Linear Algebra I

Chapter 5. Inner Product Spaces

Boqing Xue

WELCOME!

2023 Fall

上海科技大学 数学科学研究所
Institute of Mathematical Sciences, ShanghaiTech University

Chapter 5. Inner Product Spaces

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

§5.9 Polar Decomposition*

A Quick Review

- Fact: Every real n-dimensional vector space V is isomorphic to $\mathbb{R}^n$.

Observation: Vectors in $\mathbb{R}^n$ have norm, distance, angle, orthogonality, etc.
Question: What about vectors in V?

- Dot product in $\mathbb{R}^n$: $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$.
- Norm in $\mathbb{R}^n$: $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$.
- Distance in $\mathbb{R}^n$: $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$.
- Angle in $\mathbb{R}^n$: $\cos\theta = \dfrac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \, \|\mathbf{v}\|}$.

General Definition of Inner Product

Definition. Let V be an $\mathbb{F}$-vector space. An inner product on V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ such that the following axioms are satisfied.

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ $\quad (\forall \mathbf{u}, \mathbf{v} \in V)$.
(ii) $\langle k\mathbf{u} + \ell \mathbf{v}, \mathbf{w} \rangle = k\langle \mathbf{u}, \mathbf{w} \rangle + \ell \langle \mathbf{v}, \mathbf{w} \rangle$ $\quad (\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V, \forall k, l \in \mathbb{R})$.
(iii) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ $\quad (\forall \mathbf{v} \in V)$. And $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = 0$.

In this case, we call $(V, \langle \cdot, \cdot \rangle)$ a inner product space.

Remark: When $\mathbb{F} = \mathbb{R}$, we may ignore the "conjugate".
Remark: When $\mathbb{F} = \mathbb{R}$, we also call it real inner product space;
When $\mathbb{F} = \mathbb{C}$, we also call it unitary space.

Remark: One always has $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{0} \rangle = 0$.
Remark: By (i) and (ii), we have

$$\langle \mathbf{w}, k\mathbf{u} + l\mathbf{v} \rangle = \overline{k}\langle \mathbf{w}, \mathbf{u} \rangle + \overline{l}\langle \mathbf{w}, \mathbf{v} \rangle.$$

Examples

Example. Let $\mathbb{F} = \mathbb{R}$ and $V = \mathbb{R}^n$. For $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$, define

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + \ldots + u_n v_n.$$

Then $(V, \langle \cdot, \cdot \rangle)$ forms a real inner product space.

Example. Let $\mathbb{F} = \mathbb{C}$ and $V = \mathbb{C}^n$. For $\mathbf{z} = (z_1, \ldots, z_n)$ and $\mathbf{v} = (w_1, \ldots, w_n)$, define

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1 \overline{w_1} + \ldots + z_n \overline{w_n}.$$

Then $(V, \langle \cdot, \cdot \rangle)$ forms a real inner product space.

Examples

Example. Let $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2)$. Prove that $\langle \cdot, \cdot \rangle$ given below is a real inner products on $\mathbb{R}^2$: $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_1 v_2 + u_2 v_1 + 2 u_2 v_2$.

Proof:

①

Example. Let $A \in M_n(\mathbb{R})$ be invertible. Suppose that $\langle \cdot, \cdot \rangle$ is a real inner product on $\mathbb{R}^n$. Prove that $\langle \cdot, \cdot \rangle_A$ defined by (view $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ as column matrices)

$$\langle \mathbf{u}, \mathbf{v} \rangle_A = \langle A\mathbf{u}, A\mathbf{v} \rangle, \quad (\mathbf{u}, \mathbf{v} \in \mathbb{R}^n)$$

is also a real inner product on $\mathbb{R}^n$.

Proof:

②

Norm and Distance

Definition. If $(V, \langle \cdot, \cdot \rangle)$ is an inner product space, then the norm of a vector $\mathbf{v} \in V$ is

$$\|\mathbf{v}\| = \sqrt{\langle v, v \rangle}.$$

The distance between two vectors is

$$d(u, v) = \|u - v\|.$$

Vectors with norm 1 are called unit vectors.

Example. In $\mathbb{R}^2$, let $\mathbf{u} = (1, 0)$ and $\mathbf{v} = (0, 1)$. Evaluate $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ and $d(\mathbf{u}, \mathbf{v})$ with the following two inner products, respectively.
(a) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2$;
(b) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_1 v_2 + u_2 v_1 + 2 u_2 v_2$.

Solution:

3

Basic Properties

Theorem. If $\mathbf{u}, \mathbf{v}$ are vectors in an inner product space, and if k is a scalar, then

- ◇ $\|\mathbf{v}\| \geq 0$ with equality if and only if $\mathbf{v} = 0$.
- ◇ $\|k\mathbf{v}\| = |k| \cdot \|\mathbf{v}\|$.
- ◇ $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.
- ◇ $d(\mathbf{u}, \mathbf{v}) \geq 0$ with equality if and only if $\mathbf{v} = \mathbf{u}$.

Parallelogram Equation & Polarization identity

Theorem. (Parallelogram Equation) If $\mathbf{u}, \mathbf{v}$ are vectors in an inner product space, then

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\right)$$

Theorem. (Polarization identity) Let $\mathbf{u}, \mathbf{v}$ are vectors in an inner product space. When $\mathbb{F} = \mathbb{R}$, we have

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4} \sum_{k=0}^{1} \left\|\mathbf{u} + (-1)^k \mathbf{v}\right\|^2 = \frac{1}{4}\left(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2\right).$$

When $\mathbb{F} = \mathbb{C}$, it satisfies that

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle &= \frac{1}{4} \sum_{k=0}^{3} \|\mathbf{u} + \mathrm{i}^k \mathbf{v}\|^2 \\ &= \frac{1}{4}\left(\|\mathbf{u}+\mathbf{v}\|^2 + \mathrm{i}\|\mathbf{u}+\mathrm{i}\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2 - \mathrm{i}\|\mathbf{u}-\mathrm{i}\mathbf{v}\|^2\right)\end{aligned}$$

Examples

Example. Show that the following $\langle\cdot,\cdot\rangle$ is an inner products on V, respectively. And compute the inner products, norms and the distances.

(1) $\mathbb{F}=\mathbb{R}$, $V=C[-\pi,\pi]$, $\langle f,g\rangle=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)g(t)dt$.

Let $f(t)=t$, $g(t)=t^2$. Evaluate $\|f\|$, $\|g\|$ and $\langle f,g\rangle$.

(2) $\mathbb{F}=\mathbb{R}$, $V=M_n(\mathbb{R})$, $\langle A,B\rangle=\mathrm{tr}(B^TA)$.

Let $A=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$, $B=\begin{bmatrix}-1 & 0\\ 3 & 2\end{bmatrix}$. Evaluate $\langle A,B\rangle$, $\|A\|$ and $d(A,B)$.

Solution:

4

Problem*: (1)' $\mathbb{F}=\mathbb{C}$, $V=C_{\mathbb{C}}[-\pi,\pi]$ (continuous complex-valued functions),

$$\langle f,g\rangle=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\overline{g(t)}dt.$$

Problem*: (2)' $\mathbb{F}=\mathbb{C}$, $V=M_n(\mathbb{C})$, $\langle A,B\rangle=\mathrm{tr}(B^*A)$.

Examples

Example. Show that the following $\langle\cdot,\cdot\rangle$ is an inner products on P_n (with real coefficinets), respectively. And compute the inner products, norms and the distances.

(3) $\mathbb{F}=\mathbb{R}$, $V=P_n$, $\left\langle \sum_{k=0}^n a_k x^k, \sum_{k=0}^n b_k x^k \right\rangle = \sum_{k=0}^n a_k b_k$.

