

1. Recall the definition for the notion of *orthogonality* from the handout *Inner product, norm, and orthogonality*:

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. We say \mathbf{u} is orthogonal to \mathbf{v} , and write $\mathbf{u} \perp \mathbf{v}$, if and only if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Also recall these basic properties of orthogonality:

- (a) Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Then $\mathbf{u} \perp \mathbf{v}$ if and only if $\mathbf{v} \perp \mathbf{u}$.
- (b) Suppose $\mathbf{u} \in \mathbb{R}^n$. Then $\mathbf{u} \perp \mathbf{u}$ if and only if $\mathbf{u} = \mathbf{0}_n$.
- (c) Suppose $\mathbf{u} \in \mathbb{R}^n$. Then $(\mathbf{u} \perp \mathbf{v} \text{ for any } \mathbf{v} \in \mathbb{R}^n)$ if and only if $\mathbf{u} = \mathbf{0}_n$.
- (d) Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Then $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ if and only if $\mathbf{u} \perp \mathbf{v}$.

2. Theorem (A).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ be non-zero vectors in \mathbb{R}^n . Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ are pairwise orthogonal (in the sense that $\mathbf{u}_i \perp \mathbf{u}_j$ whenever $i \neq j$.)

Then the statements below hold:

- (a) $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ are linearly independent.
- (b) For any $\mathbf{v} \in \mathbb{R}^n$, if \mathbf{v} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ then $\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\|\mathbf{u}_2\|^2} \mathbf{u}_2 + \dots + \frac{\langle \mathbf{v}, \mathbf{u}_k \rangle}{\|\mathbf{u}_k\|^2} \mathbf{u}_k$.

3. Proof of Theorem (A).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ be non-zero vectors in \mathbb{R}^n . Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ are pairwise orthogonal.

- (a) Pick any $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$. Suppose $\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_k \mathbf{u}_k = \mathbf{0}$.
For each $j = 1, 2, \dots, k$, we have

$$\begin{aligned} \alpha_j \|\mathbf{u}_j\|^2 &= \alpha_1 \langle \mathbf{u}_1, \mathbf{u}_j \rangle + \alpha_2 \langle \mathbf{u}_2, \mathbf{u}_j \rangle + \dots + \alpha_k \langle \mathbf{u}_k, \mathbf{u}_j \rangle \\ &= \langle \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_k \mathbf{u}_k, \mathbf{u}_j \rangle \\ &= \langle \mathbf{0}, \mathbf{u}_j \rangle = 0 \end{aligned}$$

Since \mathbf{u}_j is not the zero vector, $\|\mathbf{u}_j\| \neq 0$. Then $\alpha_j = 0$.

It follows that $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ are linearly independent.

- (b) Exercise. (Imitate what has been done above.)

4. Definition. (Orthonormal set and orthonormal basis.)

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k \in \mathbb{R}^n$.

- (a) We say that $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal set in \mathbb{R}^n if and only if $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ are pairwise orthogonal and $\|\mathbf{u}_j\| = 1$ for each $j = 1, 2, \dots, k$.
- (b) Suppose V is a subspace of \mathbb{R}^n . Then we say that $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal basis for V if and only if $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute a basis for V and also constitute an orthonormal set.

Remark. When $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal set in \mathbb{R}^n , they constitute an orthonormal basis for $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\})$.

5. Theorem (B).

Let W be a subspace of \mathbb{R}^n .

Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal basis for W .

Suppose $\mathbf{s}, \mathbf{t} \in W$. Define $\beta_j = \langle \mathbf{s}, \mathbf{u}_j \rangle$, $\gamma_j = \langle \mathbf{t}, \mathbf{u}_j \rangle$ for each $j = 1, 2, \dots, k$.

Then the statements below hold:

- (a) $\mathbf{s} = \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \dots + \beta_k \mathbf{u}_k$.
- (b) $\|\mathbf{s}\|^2 = \beta_1^2 + \beta_2^2 + \dots + \beta_k^2$.
- (c) $\langle \mathbf{s}, \mathbf{t} \rangle = \beta_1 \gamma_1 + \beta_2 \gamma_2 + \dots + \beta_k \gamma_k$.

