6.1 Inner Products

DEFINITION 1 An *inner product* on a real vector space V is a function that associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors in V in such a way that the following axioms are satisfied for all vectors \mathbf{u}, \mathbf{v} , and \mathbf{w} in V and all scalars k.

- 1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ [Symmetry axiom]
- 2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ [Additivity axiom]
- 3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k \langle \mathbf{u}, \mathbf{v} \rangle$ [Homogeneity axiom]
- 4. $\langle \mathbf{v}, \mathbf{v} \rangle \ge 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$ [Positivity axiom]

A real vector space with an inner product is called a real inner product space.

DEFINITION 2 If V is a real inner product space, then the **norm** (or **length**) of a vector v in V is denoted by ||v|| and is defined by

$$\|v\| = \sqrt{\langle v, \, v \rangle}$$

and the *distance* between two vectors is denoted by $d(\mathbf{u}, \mathbf{v})$ and is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$$

A vector of norm 1 is called a unit vector.

THEOREM 6.1.1 If \mathbf{u} and \mathbf{v} are vectors in a real inner product space V, and if k is a scalar, then:

- (a) $\|\mathbf{v}\| \ge 0$ with equality if and only if $\mathbf{v} = \mathbf{0}$.
- (b) $||k\mathbf{v}|| = |k| ||\mathbf{v}||$.
- (c) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.
- (d) $d(\mathbf{u}, \mathbf{v}) \ge 0$ with equality if and only if $\mathbf{u} = \mathbf{v}$.

$$w_1, w_2, \ldots, w_n$$

are positive real numbers, which we will call weights, and if $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, \dots, v_n)$ are vectors in \mathbb{R}^n , then it can be shown that the formula

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \dots + w_n u_n v_n \tag{2}$$

defines an inner product on R^n that we call the weighted Euclidean inner product with weights w_1, w_2, \ldots, w_n .

DEFINITION 3 If V is an inner product space, then the set of points in V that satisfy

$$\|\mathbf{u}\| = 1$$

is called the *unit sphere* or sometimes the *unit circle* in V.

<u>matrix inner products</u>: if $u \cdot v$ is the Euclidean inner product on R^n , then the formula $\langle u, v \rangle = Au \cdot Av$ also defines an inner product. $\langle u, v \rangle = (Av)^T Au = v^T A^T Au$

If u = U and v = V are matrices in the vector space Mnxn, then the formula $\langle u, v \rangle = tr(U^TV)$ defines an inner product on Mnxn called the **standard inner product** on that space

If are polynomials in P_n , then the following formula defines an inner product on P_n (verify) that we will call the *standard inner product* on this space: $\langle p,q \rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n$

EXAMPLE 10 An Integral Inner Product on C[a, b]

Let $\mathbf{f} = f(x)$ and $\mathbf{g} = g(x)$ be two functions in C[a, b] and define

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_{a}^{b} f(x)g(x) dx$$
 (12)

We will show that this formula defines an inner product on C[a, b] by verifying the four inner product axioms for functions f = f(x), g = g(x), and h = h(x) in C[a, b]:

Axiom 1:
$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_a^b f(x)g(x) \, dx = \int_a^b g(x)f(x) \, dx = \langle \mathbf{g}, \mathbf{f} \rangle$$

Axiom 2: $\langle \mathbf{f} + \mathbf{g}, \mathbf{h} \rangle = \int_a^b (f(x) + g(x))h(x) \, dx$

$$= \int_a^b f(x)h(x) \, dx + \int_a^b g(x)h(x) \, dx$$

$$= \langle \mathbf{f}, \mathbf{h} \rangle + \langle \mathbf{g}, \mathbf{h} \rangle$$
Axiom 3: $\langle k\mathbf{f}, \mathbf{g} \rangle = \int_a^b kf(x)g(x) \, dx = k \int_a^b f(x)g(x) \, dx = k \langle \mathbf{f}, \mathbf{g} \rangle$

Axiom 4: If f = f(x) is any function in C[a, b], then

$$\langle \mathbf{f}, \mathbf{f} \rangle = \int_{a}^{b} f^{2}(x) \, dx \ge 0 \tag{13}$$

since $f^2(x) \ge 0$ for all x in the interval [a, b]. Moreover, because f is continuous on [a, b], the equality in Formula (13) holds if and only if the function f is identically zero on [a, b], that is, if and only if f = 0; and this proves that Axiom 4 holds.

