Orthogonality for Subspaces and Bases

Recall that $\mathbf{x} \perp \mathbf{y}$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$, so it's just a dot product based definition. We are going to extend that into sets.

Definition: The sets $P \in \mathbb{R}^n$ and $Q \in \mathbb{R}^n$ are orthogonal if

$$\mathbf{p} \cdot \mathbf{q} \text{ for all } \mathbf{p} \in P \text{ and all } \mathbf{q} \in Q,$$

so all of P is orthogonal to all of Q (vice-versa, etc).

You can also have this for subspaces. The subspaces U and W of V are $U \perp W$ if all $\mathbf{u} \in U$ and all $\mathbf{w} \in W$ have $\mathbf{u} \cdot \mathbf{w} = 0$.

Example: The set $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} \right\}$ is orthogonal to the set $\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$. The spans of those sets will also be orthogonal to each other.

Definition: The *Orthogonal Complement* of the subspace $U \subseteq V$ is the set

$$U^{\perp} = \{\mathbf{x} \in V \text{ such that } \mathbf{x} \perp \text{ all } \mathbf{u} \in U\}.$$

Property: The orthogonal complement of a subspace is also a subspace.

Proof: We'll have to use the subspace test here.

- $\mathbf{0} \in V$ has $\mathbf{0} \cdot \mathbf{u}$ for all $\mathbf{u} \in U$, so it's in $U^{\perp}$.
- Take $\mathbf{x}, \mathbf{y} \in U^{\perp}$.
  $$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{u} = \mathbf{x} \cdot \mathbf{u} + \mathbf{y} \cdot \mathbf{u} = 0 + 0 = 0$$
  for any $\mathbf{u} \in U$. So, $U^{\perp}$ is closed under addition.
- Take $\mathbf{x} \in U^{\perp}$ and $a \in \mathbb{R}$.
  $$(a\mathbf{x}) \cdot \mathbf{u} = a(\mathbf{x} \cdot \mathbf{u}) = a(0) = 0$$
  so it's closed under multiplication too.

all three conditions hold, it's a subspace.

Notice that the properties of the dot product are necessary here...

Property: The intersection between U and $U^{\perp}$, $U \bigcap U^{\perp}$ is just the zero vector.

Property: The double complement, $\left(U^{\perp}\right)^{\perp}$ is simply U itself.

Example: In $\mathbb{R}^3$, here we have two subspaces:

$$U = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x - 2y - 3z = 0 \right\} \qquad V = \left\{ a \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, a \in \mathbb{R} \right\},$$

What are their complements?

For U, it'll be the set of all vectors in $\mathbb{R}^3$ orthogonal to the plane itself, so the subspace $\text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ -3 \end{bmatrix}\right\}$.

For V, it'll be the set of all vectors in $\mathbb{R}^3$ orthogonal to that line, so the plane $x+y+2z = 0$.

In $\mathbb{R}^3$, the orthogonal complement of a line is a plane, the complement of a plane is a line.

This seems reasonable, since these two example are just spans of different numbers of vectors (two then one). If you want an orthogonal complement of the span of two vectors, you'd want the span of a set of vectors orthogonal to those two. In $\mathbb{R}^3$ it'll be impossible to find more than one vectors L.I. to those two, and so we get just one.

Property: If $U = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq \mathbb{R}^n$ then

$$U^\perp = \{\mathbf{x} \in V \text{ such that } \mathbf{x}\cdot\mathbf{u}_1 = \mathbf{x}\cdot\mathbf{u}_2 = \cdots = \mathbf{x}\cdot\mathbf{u}_m = 0\},$$

so it just has to be orthogonal to all the vectors.

Corollary: If $U = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ then if we arrange a matrix

$$A = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_m \end{bmatrix} \qquad \text{then} \qquad U^\perp = \{\mathbf{x} \in \mathbb{R}^n \text{ such that } A\mathbf{x} = \mathbf{0}\}.$$

This has a fairly direct effect:

Theorem: For any matrix A, $\text{Row}(A)^\perp = \text{Null}(A)$, and (using a bunch of transposes) $\text{Col}(A)^\perp = \text{Null}(A^T)$.

Example: If $U = \text{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ -2 \\ 3 \end{bmatrix}\begin{bmatrix} -1 \\ 3 \\ 0 \\ -2 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ 1 \\ -2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ -3 \\ 0 \end{bmatrix}\right\}$ then find a basis for $U^\perp$.