6. Proof of Theorem (B).

Let W be a subspace of \mathbb{R}^n .

Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal basis for W .

Suppose $\mathbf{s}, \mathbf{t} \in W$. Define $\beta_j = \langle \mathbf{s}, \mathbf{u}_j \rangle$, $\gamma_j = \langle \mathbf{t}, \mathbf{u}_j \rangle$ for each $j = 1, 2, \dots, k$.

(a) Since $\mathbf{s} \in W$ and $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute a basis for W , \mathbf{s} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$.

Then, by Theorem (A), $\mathbf{s} = \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \dots + \beta_k \mathbf{u}_k$.

(b) We have

$$\begin{aligned} \|\mathbf{s}\|^2 &= \langle \mathbf{s}, \mathbf{s} \rangle = \langle \mathbf{s}, \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \dots + \beta_k \mathbf{u}_k \rangle = \beta_1 \langle \mathbf{s}, \mathbf{u}_1 \rangle + \beta_2 \langle \mathbf{s}, \mathbf{u}_2 \rangle + \dots + \beta_k \langle \mathbf{s}, \mathbf{u}_k \rangle \\ &= \beta_1^2 + \beta_2^2 + \dots + \beta_k^2. \end{aligned}$$

(c) Exercise.

7. Theorem (C).

Let W be a subspace of \mathbb{R}^n .

Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal basis for W .

Suppose $\mathbf{z} \in \mathbb{R}^n$.

Define $\alpha_1 = \langle \mathbf{z}, \mathbf{u}_1 \rangle$, $\alpha_2 = \langle \mathbf{z}, \mathbf{u}_2 \rangle$, ..., $\alpha_k = \langle \mathbf{z}, \mathbf{u}_k \rangle$.

Define $\mathbf{v} \in W$ by $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_k \mathbf{u}_k$.

Define $\mathbf{y} \in \mathbb{R}^n$ by $\mathbf{y} = \mathbf{z} - \mathbf{v}$.

Then the statements below hold:

(a) i. $\mathbf{z} = \mathbf{v} + \mathbf{y}$.

ii. $\mathbf{y} \perp \mathbf{s}$ for any $\mathbf{s} \in W$. (In particular, $\mathbf{y} \perp \mathbf{v}$.)

(b) Suppose $\mathbf{s} \in W$. Then $\|\mathbf{z} - \mathbf{s}\| \geq \|\mathbf{z} - \mathbf{v}\|$. Equality holds if and only if $\mathbf{s} = \mathbf{v}$.

(c) The inequality $\|\mathbf{z}\|^2 \geq \alpha_1^2 + \alpha_2^2 + \dots + \alpha_k^2$ holds.

Moreover, the statements below are logically equivalent:

i. $\mathbf{z} \in W$.

ii. $\mathbf{z} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_k \mathbf{u}_k$.

iii. $\|\mathbf{z}\|^2 = \alpha_1^2 + \alpha_2^2 + \dots + \alpha_k^2$.

iv. For any $\mathbf{x} \in \mathbb{R}^n$, $\langle \mathbf{z}, \mathbf{x} \rangle = \alpha_1 \langle \mathbf{u}_1, \mathbf{x} \rangle + \alpha_2 \langle \mathbf{u}_2, \mathbf{x} \rangle + \dots + \alpha_k \langle \mathbf{u}_k, \mathbf{x} \rangle$.

8. Illustrations of the construction described in Theorem (C).

(a) Let $\mathbf{u}_1 = \begin{bmatrix} \sqrt{3}/2 \\ 1/2 \end{bmatrix}$, and $W = \text{Span}(\{\mathbf{u}_1\})$

Note that $\|\mathbf{u}_1\| = 1$.

Then \mathbf{u}_1 constitute an orthonormal basis for W .

• Suppose $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$.

Define $\alpha_1 = \langle \mathbf{z}, \mathbf{u}_1 \rangle$.

