Chapter 19: Orthogonality, Orthogonal Projections and the Gram-Schmidt Algorithm

Orthogonality:

- Recall that $\vec{u}, \vec{v} \in \mathbb{R}^n$ are orthogonal $\iff \vec{u} \cdot \vec{v} = 0$
- The dot product satisfies:
  (1) $\vec{u} \cdot \vec{u} \geq 0$ and is equal to zero iff $\vec{u} = \vec{0}$,
  (2) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (commutative)
  (3) $(a\vec{u} + b\vec{v}) \cdot (c\vec{w}) = ac\, \vec{u} \cdot \vec{w} + bc\, \vec{v} \cdot \vec{w}$, for all $a, b, c \in \mathbb{R}$ and $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$,
  (4) $(a\vec{u}) \cdot (b\vec{v} + c\vec{w}) = ab\, \vec{u} \cdot \vec{v} + ac\, \vec{u} \cdot \vec{w}$ (bi-linearity)

Example: Standard basis of $\mathbb{R}^3$, $(1,0,0), (0,1,0), (0,0,1)$ are all orthogonal pairwise.

Definition: A set of vectors in $\mathbb{R}^n$, i.e. $\{\vec{v}_1, \dots, \vec{v}_m\}$, is called "orthogonal" if

$$\boxed{\vec{v}_i \cdot \vec{v}_j = 0,}$$

for all $1 \leq i < j \leq m$ and $\vec{v}_j \neq \vec{0}$ for all $1 \leq i \leq m$.

(They are all non-zero vectors and pairwise orthogonal.)

Example: The standard basis of $\mathbb{R}^n$ is an orthogonal set.

- If $\{\vec{v}_1, \dots, \vec{v}_m\}$ is an orthogonal set, then

$$\|\vec{v}_1 + \cdots + \vec{v}_m\|^2 = \|\vec{v}_1\|^2 + \|\vec{v}_2\|^2 + \cdots + \|\vec{v}_m\|^2,$$

where $\|\vec{v}_j\|^2 = \vec{v}_j \cdot \vec{v}_j$, for $j = 1, 2, \dots, m$.

Theorem (19.1):

Any orthogonal set of vectors is linearly independent, or, any orthogonal set of vectors in $\mathbb{R}^n$ has at most n elements, or, any orthogonal set of n vectors in $\mathbb{R}^n$ is a basis for $\mathbb{R}^n$, called an "orthogonal basis".

Proof: Please refer to the textbook VSF.

Theorem (19.2): (The Expansion Theorem)

Suppose $\{\vec{w}_1, \dots, \vec{w}_m\}$ is an orthogonal basis for a subspace W of $\mathbb{R}^n$. Then any vector $\vec{w} \in W$ can be written as,

$$\vec{w} = \left(\frac{\vec{w} \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\right)\vec{w}_1 + \left(\frac{\vec{w} \cdot \vec{w}_2}{\|\vec{w}_2\|^2}\right)\vec{w}_2 + \cdots + \left(\frac{\vec{w} \cdot \vec{w}_m}{\|\vec{w}_m\|^2}\right)\vec{w}_m,$$

in which the coefficients of $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_m$ are the coordinates of $\vec{w}$ relative to the ordered basis. They are also called the "Fourier coefficients" of $\vec{w}$ relative to the orthogonal basis. This theorem gives us a simple formula to write $\vec{w}$ as a linear combination of basis vectors.

Proof: Since $\{\vec{w}_1, \cdots, \vec{w}_m\}$ form a basis of W, so

$$\vec{w} = a_1\vec{w}_1 + a_2\vec{w}_2 + \cdots + a_m\vec{w}_m$$

$$\Rightarrow \vec{w}_i \cdot \vec{w} = a_i\, \vec{w}_i \cdot \vec{w}_i \Rightarrow a_i = \frac{\vec{w}_i \cdot \vec{w}}{\vec{w}_i \cdot \vec{w}_i} = \frac{\vec{w}_i \cdot \vec{w}}{\|\vec{w}_i\|^2}. \qquad \square$$

Example: Let $W = \{(x,y,z,w) \in \mathbb{R}^4 \mid x-y-z+w=0\}$. W is a subspace of $\mathbb{R}^4$ and $\dim(W) = 3$, since $W = Null([1,-1,-1,1])$ and it has rank 1. So,

$$\dim(W) = \dim(Null([1,-1,-1,1])) = 4 - rank([1,-1,-1,1]) = 3$$

(rank-nullity theorem).

