Northeastern University, Department of Mathematics

MATH 5110: Applied Linear Algebra and Matrix Analysis. (Fall 2021)

• Instructor: **He Wang** Email: **he.wang@northeastern.edu**

§6 Inner product spaces

Contents

1. Inner Product Spaces

Recall that for vectors $\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, the **dot product** of $\vec{u}$ and $\vec{v}$ is

Theorem 1 (Properties of the dot Product)**.** *For vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ and a scalar $c \in \mathbb{R}$, the following hold:*

(1.)

(2.)

(3.)

(4.)

(5.)

Definition 2 (Inner Product)**.** Let V be a real vector space. An **inner product** on V is a binary function

$$\langle -, - \rangle : V \times V \to \mathbb{R}$$

such that for vectors $\vec{u}, \vec{v}, \vec{w} \in V$ and a scalar $c \in \mathbb{R}$, the following hold:

(1.) $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.
(2.) $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$.
(3.) $\langle c\vec{u}, \vec{v} \rangle = c\langle \vec{u}, \vec{v} \rangle$.
(4.) $\langle \vec{u}, \vec{u} \rangle \geq 0$
(5.) $\langle \vec{u}, \vec{u} \rangle = 0$ if and only if $\vec{u} = \vec{0}$.

We call V an **inner product space**.

Example 3. (Weighted dot products) Let $c_1, ..., c_n$ be positive numbers. The weighted inner product on $\mathbb{F}^n$ is

Example 4. Let $P_n(\mathbb{F})$ be the vector space of polynomials of degree at most n with coefficient in $\mathbb{F}$. An inner product on $P_n(\mathbb{R})$ can be defined as

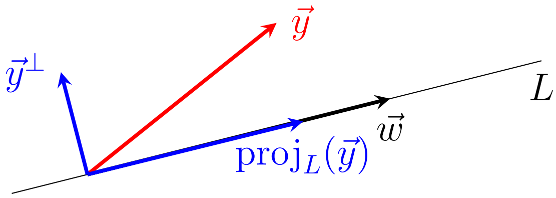
Definition 5. Two vectors $\vec{u}$ and $\vec{v}$ are called **orthogonal** if

2. Norms

Definition 6 (Norm of a Vector)**.** Let V be a inner product space over $\mathbb{F}$. The **length** or **norm** of a vector $\vec{v} \in V$, denoted by $||\vec{v}||$, is defined as

Proposition 7. *For any vector $\vec{v} \in V$ and any scalar $c \in \mathbb{F}$ one obtains*
$$||c \cdot \vec{v}|| = |c| \cdot ||\vec{v}||.$$

Theorem 8 (Pythagorean Theorem)**.** *If two vectors $\vec{u}, \vec{v} \in V$ are orthogonal, then they satisfy the* ***Pythagorean Relation***
$$||\vec{u} + \vec{v}||^2 = ||\vec{u}||^2 + ||\vec{v}||^2.$$



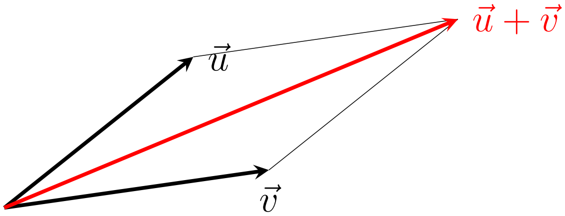
Definition 9. Let $L = \text{Span}\{\vec{w}\}$ be the subspace in V spanned by $\vec{w} \in V$. For a given vector $\vec{y} \in V$, the **orthogonal projection of $\vec{y}$ onto L**

Proposition 10. *Let $\vec{w}$ be a nonzero vector in V. Any vector $\vec{y} \in V$ can be uniquely written as the sum of a scalar product of $\vec{w} \in V$ and a vector orthogonal to $\vec{w}$.*

Theorem 11 (**Cauchy-Schwarz inequality**)**.**

Proposition 12 (Triangle Inequality)**.** *Two vectors* $\vec{u}, \vec{v} \in V$ *satisfy*

$$||\vec{u}+\vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||.$$



Definition 13. (Angles Between Vectors) The **angle between two nonzero vectors** $\vec{u}, \vec{v} \in V$ is the the angle $0 \leq \theta \leq \pi$ satisfying

A vector space V with norm is called a **normed vector space**.