(4) $\mathbb{F}=\mathbb{R}$, $V=P_n$, given distinct $x_0, x_1, \ldots, x_n \in \mathbb{R}$,

$$\langle p, q\rangle = \sum_{j=0}^n p(x_j)q(x_j) \text{ (called evaluation inner product).}$$

In (3) and (4), let $n=2$, $x_0=-2$, $x_1=0$, $x_2=2$. Take $p(x)=x^2$ and $q(x)=1+x$. Evaluate $\langle p,q\rangle$ and $\|p\|$.

Solution:

5

Problem*: (3)' $\mathbb{F}=\mathbb{R}$, $V=P_n$ (with complex coefficients), $\left\langle \sum_{k=0}^n a_k x^k, \sum_{k=0}^n b_k x^k \right\rangle = \sum_{k=0}^n a_k \overline{b_k}$.

Examples

Example.
Show that the following $\langle \cdot, \cdot \rangle$ is an inner products on V.

(5) $V = \mathbb{R}^\infty = \{(x_1, x_2, \ldots, x_n, \ldots) : \sum_{i=1}^\infty x_i^2 < \infty\}$,
$\langle (x_i), (y_i) \rangle = \sum_{i=1}^\infty x_i y_i$.

Proof:

6

Self-adjoint Hermitian Matrix

Definition. Let $\mathbf{u}, \mathbf{v} \in M_{n\times 1}(\mathbb{C})$ be column vectors. Then the (complex) inner product of $\mathbf{u}$ and $\mathbf{v}$ is given by

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^* \mathbf{u}.$$

Example. Suppose that $A \in M_{m\times n}(\mathbb{C})$, $u \in M_{n\times 1}(\mathbb{C})$ and $v \in M_{m\times 1}(\mathbb{C})$. Then $A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^*\mathbf{v}$.

Definition. A matrix $A \in M_n(\mathbb{C})$ is said to be self-adjoint/Hermitian if $A^* = A$.

Example. The eigenvalues of an Hermitian matrix are real numbers.
Proof:

7

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

§5.9 Polar Decomposition*

Norm and Distance

Now we always assume $(V, \langle\cdot,\cdot\rangle)$ is a real inner product space.

Theorem. For $\mathbf{u},\mathbf{v} \in V$, we have

$$|\langle \mathbf{u},\mathbf{v}\rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|.$$

Proof:

8

Remark: $-1 \leq \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} \leq 1$. So it is reasonable to define the angle θ between vectors u and v as

$$\theta = \arccos\left(\frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right).$$

Corollary. For $\mathbf{u},\mathbf{v} \in V$, the following inequalities hold.

◇ $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$. (Triangular inequality)

◇ $d(\mathbf{u},\mathbf{v}) \leq d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$. (Triangular inequality)

Orthogonality

- $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

Definition. We say that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Example. Prove that the following vectors are pairwise orthogonal in given inner product spaces.

(1) $V = M_2$, $\langle A, B \rangle = \mathrm{tr}(B^T A)$.

Take $A = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$.

Proof:

9

Example. Prove that the following vectors are pairwise orthogonal in given inner product spaces.

(2) $\mathbb{F}=\mathbb{R}$, $V=C[-\pi,\pi]$, $\langle f,g\rangle=\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)dt$.

Take $f_m(t)=\sin mt$, $g_n(t)=\cos nt$ and $h(t)=1/\sqrt{2}$.

(3) $\mathbb{F}=\mathbb{C}$, $V=C_{\mathbb{C}}[-\pi,\pi]$ $\langle f,g\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}dt$.
(the space of all continuous complex-valued functions)

Take $f_n(t)=e^{int}$ $(n\in\mathbb{Z})$.

Proof:

10

Gou-Gu Theorem

Theorem. (Gou-Gu Theorem) Suppose $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors in an inner product space. Then

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Example. Let $\mathbb{F}=\mathbb{R}$, $V=P_2$ and $\langle p,q\rangle = \int_{-1}^{1} p(x)q(x)dx$. Verify Gou-Gu Theorem with $p(x)=x$ and $q(x)=x^2$.

Proof:

11

Orthogonal Complements

Definition. If W is a subspace of an inner product space V, then the orthogonal complement of W is defined by

$$W^\perp = \{\mathbf{v} : \langle \mathbf{v}, \mathbf{w} \rangle = 0 \, (\forall \mathbf{w} \in W)\}.$$

Theorem. Let W be a subspace of an inner product space V.
(1) $W^\perp$ is a subspace of V.
(2) $W^\perp \cap W = \{\mathbf{0}\}$.
(3) If V is finite-dimensional, then $(W^\perp)^\perp = W$.

Proof:

12

Remark: The statement (3) may be incorrect when V is infinite-dimensional. It would be correct provided that W is a "closed" subspace.

Example. Let W be the subspace of $\mathbb{R}^6$ spanned by the vectors

$$\mathbf{w}_1 = (1, 3, -2, 0, 2, 0), \quad \mathbf{w}_2 = (2, 6, -5, -2, 4, -3)$$
$$\mathbf{w}_3 = (0, 0, 5, 10, 0, 15), \quad \mathbf{w}_4 = (2, 6, 0, 8, 4, 18).$$

Find a basis for the orthogonal complement of W under Euclidean inner product.

Solution:

13

Examples

Example. Let $V = C[-1,1]$, and

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx, \quad (f, g \in V).$$

Let V_e (V_o) be the space of even (odd, respectively) functions in V. Prove that $V_o^{\perp} = V_e$.

Proof:

14

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

§5.9 Polar Decomposition*

Orthonormal Sets

Philosophy: Basis is designed for a vector space. Now there is a inner product, the basis vectors can have “angles”.

Definition.

◇ A set of two or more vectors in an inner product space is called orthogonal if all pairs of distinct vectors in the set are orthogonal.

◇ An orthogonal set in which each vectors has norm 1 is called orthonormal.

◇ If a basis is orthogonal, then we say it is an orthogonal basis.

◇ If a basis is orthonormal, then we say it is an orthonormal basis.

Example. In $\mathbb{R}^3$ with Euclidean inner product, let $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, where

$$\mathbf{u}_1 = (0, 1, 0), \quad \mathbf{u}_2 = (1, 0, 1), \quad \mathbf{u}_3 = (1, 0, -1).$$

Verify that the set S is orthogonal. Construct an orthonormal basis by normalising these vectors.

Solution:

15

Examples

Example. Show that the following sets are orthonormal in given inner product spaces.

(1) $V = M_n$, $\langle A, B\rangle = \mathrm{tr}(B^T A)$, the standard basis.

(2) $V = P_n$, $\left\langle \sum_{k=0}^{n} a_k x^k, \sum_{k=0}^{n} b_k x^k \right\rangle = \sum_{k=0}^{n} a_k b_k$.
$B = \{1, x, x^2, \ldots, x^n\}$.

Proof:

16

Examples

Example. Show that the following sets are orthonormal in given inner product spaces.

(3) $V = C[-\pi,\pi]$, $\langle f,g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)dt$.

$$B = \left\{\frac{1}{\sqrt{2}}, \sin x, \sin 2x, \ldots, \sin kx, \cos x, \cos 2x, \ldots, \cos \ell x\right\}.$$

(4) $V = C_{\mathbb{C}}[-\pi,\pi]$, $\langle f,g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}dt$.

$$B = \left\{e^{inx} : |n| \le N\right\}.$$

Proof:

Properties of Orthogonal Basis

Theorem. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of non-zero vectors in an inner product space $(V, \langle\cdot,\cdot\rangle)$, then S is linearly independent.

