9 Orthogonality and Projections

In this section we discuss how to test if two vectors are orthogonal and how to construct vectors that are orthogonal.

9.1 Dot Products and Orthogonality

*The **dot product** (or **inner product**) of two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is denoted by $\boldsymbol{u} \cdot \boldsymbol{v}$ and defined as the sum of the product of corresponding entries: that is*

$$\boldsymbol{u} \cdot \boldsymbol{v} = \sum_i u_i v_i.$$

Example. *For example,*

$$\begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 4 \\ 7 \end{bmatrix} = 12 - 4 - 14 = -6.$$

If we view the two vectors as matrices, then the dot product $\boldsymbol{u} \cdot \boldsymbol{v}$ is the entry in the 1×1 matrix given by $\boldsymbol{u}^T \boldsymbol{v}$.

There are some immediate properties of the dot product. These include the following:

Fact 9.1 *(Commutative law)* $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$
(Distributive law) $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}$

*Two vectors are **orthogonal** if their dot product is zero. Orthogonal vectors are sometimes called **perpendicular** vectors.*

Example. *Find vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ that are orthogonal to each other and to $\boldsymbol{w} = (0, 1, 1)$. There are systematic ways to do this. But one way to proceed is to note that the vectors orthogonal to $(0, 1, 1)$ form the null space of the matrix $\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$. Using that, or just observations, we can see that one vector orthogonal to $\boldsymbol{w}$ is $(1, 0, 0)$. Another vector is to take $(0, 1, -1)$.*

*The **length** (or **norm**) of vector $\boldsymbol{v}$ is defined as $||\boldsymbol{v}|| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$. A **unit vector** has length 1.*

ALGOR *To obtain a unit vector in the same direction, divide by the length (called **normalization**).*

EXAMPLE. *For example, the vector $(3, -1, 2)$ has norm $\sqrt{14}$; a unit vector in the same direction as it is $\frac{1}{\sqrt{14}}(3, -1, 2)$.*

We next state Pythagoras' theorem as it appears in vectors. Note the converse.

Fact 9.2 PYTHAGORAS' THEOREM. *Vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal if and only if $||\boldsymbol{u}+\boldsymbol{v}||^2 = ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2$.*

PROOF. *Consider the following computation:*

$$\begin{aligned} ||\boldsymbol{u}+\boldsymbol{v}||^2 &= (\boldsymbol{u}+\boldsymbol{v}) \cdot (\boldsymbol{u}+\boldsymbol{v}) \qquad \textit{(by defn. of norm)} \\ &= \boldsymbol{u}\cdot\boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{v} + \boldsymbol{v}\cdot\boldsymbol{u} + \boldsymbol{v}\cdot\boldsymbol{v} \qquad \textit{(by distr. law)} \\ &= \boldsymbol{u}\cdot\boldsymbol{u} + \boldsymbol{v}\cdot\boldsymbol{v} + 2\boldsymbol{u}\cdot\boldsymbol{v} \qquad \textit{(by comm. law)} \\ &= ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 + 2\boldsymbol{u}\cdot\boldsymbol{v}. \end{aligned}$$

This means that $||\boldsymbol{u}+\boldsymbol{v}||^2 = ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2$ if and only if $\boldsymbol{u}\cdot\boldsymbol{v} = 0$.

9.2 Orthogonal Complements

*For a set W, the **orthogonal complement** is denoted by $W^\perp$ and is defined as the set of all vectors that are orthogonal to **all** of W.*

Fact 9.3 *For any subset W, the orthogonal complement $W^\perp$ is a subspace.*

This fact can be shown using the standard recipe for a subspace. For example, suppose both $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are orthogonal to every $\boldsymbol{w}$ in W. Then so is the sum of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, since $(\boldsymbol{x}_1 + \boldsymbol{x}_2)\cdot\boldsymbol{w} = \boldsymbol{x}_1\cdot\boldsymbol{w} + \boldsymbol{x}_1\cdot\boldsymbol{w} = 0 + 0 = 0$ for each $\boldsymbol{w}$. Thus $W^\perp$ is closed under addition.