Define $\mathbf{v} = \alpha_1 \mathbf{u}_1$.

$$\text{Then } \mathbf{v} = \left(\frac{\sqrt{3}}{2} z_1 + \frac{1}{2} z_2\right) \begin{bmatrix} \sqrt{3}/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 3z_1/4 + \sqrt{3}z_2/4 \\ \sqrt{3}z_1/4 + z_2/4 \end{bmatrix} = \begin{bmatrix} 3/4 & \sqrt{3}/4 \\ \sqrt{3}/4 & 1/4 \end{bmatrix} \mathbf{z}.$$

Define $\mathbf{y} = \mathbf{z} - \mathbf{v}$.

$$\text{Then } \mathbf{y} = \begin{bmatrix} z_1/4 - \sqrt{3}z_2/4 \\ -\sqrt{3}z_1/4 \end{bmatrix} + 3z_2/4 = \begin{bmatrix} 1/4 & -\sqrt{3}/4 \\ -\sqrt{3}/4 & 3/4 \end{bmatrix} \mathbf{z}.$$

\mathbf{z} is 'decomposed' into the sum of \mathbf{v}, \mathbf{y} which form a pair of vectors orthogonal to each other, and in which the vector \mathbf{y} is orthogonal to every vector in W .

(b) Let $\mathbf{u}_1 = \mathbf{e}_1^{(3)}$, $\mathbf{u}_2 = \mathbf{e}_2^{(3)}$, and $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2\})$.

Note that $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = 1$ and $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0$.

Then $\mathbf{u}_1, \mathbf{u}_2$ constitute an orthonormal basis for W .

- Suppose $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$.

Define $\alpha_1 = \langle \mathbf{z}, \mathbf{u}_1 \rangle$, $\alpha_2 = \langle \mathbf{z}, \mathbf{u}_2 \rangle$.

Define $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2$.

$$\text{Then } \mathbf{v} = z_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + z_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{z}.$$

Define $\mathbf{y} = \mathbf{z} - \mathbf{v}$.

$$\text{Then } \mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{z}.$$

\mathbf{z} is ‘decomposed’ into the sum of \mathbf{v}, \mathbf{y} which form a pair of vectors orthogonal to each other, and in which the vector \mathbf{y} is orthogonal to every vector in W .

- (c) Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$, and $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2\})$.

Note that $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = 1$ and $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0$.

Then $\mathbf{u}_1, \mathbf{u}_2$ constitute an orthonormal basis for W .

- Suppose $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$.

Define $\alpha_1 = \langle \mathbf{z}, \mathbf{u}_1 \rangle$, $\alpha_2 = \langle \mathbf{z}, \mathbf{u}_2 \rangle$.

Define $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2$.

$$\text{Then } \mathbf{v} = \left(\frac{z_1}{3} + \frac{2z_2}{3} + \frac{2z_3}{3}\right) \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} + \left(-\frac{2z_1}{3} - \frac{z_2}{3} + \frac{2z_3}{3}\right) \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \dots = \begin{bmatrix} 5/9 & 4/9 & -2/9 \\ 4/9 & 5/9 & 2/9 \\ -2/9 & 2/9 & 8/9 \end{bmatrix} \mathbf{z}.$$

Define $\mathbf{y} = \mathbf{z} - \mathbf{v}$.

$$\text{Then } \mathbf{y} = \dots = \begin{bmatrix} 4/9 & -4/9 & 2/9 \\ -4/9 & 4/9 & -2/9 \\ 2/9 & -2/9 & 1/9 \end{bmatrix} \mathbf{z}.$$

\mathbf{z} is ‘decomposed’ into the sum of \mathbf{v}, \mathbf{y} which form a pair of vectors orthogonal to each other, and in which the vector \mathbf{y} is orthogonal to every vector in W .

- (d) Let $\mathbf{u}_1 = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$, and $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2\})$.

Note that $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = 1$ and $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0$.

Then $\mathbf{u}_1, \mathbf{u}_2$ constitute an orthonormal basis for W .

- Suppose $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}$.

Define $\alpha_1 = \langle \mathbf{z}, \mathbf{u}_1 \rangle$, $\alpha_2 = \langle \mathbf{z}, \mathbf{u}_2 \rangle$.

