

If $v \in V$ and $\|v\| \le 1$ then

$$\begin{aligned}\|Tv\|^2 &= s_1^2\,|\langle v, e_1\rangle|^2 + \cdots + s_m^2\,|\langle v, e_m\rangle|^2 \\ &\le s_1^2\Big(|\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2\Big) \\ &\le s_1^2\,\|v\|^2 \\ &\le s_1^2,\end{aligned}$$

and thus $\|Tv\| \le s_1$.

Because $Te_1 = s_1 f_1$, we have

$$\max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\} = s_1.$$

Now we can define $\|T\|$ by

$$\begin{aligned}\|T\| &= \max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\} \\ &= \text{largest singular value of } T.\end{aligned}$$

How can we best approximate T by linear maps with lower-dimensional range?

Suppose T has singular value decomposition

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

Suppose $k \in \{1, \ldots, m-1\}$. To approximate T by linear maps whose range has dimension at most k, throw away the terms

$$s_{k+1}\langle v, e_{k+1}\rangle f_{k+1} + \cdots + s_m\langle v, e_m\rangle f_m,$$

leaving

$$s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_k\rangle f_k.$$

Let $\mathcal{L}_k(V, W)$ denote the set of linear maps $S\colon V \to W$ such that $\dim \operatorname{range} S \leq k$.

best approximation by linear map whose range has dimension $\leq k$

Suppose

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

is a singular value decomposition of T and $1 \leq k < m$. Then

$$\min\{\|T - S\| : S \in \mathcal{L}_k(V, W)\} = s_{k+1}.$$

Furthermore, if $T_k\colon V \to W$ is defined by

$$T_k v = s_1\langle v, e_1\rangle f_1 + \cdots + s_k\langle v, e_k\rangle f_k$$

for $v \in V$, then $\dim \operatorname{range} T_k \leq k$ and $\|T - T_k\| = s_{k+1}$.

approximation by lower-dimensional linear maps

best approximation by linear map whose range has dimension $\leq k$

Suppose

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

is a singular value decomposition of T and $1 \leq k < m$. Then

$$\min\{\|T - S\| : S \in \mathcal{L}_k(V, W)\} = s_{k+1}.$$

Furthermore, if $T_k \colon V \to W$ is defined by

$$T_k v = s_1\langle v, e_1\rangle f_1 + \cdots + s_k\langle v, e_k\rangle f_k$$

for $v \in V$, then $\dim \operatorname{range} T_k \leq k$ and $\|T - T_k\| = s_{k+1}$.

Proof Suppose $S \in \mathcal{L}_k(V, W)$. Hence $Se_1, \ldots, Se_{k+1}$ is linearly dependent. Thus there exist $a_1, \ldots, a_{k+1} \in \mathbf{F}$ such that

$$a_1 Se_1 + \cdots + a_{k+1} Se_{k+1} = 0$$

and $|a_1|^2 + \cdots + |a_{k+1}|^2 = 1$. We have

$$\begin{aligned}
&\|(T - S)(a_1 e_1 + \cdots + a_{k+1} e_{k+1})\|^2 \\
&\quad = \|T(a_1 e_1 + \cdots + a_{k+1} e_{k+1})\|^2 \\
&\quad = \|s_1 a_1 f_1 + \cdots + s_{k+1} a_{k+1} f_{k+1}\|^2 \\
&\quad = {s_1}^2 |a_1|^2 + \cdots + {s_{k+1}}^2 |a_{k+1}|^2 \\
&\quad \geq {s_{k+1}}^2 \left(|a_1|^2 + \cdots + |a_{k+1}|^2\right) \\
&\quad = {s_{k+1}}^2.
\end{aligned}$$

Thus $\|T - S\| \geq s_{k+1}$. ■

Let $s_1 \geq \cdots \geq s_n$ denote the singular values of T.

minimal restriction of T to subspace of dimension k

Suppose $1 \leq k \leq n$. Then

$$\min\{\|T|_U\| : U \text{ is a subspace of } V \text{ with } \dim U = k\} = s_{n-k+1}.$$

best approximation by an isometry

Suppose $\dim W \geq n$. Let $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ be orthonormal bases of V and W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

for all $v \in V$.
Define $S\colon V \to W$ by

$$Sv = \langle v, e_1 \rangle f_1 + \cdots + \langle v, e_n \rangle f_n.$$

Then

(a) S is an isometry and $\|T - S\| = \max\{|s_1 - 1|, \ldots, |s_n - 1|\}$;

(b) if $E\colon V \to W$ is an isometry, then $\|T - E\| \geq \|T - S\|$.

polar decomposition

The polar decomposition of a complex number z is

$$z = r\, e^{i\theta} = e^{i\theta} r,$$

where $\overline{e^{i\theta}}\, e^{i\theta} = 1$ and $r = |z| = \sqrt{\bar{z}\, z} \geq 0$.

