Chapter 0 Inner Product Spaces

0.1 Motivation (and some Reviews)

A. Inner Product in $\mathbf{R}^3$:

A **vector** is a quantity that has both of **magnitude** and **direction**. Let $\overrightarrow{a} =< a_1, a_2, a_3 >$. Then, the magnitude of $\overrightarrow{a}$ is

$$|\overrightarrow{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

The only vector with length zero is **zero vector**, denoted by $\overrightarrow{0} =< 0, 0 >$ in $\mathbb{R}^2$ or $\overrightarrow{0} =< 0, 0, 0 >$ in $\mathbb{R}^3$.

A vector whose length is 1 is called *unit vector.* If $\overrightarrow{a} \neq 0$, then the unit vector that has the same direction as $\overrightarrow{a}$ is

$$\overrightarrow{u} = \frac{\overrightarrow{a}}{|\overrightarrow{a}|}.$$

Definition 1. If $\overrightarrow{a} =< a_1, a_2, a_3 >$ and $\overrightarrow{b} =< b_1, b_2, b_3 >$, the **dot product** of $\overrightarrow{a}$ and $\overrightarrow{b}$ is the number $\overrightarrow{a} \cdot \overrightarrow{b}$ given by

$$\overrightarrow{a} \cdot \overrightarrow{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Note 1: A result of the dot product of two vectors is a scalar (not a vector). So, the dot product is sometimes called **inner product** or **scalar product**.

Theorem 1. *Let $\overrightarrow{a} =< a_1, a_2, a_3 >$ and $\overrightarrow{b} =< b_1, b_2, b_3 >$. If θ is the angle between the vectors $\overrightarrow{a}$ and $\overrightarrow{b}$, then*

$$\overrightarrow{a} \cdot \overrightarrow{b} = |\overrightarrow{a}||\overrightarrow{b}| \cos\theta.$$

Corollary 1. *Let $\overrightarrow{a} =< a_1, a_2, a_3 >$ and $\overrightarrow{b} =< b_1, b_2, b_3 >$. If θ is the angle between the vectors $\overrightarrow{a}$ and $\overrightarrow{b}$, then*

$$\cos\theta = \frac{\overrightarrow{a} \cdot \overrightarrow{b}}{|\overrightarrow{a}||\overrightarrow{b}|}.$$

B. Projection :

C.1. Scalar Projection : component of $\overrightarrow{b}$ along $\overrightarrow{a}$

$$\text{comp}_{\overrightarrow{a}} \overrightarrow{b} = |\overrightarrow{b}| \cos\theta = \frac{\overrightarrow{a} \cdot \overrightarrow{b}}{|\overrightarrow{a}|}.$$

C.2. Vector Projection:

$$\text{proj}_{\overrightarrow{a}} \overrightarrow{b} = \text{comp}_{\overrightarrow{a}} \overrightarrow{b} \frac{\overrightarrow{a}}{|\overrightarrow{a}|} = \left(\frac{\overrightarrow{a} \cdot \overrightarrow{b}}{|\overrightarrow{a}|} \right) \frac{\overrightarrow{a}}{|\overrightarrow{a}|}.$$

0.2 Definition of Inner Product

Definition 2. An **inner product space** on a vector space V is a function

$$< \cdot, \cdot >: V \times V \to \mathbb{C}$$

that satisfies the following properties:

- Positivity: $< v, v >> 0$ for each nonzero $v \in V$.
- Conjugate symmetry: $\overline{< v, w >} =< w, v >$ for all $v, w \in V$.
- Linearity: $< \alpha u + \beta v, w >= \alpha < u, w > + \beta < v, w >$ for all $u, v, w \in V$ and $\alpha, \beta \in \mathbb{C}$.

Remark :
• If $v, w \in V$ and $\beta \in C$,

$$< v, cw >= \bar{c} < v, w > .$$

• For a nonzero $v \in V$, the *length* or *norm* of the vector v is defined by

$$\|v\| = \sqrt{< v, v >}$$

• The distance between u and v is indicated by $\|u - v\|$. We say that $u_k \to u$ if $\|u_k - u\| \to 0$.

• See Examples 0.2, 0.3.