Define $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2$.

$$\text{Then } \mathbf{v} = \left(\frac{z_1}{2} + \frac{z_2}{2} + \frac{z_3}{2} + \frac{z_4}{2}\right) \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} + \left(-\frac{z_1}{2} - \frac{z_2}{2} + \frac{z_3}{2} + \frac{z_4}{2}\right) \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix} = \dots = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix} \mathbf{z}.$$

Define $\mathbf{y} = \mathbf{z} - \mathbf{v}$.

$$\text{Then } \mathbf{y} = \dots = \begin{bmatrix} 1/2 & -1/2 & 0 & 0 \\ -1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & -1/2 \\ 0 & 0 & -1/2 & 1/2 \end{bmatrix} \mathbf{z}.$$

\mathbf{z} is ‘decomposed’ into the sum of \mathbf{v}, \mathbf{y} which form a pair of vectors orthogonal to each other, and in which the vector \mathbf{y} is orthogonal to every vector in W .

(e) Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \\ 2/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, and $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\})$.

Note that $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = \|\mathbf{u}_3\| = 1$ and $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{u}_3 \rangle = \langle \mathbf{u}_2, \mathbf{u}_3 \rangle = 0$.

Then $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ constitute an orthonormal basis for W .

• Suppose $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}$.

Define $\alpha_1 = \langle \mathbf{z}, \mathbf{u}_1 \rangle$, $\alpha_2 = \langle \mathbf{z}, \mathbf{u}_2 \rangle$, $\alpha_3 = \langle \mathbf{z}, \mathbf{u}_3 \rangle$.

Define $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \alpha_3 \mathbf{u}_3$.

Then

$$\begin{aligned} \mathbf{v} &= \left(\frac{z_1}{3} + \frac{2z_2}{3} + \frac{2z_4}{3}\right) \begin{bmatrix} 1/3 \\ 2/3 \\ 0 \\ 2/3 \end{bmatrix} + \left(\frac{2z_1}{3} - \frac{z_2}{3} + \frac{2z_3}{3}\right) \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \\ 0 \end{bmatrix} + \left(-\frac{2z_2}{3} - \frac{z_3}{3} + \frac{2z_4}{3}\right) \begin{bmatrix} 0 \\ -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} \\ &= \cdots = \begin{bmatrix} 5/9 & 0 & 4/9 & 2/9 \\ 0 & 1 & 0 & 0 \\ 4/9 & 0 & 5/9 & -2/9 \\ 2/9 & 0 & -2/9 & 8/9 \end{bmatrix} \mathbf{z}. \end{aligned}$$

Define $\mathbf{y} = \mathbf{z} - \mathbf{v}$.

$$\text{Then } \mathbf{y} = \cdots = \begin{bmatrix} 4/9 & 0 & -4/9 & -2/9 \\ 0 & 0 & 0 & 0 \\ -4/9 & 0 & 4/9 & 2/9 \\ -2/9 & 0 & 2/9 & 1/9 \end{bmatrix} \mathbf{z}.$$

\mathbf{z} is ‘decomposed’ into the sum of \mathbf{v}, \mathbf{y} which form a pair of vectors orthogonal to each other, and in which the vector \mathbf{y} is orthogonal to every vector in W .

9. Proof of Theorem (C).

Let W be a subspace of \mathbb{R}^n .

Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal basis for W .

Suppose $\mathbf{z} \in \mathbb{R}^n$.

Define $\alpha_1 = \langle \mathbf{z}, \mathbf{u}_1 \rangle$, $\alpha_2 = \langle \mathbf{z}, \mathbf{u}_2 \rangle$, ..., $\alpha_k = \langle \mathbf{z}, \mathbf{u}_k \rangle$.

Define $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_k \mathbf{u}_k$, and $\mathbf{y} = \mathbf{z} - \mathbf{v}$.

