MATH 1030 Chapter 17

17.1 Basic properties of inner products

Definition 17.1. Given two vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^m$, we define

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{m} [\mathbf{v}]_i [\mathbf{w}]_i = [\mathbf{v}]_1 [\mathbf{w}]_1 + \cdots + [\mathbf{v}]_m [\mathbf{w}]_m. \tag{17.1}$$

It is called the **inner product** of $\mathbb{R}^m$. The vector space $\mathbb{R}^m$ together with the operation $\langle -, - \rangle$ is called an inner product space. If we regard $\mathbf{v}$ and $\mathbf{w}$ as $m \times 1$ matrices, then we can write:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^t \mathbf{w}. \tag{17.2}$$

Example 17.2. We have:

$$\left\langle \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\rangle = 1 \times 3 + 2 \times 4 = 11$$

and:

$$\left\langle \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right\rangle = 1 \times 4 + 2 \times 5 + 3 \times 6 = 32.$$

Proposition 17.3. *For any* $\mathbf{v}, \mathbf{w}, \mathbf{u} \in \mathbb{R}^m$ *and* $\alpha \in \mathbb{R}$. *We have*

1. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$.

2. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle$.

3. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$.

4. $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ *for* $\mathbf{v} \neq \mathbf{0}$.

Proof of Proposition 17.3. 1. We compute

$$\begin{aligned}\langle \mathbf{v}+\mathbf{w},\mathbf{u}\rangle &= [\mathbf{v}+\mathbf{w}]_1[\mathbf{u}]_1+[\mathbf{v}+\mathbf{w}]_2[\mathbf{u}]_2+\cdots+[\mathbf{v}+\mathbf{w}]_m[\mathbf{u}]_m\\ &= ([\mathbf{v}]_1+[\mathbf{w}]_1)[\mathbf{u}]_1+([\mathbf{v}]_2+[\mathbf{w}]_2)[\mathbf{u}]_2+\cdots+([\mathbf{v}]_n+[\mathbf{w}]_m)[\mathbf{u}]_m\\ &= [\mathbf{v}]_1[\mathbf{u}]_1+[\mathbf{w}]_1[\mathbf{u}]_1+[\mathbf{v}]_2[\mathbf{u}]_2+[\mathbf{w}]_2[\mathbf{u}]_2+\cdots+[\mathbf{v}]_m[\mathbf{u}]_m+[\mathbf{w}]_m[\mathbf{u}]_m\\ &= [\mathbf{v}]_1[\mathbf{u}]_1+\cdots+[\mathbf{v}]_m[\mathbf{u}]_m+[\mathbf{w}]_1[\mathbf{u}]_1+\cdots+[\mathbf{w}]_m[\mathbf{u}]_m\\ &= \langle \mathbf{v},\mathbf{u}\rangle+\langle \mathbf{w},\mathbf{u}\rangle\,.\end{aligned}$$

Or we can use (17.2):

$$\langle \mathbf{v}+\mathbf{w},\mathbf{u}\rangle = (\mathbf{v}+\mathbf{w})^t\mathbf{u} = (\mathbf{v}^t+\mathbf{w}^t)\mathbf{u}$$

$$= \mathbf{v}^t\mathbf{u}+\mathbf{w}^t\mathbf{u} = \langle \mathbf{v},\mathbf{u}\rangle+\langle \mathbf{w},\mathbf{u}\rangle\,.$$

2. We compute

$$\begin{aligned}\langle \alpha\mathbf{v},\mathbf{w}\rangle &= [\alpha\mathbf{v}]_1[\mathbf{w}]_1+[\alpha\mathbf{v}]_2[\mathbf{w}]_2+\cdots+[\alpha\mathbf{v}]_m[\mathbf{w}]_m\\ &= \alpha[\mathbf{v}]_1[\mathbf{w}]_1+\alpha[\mathbf{v}]_2[\mathbf{w}]_2+\cdots+\alpha[\mathbf{v}]_m[\mathbf{w}]_m\\ &= \alpha([\mathbf{v}]_1[\mathbf{w}]_1+[\mathbf{v}]_2[\mathbf{w}]_2+\cdots+[\mathbf{v}]_m[\mathbf{w}]_m)\\ &= \alpha\,\langle \mathbf{v},\mathbf{w}\rangle\,.\end{aligned}$$

