Inner Product Spaces

Elif Tan

Ankara University

Inner Product Spaces

Definition (Inner Product Space)

Let $(V, \oplus, \odot)$ be a real vector space. If the function $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ satisfies the following properties, then V is called an **inner product space** and the function $\langle\ ,\ \rangle$ is called an **inner product function**:
For all $u, v, w \in V$ and all $c \in \mathbb{R}$,
(i) $\langle u, u\rangle \geq 0$ and $\langle u, u\rangle = 0 \Leftrightarrow u = 0$
(ii) $\langle u, v\rangle = \langle v, u\rangle$
(iii) $\langle u \oplus v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
(iv) $\langle c \odot u, v\rangle = c\langle u, v\rangle$.

For the simplicity, we drop the notations $\oplus, \odot$.

Inner Product Spaces

Example

For $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$, the standard inner product (dot product) on $\mathbb{R}^n$ is defined by

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

Example: Let $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$.

1.

$$\langle u, v \rangle = u_1 v_1 - u_2 v_1 - u_1 v_2 + 3 u_2 v_2.$$

is an inner product on $\mathbb{R}^2$.
Since

$$\langle u, u \rangle = u_1^2 - 2u_1 u_2 + 3u_2^2 = (u_1 - u_2)^2 + 2u_2^2 \geq 0$$

and all other three properties hold.

2.

$$\langle u, v \rangle = u_1 v_2 + u_2 v_1$$

is not an inner product space on $\mathbb{R}^2$.
Since for $u_1 = 1, u_2 = -1$,

$$\langle u, u \rangle = \left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\rangle = 2u_1 u_2 = -2 < 0.$$

Inner Product Spaces

3. Check if

$$\langle u, v \rangle = u_1^2 + u_2^2 + v_1^2 + v_2^2$$

is an inner product space on $\mathbb{R}^2$.

4. Check if

$$\langle u, v \rangle = 4u_1 v_1 + 9u_2 v_2$$

is an inner product space on $\mathbb{R}^2$.

Theorem

Let V be an inner product space, and $S = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for the vector space V. Then the matrix $A = [a_{ij}]_{n \times n}$, where $a_{ij} := \langle a_i, a_j \rangle$ is a symmetric matrix, and for every $u, v \in V$, it determines $\langle u, v \rangle$.

Note that the matrix $A = [a_{ij}]_{n \times n}$, $a_{ij} = \langle u_i, u_j \rangle$ is called **the matrix of the inner product with respect to the ordered basis** S.

$$A = \begin{bmatrix} \langle u_1, u_1 \rangle & \langle u_1, u_2 \rangle & \cdots & \langle u_1, u_n \rangle \\ \langle u_2, u_1 \rangle & \langle u_2, u_2 \rangle & \cdots & \langle u_2, u_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle u_n, u_1 \rangle & \langle u_n, u_2 \rangle & \cdots & \langle u_n, u_n \rangle \end{bmatrix}.$$

Definition (Positive definite matrix)

The $n \times n$ symmetric matrix A is called **positive definite matrix** if it has the property that

$$x^T Ax > 0 \text{ for all } 0 \neq x \in \mathbb{R}^n$$

Theorem

Let $A = [a_{ij}]_{n\times n}$ be a positive definite matrix, and $S = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for the vector space V. Then the function $\langle , \rangle : V \times V \to \mathbb{R}$ that is defined by

$$\langle u, v\rangle := [u]_S^T A [v]_S \quad \text{for all } u, v \in V$$

is an inner product function on V.

Note that it is not easy to determine when a symmetric matrix is positive definite!

Inner Product Spaces

Example

Consider the standard inner product function $\left\langle \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix} \right\rangle = ac + bd$ on $\mathbb{R}^2$. The matrix of the inner product with respect to the ordered basis $S = \left\{ u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ is

$$A = \begin{bmatrix} \langle u_1, u_1 \rangle & \langle u_1, u_2 \rangle \\ \langle u_2, u_1 \rangle & \langle u_2, u_2 \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Conversely, consider the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.($A$ is positive definite matrix,verify it). The inner product with respect to the ordered standard basis in $\mathbb{R}^2$ is

$$\left\langle \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix} \right\rangle = \begin{bmatrix} a & b \end{bmatrix} A \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd.$$

Example: Find (if possible) the inner product on $\mathbb{R}^3$ that correspond to the matrix

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

It is clear to see that A is symmetric. Lets check if A is positive definite.