A linear map $S\colon V \to V$ is called *unitary* if

$$S^* S = S S^* = I.$$

Because V is finite-dimensional, if $S\colon V \to V$ is linear then

$$S^* S = I \iff S S^* = I.$$

Thus an operator from V to V is unitary if and only if it is an isometry.

If $S\colon V \to V$ is a positive operator, then $\sqrt{S}$ denotes the unique positive operator on V such that $(\sqrt{S})^2 = S$.

Suppose $W = V$ and we have an SVD

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \ldots + s_n \langle v, e_n \rangle f_n.$$

Then

$$T^* v = s_1 \langle v, f_1 \rangle e_1 + \ldots + s_n \langle v, f_n \rangle e_n.$$

Thus

$$T^* T\, v = s_1^2 \langle v, e_1 \rangle e_1 + \ldots + s_n^2 \langle v, e_n \rangle e_n$$

and

$$\sqrt{T^* T} v = s_1 \langle v, e_1 \rangle e_1 + \ldots + s_n \langle v, e_n \rangle e_n.$$

polar decomposition

Suppose $T\colon V \to V$ is linear. Then there exists a unitary operator $S\colon V \to V$ such that
$$T = S\sqrt{T^*T}.$$

Proof Suppose we have an SVD
$$Tv = s_1\langle v, e_1\rangle f_1 + \ldots + s_n\langle v, e_n\rangle f_n$$
for all $v \in V$. Define $S\colon V \to V$ by
$$Sv = \langle v, e_1\rangle f_1 + \ldots + \langle v, e_n\rangle f_n.$$
Then S is an isometry and hence is unitary.

We have
$$Se_k = f_k$$
for $k = 1, \ldots, n$ and
$$\sqrt{T^*T}v = s_1\langle v, e_1\rangle e_1 + \ldots + s_n\langle v, e_n\rangle e_n.$$

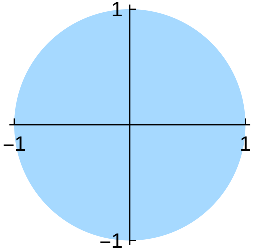
Thus
$$\begin{aligned} S(\sqrt{T^*T}v) &= s_1\langle v, e_1\rangle f_1 + \ldots + s_n\langle v, e_n\rangle f_n \\ &= Tv. \quad \blacksquare \end{aligned}$$

Note that if $T = S\sqrt{T^*T}$ where S is as in the proof above, then S is a best approximation to T among the unitary operators.

definition: *ball*, $B(r)$

The *ball* in V centered at 0 with radius 1, denoted B, is defined by

$$B = \{v \in V : \|v\| < 1\}.$$

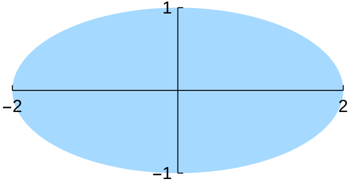


The ball B in $\mathbf{R}^2$.

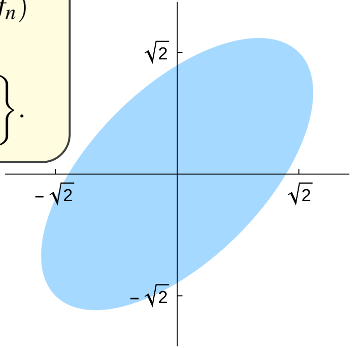
definition: *ellipsoid*, $E(s_1 f_1, \ldots, s_n f_n)$, *principal axes*

Suppose that $f_1, \ldots, f_n$ is an orthonormal basis of V and $s_1, \ldots, s_n$ are positive numbers. The *ellipsoid* $E(s_1 f_1, \ldots, s_n f_n)$ with *principal axes* $s_1 f_1, \ldots, s_n f_n$ is defined by

$$E(s_1 f_1, \ldots, s_n f_n) = \left\{ v \in V : \frac{|\langle v, f_1 \rangle|^2}{s_1^2} + \cdots + \frac{|\langle v, f_n \rangle|^2}{s_n^2} < 1 \right\}.$$