Or we can use (17.2):

$$\langle \alpha\mathbf{v},\mathbf{w}\rangle = (\alpha\mathbf{v})^t\mathbf{w} = \alpha\mathbf{v}^t\mathbf{w} = \alpha\,\langle \mathbf{v},\mathbf{w}\rangle\,.$$

3. We compute

$$\begin{aligned}\langle \mathbf{v},\mathbf{w}\rangle &= [\mathbf{v}]_1[\mathbf{w}]_1+[\mathbf{v}]_2[\mathbf{w}]_2+\cdots+[\mathbf{v}]_m[\mathbf{w}]_m.\\ &= [\mathbf{w}]_1[\mathbf{v}]_1+[\mathbf{w}]_2[\mathbf{v}]_2+\cdots+[\mathbf{w}]_m[\mathbf{v}]_m.\\ &= \langle \mathbf{w},\mathbf{v}\rangle\,.\end{aligned}$$

4. We compute

$$\langle \mathbf{v},\mathbf{v}\rangle = [\mathbf{v}]_1^2+[\mathbf{v}]_2^2+\cdots+[\mathbf{v}]_m^2 \geq 0.$$

Noting that $\langle \mathbf{v},\mathbf{v}\rangle = 0$ if and only if $[\mathbf{v}]_i = 0$ for all $1 \leq i \leq n$, we see that $\mathbf{v} = \mathbf{0}$

□

Proposition 17.4. *Let $\alpha, \beta \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w}, \mathbf{u} \in \mathbb{R}^m$. We have*

1. $\langle \alpha \mathbf{v} + \beta \mathbf{w}, \mathbf{u} \rangle = \alpha \langle \mathbf{v}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle$.
2. $\langle \mathbf{u}, \alpha \mathbf{v} + \beta \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle + \beta \langle \mathbf{u}, \mathbf{w} \rangle$.
3. $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$.
4. *If* $\langle \mathbf{v}, \mathbf{x} \rangle = 0$ *for all* $\mathbf{x} \in \mathbb{R}^m$*, then* $\mathbf{v} = \mathbf{0}$.
5. *If* $\langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{w}, \mathbf{x} \rangle$ *for all* $x \in \mathbb{R}^m$*, then* $\mathbf{v} = \mathbf{w}$.

Proof of Proposition 17.4.

1. By Proposition 17.3 item 1, we have:

$$\langle \alpha \mathbf{v} + \beta \mathbf{w}, \mathbf{u} \rangle = \langle \alpha \mathbf{v}, \mathbf{u} \rangle + \langle \beta \mathbf{w}, \mathbf{u} \rangle .$$

   By Proposition 17.3 item 2,

$$\langle \alpha \mathbf{v}, \mathbf{u} \rangle + \langle \beta \mathbf{w}, \mathbf{u} \rangle = \alpha \langle \mathbf{v}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle .$$

   Or we can also use (17.2):

$$\langle \alpha \mathbf{v} + \beta \mathbf{w}, \mathbf{u} \rangle = (\alpha \mathbf{v} + \beta \mathbf{w})^t \mathbf{u}$$

$$= \alpha (\mathbf{v})^t \mathbf{u} + \beta (\mathbf{w})^t \mathbf{u} = \alpha \langle \mathbf{v}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle .$$

2. By the previous part and Proposition Proposition 17.3 item 3, we have

$$\langle \mathbf{u}, \alpha \mathbf{v} + \beta \mathbf{w} \rangle = \langle \alpha \mathbf{v} + \beta \mathbf{w}, \mathbf{u} \rangle$$

$$= \alpha \langle \mathbf{v}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle + \beta \langle \mathbf{u}, \mathbf{w} \rangle .$$

   Or we can also use (17.2) (fill the detail).

