Inner products

Let V be a vector space over a field F. Recall that a bilinear form is a function B which takes as input two vectors $\vec{v}$ and $\vec{w}$ and is linear in each input, meaning

$$B(\vec{v}_1 + \vec{v}_2, \vec{w}) = B(\vec{v}_1, \vec{w}) + B(\vec{v}_2, \vec{w})$$

$$B(\vec{v}, \vec{w}_1 + \vec{w}_2) = B(\vec{v}, \vec{w}_1) + B(\vec{v}, \vec{w}_2)$$

$$B(c\vec{v}, \vec{w}) = B(\vec{v}, c\vec{w}) = cB(\vec{v}, \vec{w}).$$

Let V be a vector space over a field F. Recall that a bilinear form is a function B which takes as input two vectors $\vec{v}$ and $\vec{w}$ and is linear in each input, meaning

$$B(\vec{v}_1 + \vec{v}_2, \vec{w}) = B(\vec{v}_1, \vec{w}) + B(\vec{v}_2, \vec{w})$$

$$B(\vec{v}, \vec{w}_1 + \vec{w}_2) = B(\vec{v}, \vec{w}_1) + B(\vec{v}, \vec{w}_2)$$

$$B(c\vec{v}, \vec{w}) = B(\vec{v}, c\vec{w}) = cB(\vec{v}, \vec{w}).$$

The bilinear form B is called ***symmetric*** if $B(\vec{v}, \vec{w}) = B(\vec{w}, \vec{v})$.

We now switch to the field being the real numbers. In this case, we can define a special kind of symmetric bilinear form: A symmetric bilinear form is called ***positive definite*** if, for all nonzero vectors $\vec{v}$, we ave $B(\vec{v}, \vec{v}) > 0$.

A positive definite symmetric bilinear form is called an ***inner product***. We'll often denote an inner product as $(\vec{v}|\vec{w})$, $\langle \vec{v}, \vec{w} \rangle$ or $\vec{v} \cdot \vec{w}$.

We now switch to the field being the real numbers. In this case, we can define a special kind of symmetric bilinear form: A symmetric bilinear form is called ***positive definite*** if, for all nonzero vectors $\vec{v}$, we have $B(\vec{v}, \vec{v}) > 0$.

A positive definite symmetric bilinear form is called an ***inner product***. We'll often denote an inner product as $(\vec{v}|\vec{w})$, $\langle \vec{v}, \vec{w} \rangle$ or $\vec{v} \cdot \vec{w}$.

There is a variant for complex numbers coming soon.

Let V be a vector space with an inner product $B(\ ,\ )$.

For a vector $\vec{v}$ in B, we put $|\vec{v}| = \sqrt{B(\vec{v}, \vec{v})}$.

For two nonzero vectors $\vec{x}$ and $\vec{y}$, we define the angle between $\vec{x}$ and $\vec{y}$ to be $\cos^{-1} \frac{B(\vec{x},\vec{y})}{|\vec{x}|\ |\vec{y}|}$.

For two nonzero vectors $\vec{v}$ and $\vec{y}$, we define the angle between $\vec{x}$ and $\vec{y}$ to be $\cos^{-1} \frac{B(\vec{x},\vec{y})}{|\vec{x}|\ |\vec{y}|}$.

To see that this is defined, we want to prove the

Cauchy-Schwartz inequality: For any vectors $\vec{x}$ and $\vec{y}$, we have $|B(\vec{x},\vec{y})| \leq |\vec{x}|\ |\vec{y}|$. Moreover, we have equality if and only if $\vec{x}$ and $\vec{y}$ are parallel.

For two nonzero vectors $\vec{v}$ and $\vec{y}$, we define the angle between $\vec{x}$ and $\vec{y}$ to be $\cos^{-1} \frac{B(\vec{x},\vec{y})}{|\vec{x}|\ |\vec{y}|}$.

To see that this is defined, we want to prove the ***Cauchy-Schwartz inequality***: For any vectors $\vec{x}$ and $\vec{y}$, we have $|B(\vec{x},\vec{y})| \leq |\vec{x}|\ |\vec{y}|$. Moreover, we have equality if and only if $\vec{x}$ and $\vec{y}$ are parallel.

Proof: For simplicity, assume that $\vec{x}$ and $\vec{y}$ are nonzero. Put $p = |\vec{x}|$, $q = |\vec{y}|$ and put $\vec{u} = \frac{\vec{x}}{p}$ and $\vec{v} = \frac{\vec{y}}{q}$, so $\vec{x} = p\vec{u}$ and $\vec{y} = q\vec{v}$ with $|\vec{u}| = |\vec{v}| = 1$.

Then $B(\vec{u} \pm \vec{v}, \vec{u} \pm \vec{v}) = B(\vec{u},\vec{u}) \pm 2B(\vec{u},\vec{v}) + B(\vec{v},\vec{v}) \geq 0$. We have $B(\vec{u},\vec{u}) = B(\vec{v},\vec{v}) = 1$, so this gives $2 \pm 2B(\vec{u},\vec{v}) \geq 0$. So $|B(\vec{u},\vec{v})| \leq 1$ as desired. Moreover, we have equality if and only if $\vec{u} = \mp\vec{v}$. $\square$

For two nonzero vectors $\vec{v}$ and $\vec{y}$, we define the angle between $\vec{x}$ and $\vec{y}$ to be $\cos^{-1} \frac{B(\vec{x},\vec{y})}{|\vec{x}|\ |\vec{y}|}$.