Definition 14. A **norm** on V is a map from V to $\mathbb{F}$ such that

(1) $||\vec{x}|| \geq 0$ for all $\vec{x} \in V$. $||\vec{x}|| = 0$ if and only if $\vec{x} = \vec{0}$.
(2) $||c\vec{x}|| = c||\vec{x}||$ for all $\vec{x} \in V$ and $c \in \mathbb{F}$.
(3) The triangle inequality $||\vec{u}+\vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||$ holds for all vectors in V.

Definition 15 (Distance Between Vectors)**.** The **distance** $\operatorname{dist}(\vec{u}, \vec{v})$ between vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ is defined as

$$\operatorname{dist}(\vec{u}, \vec{v}) = ||\vec{u} - \vec{v}|| = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2}.$$

Example 16. (l^p spaces) Let $1 \leq p < \infty$, it is natural to define l^p norms on $\mathbb{F}^n$

Example 17. (l^∞ spaces) It is natural to define l^∞ norms on $\mathbb{F}^n$

Example 18. (Norms on $\mathbb{F}^{m\times n}$ induced by norms on $\mathbb{F}^n$) Normed matrix vector spaces $\mathbb{F}^{m\times n}$. Using norms on $\mathbb{F}^n$, one can define norms on matrix vector spaces

Example 19. Infinity norm on $\mathbb{F}^{m\times n}$.

Definition 20 (Orthogonal Set)**.** A set $\{\vec{u}_1, \dots, \vec{u}_p\}$ of vectors in a inner vector space V is called **orthogonal** if

Proposition 21.
- *Orthogonal vectors are linear independent.*
- *Orthogonal vectors $\{\vec{u}_1, \dots, \vec{u}_n\}$ in $\mathbb{R}^n$ form a basis of $\mathbb{R}^n$.*

Definition 22.
- An **orthogonal basis** for a subspace W of an inner product space V is any basis for W which is also an orthogonal set.
- If each vector is a **unit** vector in an orthogonal basis, then it is an **orthonormal basis**.

Theorem 23 (Coordinates with respect to an orthogonal basis)**.** *Let $\mathscr{B} = \{\vec{u}_1, \dots, \vec{u}_p\}$ be an **orthogonal** basis for a subspace W of an inner product space V, and let $\vec{y}$ be any vector in W. Then*

$$\vec{y} = \left(\frac{\langle \vec{y}, \vec{u}_1 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle} \right) \vec{u}_1 + \cdots + \left(\frac{\langle \vec{y}, \vec{u}_p \rangle}{\langle \vec{u}_p, \vec{u}_p \rangle} \right) \vec{u}_p$$

If $W = \mathbb{R}^n$, then the $\mathscr{B}$-coordinates of $\vec{y}$ are given by:

$$[\vec{y}]_{\mathscr{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \quad \textit{with} \quad c_i = \frac{\langle \vec{y}, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle} = \frac{\langle \vec{y}, \vec{u}_i \rangle}{||\vec{u}_i||^2}$$

*In particular, let $\mathscr{B} = \{\vec{u}_1, \dots, \vec{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let $\vec{y}$ be any vector in W. Then*

$$\vec{y} = \langle \vec{y}, \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{y}, \vec{u}_p \rangle \vec{u}_p$$

Theorem 24 (Orthogonal Decomposition)**.** *Let W be any subspace of V and let $\vec{y} \in V$ be any vector. Then there exists a unique decomposition*

$$\vec{y} = \mathrm{proj}_W(\vec{y}) + \vec{y}^{\perp}$$

with $\mathrm{proj}_W(\vec{y}) \in W$ *and* $\vec{y}^{\perp}$ *is perpendicular to* W*.*

Theorem 25 (Orthogonal Decomposition)**.** *If* $\{\vec{u}_1, \ldots, \vec{u}_p\}$ *is an* orthogonal basis *for* W*, then*

$$\mathrm{proj}_W(\vec{y}) = \left(\frac{\langle \vec{y}, \vec{u}_1 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle}\right) \vec{u}_1 + \cdots + \left(\frac{\langle \vec{y}, \vec{u}_p \rangle}{\langle \vec{u}_p, \vec{u}_p \rangle}\right) \vec{u}_p$$

and $\vec{y}^{\perp} = \vec{y} - \mathrm{proj}_W(\vec{y})$*.*

Definition 26 (Orthogonal Complements)**.** Given a nonempty **subset** (finite or infinite) W of V, its **orthogonal complement** $W^{\perp}$ is the set of all vectors $\vec{v} \in V$ orthogonal to W.

