1 Overview

Main ideas:

1. definition of orthogonal set; thm: linear independence of orthogonal set; definition of an orthogonal basis, orthogonal decomposition of a vector, definition of orthogonal projection
2. orthonormal set, orthonormal basis
3. matrices whose columns are orthonormal sets, orthogonal matrices, isometries!

Examples in text:

1. show that a given set of three vectors in $\mathbb{R}^3$ is an orthogonal set
2. orthogonal decomposition of a vector in $\mathbb{R}^3$
3. orthogonal projection of a vector in $\mathbb{R}^2$ onto another vector in $\mathbb{R}^2$
4. distance from a point to a line
5. verify that a given set is an orthonormal basis of $\mathbb{R}^3$
6. verify that a given matrix with orthonormal columns perserves lengths of vectors
7. an orthogonal matrix

2 Discussion and Worked Examples

2.1 Orthogonality, Linear Independence, and Decompositions

If each pair of vectors in a set of vectors is orthogonal, then the set of vectors is said to be *orthogonal.*

Example Determine whether the following set of vectors is orthogonal.

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix} \right\}$$

Let v_1, v_2, and v_3 denote the vectors listed above, in that order. Then we can compute

$$\begin{aligned} v_1 \cdot v_2 &= 0 - 2 + 2 = 0 \\ v_1 \cdot v_3 &= -5 + 4 + 1 = 0 \\ v_2 \cdot v_3 &= 0 - 2 + 2 = 0 \end{aligned}$$

Thus $\{v_1, v_2, v_3\}$ is an orthogonal set.

Theorem. *An orthogonal set of nonzero vectors is linearly independent.*

Proof. Suppose $\{v_1, \ldots, v_n\}$ is orthogonal, and $c_1, \ldots c_n$ are weights such that

$$c_1 v_1 + \ldots + c_n v_n = 0$$

Taking the dot product of both sides with v_1 and using bilinearity yields

$$c_1 \langle v_1, v_1 \rangle + c_2 \langle v_2, v_1 \rangle + \ldots + c_n \langle v_n, v_1 \rangle = \langle 0, v_1 \rangle = 0$$

Since v_1 is orthogonal to all of the other vectors in the given set, this reduces to

$$c_1 \|v_1\|^2 = 0$$

Since $v_1 \neq 0$, $c_1 = 0$. Similarly $c_i \|v_i\|^2 = 0$ and thus $c_i = 0$ for all $1 \leq i \leq n$. Thus $\{v_1, \ldots, v_n\}$ is linearly independent. ☐

Of course, not every linearly independent set is orthogonal; an orthogonal set is a special kind of linearly independent set. If a basis for $\mathbb{R}^n$ is orthogonal it is called an *orthogonal basis.* Orthogonal bases are very convenient, because there is a nice formula for expressing a vector as a linear combination of orthogonal vectors (i.e. finding coordinates relative to an orthogonal basis).

Suppose $\{w_1, \ldots, w_n\}$ is an orthogonal basis for a subspace W of $\mathbb{R}^n$. We wish to write w in W in terms of this basis.

$$w = c_1 w_1 + \ldots + c_n w_n$$

Taking the dot product of both sides with w_1, yields

$$\langle w, w_1 \rangle = c_1 \langle w_1, w_1 \rangle + 0$$

Thus, $c_1 = \langle w, w_1 \rangle / \langle w_1, w_1 \rangle$. We can find the other weights in the same way. Thus, we have derived the following result.

Theorem. *Let $\{w_1, \ldots, w_n\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. Let w be a vector in W. Then,*

$$w = \frac{\langle w, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \ldots + \frac{\langle w, w_n \rangle}{\langle w_n, w_n \rangle} w_n$$

We say that this is a *decomposition* of w with respect to the basis $\{w_1, \ldots, w_n\}$. The terms are called *components* of w with respect to $\{w_1, \ldots, w_n\}$ and the weights $\langle w, w_i \rangle / \langle w_i, w_i \rangle$ are called *decomposition coefficients.* Notice that the components are precisely the projections of w onto the basis vectors. If we normalize the basis (i.e. scale each vector in the basis by the reciprocal of its length), the decomposition is even simpler:

$$w = \langle w, \hat{w}_1 \rangle \hat{w}_1 + \ldots + \langle w, \hat{w}_n \rangle \hat{w}_n$$

This is why a normalized orthogonal basis, called an *orthonormal* basis, is convenient. The standard basis for $\mathbb{R}^n$ is an orthonormal basis.

Example Write the vector $v = (1, 1, 1)$ as a linear combination of the vectors in the previous example.