3. We compute

$$\langle \mathbf{0}, \mathbf{v} \rangle = 0[\mathbf{v}]_1 + \cdots + 0[\mathbf{v}]_m = 0$$

   and

$$\langle \mathbf{v}, \mathbf{0} \rangle = [\mathbf{v}]_1 0 + \cdots + [\mathbf{v}]_m 0 = 0.$$

4. Suppose $\langle \mathbf{v}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in \mathbb{R}^m$. Let $\mathbf{x} = \mathbf{v}$. Then $\langle \mathbf{v}, \mathbf{v} \rangle = 0$. By Proposition Proposition 17.3 item 4, $\mathbf{v} = \mathbf{0}$.

5. Suppose $\langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{w}, \mathbf{x} \rangle$, then $0 = \langle \mathbf{v}, \mathbf{x} \rangle - \langle \mathbf{w}, \mathbf{x} \rangle = \langle \mathbf{v} - \mathbf{w}, \mathbf{x} \rangle$ for all $\mathbf{x} \in V$. By the previous part $\mathbf{v} - \mathbf{w} = \mathbf{0}$. So $\mathbf{v} = \mathbf{w}$.

□

Definition 17.5 (Norm)**.** The **norm** (or **length**) of $\mathbf{v} \in \mathbb{R}^n$ is defined to be $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$. Note that $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$. So the symbol $\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ is meaningful.

Example 17.6. Let $V = \mathbb{R}^3$ with the standard inner product. Let

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Then

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$$

and

$$\|\mathbf{w}\| = \sqrt{\langle \mathbf{w}, \mathbf{w} \rangle} = \sqrt{1^2 + 0^2 + 1^2} = \sqrt{2}.$$

Proposition 17.7. *Let $\alpha \in \mathbb{R}$ and $\mathbf{v} \in \mathbb{R}^m$.*

1. $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$.
2. $\|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\|$.
3. *Suppose that* $\mathbf{v} \neq \mathbf{0}$ *and let* $\alpha = \frac{1}{\|\mathbf{v}\|}$. *Then* $\|\alpha \mathbf{v}\| = 1$.

Proof of Proposition 17.7.

1. $\|\mathbf{v}\| = 0 \iff 0 = \|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle$. By Proposition (Proposition 17.3), item 4, the above is true if and only if $\mathbf{v} = \mathbf{0}$.
2. $\|\alpha \mathbf{v}\| = \sqrt{\langle \alpha \mathbf{v}, \alpha \mathbf{v} \rangle} = \sqrt{\alpha \langle \mathbf{v}, \alpha \mathbf{v} \rangle} = \sqrt{\alpha^2 \langle \mathbf{v}, \mathbf{v} \rangle} = |\alpha| \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = |\alpha| \|\mathbf{v}\|$.
3. By the previous part

$$\|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\| = \frac{1}{\|\mathbf{v}\|} \|\mathbf{v}\| = 1.$$

□

Definition 17.8 (unit vector)**.** A vector $\mathbf{v} \in \mathbb{R}^m$ is said to be a **unit vector** if $\|\mathbf{v}\| = 1$. A non-zero vector $\mathbf{v}$ can be **normalized** to a unit vector $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ (see the previous proposition item 3).

Example 17.9. In Example 17.6, the vectors $\mathbf{v}$ and $\mathbf{w}$ can be normalized to:

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}}{\sqrt{14}} = \begin{bmatrix} \frac{1}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \end{bmatrix}$$

and

$$\frac{\mathbf{w}}{\|\mathbf{w}\|} = \frac{\mathbf{w}}{\sqrt{2}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

respectively.

17.2 Orthogonal sets

Definition 17.10. Two vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$ are said **orthogonal** or **perpendicular** if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. In this case we write $\mathbf{v} \perp \mathbf{w}$.