So, a subspace of $\mathbb{R}^n$ can be written as the row space of a matrix. Doing so makes the orthogonal complement equal to the null space of the matrix. This has a few implications:

Theorem: If $U \subseteq \mathbb{R}^n$ and $\text{Dim}(U) = r$ then $\text{Dim}(U^\perp) = n - r$.

How To Solve for All At Once, *OPTIONAL*

The annoyance of the previous section is that our usual reduction of A yields a basis of the row space, column space (taking the original vectors in the pivot columns) and, with some effort, the null space. These correspond to the spans of the row and columns as well as the orthogonal complement of the span of the rows. We're missing the complement of the column space. We can actually solve it this way:

1. Take your matrix A and put it into an augmented matrix of the form $[A|I]$ with I the identity matrix with the same height.
2. Row reduce the A side until in RREF. Make sure to do all the row operations on the right hand side as well.
3. (a) The basis of $\text{Row}(A)$ will be the remaining non-zero columns on the left.
   (b) The basis of $\text{Col}(A)$ will be the original A columns corresponding to the pivot columns.
   (c) The basis of $\text{Null}(A)$, of $\text{Row}(A)^\perp$, will be the vectors from the general solution if the right hand side is replaced by a column of zeros.
   (d) The basis of $\text{Null}(A^T)$, of $\text{Col}(A)^\perp$, will be the row vectors on the right corresponding to zero rows on the left.

Example: The vector set $\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1/2 \\ -1/2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} \right\}$ results in the system

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1/2 & 2 & 0 & 1 & 0 \\ 1 & -1/2 & 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow \cdots \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/2 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right].$$

There are several ways to analyze this algorithm. We should be familiar with much of it. The justification for the right hand side is that we are actually solving it for an arbitrary right hand side, with each column on the RHS representing a different coefficient of the $\mathbf{b}$ vector. The full system above, at the beginning, is

$$a_1 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 1/2 \\ -1/2 \end{bmatrix} + a_3 \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Once row reduced on the left, each zero row represents a row that COULD cause a contradiction. In the example, we get an answer if and only if $\mathbf{b}$ is orthogonal to $[1, -1, 1]^T$. So, the column space of A is the orthogonal complement to that one vector, so

$$\text{Col}(A) = \text{Span} \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}^\perp \quad \Longrightarrow \quad \text{Col}(A)^\perp = \text{Span} \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

using the $(U^\perp)^\perp = U$ equality.

Orthogonal Bases

Definition: A set of vectors is orthogonal if every pair of different vectors in the set is orthogonal.

Example: The set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is orthogonal but $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ is not.

Property: If a set of NON-ZERO vectors in $\mathbb{R}^n$ is orthogonal then the set is also linearly independent.

Proof: This is surprisingly easy. Take a set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. We need to check the old equation

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0}.$$

What do we do? We inner product $\mathbf{v}_1$ on each side for

$$a_1\mathbf{v}_1 \cdot \mathbf{v}_1 + a_2\mathbf{v}_2 \cdot \mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \cdot \mathbf{v}_1 = \mathbf{0} \cdot \mathbf{v}_1$$

which works out to be

$$a_1\mathbf{v}_1 \cdot \mathbf{v}_1 = 0$$

which makes $a_1 = 0$ since $\mathbf{v}_1 \neq 0$ which makes $\mathbf{v}_1 \cdot \mathbf{v}_1 \neq 0$.

Notes: There are some implications here. A set of n non-zero orthogonal vectors is automatically a basis of $\mathbb{R}^n$, of course. Also, a non-zero orthogonal set is going to be a basis of whatever space is spans. Note too that the basic unit vectors of $\mathbb{R}^n$ form an orthogonal basis.