0.3 The Spaces L^2 and ℓ^2

Definition 3. For any interval $a \le t \le b$, the space $L^2([a, b])$ is the set of all square integrable functions defined in $[a, b]$, that is,

$$L^2([a, b]) = \left\{ f : [a, b] \to \mathbb{C} \ : \ \|f\|_2^2 := \int_a^b |f(x)|^2 dx < \infty \right\}.$$

Remark : We say that two functions f and g in $L^2[a, b]$ are equal if $f(x) = g(x)$ for all $x \in [a, b]$ except for a finite x-values.

Definition 4. The L^2 inner product on $L^2[a, b]$ is defined as

$$< f, g >= \int_a^b f(t) \overline{g(t)} dt$$

for $f, g \in L^2[a, b]$.

Definition 5. The space $\ell^2[a,b]$ is the set of all sequences $X = (x_n \in \mathbb{C} : n \in \mathbb{Z})$ with

$$\sum_{n\in\mathbb{Z}} |x_n|^2 < \infty.$$

The inner product on this space is defined by

$$< X, Y >= \sum_{n\in\mathbb{Z}} x_n \overline{y_n}$$

for $X = (x_n)_{n\in\mathbb{Z}}$ and $Y = (y_n)_{n\in\mathbb{Z}}$.

Definition 6. A sequence f_n converges to $f \in L^2[a,b]$ if

$$\|f_n - f\|_2 \to 0, \quad \text{as} \quad n \to \infty.$$

More precisely, given any $\epsilon > 0$, there exists a positive $N \in \mathbb{N}$ such that $\|f_n - f\|_2 < \epsilon$ for all $n > N$.

Definition 7. A sequence f_n converges to f pointwise on $[a,b]$ if for any $t \in [a,b]$ and any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $|f_n(t) - f(t)| < \epsilon$ for all $n \geq N$.

Definition 8. A sequence f_n converges to f uniformy on $[a,b]$ if given any $\epsilon > 0$, there exists a positive $N \in \mathbb{N}$ such that $|f_n(t) - f(t)| < \epsilon$ for all $n > N$ and $t \in [a,b]$.

Remark : Convergence in L^2 is sometimes called *convergence in mean.*

• See Examples 0.9.

Theorem 2. *If a sequence f_n converges uniformly to f as $n \to \infty$ on $[a,b]$, then f_n converges to f in $L^2[a,b]$.*

Remark : The converse of the above theorem is not true; consider the sequence of functions on $[0,1]$:

$$f_n(x) := \begin{cases} 1, & \text{if} \quad x \in [0, 1/n], \\ 0, & \text{if} \quad x \in (1/n, 1]. \end{cases} \tag{1}$$

Also, the pointwise convergence of f_n to f on $[a,b]$ does not imply $f_n \to f$ in $L^2[a,b]$.

0.4 Schwarz and Triangular Inequalities

Theorem 3. *Suppose that V is an inner product space (either real or complex). Then, for all $X, Y \in V$, the followings hold:*

Schwarz Inequality:

$$| < X, Y > | \leq \|X\| \|Y\|.$$

Equaity holds if and only if X and Y are linearly dependent. Moreover, $< X, Y >= \|X\| \|Y\|$ if and only if X or Y is a nonnegative multiple of the other.

Triangle inequality:

$$\|X + Y\| \leq \|X\| + \|Y\|.$$

Equaity holds if and only if X or Y is a nonnegative multiple of the other.

0.5 Orthogonality

In $\mathbb{R}^3$, $< x, y >= |x||y| \cos\theta$ with θ the angle between x and y. We say $x \perp y$ if and only if $< x, y >= 0$.

Definition 9. Suppose that V is an inner product space.

- x and y in V are said to be orthogonal if $< x, y >= 0$.
- $e_1, \ldots, e_N$ in V are orthononal if $\|e_j\| = 1$ and $< e_i, e_j >= 0$ for any $j \neq j$.
- Two spaces V_1 and V_2 are orthogonal ($V_1 \perp V_2$) if for any $v_1 \in V_1$ and $v_2 \in V_2$, $v_1 \perp v_2$, i.e., $< v_1, v_2 >= 0$.

• See Examples 0.13-0.17.

Theorem 4. *Let V be an inner product space and V_0 is a subspace of V. Suppose that $\{e_1, \ldots, e_N\}$ is an orthonormal basis for V_0. Then if $v \in V_0$,*

$$v = \sum_{n=1}^{N} < v, e_j > e_j.$$

• See Examples 0.22-0.23.

