Inner Product Spaces and Orthogonality

week 13-14 Fall 2006

1 Dot product of $\mathbb{R}^n$

The **inner product** or **dot product** of $\mathbb{R}^n$ is a function $\langle\, , \rangle$ defined by

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \quad \text{for} \quad \boldsymbol{u} = [a_1, a_2, \ldots, a_n]^T,\ \boldsymbol{v} = [b_1, b_2, \ldots, b_n]^T \in \mathbb{R}^n.$$

The inner product $\langle\, , \rangle$ satisfies the following properties:

(1) **Linearity**: $\langle a\boldsymbol{u} + b\boldsymbol{v}, \boldsymbol{w} \rangle = a\langle \boldsymbol{u}, \boldsymbol{w} \rangle + b\langle \boldsymbol{v}, \boldsymbol{w} \rangle$.

(2) **Symmetric Property**: $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$.

(3) **Positive Definite Property**: For any $\boldsymbol{u} \in V$, $\langle \boldsymbol{u}, \boldsymbol{u} \rangle \geq 0$; and $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = 0$ if and only if $\boldsymbol{u} = 0$.

With the dot product we have geometric concepts such as the length of a vector, the angle between two vectors, orthogonality, etc. We shall push these concepts to abstract vector spaces so that geometric concepts can be applied to describe abstract vectors.

2 Inner product spaces

Definition 2.1. An **inner product** of a real vector space V is an assignment that for any two vectors $u, v \in V$, there is a real number $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$, satisfying the following properties:

(1) **Linearity**: $\langle a\boldsymbol{u} + b\boldsymbol{v}, \boldsymbol{w} \rangle = a\langle \boldsymbol{u}, \boldsymbol{w} \rangle + b\langle \boldsymbol{v}, \boldsymbol{w} \rangle$.

(2) **Symmetric Property**: $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$.

(3) **Positive Definite Property**: For any $\boldsymbol{u} \in V$, $\langle \boldsymbol{u}, \boldsymbol{u} \rangle \geq 0$; and $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = 0$ if and only if $\boldsymbol{u} = 0$.

The vector space V with an inner product is called a **(real) inner product space**.

Example 2.1. For $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{R}^2$, define

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 2x_1 y_1 - x_1 y_2 - x_2 y_1 + 5x_2 y_2.$$

Then $\langle\, , \rangle$ is an inner product on $\mathbb{R}^2$. It is easy to see the linearity and the symmetric property. As for the positive definite property, note that

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{x} \rangle &= 2x_1^2 - 2x_1 x_2 + 5x_2^2 \\ &= (x_1 + x_2)^2 + (x_1 - 2x_2)^2 \geq 0. \end{aligned}$$

Moreover, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ if and only if

$$x + x_2 = 0, \quad x_1 - 2x_2 = 0,$$

which implies $x_1 = x_2 = 0$, i.e., $\boldsymbol{x} = \boldsymbol{0}$. This inner product on $\mathbb{R}^2$ is different from the dot product of $\mathbb{R}^2$.

For each vector $\boldsymbol{u} \in V$, the **norm** (also called the **length**) of $\boldsymbol{u}$ is defined as the number

$$\|\boldsymbol{u}\| := \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}.$$

If $\|\boldsymbol{u}\| = 1$, we call $\boldsymbol{u}$ a **unit vector** and $\boldsymbol{u}$ is said to be **normalized**. For any nonzero vector $\boldsymbol{v} \in V$, we have the unit vector

$$\hat{\boldsymbol{v}} = \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}.$$

This process is called **normalizing $\boldsymbol{v}$**.

Let $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ be a basis of an n-dimensional inner product space V. For vectors $\boldsymbol{u}, \boldsymbol{v} \in V$, write

$$u = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \cdots + x_n\boldsymbol{u}_n,$$
$$v = y_1\boldsymbol{u}_1 + y_2\boldsymbol{u}_2 + \cdots + y_n\boldsymbol{u}_n.$$

The linearity implies

$$\begin{aligned} \langle \boldsymbol{u}, \boldsymbol{v} \rangle &= \left\langle \sum_{i=1}^n x_i\boldsymbol{u}_i, \sum_{j=1}^n y_j\boldsymbol{u}_j \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle. \end{aligned}$$

We call the $n \times n$ matrix

$$A = \begin{bmatrix} \langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle & \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle & \cdots & \langle \boldsymbol{u}_1, \boldsymbol{u}_n \rangle \\ \langle \boldsymbol{u}_2, \boldsymbol{u}_1 \rangle & \langle \boldsymbol{u}_2, \boldsymbol{u}_2 \rangle & \cdots & \langle \boldsymbol{u}_2, \boldsymbol{u}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \boldsymbol{u}_n, \boldsymbol{u}_1 \rangle & \langle \boldsymbol{u}_n, \boldsymbol{u}_2 \rangle & \cdots & \langle \boldsymbol{u}_n, \boldsymbol{u}_n \rangle \end{bmatrix}$$

the **matrix of the inner product $\langle\,,\rangle$ relative to the basis $\mathcal{B}$**. Thus, using coordinate vectors

$$[\boldsymbol{u}]_{\mathcal{B}} = [x_1, x_2, \ldots, x_n]^T, \quad [\boldsymbol{v}]_{\mathcal{B}} = [y_1, y_2, \ldots, y_n]^T,$$

we have

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = [\boldsymbol{u}]_{\mathcal{B}}^T A [\boldsymbol{v}]_{\mathcal{B}}.$$

3 Examples of inner product spaces

Example 3.1. The vector space $\mathbb{R}^n$ with the dot product

$$\boldsymbol{u} \cdot \boldsymbol{v} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n,$$

where $\boldsymbol{u} = [a_1, a_2, \ldots, a_n]^T$, $\boldsymbol{v} = [b_1, b_2, \ldots, b_n]^T \in \mathbb{R}^n$, is an inner product space. The vector space $\mathbb{R}^n$ with this special inner product (dot product) is called the **Euclidean n-space**, and the dot product is called the **standard inner product** on $\mathbb{R}^n$.

Example 3.2. The vector space $C[a, b]$ of all real-valued continuous functions on a closed interval $[a, b]$ is an inner product space, whose inner product is defined by

$$\langle f, g \rangle = \int_a^b f(t)g(t)\mathrm{d}t, \quad f, g \in C[a, b].$$

Example 3.3. The vector space $\mathbf{M}_{m,n}$ of all $m \times n$ real matrices can be made into an inner product space under the inner product

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

where $A, B \in \mathbf{M}_{m,n}$.

For instance, when $m = 3, n = 2$, and for

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix},$$

we have

$$B^T A = \begin{bmatrix} b_{11}a_{11} + b_{21}a_{21} + b_{31}a_{31} & b_{11}a_{12} + b_{21}a_{22} + b_{31}a_{32} \\ b_{12}a_{11} + b_{22}a_{21} + b_{32}a_{31} & b_{12}a_{12} + b_{22}a_{22} + b_{32}a_{32} \end{bmatrix}.$$

Thus

$$\begin{aligned} \langle A, B \rangle &= b_{11}a_{11} + b_{21}a_{21} + b_{31}a_{31} \\ &\quad + b_{12}a_{12} + b_{22}a_{22} + b_{32}a_{32} \\ &= \sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} b_{ij}. \end{aligned}$$

This means that the inner product space $\big(\mathbf{M}_{3,2}, \langle , \rangle\big)$ is **isomorphic** to the Euclidean space $\big(\mathbb{R}^{3\times 2}, \cdot\big)$.

4 Representation of inner product

Theorem 4.1. *Let V be an n-dimensional vector space with an inner product $\langle , \rangle$, and let A be the matrix of $\langle , \rangle$ relative to a basis $\mathcal{B}$. Then for any vectors $\boldsymbol{u}, \boldsymbol{v} \in V$,*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{x}^T A \boldsymbol{y}.$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are the coordinate vectors of $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively, i.e., $\boldsymbol{x} = [\boldsymbol{u}]_{\mathcal{B}}$ and $\boldsymbol{y} = [\boldsymbol{v}]_{\mathcal{B}}$.

Example 4.1. For the inner product of $\mathbb{R}^3$ defined by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 2x_1y_1 - x_1y_2 - x_2y_1 + 5x_2y_2,$$

where $\boldsymbol{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right], \boldsymbol{y} = \left[\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right] \in \mathbb{R}^2$, its matrix relative to the standard basis $\mathcal{E} = \big\{\boldsymbol{e}_1, \boldsymbol{e}_2\big\}$ is

$$A = \begin{bmatrix} \langle \boldsymbol{e}_1, \boldsymbol{e}_1 \rangle & \langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle \\ \langle \boldsymbol{e}_2, \boldsymbol{e}_1 \rangle & \langle \boldsymbol{e}_2, \boldsymbol{e}_2 \rangle \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 5 \end{bmatrix}.$$

The inner product can be written as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T A \boldsymbol{y} = [x_1, x_2] \begin{bmatrix} 2 & -1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

We may change variables so the the inner product takes a simple form. For instance, let

$$\begin{cases} x_1 = (2/3)x_1' + (1/3)x_2' \\ x_2 = (1/3)x_1' - (1/3)x_2' \end{cases},$$

$$\begin{cases} y_1 = (2/3)y_1' + (1/3)y_2' \\ y_2 = (1/3)y_1' - (1/3)y_2' \end{cases}.$$

We have

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{y} \rangle &= 2\left(\frac{2}{3}x_1' + \frac{1}{3}x_2'\right)\left(\frac{2}{3}y_1' + \frac{1}{3}y_2'\right) \\ &\quad -\left(\frac{2}{3}x_1' + \frac{1}{3}x_2'\right)\left(\frac{1}{3}y_1' - \frac{1}{3}y_2'\right) \\ &\quad -\left(\frac{1}{3}x_1' - \frac{1}{3}x_2'\right)\left(\frac{1}{3}y_1' - \frac{1}{3}y_2'\right) \\ &\quad +5\left(\frac{1}{3}x_1' - \frac{1}{3}x_2'\right)\left(\frac{1}{3}y_1' - \frac{1}{3}y_2'\right) \\ &= x_1'y_1' + x_2'y_2' = \boldsymbol{x}'^T \boldsymbol{y}'. \end{aligned}$$

This is equivalent to choosing a new basis so that the matrix of the inner product relative to the new basis is the identity matrix.