Actually, orthogonal bases are actually quite helpful. Recall that when $\mathbf{x} \in$ Span of the basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ then there are UNIQUE coefficients such that

$$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 + \cdots + a_n\mathbf{v}_n.$$

Having the $\mathbf{v}$ form an orthogonal basis makes this even easier. We can find the a using inner products. We start with the above equation and use $\mathbf{v}_1$ (as before) for

$$\begin{aligned} \mathbf{x} \cdot \mathbf{v}_1 &= a_1\mathbf{v}_1 \cdot \mathbf{v}_1 + a_2\mathbf{v}_2 \cdot \mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \cdot \mathbf{v}_1 \\ &= a_1\mathbf{v}_1 \cdot \mathbf{v}_1 \end{aligned}$$

so $a_1 = \frac{\mathbf{x} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}$. Sort of the numerical (non-vector) component of the projection. We can get this for ANY of the vectors and coefficients:

$$a_i = \frac{\mathbf{x} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \qquad i \in 1 \ldots n.$$

which works out to

$$\mathbf{x} = \text{Proj}_{\mathbf{v}_1}(\mathbf{x}) + \text{Proj}_{\mathbf{v}_2}(\mathbf{x}) + \cdots + \text{Proj}_{\mathbf{v}_n}(\mathbf{x}).$$

Example: Write $\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} \in \mathbb{R}^3$ in terms of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$.

They're orthogonal, and so linearly independent, so they span $\mathbb{R}^2$. We just use the inner product:

$$= \frac{-1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{-3}{6}\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} + \frac{6}{3}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Check:} = \begin{bmatrix} -1/2 \\ -1/2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1/2 \\ -1/2 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}.$$

The helpful thing here is that we can actually expand this into projections, projections on to subspaces (U) and orthogonal complements ($U^\perp$). The result is

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \qquad \mathbf{x}_1 \in U \quad \mathbf{x}_2 \in U^\perp.$$

with $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$, and so on.

Theorem: If $\mathbf{x} \in \mathbb{R}^n$ and $V \subseteq \mathbb{R}^n$ has orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_m\}$ then

$$\text{Proj}_V(\mathbf{x}) = \text{Proj}_{\mathbf{v}_1}(\mathbf{x}) + \text{Proj}_{\mathbf{v}_2}(\mathbf{x}) + \cdots + \text{Proj}_{\mathbf{v}_m}(\mathbf{x}).$$

Proving that this is true is surprisingly easy. Recall the section about projections. The value $\mathbf{z}$ is a projection of $\mathbf{x}$ onto $\mathbf{y}$ if $(\mathbf{x} - \mathbf{z})$ is orthogonal to $\mathbf{y}$ (this means that $\mathbf{z}$ has ALL of the $\mathbf{x}$ component in the $\mathbf{y}$ direction, all that's left is at right angles). We use the same principle. This time we need to confirm that

$$\big(\text{Proj}_{\mathbf{v}_1}(\mathbf{x}) + \text{Proj}_{\mathbf{v}_2}(\mathbf{x}) + \cdots + \text{Proj}_{\mathbf{v}_m}(\mathbf{x}) - \mathbf{x}\big)$$

is orthogonal to all elements in V (which will simultaneously prove we've found the projection onto $V^\perp$!).

Recall that V is spanned by the $\mathbf{v}$ vectors. If we get orthogonality for all of them, we get orthogonality for V. So:

$$\begin{aligned}
&\big(\text{Proj}_{\mathbf{v}_1}(\mathbf{x}) + \text{Proj}_{\mathbf{v}_2}(\mathbf{x}) + \cdots + \text{Proj}_{\mathbf{v}_m}(\mathbf{x}) - \mathbf{x}\big) \cdot \mathbf{v}_i \\
&= \text{Proj}_{\mathbf{v}_i}(\mathbf{x}) \cdot \mathbf{v}_i - \mathbf{x} \cdot \mathbf{v}_i \\
&= \left(\frac{\mathbf{v}_i \cdot \mathbf{x}}{\mathbf{v}_i \cdot \mathbf{v}_i}\mathbf{v}_i\right) \cdot \mathbf{v}_i - \mathbf{x} \cdot \mathbf{v}_i \\
&= \frac{\mathbf{v}_i \cdot \mathbf{x}}{\mathbf{v}_i \cdot \mathbf{v}_i}\mathbf{v}_i \cdot \mathbf{v}_i - \mathbf{x} \cdot \mathbf{v}_i \\
&= \mathbf{x} \cdot \mathbf{v}_i - \mathbf{x} \cdot \mathbf{v}_i = 0
\end{aligned}$$

and done.

Example Questions:

Section 4.5: 2.b), 5.b), 6.b) (Expansion Theorem: the one right above, about using projections to figure out the coefficients).

Section 4.6: 1.df),2, 9.bdf) 14, 16