For $\mathbf{0} \neq x = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$,

$$\begin{aligned} x^T A x &= \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= 2a^2 + 2b^2 + 2c^2 + 2ab + 2ac + 2bc \\ &= a^2 + b^2 + c^2 + (a+b+c)^2 > 0. \end{aligned}$$

Thus A corresponds an inner product. For all $u, v \in \mathbb{R}^3$

$$
\begin{aligned}
\langle u, v\rangle &= [u]_S^T\, A\, [v]_S \\
&= \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \\
&= (2u_1 + u_2 + u_3)\, v_1 + (u_1 + 2u_2 + u_3)\, v_2 + (u_1 + u_2 + 2u_3)\, v_3 .
\end{aligned}
$$

Definition (Lenght)

Let V be an inner product space. The lenght (norm) of $v \in V$ is defined by

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

Definition (Distance)

Let V be an inner product space. The distance between u and v in V is defined by

$$d(u, v) := ||u - v|| = \sqrt{\langle u - v, u - v \rangle}.$$

- $d(u, v) = 0 \Leftrightarrow u = v$
- $d(u, v) = d(v, u)$

Examples:

1. The Euclidean norm in $\mathbb{R}^2$ is

$$||v|| := \sqrt{\langle v, v \rangle} = \sqrt{v_1^2 + v_2^2}$$

and the distance between u and v in $\mathbb{R}^2$ is

$$d(u, v) := ||u - v|| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}.$$

2. The Euclidean norm in $\mathbb{R}^n$ is

$$||v|| := \sqrt{\langle v, v \rangle} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

and the distance between u and v in $\mathbb{R}^n$ is

$$d(u, v) := ||u - v|| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}.$$

- **Cauchy-Schwarz inequality:** Let V be an inner product space. Then

$$|\langle u, v\rangle| \leq ||u|| \, . \, ||v||$$

for $u, v \in V$.

By using Cauchy-Schwarz inequality, we define the cosine of an angle between nonzero vectors u and v in V as

$$\cos\theta := \frac{\langle u, v\rangle}{\|u\| \, \|v\|}, \ 0 \leq \theta \leq \pi.$$

Let V be an inner product space. Then $\forall u, v \in V$ and $\forall c \in \mathbb{R}$, we have the followings.

- **Homogenity**: $\|cv\| = |c| \, \|v\|$.

$$\begin{aligned} \|cv\|^2 &= \langle cv, cv \rangle \\ &= c \langle v, cv \rangle \\ &= c^2 \langle v, v \rangle \\ &= c^2 \|v\|^2 . \end{aligned}$$

- **Triangle inequality**: $\|u + v\| \leq \|u\| + \|v\|$.

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

- **Parallelogram law:**

$$||u+v||^2 + ||u-v||^2 = 2\,||u||^2 + 2\,||v||^2 \,.$$

Since

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

then we get the desired result.

Inner Product Spaces

Definition

Let V be an inner product space. The vectors u and v in V are **orthogonal** if $\langle u, v\rangle = 0$. That is,

$$u \perp v \Leftrightarrow \langle u, v\rangle = 0.$$

A set of S of vectors in V is called ortogonal if any two distinct vectors in S are orthogonal.
Additionally, if each vector in S is a unit vector ($||u|| = 1$), then S is called **orthonormal**.

- $\mathbf{0} \perp v$, for all $v \in V$.
- If $S = \{u_1, u_2, \ldots, u_n\}$ is an orthogonal set of nonzero vectors in an inner product space V, then S is linearly independent.

Inner Product Spaces

- **Pythagoraen Theorem:** If $u \perp v$, then

$$||u+v||^2 = ||u||^2 + ||v||^2 .$$

Since $u \perp v$, then $\langle u, v \rangle = 0$.

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

Gram-Schmidt Process

Definition (Orthonormal Basis)

Let V be an inner product space and $S = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for the vector space V. S is called an orthonormal basis if

$$\langle u_i, u_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

for all $i, j = 1, ..., n$.

To obtain this orthonormal basis we use a method which is called the **Gram-Schmidt process**.

Gram-Schmidt Process

- Let $T = \{w_1, w_2, \ldots, w_n\}$ be an **orthonormal basis** for the inner product space V and $v \in V$. Then

$$v = c_1 \odot w_1 \oplus c_2 \odot w_2 \oplus \cdots \oplus c_n \odot w_n$$

where

$$c_i = \langle v, w_i \rangle .$$

Because

$$\begin{aligned} \langle v, w_i \rangle &= \langle c_1 \odot w_1 \oplus c_2 \odot w_2 \oplus \cdots \oplus c_n \odot w_n, w_i \rangle \\ &= c_1 \langle w_1, w_i \rangle + \cdots + c_i \langle w_i, w_i \rangle + \cdots + c_n \langle w_n, w_i \rangle \\ &= c_i \langle w_i, w_i \rangle = c_i . \end{aligned}$$

We determine the coordinates of the vector by using inner product instead of solving linear system!