- (a) i. By definition, $\mathbf{z} = \mathbf{v} + \mathbf{y}$.
 ii. Pick any $\mathbf{s} \in W$. Define $\beta_1 = \langle \mathbf{s}, \mathbf{u}_1 \rangle$, $\beta_2 = \langle \mathbf{s}, \mathbf{u}_2 \rangle$, ..., $\beta_k = \langle \mathbf{s}, \mathbf{u}_k \rangle$. Then $\mathbf{s} = \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \cdots + \beta_k \mathbf{u}_k$.
 Note that $\langle \mathbf{v}, \mathbf{s} \rangle = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \cdots + \alpha_k \beta_k$.
 Also note that

$$\langle \mathbf{z}, \mathbf{s} \rangle = \langle \mathbf{z}, \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \cdots + \beta_k \mathbf{u}_k \rangle = \beta_1 \langle \mathbf{z}, \mathbf{u}_1 \rangle + \beta_2 \langle \mathbf{z}, \mathbf{u}_2 \rangle + \cdots + \beta_k \langle \mathbf{z}, \mathbf{u}_k \rangle = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \cdots + \alpha_k \beta_k.$$

Then $\langle \mathbf{y}, \mathbf{s} \rangle = \langle \mathbf{z} - \mathbf{v}, \mathbf{s} \rangle = \langle \mathbf{z}, \mathbf{s} \rangle - \langle \mathbf{v}, \mathbf{s} \rangle = 0$.

Therefore $\mathbf{y} \perp \mathbf{s}$.

- (b) Suppose $\mathbf{s} \in W$.

Note that $\mathbf{v} \in W$. Then $\mathbf{v} - \mathbf{s} \in W$. Then $\mathbf{z} - \mathbf{v} \perp \mathbf{v} - \mathbf{s} \in W$.

- We have $\|\mathbf{z} - \mathbf{s}\|^2 = \|(\mathbf{z} - \mathbf{v}) + (\mathbf{v} - \mathbf{s})\|^2 = \|\mathbf{z} - \mathbf{v}\|^2 + \|\mathbf{v} - \mathbf{s}\|^2$.

Since $\|\mathbf{v} - \mathbf{s}\|^2 \geq 0$, we have $\|\mathbf{z} - \mathbf{s}\|^2 \geq \|\mathbf{z} - \mathbf{v}\|^2$.

Then $\|\mathbf{z} - \mathbf{s}\| \geq \|\mathbf{z} - \mathbf{v}\|$.

- Suppose $\mathbf{s} = \mathbf{v}$. Then $\|\mathbf{z} - \mathbf{s}\| = \|\mathbf{z} - \mathbf{v}\|$.

- Suppose $\|\mathbf{z} - \mathbf{s}\| = \|\mathbf{z} - \mathbf{v}\|$. Then $\|\mathbf{v} - \mathbf{s}\|^2 = 0$. Therefore $\mathbf{s} - \mathbf{v} = \mathbf{0}$. Hence $\mathbf{s} = \mathbf{v}$.

- (c) Exercise. (Apply the definition of \mathbf{v} and \mathbf{y} . The inequality concerned is simply ‘ $\|\mathbf{z}\| \geq \|\mathbf{v}\|$ ’ in disguise. Equality holds if and only if $\mathbf{y} = \mathbf{0}$.)

10. Recall the definition for the notion of *orthogonal complement of a subspace of \mathbb{R}^n* from the handout *Orthogonal complement*.

Suppose W is a subspace of \mathbb{R}^n .

The perp of W , which as a set is given by $W^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \perp \mathbf{u} \text{ for any } \mathbf{u} \in W\}$, is called the *orthogonal complement of W in \mathbb{R}^n* .

Also recall the result (\star) from the same handout:

Suppose W is a subspace of \mathbb{R}^n . Then for any $\mathbf{z} \in \mathbb{R}^n$, there exist some unique $\mathbf{s} \in W$, $\mathbf{t} \in W^\perp$ such that $\mathbf{z} = \mathbf{s} + \mathbf{t}$.

With the help of the result (\star), we can enrich the content of part (a) in Theorem (C) by appending a ‘uniqueness part’.

11. **Theorem (D).**

Let W be a subspace of \mathbb{R}^n .

Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal basis for W .

Suppose $\mathbf{z} \in \mathbb{R}^n$.

Define $\alpha_1 = \langle \mathbf{z}, \mathbf{u}_1 \rangle$, $\alpha_2 = \langle \mathbf{z}, \mathbf{u}_2 \rangle$, ..., $\alpha_k = \langle \mathbf{z}, \mathbf{u}_k \rangle$.