Theorem 27. *Let S be a subset of V. Let $W = \mathrm{Span}(S)$, then*

Theorem 28. *Let W be a subspace of V, then*

$$V = W \oplus W^{\perp}$$

Theorem 29. *Let A be an $m \times n$ matrix, then*

$$(\mathrm{Row}\, A)^{\perp} = \ker(A) \quad \textit{and} \quad (\mathrm{im}\, A)^{\perp} = \ker A^T.$$

More over,

$$\mathbb{F}^m = \ker A^T \oplus \mathrm{im}\, A$$

4. Gram-Schmidt process and QR-factorization

The **Gram-Schmidt process** is an algorithm that produces an orthogonal (or orthonormal) basis for any subspace W of V by starting with any basis for W.

Theorem 30 (Gram-Schmidt (Orthogonalize))**.** *Let W be a subspace of V and let $\vec{b}_1, \cdots, \vec{b}_p$ be a basis for W. Define vectors $\vec{v}_1, \ldots, \vec{v}_p$ as*

Theorem 31 (Gram-Schmidt (Normalize))**.** *If $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W, then*

$$\text{Basis} \xrightarrow{\text{orthogonalize}} \text{Orthogonal basis} \xrightarrow{\text{normalize}} \text{Orthonormal basis.}$$

QR-Factorization.

QR**-Factorization** is the matrix version of Gram-Schmidt process for a subspace W of $\mathbb{F}^n$:

$$\text{Basis } \mathscr{B} = \{\vec{b}_1, \dots, \vec{b}_p\} \xrightarrow{\text{orthogonalize}} \text{Orthogonal basis } \mathscr{V} = \{\vec{v}_1, \dots, \vec{v}_p\}$$
$$\xrightarrow{\text{normalize}} \text{Orthonormal basis } \mathscr{U} = \{\vec{u}_1, \dots, \vec{u}_p\}.$$

Given a vector in W, let's compare their coordinates:

$$\begin{tikzcd}
{[\vec{x}]_{\mathscr{B}}} \arrow[dr, "M"] \arrow[dd, "R"'] & \\
& \vec{x} \\
{[\vec{x}]_{\mathscr{U}}} \arrow[ur, "Q"'] &
\end{tikzcd}$$

Each matrix defines an isomorphism. So, $M = QR$.

Here $M = [\vec{b}_1 \ \dots \ \vec{b}_p]$ and $Q = [\vec{u}_1, \dots, \vec{u}_p]$.

Theorem 32. *Given a $n \times p$ matrix $M = [\vec{b}_1 \ \dots \ \vec{b}_p]$ with independent columns. There is a unique decomposition*

$$M = QR$$

where, $Q = [\vec{u}_1, \dots, \vec{u}_p]$ has orthonormal columns and R is an $p \times p$ upper triangular matrix with

$$r_{ii} = ||\vec{v}_i|| \text{ for } i = 1, \dots, p \text{ and } r_{ij} = \langle \vec{u}_i, \vec{b}_j \rangle \text{ for } i < j.$$

Proof. Proof(for $p = 3$): From Gram-Schmidt process, write $\vec{b}_i$ as linear combinations of $\vec{u}_i$.

$$\vec{b}_1 = \vec{v}_1 = ||\vec{v}_1||\vec{u}_1$$

$$\vec{b}_2 = \vec{v}_2 + \frac{\langle \vec{b}_2, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 = ||\vec{v}_2||\vec{u}_2 + \langle \vec{b}_2, \vec{u}_1 \rangle \vec{u}_1$$

$$\vec{b}_3 = \vec{v}_3 + \frac{\langle \vec{b}_3, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 + \frac{\langle \vec{b}_3, \vec{v}_2 \rangle}{||\vec{v}_2||^2}\vec{v}_2 = ||\vec{v}_3||\vec{u}_3 + \langle \vec{b}_3, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{b}_3, \vec{u}_2 \rangle \vec{u}_2$$