In fact, the matrix of the inner product relative to the basis

$$\mathcal{B} = \left\{ \boldsymbol{u}_1 = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 1/3 \\ -1/3 \end{bmatrix} \right\}$$

is the identity matrix, i.e.,

$$\begin{bmatrix} \langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle & \langle \boldsymbol{u}_2, \boldsymbol{u}_1 \rangle \\ \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle & \langle \boldsymbol{u}_2, \boldsymbol{u}_2 \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Let P be the transition matrix from the standard basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ to the basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, i.e.,

$$[\boldsymbol{u}_1, \boldsymbol{u}_2] = [\boldsymbol{e}_1, \boldsymbol{e}_2]P = [\boldsymbol{e}_1, \boldsymbol{e}_2] \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix}.$$

Let $\boldsymbol{x}'$ be the coordinate vector of the vector $\boldsymbol{x}$ relative the basis $\mathcal{B}$. (The coordinate vector of $\boldsymbol{x}$ relative to the standard basis is itself $\boldsymbol{x}$.) Then

$$\boldsymbol{x} = [\boldsymbol{e}_1, \boldsymbol{e}_2]\boldsymbol{x} = [\boldsymbol{u}_1, \boldsymbol{u}_2]\boldsymbol{x}' = [\boldsymbol{e}_1, \boldsymbol{e}_2]P\boldsymbol{x}'.$$

It follows that

$$\boldsymbol{x} = P\boldsymbol{x}'.$$

Similarly, let $\boldsymbol{y}'$ be the coordinate vector of $\boldsymbol{y}$ relative to $\mathcal{B}$. Then

$$\boldsymbol{y} = P\boldsymbol{y}'.$$

Note that $\boldsymbol{x}^T = \boldsymbol{x}'^T P^T$. Thus, on the one hand by Theorem,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}'^T I_n \boldsymbol{y}' = \boldsymbol{x}'^T \boldsymbol{y}'.$$

On the other hand,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T A \boldsymbol{y} = \boldsymbol{x}'^T P^T A P \boldsymbol{y}'.$$

Theorem 4.2. *Let V be a finite-dimensional inner product space. Let A, B be matrices of the inner product relative to bases $\mathcal{B}, \mathcal{B}'$ of V, respectively. If P is the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$. Then*

$$B = P^T A P.$$

5 Cauchy-Schwarz inequality

Theorem 5.1 (Cauchy-Schwarz Inequality). *For any vectors $\boldsymbol{u}, \boldsymbol{v}$ in an inner product space V,*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \leq \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle.$$

Equivalently,

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

Proof. Consider the function

$$y = y(t) := \langle \boldsymbol{u} + t\boldsymbol{v}, \boldsymbol{u} + t\boldsymbol{v} \rangle, \quad t \in \mathbb{R}.$$

Then $y(t) \geq 0$ by the third property of inner product. Note that $y(t)$ is a quadratic function of t. In fact,

$$\begin{aligned} y(t) &= \langle \boldsymbol{u}, \boldsymbol{u} + t\boldsymbol{v} \rangle + \langle t\boldsymbol{v}, \boldsymbol{u} + t\boldsymbol{v} \rangle \\ &= \langle \boldsymbol{u}, \boldsymbol{u} \rangle + 2\langle \boldsymbol{u}, \boldsymbol{v} \rangle t + \langle \boldsymbol{v}, \boldsymbol{v} \rangle t^2. \end{aligned}$$

Thus the quadratic equation

$$\langle \boldsymbol{u}, \boldsymbol{u} \rangle + 2\langle \boldsymbol{u}, \boldsymbol{v} \rangle t + \langle \boldsymbol{v}, \boldsymbol{v} \rangle t^2 = 0$$

has at most one solution as $y(t) \geq 0$. This implies that its discriminant must be less or equal to zero, i.e.,

$$\big(2\langle \boldsymbol{u}, \boldsymbol{v} \rangle\big)^2 - 4\langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle \leq 0.$$

The Cauchy-Schwarz inequality follows. $\square$

Theorem 5.2. *The norm in an inner product space V satisfies the following properties:*

(**N1**) $\|\boldsymbol{v}\| \geq 0$; *and* $\|\boldsymbol{v}\| = 0$ *if and only if* $\boldsymbol{v} = 0$.

(**N2**) $\|c\boldsymbol{v}\| = |c| \, \|\boldsymbol{v}\|$.

(**N3**) $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$.

For nonzero vectors $\boldsymbol{u}, \boldsymbol{v} \in V$, the Cauchy-Schwarz inequality implies

$$-1 \leq \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|} \leq 1.$$

The **angle θ between $\boldsymbol{u}$ and $\boldsymbol{v}$** is defined by

$$\cos\theta = \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|}.$$

The angle exists and is unique.

6 Orthogonality

Let V be an inner product space. Two vectors $\boldsymbol{u}, \boldsymbol{v} \in V$ are said to be **orthogonal** if

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0.$$

Example 6.1. For inner product space $C[-\pi, \pi]$, the functions $\sin t$ and $\cos t$ are orthogonal as

$$\begin{aligned} \langle \sin t, \cos t \rangle &= \int_{-\pi}^{\pi} \sin t \, \cos t \, \mathrm{d}t \\ &= \frac{1}{2} \sin^2 t \Big|_{-\pi}^{\pi} = 0 - 0 = 0. \end{aligned}$$

Example 6.2. Let $\boldsymbol{u} = [a_1, a_2, \ldots, a_n]^T \in \mathbb{R}^n$. The set of all vector of the Euclidean n-space $\mathbb{R}^n$ that are orthogonal to $\boldsymbol{u}$ is a subspace of $\mathbb{R}^n$. In fact, it is the solution space of the single linear equation

$$\langle \boldsymbol{u}, \boldsymbol{x} \rangle = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0.$$

Example 6.3. Let $\boldsymbol{u} = [1, 2, 3, 4, 5]^T$, $\boldsymbol{v} = [2, 3, 4, 5, 6]^T$, and $\boldsymbol{w} = [1, 2, 3, 3, 2]^T \in \mathbb{R}^5$. The set of all vectors of $\mathbb{R}^5$ that are orthogonal to $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ is a subspace of $\mathbb{R}^5$. In fact, it is the solution space of the linear system

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 0 \\ 2x_1 + 3x_2 + 4x_3 + 5x_4 + 6x_5 = 0 \\ x_1 + 2x_2 + 3x_3 + 3x_4 + 2x_5 = 0 \end{cases}$$

Let S be a nonempty subset of an inner product space V. We denote by $S^\perp$ the set of all vectors of V that are orthogonal to every vector of S, called the **orthogonal complement** of S in V. In notation,

$$S^\perp := \Big\{ \boldsymbol{v} \in V \;\big|\; \langle \boldsymbol{v}, \boldsymbol{u} \rangle = 0 \text{ for all } \boldsymbol{u} \in S \Big\}.$$

If S contains only one vector $\boldsymbol{u}$, we write

$$\boldsymbol{u}^\perp = \Big\{ \boldsymbol{v} \in V \;\big|\; \langle \boldsymbol{v}, \boldsymbol{u} \rangle = 0 \Big\}.$$

Proposition 6.1. *Let S be a nonempty subset of an inner product space V. Then the orthogonal complement $S^\perp$ is a subspace of V.*

Proof. To show that $S^{\perp}$ is a subspace. We need to show that $S^{\perp}$ is closed under addition and scalar multiplication. Let $\boldsymbol{u}, \boldsymbol{v} \in S^{\perp}$ and $c \in \mathbb{R}$. Since $\langle \boldsymbol{u}, \boldsymbol{w} \rangle = 0$ and $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$ for all $\boldsymbol{w} \in S$, then

$$\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0,$$

$$\langle c\boldsymbol{u}, \boldsymbol{w} \rangle = c\langle \boldsymbol{u}, \boldsymbol{w} \rangle = 0$$

for all $\boldsymbol{w} \in S$. So $\boldsymbol{u} + \boldsymbol{v}, c\boldsymbol{u} \in S^{\perp}$. Hence $S^{\perp}$ is a subspace of $\mathbb{R}^n$. □

Proposition 6.2. *Let S be a subset of an inner product space V. Then every vector of $S^{\perp}$ is orthogonal to every vector of* $\operatorname{Span}(S)$*, i.e.,*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0, \quad \text{for all} \quad \boldsymbol{u} \in \operatorname{Span}(S), \ \boldsymbol{v} \in S^{\perp}.$$

