Orthogonality and Approximations

Inner Product Spaces.

We will be considering a real vector space V on which an inner product is defined. An inner product is a function (x, y) from $V \times V$ to $\mathbb{R}$ with the following properties:

(1) $\quad (x, y) = (y, x)$

(2) $\quad (a_1x_1 + a_2x_2, y) = a_1(x_1, y) + a_2(x_2, y)$

(3) $\quad (x, x) > 0$ for $x \neq 0$

It is a **symmetric positive definite bilinear form**.

You should be familiar with the *dot product*

$$(x, y) = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

which is the standard inner product on $\mathbb{R}^n$.

When dealing with functions, the usual inner product takes the form

$$(f, g) = \int_{\mathcal{D}} f(s)g(s)\,ds \ .$$

In terms of the inner product we can define the length or norm of a vector/function by

$$||x|| = \sqrt{(x, x)} \ .$$

Finally, if $(x, y) = 0$, we say that x and y are orthogonal.

Approximations.

Suppose that W is a subspace of V, and $\{e_1, e_2, \ldots e_m\}$ is a basis for W.

Given a vector $x \in V$, we can look for the vector in W which is closest to x; that is, the best approximation to x in W. This vector is found geometrically by dropping a perpendicular from x onto W. The resulting vector is called the orthogonal projection of x onto W. If $x_a = x_1e_1 + x_2e_2 + \ldots x_me_m$ is this vector, then $x - x_a$ is orthogonal to every vector in W. Therefore, if y is any other vector in W (so that $x_a - y$ is also in W),

$$\begin{aligned}
||x - y||^2 &= (x - y, x - y) \\
&= ([x - x_a] + [x_a - y], [x - x_a] + [x_a - y]) \\
&= (x - x_a, x - x_a) + 2(x - x_a, x_a - y) + (x_a - y, x_a - y) \\
&= ||x - x_a||^2 + 0 + ||x_a - y||^2 \\
&> ||x - x_a||^2 \\
||x - y|| &> ||x - x_a||
\end{aligned}$$

showing that the orthogonal projection is the closest vector in W to x.

In particular, $x - x_a$ is orthogonal to each of the basis vectors e_i. The equations

$$\begin{aligned}
(x - x_a, e_i) &= 0 \\
(x, e_i) &= (x_a, e_i) \ \ ; \ i = 1, m
\end{aligned}$$

provide m simultaneous equations

$$\begin{aligned}
(e_1, e_1)x_1 + (e_2, e_1)x_2 + \ldots (e_m, e_1)x_m &= (x, e_1) \\
(e_1, e_2)x_1 + (e_2, e_2)x_2 + \ldots (e_m, e_2)x_m &= (x, e_2) \\
&\cdots \\
(e_1, e_m)x_1 + (e_2, e_m)x_2 + \ldots (e_m, e_m)x_m &= (x, e_m)
\end{aligned}$$

for evaluating the coefficients x_i.

These equations simplify if the basis vectors e_i are mutually orthogonal; that is $(e_i, e_j) = 0$ if $i \neq j$. In this case we obtain the equations

$$(e_i, e_i)x_i = (x, e_i)$$

and the solutions

$$x_i = \frac{(x, e_i)}{(e_i, e_i)}$$

are called the Fourier coefficients.

As well as simplifying the calculations, orthogonal basis vectors have the advantage that we can improve the approximation without having to recalculate everything. If we want to improve the approximation by extending the space W to $W+$ (say) by adding another basis vector e_{m+1}, then if the basis vectors are not orthogonal, we have to solve the $(m+1) \times (m+1)$ system

$$\begin{aligned}
(e_1, e_1)x_1 + (e_2, e_1)x_2 + \ldots (e_m, e_1)x_m + (e_{m+1}, e_1)x_{m+1} &= (x, e_1) \\
(e_1, e_2)x_1 + (e_2, e_2)x_2 + \ldots (e_m, e_2)x_m + (e_{m+1}, e_2)x_{m+1} &= (x, e_2) \\
&\cdots \\
(e_1, e_m)x_1 + (e_2, e_m)x_2 + \ldots (e_m, e_m)x_m + (e_{m+1}, e_m)x_{m+1} &= (x, e_m) \\
(e_1, e_{m+1})x_1 + (e_2, e_{m+1})x_2 + \ldots (e_m, e_{m+1})x_m + (e_{m+1}, e_{m+1})x_{m+1} &= (x, e_{m+1})
\end{aligned}$$

for potentially different values of $x_i, i = 1, m$ as well as x_{m+1}.

However, if the basis is orthogonal, the values of $x_i, i = 1, m$ are unchanged, and $x_{m+1} = (x, e_{m+1})/(e_{m+1}, e_{m+1})$ can be calculated without reference to the other coefficients. This is of considerable practical importance when the number of basis elements is potentially infinite.

For example, suppose that we are looking for a polynomial approximation to some function f on the interval $[-1, 1]$ using the inner product

$$(f, g) = \int_{-1}^{1} f(t)g(t)\, dt \ .$$

The obvious basis functions $\{1, t, t^2, \ldots t^n\}$ are not orthogonal; for example

$$(1, t^2) = \int_{-1}^{1} t^2\, dt = \frac{1}{3}t^3\Big|_{-1}^{1} = \frac{2}{3} \ .$$

Therefore, if for example we wish to approximate $f(t) = 1/(t^2+2t+2)$ on $[-1,1]$ by means of a polynomial of degree 1, we have

$$\begin{aligned}
(1,1) &= \int_{-1}^{1} dt = 2 \\
(1,t) &= \int_{-1}^{1} t\,dt = 0 \\
(t,t) &= \int_{-1}^{1} t^2\,dt = \frac{2}{3} \\
(f,1) &= \int_{-1}^{1} \frac{dt}{(t+1)^2+1} = \arctan 2 \\
(f,t) &= \int_{-1}^{1} \frac{(t+1)-1}{(t+1)^2+1}\,dt = \frac{1}{2}\log(5) - \arctan 2
\end{aligned}$$

so that the *best* linear approximation is

$$f \sim \frac{1}{2}\arctan 2 + \left(\frac{3}{4}\log(5) - \frac{3}{2}\arctan 2\right) t \;=\; 0.5535 - 0.4536\,t$$

while if we wish to use a second degree polynomial approximation,

$$\begin{aligned}
(1,t^2) &= \int_{-1}^{1} t^2\,dt = \frac{2}{3} \\
(t,t^2) &= \int_{-1}^{1} t^3\,dt = 0 \\
(t^2,t^2) &= \int_{-1}^{1} t^4\,dt = \frac{2}{5} \\
(f,t^2) &= \int_{-1}^{1} \frac{(t^2+2t+2)-2(t+1)}{(t+1)^2+1}\,dt = 2 - \log(5)
\end{aligned}$$

so that if the approximation is $x_0 + x_1 t + x_2 t^2$, we have to solve

$$\begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \arctan 2 \\ \frac{1}{2}\log(5) - \arctan 2 \\ 2 - \log(5) \end{pmatrix}$$

which gives

$$\begin{aligned}
x_0 &= \frac{9}{8}\arctan 2 - \frac{15}{4} + \frac{15}{8}\log(5) \\
x_1 &= \frac{3}{4}\log(5) - \frac{3}{2}\arctan 2 \\
x_2 &= -\frac{15}{8}\arctan 2 + \frac{45}{4} - \frac{45}{8}\log(5) \\
f(t) &\sim 0.5132 - 0.4536\,t + 0.121\,t^2
\end{aligned}$$

Fortunately, the functions which we normally use in this course for such approximations are orthogonal.