18.700 - LINEAR ALGEBRA, DAY 19
ISOMETRIES, UNITARY OPERATORS, AND MATRIX FACTORIZATIONS
SINGULAR VALUE DECOMPOSITION

SAM SCHIAVONE

CONTENTS

I. PRE-CLASS PLANNING

I.1. Goals for lesson.

(1) Students will learn the defintion of isometries and unitary operators.
(2) Students will learn QR Decomposition.
(3) Students will learn Singular Value Decomposition.

I.2. Methods of assessment.

(1) Student responses to questions posed during lecture
(2) Student responses to worksheet

I.3. Materials to bring.

(1) Laptop + adapter (2) Worksheets (3) Chalk

(0:00)

II. Lesson Plan

II.1. **Last time.**

- Proved the Spectral Theorem over $\mathbb{R}$.

  Theorem 1 (Spectral Theorem over $\mathbb{R}$). *Suppose $\mathbb{F} = \mathbb{R}$ and $T \in \mathcal{L}(V)$. TFAE.*
  (i) T is self-adjoint.
  (ii) There is an orthonormal basis of V of eigenvectors of T.

- Defined positive linear operators.
- Proved properties of positive linear operators.
- Defined of isometries and unitary operators.

II.2. **7D: Isometries, Unitary Operators, and Matrix Factorizations.** Throughout today, let V and W be nonzero finite-dimensional inner product spaces.

Definition 2. A linear map $S \in \mathcal{L}(V, W)$ is an *isometry* if

$$\|S(v)\| = \|v\|$$

for all $v \in V$.

Theorem 3. *Let $S \in \mathcal{L}(V, W)$, and let $\mathcal{E} := (e_1, \dots, e_n)$ and $\mathcal{F} := (f_1, \dots, f_m)$ be orthonormal bases for V and W, respectively. TFAE.*

(a) S is an isometry.
(b) $S^*S = I$.
(c) S preserves inner products, i.e.,

$$\langle S(u), S(v)\rangle = \langle u, v\rangle$$

for all $u, v \in V$.
(d) $S(e_1), \dots, S(e_n)$ is an orthonormal list in W.
(e) The columns of ${}_{\mathcal{F}}[S]_{\mathcal{E}}$ form an orthonormal list in $\mathbb{F}^m$ with respect to the usual inner product.

Lemma 4. *Let $T \in \mathcal{L}(V)$ be self-adjoint. If $\langle T(v), v\rangle = 0$ for all $v \in V$, then $T = 0$.*

Proof. Given $v \in V$, let $u = v + T(v)$. Then

$$\begin{aligned} 0 = \langle T(u), u\rangle &= \langle T(v + T(v)), v + T(v)\rangle = \langle T(v) + T^2(v), v + T(v)\rangle \\ &= \cancelto{0}{\langle T(v), v\rangle} + \langle T(v), T(v)\rangle + \langle T^2(v), v\rangle + \langle T^2(v), T(v)\rangle \\ &= \langle T(v), T(v)\rangle + \langle T(v), T(v)\rangle + \cancelto{0}{\langle T(T(v)), T(v)\rangle} = 2\|T(v)\|^2 . \end{aligned}$$

Thus $T(v) = 0$. Since v was arbitrary, then $T = 0$. ∎

Proof of Theorem. (a) $\implies$ (b): Assume S is an isometry. Given $v \in V$, then

$$\begin{aligned} \langle (I - S^*S)(v), v\rangle &= \langle v, v\rangle - \langle S^*S(v), v\rangle = \|v\|^2 - \langle S(v), S(v)\rangle = \|v\|^2 - \|S(v)\|^2 \\ &= \|v\|^2 - \|v\|^2 = 0 . \end{aligned}$$

By the lemma, then $I - S^*S = 0$, so $S^*S = I$.

(b) $\implies$ (c): Assume $S^*S = I$. Given $u, v \in V$, then

$$\langle S(u), S(v) \rangle = \langle S^*S(u), v \rangle = \langle I(u), v \rangle = \langle u, v \rangle .$$

(c) $\implies$ (d): Assume $\langle S(u), S(v) \rangle = \langle u, v \rangle$ for all $u, v \in V$. Then

$$\langle S(e_j), S(e_k) \rangle = \langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise} \end{cases}$$

for each $j, k \in \{1, \ldots, n\}$.