Example. *Consider in $\mathbb{R}^3$ the plane P given by $3x + 4y - z = 0$. Then if we take any vector (x, y, z) in the plane P, it is orthogonal to the vector $(3, 4, -1)$: just computes their dot product and note that it is zero. This means that the orthogonal complement of the plane P contains all multiples of the vector $(3, 4, -1)$. Indeed, these are the only elements: the orthogonal complement to plane P in $\mathbb{R}^3$ is a line.*

The concept of orthogonal complement connects our three spaces of a matrix:

Fact 9.4 *For any matrix A:*

$$(\textit{Row } A)^{\perp} = \textit{Nul } A \qquad \textit{and} \qquad (\textit{Col } A)^{\perp} = \textit{Nul } A^T$$

9.3 Orthogonal and Orthonormal Sets

*An **orthogonal set** is a collection of vectors that are pairwise orthogonal. An **orthonormal set** is an orthogonal set of unit vectors.*

(The "pairwise" comment means that for every pair of distinct vectors, the two vectors are orthogonal to each other.)

A key result is that orthogonality implies independence. We omit the proof:

Fact 9.5 *If S is an orthogonal set of nonzero vectors, then S is linearly independent.*

If matrix U has orthonormal columns, then $U^T U = I$. As a matrix transform, such a matrix U preserves lengths and orthogonality. It can be shown that such a matrix must also have orthonormal rows. Thus we can speak of an **orthonormal matrix**.

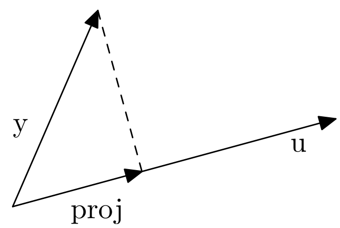
Fact 9.6 *If $B = \{\boldsymbol{w}_i\}$ is an orthonormal basis, then the coordinates of vector $\boldsymbol{v}$ relative to B are the dot-products of $\boldsymbol{v}$ with each $\boldsymbol{w}_i$.*

We will eventually show that:

Fact 9.7 *Every vector space has an orthonormal basis.*

9.4 Projections

The (orthogonal) ***projection*** *of vector* $\boldsymbol{y}$ *onto vector* $\boldsymbol{u}$ *is its "shadow". It is denoted by* $\mathrm{proj}_{\boldsymbol{u}}(\boldsymbol{y})$.



If one lets the projection be $\alpha\boldsymbol{u}$ and requires that $y - \mathrm{proj}_{\boldsymbol{u}}(\boldsymbol{y})$ is orthogonal to $\boldsymbol{u}$, some algebra produces the following formula:

Fact 9.8 *For vectors* $\boldsymbol{y}$ *and* $\boldsymbol{u}$*, the projection of* $\boldsymbol{y}$ *onto* $\boldsymbol{u}$ *is given by:*

$$\mathrm{proj}_{\boldsymbol{u}}(\boldsymbol{y}) = \frac{\boldsymbol{y} \cdot \boldsymbol{u}}{\boldsymbol{u} \cdot \boldsymbol{u}} \boldsymbol{u}$$

Example. *Calculate* $\mathrm{proj}_{\boldsymbol{b}}(\boldsymbol{a})$ *and* $\mathrm{proj}_{\boldsymbol{a}}(\boldsymbol{b})$ *for* $\boldsymbol{a} = (3, 4)$ *and* $\boldsymbol{b} = (-5, 2)$*. Then* $\mathrm{proj}_{\boldsymbol{b}}(\boldsymbol{a}) = (35/29, -14/29)$ *and* $\mathrm{proj}_{\boldsymbol{a}}(\boldsymbol{b}) = (-21/25, -28/25)$.

One can also define the (orthogonal) projection $\mathrm{proj}_W(\boldsymbol{y})$ of the vector $\boldsymbol{y}$ onto the vector space W. If we think of $\boldsymbol{y}$ as a point, then the projection of it onto W is the closest point of W to it.

It can be shown that, if W has an orthonormal basis, then the projection of $\boldsymbol{y}$ onto W is given by a simple formula: the coefficients of the basis are just the dot-product of $\boldsymbol{y}$ with each of them.

Fact 9.9 *If* W *is a subspace with orthonormal basis* $\{\boldsymbol{w}_i\}$*, then*

$$\mathrm{proj}_W(\boldsymbol{y}) = \sum_i (\boldsymbol{y} \cdot \boldsymbol{w}_i)\, \boldsymbol{w}_i$$

We omit the proof of the following result:

Fact 9.10 *If* W *is a subspace of* V*, then every vector* $\boldsymbol{y}$ *in* V *can be written uniquely in the form* $\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{z}$*, where* $\hat{\boldsymbol{y}}$ *in* W *and* $\boldsymbol{z}$ *in* $W^{\perp}$.

In the above result, $\hat{\boldsymbol{y}} = \text{proj}_W(\boldsymbol{y})$.

Example. *Consider in* $\mathbb{R}^3$ *the plane* P *given by* $3x + 4y - z = 0$. *Write the vector* $\boldsymbol{v} = (9, 9, 11)$ *as the sum of vector in* P *and vector in* $P^\perp$.