Proof. For any $\boldsymbol{u} \in \operatorname{Span}(S)$, the vector $\boldsymbol{u}$ must be a linear combination of some vectors in S, say,

$$\boldsymbol{u} = a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + \cdots + a_k\boldsymbol{u}_k.$$

Then for any $\boldsymbol{v} \in S^{\perp}$,

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = a_1\langle \boldsymbol{u}_1, \boldsymbol{v} \rangle + a_2\langle \boldsymbol{u}_2, \boldsymbol{v} \rangle + \cdots + a_n\langle \boldsymbol{u}_n, \boldsymbol{v} \rangle = 0.$$

□

Example 6.4. Let A be an $m \times n$ real matrix. Then $\operatorname{Nul} A$ and $\operatorname{Row} A$ are orthogonal complements of each other in $\mathbb{R}^n$, i.e.,

$$\operatorname{Nul} A = \big(\operatorname{Row} A\big)^{\perp}, \quad \big(\operatorname{Nul} A\big)^{\perp} = \operatorname{Row} A.$$

7 Orthogonal sets and bases

Let V be an inner product space. A subset $S = \big\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\big\}$ of nonzero vectors of V is called an **orthogonal set** if every pair of vectors are orthogonal, i.e.,

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = 0, \quad 1 \leq i < j \leq k.$$

An orthogonal set $S = \big\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\big\}$ is called an **orthonormal set** if we further have

$$\|\boldsymbol{u}_i\| = 1, \quad 1 \leq i \leq k.$$

An **orthonormal basis** of V is a basis which is also an orthonormal set.

Theorem 7.1 (Pythagoras). *Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ be mutually orthogonal vectors. Then*

$$\|\boldsymbol{v}_1 + \boldsymbol{v}_2 + \cdots + \boldsymbol{v}_k\|^2 = \|\boldsymbol{v}_1\|^2 + \|\boldsymbol{v}_2\|^2 + \cdots + \|\boldsymbol{v}_k\|^2.$$

Proof. For simplicity, we assume $k = 2$. If $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal, i.e., $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$, then

$$\begin{aligned} \|\boldsymbol{u} + \boldsymbol{v}\|^2 &= \langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{u} + \boldsymbol{v} \rangle \\ &= \langle \boldsymbol{u}, \boldsymbol{u} \rangle + \langle \boldsymbol{v}, \boldsymbol{v} \rangle \\ &= \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2. \end{aligned}$$

□

Example 7.1. The three vectors

$$\boldsymbol{v}_1 = [1, 2, 1]^T, \quad \boldsymbol{v}_2 = [2, 1, -4]^T, \quad \boldsymbol{v}_3 = [3, -2, 1]^T$$

are mutually orthogonal. Express the the vector $\boldsymbol{v} = [7, 1, 9]^T$ as a linear combination of $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$.

Set

$$x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + x_3\boldsymbol{v}_3 = \boldsymbol{v}.$$

There are two ways to find x_1, x_2, x_3.

Method 1: Solving the linear system by performing row operations to its augmented matrix

$$\big[\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \mid \boldsymbol{v}\big],$$

we obtain $x_1 = 3$, $x_2 = -1$, $x_3 = 2$. So $\boldsymbol{v} = 3\boldsymbol{v}_1 - \boldsymbol{v}_2 + 2\boldsymbol{v}_3$.

Method 2: Since $\boldsymbol{v}_i \perp \boldsymbol{v}_j$ for $i \neq j$, we have

$$\langle \boldsymbol{v}, \boldsymbol{v}_i \rangle = \langle x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + x_3\boldsymbol{v}_3, \boldsymbol{v}_i \rangle = x_i \langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle,$$

where $i = 1, 2, 3$. Then

$$x_i = \frac{\langle \boldsymbol{v}, \boldsymbol{v}_i \rangle}{\langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle}, \quad i = 1, 2, 3.$$

We then have

$$\begin{aligned} x_1 &= \frac{7+2+9}{1+4+1} = \frac{18}{6} = 3, \\ x_2 &= \frac{14+1-36}{4+1+16} = \frac{-21}{21} = -1, \\ x_3 &= \frac{21-2+9}{9+4+1} = \frac{28}{14} = 2. \end{aligned}$$

Theorem 7.2. *Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ be an orthogonal basis of a subspace W. Then for any $\boldsymbol{w} \in W$,*

$$\boldsymbol{w} = \frac{\langle \boldsymbol{v}_1, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{v}_2, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{v}_k, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}_k, \boldsymbol{v}_k \rangle}\boldsymbol{v}_k.$$

Proof. Trivial. □

8 Orthogonal projection

Let V be an inner product space. Let $\boldsymbol{v}$ be a nonzero vector of V. We want to decompose an arbitrary vector $\boldsymbol{y}$ into the form

$$\boldsymbol{y} = \alpha\boldsymbol{v} + \boldsymbol{z}, \quad \text{where} \quad \boldsymbol{z} \in \boldsymbol{v}^{\perp}.$$

Since $\boldsymbol{z} \perp \boldsymbol{v}$, we have

$$\langle \boldsymbol{v}, \boldsymbol{y} \rangle = \langle \alpha\boldsymbol{v}, \boldsymbol{v} \rangle = \alpha \langle \boldsymbol{v}, \boldsymbol{v} \rangle.$$

This implies that

$$\alpha = \frac{\langle \boldsymbol{v}, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}.$$

We define the vector

$$\mathrm{Proj}_{\boldsymbol{v}}(\boldsymbol{y}) = \frac{\langle \boldsymbol{v}, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}\boldsymbol{v},$$

called the **orthogonal projection of $\boldsymbol{y}$ along $\boldsymbol{v}$**. The linear transformation $\mathrm{Proj}_{\boldsymbol{u}} : V \to V$ is called the **orthogonal projection of V onto the direction $\boldsymbol{v}$**.

Proposition 8.1. *Let $\boldsymbol{v}$ be a nonzero vector of the Euclidean n-space $\mathbb{R}^n$. Then the orthogonal projection $\mathrm{Proj}_{\boldsymbol{u}} : \mathbb{R}^n \to \mathbb{R}^n$ is given by*

$$\mathrm{Proj}_{\boldsymbol{v}}(\boldsymbol{y}) = \frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\boldsymbol{v}\boldsymbol{v}^T\boldsymbol{y};$$

and the orthogonal projection $\mathrm{Proj}_{\boldsymbol{v}^{\perp}} : \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\mathrm{Proj}_{\boldsymbol{v}^{\perp}}(\boldsymbol{y}) = \left(I - \frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\boldsymbol{v}\boldsymbol{v}^T\right)\boldsymbol{y}.$$

Write the vector $\boldsymbol{v}$ as $\boldsymbol{v} = [a_1, a_2, \ldots, a_n]^T$. The for any scalar c,

$$c\boldsymbol{v} = \begin{bmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_n \end{bmatrix} = \begin{bmatrix} a_1 c \\ a_2 c \\ \vdots \\ a_n c \end{bmatrix} = \boldsymbol{v}[c],$$

where $[c]$ is the 1×1 matrix with the only entry c. Note that

$$[\boldsymbol{v} \cdot \boldsymbol{y}] = \boldsymbol{v}^T \boldsymbol{y}.$$

Then the orthogonal projection $\mathrm{Proj}_{\boldsymbol{v}}$ can be written as

$$\begin{aligned} \mathrm{Proj}_{\boldsymbol{v}}(\boldsymbol{y}) &= \left(\frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\right)(\boldsymbol{v} \cdot \boldsymbol{y})\boldsymbol{v} \\ &= \left(\frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\right)\boldsymbol{v}[\boldsymbol{v} \cdot \boldsymbol{y}] \\ &= \left(\frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\right)\boldsymbol{v}\boldsymbol{v}^T\boldsymbol{y}. \end{aligned}$$

This means that the standard matrix of $\mathrm{Proj}_{\boldsymbol{v}}$ is

$$\left(\frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\right)\boldsymbol{v}\boldsymbol{v}^T.$$

Indeed, $\boldsymbol{v}$ is an $n \times 1$ matrix and $\boldsymbol{v}^T$ is a $1 \times n$ matrix, the product $\boldsymbol{v}\boldsymbol{v}^T$ is an $n \times n$ matrix.

The orthogonal projection $\mathrm{Proj}_{\boldsymbol{v}^\perp} : \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\mathrm{Proj}_{\boldsymbol{v}^\perp}(\boldsymbol{y}) = \boldsymbol{y} - \mathrm{Proj}_{\boldsymbol{v}}(\boldsymbol{y}) = \left(I - \frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\boldsymbol{v}\boldsymbol{v}^T\right)\boldsymbol{y}.$$

This means that the standard matrix of $\mathrm{Proj}_{\boldsymbol{v}^\perp}$ is

$$I - \left(\frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\right)\boldsymbol{v}\boldsymbol{v}^T.$$

Example 8.1. Find the linear mapping from $\mathbb{R}^3$ to $\mathbb{R}^3$ that is a the orthogonal projection of $\mathbb{R}^3$ ont the plane $x_1 + x_2 + x_3 = 0$.

