7 Orthogonal Bases

7.1 Gram-Schmidt Proceess

Definition 7.1 A set of vectors in an inner product space is called an *orthogonal set* if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is called *orhonormal*.

Proposition 7.1 *Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be an orthogonal set of nonzero vectors in an inner product space.*

(a) *S is a linearly independent set.*

(b) *Let $W = \operatorname{Span}(S)$ and $\boldsymbol{w} \in W$, then*

$$\boldsymbol{w} = \frac{\langle \boldsymbol{w}, \boldsymbol{v}_1\rangle}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{w}, \boldsymbol{v}_2\rangle}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{w}, \boldsymbol{v}_m\rangle}{\|\boldsymbol{v}_m\|^2}\boldsymbol{v}_m.$$

(c) *If $\boldsymbol{v} \in V$, then*

$$\left\langle \boldsymbol{v} - \Big(\frac{\langle \boldsymbol{v}, \boldsymbol{v}_1\rangle}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{v}, \boldsymbol{v}_2\rangle}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{v}, \boldsymbol{v}_m\rangle}{\|\boldsymbol{v}_m\|^2}\boldsymbol{v}_m\Big), \boldsymbol{w} \right\rangle = 0 \text{ for all } \boldsymbol{w} \in W.$$

Proof. (a): Suppose $k_1\boldsymbol{v} + k_2\boldsymbol{v}_2 + \cdots + k_m\boldsymbol{v}_m = \boldsymbol{0}$. Since $\langle \boldsymbol{0}, \boldsymbol{v}_i\rangle = 0$,

$$0 = \langle k_1\boldsymbol{v} + k_2\boldsymbol{v}_2 + \cdots + k_m\boldsymbol{v}_m, \boldsymbol{v}_i\rangle = k_i\langle \boldsymbol{v}_i, \boldsymbol{v}_i\rangle = k_i\|\boldsymbol{v}_i\|^2.$$

Since $\boldsymbol{v}_i \neq \boldsymbol{0}$, $\|\boldsymbol{v}_i\| \neq 0$. We have $k_i = 0$ for all $i = 1, 2, \ldots, m$. Hence S is a linearly independent set.

(b): Let $\boldsymbol{w} = k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_m\boldsymbol{v}_m \in \operatorname{Span}(S)$. Then $\langle \boldsymbol{w}, \boldsymbol{v}_i\rangle = k_i\|\boldsymbol{v}_i\|^2$. Hence $k_i = \langle \boldsymbol{w}, \boldsymbol{v}_i\rangle/\|\boldsymbol{v}_i\|^2$.

(c): It is straightforward to show that

$$\left\langle \boldsymbol{v} - \Big(\frac{\langle \boldsymbol{v}, \boldsymbol{v}_1\rangle}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{v}, \boldsymbol{v}_2\rangle}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{v}, \boldsymbol{v}_m\rangle}{\|\boldsymbol{v}_m\|^2}\boldsymbol{v}_m\Big), \boldsymbol{v}_i \right\rangle = 0 \text{ for all } i = 1, 2, \ldots, m.$$

Hence we have the assertion, as every vector in W is a linear combination of S. ∎

If S is an orthogonal basis of $W = \operatorname{Span}(S)$, the following vector is called the *orthogonal projection* of $\boldsymbol{v}$ on W and denoted by

$$\operatorname{proj}_W(\boldsymbol{v}) = \frac{\langle \boldsymbol{w}, \boldsymbol{v}_1\rangle}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{w}, \boldsymbol{v}_2\rangle}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{w}, \boldsymbol{v}_m\rangle}{\|\boldsymbol{v}_m\|^2}\boldsymbol{v}_m.$$

Definition 7.2 Let W be a subspace of an inner product space V. A vector $\boldsymbol{u}$ in V is said to be *orthogonal to W* if it is orthogonal to every vector in W, and the set of all vectors in V that are orthogonal to W is called the *orthogonal complement* of W and is denoted by $W^\perp$. Hence $W^\perp = \{\boldsymbol{v} \in V \mid \langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0 \text{ for all } \boldsymbol{w} \in W\}$.

Using the terminology defined above, if $\boldsymbol{v} \in V$, then Proposition 7.1 (c) asserts:

$$\boldsymbol{v} - \mathrm{proj}_W(\boldsymbol{v}) \in W^{\perp}.$$

Theorem 7.2 ((6.3.6) Gram-Schmidt Process) *Every nonzero finite-dimensional inner product space has an orthonormal basis.*

Proof. It suffices to show that when $\{\boldsymbol{u}_1, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ is a basis of V, there is an orthonormal basis $\{\boldsymbol{v}_1, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$.

Let $W_i = \mathrm{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_i\}$. We construct an orthogonal basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_i\}$ of W_i by induction on i.

Step 1. Let $\boldsymbol{v}_1 = \boldsymbol{u}_1/\|\boldsymbol{u}_1\|$. Then $\{\boldsymbol{v}_1\}$ is an orthonormal basis of W_1.