Define $\mathbf{v} \in W$ by $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_k \mathbf{u}_k$.

Define $\mathbf{y} \in \mathbb{R}^n$ by $\mathbf{y} = \mathbf{z} - \mathbf{v}$.

Then the statements below hold:

- (a) i. $\mathbf{z} = \mathbf{v} + \mathbf{y}$.
 ii. $\mathbf{y} \perp \mathbf{s}$ for any $\mathbf{s} \in W$. (In particular, $\mathbf{y} \perp \mathbf{v}$.)
- (b) Suppose $\mathbf{v}', \mathbf{y}' \in \mathbb{R}^n$.
 Suppose $\mathbf{v}' \in W$, $\mathbf{z} = \mathbf{v}' + \mathbf{y}'$, and $\mathbf{y}' \perp \mathbf{s}$ for any $\mathbf{s} \in W$. Then $\mathbf{v}' = \mathbf{v}$ and $\mathbf{y}' = \mathbf{y}$.

Remarks.

- In plain words, statement (b) is saying that \mathbf{z} is decomposed in a unique way as a sum of two vectors, one in W and the other in W' . The two vectors are \mathbf{v} and \mathbf{y} respectively.

The vector \mathbf{v} is determined independent of the choice of orthonormal bases for W :

Suppose that $\mathbf{u}'_1, \mathbf{u}'_2, \dots, \mathbf{u}'_k$ also constitute an orthonormal basis for W , and $\alpha'_1 = \langle \mathbf{z}, \mathbf{u}'_1 \rangle$, $\alpha'_2 = \langle \mathbf{z}, \mathbf{u}'_2 \rangle$, ..., $\alpha'_k = \langle \mathbf{z}, \mathbf{u}'_k \rangle$.

Further suppose that $\mathbf{v}' = \alpha'_1 \mathbf{u}'_1 + \alpha'_2 \mathbf{u}'_2 + \dots + \alpha'_k \mathbf{u}'_k$ and $\mathbf{y}' = \mathbf{z} - \mathbf{v}'$.

Then it happens that $\mathbf{v}' = \mathbf{v}$ and $\mathbf{y}' = \mathbf{y}$.

- Terminology.* This uniqueness makes sense of naming the vectors \mathbf{v}, \mathbf{y} with reference to \mathbf{z} and W . The vector \mathbf{v} is called the *orthogonal projection of the vector \mathbf{z} onto W* . It is denoted by $\text{pr}_W(\mathbf{z})$. The vector \mathbf{y} is called the *orthogonal complement of \mathbf{z} with respect to W* .

The other parts of Theorem (C) can be re-stated in terms of orthogonal projections.

12. **Theorem (E).**

Let W be a subspace of \mathbb{R}^n , and $\mathbf{z} \in \mathbb{R}^n$.

- (a) Suppose $\mathbf{s} \in W$. Then $\|\mathbf{z} - \mathbf{s}\| \geq \|\mathbf{z} - \text{pr}_W(\mathbf{z})\|$. Equality holds if and only if $\mathbf{s} = \text{pr}_W(\mathbf{z})$.
- (b) The inequality $\|\mathbf{z}\| \geq \|\text{pr}_W(\mathbf{z})\|$ holds. Equality holds if and only if $\mathbf{z} \in W$.

Remarks.

- Statement (a) says that amongst all vectors in W , it is $\text{pr}_W(\mathbf{z})$ whose distance with \mathbf{z} is the smallest. In plain words, $\text{pr}_W(\mathbf{z})$ is the ‘closest (or best) approximation’ to \mathbf{z} amongst all vectors in W . This result is the corner stone of the ‘least square method’ for finding approximations.
- Statement (b) says that the ‘length’ of the vector \mathbf{v} is no less than that of its projection onto W , which is $\text{pr}_W(\mathbf{z})$. This inequality is known as Bessel’s Inequality.

13. **Theorem (F).**

Let W be a subspace of \mathbb{R}^n .

Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal basis for W .

Define the $(n \times k)$ -matrix U by $U = [\mathbf{u}_1 \mid \mathbf{u}_2 \mid \dots \mid \mathbf{u}_k]$.