(d) $\implies$ (e): Assume $S(e_1), \ldots, S(e_n)$ is an orthonormal list. Let $A = {}_{\mathcal{F}}[S]_{\mathcal{E}}$. Then

$$\langle A_{\cdot,j}, A_{\cdot,k} \rangle = \sum_{i=1}^{m} A_{i,j} \overline{A_{i,k}} = \left\langle \sum_{i=1}^{m} A_{i,j} f_i, \sum_{i=1}^{m} A_{i,k} f_i \right\rangle = \langle S(e_j), S(e_k) \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where the second equality follows from the Pythagorean Theorem.

(e) $\implies$ (a): Assume the columns of ${}_{\mathcal{F}}[S]_{\mathcal{E}}$ form an orthonormal list. Given $v \in V$, then

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n ,$$

so

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2$$

by the Pythagorean Theorem. By a similar calculation to (5), then $S(e_1), \ldots, S(e_n)$ is an orthonormal list. Then

$$S(v) = S(\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n) = \langle v, e_1 \rangle S(e_1) + \cdots + \langle v, e_n \rangle S(e_n)$$

so

$$\|S(v)\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 = \|v\|^2 .$$

Thus $\|S(v)\| = \|v\|$, so S is an isometry. □

Definition 6. An operator $S \in \mathcal{L}(V)$ is *unitary* if S is an invertible isometry.

Theorem 7. *Let $S \in \mathcal{L}(V)$, and let $\mathcal{E} := (e_1, \ldots, e_n)$ be an orthonormal basis of V. TFAE.*

(a) S is a unitary operator.
(b) $S^*S = SS^* = I$.
(c) S is invertible and $S^{-1} = S^*$.
(d) $S(e_1), \ldots, S(e_n)$ is an orthonormal basis of V.
(e) The rows of $[S]_{\mathcal{E}}$ form an orthonormal basis of $\mathbb{F}^n$.
(f) S^* is a unitary operator.

Proof. Similar to the previous theorem. See 7.53 in text for details. □

Proposition 8. *Let $S \in \mathcal{L}(V)$ be a unitary operator and suppose λ is an eigenvalue of S. Then $|\lambda| = 1$.*

Proof. Let $0 \neq v \in V$ be a corresponding eigenvector. Then

$$|\lambda| \|v\| = \|\lambda v\| = \|S(v)\| = \|v\| .$$

Since $v \neq 0$, then $\|v\| \neq 0$. Dividing, then $|\lambda| = 1$. □

Definition 9. A matrix $Q \in M_{n\times n}(\mathbb{F})$ is *unitary* if the associated linear operator

$$\begin{aligned} L_Q : \mathbb{F}^n &\to \mathbb{F}^n \\ v &\mapsto Qv \end{aligned}$$

is unitary. Equivalently, if the columns of Q form an orthonormal basis of $\mathbb{F}^n$.

Theorem 10 (QR Factorization). *Suppose $A \in M_{n\times n}(\mathbb{F})$ is a square matrix with linearly independent columns. Then there exist unique matrices $Q, R \in M_{n\times n}(\mathbb{F})$ such that*

(i) Q is unitary;
(ii) R is upper triangular with positive diagonal entries; and
(iii) $A = QR$.

Proof. This follows from a matrix interpretation of the Gram-Schmidt procedure. Let $v_1, \dots, v_n$ be the columns of A. Let $e_1, \dots, e_n$ be the orthonormal list resulting from the Gram-Schmidt procedure, and let Q be the matrix whose columns are $e_1, \dots, e_n$. The equations

$$\begin{aligned} f_k &:= v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, f_j\rangle}{\|f_j\|^2} f_j \\ e_k &:= \frac{1}{\|f_k\|} f_k \end{aligned}$$

give the entries for the upper triangular matrix R^{-1} such that $Q = AR^{-1}$. Details left as an exercise. □

II.3. **7E: Singular Value Decomposition (SVD).** Given a linear map $T : V \to W$, SVD will provide orthonormal bases $\mathcal{E}$ of V and $\mathcal{F}$ of W such that ${}_{\mathcal{F}}[T]_{\mathcal{E}}$ is particularly simple.