Now, let $\vec{w}_1 = (1,1,1,1)$, $\vec{w}_2 = (1,-1,1,-1)$ and $\vec{w}_3 = (1,1,-1,-1)$. Check that $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ is an orthogonal set, and hence LI, so they form an orthogonal basis for W.

Let's write $\vec{w} = (1,2,3,4) \in W$ (check!) as a linear combination of $\vec{w}_1, \vec{w}_2, \vec{w}_3$, using the theorem (19.2):

$$\frac{\vec{w}\cdot\vec{w}_1}{\|\vec{w}_1\|^2} = \frac{10}{4} = 5/2, \quad \frac{\vec{w}\cdot\vec{w}_2}{\|\vec{w}_2\|^2} = \frac{-2}{4} = -\frac{1}{2}, \quad \frac{\vec{w}\cdot\vec{w}_3}{\|\vec{w}_3\|^2} = \frac{-4}{4} = -1,$$

Therefore,

$$\boxed{\vec{w} = \frac{5}{2}\vec{w}_1 - \frac{1}{2}\vec{w}_2 - \vec{w}_3.}$$

Note: orthogonal basis are meant to simplify calculations, but without them we have to use the usual methods, such as the elementary row reductions, etc.

Definition (19.2): Orthogonal projection onto a subspace

Let W be a subspace of $\mathbb{R}^n$ and $\{\vec{\omega}_1, \vec{\omega}_2, \dots, \vec{\omega}_m\}$ an orthogonal basis for W. Then for any $\vec{v} \in \mathbb{R}^n$, the orthogonal projection of $\vec{v}$ onto W is defined by:

$$\operatorname{proj}_W(\vec{v}) = \left(\frac{\vec{v}\cdot\vec{\omega}_1}{\|\vec{\omega}_1\|^2}\right)\vec{\omega}_1 + \dots + \left(\frac{\vec{v}\cdot\vec{\omega}_m}{\|\vec{\omega}_M\|^2}\right)\vec{\omega}_m$$

Example: $W = \{(x,y,z,\omega) \in \mathbb{R}^4 \mid x - y - z + \omega = 0\}$, $\vec{v} = (1,2,3,5)$

$\vec{\omega}_1 = (1,1,1,1)$, $\vec{\omega}_2 = (1,-1,1,-1)$, $\vec{\omega}_3 = (1,1,-1,-1)$: these are orthogonal basis for W:

$$\frac{\vec{v}\cdot\vec{\omega}_1}{\|\vec{\omega}_1\|^2} = \frac{11}{4}, \quad \frac{\vec{v}\cdot\vec{\omega}_2}{\|\vec{\omega}_2\|^2} = -\frac{3}{4}, \quad \frac{\vec{v}\cdot\vec{\omega}_3}{\|\vec{\omega}_3\|^2} = -\frac{5}{4}, \text{ therefore}$$

$$\operatorname{proj}_W(\vec{v}) = \frac{11}{4}\vec{\omega}_1 - \frac{3}{4}\vec{\omega}_2 - \frac{5}{4}\vec{\omega}_3 = \frac{1}{4}(3,9,13,19).$$

Also, $\vec{v} - \operatorname{proj}_W(\vec{v}) = \frac{1}{4}(1,-1,-1,1)$, which is orthogonal to every vector in W, since $(x,y,z,\omega) \in W \Leftrightarrow$

$$(x,y,z,\omega)\cdot(1,-1,-1,1) = x - y - z + \omega = 0 \checkmark$$

Theorem (19.3): "Best Approximation Theorem"

Let W be a subspace of $\mathbb{R}^n$ and $\vec{v} \in \mathbb{R}^n$. Then,

(1) $\text{proj}_W(\vec{v}) \in W$,

(2) $\vec{v} - \text{proj}_W(\vec{v})$ is orthogonal to every vector in W,

(3) $\text{proj}_W(\vec{v})$ is the best approximation to $\vec{v}$ by vectors in W,

($\text{proj}_W(\vec{v})$ is the closest vector in W to $\vec{v}$),

(4) The vector $\text{proj}_W(\vec{v})$ is the "only" vector in $\mathbb{R}^n$, which satisfies (1) and (2). (orthogonal projection is uniquely characterized by the two properties (1) and (2)).

Proof: please refer to the textbook.

Note: Part (4) is saying that the orthogonal projection does not depend on a choice of orthogonal basis and with different bases we will obtain exactly the same projection.

⊛ The Gram-Schmidt Algorithm (Finding orthogonal basis):

Every subspace W of $\mathbb{R}^n$ has an orthogonal basis. The Gram-Schmidt algorithm takes any basis of W as input and output an orthogonal basis of W.

