Abstract inner product spaces

Definition An *inner product space* is a vector space V over the real field $\mathbb{R}$ equipped with a rule for multiplying vectors, such that the product of two vectors is a scalar, and the following properties hold:

IP1) $(\underset{\sim}{u}, \underset{\sim}{v}) = (\underset{\sim}{v}, \underset{\sim}{u})$ for all $\underset{\sim}{u}$, $\underset{\sim}{v} \in V$;

IP2) $(\underset{\sim}{u} + \underset{\sim}{v}, \underset{\sim}{w}) = (\underset{\sim}{u}, \underset{\sim}{w}) + (\underset{\sim}{v}, \underset{\sim}{w})$ for all $\underset{\sim}{u}$, $\underset{\sim}{v}\,\underset{\sim}{w} \in V$;

IP3) $(k\underset{\sim}{u}, \underset{\sim}{v}) = k(\underset{\sim}{u}, \underset{\sim}{v})$ for all $\underset{\sim}{u}$, $\underset{\sim}{v} \in V$ and all $k \in \mathbb{R}$;

IP4) $(\underset{\sim}{u}, \underset{\sim}{u}) \geq 0$ for all $\underset{\sim}{u} \in V$, and if $(\underset{\sim}{u}, \underset{\sim}{u}) = 0$ then $\underset{\sim}{v} = \underset{\sim}{0}$.

The scalar (u, v) is called the *inner product* of the vectors u and v.

It was pointed out in Lecture 6 that if A is any $n \times n$ symmetric matrix with positive eigenvalues, then the rule

$$(\underset{\sim}{u}, \underset{\sim}{v}) = \underset{\sim}{u}^T A \underset{\sim}{v}$$

defines an inner product on $\mathbb{R}^n$. In the case that A is the identity matrix then this construction gives the standard inner product on $\mathbb{R}^n$; that is, the dot product.

Now consider $\mathcal{C}[a, b]$, the space of continuous functions from the interval $[a, b]$ to $\mathbb{R}$, and for f, $g \in \mathcal{C}[a, b]$ define

$$(f, g) = \int_a^b f(x)g(x)\,dx.$$

Checking that (IP1), (IP2) and (IP3) are satisfied is quite straightforward. For example, if f, g and h are arbitrary elements of $\mathcal{C}[a, b]$, then

$$\begin{aligned}(f + g, h) &= \int_a^b (f + g)(x)\,h(x)\ dx = \int_a^b (f(x) + g(x))h(x)\,dx \\ &= \int_a^b f(x)h(x) + g(x)h(x)\,dx = \int_a^b f(x)h(x)\,dx + \int_a^b g(x)h(x)\,dx = (f, h) + (g, h),\end{aligned}$$

which shows that (IP2) holds. The first part of (IP4) is also clear:

$$(f, f) = \int_a^b f(x)^2\,dx$$

is obviously nonnegative. The other assertion of (IP4)—that $(f, f) = 0$ implies $f = 0$— is also intuitively reasonable: if $f(x)$ is nonzero at any point in the interval $[a, b]$, then continuity of f guarantees that there is some subinterval of $[a, b]$ and some $\alpha > 0$ such that $f(x)^2 > \alpha$ at all points of this subinterval, and it follows that $\int_a^b f(x)^2\,dx > 0$. We omit further details of the proof, since the calculus involved is somewhat removed from the topics that are the main focus of this course.

We can also use the formula $\int_a^b f(x)g(x)\,dx$ to define an inner product on $\mathcal{P}[a, b]$, the polynomial functions on $[a, b]$ (which is a subspace of $\mathcal{C}[a, b]$).

If V is an inner product space and $f \in V$, then, as in $\mathbb{R}^n$, we define $\|f\| = \sqrt{(f, f)}$, and call this the *length* of f. The Cauchy-Schwarz Inequality says that $|(f, g)| \leq \|f\|\,\|g\|$ for all f, $g \in V$, and the proof of this inequality for inner product spaces in general is just the same as its proof in $\mathbb{R}^n$, since the proof uses nothing beyond the properties

(IP1)–(IP4). As in $\mathbb{R}^n$, we define the angle between elements f and g of an inner product space to be $\arccos((f,g)/\|f\|\,\|g\|)$. We say that f and g are orthogonal if $(f,g)=0$. Pythagoras' Theorem holds for any inner product space, and is proved in the same way as it is for $\mathbb{R}^n$. And our discussion of projections also goes through unchanged.

Example 1. Let f_1, f_2 and f_3 be the functions $[0,1]\to\mathbb{R}$ defined by

$$f_1(x)=1, \qquad f_2(x)=1-2x, \qquad f_3(x)=1-6x+6x^2.$$

These form an orthogonal set in $\mathcal{P}[0,1]$. To see this, we must check that (f_1,f_2), (f_1,f_3) and (f_2,f_3) are all zero. We have

$$\begin{aligned}
(f_1,f_2) &= \int_0^1 (1-2x)\,dx = x-x^2\Big]_0^1 = 0-0=0,\\
(f_1,f_3) &= \int_0^1 (1-6x+6x^2)\,dx = 1x-3x^2+2x^3\Big]_0^1 = 0-0=0,\\
(f_2,f_3) &= \int_0^1 (1-2x)(1-6x+6x^2)\,dx = \int_0^1 1-8x+18x^2-12x^3)\,dx\\
&\qquad\qquad\qquad\qquad\qquad\qquad = x-4x^2+6x^3-3x^4\Big]_0^1 = 0-0=0,
\end{aligned}$$

as required.