Well, one way to proceed is to calculate the vector in P, *which by the above result is* $\text{proj}_P(\boldsymbol{v})$; *if we use the above method we need an orthonormal basis of* P *(which we don't yet know how to find).*

An alternative approach is that we noted that any vector in $P^\perp$ *is a multiple of* $\boldsymbol{w} = (3, 4, -1)$. *So if we assume the requisite vector in* $P^\perp$ *is* $a\boldsymbol{w}$, *then we need* $(\boldsymbol{v} - a\boldsymbol{w}) \cdot \boldsymbol{v} = 0$. *This solves to* $a = 2$; *so* $\boldsymbol{v} = 2\boldsymbol{w} + (3, 1, 13)$.

9.5 Iterative Orthogonalization

There is a famous process (usually called Gram-Schmidt) that can be used to take a set of vectors and produce a set of vectors with the same span as the original but whose vectors are pairwise orthogonal. The process works through the set of vectors one at a time. Each time it "straightens" the vector with respect to the previous ones. The key point is that it suffices to subtract the projection of the vector onto the previous vectors.

ALGOR *Input: collection* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ *of linearly independent vectors.*
Output: collection $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ *of orthogonal vectors that span the same space.*
Process: Generate vectors $\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\boldsymbol{y}_3$, $\ldots$ *by*

$$\boldsymbol{y}_i = \boldsymbol{x}_i - \sum_{j=1}^{i-1} \text{proj}_{\boldsymbol{y}_j}(\boldsymbol{x}_i)$$

The process can be justified by checking that each requisite dot product is zero.

Example. *Find an orthonormal basis of the subspace spanned by the following three vectors.*

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{x}_2 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix} \qquad \boldsymbol{x}_3 = \begin{bmatrix} 3 \\ 1 \\ 1 \\ -7 \end{bmatrix}$$

The above process produces an orthogonal basis:

$$\boldsymbol{y}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{y}_2 = \boldsymbol{x}_2 - \frac{2}{2}\boldsymbol{y_1} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 3 \end{bmatrix} \qquad \boldsymbol{y}_3 = \boldsymbol{x}_3 - \frac{4}{2}\boldsymbol{y}_1 - \frac{-20}{12}\boldsymbol{y}_2 = \begin{bmatrix} 8/3 \\ -8/3 \\ -2/3 \\ -2 \end{bmatrix}$$

So after normalization we have the vectors
$\frac{1}{\sqrt{2}}(1,1,0,0)$, $\frac{1}{\sqrt{12}}(1,-1,-1,3)$, *and* $\frac{1}{\sqrt{42}}(4,-4,-1,-3)$

Since one can take an orthogonal basis and normalize each vector, the above algorithm provides a constructive proof of the earlier claim that every vector space has an orthonormal basis.

Practice

9.1. For each of the following triples, determine h such that the triple is orthogonal.
  (a) $(1,0,0)$, $(0,1,1)$, $(0,1,h)$
  (b) $(1,2,-2)$, $(2,3,4)$, $(0,1,h)$
  (c) $(h,h,0)$, $(h,0,h)$, $(0,h,h)$

9.2. Prove that the formula in Fact 9.8 for the orthogonal projection of one vector onto another is correct.

9.3. (a) Find an orthonormal basis for the plane through the points $(1,0,-1)$, $(3,2,5)$, and $(0,0,0)$.
  (b) Find the closest point on the plane to $(2,1,3)$.

Solutions to Practice Exercieses

9.1. (a) $h=-1$
  (b) DNE
  (c) $h=0$

9.2. Say projection $\boldsymbol{P} = \alpha\boldsymbol{u}$. Then we need $(\boldsymbol{y}-\boldsymbol{P})\cdot\boldsymbol{u} = 0$. But $(\boldsymbol{y}-\boldsymbol{P})\cdot\boldsymbol{u} = \boldsymbol{y}\cdot\boldsymbol{u} - \alpha\boldsymbol{u}\cdot\boldsymbol{u}$. Solve for α.

9.3. (a) The plane is generated by the vectors $(1, 0, -1)$ and $(3, 2, 5)$. From this we calculate that an orthogonal basis for this plane is $(1, 0, -1)$ and $(4, 2, 4)$ and normalized this is $\frac{1}{\sqrt{2}}(1, 0, -1)$ and $\frac{1}{6}(4, 2, 4)$.
(b) Using Fact 9.9 we get: $(\frac{35}{18}, \frac{11}{9}, \frac{53}{18})$.