THEOREM 6.1.2 If \mathbf{u} , \mathbf{v} , and \mathbf{w} are vectors in a real inner product space V, and if k is a scalar, then:

- (a) $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$
- (b) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
- (c) $\langle \mathbf{u}, \mathbf{v} \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle \langle \mathbf{u}, \mathbf{w} \rangle$
- (d) $\langle \mathbf{u} \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{w} \rangle$
- (e) $k\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, k\mathbf{v} \rangle$

6.2 Angle and Orthogonality in Inner Product Spaces

Recall from Formula (20) of Section 3.2 that the angle θ between two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n is

$$\theta = \cos^{-1}\left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}\right) \tag{1}$$

We were assured that this formula was valid because it followed from the Cauchy-Schwarz inequality (Theorem 3.2.4) that

$$-1 \le \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \le 1 \tag{2}$$

as required for the inverse cosine to be defined. The following generalization of the Cauchy-Schwarz inequality will enable us to define the angle between two vectors in *any* real inner product space.

THEOREM 6.2.1 Cauchy–Schwarz Inequality

If **u** and **v** are vectors in a real inner product space V, then

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \le \|\mathbf{u}\| \|\mathbf{v}\| \tag{3}$$

(Figure 6.2.1). This enables us to define the angle θ between u and v to be

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}\right)$$

THEOREM 6.2.2 If \mathbf{u} , \mathbf{v} , and \mathbf{w} are vectors in a real inner product space V, and if k is any scalar, then:

- (a) $\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$ [Triangle inequality for vectors]
- (b) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ [Triangle inequality for distances]

DEFINITION 1 Two vectors **u** and **v** in an inner product space V called *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

THEOREM 6.2.3 Generalized Theorem of Pythagoras

If u and v are orthogonal vectors in a real inner product space, then

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

DEFINITION 2 If W is a subspace of a real inner product space V, then the set of all vectors in V that are orthogonal to every vector in W is called the *orthogonal* complement of W and is denoted by the symbol W^{\perp} .

THEOREM 6.2.4 If W is a subspace of a real inner product space V, then:

- (a) W[⊥] is a subspace of V.
- (b) $W \cap W^{\perp} = \{0\}.$

EXAMPLE 6 Basis for an Orthogonal Complement

Let W be the subspace of R6 spanned by the vectors

$$\mathbf{w}_1 = (1, 3, -2, 0, 2, 0),$$
 $\mathbf{w}_2 = (2, 6, -5, -2, 4, -3),$ $\mathbf{w}_3 = (0, 0, 5, 10, 0, 15),$ $\mathbf{w}_4 = (2, 6, 0, 8, 4, 18)$

Find a basis for the orthogonal complement of W.

Solution The subspace W is the same as the row space of the matrix

$$A = \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix}$$

Since the row space and null space of A are orthogonal complements, our problem reduces to finding a basis for the null space of this matrix. In Example 4 of Section 4.7 we showed that

$$\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for this null space. Expressing these vectors in comma-delimited form (to match that of w_1 , w_2 , w_3 , and w_4), we obtain the basis vectors

$$\mathbf{v}_1 = (-3, 1, 0, 0, 0, 0), \quad \mathbf{v}_2 = (-4, 0, -2, 1, 0, 0), \quad \mathbf{v}_3 = (-2, 0, 0, 0, 1, 0)$$

You may want to check that these vectors are orthogonal to w_1 , w_2 , w_3 , and w_4 by computing the necessary dot products.