Example 17.11.

1. Let $V = \mathbb{R}^3$. Then

   $$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \perp \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix},$$

   as

   $$\left\langle \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\rangle = 1 \times (-1) + 2 \times (-1) + 3 \times 1 = 0.$$

2. Let $V = \mathbb{R}^m$. Then $\mathbf{e}_i \perp \mathbf{e}_j$ if $i \neq j$.

Definition 17.12. A subset $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of $\mathbb{R}^m$ is said to be **orthogonal** if the following conditions hold:

1. $\mathbf{0} \notin S$, i.e. $\mathbf{v}_i \neq \mathbf{0}$ for $i = 1, \dots, k$.
2. $\mathbf{v}_i \perp \mathbf{v}_j$ for $i \neq j$, i.e., $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for $i \neq j$.

Example 17.13.

1. $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$ is orthogonal.

2. $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}$ is orthogonal.

3. For any $k \le m$, the set $S = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k\} \subset \mathbb{R}^m$ is orthogonal.

Proposition 17.14. *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an orthogonal subset of $\mathbb{R}^m$. Let*

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k,$$

$$\mathbf{w} = \beta_1 \mathbf{v}_1 + \cdots + \beta_k \mathbf{v}_k.$$

Then, for $\alpha_i, \beta_i \in \mathbb{R}$, $i = 1, \dots k$, we have

$$\langle \mathbf{v}, \mathbf{w} \rangle = \alpha_1 \beta_1 \|\mathbf{v}_1\|^2 + \cdots + \alpha_k \beta_k \|\mathbf{v}_k\|^2.$$

Proof of Proposition 17.14. First for $1 \le i \le k$, we compute

$$\begin{aligned} \langle \mathbf{v}, \mathbf{v}_i \rangle &= \langle \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k, \mathbf{v}_i \rangle \\ &= \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_i \rangle + \cdots + \alpha_k \langle \mathbf{v}_k, \mathbf{v}_i \rangle \\ &= \alpha_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle = \alpha_i \|\mathbf{v}_i\|^2. \end{aligned}$$

The last step follows from the fact that $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0$ for $j \neq i$. But then

$$\begin{aligned} \langle \mathbf{v}, \mathbf{w} \rangle &= \langle \mathbf{v}, \beta_1 \mathbf{v}_1 + \cdots + \beta_k \mathbf{v}_k \rangle \\ &= \beta_1 \langle \mathbf{v}, \mathbf{v}_1 \rangle + \cdots + \beta_k \langle \mathbf{v}, \mathbf{v}_k \rangle \\ &= \alpha_1 \beta_1 \|\mathbf{v}_1\|^2 + \cdots + \alpha_k \beta_k \|\mathbf{v}_k\|^2. \end{aligned}$$

□

Theorem 17.15. *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an orthogonal subset of $\mathbb{R}^m$. Then S is linearly independent.*

Proof of Theorem 17.15. Suppose that we have a relation of linear dependence:

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0}.$$

For $1 \le i \le k$ we have

$$\langle \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k, \mathbf{v}_i \rangle = \langle \mathbf{0}, \mathbf{v}_i \rangle = 0.$$

i.e. for $1 \le i \le k$,

$$\begin{aligned} \langle \mathbf{v}, \mathbf{v}_i \rangle &= \langle \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k, \mathbf{v}_i \rangle \\ &= \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_i \rangle + \cdots + \alpha_k \langle \mathbf{v}_k, \mathbf{v}_i \rangle \\ &= \alpha_i \|\mathbf{v}_i\|^2 = 0. \end{aligned}$$

So for $1 \le i \le k$ we have

$$\alpha_i = 0.$$

Therefore the relation of linear dependence is trivial. Hence S is linearly independent. □

Theorem 17.16. *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an orthogonal subset of $\mathbb{R}^m$. Suppose that $\mathbf{v} \in \langle S \rangle$. Write*

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k,$$

for some $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, k$. Then

$$\alpha_i = \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2},$$

i.e.