Proposition 11. *Let $T \in \mathcal{L}(V, W)$. Then*

(a) T^*T is positive;
(b) $\ker(T^*T) = \ker(T)$;
(c) $\operatorname{img}(T^*T) = \operatorname{img}(T^*)$;
(d) $\dim(\operatorname{img}(T)) = \dim(\operatorname{img}(T^*)) = \dim(\operatorname{img}(T^*T))$.

Proof. (a) [Ask students for the two conditions.] Since

$$(T^*T)^* = T^*(T^*)^* = T^*T\,,$$

then T^*T is self-adjoint. Given $v \in V$, then

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

so T^*T is positive.

(b) ($\supseteq$): $T(v) = 0 \implies T^*(T(v)) = T^*(0) = 0$.

($\subseteq$): Given $v \in \ker(T^*T)$, then $T^*T(v) = 0$. <u>Goal</u>: $T(v) = 0 \iff \|T(v)\|^2 = 0$.

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

(c) Recall that $\operatorname{img}(S) = \ker(S^*)^\perp$ for all $S \in \mathcal{L}(V, W)$. Since T^*T is self-adjoint, then

$$\operatorname{img}(T^*T) = \ker(T^*T)^\perp = \ker(T)^\perp = \operatorname{img}(T^*)$$

where the second equality comes from part (b).

(d)

$$\dim(\text{img}(T)) = \dim(\ker(T^*)^\perp) = \dim(W) - \dim(\ker(T^*)) = \dim(\text{img}(T^*)$$

by facts about orthogonal complements, and Rank-Nullity.

□

Definition 12. Let $T \in \mathcal{L}(V, W)$. The *singular values* of T are the nonnegative square roots of the eigenvalues of T^*T, listed in decreasing order, and with multiplicity.

Example 13. Defined $T \in \mathcal{L}(\mathbb{F}^4)$ by $T(v) = Av$ where

$$A := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

[Compute A^* and A^*A. Should get

$$A^*A = \begin{pmatrix} 9 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}.]$$

Thus the eigenvalues of T^*T are $9, 9, 4, 0$, so the singular values of T are $3, 3, 2, 0$.

Proposition 14. *Let $T \in \mathcal{L}(V, W)$.*

(a) T is injective iff 0 is not a singular value of T.
(b) The number of positive (i.e., nonzero) singular values of T equals $\dim(\text{img}(T))$.

Proof. (a) T is injective iff [ask students]

$$0 = \ker(T) = \ker(T^*T)$$

iff T^*T is injective iff 0 is not an eigenvalue of T^*T iff 0 is not a singular value of T. (Here we used the previous proposition for the second equality.)

(b) Exercise.

□

Proposition 15. *Suppose $S \in \mathcal{L}(V, W)$. Then S is an isometry iff all its singular values are 1.*

Proof. (⇒): Assume S is an isometry. Then $S^*S = I$, so all the eigenvalues of S^*S are 1. Thus all the singular values of S are 1.

(⇐): Exercise. □

Theorem 16 (Singular Value Decomposition (SVD))**.** *Suppose $T \in \mathcal{L}(V, W)$. Let $s_1, \ldots, s_n$ be the singular values of T, and let $s_1, \ldots, s_m$ be the positive ones. Then there exist orthonormal lists $e_1, \ldots, e_m$ in V and $f_1, \ldots, f_m$ in W such that*

$$T(v) = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for all $v \in V$.

Proof. Since T^*T is positive, then by the Spectral Theorem there is an orthonormal basis $e_1, \dots, e_n$ of V such that

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

for each $k = 1, \dots, n$. Now, for each $k = 1, \dots, m$, define

$$f_k := \frac{T(e_k)}{s_k}.$$

We show that $f_1, \dots, f_m$ is orthonormal:

$$\begin{aligned}\langle f_j, f_k \rangle &= \left\langle \frac{1}{s_j} T(e_j), \frac{1}{s_k} T(e_k) \right\rangle = \frac{1}{s_j s_k} \langle T(e_j), T(e_k) \rangle = \frac{1}{s_j s_k} \langle e_j, T^*T(e_k) \rangle = \frac{1}{s_j s_k} \langle e_j, s_k^2 e_k \rangle \\ &= \frac{s_k}{s_j} \langle e_j, e_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k \end{cases}\end{aligned}$$

for all $j, k \in \{1, \dots, m\}$.