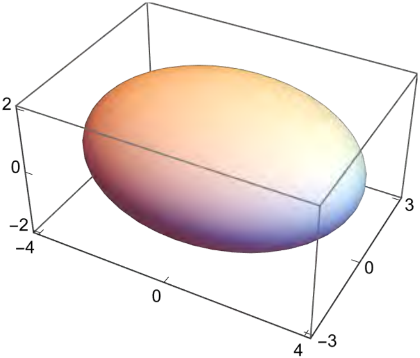


The ellipsoid $E(2f_1, f_2)$ *in* $\mathbf{R}^2$*, where* f_1, f_2 *is the standard basis of* $\mathbf{R}^2$*.*



The ellipsoid $E(2f_1, f_2)$ *in* $\mathbf{R}^2$*, where* $f_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ *and* $f_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$*.*

operators applied to ellipsoids



The ellipsoid $E(4f_1, 3f_2, 2f_3)$ *in* $\mathbf{R}^3$*,*
where f_1, f_2, f_3 *is the standard basis of* $\mathbf{R}^3$*.*

operators applied to ellipsoids

invertible operator takes ball to ellipsoid

Suppose $T\colon V \to V$ is invertible. Then T maps the ball B in V onto an ellipsoid in V.

Proof   Suppose T has SVD

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for all $v \in V$. We will show that

$$T(B) = E(s_1 f_1, \ldots, s_n f_n).$$

First suppose $v \in B$. Then

$$\frac{|\langle Tv, f_1\rangle|^2}{s_1^2} + \cdots + \frac{|\langle Tv, f_n\rangle|^2}{s_n^2} = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 < 1.$$

Thus $Tv \in E(s_1 f_1, \ldots, s_n f_n)$, and hence $T(B) \subset E(s_1 f_1, \ldots, s_n f_n)$.

To prove inclusion in the other direction, suppose $w \in E(s_1 f_1, \ldots, s_n f_n)$. Let

$$v = \frac{\langle w, f_1\rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_n\rangle}{s_n} e_n.$$

Then $\|v\| < 1$; hence $v \in B$. Because $Te_k = s_k f_k$, we have

$$\begin{aligned} Tv &= \langle w, f_1\rangle f_1 + \cdots + \langle w, f_n\rangle f_n \\ &= w. \end{aligned}$$

Thus $T(B) \supset E(s_1 f_1, \ldots, s_n f_n)$. ∎

invertible operator takes ellipsoids to ellipsoids

Suppose $T\colon V \to V$ is invertible and E is an ellipsoid in V. Then $T(E)$ is an ellipsoid in V.

Proof Let $S\colon V \to V$ be an invertible operator such that $E = S(B)$. Then

$$\begin{aligned} T(E) &= T\bigl(S(B)\bigr) \\ &= (TS)(B). \end{aligned}$$

By our previous result, the invertible linear map TS maps the ball to an ellipsoid. ∎

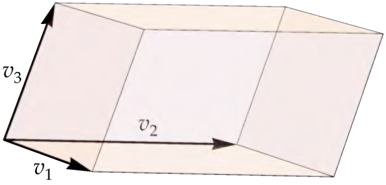
operators applied to parallelepipeds

definition: $P(v_1,\ldots,v_n)$, *parallelepiped*

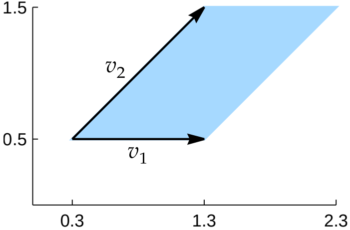
Suppose $v_1,\ldots,v_n$ is a basis of V. Let

$$P(v_1,\ldots,v_n) = \{a_1v_1+\cdots+a_nv_n : a_1,\ldots,a_n \in (0,1)\}.$$

A *parallelepiped* is a set of the form $v + P(v_1,\ldots,v_n)$ for some $v \in V$. The vectors $v_1,\ldots,v_n$ are called the *edges* of this parallelepiped.



A parallelepiped in $\mathbf{R}^3$.