To find the orthogonal projection of $\mathbb{R}^3$ onto the subspace $\boldsymbol{v}^\perp$, where $\boldsymbol{v} = [1, 1, 1]^T$, we find the following orthogonal projection

$$\begin{aligned} \mathrm{Proj}_{\boldsymbol{v}}(\boldsymbol{y}) &= \left(\frac{\boldsymbol{v} \cdot \boldsymbol{y}}{\boldsymbol{v} \cdot \boldsymbol{v}}\right)\boldsymbol{v} = \frac{y_1 + y_2 + y_3}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}. \end{aligned}$$

Then the orthogonal projection of $\boldsymbol{y}$ onto $\boldsymbol{v}^\perp$ is given by

$$\begin{aligned} \mathrm{Proj}_{\boldsymbol{v}^\perp}\boldsymbol{y} &= \boldsymbol{y} - \mathrm{Proj}_{\boldsymbol{v}}(\boldsymbol{y}) = \left(I - \frac{1}{\boldsymbol{v} \cdot \boldsymbol{v}}\boldsymbol{v}\boldsymbol{v}^T\right)\boldsymbol{y} \\ &= \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}. \end{aligned}$$

Let W be a subspace of V, and let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ be an orthogonal basis of W. We want to decompose an arbitrary vector $\boldsymbol{y} \in V$ into the form
$$\boldsymbol{y} = \boldsymbol{w} + \boldsymbol{z}$$
with $\boldsymbol{w} \in W$ and $\boldsymbol{z} \in W^{\perp}$. Then there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_k$ such that
$$\hat{\boldsymbol{y}} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_k \boldsymbol{v}_k.$$
Since $\boldsymbol{z} \perp \boldsymbol{v}_1$, $z \perp \boldsymbol{v}_2$, …, $\boldsymbol{z} \perp \boldsymbol{v}_k$, we have
$$\langle \boldsymbol{v}_i, \boldsymbol{y} \rangle = \langle \boldsymbol{v}_i, \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_r \boldsymbol{v}_k + \boldsymbol{z} \rangle = \alpha_i \langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle.$$
Then
$$\alpha_i = \frac{\langle \boldsymbol{v}_i, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle}, \quad 1 \le i \le k.$$
We thus define
$$\mathrm{Proj}_W(\boldsymbol{y}) = \frac{\langle \boldsymbol{v}_1, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} \boldsymbol{v}_1 + \frac{\langle \boldsymbol{v}_2, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle} \boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{v}_k, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}_k, \boldsymbol{v}_k \rangle} \boldsymbol{v}_k,$$
called the **orthogonal projection of $\boldsymbol{v}$ along W**. The linear transformation
$$\mathrm{Proj}_W : V \to V$$
is called the **orthogonal projection of V onto W**.

Theorem 8.2. *Let V be an n-dimensional inner product space. Let W be a subspace with an orthogonal basis $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$. Then for any $v \in V$,*
$$\mathrm{Proj}_W(\boldsymbol{y}) = \frac{\langle \boldsymbol{v}_1, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} \boldsymbol{v}_1 + \frac{\langle \boldsymbol{v}_2, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle} \boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{v}_k, \boldsymbol{y} \rangle}{\langle \boldsymbol{v}_k, \boldsymbol{v}_k \rangle} \boldsymbol{v}_k,$$
$$\mathrm{Proj}_{W^{\perp}}(\boldsymbol{y}) = \boldsymbol{y} - \mathrm{Proj}_W(\boldsymbol{y}).$$
In particular, if $\mathcal{B}$ is an orthonormal basis of W, then
$$\mathrm{Proj}_W(\boldsymbol{y}) = \langle \boldsymbol{v}_1, \boldsymbol{y} \rangle \boldsymbol{v}_1 + \langle \boldsymbol{v}_2, \boldsymbol{y} \rangle \boldsymbol{v}_2 + \cdots + \langle \boldsymbol{v}_k, \boldsymbol{y} \rangle \boldsymbol{v}_k.$$

Proposition 8.3. *Let W be a subspace of $\mathbb{R}^n$. Let $U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k]$ be an $n \times k$ matrix, whose columns form an orthonormal basis of W. Then the orthogonal projection $\mathrm{Proj}_W : \mathbb{R}^n \to \mathbb{R}^n$ is given by*
$$\mathrm{Proj}_W(\boldsymbol{y}) = UU^T \boldsymbol{y}.$$

Proof. For any $\boldsymbol{y} \in \mathbb{R}^n$, we have
$$\mathrm{Proj}_W(\boldsymbol{y}) = (\boldsymbol{u}_1 \cdot \boldsymbol{y})\boldsymbol{u}_1 + (\boldsymbol{u}_2 \cdot \boldsymbol{y})\boldsymbol{u}_2 + \cdots + (\boldsymbol{u}_k \cdot \boldsymbol{y})\boldsymbol{u}_k.$$
Note that
$$U^T \boldsymbol{y} = \begin{bmatrix} \boldsymbol{u}_1^T \\ \boldsymbol{u}_2^T \\ \vdots \\ \boldsymbol{u}_k^T \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} \boldsymbol{u}_1^T \boldsymbol{y} \\ \boldsymbol{u}_2^T \boldsymbol{y} \\ \vdots \\ \boldsymbol{u}_k^T \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 \cdot \boldsymbol{y} \\ \boldsymbol{u}_2 \cdot \boldsymbol{y} \\ \vdots \\ \boldsymbol{u}_k \cdot \boldsymbol{y} \end{bmatrix}.$$
Then
$$\begin{aligned} UU^T \boldsymbol{y} &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k] \begin{bmatrix} \boldsymbol{u}_1 \cdot \boldsymbol{y} \\ \boldsymbol{u}_2 \cdot \boldsymbol{y} \\ \vdots \\ \boldsymbol{u}_k \cdot \boldsymbol{y} \end{bmatrix} \\ &= \mathrm{Proj}_W(\boldsymbol{y}). \end{aligned}$$
□

Example 8.2. Find the orthogonal projection

$$\mathrm{Proj}_W : \mathbb{R}^3 \to \mathbb{R}^3,$$

where W is the plane $x_1 + x_2 + x_3 = 0$.

By inspection, the following two vectors

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

form an orthogonal basis of W. Then

$$\begin{aligned}
\mathrm{Proj}_W(\boldsymbol{y}) &= \left(\frac{\boldsymbol{v}_1 \cdot \boldsymbol{y}}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right)\boldsymbol{v}_1 + \left(\frac{\boldsymbol{v}_2 \cdot \boldsymbol{y}}{\boldsymbol{v}_2 \cdot \boldsymbol{v}2}\right)\boldsymbol{v}_2 \\
&= \frac{y_1 - y_2}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \\
&\quad + \frac{y_1 + y_2 - 2y_3}{6}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \\
&= \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.
\end{aligned}$$

Example 8.3. Find the matrix of the orthogonal projection

$$\mathrm{Proj}_W : \mathbb{R}^3 \to \mathbb{R}^3,$$

where

$$W = \mathrm{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right\}.$$

The following two vectors

$$\boldsymbol{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \boldsymbol{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}$$

form an orthonormal basis of W. Then the standard matrix of Proj_W is the product

$$\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{2} \\ 1/\sqrt{3} & 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix},$$

which results the matrix

$$\begin{bmatrix} 5/6 & -1/6 & 1/3 \\ -1/6 & 5/6 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Alternatively, the matrix can be found by computing the orthogonal projection:

$$\begin{aligned}
\mathrm{Proj}_W(\boldsymbol{y}) &= \frac{y_1 + y_2 + y_3}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{y_1 - y_2}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \\
&= \frac{1}{6}\begin{bmatrix} 5y_1 - y_2 + 2y_3 \\ -y_1 + 5y_2 + 2y_3 \\ 2y_1 + 2y_2 + 2y_3 \end{bmatrix} \\
&= \begin{bmatrix} 5/6 & -1/6 & 1/3 \\ -1/6 & 5/6 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.
\end{aligned}$$

9 Gram-Schmidt process

Let W be a subspace of an inner product space V. Let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ be a basis of W, not necessarily orthogonal. An orthogonal basis $\mathcal{B}' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$ may be constructed from $\mathcal{B}$ as follows:

$$\begin{aligned}
\boldsymbol{w}_1 &= \boldsymbol{v}_1, & W_1 &= \mathrm{Span}\,\{\boldsymbol{w}_1\}, \\
\boldsymbol{w}_2 &= \boldsymbol{v}_2 - \mathrm{Proj}_{W_1}(\boldsymbol{v}_2), & W_2 &= \mathrm{Span}\,\{\boldsymbol{w}_1, \boldsymbol{w}_2\}, \\
\boldsymbol{w}_3 &= \boldsymbol{v}_3 - \mathrm{Proj}_{W_2}(\boldsymbol{v}_3), & W_3 &= \mathrm{Span}\,\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}, \\
&\vdots \\
\boldsymbol{w}_{k-1} &= \boldsymbol{v}_{k-1} - \mathrm{Proj}_{W_{k-1}}(\boldsymbol{v}_{k-1}), & W_{k-1} &= \mathrm{Span}\,\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{k-1}\}, \\
\boldsymbol{w}_k &= \boldsymbol{v}_k - \mathrm{Proj}_{W_{k-1}}(\boldsymbol{v}_k).
\end{aligned}$$

More precisely,

$$\begin{aligned}
\boldsymbol{w}_1 &= \boldsymbol{v}_1, \\
\boldsymbol{w}_2 &= \boldsymbol{v}_2 - \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_2\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle}\boldsymbol{w}_1, \\
\boldsymbol{w}_3 &= \boldsymbol{v}_3 - \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle}\boldsymbol{w}_1 - \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2\rangle}\boldsymbol{w}_2, \\
&\vdots \\
\boldsymbol{w}_k &= \boldsymbol{v}_k - \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle}\boldsymbol{w}_1 - \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2\rangle}\boldsymbol{w}_2 - \cdots - \frac{\langle \boldsymbol{w}_{k-1}, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_{k-1}, \boldsymbol{w}_{k-1}\rangle}\boldsymbol{w}_{k-1}.
\end{aligned}$$

The method of constructing the orthogonal vector $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k$ is known as the **Gram-Schmidt process**.