Note that for $k > m$ we have

$$T^*T(e_k) = s_k^2 e_k = 0 \implies e_k \in \ker(T^*T) = \ker(T) \implies T(e_k) = 0$$

by a previous result. Given $v \in V$, since $e_1, \dots, e_n$ is orthonormal, then

$$\begin{aligned}T(v) &= T(\langle v, e_1 \rangle e_1 + \cdots \langle v, e_n \rangle e_n = \langle v, e_1 \rangle T(e_1) + \cdots \langle v, e_m T(e_m) \rangle \\ &= \langle v, e_1 \rangle s_1 f_1 + \cdots + \langle v, e_m \rangle s_m f_m.\end{aligned}$$

□

Proposition 17 (SVD of adjoint). *Suppose $T \in \mathcal{L}(V, W)$ and $s_1, \dots, s_m$, $e_1, \dots, e_m$, and $f_1, \dots, f_m$ are as before, so*

$$T(v) = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for all $v \in V$. Then

$$T^*(w) = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle v, e_m \rangle e_m.$$

Definition 18. Let $A \in M_{m \times n}(\mathbb{F})$. A is *(generalized) diagonl* if $A_{ij} = 0$ for all $i, j \in \{1, \dots, \min(m, n)\}$ with $i \neq j$.

Theorem 19 (SVD, matrix version). *Let $A \in M_{m \times n}(\mathbb{F})$ have rank r. Then there exists*

- *a generalized diagonal matrix Σ whose diagonal entries are the positive singular values of A;*
- *a unitary matrix $U \in M_{m \times m}(\mathbb{F})$; and*
- *a unitary matrix $V \in M_{n \times n}(\mathbb{F})$*

such that $A = U \Sigma V^t$.

Proof sketch. Let $v_1, \dots, v_n$ be an orthonormal basis of $\mathbb{F}^n$ consisting of eigenvectors of A^*A. Let V be the matrix whose columns are $v_1, \dots, v_n$.

Suppose A has r nonzero singular values. Then one can show that $Av_1, \dots, Av_r$ is an orthogonal basis for $\mathrm{Col}(A)$. Normalize this to obtain an orthonormal basis $u_1, \dots, u_r$ of $\mathrm{Col}(A)$, where

$$u_i := \frac{1}{s_i} A v_i$$

for $i = 1, \dots, r$. Extend this to a basis $u_1, \dots, u_m$ of $\mathbb{F}^m$. Let U be the matrix with columns $u_1, \dots, u_m$. We claim that

$$A = U\Sigma V^* .$$

□

II.4. **Worksheet.**

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix} \implies A^*A = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix} .$$

Then $\operatorname{minpoly}(A^*A) = x^2 - 34x + 225 = (x-9)(x-25)$.

$$V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

$$Av_1 = \begin{pmatrix} 5/\sqrt{2} \\ 5/\sqrt{2} \\ 0 \end{pmatrix} \qquad Av_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 2\sqrt{2} \end{pmatrix}$$

$$u_1 = \frac{1}{5}Av_1$$
$$u_2 = \frac{1}{3}Av_2$$
$$u_3 = \begin{pmatrix} -2/3 \\ 2/3 \\ 1/3 \end{pmatrix}$$

https://sagecell.sagemath.org/?z=eJxtj0FugzAQRfdI3MFL2x0g2M2uqQQ3oBVsLLeyEpqgNqEFy-X4H
vZ_M_MnIztyNrZrBloUoJQEAUKDEiAhElqzMMgQyWLbmUv_2fY1xa_-1H7TDEX-3-PZ1c3RDYOn-GF4RKYolA6DM
5QO5IHh9qS5mnXYqoqz2k0qT_6iwVDMgsxzhOrEPRDfWbkLt0oQXq6m9f5VJwQl8_q5UNKxTloiBNtnzsOSaSoxI
E6DdmC1Avf33Fvzf6dTrM2sNFsauaBuTP3T05v1mHsBwcyiHY=&lang=sage&interacts=eJyLjgUAARUAuQ=
=