Proof:

17

Theorem. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for an inner product space V, and if $\mathbf{u}$ is any vector in V, then

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_n\rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n.$$

Proof:

18

Remark: When S is an orthonormal basis,

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle\mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle\mathbf{v}_2 + \cdots \langle \mathbf{u}, \mathbf{v}_n\rangle\mathbf{v}_n.$$

Examples

Example. Let $\mathbf{v}_1 = (0, 2, 0)$, $\mathbf{v}_2 = (3, 0, 3)$, $\mathbf{v}_3 = (-4, 0, 4)$ and $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

(1) Verify that S form an orthogonal basis for $\mathbb{R}^3$ with Euclidean inner product.
(2) Find $[\mathbf{u}]_B$, where $\mathbf{u} = (1, 2, 4)$.

Solution:

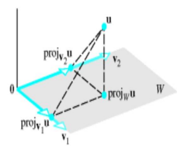
19

Orthogonal Projections

Theorem. (Projection Theorem) If W is a subspace of a finite-dimensional inner product space V, then every vector $\mathbf{u}$ in V can be expressed in exactly one way as

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2,$$

where $\mathbf{w}_1 \in W$ and $\mathbf{w}_2 \in W^{\perp}$.



Remark: We denote $\mathbf{w}_1 = \text{proj}_W(\mathbf{u})$ and $\mathbf{w}_2 = \text{proj}_{W^\perp}(\mathbf{u})$.

Theorem. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and W be a subspace of V with an orthogonal basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$. Then for any $\mathbf{u} \in V$,

$$\text{proj}_W \mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_1 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_r \rangle}{\|\mathbf{v}_r\|^2}\mathbf{v}_r$$

Proof:

20

Remark: When S is an orthonormal basis for W,

$$\text{proj}_W \mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots, + \langle \mathbf{u}, \mathbf{v}_r \rangle \mathbf{v}_r.$$

Examples

Example. Let $\mathbb{F} = \mathbb{R}$, $V = C[-\pi, \pi]$ and

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)g(t)dt, \quad (\forall f, g \in V).$$

Find the orthogonal projection of $h(t) = t$ on the subspace $W = \text{span}\{\cos t, \sin t\}$.

Solution:

21

Gram-Schmidt Process

Theorem. Every nonzero finite-dimensional inner product space has an orthonormal basis.

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for V.

Step 1-1. Take $\mathbf{u}_1 = \mathbf{v}_1$.

Step 1-2. Take $\mathbf{u}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 = \mathrm{proj}_{\mathrm{span}\{\mathbf{u}_1\}^\perp}(\mathbf{v}_2)$.

Step 1-3. Take $\mathbf{u}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 - \frac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\|\mathbf{u}_2\|^2}\mathbf{u}_2 = \mathrm{proj}_{\mathrm{span}\{\mathbf{u}_1, \mathbf{u}_2\}^\perp}(\mathbf{v}_3)$.

......

Step 1-n. Take

$$\mathbf{u}_n = \mathbf{v}_n - \frac{\langle \mathbf{v}_n, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 - \ldots - \frac{\langle \mathbf{v}_n, \mathbf{u}_{n-1} \rangle}{\|\mathbf{u}_{n-1}\|^2}\mathbf{u}_{n-1} = \mathrm{proj}_{\mathrm{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}\}^\perp}(\mathbf{v}_n).$$

Conclusion of step 1:

We obtain an orthogonal basis $S' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ for V.

Step 2. Let $\mathbf{w}_i = \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|}$ $(1 \leq i \leq n)$.

Conclusion of step 2:

We obtain an orthonormal basis $S'' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ for V.

Examples

Example. Consider $\mathbb{F} = \mathbb{R}$ and P_2 with the inner product given by

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)dx, \quad (\forall p, q \in P_2).$$

Apply the Gram-Schmidt process to transform the standard basis $\{1, x, x^2\}$ for P_2 into an orthonormal basis $\{q_1(x), q_2(x), q_3(x)\}$.

Solution:

22

Remark: These polynomials are called Legendre polynomials.

Extention of Orthonormal Basis

By the Gram-Schmidt process, orthonormal sets can be extened to orthonormal bases.

Theorem. If W is a finite-dimensional inner product space, then:
(a) Every orthogonal set of nonzero vectors in W can be enlarged to an orthogonal basis for W.
(b) Every orthonormal set in W can be enlarged to an orthonormal basis for W.

QR-Decomposition

Theorem. Suppose that $A \in M_{m\times n}$ and $\operatorname{rank}(A) = n$. Then A can be factored as $A = QR$, where $Q \in M_{m\times n}$ is an matrix with orthonormal column vectors, and R is an invertible upper triangular matrix.

Indeed, let $A = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \end{bmatrix}$. Suppose that we obtain orthonormal vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$ by applying the Gram-Schmidt process to $\mathbf{c}_1, \ldots, \mathbf{c}_n$. Then

$$A = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \langle \mathbf{c}_1, \mathbf{q}_1 \rangle & \langle \mathbf{c}_2, \mathbf{q}_1 \rangle & \cdots & \langle \mathbf{c}_n, \mathbf{q}_1 \rangle \\ \langle \mathbf{c}_1, \mathbf{q}_2 \rangle & \langle \mathbf{c}_2, \mathbf{q}_2 \rangle & \cdots & \langle \mathbf{c}_n, \mathbf{q}_2 \rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{c}_1, \mathbf{q}_n \rangle & \langle \mathbf{c}_2, \mathbf{q}_n \rangle & \cdots & \langle \mathbf{c}_n, \mathbf{q}_n \rangle \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \langle \mathbf{c}_1, \mathbf{q}_1 \rangle & \langle \mathbf{c}_2, \mathbf{q}_1 \rangle & \cdots & \langle \mathbf{c}_n, \mathbf{q}_1 \rangle \\ & \langle \mathbf{c}_2, \mathbf{q}_2 \rangle & \cdots & \langle \mathbf{c}_n, \mathbf{q}_1 \rangle \\ & & \ddots & \vdots \\ & & & \langle \mathbf{c}_n, \mathbf{q}_n \rangle \end{bmatrix} = QR.$$

An Example

Example. Find a QR-decomposition of $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$.

Solution:

23

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

§5.9 Polar Decomposition*

Definition of Orthogonal Matrix

Recall: "Orthonormal basis" play an important role in an inner product space.

Question: What is the relationship between two orthonormal bases?

How does the transition matrix look like?

(from an orthonormal basis to standard bases)

Definition. A matrix $A \in M_n(\mathbb{R})$ is said to be orthogonal if one of the following equivalent conditions holds:

(1) $A^T A = AA^T = I_n$;

(2) $A^{-1} = A^T$;

(3) The row vectors of A form an orthonormal set in $\mathbb{R}^n$ with the Euclidean inner product;

(4) The column vectors of A form an orthonormal set in $\mathbb{R}^n$ with the Euclidean inner product

Proof of the equivalence:

24

Examples

Example. Show that the matrix

$$\begin{bmatrix} 3/7 & 2/7 & 6/7 \\ -6/7 & 3/7 & 2/7 \\ 2/7 & 6/7 & -3/7 \end{bmatrix}$$

is orthogonal.

Solution:

25

Example. Let $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Whether the matrices R_θ, R_θ^T, R_θ^{-1}, $R_\theta R_{\theta'}$ orthogonal?

Solution:

26

Definition of Unitary Matrix

Definition. A matrix $A \in M_n(\mathbb{C})$ is said to be unitary if one of the following equivalent conditions holds:
(1) $A^*A = AA^* = I_n$;
(2) $A^{-1} = A^*$;
(3) The row vectors of A form an orthonormal set in $\mathbb{C}^n$ with the Euclidean inner product;
(4) The column vectors of A form an orthonormal set in $\mathbb{C}^n$ with the Euclidean inner product

Example. Show that a 1×1 matrix $[z] \in M_1(\mathbb{C})$ is unitary if and only if $|z| = 1$.

