

1 Inner Product and Orthogonality

Definition 1: The *inner product* of two vectors \mathbf{x} and \mathbf{y} ,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$$

is denoted $\langle \mathbf{x}, \mathbf{y} \rangle$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \dots x_n y_n$$

Note that

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x}$$

Definition 2: Two vectors \mathbf{x} and \mathbf{y} are *orthogonal* if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Example 1: Are the following vectors orthogonal?

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Draw them on the real plane.

Answer:

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = (1)(1) + (1)(-1) = 0$$

Yes, they are orthogonal.

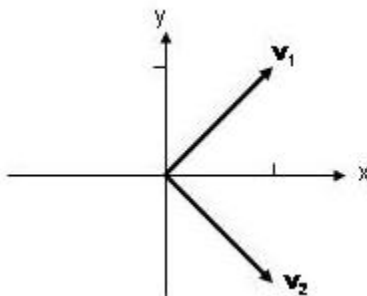


Figure 1: Orthogonal Vectors

We can say a larger set of vectors is orthogonal as well, if each vector is orthogonal to all others.

Definition 3: A set of vectors $\{\mathbf{v}_i\}_{i=1}^m$ is *orthogonal* if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \forall i \neq j$.

Example 3: Find a non-zero vector that is orthogonal to both $(1, 1, 1)$ and $(1, -1, 0)$. Are there many such vectors, or only one?

Answer: Suppose the vector is (x, y, z) . Then

$$x(1) + y(1) + z(1) = 0$$

and

$$x(1) + y(-1) + z(0) = 0$$

That is, there are two equations in the three unknowns:

$$x + y + z = 0$$

and

$$x - y = 0$$

Gaussian elimination produces:

$$\begin{array}{rcl} x & +y & +z = 0 \\ & -2y & -z = 0 \end{array}$$

And there are several solutions: $z, y = -\frac{z}{2}, x = \frac{-z}{2}$. One solution is $(-1, -1, 2)$. There are several such *non-zero* vectors.

2 Linear Independence and Orthogonality

Orthogonality implies linear independence if all vectors are non-zero. Note that the converse is not true; Example 3 shows linearly independent vectors that are not orthogonal.

Theorem If a set of non-zero vectors $\{\mathbf{v}_i\}_{i=1}^m$ is orthogonal, it is linearly independent.

Proof: To show linear independence, we first suppose

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots c_n\mathbf{v}_n = \mathbf{0}$$

(now we need to show that this implies that all c_i are zero).

Take the inner product of both sides with any of the \mathbf{v}_i , say \mathbf{v}_1 :

$$\mathbf{v}_1^T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots c_n\mathbf{v}_n) = \mathbf{v}_1^T(\mathbf{0}) = 0$$

$$\Rightarrow c_1\mathbf{v}_1^T\mathbf{v}_1 + c_2\mathbf{v}_1^T\mathbf{v}_2 + \dots c_n\mathbf{v}_1^T\mathbf{v}_n = 0$$

As $\{\mathbf{v}_i\}_{i=1}^m$ are orthogonal,

$$\mathbf{v}_1^T\mathbf{v}_i = 0 \quad i \neq 1$$

and

$$c_1\mathbf{v}_1^T\mathbf{v}_1 = 0$$

As the vectors $\{\mathbf{v}_i\}_{i=1}^m$ are non-zero,

$$\mathbf{v}_1^T\mathbf{v}_1 \neq 0 \Rightarrow c_1 = 0$$

In this manner, by taking inner products of the first equation with any vector \mathbf{v}_i , we can show that $c_i = 0$. Thus,

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots c_n\mathbf{v}_n = \mathbf{0} \Rightarrow c_1 = c_2 = \dots c_n = 0$$

Example 3: Are the following vectors linearly independent? Are they orthogonal?

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

Answer: Yes, they are linearly independent because:

$$\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

gives $a_1 = a_2 = a_3 = 0$ as the only solution.