Then the statements below hold:

- (a) For any $\mathbf{z} \in \mathbb{R}^n$, $\text{pr}_W(\mathbf{z}) = UU^t\mathbf{z}$.
- (b) UU^t is symmetric and idempotent.
- (c) $\mathcal{C}(UU^t) = W$.
- (d) $\mathcal{N}(UU^t) = W^\perp$.

Remarks.

- When $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_k$ constitute an orthonormal basis for W and $S = [\mathbf{s}_1 \mid \mathbf{s}_2 \mid \dots \mid \mathbf{s}_k]$, we have $\text{pr}_W(\mathbf{z}) = SS^t\mathbf{z}$ for any $\mathbf{z} \in \mathbb{R}^n$. It follows that $UU^t = SS^t$.

This $(n \times n)$ -square matrix is independent of the choice of orthonormal bases for W .

- *Terminology.* This uniqueness makes sense of naming the matrix UU^t with reference to W . The matrix UU^t is called the projection matrix from \mathbb{R}^n onto W . Multiplication by this matrix from the left to a vector in \mathbb{R}^n results in the orthogonal projection of that vector onto W .

14. **Proof of Theorem (F).**

Let W be a subspace of \mathbb{R}^n .

Suppose $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ constitute an orthonormal basis for W .

Define the $(n \times k)$ -matrix U by $U = [\mathbf{u}_1 \mid \mathbf{u}_2 \mid \dots \mid \mathbf{u}_k]$.

- (a) Pick any $\mathbf{z} \in \mathbb{R}^n$. We have

$$\begin{aligned} UU^t\mathbf{z} &= U \begin{bmatrix} \frac{\mathbf{u}_1^t \mathbf{z}}{\mathbf{u}_2^t \mathbf{z}} \\ \vdots \\ \frac{\mathbf{u}_k^t \mathbf{z}}{\mathbf{u}_k^t \mathbf{z}} \end{bmatrix} = U \begin{bmatrix} \mathbf{u}_1^t \mathbf{z} \\ \mathbf{u}_2^t \mathbf{z} \\ \vdots \\ \mathbf{u}_k^t \mathbf{z} \end{bmatrix} = [\mathbf{u}_1 \mid \mathbf{u}_2 \mid \dots \mid \mathbf{u}_k] \begin{bmatrix} \langle \mathbf{z}, \mathbf{u}_1 \rangle \\ \langle \mathbf{z}, \mathbf{u}_2 \rangle \\ \vdots \\ \langle \mathbf{z}, \mathbf{u}_k \rangle \end{bmatrix} \\ &= \langle \mathbf{z}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{z}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \dots + \langle \mathbf{z}, \mathbf{u}_k \rangle \mathbf{u}_k = \text{pr}_W(\mathbf{z}) \end{aligned}$$

- (b) We have $(UU^t)^t = (U^t)^t U^t = UU^t$. Then UU^t is symmetric.

We have $(UU^t)^2 = (UU^t)(UU^t) = U(U^t U)U^t = UI_k U^t = UU^t$. Then UU^t is idempotent.

- (c) We verify that $W = \mathcal{C}(UU^t)$:

- [We verify that for any $\mathbf{x} \in \mathbb{R}^n$, if $\mathbf{x} \in W$ then $\mathbf{x} \in \mathcal{C}(UU^t)$.]

Pick any $\mathbf{x} \in \mathbb{R}^n$. Suppose $\mathbf{x} \in W$.

Since $\mathbf{x} \in W$, We have $\mathbf{x} = \text{pr}_W(\mathbf{x})$.

By the result in part (a), we have $\text{pr}_W(\mathbf{x}) = UU^t\mathbf{x}$.

Then $\mathbf{x} = UU^t\mathbf{x}$. Therefore, by definition, $\mathbf{x} \in \mathcal{C}(UU^t)$.

- [We verify that for any $\mathbf{x} \in \mathbb{R}^n$, if $\mathbf{x} \in \mathcal{C}(UU^t)$ then $\mathbf{x} \in W$.]

Pick any $\mathbf{x} \in \mathbb{R}^n$. Suppose $\mathbf{x} \in \mathcal{C}(UU^t)$.