The set of vectors in the previous example is an orthogonal basis for $\mathbb{R}^3$. (The fact that the set is orthogonal implies that it is linearly independent, and any set of three linearly independent vectors in $\mathbb{R}^3$ spans $\mathbb{R}^3$ and is a basis for $\mathbb{R}^3$.) Thus there are weights such that $v = c_1 v_1 + c_2 v_2 + c_3 v_3$. By the orthogonal decomposition theorem, we can compute the weights using ratios of dot products:

$$c_1 = \frac{v \cdot v_1}{v_1 \cdot v_1} = \frac{1 - 2 + 1}{1 + 4 + 1} = 0$$

$$c_2 = \frac{v \cdot v_2}{v_2 \cdot v_2} = \frac{0 + 1 + 2}{0 + 1 + 5} = \frac{3}{5}$$

$$c_3 = \frac{v \cdot v_3}{v_3 \cdot v_3} = \frac{-5 - 2 + 1}{25 + 4 + 1} = -\frac{1}{5}$$

Thus,

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \frac{3}{5} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} - \frac{1}{5} \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$$

Example Normalize the vectors v_1, v_2, and v_3 to obtain an orthonormal basis for $\mathbb{R}^3$. Then decompose $v = (1, 1, 1)$ and $w = (1, 0, 2)$ with respect to the orthonormal basis.

To normalize the vectors, we simply scale by the reciprocals of the lengths:

$$\hat{v}_1 = \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \qquad \hat{v}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \qquad \hat{v}_3 = \begin{bmatrix} -5/\sqrt{30} \\ -2/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}$$

The decompositions coefficients for v with respect to the orthonormal basis are $\langle v, \hat{v}_i \rangle$, so

$$v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0 \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} + \frac{3}{\sqrt{5}} \begin{bmatrix} 0 \\ 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} - \frac{6}{\sqrt{30}} \begin{bmatrix} -5/\sqrt{30} \\ -2/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}$$

Similarly,

$$w = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \frac{3}{\sqrt{6}} \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} + \frac{4}{\sqrt{5}} \begin{bmatrix} 0 \\ 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} - \frac{3}{\sqrt{30}} \begin{bmatrix} -5/\sqrt{30} \\ -2/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}$$

Example Compute the dot product of the vectors v and w from the previous example in two ways: first using their standard coordinates and second using their decompositions in terms of the orthonormal basis in the previous example.

Using standard coordinates:

$$\langle v, w \rangle = (1, 1, 1) \cdot (1, 0, 2) = 1 + 0 + 2 = 3$$

Using the orthogonal decompositions with respect to $\{v_1, v_2, v_3\}$,

$$\langle v, w \rangle = \left(0\,\hat{v}_1 + (3/\sqrt{5})\hat{v}_2 + (-6/\sqrt{30})\hat{v}_3\right) \cdot \left((3/\sqrt{6})\hat{v}_1 + (4/\sqrt{5})\hat{v}_2 + (-3/\sqrt{30})\hat{v}_3\right)$$

We take advantage of the fact that $\hat{v}_1 \cdot \hat{v}_2 = 0$, $\hat{v}_1 \cdot \hat{v}_3 = 0$, and $\hat{v}_2 \cdot \hat{v}_3 = 0$ (so $\hat{v}_i \cdot \hat{v}_j = 0$ when $i \neq j$) and $\hat{v}_i \cdot \hat{v}_i = 1$. Thus,

$$\langle v, w \rangle = 0 + 12/5 + 18/30 = 3$$

2.2 Orthogonal Matrices

Since matrix multiplication can be computed by taking dot products of rows with columns, one might suppose that something interesting will happen when we take a matrix whose columns are an orthogonal set and multiply on the left by its transpose.

Let A be the matrix whose columns are the vectors v_1, v_2, v_3 from above. Multiplying on the left by the transpose yields:

$$A^T A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 2 \\ -5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -5 \\ -2 & 1 & -2 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 30 \end{bmatrix}$$

This is the diagonal matrix whose entries are the norms-squared of the vectors v_1, v_2 and v_3. This is because when we perform the matrix multiplication we are taking the dot products of pairs of vectors in the set $\{v_1, v_2, v_3\}$.

If the columns of a matrix are an orthonormal set, the norm-squared of each column (considered as a vector) is one. For example, let U be the matrix whose columns are $\hat{v}_1$, $\hat{v}_2$, and $\hat{v}_3$, from the examples above. Then

$$U^T U = \begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ -5/\sqrt{30} & -2/\sqrt{30} & 1/\sqrt{30} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 0 & -5/\sqrt{30} \\ -2/\sqrt{6} & 1/\sqrt{5} & -2/\sqrt{30} \\ 1/\sqrt{6} & 2/\sqrt{5} & 1/\sqrt{30} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus if U is a matrix whose columns are an orthonormal set, $U^T U = \mathbb{I}$. The converse is also true, since the matrix product $U^T U$ being equal to the identity matrix means that: $\langle u_i, u_i \rangle = 1$ and $\langle u_i, u_j \rangle = 0$ if $i \neq j$. This is summarized in the following theorem.