Thus, $\angle(\vec{x},\vec{y})$ is ***acute*** if $B(\vec{x},\vec{y}) > 0$, is ***obtuse*** if $B(\vec{x},\vec{y}) < 0$ and is ***a right angle*** if $B(\vec{x},\vec{y}) = 0$.

We say that $\vec{x}$ and $\vec{y}$ are ***perpendicular*** or ***orthogonal*** if $B(\vec{x},\vec{y}) = 0$. We say that two subspaces X and Y of V are ***perpendicular*** or ***orthogonal*** if $B(\vec{x},\vec{y}) = 0$ for all $\vec{x} \in X$ and $\vec{y} \in Y$.

We say that two subspaces X and Y of V are ***perpendicular*** or ***orthogonal*** if $B(\vec{x}, \vec{y}) = 0$ for all $\vec{x} \in X$ and $\vec{y} \in Y$.

Problem Check that, if X and Y are orthogonal, then $X \cap Y = \{\vec{0}\}$.

We say that two subspaces X and Y of V are ***perpendicular*** or ***orthogonal*** if $B(\vec{x}, \vec{y}) = 0$ for all $\vec{x} \in X$ and $\vec{y} \in Y$.

Problem Check that, if X and Y are orthogonal, then $X \cap Y = \{\vec{0}\}$.

This suggests that we will often have orthogonal direct sums. Recall that, when we have a direct sum $V = X \oplus Y$, we get linear maps $p_X : V \to X$ and $p_Y : V \to Y$ such that $v = p_X(v) + p_X(v)$.

Problem If $V = X \oplus X^{\perp}$, show that $p_X(v)$ is the closest point to v in X.

In order to compute and discuss orthogonal projections, we introduce the notion of ***orthonormal bases***. A list of vector $\vec{u}_1$, $\vec{u}_2$, ..., is called ***orthonormal*** if

$$B(\vec{u}_i, \vec{u}_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Here are the main results about orthonormal vectors:

- If $\vec{u}_1$, $\vec{u}_2$, ..., are orthonormal, then they are linearly independent.
- If $\vec{u}_1$, $\vec{u}_2$, ..., $\vec{u}_n$ is an orthonormal basis of $X \subset V$, then p_X is given by the formula

$$p_X(\vec{v}) = \sum_{i=1}^{n} B(\vec{u}_i, \vec{v})\vec{u}_i.$$

- In the above case, $X^\perp = \mathrm{Ker}\; p_X$, and $V = X \oplus X^\perp$.
- A finite dimensional vector space, with an inner product, always has an orthonormal basis.

We'll head to Miro to prove the first three ...

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

So, take an $n-1$ dimensional subspace X of V. By induction, X has an orthonormal basis $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$.

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

So, take an $n-1$ dimensional subspace X of V. By induction, X has an orthonormal basis $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$.

Let $\vec{v}$ be any vector not in X. Put $\vec{u} = \vec{v} - p_X(\vec{v})$. Then $\vec{u}$ is orthogonal to X and (since $\vec{v} \notin X$), $\vec{u}$ is not 0. This means that we have

$$B(\vec{u}_1, \vec{u}) = B(\vec{u}_2, \vec{u}) = \cdots = B(\vec{u}_{n-1}, \vec{u}) = 0$$

$$B(\vec{u}, \vec{u}) > 0.$$

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

So, take an $n-1$ dimensional subspace X of V. By induction, X has an orthonormal basis $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$.

Let $\vec{v}$ be any vector not in X. Put $\vec{u} = \vec{v} - p_X(\vec{v})$. Then $\vec{u}$ is orthogonal to X and (since $\vec{v} \notin X$), $\vec{u}$ is not 0. This means that we have

$$B(\vec{u}_1, \vec{u}) = B(\vec{u}_2, \vec{u}) = \cdots = B(\vec{u}_{n-1}, \vec{u}) = 0$$

$$B(\vec{u}, \vec{u}) > 0.$$

Rescaling $\vec{u}$ by $\sqrt{B(\vec{u}, \vec{u})}$, we get an orthonormal basis for X. $\square$

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

So, take an $n-1$ dimensional subspace X of V. By induction, X has an orthonormal basis $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$.

Let $\vec{v}$ be any vector not in X. Put $\vec{u} = \vec{v} - p_X(\vec{v})$. Then $\vec{u}$ is orthogonal to X and (since $\vec{v} \notin X$), $\vec{u}$ is not 0. This means that we have

$$B(\vec{u}_1, \vec{u}) = B(\vec{u}_2, \vec{u}) = \cdots = B(\vec{u}_{n-1}, \vec{u}) = 0$$

$$B(\vec{u}, \vec{u}) > 0.$$

Rescaling $\vec{u}$ by $\sqrt{B(\vec{u}, \vec{u})}$, we get an orthonormal basis for X. $\square$