Proof:

27

Properties of Orthogonal/Unitary Matrix

Theorem. Let $A, B \in M_n(\mathbb{R})$ be orthogonal matrices. Then
- ◇ A^T is orthogonal;
- ◇ A^{-1} is orthogonal;
- ◇ AB is orthogonal;
- ◇ $\det(A) = \pm 1$. $|\det(A)| = 1$

Proof:

28

Theorem. Let $A, B \in M_n(\mathbb{C})$ be unitary matrices. Then
- ◇ A^* is unitary;
- ◇ A^{-1} is unitary;
- ◇ AB is unitary;
- ◇ $\det(A) = e^{i\theta}$ for some $0 \le \theta < 2\pi$. $|\det(A)| = 1$

Properties of Orthogonal/Unitary Matrix

Theorem. If $A \in M_n(\mathbb{R})$, then the following are equivalent.
(1) A is orthogonal.
(2) $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$ in $\mathbb{R}^n$.
(3) $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
Proof:

29

Remark: As a linear transformation, T_A preserves norm, angle and area. And it preserve (or reverses) the orientation if its determinant is 1 (or -1, respectively).

Corollary. All possible eigenvalues of an orthogonal matrix are ± 1.
Proof:

30

Problem*: Suppose that A is an $n \times n$ orthogonal matrix, where n is odd.

Also suppose that $\det(A) = 1$. Prove that 1 is an eigenvalue of A.

Theorem. If $A \in M_n(\mathbb{C})$, then the following are equivalent.
(1) A is unitary.
(2) $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$ in $\mathbb{C}^n$.
(3) $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$.

Corollary. Any possible eigenvalue λ of a unitary matrix satisfies that $|\lambda| = 1$.

Transition Matrix

Theorem. Let S be an orthonormal basis for an n-dimensional inner product space V over $\mathbb{F}$. Let $(\mathbf{u})_S = (u_1, u_2, \cdots, u_n)$ and $(\mathbf{v})_S = (v_1, v_2, \cdots, v_n)$. Then

- ◇ $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 \overline{v_1} + \cdots u_n \overline{v_n}$;
- ◇ $\|\mathbf{u}\| = \sqrt{|u_1|^2 + \cdots |u_n|^2}$;
- ◇ $d(\mathbf{u}, \mathbf{v}) = \sqrt{|u_1 - v_1|^2 + \ldots + |u_n - v_n|^2}$.

Proof:

31

Remark: When $\mathbb{F} = \mathbb{R}$, we may ignore the "conjugate".

Theorem. Let V be a finite-dimensional inner product space. If P is the transition matrix from one orthonormal basis to another orthonormal basis for V, then P is an orthogonal/unitary matrix.

Proof:

32

Adjoint Operator*

Theorem. Let $(V, \langle\cdot,\cdot\rangle_V)$ be a finite-dimensional complex inner product spaces. Suppose that $T : V \to V$ is a linear operator. Then there exists a unique linear transformation $T^* : V \to V$ such that

$$\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, T^*\mathbf{y}\rangle, \quad (\forall \mathbf{x}, \mathbf{y} \in V).$$

Remark: We do not prove the above theorem here.

Remark: Under a given orthonormal bases, we have $[T^*] = [T]^*$.

The main purpose of this page is:

"(conjugate) transpose" can be defined in an abstract way.

Remark: A linear operator $U : V \to V$ is said to be unitary if

$$UU^* = U^*U = I.$$

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

§5.9 Polar Decomposition*

Concept of Diagonalisazion

Recall: "Orthonormal basis" play an important role in an inner product space.

Recall: A "diagonalizable" matrix is relatively easy to determine.

Our interest: an operator T on V such that eigenvectors form an orthonormal basis for V.

Example. The projection on the xz-plane in $\mathbb{R}^3$ is given by

$$T(x, y, z) = (x, 0, z), \quad (x, y, z) \in \mathbb{R}^3.$$

Indeed, it has standard matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

The vectors $\mathbf{e}_1$ and $\mathbf{e}_3$ are eigenvectors corresponding to the eigenvalue 1; and $\mathbf{e}_2$ is an eigenvector corresponding to 0.

Conditions for Orthogonal Diagonalizability

Definition. Let $A \in M_n(\mathbb{R})$. If

$$P^T AP = P^{-1}AP = D$$

for some orthogonal matrix P and some diagonal matrix D, then we say that A is orthogonally diagonalizable and P orthogonally diagonalizes A.

Remark: In numerical algorithms, the complexity of dealing with P^T is much less than with P^{-1}.

Theorem. Let $A \in M_n(\mathbb{R})$, then the following are equivalent.
(1) A is orthogonally diagonalizable.
(2) A has an orthonormal set of n eigenvectors.
(3) A is symmetric.

Problem*: Prove (1)⇔(2) and (1)⇒(3).

Remark: The proof of (3)⇒(1) is beyond the scope of the course.

Remark: $P = \left[\begin{array}{c:c:c:c} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{array}\right]$, $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$.

Here $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ $(1 \le i \le n)$, and the set $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is orthonormal.

Conditions for Unitary Diagonalizability

Definition. Let $A \in M_n(\mathbb{C})$. If

$$U^*AU = U^{-1}AU = D$$

for some unitary matrix U and diagonal matrix D, then we say that A is unitarily diagonalizable and U unitarily diagonalizes A.

Theorem. Let $A \in M_n(\mathbb{C})$, then the following are equivalent.
(1) A is unitarily diagonalizable.
(2) A has an orthonormal set of n eigenvectors.
(3) A is normal, i.e., $A^*A = AA^*$.

Remark: Hermitian, unitary matrices are all normal.
Remark: The proof of this theorem is beyond the scope of the course.

Orthogonal Diagonalization

Theorem. Let $A \in M_n(\mathbb{R})$ be a symmetric matrix, then eigenvectors from different eigenspaces are orthogonal.

Proof:

33

Remark: Same holds for an Hermitian matrix in the complex case.

Example. Suppose that $A \in M_3(\mathbb{R})$ is symmetric and $\text{rank}(A) = 2$. Suppose that

$$A \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix}$$

Find all the eigenvalues of A. For each eigenvalue, find an eigenvector corresponding to it.

Proof:

34

An Example

Fact: Suppose that all the distinct eigenvalues of $A \in M_n$ are $\lambda_1, \ldots, \lambda_k$, and $\{\mathbf{u}_{i,1}, \ldots, \mathbf{u}_{i,i_{n_i}}\}$ are an orthonormal basis for $E_{\lambda_i}(A)$ $(1 \le i \le k)$. Then

$$\{\mathbf{u}_{1,1}, \ldots, \mathbf{u}_{1,n_1}, \mathbf{u}_{2,1}, \ldots, \mathbf{u}_{2,n_2}, \ldots, \mathbf{u}_{k,1}, \ldots, \mathbf{u}_{k,n_k}\}$$

form an orthonormal basis for $\mathbb{R}^n$.

Remark: We only need to apply Gram-Schmidt process on each eigenspaces.

Example. Orthogonal diagonalize the matrix $A = \begin{bmatrix} 2 & 3 & 3 \\ 3 & 2 & 3 \\ 3 & 3 & 2 \end{bmatrix}$.

Proof:

35

Problem*: Summarize the main steps of orthogonal diagonalization.

An Example

Example. Consider

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in M_2(\mathbb{C}).$$

Find a unitary matrix U and unitarily diagonalize A.

Proof:

36

Spectral Decomposition of Symmetric Matrix

Theorem. Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. Then there are $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ and orthonormal vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \in \mathbb{R}^n$ such that

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots \lambda_n \mathbf{u}_n \mathbf{u}_n^T.$$

Proof:

37

Remark: The above equality is called a spectral decomposition of A.

Remark: In above theorem, $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$ $(1 \le i \le n)$.
And A is orthogonally diagonalized by $P = \left[\begin{array}{c:c:c:c} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n \end{array}\right]$.

Problem: Prove that the square matrix $\mathbf{u}_1 \mathbf{u}_1^T$ is the orthogonal projection on $\operatorname{span}\{\mathbf{u}_1\}$.