Step 2. Suppose $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{i-1}\}$ is an orthonormal basis of W_{i-1}. Let

$$\begin{aligned}
\boldsymbol{v}_i' &= \boldsymbol{u}_i - \mathrm{proj}_{W_{i-1}}(\boldsymbol{u}_i) \\
&= \boldsymbol{u}_i - (\langle \boldsymbol{u}_i, \boldsymbol{v}_1\rangle \boldsymbol{v}_1 + \langle \boldsymbol{u}_i, \boldsymbol{v}_1\rangle \boldsymbol{v}_2 + \cdots + \langle \boldsymbol{u}_i, \boldsymbol{v}_{i-1}\rangle \boldsymbol{v}_{i-1}), \text{ and} \\
\boldsymbol{v}_i &= \frac{\boldsymbol{v}_i'}{\|\boldsymbol{v}_i'\|}.
\end{aligned}$$

This gives an orthonormall basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_i\}$ of W_i. Note that

$$W_i = \mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_i\} = \mathrm{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_i\}.$$

■

Theorem 7.3 (6.2.5, 6.2.6, 6.3.4) *If W is a subspace of a finite-dimensional inner product space V, then*

(a) *$W^{\perp}$ is a subspace of V.*

(b) *The only vector common to W and $W^{\perp}$ is $\mathbf{0}$.*

(c) *$(W^{\perp})^{\perp} = W$.*

(d) *$\dim V = \dim W + \dim W^{\perp}$.*

(e) *Every vector $\boldsymbol{v} \in V$ is expressed as a sum $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{u}$ such that $\boldsymbol{w} \in W$ and $\boldsymbol{u} \in W^{\perp}$.*

Proof. (a): Recall that $W^{\perp} = \{\boldsymbol{v} \in V \mid \langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0 \text{ for all } \boldsymbol{w} \in W\}$. Since $\mathbf{0} \in W^{\perp}$, $W^{\perp} \neq \emptyset$. Let $\boldsymbol{u}, \boldsymbol{v} \in W^{\perp}$ and k a scalar. Then by definition of $W^{\perp}$, $\langle \boldsymbol{u}, \boldsymbol{w}\rangle = \langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0$ for all $\boldsymbol{w} \in W$. Now for all $\boldsymbol{w} \in W$,

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

Hence $\boldsymbol{u} + \boldsymbol{v} \in W^{\perp}$ and $k\boldsymbol{u} \in W^{\perp}$. Thus $W^{\perp}$ is a subspace of V by Theorem 3.2.

(b): Let $\boldsymbol{v}$ be a vector common to W and $W^{\perp}$, in this case we write $\boldsymbol{v} \in W \cap W^{\perp}$. Consider $\|\boldsymbol{v}\|^2 = \langle \boldsymbol{v}, \boldsymbol{v}\rangle$. Regarding the first $\boldsymbol{v}$ is in $W^{\perp}$ and the second $\boldsymbol{v}$ in W, we have $\|\boldsymbol{v}\|^2 = \langle \boldsymbol{v}, \boldsymbol{v}\rangle = 0$. Hence $\boldsymbol{v} = \mathbf{0}$ and $W \cap W^{\perp} = \{\mathbf{0}\}$.

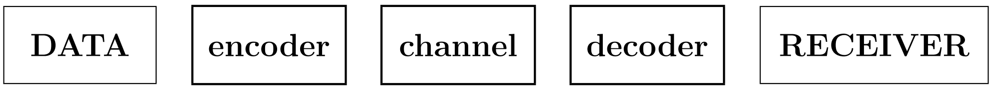
(c): Let $\boldsymbol{w} \in W$ and $\boldsymbol{v} \in W^{\perp}$. Then clearly $\langle \boldsymbol{w}, \boldsymbol{v} \rangle = 0$. Hence $W \subset (W^{\perp})^{\perp}$. Let $\boldsymbol{v} \in (W^{\perp})^{\perp}$. By Proposition 7.1 (c), $\boldsymbol{u} = \boldsymbol{v} - \mathrm{proj}_W(\boldsymbol{v}) \in W^{\perp}$. Since $\boldsymbol{v} \in (W^{\perp})^{\perp}$ and $\mathrm{proj}_W(\boldsymbol{v}) \in W \subset (W^{\perp})^{\perp}$, $\boldsymbol{u} \in (W^{\perp})^{\perp}$ as $(W^{\perp})^{\perp}$ is a subspace by (a). Thus $\boldsymbol{u} \in W^{\perp} \cap (W^{\perp})^{\perp} = \{\boldsymbol{0}\}$ by (b). Therefore $\boldsymbol{v} = \mathrm{proj}_W(\boldsymbol{v}) \in W$ and $(W^{\perp})^{\perp} \subset W$. We have $W = (W^{\perp})^{\perp}$.