No, they are not orthogonal: $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = -2$, $\langle \mathbf{v}_1, \mathbf{v}_3 \rangle = 0$, $\langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$

Thus Example 3 provides an example of vectors that are linearly independent but not orthogonal.

3 Norms and Orthonormality

Definition 4: The norm of the vector \mathbf{x} is:

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{x_1^2 + x_2^2 + \dots + x_i^2 + \dots + x_n^2}$$

Observe that

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

An orthonormal set of vectors is orthogonal, and, additionally, each vector has a norm of one.

Definition 5: A set of vectors $\{\mathbf{v}_i\}_{i=1}^m$ is *orthonormal* if it is orthogonal, and, additionally, $\|\mathbf{v}_i\| = 1 \forall i$.

In particular, one can have orthonormal bases that are very useful.

Example 4: Are the vectors in Example 1 orthonormal? Is there a simple way to make them orthonormal?

Answer: No, they are not orthonormal, because $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \sqrt{2}$. Yes, one can make them orthonormal by dividing each vector by $\sqrt{2}$. So,

$$\mathbf{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

is an orthonormal set.

Example 5: Consider the matrix \mathbf{U} , whose columns are orthonormal. Show that $\mathbf{U}^{-1} = \mathbf{U}^T$.

Answer: Let \mathbf{u}_i denote the i^{th} column of \mathbf{U} . The $(i, j)^{\text{th}}$ element of the matrix product $\mathbf{U}^T \mathbf{U}$ is $\langle \mathbf{u}_i, \mathbf{u}_j \rangle$. From the definition of orthonormality, $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ $i \neq j$ and $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 1$ $i = j$. Hence only the diagonal entries of $\mathbf{U}^T \mathbf{U}$ are non-zero, and their values are 1. Hence $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ and $\mathbf{U}^{-1} = \mathbf{U}^T$.

4 Coefficients wrt an orthonormal Basis

Suppose $\{\mathbf{v}_i\}_{i=1}^n$ is an orthonormal basis of \mathbb{R}^n . Then the coefficients a_i of \mathbf{x} with respect to (wrt) this basis are simply $a_i = \langle \mathbf{x}, \mathbf{v}_i \rangle$.

To see this, let

$$\mathbf{x} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_n \mathbf{v}_n$$

and

$$\langle \mathbf{v}_1, \mathbf{x} \rangle = \mathbf{v}_1^T \mathbf{x} = a_1 \mathbf{v}_1^T \mathbf{v}_1 + a_2 \mathbf{v}_1^T \mathbf{v}_2 + \dots + a_n \mathbf{v}_1^T \mathbf{v}_n$$

As $\{\mathbf{v}_i\}_{i=1}^n$ is orthonormal,

$$\langle \mathbf{v}_1, \mathbf{x} \rangle = \mathbf{v}_1^T \mathbf{x} = a_1 \mathbf{v}_1^T \mathbf{v}_1 + 0 + \dots + 0 = a_1(1)$$

Similarly, one can show that

$$\langle \mathbf{v}_i, \mathbf{x} \rangle = a_i \quad \forall i$$

Example 6 What are the coefficients of $\mathbf{x} = (1, 0, 1, 1)$ with respect to orthonormal basis

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix},$$

(A set of multiples of these vectors was shown to be a basis of \mathbb{R}^4 in the notes on Vector Spaces, Example 11. It is easy to check that these are orthonormal). The coefficients of \mathbf{x} wrt this basis are easily found using inner products: $a_1 \langle \mathbf{v}_1, \mathbf{x} \rangle = \mathbf{v}_1^T \mathbf{x} = \frac{3}{2}$, and, similarly, $a_2 = \frac{1}{2}$, $a_3 = -\frac{1}{2}$, $a_4 = -\frac{1}{2}$. You may check that these are correct.

5 Orthonormal Eigenvectors as a Basis

It can be shown that eigenvectors corresponding to distinct eigenvalues are linearly independent. Eigenvectors corresponding to the same eigenvalues can be made into a linearly independent set. So, it is possible to use eigenvectors of a matrix of size $n \times n$ to form a basis of \mathbb{R}^n . Why might this be useful?

