24. Orthogonal Complements and Gram-Schmidt

Given a vector u and some other vector v not in the span of u, we can construct a new vector: Put pictures here?

$$v^{\perp} = v - \frac{u \cdot v}{u \cdot u} u.$$

This new vector $v^{\perp}$ is orthogonal to u because

$$u \bullet v^{\perp} = u \bullet v - \frac{u \cdot v}{u \cdot u} u \bullet u = 0.$$

Hence, $\{u, v^{\perp}\}$ is an orthogonal basis for $\operatorname{span}\{u, v\}$. When v is not parallel to u, $v^{\perp} \neq 0$, and normalizing these vectors we obtain $\{\frac{u}{|u|}, \frac{v^{\perp}}{|v^{\perp}|}\}$, an orthonormal basis.

Sometimes we write $v = v^{\perp} + v^{\parallel}$ where:

$$\begin{aligned} v^{\perp} &= v - \frac{u \cdot v}{u \cdot u} u \\ v^{\parallel} &= \frac{u \cdot v}{u \cdot u} u. \end{aligned}$$

This is called an *orthogonal decomposition* because we have decomposed v into a sum of orthogonal vectors. It is significant that we wrote this decomposition with u in mind; $v^{\parallel}$ is parallel to u.

If u, v are linearly independent vectors in $\mathbb{R}^3$, then the set $\{u, v^{\perp}, u \times v^{\perp}\}$ would be an orthogonal basis for $\mathbb{R}^3$. This set could then be normalized by dividing each vector by its length to obtain an orthonormal basis.

However, it often occurs that we are interested in vector spaces with dimension greater than 3, and must resort to craftier means than cross products to obtain an orthogonal basis. [1]

Given a third vector w, we should first check that w does not lie in the span of u and v, *i.e.* check that u, v and w are linearly independent. We then can define:

$$w^{\perp} = w - \frac{u \cdot w}{u \cdot u}\, u - \frac{v^{\perp} \cdot w}{v^{\perp} \cdot v^{\perp}}\, v^{\perp}.$$

One can check by directly computing $u \bullet w^{\perp}$ and $v^{\perp} \bullet w^{\perp}$ that $w^{\perp}$ is orthogonal to both u and $v^{\perp}$; as such, $\{u, v^{\perp}, w^{\perp}\}$ is an orthogonal basis for $\operatorname{span}\{u, v, w\}$.

[1] Actually, given a set T of $(n-1)$ independent vectors in n-space, one can define an analogue of the cross product that will produce a vector orthogonal to the span of T, using a method exactly analogous to the usual computation for calculating the cross product of two vectors in $\mathbb{R}^3$. This only gets us the *last* orthogonal vector, though; the process in this Section gives a way to get a full orthogonal basis.

In fact, given a collection $\{v_1, v_2, \ldots\}$ of linearly independent vectors, we can produce an orthogonal basis for $\operatorname{span}\{v_1, v_2, \ldots\}$ consisting of the following vectors:

$$
\begin{aligned}
v_1^\perp &= v_1 \\
v_2^\perp &= v_2 - \frac{v_1 \cdot v_2}{v_1 \cdot v_1} v_1 \\
&\vdots \\
v_3^\perp &= v_3 - \frac{v_1^\perp \cdot v_3}{v_1^\perp \cdot v_1^\perp} v_1^\perp - \frac{v_2^\perp \cdot v_3}{v_2^\perp \cdot v_2^\perp} v_2^\perp \\
&\vdots \\
v_i^\perp &= v_i - \sum_{j<i} \frac{v_j \cdot v_i}{v_j \cdot v_j} v_j^\perp \\
&\vdots \\
v_n^\perp &= v_n - \frac{v_1^\perp \cdot v_n}{v_1^\perp \cdot v_1^\perp} v_1^\perp - \ldots - \frac{v_{n-1}^\perp \cdot v_n}{v_{n-1}^\perp \cdot v_{n-1}^\perp} v_{n-1}^\perp
\end{aligned}
$$

Notice that each $v_i^\perp$ here depends on the existence of $v_j^\perp$ for every $j < i$. This allows us to inductively/algorithmically build up an orthogonal set of vectors whose span is $\operatorname{span}\{v_1, v_2, \ldots\}$.

Example Let $u = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$, $v = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$, and $w = \begin{pmatrix} 3 & 1 & 1 \end{pmatrix}$. We'll apply Gram-Schmidt to obtain an orthogonal basis for $\mathbb{R}^3$.

First, we set $u^\perp = u$. Then:

$$
\begin{aligned}
v^\perp &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} - \frac{2}{2}\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \\
w^\perp &= \begin{pmatrix} 3 & 1 & 1 \end{pmatrix} - \frac{4}{2}\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} - \frac{1}{1}\begin{pmatrix} 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}.
\end{aligned}
$$

Then the set

$$\{\begin{pmatrix} 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}\}$$

is an orthogonal basis for $\mathbb{R}^3$. To obtain an orthonormal basis, we simply divide each of these vectors by its length, yielding:

$$\{\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{pmatrix}\}.$$

Let U and V be subspaces of a vector space W. We saw as a review exercise that $U \cap V$ is a subspace of W, and that $U \cup V$ was not a subspace. However, $\operatorname{span} U \cup V$ is a subspace[2].