(e): Let $\boldsymbol{v} \in V$, $\boldsymbol{w} = \mathrm{proj}_W(\boldsymbol{v})$ and $\boldsymbol{u} = \boldsymbol{v} - \mathrm{proj}_W(\boldsymbol{v})$. Then $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{u}$ and $\boldsymbol{w} \in W$, $\boldsymbol{u} \in W^{\perp}$ by Proposition 7.1 (c).

(d): Let $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_s\}$ be an orthonormal basis of W and $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_t\}$ be an orthonormal basis of $W^{\perp}$. Then $S = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_s, \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_t\}$ is orhonormal set of vectors and by (e) every vector of V can be expressed as a linear combination of these vectors, S is a basis of V. Hence we have the assertion. ■

7.2 An Application to Coding Theory

When we want to transmit some information through a channel, we add extra information to the original data by an encoder so that we can recover the original information by an decoder even if some changes may occur.



We consider binary data only and computation in $F = \{0, 1\}$:

$$0 + 0 = 0,\ 0 + 1 = 1,\ 1 + 0 = 1,\ 1 + 1 = 0.$$

$$0 \cdot 0 = 0,\ 0 \cdot 1 = 0,\ 1 \cdot 0 = 0,\ 1 \cdot 1 = 1.$$

Thus we consider vector spaces over F and all entries of matrices are in F. We need two matrices.

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad H^T = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

We send a binary data $\boldsymbol{a} = (a_1, a_2, a_3.a_4) \in F^4$ of length four..

Instead of sending $\boldsymbol{a}$ we send $\boldsymbol{a} \cdot G$.

$$12 : (1100) \cdot G = (1100110), \quad 4 : (0100) \cdot G = (0100101).$$

Suppose there was some noise and it changed a part. As a result we got $\boldsymbol{x}$. Then we compute $\boldsymbol{x} \cdot H^T$. The decimal number of it tells where the error occurred.

For example if the fourth digit of (1100110) changed from 0 to 1 and (1101110) is received. Since $(1101110)H^T = (100)$ and the decimal number of (100) is 4, it indicates that the error occurred in the fourth digit. Hence we can recover (1100110) by changing the fourth digit from 1 to 0. For the second, suppose we received (0100100). Since $(0100100)H^T = (111)$, it suggests that the error occurred in the seventh digit. We recover the original information (0100101).

Let us think why it worked. Let

$$C = \{\boldsymbol{x} \cdot G \mid \boldsymbol{x} \in F^4\} \subset V = F^7.$$

Then we find

$$\boldsymbol{c} + \boldsymbol{c}' \in C \text{ for every } \boldsymbol{c},\ \boldsymbol{c}' \in C$$

and C is a subspace of V.

It is clear as we multiplied the vector in the data set F^4 by the matrix G. C is the range (or the image) of the linear transformation defined by G. We call these codes binary linear codes of length 7.

Moreover, let us observe that we have

$$G \cdot H^T = O,$$

where O denotes the zero matrix (of size 4×3 in this case). When we send $\boldsymbol{a}$ (a binary information of length 4), we compute $\boldsymbol{c} = \boldsymbol{a} \cdot G$ and send this word of length 7 instead. Hence,

$$\boldsymbol{c} \cdot H^T = \boldsymbol{a} \cdot G \cdot H^T = \boldsymbol{a} \cdot O = (000) \text{ for every } \boldsymbol{c} \in C$$

Let $\boldsymbol{e}_1 = (1000000), \boldsymbol{e}_2 = (0100000), \ldots, \boldsymbol{e}_7 = (0000001)$. Let $\boldsymbol{c} \in C$. If there is an error in ith bit (or position), then we have $\boldsymbol{c}+\boldsymbol{e}_i$. Then we have $(\boldsymbol{c}+\boldsymbol{e}_i)\cdot H^T = \boldsymbol{e}_i \cdot H^T$, which is nothing but the ith column of H^T. But the ith column of H^T is the binary expression of i. We can tell the position the error occurred. Hence we can recover the correct information as far as the error occurred at only one place. Then what happens if no error occurred? In that case we conclude that the error occurred at 0th position.

Exercise 7.1 [Quiz 7] Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{e}_1, \boldsymbol{e}_2$ and $\boldsymbol{e}_3$ be vectors in $\boldsymbol{R}^3$ given below.

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix},\ \boldsymbol{v}_2 = \begin{bmatrix} -2 \\ 7 \\ 4 \end{bmatrix},\ \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ -8 \\ -6 \end{bmatrix}, \boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},\ \boldsymbol{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

For $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{R}^3$, let $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T \boldsymbol{v}$ be the inner product and $U = \mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$. You may quote the facts shown in previous quizzes.

1. Compute $\boldsymbol{v}_2 - \frac{\langle \boldsymbol{v}_2, \boldsymbol{v}_1 \rangle}{\|\boldsymbol{v}_1\|^2} \boldsymbol{v}_1$.

2. Find an orthonormal basis of U.

3. Find an orthonormal basis of $\boldsymbol{R}^3$ containing the basis constructed in 2.

4. Find a basis of $U^\perp$.

5. Express each of $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ as a linear combination of the orthonormal basis constructed in 3.