6.3 Gram-Schmidt Process; QR-Decomposition

DEFINITION 1 A set of two or more vectors in a real inner product space is said to be *orthogonal* if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is said to be *orthonormal*.

THEOREM 6.3.1 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space, then S is linearly independent.

THEOREM 6.3.2

(a) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthogonal basis for an inner product space V, and if \mathbf{u} is any vector in V, then

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2 + \dots + \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2} \mathbf{v}_n$$
(3)

(b) If $S = \{v_1, v_2, ..., v_n\}$ is an orthonormal basis for an inner product space V, and if u is any vector in V, then

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \dots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n \tag{4}$$

Using the terminology and notation from Definition 2 of Section 4.4, it follows from Theorem 6.3.2 that the coordinate vector of a vector \mathbf{u} in V relative to an orthogonal basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is

$$(\mathbf{u})_{S} = \left(\frac{\langle \mathbf{u}, \mathbf{v}_{1} \rangle}{\|\mathbf{v}_{1}\|^{2}}, \frac{\langle \mathbf{u}, \mathbf{v}_{2} \rangle}{\|\mathbf{v}_{2}\|^{2}}, \dots, \frac{\langle \mathbf{u}, \mathbf{v}_{n} \rangle}{\|\mathbf{v}_{n}\|^{2}}\right)$$
(6)

and relative to an orthonormal basis $S = \{v_1, v_2, \dots, v_n\}$ is

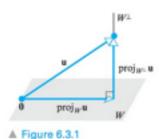
$$(\mathbf{u})_S = (\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \dots, \langle \mathbf{u}, \mathbf{v}_n \rangle)$$
 (7)

THEOREM 6.3.3 Projection Theorem

If W is a finite-dimensional subspace of an inner product space V, then every vector u in V can be expressed in exactly one way as

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \tag{8}$$

where \mathbf{w}_1 is in W and \mathbf{w}_2 is in W^{\perp}.



These are called the orthogonal projection of u on W and the orthogonal projection of u on W^{\perp} , respectively. The vector \mathbf{w}_2 is also called the component of u orthogonal to W. Using the notation in (9), Formula (8) can be expressed as

$$\mathbf{u} = \operatorname{proj}_{\mathbf{w}} \mathbf{u} + \operatorname{proj}_{\mathbf{w}^{\perp}} \mathbf{u}$$
 (10)

(Figure 6.3.1). Moreover, since $proj_{W^{\perp}}\mathbf{u} = \mathbf{u} - proj_{W}\mathbf{u}$, we can also express Formula (10) as

$$\mathbf{u} = \operatorname{proj}_{\mathbf{W}} \mathbf{u} + (\mathbf{u} - \operatorname{proj}_{\mathbf{W}} \mathbf{u}) \tag{11}$$

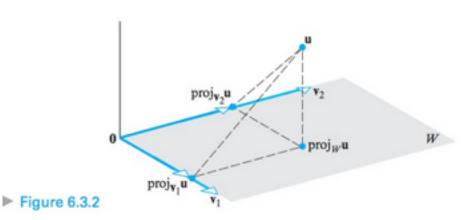
THEOREM 6.3.4 Let W be a finite-dimensional subspace of an inner product space V.