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \cdots + \frac{\langle \mathbf{v}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2}\mathbf{v}_k.$$

Proof of Theorem 17.16. Suppose that $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k$. Then, for $1 \le i \le k$, we compute

$$\begin{aligned} \langle \mathbf{v}, \mathbf{v}_i \rangle &= \langle \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k, \mathbf{v}_i \rangle \\ &= \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_i \rangle + \cdots + \alpha_k \langle \mathbf{v}_k, \mathbf{v}_i \rangle \\ &= \alpha_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle = \alpha_i \|\mathbf{v}_i\|^2. \end{aligned}$$

Hence

$$\alpha_i = \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2}.$$

□

Remark. The advantage of using the above method is that we don't have to solve linear equations to find the linear combination.

In order to use the theorem, we need to ensure that $\mathbf{v} \in \langle S \rangle$.

Example 17.17. We use Example 17.13. Let $S = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}$.

Given that

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

is in $\langle S\rangle$, we find the following linear combinations:

$$\alpha_1 = \frac{\langle \mathbf{v}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2} = \frac{6}{3} = 2.$$

$$\alpha_2 = \frac{\langle \mathbf{v}, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2} = \frac{-1}{2} = -\frac{1}{2}.$$

$$\alpha_3 = \frac{\langle \mathbf{v}, \mathbf{v}_3\rangle}{\|\mathbf{v}_3\|^2} = \frac{-3}{6} = -\frac{1}{2}.$$

Hence

$$\mathbf{v} = 2\mathbf{v}_1 - \frac{1}{2}\mathbf{v}_2 - \frac{1}{2}\mathbf{v}_3.$$

Definition 17.18. Let V be a subspace of $\mathbb{R}^m$. A subset S of V is said to be an **orthogonal basis** for V if S is a basis of V and S is orthogonal.

If S is an orthogonal subset of V, then by Theorem Theorem 17.15, it is automatically linearly independent. So in order to check if S is an orthogonal basis, we need only check that $\langle S\rangle = V$. So we have the following result.

Theorem 17.19. *Let V be a subspace of $\mathbb{R}^m$. Suppose that S is an orthogonal subset of V. Then S is an orthogonal basis if and only if $\langle S\rangle = V$.*

Corollary 17.20. *Suppose that S is an orthogonal subset of $\mathbb{R}^m$. Then S is a basis of $\langle S\rangle$.*

Corollary 17.21. *Let V be a subspace of $\mathbb{R}^m$. Suppose that $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthogonal basis of V. Then for any $\mathbf{v} \in V$, we have*

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \cdots + \frac{\langle \mathbf{v}, \mathbf{v}_n\rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n$$

Proof of Corollary 17.21. This follows from Theorem 17.16. □

Example 17.22.

1. The set $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$ an orthogonal basis of $\mathbb{R}^2$.

2. The set $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}$ is an orthogonal basis of $\mathbb{R}^3$. Indeed, $\dim \mathbb{R}^3 = 3$ and S, with 3 vectors, is linearly independent.

3. The set $S = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$ is an orthogonal basis of $\mathbb{R}^m$. It is called **the standard basis** for V.

Definition 17.23. A subset $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of $\mathbb{R}^m$ is said to be **orthonormal** if it is orthogonal and every vector in S is a unit vector, i.e.

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Let V be a subspace of $\mathbb{R}^m$. The subset S is said to be an **orthonormal basis** for V if it is orthonormal and is a basis of V.

Because an orthonormal set S is orthogonal, the above theorems regarding orthogonal sets are also true for orthonormal sets. In particular we have the following result.

Theorem 17.24. *Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be an orthonormal subset of $\mathbb{R}^m$ and let $\mathbf{v} \in \langle S \rangle$. Then*

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{v}, \mathbf{v}_k \rangle \mathbf{v}_k.$$

Proof of Theorem 17.24. By Theorem Theorem 17.16 and $\|\mathbf{v}_i\| = 1$ for $i = 1, \dots, k$. □

If $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is an orthogonal subset of $\mathbb{R}^m$, then $\left\{ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \dots, \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|} \right\}$ is an orthonormal subset. The process is called **normalization**.