Well, it makes operations with the matrix particularly easy. Suppose \mathbf{x} is represented as a linear combination of the eigenvectors of matrix \mathbf{A} , denoted \mathbf{v}_i (eigenvalue of \mathbf{v}_i is λ_i):

$$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots a_n\mathbf{v}_n$$

Then,

$$\mathbf{Ax} = \mathbf{A}(a_1\mathbf{v}_1) + \mathbf{A}(a_2\mathbf{v}_2) + \dots \mathbf{A}(a_n\mathbf{v}_n) = a_1\mathbf{A}\mathbf{v}_1 + a_2\mathbf{A}\mathbf{v}_2 \dots a_n\mathbf{A}\mathbf{v}_n$$

because the \mathbf{v}_i are eigenvectors of \mathbf{A} ,

$$\mathbf{Ax} = a_1\lambda_1\mathbf{v}_1 + a_2\lambda_2\mathbf{v}_2 \dots a_n\lambda_n\mathbf{v}_n$$

Additionally, if the eigenvectors are orthonormal, finding the coefficients is very easy.

Example 7: Previously, in class, we have shown that the eigenvectors of matrix:

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

are $(1, 1)$ and $(1, -1)$ corresponding to eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$ respectively. Example 1 shows that these are orthogonal, and Example 4 provides a corresponding orthonormal set:

$$\mathbf{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

What are the coefficients of $\mathbf{x} = (3, 2)$ wrt to basis $\{\mathbf{v}_3, \mathbf{v}_4\}$? Using these values, determine \mathbf{Ax} .

Answer: Because the basis is orthonormal, inner products provide the coefficients: $a_1 = \frac{5}{\sqrt{2}}$, $a_2 = \frac{1}{\sqrt{2}}$. Thus,

$$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2$$

and:

$$\mathbf{Ax} = a_1\mathbf{A}\mathbf{v}_1 + a_2\mathbf{A}\mathbf{v}_2 = a_1\lambda_1\mathbf{v}_1 + a_2\lambda_2\mathbf{v}_2 = a_1\mathbf{v}_1 - a_2\lambda_2\mathbf{v}_2$$

Substituting for the values gives:

$$= (2, 3)$$

And this can be checked to be correct.

Exercises

1. Find the norms and the inner product of $x = (1, 4, 0, 2)$ and $y = (2, -2, 1, 3)$.
2. According to analytic geometry, two lines are orthogonal when the product of their slopes is -1 . Consider the line joining the origin to point (x_1, y_1) . What is its slope? Consider the line joining the origin to another point (x_2, y_2) . What is its slope? When are these two lines orthogonal (that is, when is the product of the slopes equal to -1)?

Now consider these two lines as vectors. Using the definition in this section on vector orthogonality, when are they orthogonal? Is this condition identical to that obtained by considering the vectors as lines through the origin?

3. For invertible matrix \mathbf{X} , is the i^{th} row of \mathbf{X}^{-1} orthogonal to the j^{th} column of \mathbf{X} ? Is the i^{th} column of \mathbf{X}^{-1} orthogonal to the j^{th} row of \mathbf{X} ?

4. Which pairs are orthogonal among the vectors:

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

5. In \mathbb{R}^3 find all vectors that are orthogonal to $(1, 1, 1)$ and $(1, -1, 0)$.

6. Find a non-zero vector orthogonal to all rows of

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{bmatrix}$$

Find another non-zero vector orthogonal to all columns of the above matrix.

7. Find all vectors in \mathbb{R}^4 that are orthogonal to $(1, 4, 4, 1)$ and $(2, 9, 8, 2)$.

Exercise Solutions

1. Find the norms and the inner product of $\mathbf{x} = (1, 4, 0, 2)$ and $\mathbf{y} = (2, -2, 1, 3)$.

Answer:

$$\|\mathbf{x}\| = \sqrt{21}; \|\mathbf{y}\| = \sqrt{18} = 3\sqrt{2}$$

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0$$

2. According to analytic geometry, two lines are orthogonal when the product of their slopes is -1 . Consider the line joining the origin to point (x_1, y_1) . What is its slope? Consider the line joining the origin to another point (x_2, y_2) . What is its slope? When are these two lines orthogonal (that is, when is the product of the slopes equal to -1)?