(a) If $\{v_1, v_2, ..., v_r\}$ is an orthogonal basis for W, and u is any vector in V, then

$$\operatorname{proj}_{W} \mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_{1} \rangle}{\|\mathbf{v}_{1}\|^{2}} \mathbf{v}_{1} + \frac{\langle \mathbf{u}, \mathbf{v}_{2} \rangle}{\|\mathbf{v}_{2}\|^{2}} \mathbf{v}_{2} + \dots + \frac{\langle \mathbf{u}, \mathbf{v}_{r} \rangle}{\|\mathbf{v}_{r}\|^{2}} \mathbf{v}_{r}$$
(12)

(b) If $\{v_1, v_2, \dots, v_r\}$ is an orthonormal basis for W, and u is any vector in V, then

$$\operatorname{proj}_{W} \mathbf{u} = \langle \mathbf{u}, \mathbf{v}_{1} \rangle \mathbf{v}_{1} + \langle \mathbf{u}, \mathbf{v}_{2} \rangle \mathbf{v}_{2} + \dots + \langle \mathbf{u}, \mathbf{v}_{r} \rangle \mathbf{v}_{r}$$
(13)



THEOREM 6.3.5 Every nonzero finite-dimensional inner product space has an orthonormal basis.

The Gram-Schmidt Process

To convert a basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$, perform the following computations:

Step 1.
$$\mathbf{v}_1 = \mathbf{u}_1$$

Step 2.
$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1$$

Step 3.
$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2$$

Step 4.
$$\mathbf{v}_4 = \mathbf{u}_4 - \frac{\langle \mathbf{u}_4, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{u}_4, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2 - \frac{\langle \mathbf{u}_4, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2} \mathbf{v}_3$$

:

(continue for r steps)

Optional Step. To convert the orthogonal basis into an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_r\}$, normalize the orthogonal basis vectors.

Assume that the vector space R³ has the Euclidean inner product. Apply the Gram-Schmidt process to transform the basis vectors

$$\mathbf{u}_1 = (1, 1, 1), \quad \mathbf{u}_2 = (0, 1, 1), \quad \mathbf{u}_3 = (0, 0, 1)$$

into an orthogonal basis $\{v_1, v_2, v_3\}$, and then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{q_1, q_2, q_3\}$.

Solution

Step 1.
$$\mathbf{v}_1 = \mathbf{u}_1 = (1, 1, 1)$$

Step 2.
$$\mathbf{v}_2 = \mathbf{u}_2 - \operatorname{proj}_{W_1} \mathbf{u}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1$$

$$= (0, 1, 1) - \frac{2}{3} (1, 1, 1) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3} \right)$$
Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \operatorname{proj}_{W_2} \mathbf{u}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2$

$$= (0, 0, 1) - \frac{1}{3} (1, 1, 1) - \frac{1/3}{2/3} \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3} \right)$$

$$= \left(0, -\frac{1}{2}, \frac{1}{2} \right)$$

Thus,

$$\mathbf{v}_1 = (1, 1, 1), \quad \mathbf{v}_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right), \quad \mathbf{v}_3 = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

THEOREM 6.3.6 If W is a finite-dimensional inner product space, then:

- (a) Every orthogonal set of nonzero vectors in W can be enlarged to an orthogonal basis for W.
- (b) Every orthonormal set in W can be enlarged to an orthonormal basis for W.

THEOREM 6.3.7 QR-Decomposition

If A is an $m \times n$ matrix with linearly independent column vectors, then A can be factored as

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal column vectors, and R is an $n \times n$ invertible upper triangular matrix.

It follows from Theorem 6.3.7 that every invertible matrix has a QR-decomposition.

Find a QR-decomposition of

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Solution The column vectors of A are

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Applying the Gram-Schmidt process with normalization to these column vectors yields the orthonormal vectors (see Example 8)

$$\mathbf{q}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Thus, it follows from Formula (16) that R is

$$R = \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{q}_1 \rangle & \langle \mathbf{u}_2, \mathbf{q}_1 \rangle & \langle \mathbf{u}_3, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{u}_2, \mathbf{q}_2 \rangle & \langle \mathbf{u}_3, \mathbf{q}_2 \rangle \\ 0 & 0 & \langle \mathbf{u}_3, \mathbf{q}_3 \rangle \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

from which it follows that a QR-decomposition of A is

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \blacktriangleleft$$

$$A = 0$$