Definition 10. Let V be an inner product space and V_0 is a finite dimensional subspace of V. For any $v \in V$, the orthogonal projection of v onto V_0 is the unique vector $v_0 \in V_0$ such that

$$\|v - v_0\| = \min_{w \in V_0} \|v - w\|.$$

Theorem 5. *Let V be an inner product space and V_0 is a finite dimensional subspace of V. Let $v \in V$ and v_0 is its othogonal projection onto V_0 if and only if $v - v_0 \perp V_0$ (i.e., $v - v_0 \perp w$ for any $w \in V_0$.)*

Theorem 6. *Let V be an inner product space and V_0 is a finite dimensional subspace of V. Suppose that $\{e_1, \ldots, e_N\}$ is an orthonormal basis for V_0. If v_0 is the othogonal projection of $v \in V$ onto V_0, v_0 is of the form*

$$v_0 = \sum_{j=1}^{N} c_j e_j, \quad c_j =< v, e_j > .$$

Definition 11. Let V be an inner product space and V_0 is a finite dimensional subspace of V. The orthogonal complement of V_0, say $V_0^{\perp}$, is defined by

$$V_0^{\perp} = \{v \in V :< v, w >= 0, \quad \forall w \in V_0\}.$$

Theorem 7. *Let V be an inner product space and V_0 is a finite dimensional subspace of V. For any $v \in V$, there $\exists v_0 \in V_0$ and $v_1 \in V_0^{\perp}$ such that*

$$v = v_0 + v_1, \quad \text{i.e.,} \quad V = V_0 + V_0^{\perp}.$$

• See Examples 0.26.

Theorem 8. *Let $\{v_1, \ldots, v_N\}$ be a basis for V_0, where V_0 is subspace of V. Then, there $\exists\{e_1, \ldots, e_N\}$: orthonormal basis for V_0 such that e_j is a linear combination of v_j, $j = 1, \ldots, j-1$. Practically, e_j, $j = 1, \ldots, N$, is given (inductively) by*

$$e_j = v_j - \sum_{\ell=1}^{j-1} < v_j, e_\ell > e_\ell / \|v_j - \sum_{\ell=1}^{j-1} < v_j, e_\ell > e_\ell\|.$$

0.7 Least Squares and Linear Predictive Coding

Goal : For the given data (x_j, y_j) with $j = 1, \ldots, N$, find m and b such that

$$\text{minimize} \sum_{j=1}^{N} |y_j - (mx_j + b)|^2.$$

Solution: Let $\mathbf{y} = (y_1, \ldots, y_N)^T$ and $\mathbf{c} = (b,\ m)^T$. Then the solution $\mathbf{c}$ is given as

$$\mathbf{c} = (\mathbf{E}\mathbf{E}^T)^{-1}\mathbf{E}\mathbf{y} \tag{2}$$

where

$$\mathbf{E} = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_N \end{pmatrix}.$$

Put

$$Y = (y_1, \cdots, y_N)^T, \quad X = (x_1, \cdots, x_N)^T, \quad U = (1, \ldots, 1)^T.$$

Let

$$V_0 = \text{span}\{U, X\} = \{mx + bU : m, b \in \mathbb{R}\}.$$

Then the solution $\mathbf{c}$ in (2) can be understood as the coefficient of the projection $P_0(x) = mx + b$ of the vector Y onto V_0.

Generalization: For the given data (x_j, y_j) with $j = 1, \ldots, N$, find a polynomial

$$p(x) = a_0 + \cdots + a_k x^k$$

such that

$$\text{minimize} \sum_{j=1}^{N} |y_j - p(x)|^2.$$

Solution: Let $\mathbf{y} = (y_1, \ldots, y_N)^T$ and $\mathbf{a} = (a_0, \cdots a_k)^T$. Then the solution $\mathbf{a}$ is given as

$$\mathbf{a} = (\mathbf{E}\mathbf{E}^T)^{-1}\mathbf{E}\mathbf{y}$$

where

$$\mathbf{E} = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_N \\ \vdots & \cdots & \vdots \\ x_1^k & \cdots & x_N^k \end{pmatrix}.$$

- See Examples 0.36