Example 17.25. 1. The set $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$ is an orthogonal basis of $\mathbb{R}^2$. Normalizing it, we obtain an orthonormal basis

$$S' = \left\{ \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}.$$

2. The set $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}$ is an orthogonal basis of $\mathbb{R}^3$. Normalizing it, we obtain an orthonormal basis

$$S' = \left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}.$$

17.3 Gram-Schmidt Orthogonalization process

Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be an orthogonal subset of $\mathbb{R}^m$. If $\mathbf{w} \in \langle S \rangle$, then

$$\mathbf{w} = \frac{\langle \mathbf{w}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \cdots + \frac{\langle \mathbf{w}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2}\mathbf{v}_k.$$

But what if $\mathbf{w}$ is not in $\langle S \rangle$? Let's compare the difference. We have the following theorem.

Theorem 17.26. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be an orthogonal subset of $\mathbb{R}^m$ and let $\mathbf{w} \in \mathbb{R}^m$. Then, for each $i = 1, \ldots, k$, the vector*

$$\mathbf{v} = \mathbf{w} - \frac{\langle \mathbf{w}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{w}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2}\mathbf{v}_k$$

is perpendicular to $\mathbf{v}_i$.

Proof of Theorem 17.26. For $1 \le i \le k$, we compute

$$\langle \mathbf{v}, \mathbf{v}_i \rangle = \langle \mathbf{w}, \mathbf{v}_i \rangle - \frac{\langle \mathbf{w}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \langle \mathbf{v}_1, \mathbf{v}_i \rangle - \cdots - \frac{\langle \mathbf{w}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2} \langle \mathbf{v}_k, \mathbf{v}_i \rangle .$$

Because $\langle \mathbf{v}_j, \mathbf{v}_i \rangle$ is 0 unless $j = i$, the above becomes

$$\langle \mathbf{w}, \mathbf{v}_i \rangle - \frac{\langle \mathbf{w}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \langle \mathbf{v}_i, \mathbf{v}_i \rangle = \langle \mathbf{w}, \mathbf{v}_i \rangle - \langle \mathbf{w}, \mathbf{v}_i \rangle = 0.$$

Hence $\mathbf{v} \perp \mathbf{v}_i$ for $i = 1, \ldots, k$. □

Theorem 17.27 (Gram-Schmidt Orthogonalization Process)**.** *Let $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ be a linearly independent subset of V. Let $\mathbf{v}_1 = \mathbf{w}_1$ and set, recursively,*

$$\mathbf{v}_\ell = \mathbf{w}_\ell - \frac{\langle \mathbf{w}_\ell, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{w}_\ell, \mathbf{v}_{\ell-1} \rangle}{\|\mathbf{v}_{\ell-1}\|^2}\mathbf{v}_{\ell-1} \text{ for } 2 \le \ell \le k.$$

Then $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthogonal set.

Moreover, $\operatorname{Span}\{\mathbf{w}_1, \ldots, \mathbf{w}_\ell\} = \operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ *for $\ell = 1, \ldots, k$. In particular* $\operatorname{Span} S = \operatorname{Span} S'$.

The process of obtaining S' by the above procedure is called the **Gram-Schmidt Orthogonalization Process**.

Proof of Gram-Schmidt Orthogonalization Process. We have $\langle \{\mathbf{w}_1\} \rangle = \langle \{\mathbf{v}_1\} \rangle$. We are going to add one vector at a time.

Suppose that $\langle \{\mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1}\} \rangle = \langle \{\mathbf{w}_1, \ldots, \mathbf{w}_{\ell-1}\} \rangle$ and that the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1}\}$ is orthogonal.