Now consider these two lines as vectors. Using the definition in this section on vector orthogonality, when are they orthogonal? Is this condition identical to that obtained by considering the vectors as lines through the origin?

Answer: The slope of the line joining (x_1, y_1) to $(0, 0)$ is $\frac{y_1}{x_1}$. That of the line joining (x_2, y_2) to $(0, 0)$ is $\frac{y_2}{x_2}$. The two lines are orthogonal when

$$\frac{y_1}{x_1} \times \frac{y_2}{x_2} = -1 \text{ OR } x_1x_2 + y_1y_2 = 0$$

Vector orthogonality gives the same result.

3. For invertible matrix \mathbf{X} , is the i^{th} row of \mathbf{X}^{-1} orthogonal to the j^{th} column of \mathbf{X} ? Is the i^{th} column of \mathbf{X}^{-1} orthogonal to the j^{th} row of \mathbf{X} ?

Answer: The answer to both questions is *yes*.

Because $\mathbf{X}^{-1}\mathbf{X} = \mathbf{I}$, the i^{th} row of \mathbf{X}^{-1} times the j^{th} column of \mathbf{X} gives the ij^{th} element of \mathbf{I} , which is zero unless $i = j$. That is, the inner product of the i^{th} row of \mathbf{X}^{-1} and the j^{th} column of \mathbf{X} is zero, or they are orthogonal.

Similarly, because $\mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$, the j^{th} row of \mathbf{X} times the i^{th} column of \mathbf{X}^{-1} gives the ji^{th} element of \mathbf{I} , which is zero unless $i = j$. This implies that the j^{th} row of \mathbf{X} and the i^{th} column of \mathbf{X}^{-1} are orthogonal too.

4. Which pairs are orthogonal among the vectors:

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

Answer:

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = -4; \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = 0; \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0;$$

\mathbf{v}_1 and \mathbf{v}_3 are orthogonal, as are \mathbf{v}_2 and \mathbf{v}_3 .

5. Find a non-zero vector orthogonal to all rows of

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{bmatrix}$$

Find another non-zero vector orthogonal to all columns of the above matrix.

Answer: Consider a vector (x, y, z) . Orthogonality to the rows of the above matrix give:

$$\begin{aligned}x + 2y + z &= 0 \\2x + 4y + 3z &= 0 \\3x + 6y + 4z &= 0\end{aligned}$$

Gaussian elimination gives:

$$\begin{aligned}x + 2y + z &= 0 \\z &= 0 \\z &= 0\end{aligned}$$

and any vector with $z = 0$ and $x = -2y$ will be orthogonal to all rows of the given matrix. One such vector is $(-2, 1, 0)$.

Orthogonality to the columns of the above matrix give:

$$\begin{aligned}x + 2y + 3z &= 0 \\2x + 4y + 6z &= 0 \\x + 3y + 4z &= 0\end{aligned}$$

Gaussian elimination gives:

$$\begin{aligned}x + 2y + 3z &= 0 \\0z &= 0 \\y + z &= 0\end{aligned}$$

and any vector with $y = -z$ and $x = -z$ will be orthogonal to all rows of the given matrix. One such vector is $(-1, -1, 1)$.

6. Find all vectors in \mathbb{R}^4 that are orthogonal to $(1, 4, 4, 1)$ and $(2, 9, 8, 2)$.

Answer: Consider a vector (w, x, y, z) . The conditions give:

$$\begin{aligned}w + 4x + 4y + z &= 0 \\2w + 9x + 8y + 2z &= 0\end{aligned}$$

Gaussian elimination gives:

$$\begin{aligned}w + 4x + 4y + z &= 0 \\x &= 0\end{aligned}$$

and vectors with $x = 0$ and $w + 4y + z = 0$ represent all the required vectors.