Theorem. *A matrix has orthonormal columns if and only if its transpose is its inverse*

A square matrix whose columns are an orthonormal basis for $\mathbb{R}^n$ is called an *orthogonal matrix*. Non-square matrices whose columns are an orthonormal set also have some nice properties, but there is no particular term for such matrices.

(Recall that matrices whose transposes are their inverses play a role in the singular value decomposition discussed in the exercises of Section 2.5.)

Recall that matrix-vector multiplication can be interpreted as taking a linear combination of the columns of the matrix, with the entries of the vectors as weights. Thus we may interpret the orthogonal decompositions of the vectors v and w in the example above as matrix vector products:

$$v = \begin{bmatrix}1\\1\\1\end{bmatrix} = 0\begin{bmatrix}1/\sqrt{6}\\-2/\sqrt{6}\\1/\sqrt{6}\end{bmatrix} + \frac{3}{\sqrt{5}}\begin{bmatrix}0\\1/\sqrt{5}\\2/\sqrt{5}\end{bmatrix} - \frac{6}{\sqrt{30}}\begin{bmatrix}-5/\sqrt{30}\\-2/\sqrt{30}\\1/\sqrt{30}\end{bmatrix} = \begin{bmatrix}1/\sqrt{6} & 0 & -5/\sqrt{30}\\-2/\sqrt{6} & 1/\sqrt{5} & -2/\sqrt{30}\\1/\sqrt{6} & 2/\sqrt{5} & 1/\sqrt{30}\end{bmatrix}\begin{bmatrix}0\\3/\sqrt{5}\\-6/\sqrt{30}\end{bmatrix}$$

$$w = \begin{bmatrix}1\\0\\2\end{bmatrix} = \frac{3}{\sqrt{6}}\begin{bmatrix}1/\sqrt{6}\\-2/\sqrt{6}\\1/\sqrt{6}\end{bmatrix} + \frac{4}{\sqrt{5}}\begin{bmatrix}0\\1/\sqrt{5}\\2/\sqrt{5}\end{bmatrix} - \frac{3}{\sqrt{30}}\begin{bmatrix}-5/\sqrt{30}\\-2/\sqrt{30}\\1/\sqrt{30}\end{bmatrix} = \begin{bmatrix}1/\sqrt{6} & 0 & -5/\sqrt{30}\\-2/\sqrt{6} & 1/\sqrt{5} & -2/\sqrt{30}\\1/\sqrt{6} & 2/\sqrt{5} & 1/\sqrt{30}\end{bmatrix}\begin{bmatrix}3/\sqrt{6}\\4/\sqrt{5}\\-3/\sqrt{30}\end{bmatrix}$$

Letting $x = (c_1, c_2, c_3)$ be the vector whose entries are the weights for the orthogonal decomposition of v with respect to $\{v_1, v_2, v_3\}$ and letting $y = (d_1, d_2, d_3)$ be the vector whose entries are the weights for the orthogonal decomposition of w, we can rewrite the above observations as

$$v = Ux \quad \text{and} \quad w = Uy$$

Thus, as computed explicitly above,

$$Ux \cdot Uy = v \cdot w = \left(c_1\hat{v}_1 + c_2\hat{v}_2 + c_3\hat{v}_3\right) \cdot \left(d_1\hat{v}_1 + d_2\hat{v}_2 + d_3\hat{v}_3\right) = c_1d_2 + c_2d_2 + c_3d_3 = x \cdot y$$

It turns out that this is true generally: when the columns of an $m \times n$ matrix U are an orthonormal set, then $Ux \cdot Uy = x \cdot y$ for all vectors x and y in $\mathbb{R}^n$. Since the dot product is the key to lengths, distances, and orthogonality, the fact that the map $x \mapsto Ux$ preserves the dot product implies that it also preserves lengths, distances, and orthogonality.

Theorem. *Let U be an $m \times n$ matrix with orthonormal columns. Then the linear transformation $x \mapsto Ux$ preserves the inner product in the following sense: $Ux \cdot Uy = x \cdot y$ for all x and y in $\mathbb{R}^n$. Thus the linear transformation preserves length, distance, and orthogonality, i.e.*

1. $\|Ux\| = \|x\|$ *i.e. the length of x is the same as the length of its image Ux*
2. $\|Ux - Uy\| = \|x - y\|$, *i.e. the distance between two vectors x and y is the same as the distance between their images Ux and Uy*
3. $Ux \cdot Uy = 0 \Leftrightarrow x \cdot y = 0$, *i.e. the vectors x and y are orthogonal if and only if their images Ux and Uy are orthogonal.*

For example, if we consider linear transformations of the plane, the linear transformations that preserve the dot product are precisely the linear transformations that are rigid motions, i.e. rotations and reflections. Since dilations, expansions, shears, and projections "stretch," "shrink," or "collapse" the plane they do not preserve distances or angles.