- Let V be an inner product space and $\{0\} \neq W < V$ and $dimW = m$. Then there exists an orthonormal basis $T = \{w_1, w_2, \ldots, w_n\}$ for W. To find this basis we will give a procedure called as **Gram-Schmidt.**

Gram-Schmidt Process

The matrix of the inner product with respect to the ordered **orthogonal** basis K is

$$\begin{bmatrix} \langle v_1, v_1 \rangle & 0 & \cdots & 0 \\ 0 & \langle v_2, v_2 \rangle & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \langle v_n, v_n \rangle \end{bmatrix}.$$

Gram-Schmidt Process

The matrix of the inner product with respect to the ordered **orthonormal** basis T is

$$I_m = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Gram-Schmidt Process

For $u, v \in V$, we have $[u]_T = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$ and $[v]_T = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$.

If we use the orthonormal basis T, then

$$\begin{aligned} \langle u, v \rangle &= \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} I_m \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \\ &= a_1 b_1 + a_2 b_2 + \cdots + a_m b_m \end{aligned}$$

which is the standard inner product.

Gram-Schmidt Process

Gram-Schmidt Process: Let $S = \{u_1, u_2, \ldots, u_m\}$ be any basis for W.

- $v_1 := u_1$
- $v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1$ (Note that $v_1 \perp v_2$)
- $v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2$ (Note that $v_3 \perp v_1, v_3 \perp v_2$)
- Similarly, we obtain m-orthogonal vectors as

$$v_m = u_m - \frac{\langle u_m, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_m, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \cdots - \frac{\langle u_m, v_{m-1} \rangle}{\langle v_{m-1}, v_{m-1} \rangle} v_{m-1}.$$

- Thus, we obtain an orthogonal set $K = \{v_1, v_2, \ldots, v_m\}$. Since K is a linear independent set in m-dimensional vector space, K is a basis for W.

- **Normalization:** If we let

$$w_i = \frac{v_i}{\|v_i\|}$$

for $i = 1, 2, \ldots, m$, then $T = \{w_1, w_2, \ldots, w_m\}$ is an orthonormal basis for W.

Gram-Schmidt Process

Example: Let $S = \left\{ u_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}$ be a basis for $\mathbb{R}^2$. Transform S to the orthonormal basis.

- $v_1 = u_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$
-

$$\begin{aligned} v_2 &= u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 \\ &= \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \frac{\left\langle \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\rangle} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \frac{2.3 + 2.1}{3.3 + 1.1} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 6/5 \end{bmatrix} \end{aligned}$$

Gram-Schmidt Process

- Thus $K = \left\{ v_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} \frac{-2}{5} \\ \frac{6}{5} \end{bmatrix} \right\}$ is orthogonal basis.
- If we normalize v_1 and v_2, we get

$$w_1 = \frac{v_1}{\|v_1\|} = \frac{\begin{bmatrix} 3 \\ 1 \end{bmatrix}}{\sqrt{\left\langle \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\rangle}} = \begin{bmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}$$

$$w_2 = \frac{v_2}{\|v_2\|} = \frac{\begin{bmatrix} \frac{-2}{5} \\ \frac{6}{5} \end{bmatrix}}{\sqrt{\left\langle \begin{bmatrix} \frac{-2}{5} \\ \frac{6}{5} \end{bmatrix}, \begin{bmatrix} \frac{-2}{5} \\ \frac{6}{5} \end{bmatrix} \right\rangle}} = \begin{bmatrix} \frac{-1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{bmatrix}$$

- Thus $T = \{w_1, w_2\}$ is an orthonormal basis.

Projection

- Let V be an inner product space and $W < V$. A vector $u \in V$ is said to be **orthogonal to W** if it is orthogonal to every vector in W. The set of all vectors in V that are orthogonal to W is called the **orthogonal complement** of W in V and denoted $W^{\perp}$. That is

$$W^{\perp} = \{u \in V \mid \langle u, w \rangle = 0, \text{ for all } w \in W\} .$$

$W^{\perp} < V$, $W \cap W^{\perp} = \{\mathbf{0}\}$.

- Let V be an inner product space and $W < V$ with orthonormal basis $\{w_1, w_2, \ldots, w_m\}$ .For $v \in V$, there exist unique vector $w \in W$ and $u \in W^{\perp}$ such that $v = w + u$.

Projection

If $\{w_1, w_2, \ldots, w_m\}$ is orthogonal basis for the subspace W, then

- The **projection** of v on w is defined by

$$Proj_w v = \frac{\langle v, w \rangle}{\langle w, w \rangle} w.$$

- The **projection** of v on W is defined by

$$Proj_W v = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 - \cdots - \frac{\langle v, w_m \rangle}{\langle w_m, w_m \rangle} w_m$$

- $Proj_W v \in W$ is the closest vector to v, so $\|v - Proj_W v\|$ represents the distance from v to W.