Clearly, the vector $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k$ are linear combinations of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$. Conversely, the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are also linear combinations of $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k$:

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{w}_1, \\
\boldsymbol{v}_2 &= \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_2\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle}\boldsymbol{w}_1 + \boldsymbol{w}_2, \\
\boldsymbol{v}_3 &= \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle}\boldsymbol{w}_1 + \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2\rangle}\boldsymbol{w}_2 + \boldsymbol{w}_3, \\
&\vdots \\
\boldsymbol{v}_k &= \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle}\boldsymbol{w}_1 + \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2\rangle}\boldsymbol{w}_2 + \cdots + \frac{\langle \boldsymbol{w}_{k-1}, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_{k-1}, \boldsymbol{w}_{k-1}\rangle}\boldsymbol{w}_{k-1} + \boldsymbol{w}_k.
\end{aligned}$$

Hence

$$\mathrm{Span}\,\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} = \mathrm{Span}\,\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}.$$

Since $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is a basis for W, so is the set $\mathcal{B}' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$.

Theorem 9.1. *The basis $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$ constructed by the Gram-Schmidt process is an orthogonal basis of W. Moreover,*

$$\big[\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_k\big] = \big[\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \ldots, \boldsymbol{w}_k\big]R,$$

where R is the $k \times k$ upper triangular matrix

$$\begin{bmatrix}
1 & \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_2\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle} & \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle} & \cdots & \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle} \\
0 & 1 & \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2\rangle} & \cdots & \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2\rangle} \\
0 & 0 & 1 & \cdots & \frac{\langle \boldsymbol{w}_3, \boldsymbol{v}_k\rangle}{\langle \boldsymbol{w}_3, \boldsymbol{w}_3\rangle} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}.$$

Example 9.1. Let W be the subspace of $\mathbb{R}^4$ spanned by

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \ \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \ \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Construct an orthogonal basis for W.

Set $\boldsymbol{w}_1 = \boldsymbol{v}_1$. Let $W_1 = \operatorname{Span}\{\boldsymbol{w}_1\}$. To find a vector $\boldsymbol{w}_2$ in W that is orthogonal to W_1, set

$$\begin{aligned} \boldsymbol{w}_2 &= \boldsymbol{v}_2 - \operatorname{Proj}_{W_1}\boldsymbol{v}_2 = \boldsymbol{v}_2 - \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_2\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle}\boldsymbol{w}_1 \\ &= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ -3 \end{bmatrix}. \end{aligned}$$

Let $W_2 = \operatorname{Span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$. To find a vector $\boldsymbol{w}_3$ in W that is orthogonal to W_2, set

$$\begin{aligned} \boldsymbol{w}_3 &= \boldsymbol{v}_3 - \operatorname{Proj}_{W_2}\boldsymbol{v}_3 \\ &= \boldsymbol{v}_3 - \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle}\boldsymbol{w}_1 - \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2\rangle}\boldsymbol{w}_2 \\ &= \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{\frac{2}{4}}{\frac{12}{16}} \cdot \frac{1}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ -3 \end{bmatrix} \\ &= \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \\ 0 \end{bmatrix}. \end{aligned}$$

Then the set $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ is an orthogonal basis for W.

Theorem 9.2. *Any $m \times n$ real matrix A can be written as*

$$A = QR,$$

*called a QR-**decomposition**, where Q is an $m \times n$ matrix whose columns are mutually orthogonal, and R is an $n \times n$ upper triangular matrix whose diagonal entries are* 1.

Example 9.2. Find a QR-decomposition of the matrix

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

Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$ be the column vectors of A. Set

$$\boldsymbol{w}_1 = \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Then $\boldsymbol{v}_1 = \boldsymbol{w}_1$. Set

$$\begin{aligned}
\boldsymbol{w}_2 &= \boldsymbol{v}_2 - \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_2 \rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1 \rangle} \boldsymbol{w}_1 \\
&= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \end{bmatrix}.
\end{aligned}$$

Then $\boldsymbol{v}_2 = (1/2)\boldsymbol{w}_1 + \boldsymbol{w}_2$. Set

$$\begin{aligned}
\boldsymbol{w}_3 &= \boldsymbol{v}_3 - \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_3 \rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1 \rangle} \boldsymbol{w}_1 - \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_3 \rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2 \rangle} \boldsymbol{w}_2 \\
&= \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \end{bmatrix} = \boldsymbol{0}.
\end{aligned}$$

Then $\boldsymbol{v}_3 = (3/2)\boldsymbol{w}_1 + \boldsymbol{w}_2 + \boldsymbol{w}_3$. Set

$$\begin{aligned}
\boldsymbol{w}_4 &= \boldsymbol{v}_4 - \frac{\langle \boldsymbol{w}_1, \boldsymbol{v}_4 \rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1 \rangle} \boldsymbol{w}_1 - \frac{\langle \boldsymbol{w}_2, \boldsymbol{v}_4 \rangle}{\langle \boldsymbol{w}_2, \boldsymbol{w}_2 \rangle} \boldsymbol{w}_2 \\
&= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \end{bmatrix} \\
&= \begin{bmatrix} -2/3 \\ 2/3 \\ 2/3 \end{bmatrix}.
\end{aligned}$$

Then $\boldsymbol{v}_4 = (1/2)\boldsymbol{w}_1 + (1/3)\boldsymbol{w}_2 + \boldsymbol{w}_4$. Thus matrixes Q and R for QR-decomposition of A are as follows:

$$Q = \begin{bmatrix} 1 & 1/2 & 0 & -2/3 \\ 0 & 1 & 0 & 2/3 \\ 1 & -1/2 & 0 & 2/3 \end{bmatrix},$$

$$R = \begin{bmatrix} 1 & 1/2 & 3/2 & 1/2 \\ 0 & 1 & 1 & 1/3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

10 Orthogonal matrix

Let V be an n-dimensional inner product space. A linear transformation $T : V \to V$ is called an **isometry** if for any $\boldsymbol{v} \in V$,

$$\|T(\boldsymbol{v})\| = \|\boldsymbol{v}\|.$$

Example 10.1. For the Euclidean n-space $\mathbb{R}^n$ with the dot product, rotations and reflections are isometries.

Theorem 10.1. *A linear transformation $T : V \to V$ is an isometry if and only if T preserving inner product, i.e., for $\boldsymbol{u}, \boldsymbol{v} \in V$,*

$$\langle T(\boldsymbol{u}), T(\boldsymbol{v}) \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle.$$

Proof. Note that for vectors $\boldsymbol{u}, \boldsymbol{v} \in V$,

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

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v} \rangle = \langle \boldsymbol{u}, \boldsymbol{u} \rangle + \langle \boldsymbol{v}, \boldsymbol{v} \rangle + 2\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 + 2\langle \boldsymbol{u}, \boldsymbol{v} \rangle.$$

It is clear that the length preserving is equivalent to the inner product preserving. □

An $n \times n$ matrix Q is called **orthogonal** if $QQ^T = I$, i.e.,

$$Q^{-1} = Q^T.$$

Theorem 10.2. *Let Q be an $n \times n$ matrix. The following are equivalent.*

(a) *Q is orthogonal.*

(b) *Q^T is orthogonal.*

(c) *The column vectors of Q are orthonormal.*

(d) *The row vectors of Q are orthonormal.*

Proof. "$(a) \Leftrightarrow (b)$": If $QQ^T = I$, then $Q^{-1} = Q^T$. So $Q^T(Q^T)^T Q^T Q = I$. This means that Q^T is orthogonal.

"$(a) \Leftrightarrow (c)$": Let $Q = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n]$. Note that

$$Q^T Q = \begin{bmatrix} \boldsymbol{u}_1^T \\ \boldsymbol{u}_2^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix} [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n] = \begin{bmatrix} \boldsymbol{u}_1^T\boldsymbol{u}_1 & \boldsymbol{u}_1^T\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_1^T\boldsymbol{u}_n \\ \boldsymbol{u}_2^T\boldsymbol{u}_1 & \boldsymbol{u}_2^T\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_2^T\boldsymbol{u}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{u}_n^T\boldsymbol{u}_1 & \boldsymbol{u}_n^T\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n^T\boldsymbol{u}_n \end{bmatrix}.$$

Thus $Q^T Q = I$ is equivalent to $\boldsymbol{u}_i^T\boldsymbol{u}_j = 1$ for $i = j$ and $\boldsymbol{u}_i^T\boldsymbol{u}_j = 0$ for $i \neq j$. This means that Q is orthogonal if and only if $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ is an orthonormal basis of V. $\square$

Theorem 10.3. *Let V be an n-dimensional inner product space with an orthonormal basis $\mathcal{B} = \big\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big\}$. Let P be an $n \times n$ real matrix, and*

$$\big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big] = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big]P.$$

Then $\mathcal{B}' = \big\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big\}$ is an orthonormal basis if and only if P is an orthogonal matrix.

