

Orthonormal sets

Definition. A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is orthonormal if it is orthogonal and $\mathbf{v}_i \cdot \mathbf{v}_i = 1$ for all i .

Example. Consider the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}.$$

We can use matrices to express the fact that a set is orthonormal.

Theorem. Let \mathbf{A} be an $n \times n$ matrix. The following three conditions are equivalent.

1. $\mathbf{A}^T = \mathbf{A}^{-1}$
2. The columns of \mathbf{A} form an orthonormal basis of \mathbb{R}^n .
3. The rows of \mathbf{A} form an orthonormal basis of \mathbb{R}^n .

Definition. Whenever a matrix satisfies the above theorem, it is said to be an orthogonal matrix.

Example. We can use the orthonormal basis of \mathbb{R}^3 that we derived earlier to produce an orthogonal matrix.

Why are orthogonal matrices special?

Orthogonal projection

How do we project a vector \mathbf{v} onto a subspace W ?

Theorem. (Orthogonal Decomposition Theorem)

1. Each vector \mathbf{v} in \mathbb{R}^n can be written uniquely as

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp,$$

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

2. Given an orthogonal basis $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ of W , then

$$\mathbf{w} = \left(\frac{\mathbf{v} \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1} \right) \mathbf{w}_1 + \dots + \left(\frac{\mathbf{v} \cdot \mathbf{w}_k}{\mathbf{w}_k \cdot \mathbf{w}_k} \right) \mathbf{w}_k$$

and $\mathbf{w}^\perp = \mathbf{v} - \mathbf{w}$.

Important consequence: If we want to find the distance of a vector \mathbf{v} to a subspace W , then we compute

$$\|\mathbf{w}^\perp\| = \|\mathbf{v} - \mathbf{w}\|.$$

Example. Find the point closest to

$$\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix}$$

in the subspace W spanned by the two vectors

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix}.$$

Theorem. If $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthonormal basis for a subspace W , then

$$\mathbf{w} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + \dots + (\mathbf{v} \cdot \mathbf{u}_k)\mathbf{u}_k.$$

If

$$\mathbf{U} = \left[\begin{array}{c|c|c|c} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_k \end{array} \right],$$

then $\mathbf{w} = \mathbf{U}\mathbf{U}^T\mathbf{v}$.