Observe that $\mathbf{v}_l$ is a linear combination of $\mathbf{v}_1, \cdots, \mathbf{v}_{l-1}$ and $\mathbf{w}_l$. Since $\langle\{\mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1}\}\rangle = \langle\{\mathbf{w}_1, \ldots, \mathbf{w}_{\ell-1}\}\rangle$, we have:

$$\mathbf{v}_\ell \in \langle\{\mathbf{w}_1, \ldots, \mathbf{w}_{\ell-1}, \mathbf{w}_\ell\}\rangle .$$

Hence,

$$\langle\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}\rangle \subseteq \langle\{\mathbf{w}_1, \ldots, \mathbf{w}_\ell\}\rangle$$

Likewise, by the formula defining $\mathbf{v}_\ell$ we have:

$$\mathbf{w}_\ell \in \langle\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}\rangle .$$

Hence,

$$\langle\{\mathbf{w}_1, \ldots, \mathbf{w}_\ell\}\rangle \subseteq \langle\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}\rangle .$$

It now follows that:

$$\langle\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}\rangle = \langle\{\mathbf{w}_1, \ldots, \mathbf{w}_\ell\}\rangle .$$

By Theorem Theorem 17.26, $\mathbf{v}_\ell \perp \mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1}$. Thus $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ is orthogonal. Repeating the process for the incrementing ℓ reaches $\ell = k$, we complete the proof. □

Corollary 17.28. *Suppose that V is a subspace of $\mathbb{R}^m$. Then there exists an orthogonal basis of V.*

Moreover, the orthogonal basis can be further normalized to be an orthonormal basis.

Proof of Corollary 17.28. By Theorem 14.12, there exists a basis $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for V. Applying Gram-Schmidt orthogonalization process to S, we obtain an orthogonal set $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.

By Theorem Theorem 17.27 (Gram-Schmidt Orthogonalization Process), $\langle S'\rangle = \langle S\rangle = V$. By Theorem Theorem 17.19, S' is an orthogonal basis. Normalizing S', we can also obtain an orthonormal basis. □

The above proof actually describes a method to find orthogonal (orthonormal) basis of V.

Example 17.29. Let $V = \mathbb{R}^4$ with the standard inner product. Let

$$\mathbf{w}_1 = \begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix}0\\1\\2\\1\end{bmatrix}.$$

Then $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is linearly independent. We can apply Gram-Schmidt orthogonalization process to this set of vectors. Take $\mathbf{v}_1 = \mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$. Then

$$\mathbf{v}_2 = \mathbf{w}_2 - \frac{\langle \mathbf{w}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Also

$$\mathbf{v}_3 = \mathbf{w}_3 - \frac{\langle \mathbf{w}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{w}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} - \frac{2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{2}\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

The set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis of $\langle\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}\rangle$. To obtain an orthonormal basis of $\langle\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}\rangle$, we can normalized the vectors

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Example 17.30. Let $V = \mathcal{N}\left(\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}\right)$. Find an orthonormal basis of V. The set

$$S = \left\{ \mathbf{w}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis of V. We apply Gram-Schmidt orthogonalization process to the set S:

$$\mathbf{v}_1 = \mathbf{w}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{v}_2 = \mathbf{w}_2 - \frac{\langle \mathbf{w}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix}.$$

$$\mathbf{v}_3 = \mathbf{w}_3 - \frac{\langle \mathbf{w}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{w}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/3 \\ -1/3 \\ -1/3 \\ 1 \end{bmatrix}.$$

So

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/3 \\ -1/3 \\ -1/3 \\ 1 \end{bmatrix} \right\}.$$

is an orthogonal basis of V. Normalizing it, we can obtain an orthonormal basis of V:

$$\left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{12}}\begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix} \right\}.$$

The above process will be easier if we start with another basis:

$$S = \left\{ \mathbf{w}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\}$$