Proof. For simplicity, we assume $n = 3$. Since

$$\big[\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\big] = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\big]P, \quad \text{i.e.,}$$

$$\boldsymbol{v}_1 = p_{11}\boldsymbol{u}_1 + p_{21}\boldsymbol{u}_2 + p_{31}\boldsymbol{u}_3,$$

$$\boldsymbol{v}_2 = p_{12}\boldsymbol{u}_1 + p_{22}\boldsymbol{u}_2 + p_{32}\boldsymbol{u}_3,$$

$$\boldsymbol{v}_3 = p_{13}\boldsymbol{u}_1 + p_{23}\boldsymbol{u}_2 + p_{33}\boldsymbol{u}_3.$$

We then have

$$\begin{aligned} \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle &= \langle p_{1i}\boldsymbol{u}_1 + p_{2i}\boldsymbol{u}_2 + p_{3i}\boldsymbol{u}_3, p_{1j}\boldsymbol{u}_1 + p_{2j}\boldsymbol{u}_2 + p_{3j}\boldsymbol{u}_3 \rangle \\ &= p_{1i}p_{1j} + p_{2i}p_{2j} + p_{3i}p_{3j}. \end{aligned}$$

Note that $\mathcal{B}'$ is an orthonormal basis is equivalent to

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \delta_{ij},$$

and P is an orthogonal matrix if and only if

$$p_{1i}p_{1j} + p_{2i}p_{2j} + p_{3i}p_{3j} = \delta_{ij}.$$

The proof is finished. $\square$

Theorem 10.4. *Let V be an n-dimensional inner product space with an orthonormal basis $\mathcal{B} = \big\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big\}$. Let $T : V \to V$ be a linear transformation. Then T is an isometry if and only if the matrix of T relative to $\mathcal{B}$ is an orthogonal matrix.*

Proof. Let A be the matrix of T relative to the basis $\mathcal{B}$. Then

$$[T(\boldsymbol{u}_1), T(\boldsymbol{u}_2), \ldots, T(\boldsymbol{u}_n)] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n]A.$$

Note that T is an isometry if and only if $\{T(\boldsymbol{u}_1), T(\boldsymbol{u}_2), \ldots, T(\boldsymbol{u}_n)\}$ is an orthonormal basis of V, and that $\{T(\boldsymbol{u}_1), T(\boldsymbol{u}_2), \ldots, T(\boldsymbol{u}_n)\}$ is an orthonormal basis if and only the transition matrix A is an orthogonal matrix. □

Example 10.2. The matrix $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & -2 \end{bmatrix}$ is not an orthogonal matrix. The set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}$$

of the column vectors of A is an orthogonal basis of $\mathbb{R}^3$. However, the set of the row vectors of A is *not* an orthogonal set.

The matrix

$$U = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{bmatrix}$$

is an orthogonal matrix. For the vector $\boldsymbol{v} = [3, 0, 4]^T$, we have

$$U\boldsymbol{v} = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} \sqrt{3} + 4/\sqrt{6} \\ \sqrt{3} + 4/\sqrt{6} \\ \sqrt{3} - 8/\sqrt{6} \end{bmatrix}.$$

The length of $\boldsymbol{v}$ is

$$\|\boldsymbol{v}\| = \sqrt{3^2 + 0^2 + 4^2} = \sqrt{25} = 5$$

and the length of $U\boldsymbol{v}$ is

$$\|U\boldsymbol{v}\| = \sqrt{2\left(\sqrt{3} + 4/\sqrt{6}\right)^2 + \left(\sqrt{3} - 8/\sqrt{6}\right)^2} = 5.$$

11 Diagonalizing real symmetric matrices

Let V be an n-dimensional real inner product space. A linear mapping $T : V \to V$ is said to be **symmetric** if

$$\langle T(\boldsymbol{u}), \boldsymbol{v} \rangle = \langle \boldsymbol{u}, T(\boldsymbol{v}) \rangle \quad \text{for all} \quad \boldsymbol{u}, \boldsymbol{v} \in V.$$

Example 11.1. Let A be a real symmetric $n \times n$ matrix. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $T(\boldsymbol{x}) = A\boldsymbol{x}$. Then T is symmetric for the Euclidean n-space. In fact, for $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, we have

$$\begin{aligned} T(\boldsymbol{u}) \cdot \boldsymbol{v} &= (A\boldsymbol{u}) \cdot \boldsymbol{v} = (A\boldsymbol{u})^T \boldsymbol{v} = \boldsymbol{u}^T A^T \boldsymbol{v} \\ &= \boldsymbol{u}^T A\boldsymbol{v} = \boldsymbol{u} \cdot A\boldsymbol{v} = \boldsymbol{u} \cdot T(\boldsymbol{v}). \end{aligned}$$

Proposition 11.1. *Let V be an n-dimensional real inner product space with an orthonormal basis $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$. Let $T : V \to V$ be a linear mapping whose matrix relative to $\mathcal{B}$ is A. Then T is symmetric if and only the matrix A is symmetric.*

Proof. Note that

$$\big[T(\boldsymbol{u}_1), T(\boldsymbol{u}_2), \ldots, T(\boldsymbol{u}_n)\big] = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Alternatively,

$$T(\boldsymbol{u}_j) = \sum_{i=1}^{n} a_{ij}\boldsymbol{u}_i, \quad 1 \le j \le n.$$

If T is symmetric, then

$$a_{ij} = \langle \boldsymbol{u}_i, T(\boldsymbol{u}_j)\rangle = \langle T(\boldsymbol{u}_i), \boldsymbol{u}_j\rangle = a_{ji}.$$

So A is symmetric.

Conversely, if A is symmetric, then for vectors $\boldsymbol{u} = \sum_{i=1}^{n} a_i\boldsymbol{u}_i$, $\boldsymbol{v} = \sum_{i=1}^{n} b_i\boldsymbol{u}_i$, we have

$$\begin{aligned}
\langle T(\boldsymbol{u}), \boldsymbol{v}\rangle &= \sum_{i,j=1}^{n} a_i b_j \langle T(\boldsymbol{u}_i), \boldsymbol{u}_j\rangle = \sum_{i,j=1}^{n} a_i b_j a_{ji} \\
&= \sum_{i,j=1}^{n} a_i b_j a_{ij} = \sum_{i,j=1}^{n} a_i b_j \langle \boldsymbol{u}_i, T(\boldsymbol{u}_j)\rangle \\
&= \langle \boldsymbol{u}, T(\boldsymbol{v})\rangle.
\end{aligned}$$

So T is symmetric. □

Theorem 11.2. *The roots of characteristic polynomial of a real symmetric matrix A are all real numbers.*

Proof. Let λ be a (possible complex) root of the characteristic polynomial of A, and let $\boldsymbol{v}$ be a (possible complex) eigenvector for the eigenvalue λ. Then

$$A\boldsymbol{v} = \lambda\boldsymbol{v}.$$

Note that

$$\begin{aligned}
\lambda\|\boldsymbol{v}\|^2 &= \lambda\boldsymbol{v}\cdot\bar{\boldsymbol{v}} = (A\boldsymbol{v})\cdot\bar{\boldsymbol{v}} = (A\boldsymbol{v})^T\bar{\boldsymbol{v}} = \boldsymbol{v}^T A^T\bar{\boldsymbol{v}} \\
&= \boldsymbol{v}^T A\bar{\boldsymbol{v}} = \boldsymbol{v}^T\bar{A}\bar{\boldsymbol{v}} = \boldsymbol{v}^T\overline{A\boldsymbol{v}} = \boldsymbol{v}^T\overline{\lambda\boldsymbol{v}} \\
&= \boldsymbol{v}^T(\bar{\lambda}\bar{\boldsymbol{v}}) = \bar{\lambda}\boldsymbol{v}^T\bar{\boldsymbol{v}} = \bar{\lambda}\boldsymbol{v}\cdot\bar{\boldsymbol{v}} = \bar{\lambda}\|\boldsymbol{v}\|^2.
\end{aligned}$$

Thus

$$(\lambda - \bar{\lambda})\|\boldsymbol{v}\|^2 = 0.$$

Since $\|\boldsymbol{v}\| \neq 0$, it follows that $\lambda = \bar{\lambda}$. So λ is a real number. □

Theorem 11.3. *Let λ and μ be distinct eigenvalues of a symmetric linear transformation $T : V \to V$. Then eigenvectors for λ are orthogonal to eigenvectors for μ.*

Proof. Let $\boldsymbol{u}$ be an eigenvector for λ, and let $\boldsymbol{v}$ be an eigenvectors for μ, i.e., $T(\boldsymbol{u}) = \lambda\boldsymbol{u}$, $T(\boldsymbol{v}) = \mu\boldsymbol{v}$. Then

$$\begin{aligned}
\lambda\langle \boldsymbol{u}, \boldsymbol{v}\rangle &= \langle \lambda\boldsymbol{u}, \boldsymbol{v}\rangle = \langle T(\boldsymbol{u}), \boldsymbol{v}\rangle \\
&= \langle \boldsymbol{u}, T(\boldsymbol{v})\rangle = \langle \boldsymbol{u}, \mu\boldsymbol{v}\rangle = \mu\langle \boldsymbol{u}, \boldsymbol{v}\rangle.
\end{aligned}$$

