Inner products

Definition: An *inner product* on a real vector space V is an operation (function) that assigns to each pair of vectors $(\vec{u}, \vec{v})$ in V a scalar $\langle \vec{u}, \vec{v} \rangle$ satisfying the following axioms:

1. $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$ (Symmetry)
2. $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$ (Additivity)
3. $\langle k\, \vec{u}, \vec{v} \rangle = k \, \langle \vec{u}, \vec{v} \rangle$ (Homogeneity)
4. $\langle \vec{v}, \vec{v} \rangle \geq 0$ and $\langle \vec{v}, \vec{v} \rangle = 0$ iff $\vec{v} = \vec{0}$ (Positivity)

Theorem (basic properties): Given vectors $\vec{u}, \vec{v}, \vec{w}$ in an inner product space V, and a scalar k, the following properties hold:

- $\langle \vec{o}, \vec{v} \rangle = \langle \vec{v}, \vec{o} \rangle = 0$
- $\langle \vec{u} - \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle - \langle \vec{v}, \vec{w} \rangle$
- $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$
- $\langle \vec{u}, \vec{v} - \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle - \langle \vec{u}, \vec{w} \rangle$
- $\langle \vec{u}, k\vec{v} \rangle = k \, \langle \vec{u}, \vec{v} \rangle$

Norm and distance in an inner product space

Definition: If V is a real inner product space then we define

- The norm (or length) of $\vec{v}$:

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

- The distance between $\vec{u}$ and $\vec{v}$:

$$d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \sqrt{\langle \vec{u} - \vec{v}, \vec{u} - \vec{v} \rangle}$$

Theorem (basic properties): Given vectors $\vec{u}, \vec{v}$ in an inner product space V, and a scalar k, the following properties hold:

- $\|\vec{v}\| \geq 0$ and $\|\vec{v}\| = 0$ iff $\vec{v} = \vec{0}$.
- $\|k\vec{v}\| = |k| \, \|\vec{v}\|$
- $d(\vec{u}, \vec{v}) = d(\vec{v}, \vec{u})$
- $d(\vec{u}, \vec{v}) \geq 0$ and $d(\vec{u}, \vec{v}) = 0$ iff $\vec{u} = \vec{v}$.

Angle between vectors

Theorem (Cauchy-Schwarz): If u and v are vectors in an inner vector space, then

$$\left|\langle u, v\rangle\right| \leq \|u\| \, \|v\|$$

Definition: The angle between two vectors u and v in an inner vector space is defined as

$$\theta = \cos^{-1} \frac{\langle u, v\rangle}{\|u\| \, \|v\|}$$

Theorem (the triangle inequality): If u, v and w are vectors in an inner vector space, then

- $\|u + v\| \leq \|u\| + \|v\|$
- $d(u, v) \leq d(u, w) + d(w, v)$

Orthogonality

Definition: Two vectors u and v in an inner vector space are called *orthogonal* if $\langle u, v \rangle = 0$.

Clearly $u \perp v$ iff the angle between them is $\theta = \frac{\pi}{2}$.

Theorem (the Pythagorean theorem): If u and v are orthogonal vectors in an inner vector space, then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Definition: Let W be a subspace of an inner product space V. The set of vectors in V which are orthogonal to every vector in W is called the *orthogonal complement* of W and it is denoted by $W^\perp$.

Theorem: The orthogonal complement has the following properties:

- $W^\perp$ is a subspace of V.
- $W \cap W^\perp = \{\vec{o}\}$.
- If V has finite dimension then $(W^\perp)^\perp = W$.

Orthogonal sets, orthonormal sets

Let $(V, \langle\ \rangle)$ be an inner product space and let S be a set of vectors in V.

Definition: The set S is called *orthogonal* if any two vectors in S are orthogonal.
The set S is called *orthonormal* if it is orthogonal and any vector in S has norm 1.

Theorem: Every orthogonal set of nonzero vectors is linearly independent.

Definition: A set of vectors S is called an *orthogonal* basis (OGB) for V if S is a basis and an orthogonal set (that is, S is a basis where all vectors are perpendicular).
A set of vectors S is called an *orthonormal* basis (ONB) for V if S is a basis and an orthonormal set (that is, S is a basis where all vectors are perpendicular and have norm 1).

Orthogonal sets, orthonormal sets

Let $(V, \langle\ \rangle)$ be an inner product space.

Theorem: If $S = \{v_1, v_2, \ldots, v_n\}$ is an orthogonal basis in V and u is any vector in V, then

$$u = \frac{\langle u, v_1 \rangle}{\|v_1\|^2}\, v_1 + \frac{\langle u, v_2 \rangle}{\|v_2\|^2}\, v_2 + \ldots + \frac{\langle u, v_n \rangle}{\|v_n\|^2}\, v_n$$

If $S = \{v_1, v_2, \ldots, v_n\}$ is an orthonormal basis in V and u is any vector in V, then

$$u = \langle u, v_1 \rangle\ v_1 + \langle u, v_2 \rangle\ v_2 + \ldots + \langle u, v_n \rangle\ v_n$$

Projection onto a subspace

Let $(V, \langle\ \rangle)$ be an inner product space.
Let W be a finite dimensional subspace.