Now the first two vectors are perpendicular. Apply Gram-Schmidt orthogonalization process to it:

$$\mathbf{v}_1 = \mathbf{w}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

$$\mathbf{v}_2 = \mathbf{w}_2 - \frac{\langle \mathbf{w}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \mathbf{w}_2 - 0\mathbf{v}_1 = \mathbf{w}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

$$\mathbf{v}_3 = \mathbf{w}_3 - \frac{\langle \mathbf{w}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{w}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix}.$$

So

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix} \right\}$$

is an orthogonal basis of V. Normalizing it, we obtain an orthonormal basis

$$\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix} \right\}.$$

17.4 Cauchy-Schwarz Inequality

Can be skipped, will not appear in final exam

Theorem 17.31 (Cauchy-Schwarz Inequality)**.** *For* $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$*,*

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|.$$

Proof of Cauchy-Schwarz Inequality. The statement is trivial if $\mathbf{w} = \mathbf{0}$. Suppose $\mathbf{w} \neq \mathbf{0}$. Let $t \in \mathbb{R}$, then

$$0 \leq \|\mathbf{v} - t\mathbf{w}\|^2 = \langle \mathbf{v} - t\mathbf{w}, \mathbf{v} - t\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} - t\mathbf{w} \rangle - t \langle \mathbf{w}, \mathbf{v} - t\mathbf{w} \rangle$$

$$= \langle \mathbf{v}, \mathbf{v} \rangle - t \langle \mathbf{v}, \mathbf{w} \rangle - t \langle \mathbf{w}, \mathbf{v} \rangle + t^2 \langle \mathbf{w}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle - 2t \langle \mathbf{v}, \mathbf{w} \rangle + t^2 \langle \mathbf{w}, \mathbf{w} \rangle$$

Substituting

$$t = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle}$$

into the above, we obtain

$$0 \leq \langle \mathbf{v}, \mathbf{v} \rangle - \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\langle \mathbf{w}, \mathbf{w} \rangle} = \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\|\mathbf{w}\|^2}.$$

Hence

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \sqrt{\|\mathbf{v}\|^2 \|\mathbf{w}\|^2} = \|\mathbf{v}\| \|\mathbf{w}\|.$$

□

Remark. The t above is obtained by minimizing the quadratic equation $\langle \mathbf{v}, \mathbf{v} \rangle - 2t \langle \mathbf{v}, \mathbf{w} \rangle + t^2 \langle \mathbf{w}, \mathbf{w} \rangle$.

Following the proof, the equality occurs if (i) $\mathbf{v} = \mathbf{0}$ or (ii) $\mathbf{w} = \mathbf{0}$ or (iii) $\mathbf{v} - t\mathbf{w} = 0 \Leftrightarrow \mathbf{v}$ and $\mathbf{w}$ are parallel, i.e., $\mathbf{v} = \alpha \mathbf{w}$ for some scalar $\alpha \in \mathbb{R}$.

Theorem 17.32 (Triangle Inequality)**.** *For any* $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$*, we have:*

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$

Proof of Triangle Inequality.

$$\|\mathbf{v} + \mathbf{w}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle = \|\mathbf{v}\|^2 + 2 \langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2.$$

By the Cauchy-Schwarz inequality

$$| \langle \mathbf{v}, \mathbf{w} \rangle | \leq \|\mathbf{v}\| \|\mathbf{w}\|,$$

thus

$$\|\mathbf{v} + \mathbf{w}\|^2 \leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2 = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2.$$

The result follows by taking square roots on both sides. □

Example 17.33. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$ with the standard inner product. Let

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}$$

Cauchy-Schwarz inequality:

$$|v_1 w_1 + \cdots + v_m w_m| \leq \sqrt{v_1^2 + \cdots + v_m^2} \sqrt{w_1^2 + \cdots + w_m^2}.$$

Triangle inequality:

$$\sqrt{(v_1 + w_1)^2 + \cdots + (v_m + w_m)^2} \leq \sqrt{v_1^2 + \cdots + v_m^2} + \sqrt{w_1^2 + \cdots + w_m^2}.$$