Thus

$$(\lambda - \mu)\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0.$$

Since $\lambda - \mu \neq 0$, it follows that $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0$. □

Theorem 11.4. *Let V be an n-dimensional real inner product space. Let $T : V \to V$ be a symmetric linear mapping. Then V has an orthonormal basis of eigenvectors of T.*

Proof. We proceed by induction on n. For $n = 1$, it is obviously true. Let λ_1 be an eigenvalue of T, and let $\boldsymbol{u}_1$ be a unit eigenvector of T for the eigenvalue λ_1. Let $W := \boldsymbol{u}_1^\perp$. For any $\boldsymbol{w} \in W$,

$$\begin{aligned}
\langle T(\boldsymbol{w}), \boldsymbol{u}_1\rangle &= \langle \boldsymbol{w}, T(\boldsymbol{u}_1)\rangle = \langle \boldsymbol{w}, \lambda_1\boldsymbol{u}_1\rangle \\
&= \lambda_1\langle \boldsymbol{w}, \boldsymbol{u}_1\rangle = 0.
\end{aligned}$$

This means that $T(\boldsymbol{w}) \in W$. Thus the restriction $T|W : W \to W$ is a symmetric linear transformation. Since $\dim W = n - 1$, by induction hypothesis, W has an orthonormal basis $\{\boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ of eigenvectors of $T|W$. Clearly, $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ is an orthonormal basis of V, and $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ are eigenvectors of T. □

Theorem 11.5 (Real Spectral Theorem). *Any real symmetric matrix A can be diagonalized by an orthogonal matrix. More specifically, there exists an orthonormal basis $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ of $\mathbb{R}^n$ such that*

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i, \quad 1 \le i \le n,$$

$$Q^{-1}AQ = Q^TAQ = \mathrm{Diag}\big[\lambda_1, \lambda_2, \ldots, \lambda_n\big],$$

where $Q = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big]$; and **spectral decomposition**

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

Proof. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $T(\boldsymbol{x}) = A\boldsymbol{x}$. For vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$,

$$\begin{aligned} T(\boldsymbol{u}) \cdot \boldsymbol{v} &= (A\boldsymbol{u}) \cdot \boldsymbol{v} = (A\boldsymbol{u})^T \boldsymbol{v} = \boldsymbol{u}^T A^T \boldsymbol{v} \\ &= \boldsymbol{u}^T A \boldsymbol{v} = \boldsymbol{u} \cdot (A\boldsymbol{v}) = \boldsymbol{u} \cdot T(\boldsymbol{v}). \end{aligned}$$

Then T is symmetric. Thus $\mathbb{R}^n$ has an orthonormal basis $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ of eigenvectors of T. Let

$$T(\boldsymbol{u}_i) = \lambda_i \boldsymbol{u}_i, \quad 1 \le i \le n;$$

and $Q = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big]$. Then

$$Q^{-1}AQ = Q^TAQ = \mathrm{Diag}\big[\lambda_1, \lambda_2, \ldots, \lambda_n\big] = D.$$

Alternatively,

$$\begin{aligned} A &= QDQ^{-1} = QDQ^T \\ &= \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^T \\ \boldsymbol{u}_2^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix} \\ &= [\lambda_1 \boldsymbol{u}_1, \lambda \boldsymbol{u}_2, \ldots, \lambda_n \boldsymbol{u}_n] \begin{bmatrix} \boldsymbol{u}_1^T \\ \boldsymbol{u}_2^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix} \\ &= \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^T + \lambda_2 \boldsymbol{u}_2 \boldsymbol{u}_2^T + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^T. \end{aligned}$$

□

Note. It is clear that if a real square matrix A is orthogonally diagonalizable, then A is symmetric.

Example 11.2. Is the matrix $A = \begin{bmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ 3 & 3 & 1 \end{bmatrix}$ orthogonally diagonalizable?

The characteristic polynomial of A is

$$\Delta(t) = (t+2)^2(t-7).$$

There are eigenvalues $\lambda_1 = -2$ and $\lambda_2 = 7$.

For $\lambda_1 = -2$, there are two independent eigenvectors

$$\boldsymbol{v}_1 = [-1, 1, 0]^T, \quad \boldsymbol{v}_2 = [-1, 0, 1]^T.$$

Set $\boldsymbol{w}_1 = \boldsymbol{v}_1$,

$$\boldsymbol{w}_2 = \boldsymbol{v}_2 - \frac{\boldsymbol{v}_2 \cdot \boldsymbol{w}_1}{\boldsymbol{w}_1 \cdot \boldsymbol{w}_1} \boldsymbol{w}_1 = [-1/2, 1/2, 1]^T.$$

Then

$$\boldsymbol{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \boldsymbol{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

form an orthonormal basis of E_{λ_1}.

For $\lambda_2 = -7$, there is one independent eigenvector

$$\boldsymbol{v}_3 = [1, 1, 1]^T.$$

The orthonormal basis of E_{λ_2} is

$$\boldsymbol{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}.$$

Then the orthogonal matrix

$$Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

diagonalizes the symmetric matrix A.

An $n \times n$ real symmetric matrix A is called **positive definite** if, for any nonzero vector $\boldsymbol{u} \in \mathbb{R}^n$,

$$\langle \boldsymbol{u}, A\boldsymbol{u} \rangle = \boldsymbol{u}^T A \boldsymbol{u} > 0.$$

Theorem 11.6. *Let A be an $n \times n$ real symmetric matrix. Let $\langle \, , \rangle$ be defined by*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u}^T A \boldsymbol{v}, \quad \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n.$$

Then $\langle \, , \rangle$ is an inner product on $\mathbb{R}^n$ if and only if the matrix A is positive definite.

Theorem 11.7. *Let A be the matrix of an n-dimensional inner product space V relative to a basis $\mathcal{B}$. Then for $\boldsymbol{u}, \boldsymbol{v} \in V$,*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = [\boldsymbol{u}]_{\mathcal{B}}^T A [\boldsymbol{v}]_{\mathcal{B}}.$$

Moreover, A is positive definite.

12 Complex inner product spaces

Definition 12.1. Let V be a complex vector space. An **inner product** of V is a function $\langle \, , \rangle : V \times V \to \mathbb{C}$ satisfying the following properties:

(1) **Linearity**: $\langle a\boldsymbol{u} + b\boldsymbol{v}, w \rangle = a\langle \boldsymbol{u}, \boldsymbol{w} \rangle + b\langle \boldsymbol{v}, \boldsymbol{w} \rangle$.

(2) **Conjugate Symmetric Property**: $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \overline{\langle \boldsymbol{v}, \boldsymbol{u} \rangle}$.

(3) **Positive Definite Property**: For any $\boldsymbol{u} \in V$, $\langle \boldsymbol{u}, \boldsymbol{u} \rangle \geq 0$; and $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = 0$ if and only if $\boldsymbol{u} = 0$.

The complex vector space V with an inner product is called a **complex inner product space**.

Note that the Conjugate Symmetric Property implies that

$$\langle \boldsymbol{u}, a\boldsymbol{v} + b\boldsymbol{w} \rangle = \bar{a}\langle \boldsymbol{u}, \boldsymbol{v} \rangle + \bar{b}\langle \boldsymbol{u}, \boldsymbol{w} \rangle.$$

For any $\boldsymbol{u} \in V$, since $\langle \boldsymbol{u}, \boldsymbol{u} \rangle$ is a nonnegative real number, we define the **length** of $\boldsymbol{u}$ to be the real number

$$\|\boldsymbol{u}\| = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}.$$

Theorem 12.2 (Cauchy-Schwarz Inequality). *Let V be a complex inner product space. Then for any $\boldsymbol{u}, \boldsymbol{v} \in V$,*

$$\big|\langle \boldsymbol{u}, \boldsymbol{v}\rangle\big| \le \|\boldsymbol{u}\|\,\|\boldsymbol{v}\|.$$

Like real inner product space, one can similarly define **orthogonality**, the **angle between two vectors**, **orthogonal set**, **orthonormal basis**, **orthogonal projection**, and **Gram-Schmidt process**, etc.

Two vectors $\boldsymbol{u}, \boldsymbol{v}$ in a complex inner product space V are called **orthogonal** if

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0.$$

When both $\boldsymbol{u}, \boldsymbol{v} \in V$ are nonzero, the **angle** θ between $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined by

$$\cos\theta = \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}.$$

A set $\big\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\big\}$ of nonzero vectors of V is called an **orthogonal set** if the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are mutually orthogonal. A basis of V is called an **orthogonal basis** if their vectors are mutually orthogonal; a basis $\mathcal{B}$is called an **orthonormal basis** if $\mathcal{B}$ is an orthogonal basis and every vector of $\mathcal{B}$ has unit length.