So,

$$[\vec{b}_1 \ \vec{b}_2 \ \vec{b}_3] = [\vec{u}_1 \ \vec{u}_2 \ \vec{u}_3] \begin{bmatrix} ||\vec{v}_1|| & \langle \vec{u}_1, \vec{b}_2 \rangle & \langle \vec{u}_1, \vec{b}_3 \rangle \\ 0 & ||\vec{v}_2|| & \langle \vec{u}_2, \vec{b}_3 \rangle \\ 0 & 0 & ||\vec{v}_3|| \end{bmatrix}$$

□

5. **Orthogonal Transformations and Orthogonal Matrices**

Let V be a inner product space.

Definition 33. A linear transformation $T : V \to V$ is called **orthogonal** if

$$||T(\vec{x})|| = ||\vec{x}|| \text{ for all } \vec{x} \in V$$

that is, T preserves the length of vectors.

Example 34. Whether or not the following transformations are orthogonal.

(1.) Rotations $S : \mathbb{R}^2 \to \mathbb{R}^2$ are orthogonal transformations.

The matrix of rotation $S = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal.

(2.) Reflections $R : \mathbb{R}^2 \to \mathbb{R}^2$ are orthogonal transformations.

The matrix of reflection matrix $R = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$ with $a^2 + b^2 = 1$ is orthogonal.

(3.) Orthogonal projections $P : \mathbb{R}^2 \to \mathbb{R}^2$ are NOT orthogonal transformations.

The matrix of an orthogonal transformation $T : \mathbb{F}^n \to \mathbb{F}^n$ is called an **orthogonal matrix**.

Theorem 35. *Let U be an $n \times n$ orthogonal matrix and let $\vec{x}$ and $\vec{y}$ be any vectors in $\mathbb{F}^n$. Then*

(1) $||U\vec{x}|| = ||\vec{x}||$.
(2) $\langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$.
(3) $\langle U\vec{x}, U\vec{y}\rangle = 0$ *if and only if* $\langle \vec{x}, \vec{y}\rangle = 0$.

Proposition 36. *U is an orthogonal matrix if and only if $\langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$ for any $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$.*

Using the geometric meaning of the orthogonal transformation, we have

Theorem 37. *1. If A is orthogonal, then A is invertible and A^{-1} is orthogonal.*
2. If A and B are orthogonal, then AB is orthogonal.

Theorem 38. *The $n \times n$ matrix U is orthogonal if and only if $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is an orthonormal set.*

Application to real matrix A.

Recall the transpose of a matrix: Given an $m \times n$ matrix A, we define the ***transpose matrix*** A^T as the $n \times m$ matrix whose (i, j)-th entry is the (j, i)-th entry of A. The dot product can be written as matrix product

$$\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w}$$

Theorem 39. *The $n \times n$ matrix A is orthogonal if and only if $A^T A = I_n$; if and only if $A^{-1} = A^T$.*

Theorem 40. *Let W be any subspace of $\mathbb{R}^n$ with an orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_p\}$. Let $U = [\vec{u}_1\ \vec{u}_2 \cdots \vec{u}_p]$. For any $\vec{y} \in \mathbb{R}^n$,*

$$\text{proj}_W(\vec{y}) = UU^T\vec{y}.$$

That is, the ***matrix of the projection*** *onto W is*

$$P = UU^T$$

Application to complex matrix A.

The **conjugate transpose** of a complex matrix A is $A^* := \overline{A}^T$

A complex matrix A is called **Hermiltian** if $A^* = A$.

6. **The adjoint of a linear operator**

An generalization of transpose of a real matrix and conjugate transpose of a complex is the joint operator.

Definition 41. Let $T : V \to W$ be a linear map between two inner product spaces. A **joint operator** of T is a linear map $T^* : W \to V$ such that

$$\langle T(\vec{v}), \vec{w}\rangle_W = \langle \vec{v}, T^*(\vec{w})\rangle_V.$$

Theorem 42. *Let $T : V \to W$ be a linear map between two inner product spaces. Then T has a unique joint operator T^*.*

Theorem 43. *Let $T : V \to W$ be a invertible linear map between two inner product spaces. Then its joint operator T^* is also invertible and $(T^*)^{-1} = (T^{-1})^*$.*