The parallelepiped

$$(0.3, 0.5) + P\big((1,0),(1,1)\big)$$

in $\mathbf{R}^2$.

invertible operator takes parallelepipeds to parallelepipeds

Suppose $v \in V$ and $v_1, \ldots, v_n$ is a basis of V. Suppose $T\colon V \to V$ is invertible. Then

$$T\big(v + P(v_1, \ldots, v_n)\big) = Tv + P(Tv_1, \ldots, Tv_n).$$

Proof Because T is invertible, the list $Tv_1, \ldots, Tv_n$ is a basis of V. The linearity of T implies that

$$T(v + a_1v_1 + \cdots + a_nv_n) = Tv + a_1Tv_1 + \cdots + a_nTv_n$$

for all $a_1, \ldots, a_n \in (0, 1)$. Thus

$$T\big(v + P(v_1, \ldots, v_n)\big) = Tv + P(Tv_1, \ldots, Tv_n),$$

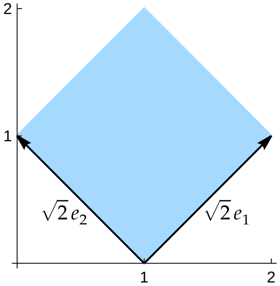
as desired. ∎

definition: *box*

A *box* in V is a set of the form

$$v + P(r_1 e_1, \dots, r_n e_n),$$

where $v \in V$ and $r_1, \dots, r_n$ are positive numbers and $e_1, \dots, e_n$ is an orthonormal basis of V.

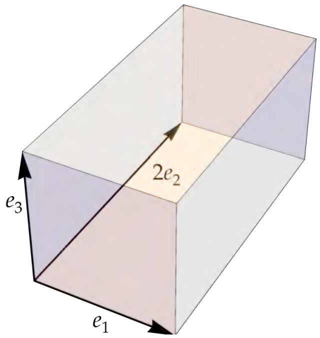


The box

$$(1, 0) + P(\sqrt{2}\, e_1, \sqrt{2}\, e_2),$$

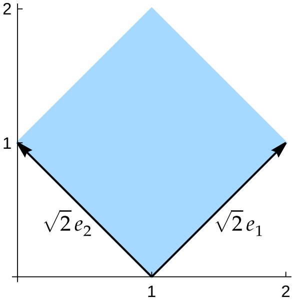
where

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

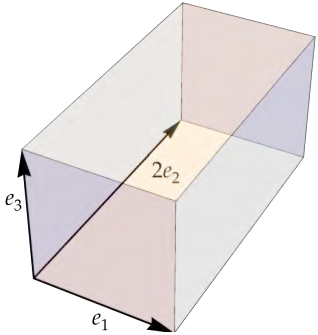


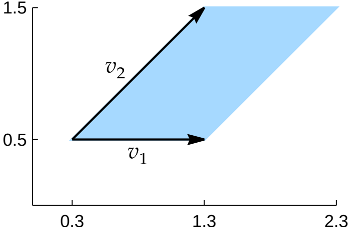
The box $P(e_1, 2e_2, e_3)$*, where* e_1, e_2, e_3 *is the standard basis of* $\mathbf{R}^3$.

operators applied to parallelepipeds

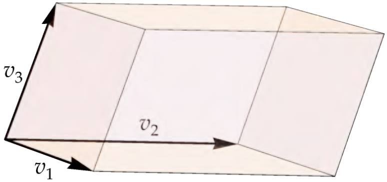


boxes





parallelepipeds



operators applied to parallelepipeds

each operator takes some boxes to boxes

Suppose $T\colon V \to V$ is invertible. Suppose T has singular value decomposition

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n,$$

where $e_1, \dots, e_n$ and $f_1, \dots, f_n$ are orthonormal bases of V and the equation above holds for all $v \in V$. Then T maps the box $v + P(r_1e_1, \dots, r_ne_n)$ onto the box $Tv + P(r_1s_1f_1, \dots, r_ns_nf_n)$ for all positive numbers $r_1, \dots, r_n$ and all $v \in V$.

Proof If $a_1, \dots, a_n \in (0, 1)$ and $r_1, \dots, r_n$ are positive numbers and $v \in V$, then

$$T(v + a_1r_1e_1 + \cdots + a_nr_ne_n) = Tv + a_1r_1s_1f_1 + \cdots + a_nr_ns_nf_n.$$

Thus $T\big(v + P(r_1e_1, \dots, r_ne_n)\big) = Tv + P(r_1s_1f_1, \dots, r_ns_nf_n)$. ■

volume via singular values

For this topic, assume $\mathbf{F} = \mathbf{R}$.