Example 12.1 (Complex Euclidean space $\mathbb{C}^n$). For vectors $\boldsymbol{u} = [z_1, \ldots, z_n]^T, , \boldsymbol{v} = [w_1, \ldots, w_n]^T \in \mathbb{C}^n$, define

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \boldsymbol{u}^T \bar{\boldsymbol{v}} = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n.$$

Then $\langle\, , \rangle$ is an inner product of $\mathbb{C}^n$, called the **standard inner product** on $\mathbb{C}^n$. The vector space $\mathbb{C}^n$ with the standard inner product is called the **complex Euclidean n-space**.

Theorem 12.3. *Let $\mathcal{B} = \big\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big\}$ be an orthogonal basis of an inner product space V. Then for any $\boldsymbol{v} \in V$,*

$$\boldsymbol{v} = \frac{\langle \boldsymbol{v}, \boldsymbol{v}_1\rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1\rangle}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{v}, \boldsymbol{v}_2\rangle}{\langle \boldsymbol{v}_2, \boldsymbol{v}_2\rangle}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{v}, \boldsymbol{v}_n\rangle}{\langle \boldsymbol{v}_n, \boldsymbol{v}_n\rangle}\boldsymbol{v}_n.$$

Let W be a subspace of V with an orthogonal basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$. Then the **orthogonal projection** $\mathrm{Proj}_W : V \to V$ is given by

$$\mathrm{Proj}_W(\boldsymbol{v}) = \frac{\langle \boldsymbol{v}, \boldsymbol{u}_1\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u}_1\rangle}\boldsymbol{u}_1 + \frac{\langle \boldsymbol{v}, \boldsymbol{u}_2\rangle}{\langle \boldsymbol{u}_2, \boldsymbol{u}_2\rangle}\boldsymbol{u}_2 + \cdots + \frac{\langle \boldsymbol{v}, \boldsymbol{u}_k\rangle}{\langle \boldsymbol{u}_k, \boldsymbol{u}_k\rangle}\boldsymbol{u}_k.$$

In particular, if $V = \mathbb{C}^n$ and $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is an orthonormal basis of W, then the standard matrix of the linear mapping Proj_W is $U\bar{U}^T$, i.e.,

$$\mathrm{Proj}_W(\boldsymbol{x}) = U\bar{U}^T\boldsymbol{x},$$

where $U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k]$ is an $n \times k$ complex matrix.

Theorem 12.4. *Let $\mathcal{B} = \big\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big\}$ be basis of an inner product space V. Let A be the matrix of the inner product of V, i.e., $A = [a_{ij}]$, where $a_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle$. Then for $\boldsymbol{u}, \boldsymbol{v} \in V$,*

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = [\boldsymbol{u}]^T A[\bar{\boldsymbol{v}}],$$

where $[\boldsymbol{u}]$ and $[\boldsymbol{v}]$ are the $\mathcal{B}$-coordinate vectors of $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively.

A complex square matrix A is called **Hermitian** if

$$A^* = A,$$

where

$$A^* := \bar{A}^T.$$

A complex square matrix A is called **positive definite** if for any $\boldsymbol{x} \in \mathbb{C}^n$,

$$\boldsymbol{x}^T A\boldsymbol{x} \ge 0.$$

Theorem 12.5. *Let A be an $n \times n$ complex matrix. Then A is the matrix of an n-dimensional inner product complex vector space relative to a basis if and only if A is Hermitian and positive definite.*

A complex square matrix A is called **unitary** if

$$A^* = A^{-1},$$

or equivalently,

$$AA^* = A^*A = I.$$

Let $A = [a_{ij}]$. Then A is unitary means that

$$a_{1i}\bar{a}_{1j} + a_{2i}\bar{a}_{2j} + \cdots + a_{ni}\bar{a}_{nj} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}.$$

Theorem 12.6. *Let A be a complex square matrix. Then the following statements are equivalent:*

(a) *A is unitary.*

(b) *The rows of A form an orthonormal set.*

(c) *The columns of A form an orthonormal set.*

Theorem 12.7. *Let V be an n-dimensional complex inner product space with an orthonormal basis $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$. Let U be an $n \times n$ real matrix, and*

$$[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n]A.$$

Then $\mathcal{B}' = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is an orthonormal basis if and only if A is a unitary matrix.

Proof. For simplicity, we assume $n = 3$. Since

$$[\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3]A,$$

i.e.,

$$\boldsymbol{v}_1 = a_{11}\boldsymbol{u}_1 + a_{21}\boldsymbol{u}_2 + a_{31}\boldsymbol{u}_3,$$

$$\boldsymbol{v}_2 = a_{12}u_1 + a_{22}u_2 + a_{32}u_3,$$

$$\boldsymbol{v}_3 = a_{13}\boldsymbol{u}_1 + a_{23}\boldsymbol{u}_2 + a_{33}\boldsymbol{u}_3.$$

We then have

$$\begin{aligned} \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle &= \langle a_{1i}\boldsymbol{u}_1 + a_{2i}\boldsymbol{u}_2 + a_{3i}\boldsymbol{u}_3, \\ &\qquad a_{1j}\boldsymbol{u}_1 + a_{2j}\boldsymbol{u}_2 + a_{3j}\boldsymbol{u}_3 \rangle \\ &= a_{1i}\bar{a}_{1j} + a_{2i}\bar{a}_{2j} + a_{3i}\bar{a}_{3j}. \end{aligned}$$

Note that $\mathcal{B}'$ is an orthonormal basis is equivalent to

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \delta_{ij},$$

and A is an unitary matrix if and only if

$$a_{1i}\bar{a}_{1j} + a_{2i}\bar{a}_{2j} + a_{3i}\bar{a}_{3j} = \delta_{ij}.$$

The proof is finished. □

Let V be an n-dimensional complex inner product space. A linear mapping $T : V \to V$ is a called an **isometry** of V if T preserves length of vector, i.e., for any $\boldsymbol{v} \in V$,

$$\|T(\boldsymbol{v})\| = \|\boldsymbol{v}\|.$$

Theorem 12.8. *Let V be an n-dimensional inner product space, and let $T : V \to V$ be a linear transformation. The following statements are equivalent:*

(a) *T preserves length, i.e., for any $\boldsymbol{v} \in V$,*

$$\|T(\boldsymbol{v})\| = \|\boldsymbol{v}\|.$$

(b) *T preserves inner product, i.e., for $\boldsymbol{u}, \boldsymbol{v} \in V$,*

$$\langle T(\boldsymbol{u}), T(\boldsymbol{v})\rangle = \langle \boldsymbol{v}, \boldsymbol{v}\rangle.$$

(c) *T preserves orthogonality, i.e., for $\boldsymbol{u}, \boldsymbol{v} \in V$,*

$$\langle T(\boldsymbol{u}), T(\boldsymbol{v})\rangle = 0 \Longleftrightarrow \langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0.$$

Proof. note that if T preserves inner product, of course it preserves length.

Now for vectors $\boldsymbol{u}, \boldsymbol{v} \in V$, we have

$$\begin{aligned} \|T(\boldsymbol{u}+\boldsymbol{v})\|^2 &= \langle T(\boldsymbol{u}+\boldsymbol{v}), T(\boldsymbol{u}+\boldsymbol{v})\rangle \\ &= \langle T(\boldsymbol{u}), T(\boldsymbol{u})\rangle + \langle T(\boldsymbol{v}), T(\boldsymbol{v})\rangle + \langle T(\boldsymbol{u}), T(\boldsymbol{v})\rangle + \langle T(\boldsymbol{v}), T(\boldsymbol{u})\rangle \\ &= \|T(\boldsymbol{u})\|^2 + \|T(\boldsymbol{v})\|^2 + 2\Re\langle T(\boldsymbol{u}), T(\boldsymbol{v})\rangle; \end{aligned}$$

$$\begin{aligned} &\left\|T\big(\boldsymbol{u}+\sqrt{-1}\boldsymbol{v}\big)\right\|^2 = \\ &\qquad \|T(\boldsymbol{u})\|^2 + \|T(\boldsymbol{v})\|^2 + 2\Im\langle T(\boldsymbol{u}), T(\boldsymbol{v})\rangle. \end{aligned}$$

On the other hand,

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 + 2\Re\langle \boldsymbol{u}, \boldsymbol{v}\rangle;$$
$$\left\|\boldsymbol{u}+\sqrt{-1}\boldsymbol{v}\right\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 + 2\Im\langle \boldsymbol{u}, \boldsymbol{v}\rangle.$$

If T preserves length, then

$$\|T(\boldsymbol{u})\|^2 = \|\boldsymbol{u}\|^2, \quad \|T(\boldsymbol{v})\|^2 = \|\boldsymbol{v}\|^2,$$
$$\|T(\boldsymbol{u}+\boldsymbol{v})\|^2 = \|\boldsymbol{u}+\boldsymbol{v}\|^2,$$
$$\left\|T\big(\boldsymbol{u}+\sqrt{-1}\boldsymbol{v}\big)\right\|^2 = \left\|\boldsymbol{u}+\sqrt{-1}\boldsymbol{v}\right\|^2.$$

It follows that

$$\Re\langle T(\boldsymbol{u}), T(\boldsymbol{v})\rangle = \Re\langle \boldsymbol{u}, \boldsymbol{v}\rangle, \quad \Im\langle T(\boldsymbol{u}), T(\boldsymbol{v})\rangle = \Im\langle \boldsymbol{u}, \boldsymbol{v}\rangle.$$

Equivalently,

$$\langle T(\boldsymbol{u}), T(\boldsymbol{v})\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle.$$

□