Theorem. Let $A \in M_n(\mathbb{C})$ be a normal matrix. Then there are $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and orthonormal vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \in \mathbb{C}^n$ such that

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^* + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^* + \cdots \lambda_n \mathbf{u}_n \mathbf{u}_n^*.$$

Remark: When A is normal, there are matrices $E_1, \ldots, E_k$ with
(i) $E_i^2 = E_i = E^*$ $(1 \leq i \leq k)$;
(ii) $E_i E_j = 0$ $(1 \leq i \neq j \leq k)$;
(iii) $\sum_{i=1}^k E_i = I$,
and distinct numbers $\mu_1, \ldots, \mu_k \in \mathbb{C}$ such that

$$A = \sum_{i=1}^{k} \mu_i E_i.$$

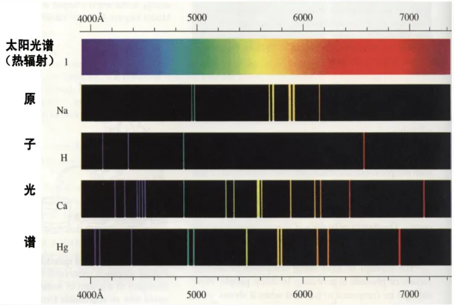
Indeed, the numbers $\mu_1, \ldots, \mu_k$ are all different eigenvalues of A, and E_i is the orthogonal projection on the eigenspace $E_{\mu_i}(A)$.

Spectral Decomposition of Hermitian Matrix*

In quantum physics, each observable can be viewed as an Hermitian (self-adjoint) operator, whose spectral are real. Let

$$A = \sum_{i=1}^{k} \mu_i E_i, \qquad (1)$$

where $\mu_1 < \mu_2 < \ldots < \mu_n$.



Remark: For an operator on infinite-dimensional space, there spectrum may contain "continuous" part. Then the summation '$\sum$' in (1) becomes integration '$\int$'.

Common (Complex) Eigenvectors*

Theorem. Suppose that the two matrices $A, B \in M_n(\mathbb{C})$ commute, i.e., $AB = BA$. Then they have a common eigenvector.

Remark: The eigenvalues of A and B, to which this common eigenvector corresponds, can be different.

Remark: In physics, two commutative observables (i.e., Hermitian operators) have common eigenvectors that forming an orthonormal basis, which means that they can be observed simultaneously.

The Nondiagonalizable Real Matrix*

Theorem. (Schur's Theorem) If the characteristic polynomial of $A \in M_n(\mathbb{R})$ has n roots $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, then there is an orthogonal matrix P such that P^TAP is an upper triangular matrix of the form

$$P^TAP = \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} := S$$

In particular, $A = PSP^T$ is called a Schur decomposition of A.

Remark: In numerical algorithms,

◇ the matrix S is much simpler than A;

◇ the orthogonal matrix does not magnify round-off error, i.e.,

$$\mathbf{x} = \hat{\mathbf{x}} + \mathbf{e} \quad \Longrightarrow \quad \|P\mathbf{x} - P\hat{\mathbf{x}}\| = \|P\mathbf{e}\| = \|\mathbf{e}\|.$$

The Nondiagonalizable Real Matrix*

Theorem. (Hessenberg's Theorem) If $A \in M_n(\mathbb{R})$, then there is an orthogonal matrix P such that

$$P^TAP = \begin{bmatrix} * & * & \dots & * & * & * \\ * & * & \dots & * & * & * \\ 0 & * & \ddots & * & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & * & * & * \\ 0 & 0 & \dots & 0 & * & * \end{bmatrix} := H.$$

In particular, $A = PHP^T$ is called an upper Hessenberg decomposition of A.

Remark: In numerical algorithms,

◇ the matrix H is much simpler than A;

◇ the orthogonal matrix does not magnify round-off error.

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

§5.9 Polar Decomposition*

Property of A^TA

Background: In application, data usually form a matrix A in $M_{m\times n}(\mathbb{R})$, where m and n may be different.

Idea: Find square matrices that contain as much information as A does.
Direction: $A^TA \in M_n(\mathbb{R})$; $AA^T \in M_m(\mathbb{R})$!

Theorem. Let $A \in M_{m\times n}(\mathbb{R})$. Then

- $\diamond$ $\mathrm{Null}(A) = \mathrm{Null}(A^TA)$;
- $\diamond$ $\mathrm{Row}(A) = \mathrm{Row}(A^TA)$;
- $\diamond$ $\mathrm{Col}(A^T) = \mathrm{Col}(A^TA)$;
- $\diamond$ $\mathrm{rank}(A) = \mathrm{rank}(A^TA)$.

Proof:

38

Singular Value

Theorem. Let $A \in M_{m\times n}(\mathbb{R})$. Then:
◇ A^TA is orthogonally diagonalizable;
◇ The eigenvalues of A^TA are nonnegative.
Proof:

39

Definition. Let $A \in M_{m\times n}(\mathbb{R})$, and $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A^TA. Then the numbers

$$\sigma_1 = \sqrt{\lambda_1},\ \sigma_2 = \sqrt{\lambda_2},\ \ldots,\ \sigma_n = \sqrt{\lambda_n}.$$

are called the singular values of A.

Remark: We usually assume that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, and so $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$.

Examples

Example. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$. Find the singular values of A and A^T, respectively.

Solution:

40

Question*: Prove that the singular values of A and A^T are same, except for some zero values.

Singular Value Decomposition (A Quick View)

A quick view on "SVD":

Let $A \in M_{m\times n}(\mathbb{R})$ and $\text{rank}(A) = r$. Then there are orthogonal matrix $U \in M_m$ and $V \in M_n$ such that

$$A = U\Sigma V,$$

where

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & & \end{bmatrix}_{m\times n}.$$

Notations occur in next several pages: Let $A \in M_{m\times n}(\mathbb{R})$ and $\text{rank}(A) = r$. Eigenvalues of A^TA are

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_n = 0,$$

with corresponding orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. The singular values of A are $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_n = 0$.

Theorem. . Let $A \in M_{m\times n}(\mathbb{R})$ and $\operatorname{rank}(A) = r$. There there is

- ⋄ an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for $\mathbb{R}^n$;
- ⋄ an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ for $\mathbb{R}^m$;
- ⋄ positive scalars $\sigma_1, \ldots, \sigma_r \in \mathbb{R}$, such that

$$A\mathbf{v}_i = \begin{cases} \sigma_i \mathbf{u}_i, & (1 \le i \le r), \\ 0, & (r < i \le n), \end{cases} \quad A^T\mathbf{u}_i = \begin{cases} \sigma_i \mathbf{v}_i, & (1 \le i \le r), \\ 0, & (r < i \le m). \end{cases}$$

Proof:

Take $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$. Here $\|A\mathbf{v}_i\| = \sigma_i$.

Then extend it to an orthonormal basis.

41

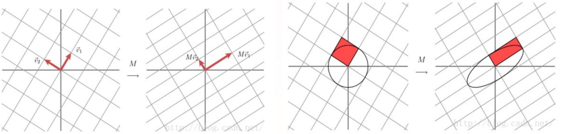
Remark: There exist orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ such that $A\mathbf{v}_1, \ldots, A\mathbf{v}_n$ are orthogonal. The quantities σ_i turns $A\mathbf{v}_1, \ldots, A\mathbf{v}_r$ into unit vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$.

Remark: In matrix notation, we have $A = U_r \Sigma_r V_r^T$, i.e.,

$$A_{m\times n} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_r \end{bmatrix}_{n\times r} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_r \end{bmatrix}_{m\times r} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}_{r\times r}.$$

Effect of a 2×2 Matrix

By geometric SVD, there must be orthonormal vectors $\mathbf{v}_1, \mathbf{v}_2$ such that $A\mathbf{v}_1, A\mathbf{v}_2$ are orthogonal.