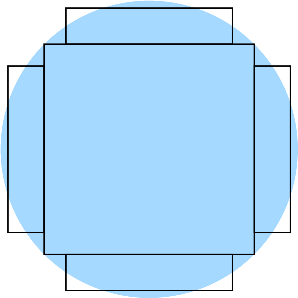
definition: *volume of a box*

If $v \in V$ and $r_1, \dots, r_n$ are positive numbers and $e_1, \dots, e_n$ is an orthonormal basis of V, then

$$\operatorname{volume}\bigl(v + P(r_1 e_1, \dots, r_n e_n)\bigr) = r_1 \times \cdots \times r_n.$$

definition: *volume*

Suppose $\Omega \subset V$. Then the *volume* of Ω, denoted $\operatorname{volume} \Omega$, is approximately the sum of the volumes of a collection of disjoint boxes that approximate Ω.



Volume of this ball ≈ sum of the volumes of the five boxes.

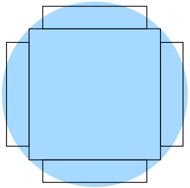
Example: Suppose $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ is defined by

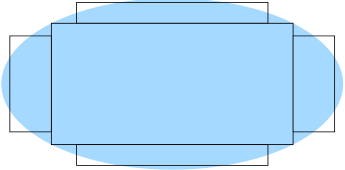
$$Tv = 2\langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2,$$

where e_1, e_2 is the standard basis of $\mathbf{R}^2$.
T stretches by a factor of 2 along the e_1 axis.
The ball gets mapped by T to the ellipsoid.

The five boxes in the top figure get mapped to boxes with twice the width and the same height. Hence each box in the top figure gets mapped to a box with twice the volume (area). The sum of the volumes of the five new boxes approximates the volume of the ellipsoid.

Thus T changes the volume of the ball by a factor of 2.





volume via singular values

volume changes by a factor of the product of the singular values

Suppose $T\colon V \to V$ is invertible and $\Omega \subset V$. Then

$$\operatorname{volume} T(\Omega) = (\text{product of singular values of } T)(\operatorname{volume} \Omega).$$

Proof Suppose T has singular value decomposition

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n,$$

where $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ are orthonormal bases of V.
Approximate Ω by boxes of the form $v + P(r_1 e_1, \ldots, r_n e_n)$, which have volume $r_1 \times \cdots \times r_n$. The operator T maps each box $v + P(r_1 e_1, \ldots, r_n e_n)$ onto the box $Tv + \mathcal{P}(r_1 s_1 f_1, \ldots, r_n s_n f_n)$, which has volume $(s_1 \times \cdots \times s_n)(r_1 \times \cdots \times r_n)$.
Because T changes the volume of each box in a collection that approximates Ω by a factor of $s_1 \times \cdots \times s_n$, the linear map T changes the volume of Ω by the same factor. ■

volume via singular values

product of singular values of $T = |\det T|$

The product of the singular values of T equals $|\det T|$.

Proof By the polar decomposition, there is a unitary operator $S\colon V \to V$ such that

$$T = S\sqrt{T^*T}.$$

Thus

$$\begin{aligned} |\det T| &= |\det S| \det \sqrt{T^*T} \\ &= \det \sqrt{T^*T} \\ &= \text{product of singular values of } T, \end{aligned}$$

as desired. ∎

SVD formula for pseudoinverse

$T: V \to W$ is linear.

$T|_{(\text{null}\, T)^\perp}$ is a one-to-one map of $(\text{null}\, T)^\perp$ onto $\text{range}\, T$.

Let $P_{\text{range}\, T}$ denote the orthogonal projection of W onto $\text{range}\, T$.

definition: *pseudoinverse, $T^\dagger$*

The *pseudoinverse* $T^\dagger: W \to V$ of T is the linear map from W to V defined by

$$T^\dagger w = (T|_{(\text{null}\, T)^\perp})^{-1} P_{\text{range}\, T}\, w$$

for $w \in W$.

Given $b \in W$, find $x \in V$ such that $Tx = b$.

If T is invertible, $x = T^{-1}b$. But perhaps no solutions; perhaps ∞ many solutions.

pseudoinverse is best solution

Suppose $b \in W$.

(a) If $x \in V$, then

$$\|T(T^\dagger b) - b\| \le \|Tx - b\|,$$

with equality if and only if $x \in T^\dagger b + \text{null}\, T$.

(b) If $x \in T^\dagger b + \text{null}\, T$, then

$$\|T^\dagger b\| \le \|x\|,$$

with equality if and only if $x = T^\dagger b$.

SVD formula for pseudoinverse

formula for pseudoinverse

Suppose $s_1, \ldots, s_m$ are the positive eigenvalues 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

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

for every $w \in W$.

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.