Suppose $\{\vec{u}_1, \ldots, \vec{u}_m\}$ is any basis for W.

Theorem (19.4): Gram-Schmidt Algorithm

Let $\{\vec{u}_1, \dots, \vec{u}_m\}$ be any basis of W. Define

(1) $\vec{w}_1 = \vec{u}_1$ and $V_1 = \text{span}\{\vec{w}_1\}$.

(2) $\vec{w}_2 = \vec{u}_2 - \text{proj}_{V_1}(\vec{u}_2)$ and $V_2 = \text{Span}\{\vec{w}_1, \vec{w}_2\}$,

(3) $\vec{w}_3 = \vec{u}_3 - \text{proj}_{V_2}(\vec{u}_3)$ and $V_3 = \text{Span}\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$.

⋮

(m) $\vec{w}_m = \vec{u}_m - \text{proj}_{V_{m-1}}(\vec{u}_m)$ and $V_m = \text{Span}\{\vec{w}_1, \dots, \vec{w}_m\}$.

Then, $W = V_m$ and $\{\vec{w}_1, \dots, \vec{w}_m\}$ is an orthogonal basis for W.

or,

(1) $\vec{w}_1 = \vec{u}_1$,

(2) $\vec{w}_2 = \vec{u}_2 - \text{proj}_{\vec{w}_1}(\vec{u}_2)$,

(3) $\vec{w}_3 = \vec{u}_3 - \text{proj}_{\vec{w}_1}(\vec{u}_3) - \text{proj}_{\vec{w}_2}(\vec{u}_3)$,

⋮

(m) $\vec{w}_m = \vec{u}_m - \sum_{i=1}^{m-1} \text{proj}_{\vec{w}_i}(\vec{u}_m)$.

Then, $\{\vec{w}_1, \dots, \vec{w}_m\}$ is an orthogonal basis for W.

* one could also scale each of the vectors $\vec{w}_i$ by dividing by their norm to produce an "orthonormal" basis for W.

Example : $W = \{(1,0,0,1), (1,1,1,0), (2,1,-1,1)\}$

$\vec{w}_1 = (1,0,0,1)$,

$$\vec{w}_2 = \vec{u}_2 - \text{Proj}_{u_1}(\vec{u}_2) = \vec{u}_2 - \frac{\vec{u}_2 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 = (1,1,1,0) - \frac{1}{2}(1,0,0,1)$$
$$= \left(\frac{1}{2}, 1, 1, -\frac{1}{2}\right).$$

$$\vec{w}_3 = \vec{u}_3 - \frac{\vec{u}_3 \cdot \vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1 - \frac{\vec{u}_3 \cdot \vec{w}_2}{\|\vec{w}_2\|^2}\vec{w}_2 = \frac{5}{2}(1,2,-3,-1)$$

So, the resulting orthogonal basis is: (after scaling them to avoid fractions)

$$\{\vec{w}_1, \vec{w}_2, \vec{w}_3\} = \{(1,0,0,1)\ (1,2,2,-1), (1,2,-3,-1)\}$$

To find the orthonormal basis, we have to divide each of these basis vectors by their norm. which would result in: (orthonormal basis)

$$\left\{\frac{\sqrt{2}}{2}(1,0,0,1), \frac{\sqrt{10}}{10}(1,2,2,-1), \frac{\sqrt{15}}{15}(1,2,-3,-1)\right\}$$

Example : Find the best approximation of $\vec{v} = (1,2,3,4)$ to W, where

$$W = \text{span}\{(1,0,0,1), (1,1,1,0), (2,1,-1,1)\} = \{(x,y,z,w) \in \mathbb{R}^4 \mid x-y-w=0\}.$$

orthogonal basis of $W = \{\underbrace{(1,0,0,1)}_{\vec{n}_1}, \underbrace{(1,2,2,-1)}_{\vec{n}_2}, \underbrace{(1,2,-3,-1)}_{\vec{n}_3}\}$. from previous example, so

$$\text{Proj}_W(\vec{v}) = \frac{\vec{v}\cdot\vec{n}_1}{\|\vec{n}_1\|^2}\vec{n}_1 + \frac{\vec{v}\cdot\vec{n}_2}{\|\vec{n}_2\|^2}\vec{n}_2 + \frac{\vec{v}\cdot\vec{n}_3}{\|\vec{n}_3\|^2}\vec{n}_3 = \frac{1}{3}(8,1,9,7).$$