Example. Shears of $\mathbb{R}^2$:

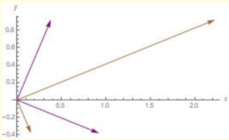
$$M = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, \quad M^T = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}, \qquad M^T M = \begin{bmatrix} 1 & k \\ k & 1+k^2 \end{bmatrix}.$$

When $k = 2$, one has

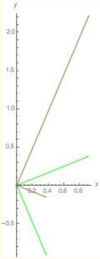
$$\lambda_{1,2} = 3 \pm 2\sqrt{2}, \quad \sigma_{1,2} = \sqrt{2} \pm 1,$$

$$\mathbf{v}_{1,2} = \left(\frac{1}{\sqrt{4 \pm 2\sqrt{2}}}, \frac{1 \pm \sqrt{2}}{\sqrt{4 + 2\sqrt{2}}} \right),$$

$$\mathbf{u}_{1,2} = \left(\frac{\pm 1 + \sqrt{2}}{\sqrt{4 \pm 2\sqrt{2}}}, \frac{\pm 1}{\sqrt{4 \pm 2\sqrt{2}}} \right).$$



$\mathbf{v}_i,\ A\mathbf{v}_i$



$\mathbf{u}_i,\ A^T\mathbf{u}_i$

Singular Value Decomposition (Full Version)

Theorem. Let $A \in M_{m\times n}(\mathbb{R})$ and $\text{rank}(A) = r$. Then there are

- ◇ an orthogonal matrix $U \in M_m(\mathbb{R})$;
- ◇ an orthogonal matrix $V \in M_n(\mathbb{R})$;
- ◇ positive scalars $\sigma_1, \ldots, \sigma_r \in \mathbb{R}$,

such that $A = U\Sigma V^T$, where

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & & \end{bmatrix}_{m\times n}.$$

Proof:

42

Remark: We have

$$U = \left[\begin{array}{c:c:c:c} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \end{array}\right]_{m\times m}, \qquad V^T = \begin{bmatrix} \mathbf{v}_1^T \\ \hline \mathbf{v}_2^T \\ \hline \vdots \\ \hline \mathbf{v}_n^T \end{bmatrix}_{n\times n}.$$

Singular Value Decomposition (Condensed Version)

Theorem. Let $A \in M_{m\times n}(\mathbb{R})$ and $\text{rank}(A) = r$. Then there are

◇ an orthonormal set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ in $\mathbb{R}^n$;

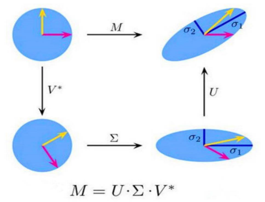
◇ an orthonormal set $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ in $\mathbb{R}^m$;

◇ positive scalars $\sigma_1, \ldots, \sigma_r \in \mathbb{R}$,

such that
$$A = \mathbf{u}_1\sigma_1\mathbf{v}_1^T + \cdots + \mathbf{u}_r\sigma_r\mathbf{v}_r^T,$$

or equivalently, $A = U_r\Sigma_r V_r^T$

$$= \begin{bmatrix} \mathbf{u}_1 & \vdots & \mathbf{u}_2 & \vdots & \cdots & \vdots & \mathbf{u}_r \end{bmatrix}_{m\times r} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}_{r\times r} \begin{bmatrix} \mathbf{v}_1^T \\ \hline \mathbf{v}_2^T \\ \hline \vdots \\ \hline \mathbf{v}_r^T \end{bmatrix}_{r\times n}.$$



Example. The shear of $\mathbb{R}^2$ with $k = 2$ has SVD:

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}+1}{\sqrt{4+2\sqrt{2}}} & \frac{\sqrt{2}-1}{\sqrt{4-2\sqrt{2}}} \\ \frac{1}{\sqrt{4+2\sqrt{2}}} & \frac{-1}{\sqrt{4-2\sqrt{2}}} \end{bmatrix} \begin{bmatrix} \sqrt{2}+1 & 0 \\ 0 & \sqrt{2}-1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{4+2\sqrt{2}}} & \frac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} \\ \frac{1}{\sqrt{4-2\sqrt{2}}} & \frac{1-\sqrt{2}}{\sqrt{4-2\sqrt{2}}} \end{bmatrix}.$$

More Properties of SVD*

Fundamental spaces of a matrix:

- ◇ $\mathrm{Col}(A) = \mathrm{Col}(U_r)$.
- ◇ $\mathrm{Null}(A) = \mathrm{Col}(V_{n-r})$.
- ◇ $\mathrm{Row}(A) = \mathrm{Col}(A^T) = \mathrm{Col}(V_r)$.

Approximation of a matrix:

- ◇ When $k \leq r$, the best approximation of matrix is given by

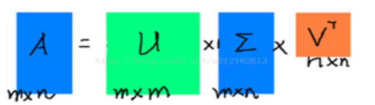
$$\|A - A_k\| = \min_{\mathrm{rank}(B) \leq k} \|A - B\|.$$

Here $\|\cdot\|$ is some certain norm on M_n, and

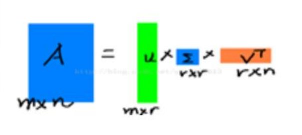
$$A_k = \mathbf{u}_1 \sigma_1 \mathbf{v}_1^T + \ldots + \mathbf{u}_k \sigma_k \mathbf{v}_k^T.$$

The most important information of A gathers at upper left part of SVD!

Applications of SVD



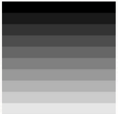
Full Version



Condensed Version

Remark: For applications, the condensed version only remains the most important part, which saves space and simplified computations.

Remark: Recall that $\sigma_1 \geq \sigma_2 \geq \dots$.
One can even discard the smaller singular values: denoise!



$$uv^H = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0.1 & 0.1 & \cdots & 0.1 \\ 0.2 & 0.2 & \cdots & 0.2 \\ \vdots & \vdots & & \vdots \\ 0.9 & 0.9 & \cdots & 0.9 \end{pmatrix}$$

Examples

Example. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Find a singular value decomposition of A .

Solution:

43

Chapter 5. Inner Product Spaces

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

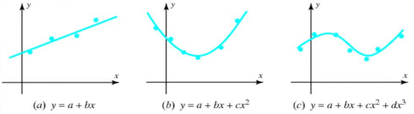
§5.9 Polar Decomposition*

Least Squares Fitting to Data

In this section, we only consider real case, i.e., $\mathbb{F} = \mathbb{R}$.

Problem: To obtain a mathematical relationship $y = f(x)$ by "fitting" a curve to points

$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_n, y_n).$$



Which coefficients $c_0, c_1, \ldots, c_{n-1}$ of a polynomial of $n-1$ minimizes

$$\sum_{k=1}^{m} \left| y_k - (c_0 + c_1 x_k + \ldots + c_{n-1} x_k^{n-1}) \right|^2?$$

Matrix Form:

Least Squares Solution

Fact: A linear system $A\mathbf{x} = \mathbf{b}$ may be inconsistent.

Aim: Find a vector that is as "close" to a solution as possible.

Problem: Find a vector $\hat{\mathbf{x}}$ that minimizes $\|A\mathbf{x} - \mathbf{b}\|$, i.e.,

$$\sum_{i=1}^{m}(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - b_i)^2$$

We call such a vector a least squares solution of the system

If the least squares solution $\hat{\mathbf{x}}$ exists, then

$$\|A\hat{\mathbf{x}} - \mathbf{b}\| = \min_{\mathbf{w}\in \mathrm{Col}(A)} \|\mathbf{w} - \mathbf{b}\|.$$

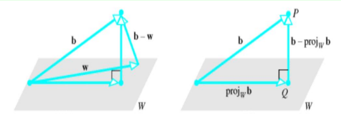
Best Approximation

Theorem. (Best Approximation) If W is a subspace of a finite-dimensional inner product space V, and if $\mathbf{b}$ is a vector in V, then $\text{proj}_W(\mathbf{b})$ is the best approximation to $\mathbf{b}$ from W in the sense that

$$\left\|\mathbf{b} - \text{proj}_W(\mathbf{b})\right\| < \|\mathbf{b} - \mathbf{w}\|$$

for any $\mathbf{w} \in W \setminus \{\text{proj}_W(\mathbf{b})\}$.