Notice that all elements of $\operatorname{span} U \cup V$ take the form $u + v$ with $u \in U$ and $v \in V$. When $U \cap V = \{0_W\}$, we call the subspace $\operatorname{span} U \cup V$ the *direct sum* of U and V, written:

$$U \oplus V = \operatorname{span} U \cup V$$

Definition Given two subspaces U and V of a space W such that $U \cap V = \{0_W\}$, the *direct sum* of U and V is defined as:

$$U \oplus V = \operatorname{span} U \cup V = \{u + v | u \in U, v \in V\}.$$

Let $w = u + v \in U \oplus V$. Then we claim that the expression $w = u + v$ is unique. To wit, suppose that $u + v = u' + v'$, with $u, u' \in U$, and $v, v' \in V$. Then we could express $0 = (u - u') + (v - v')$. Then $(u - u') = -(v - v')$. Since U and V are subspaces, we have $(u - u') \in U$ and $-(v - v') \in V$. But since these elements are equal, we also have $(u - u') \in V$. Since $U \cap V = \{0\}$, then $(u - u') = 0$. Similarly, $(v - v') = 0$, proving the claim.

Definition Given a subspace U of a vector space W, define:

$$U^\perp = \{w \in W | w \cdot u = 0 \text{ for all } u \in U\}.$$

The set $U^\perp$ (pronounced 'U-perp') is the set of all vectors in W orthogonal to *every* vector in U. This is also often called the orthogonal complement of U.

Theorem. *Let U be a subspace of a finite-dimensional vector space W. Then the set $U^\perp$ is a subspace of W, and $W = U \oplus U^\perp$.*

Proof. To see that $U^\perp$ is a subspace, we only need to check closure, which requires a simple check.

We have $U \cap U^\perp = \{0\}$, since if $u \in U$ and $u \in U^\perp$, we have:

$$u \cdot u = 0 \Leftrightarrow u = 0.$$

Finally, we show that any vector $w \in W$ is in $U \oplus U^\perp$. (This is where we use the assumption that W is finite-dimensional.) Set:

$$\begin{aligned} u &= (w \cdot e_1)e_1 + \ldots + (w \cdot e_n)e_n \in U \\ u^\perp &= w - u \end{aligned}$$

[2]This shouldn't be too surprising, though, since the span of *any* subset is a subspace.

It is easy to check that $u^{\perp} \in U^{\perp}$. Then $w = u + u^{\perp}$, so $w \in U \oplus U^{\perp}$, and we are done. $\square$

Example Consider any plane P through the origin in $\mathbb{R}^3$. Then P is a subspace, and $P^{\perp}$ is the line through the origin orthogonal to P. For example, if P is the xy-plane, then

$$\mathbb{R}^3 = P \oplus P^{\perp} = \{(x, y, 0) | x, y \in \mathbb{R}\} \oplus \{(0, 0, z) | z \in \mathbb{R}\}.$$

Notice that for any subspace U, the subspace $(U^{\perp})^{\perp}$ is just U again. As such, $\perp$ is an involution on the set of subspaces of a vector space.

References

- Hefferon, Chapter Three, Section VI.2: Gram-Schmidt Orthogonalization

Wikipedia:

- Gram-Schmidt Process
- Orthonormal Basis
- Direct Sum

Review Questions

1. Suppose u and v are linearly independent. Show that u and $v^{\perp}$ are also linearly independent. Explain why $\{u, v^{\perp}\}$ are a basis for span$\{u, v\}$.

2. Repeat the previous problem, but with three independent vectors u, v, w, and $v^{\perp}$ and $w^{\perp}$ as defined in the lecture.

3. Given any three vectors u, v, w, when do $v^{\perp}$ or $w^{\perp}$ vanish?

4. This question will answer the question, 'If I choose a vector *at random*, what is the probability that it lies in the span of some other vectors?'

    i. Given a collection S of k vectors in $\mathbb{R}^n$, consider the matrix M whose columns are the vectors in S. Show that S is linearly independent if and only if the kernel of M is trivial.

ii. Give a method for choosing a random vector v. Suppose S is a collection of k linearly independent vectors in $\mathbb{R}^n$. How can we tell whether $S \cup \{v\}$ is linearly independent? Do you think it is likely or unlikely that $S \cup \{v\}$ is linearly independent? Explain your reasoning.

iii. Let M be an $n \times n$ diagonalizable matrix whose eigenvalues are chosen uniformly at random. (*i.e.* every real number has equal chance of being an eigenvalue.) What is the probability that the columns of M form a basis for $\mathbb{R}^n$? (Hint: What is the relationship between the kernel of M and its eigenvalues?)