<u>Theorem</u>: If $S = \{v_1, v_2, \ldots, v_r\}$ is an orthogonal basis in W and u is any vector in V, then

$$\mathrm{proj}_W\, u = \frac{\langle u, v_1\rangle}{\|v_1\|^2}\, v_1 + \frac{\langle u, v_2\rangle}{\|v_2\|^2}\, v_2 + \ldots + \frac{\langle u, v_r\rangle}{\|v_r\|^2}\, v_r$$

If $S = \{v_1, v_2, \ldots, v_r\}$ is an orthonormal basis in W and u is any vector in V, then

$$\mathrm{proj}_W\, u = \langle u, v_1\rangle\ v_1 + \langle u, v_2\rangle\ v_2 + \ldots + \langle u, v_r\rangle\ v_r$$

Gram-Schmidt process

Theorem: Every nonzero finite dimensional inner product space has an orthonormal basis.

Given a basis $\{u_1, u_2, \ldots, u_n\}$, to find an orthogonal basis $\{v_1, v_2, \ldots, v_n\}$ we use the following procedure:

Step 1. $v_1 = u_1$

Step 2. $v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1$

Step 3. $v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2$

Step 4. $v_4 = u_4 - \frac{\langle u_4, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_4, v_2 \rangle}{\|v_2\|^2} v_2 - \frac{\langle u_4, v_3 \rangle}{\|v_3\|^2} v_3$

and so on for n steps, where $n = \dim(V)$.

To obtain an orthonormal basis, we simply normalize the orthogonal basis obtained above.

Formulation of the least squares problem

Given an *inconsistent* system $Ax = b$, find a vector x that comes "as close as possible" to being a solution.

In other words: find a vector x that *minimizes* the distance beyween b and Ax that is, a vector that minimizes $\|b - Ax\|$ (with respect to the Euclidian inner product).

We call such a vector x a *least squares solution* to the system $Ax = b$.

We call $b - Ax$ the corresponding *least squares vector* and $\|b - Ax\|$ the corresponding *least squares error*.

<u>Theorem</u>: If x is a least squares solution to the inconsistent system $Ax = b$, and if W is the column space of A, then x is a solution to the consistent system

$$Ax = \mathrm{proj}_W b$$

<u>Note:</u> The above theorem is not always practical, because finding the orthogonal projection $\mathrm{proj}_W b$ may take time (by using Gram-Schmidt).

Solution of the least squares problem

Theorem: For every inconsistent system $Ax = b$, the associated normal system

$$A^T A x = A^T b$$

is consistent and its solutions are *least squares solutions* of $Ax = b$.

Moreover, if W is the column space of A and if x is such a least squares solution to $Ax = b$, then

$$\mathrm{proj}_W b = Ax$$

Theorem: For an inconsistent system $Ax = b$ the following statements are equivalent:

a) There is a *unique* least squares solution.
b) The columns of A are linearly independent.
c) The matrix $A^T A$ is invertible.

Theorem: If an inconsistent system $Ax = b$ has a unique least squares solution, then it can be computed as

$$x^* = (A^T A)^{-1} A^T b$$

Function approximation

Problem: Given a function f on the interval $[a, b]$, a subspace W of $C[a, b]$, find the *best approximation* of f by a function g in W.

Best approximation is meant as minimizing the *mean square error*, where

$$\text{mean square error } = \int_a^b [f(x) - g(x)]^2 \, dx$$

If we consider the (standard) inner product on $C[a, b]$, defined by

$$\langle f_1, f_2 \rangle = \int_a^b f_1(x) \cdot f_2(x) \, dx$$

and the corresponding norm, then it is easy to see that

$$\text{mean square error } = \langle f - g, f - g \rangle = \|f - g\|^2$$

Therefore, the approximation problem can be reformulated as: find a function in W that minimizes $\|f - g\|^2$.

Solution: The best approximation of f by a function in W is

$$g = \mathrm{proj}_W f$$

Fourier series

We want to approximate functions by *trigonometric polynomials* of a certain order. In this case, the subspace W is $\mathbf{T}_n$, the set of all trigonometric polynomials of order $\leq n$. By definition,

$$\mathbf{T}_n = \text{ span } \{1, \cos x, \cos 2x, \ldots, \cos nx, \sin x, \sin 2x, \ldots, \sin nx\}$$

The trigonometric functions above that span $\mathbf{T}_n$ are *orthogonal*, so the set $\{1, \cos x, \cos 2x, \ldots, \cos nx, \sin x, \sin 2x, \ldots, \sin nx\}$ forms an orthogonal basis in $\mathbf{T}_n$.

Therefore, to compute $\mathrm{proj}_W f$, we can use the formula on the slide "Projection onto a subspace" and we get:

$$\begin{aligned} f(x) \approx \mathrm{proj}_{\mathbf{T}_n} f = \frac{a_0}{2} &+ [a_1 \cos x + a_2 \cos 2x + \ldots + a_n \cos nx] \\ &+ [b_1 \sin x + b_2 \sin 2x + \ldots + b_n \sin nx] \end{aligned}$$

where for $k = 0, 1, \ldots, n$, the numbers a_k and b_k are called the *Fourier coefficients* of f and they are computed as

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, dx \qquad b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx$$