Proof:



44

Remark: $\left\|\mathbf{b} - \text{proj}_W(\mathbf{b})\right\| = \min_{\mathbf{w}\in W} \|\mathbf{b} - \mathbf{w}\|$. The minimum takes uniquely.
Remark: A least square solution $\hat{\mathbf{x}}$ of the system $A\mathbf{x} = \mathbf{b}$ satisfies

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \min_{\mathbf{w}\in \text{Col}(A)} \|\mathbf{b} - \mathbf{w}\| \iff A\hat{\mathbf{x}} = \text{proj}_{\text{Col}(A)}(\mathbf{b}).$$

Method of Least Squares

Theorem. (1) Least square solutions of a linear system $A\mathbf{x} = \mathbf{b}$ always exist. For any least squares solution $\hat{\mathbf{x}}$ of $A\mathbf{x} = \mathbf{b}$, the orthogonal projection of $\mathbf{b}$ on $\mathrm{Col}(A)$ is

$$\mathrm{proj}_{\mathrm{Col}(A)}\mathbf{b} = A\hat{\mathbf{x}}.$$

And it satisfies that $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$.

(2) For every linear system $A\mathbf{x} = \mathbf{b}$, the associated normal system

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

is consistent, and all solutions are least squares solutions of $A\mathbf{x} = \mathbf{b}$.

Proof:

45

Remark: $A^T\mathbf{u} = \mathbf{0}$ if and only if $\mathbf{u} \in \mathrm{Col}(A)^{\perp}$

An Example

Example. Find all least squares solutions of the linear system

$$\begin{cases} x_1 + x_2 & = 0 \\ \phantom{x_1 + {}} x_2 & = 1 \\ x_1 & = 2 \end{cases}$$

Solution:

46

Remark: The matrix $A^T A$ may be singular.

An Example

Example. Find the orthogonal projection of the vector $\mathbf{b} = (0, 1, 2)$ on the subspace W of $\mathbb{R}^3$ spanned by the vectors

$$\mathbf{w}_1 = (1, 0, 1), \quad \mathbf{w}_2 = (1, 1, 0).$$

Solution:

47

Least Squares Fitting to Data

Which coefficients $c_0, c_1, \ldots, c_{n-1}$ minimizes

$$\sum_{k=1}^{m} \left| y_k - (C_0 + c_1 x_1 + \ldots + c_{m-1} x_{m-1}^{n}) \right|^2 ?$$

Method. Consider fitting a polynomial of given degree m, i.e.,

$$y(x) = c_0 + c_1 x + \ldots + c_{n-1} x^{n-1},$$

to m points $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$. The coefficients $c_0, c_1, \ldots, c_{n-1}$ can be determined by

$$M^T M \mathbf{c} = M^T \mathbf{y},$$

where the vector $\mathbf{c}$ minimizes $\|\mathbf{y} - M\mathbf{v}\|$ for $\mathbf{v} \in \mathbb{R}^n$. Here

$$M = \begin{bmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ 1 & x_2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & \cdots & x_m^{n-1} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix}.$$

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

§5.9 Polar Decomposition*

Definition of Quadratic Form

Definition. A polynomial with terms all of degree 2, i.e.,

$$Q(x_1, x_2, \ldots, x_n) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j,$$

is called a quadratic form.

Example. The following are quadratic forms.

- ◇ $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^3$.
- ◇ $Q(x, y) = (x/a)^2 + (y/b)^2$.
- ◇ $Q(x, y) = (x/a)^2 - (y/b)^2$.
- ◇ $Q(x, y) = xy$.
- ◇ $Q(x_1, x_2, x_3) = 2x_1^2 + 6x_1x_2 - 5x_2^2 + 3x_2x_3$.

Quadratic Form and Symmetric Matrix

- Since $x_1x_2 = x_2x_1$ as polynomials, we rewrite

$$c_{12}x_1x_2 = \frac{c_{12}}{2}x_1x_2 + \frac{c_{12}}{2}x_2x_1.$$

- Now

$$Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j,$$

where $a_{ii} = c_{ii}$ for $1 \leq i \leq n$ and

$$a_{ij} = \frac{1}{2}c_{ij},\ (i < j), \quad a_{ij} = \frac{1}{2}c_{ji},\ (i > j).$$

- The matrices appear as

$$Q(\mathbf{x}) = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} = \mathbf{x}^T A\,\mathbf{x}.$$

Here A is a symmetric matrix.

Quadratic Form and Symmetric Matrix

- For any matrix $C = [c_{ij}] \in M_n(\mathbb{R})$,
$$Q_C(\mathbf{x}) := \mathbf{x}^T C\,\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_ix_j$$
defines a quadratic form.

- Fact: One has $C = A + B$, where
$A = \frac{1}{2}(C + C^T)$ is symmetric; $B = \frac{1}{2}(C - C^T)$ is skew-symmetric.

- Fact: One has $Q_B(\mathbf{x}) = 0$, since $B = -B^T$. Why?

- Conclusion: $Q_C(\mathbf{x}) = Q_A(\mathbf{x})$.
The effective part is the symmetric matrix A.

Definition. Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. The quadratic form $Q_A(\mathbf{x})$ associated with A is defined to be
$$Q_A(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x}, \quad (\forall \mathbf{x} \in \mathbb{R}^n).$$

Remark: If V is a general inner product space, and L is a general operator such that $L^T = L$, then we may also define a quadratic form as $Q_L(\mathbf{v}) = \langle L\mathbf{v}, \mathbf{v}\rangle$ $(\mathbf{v} \in V)$.

Examples

Example. Express the quadratic form in the matrix notation $\mathbf{x}^T A\mathbf{x}$, where A is symmetric.
(1) $2x^2 + 6xy - 5y^2$.
(2) $x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3$.

Solution:

48

Change of Variable in a Quadratic Form

Definition. ◇ The substitution $\mathbf{x} = P\mathbf{y}$ is called a change of variable if P is invertible.
◇ The substitution $\mathbf{x} = P\mathbf{y}$ is called an orthogonal change of variable if P is orthogonal.

- We have

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T(P^T AP)\mathbf{y}.$$

Remark: When P is orthogonal, the right-hand side equals $\mathbf{y}^T(P^{-1}AP)\mathbf{y}$.

Recall: Any symmetric matrix in $M_n(\mathbb{R})$ is orthogonal diagonalizable.

Theorem. If $A \in M_n(\mathbb{R})$ is a symmetric matrix, then there is an orthogonal change of variable $\mathbf{x} = P\mathbf{y}$ that transforms the quadratic form $Q_A(\mathbf{x})$ into a quadratic form $Q_A(\mathbf{y})$ with no cross product terms, i.e.,

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T(P^T AP)\mathbf{y} = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2.$$

Here $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A corresponding to the eigenvectors that form the successive columns of P.

Examples

Example. Find a variable change that diagonalizes the quadratic form

$$Q = 2x_1^2 + 2x_2^2 + 2x_3^2 + 6x_1x_2 + 6x_1x_3 + 6x_2x_3,$$

and express Q in terms of the new variables.

Solution:

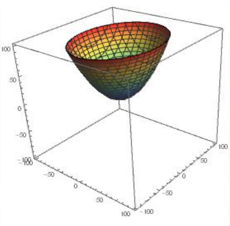
49

Problem*: Summarize the main steps of the solution.

