

Chapter 6 Orthogonality and Least Squares

David Robinson

Inner Product, Length, and Orthogonality

Theorem 1

Let \mathbf{u} , \mathbf{v} , and \mathbf{w} be vectors in \mathbb{R}^n , and let c be a scalar. Then

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
3. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$

The Length of a Vector

The length (or norm) of \mathbf{v} is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$$

Distance in \mathbb{R}^n

For \mathbf{u} and \mathbf{v} in \mathbb{R}^n , the **distance between \mathbf{u} and \mathbf{v}** , written as $\text{dist}(\mathbf{u}, \mathbf{v})$, is the length of the vector $\mathbf{u} - \mathbf{v}$.

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Orthogonal Vectors

Two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = 0$.

Theorem 2 — The Pythagorean Theorem

Two vectors \mathbf{u} and \mathbf{v} are orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

Theorem 3

Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A , and the orthogonal complement of the column space of A is the null space of A^T

$$(\text{Row } A)^\perp = \text{Nul } A \quad \text{and} \quad (\text{Col } A)^\perp = \text{Nul } A^T$$

Key Points

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \times \mathbf{v}$
2. A unit vector in the direction of a vector can be determined by dividing that vector by its length.
3. $\|c\mathbf{v}\|$ is not always equal to $c\|\mathbf{v}\|$. Since length is always positive, the value of $\|c\mathbf{v}\|$ is positive for all values of c . However, $c\|\mathbf{v}\|$ is negative if c is negative.

Orthogonal Sets

Theorem 4

If $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in \mathbb{R}^n , then S is linearly independent and hence is a basis for the subspace spanned by S .

Orthogonal Basis

An **orthogonal basis** for a subspace W of \mathbb{R}^n is a basis for W that is also an orthogonal set.

Theorem 5

Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of \mathbb{R}^n . For each \mathbf{y} in W , the weights in the linear combination are

$$\mathbf{y} = c_1\mathbf{u}_1 + \dots + c_p\mathbf{u}_p \quad \text{given by} \quad c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \quad (j = 1, \dots, p)$$

Theorem 6

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

Theorem 7

Let U be an $m \times n$ matrix with orthonormal columns, and let \mathbf{x} and \mathbf{y} be in \mathbb{R}^n . Then

1. $\|U\mathbf{x}\| = \|\mathbf{x}\|$
2. $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$
3. $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$

Key Points

1. A set of vectors is orthogonal if each pair of distinct vectors from the set is orthogonal.
2. The vector $\hat{\mathbf{y}}$ is the orthogonal projection of \mathbf{y} onto \mathbf{u} .

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \right) \mathbf{u}$$

3. \mathbf{y} can be written as the sum of a vector in $\text{Span}\{\mathbf{u}\}$ and a vector orthogonal to \mathbf{u} .

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

4. An orthonormal set is an orthogonal set where all of the vectors are unit vectors.
5. If A is a matrix with orthonormal columns, then $\|A\mathbf{x}\| = \|\mathbf{x}\|$.
6. If U is an orthogonal matrix, $U^T = U^{-1}$.
7. A matrix with orthogonal columns is an orthonormal matrix if the matrix is also square.
8. If U is a $m \times n$ matrix with orthonormal columns, \mathbf{x} is in \mathbb{R}^n , $\|U\mathbf{x}\| = (U\mathbf{x})^T(U\mathbf{x})$

Orthogonal Projections

Theorem 8 — The Orthogonal Decomposition Theorem

Let W be a subspace of \mathbb{R}^n . Then each \mathbf{y} in \mathbb{R}^n can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and \mathbf{z} is in W^\perp . In fact, if $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is any orthogonal basis of W , then

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \dots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p} \mathbf{u}_p$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$

Theorem 9 — The Best Approximation Theorem

Let W be a subspace of \mathbb{R}^n , let \mathbf{y} be any vector in \mathbb{R}^n , and let $\hat{\mathbf{y}}$ be the orthogonal projection of \mathbf{y} onto W . Then $\hat{\mathbf{y}}$ is the closest point in W to \mathbf{y} , in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for all \mathbf{v} in W distinct from $\hat{\mathbf{y}}$.

Theorem 10

If $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of \mathbb{R}^n , then

$$\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2 + \dots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p$$

If $U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_p]$, then

$$\text{proj}_W \mathbf{y} = UU^T \mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n$$

Key Points

1. If \mathbf{u}_1 and \mathbf{u}_2 are orthogonal but \mathbf{u}_3 is not orthogonal to \mathbf{u}_1 or \mathbf{u}_2 , a nonzero vector \mathbf{v} in \mathbb{R}^3 that is orthogonal to \mathbf{u}_1 and \mathbf{u}_2 can be constructed through $\mathbf{v} = \hat{\mathbf{u}}_3 - \mathbf{u}_3$.

The Gram-Schmidt Process**Theorem 11 — The Gram-Schmidt Process**

Given a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ for a nonzero subspace W of \mathbb{R}^n , define

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{x}_1 \\ \mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 \\ \mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 \\ &\vdots \\ \mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \dots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}} \mathbf{v}_{p-1} \end{aligned}$$

Then $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is an orthogonal basis for W . In addition

$$\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \quad \text{for } 1 \leq k \leq p$$

Theorem 12 — The QR Factorization

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for $\text{Col } A$ and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.