Then there exists some $\mathbf{s} \in \mathbb{R}^k$ such that $\mathbf{x} = UU^t\mathbf{s}$.

Define $\mathbf{p} \in \mathbb{R}^k$ by $\mathbf{p} = U^t\mathbf{s}$.

Then $\mathbf{x} = U\mathbf{p}$.

Therefore, by definition, $\mathbf{x} \in \mathcal{C}(U)$.

By definition, $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}) = \mathcal{C}(U)$. Hence $\mathbf{x} \in W$.

- (d) We have verified that $\mathcal{C}(UU^t) = W$.

By part (b), UU^t is symmetric.

Then $\mathcal{N}((UU^t)) = \mathcal{N}((UU^t)^t) = (\mathcal{C}(UU^t))^\perp = W^\perp$.

15. **Illustrations of the content of Theorem (F).**

- (a) Let $\mathbf{u}_1 = \begin{bmatrix} \sqrt{3}/2 \\ 1/2 \end{bmatrix}$, and $W = \text{Span}(\{\mathbf{u}_1\})$

\mathbf{u}_1 constitute an orthonormal basis for W .

Define $U = \mathbf{u}_1$.

$$\text{We have } UU^t = \begin{bmatrix} 3/4 & \sqrt{3}/4 \\ \sqrt{3}/4 & 1/4 \end{bmatrix}.$$

UU^t is the projection matrix from \mathbb{R}^2 onto W : for any $\mathbf{z} \in \mathbb{R}^2$, $\text{pr}_W(\mathbf{z}) = UU^t\mathbf{z}$.

- (b) Let $\mathbf{u}_1 = \mathbf{e}_1^{(3)}$, $\mathbf{u}_2 = \mathbf{e}_2^{(3)}$, and $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2\})$.

$\mathbf{u}_1, \mathbf{u}_2$ constitute an orthonormal basis for W .

Define $U = [\mathbf{u}_1 \mid \mathbf{u}_2]$.

$$\text{We have } UU^t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

UU^t is the projection matrix from \mathbb{R}^3 onto W : for any $\mathbf{z} \in \mathbb{R}^3$, $\text{pr}_W(\mathbf{z}) = UU^t\mathbf{z}$.

- (c) Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$, and $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2\})$.

$\mathbf{u}_1, \mathbf{u}_2$ constitute an orthonormal basis for W .

Define $U = [\mathbf{u}_1 \mid \mathbf{u}_2]$.

$$\text{We have } UU^t = \begin{bmatrix} 5/9 & 4/9 & -2/9 \\ 4/9 & 5/9 & 2/9 \\ -2/9 & 2/9 & 8/9 \end{bmatrix}.$$

UU^t is the projection matrix from \mathbb{R}^3 onto W : for any $\mathbf{z} \in \mathbb{R}^3$, $\text{pr}_W(\mathbf{z}) = UU^t\mathbf{z}$.

- (d) Let $\mathbf{u}_1 = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$, and $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2\})$.

$\mathbf{u}_1, \mathbf{u}_2$ constitute an orthonormal basis for W .

Define $U = [\mathbf{u}_1 \mid \mathbf{u}_2]$.

$$\text{We have } UU^t = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}.$$

UU^t is the projection matrix from \mathbb{R}^4 onto W : for any $\mathbf{z} \in \mathbb{R}^4$, $\text{pr}_W(\mathbf{z}) = UU^t\mathbf{z}$.

- (e) Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \\ 2/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, and $W = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\})$.

$\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ constitute an orthonormal basis for W .

Define $U = [\mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3]$.

$$\text{We have } UU^t = \begin{bmatrix} 5/9 & 0 & 4/9 & 2/9 \\ 0 & 1 & 0 & 0 \\ 4/9 & 0 & 5/9 & -2/9 \\ 2/9 & 0 & -2/9 & 8/9 \end{bmatrix}.$$

UU^t is the projection matrix from \mathbb{R}^4 onto W : for any $\mathbf{z} \in \mathbb{R}^4$, $\text{pr}_W(\mathbf{z}) = UU^t\mathbf{z}$.