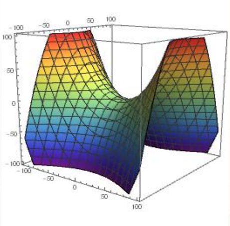
Quadratic Forms in Two Variables

- After orthogonal change of variables, a quadratic form in two variables becomes

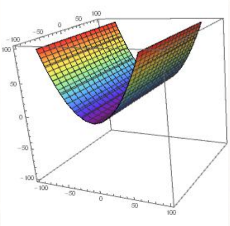
$$Q(x, y) = \lambda_1 x^2 + \lambda_2 y^2.$$



$\lambda_1 > 0,\ \lambda_2 > 0$



$\lambda_1 > 0,\ \lambda_2 < 0$



$\lambda_1 > 0,\ \lambda_2 = 0$

Case 1. $\lambda_1 > 0, \lambda_2 > 0$: $Q(x, y) \geq 0$, but $= 0$ only for $(x, y) = (0, 0)$.
Case 2. $\lambda_1 > 0, \lambda_2 < 0$: $Q(1, 0) > 0$ and $Q(0, 1) < 0$.
Case 3. $\lambda_1 > 0, \lambda_2 = 0$: $Q(x, y) \geq 0$, but $Q(0, 1) = 0$.
For other cases, we reverse the signs of λ_1 or/and λ_2.

Definition. A symmetric matrix $A \in M_n(\mathbb{R})$, or a quadratic form $\mathbf{x}^T A\mathbf{x}$, is said to be

- ◇ positive definite if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$;
- ◇ negative definite if $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \neq 0$;
- ◇ semi-positive definite if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x}$;
- ◇ semi-negative definite if $\mathbf{x}^T A\mathbf{x} \leq 0$ for all $\mathbf{x}$;
- ◇ indefinite if it has both positive and negative values.

Theorem. If A is a symmetric matrix, then:
(a) $\mathbf{x}^T A\mathbf{x}$ is positive definite if and only if all eigenvalues of A are positive.
(b) $\mathbf{x}^T A\mathbf{x}$ is semi-positive definite if and only if all eigenvalues of A are non-negative.
(c) $\mathbf{x}^T A\mathbf{x}$ is negative definite if and only if all eigenvalues of A are negative.
(d) $\mathbf{x}^T A\mathbf{x}$ is semi-negative definite if and only if all eigenvalues of A are non-positive.
(e) $\mathbf{x}^T A\mathbf{x}$ is indefinite if and only if A has both positive and negative eigenvalues.

Equivalent Definition of Positive Definite Matrix*

Theorem. Let $A \in M_n(\mathbb{R})$ be symmetric. Then the following statements are equivalent.

(i) A is positive definite, i.e., $x^T Ax > 0$ for any $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$.
(ii) All eigenvalues of A are positive.
(iii) There is some invertible matrix C such that $A = C^T C$.
(iv) There is some positive definite matrix B such that $A = B^2$.
(v...) ...

Theorem. Let $A = [a_{ij}] \in M_n(\mathbb{R})$ be a positive definite matrix. Then
(1) $\det(A) > 0$.
(2) $a_{ii} > 0$ for any $1 \leq i \leq n$.

An Example

Example. Prove that the matrix

$$H = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

is positive definite.

Solution:

51

Complex Case*

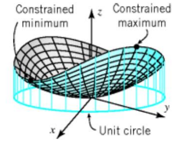
Definition. A matrix $A \in M_n(\mathbb{C})$, or a quadratic form $\mathbf{x}^* A\mathbf{x}$, is said to be

- ◇ positive definite if $\mathbf{x}^* A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$;
- ◇ negative definite if $\mathbf{x}^* A\mathbf{x} < 0$ for all $\mathbf{x} \neq 0$;
- ◇ semi-positive definite if $\mathbf{x}^* A\mathbf{x} \geq 0$ for all $\mathbf{x}$;
- ◇ semi-negative definite if $\mathbf{x}^* A\mathbf{x} \leq 0$ for all $\mathbf{x}$;
- ◇ indefinite if it has both positive and negative values.

Theorem. If A is a normal matrix, i.e., $AA^* = A^*A$, then:
(a) $\mathbf{x}^* A\mathbf{x}$ is positive definite if and only if all eigenvalues of A are positive.
(b) $\mathbf{x}^* A\mathbf{x}$ is semi-positive definite if and only if all eigenvalues of A are non-negative.

Constrained Extremum Problems

Main Problem:
To find the maximum or minimum values of
a quadratic form $\mathbf{x}^T A\mathbf{x}$ subject of the constraint $\|\mathbf{x}\| = 1$.



Theorem. Let $A(\mathbb{R})$ be a symmetric matrix whose eigenvalues in order of decreasing size are $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Then
(1) the quadratic form $\mathbf{x}^T A\mathbf{x}$ attains a maximum value and a minimum value on the set of all unit vectors;
(2) the maximum value attained in part (1) occurs at a unit vector corresponding to the eigenvalue λ_1;
(3) the minimum value attained in part (1) occurs at a unit vector corresponding to the eigenvalue λ_n.

Proof:

51

Example. Find the maximum and minimum values of the quadratic form $z = 5x^2 + 5y^2 + 4xy$ subject to the constraint $x^2 + y^2 = 1$.

Solution:

52

Chapter 5. Inner Product Spaces

§5.1 Inner Products

§5.2 Angle and Orthogonality in Inner Product Spaces

§5.3 Gram-Schmidt Process; QR-Decomposition

§5.4 Orthogonal and Unitary Matrices

§5.5 Orthogonal Diagonalization

§5.6 Singular Value Decomposition

§5.7 Best Approximation; Least Squares

§5.8 Quadratic Forms

§5.9 Polar Decomposition*

Polar Decomposition

Observation: Complex number $z = re^{i\theta}$, where $r \geq 0$ and $|e^{i\theta}| = 1$.

Theorem. Let $A \in M_n(\mathbb{R})$. Then there an orthogonal matrices P and two semi-positive definite matrix H, H' such that

$$A = HP = PH'.$$

Proof:

53

Remark: If $A = HP$, then $A^T A = P^T H^T HP = P^T H^2 P$. Process to find P and H:

◇ Orthogonally diagonalizion: $A^T A = PDP^T$, where D is a diagonal elements with non-negative diagonal elements $\sigma_1, \ldots, \sigma_n$.

◇ $H = P\sqrt{D}P^T$, where $\sqrt{D}$ is the diagonal elements with non-negative diagonal elements $\sqrt{\sigma_1}, \ldots, \sqrt{\sigma_n}$.

Problem*: Find a process to find P and H' directly.

An Overview for Complex Square Matrix

	$\mathbb{C}$	$M_n(\mathbb{C})$	Eigenvalue
Unitary	$\bar{z}z = z\bar{z} = 1$ $\lvert z\rvert = 1$	$U^*U = UU^* = I_n$	$\lvert\lambda\rvert = 1.$
Hermitian	$\bar{z} = z$ $z \in \mathbb{R}$	$H^* = H$	$\lambda \in \mathbb{R}$
Semi-positive	$z \geq 0$	$\forall x \in \mathbb{C}^n,\ x^*Qx \geq 0$ $\exists B \in M_n(\mathbb{C}),\ Q = B^*B$	$\lambda \geq 0$
Orthogonal Projection	$z = 0$ or 1.	$E^2 = E = E^*$	$\lambda \in \{0, 1\}$
Rectangular Decomposition	$z = x + iy$ $(x, y \in \mathbb{R})$	$A = H + iK$ (H, K Hermitian)	
Polar Decomposition	$z = re^{i\theta}$ $(r \geq 0, 0 \leq \theta < 2\pi)$	$A = QU$ Q semi-positive, U unitary	