Projections in inner product spaces

Let $\{\utilde{a}_1,\utilde{a}_2,\ldots,\utilde{a}_k\}$ be a basis for a finite-dimensional subspace W of an inner product space V. For each $\utilde{v}\in V$ there is a unique $\utilde{p}\in W$ with the property that $(\utilde{x},\utilde{v}-\utilde{p})=0$ for all $\utilde{x}\in W$; that is, $\utilde{v}-\utilde{p}$ is orthogonal to all elements of W. The element $\utilde{p}$ is called the projection of $\utilde{v}$ onto W. The condition that $(\utilde{x},\utilde{v}-\utilde{p})=0$ for all $\utilde{x}\in W$ can be reformulated as $(\utilde{x},\utilde{v})=(\utilde{x},\utilde{p})$ for all $\utilde{x}\in W$, and this in turn is equivalent to the condition that $(\utilde{a}_i,\utilde{v})=(\utilde{a}_i,\utilde{p})$ for all $i=1,\,2,\,\ldots,\,k$. We may write $\utilde{p}$ as a linear combination of the basis elements of W,

$$\utilde{p}=\lambda_1\utilde{a}_1+\lambda_2\utilde{a}_2+\cdots+\lambda_k\utilde{a}_k, \tag{1}$$

and then the condition becomes

$$\begin{aligned}
\begin{pmatrix}(\utilde{a}_1,\utilde{v})\\(\utilde{a}_2,\utilde{v})\\\vdots\\(\utilde{a}_k,\utilde{v})\end{pmatrix}
=\begin{pmatrix}(\utilde{a}_1,\utilde{p})\\(\utilde{a}_2,\utilde{p})\\\vdots\\(\utilde{a}_k,\utilde{p})\end{pmatrix}
&=\begin{pmatrix}(\utilde{a}_1,\lambda_1\utilde{a}_1+\lambda_2\utilde{a}_2+\cdots+\lambda_k\utilde{a}_k)\\(\utilde{a}_2,\lambda_1\utilde{a}_1+\lambda_2\utilde{a}_2+\cdots+\lambda_k\utilde{a}_k)\\\vdots\\(\utilde{a}_k,\lambda_1\utilde{a}_1+\lambda_2\utilde{a}_2+\cdots+\lambda_k\utilde{a}_k)\end{pmatrix}\\
&=\begin{pmatrix}(\utilde{a}_1,\utilde{a}_1)\lambda_1+(\utilde{a}_1,\utilde{a}_2)\lambda_2+\cdots+(\utilde{a}_1,\utilde{a}_k)\lambda_k\\(\utilde{a}_2,\utilde{a}_1)\lambda_1+(\utilde{a}_2,\utilde{a}_2)\lambda_2+\cdots+(\utilde{a}_2,\utilde{a}_k)\lambda_k\\\vdots\\(\utilde{a}_k,\utilde{a}_1)\lambda_1+(\utilde{a}_k,\utilde{a}_2)\lambda_2+\cdots+(\utilde{a}_k,\utilde{a}_k)\lambda_k\end{pmatrix}\\
&=\begin{pmatrix}(\utilde{a}_1,\utilde{a}_1)&(\utilde{a}_1,\utilde{a}_2)&\ldots&(\utilde{a}_1,\utilde{a}_k)\\(\utilde{a}_2,\utilde{a}_1)&(\utilde{a}_2,\utilde{a}_2)&\ldots&(\utilde{a}_2,\utilde{a}_k)\\\vdots&\vdots&&\vdots\\(\utilde{a}_k,\utilde{a}_1)&(\utilde{a}_k,\utilde{a}_2)&\ldots&(\utilde{a}_k,\utilde{a}_k)\end{pmatrix}\begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_k\end{pmatrix}.
\end{aligned}$$

In other words, we can find $\underset{\sim}{p}$ by solving the system of linear equations

$$G \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix} = \begin{pmatrix} (\underset{\sim}{a}_1, \underset{\sim}{v}) \\ (\underset{\sim}{a}_2, \underset{\sim}{v}) \\ \vdots \\ (\underset{\sim}{a}_k, \underset{\sim}{v}) \end{pmatrix},$$

where the matrix G is the Gram matrix of the basis $\{\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_k\}$, and then $\underset{\sim}{p}$ is given by Eq. (1) above.

The above calculation is essentially the same as a calculation we did when discussing projections in the context of $\mathbb{R}^n$ in the first week of lectures. The formula above is the same as the formula we derived then, since the matrix G above coincides with the matrix $A^T A$ that we had before.

As in $\mathbb{R}^n$, projections become simpler when orthogonal bases are used. So you should generally not use the above formula to compute projections. Instead, first apply the Gram-Schmidt process to obtain an orthogonal basis $\{\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_k\}$ for the subspace W, and then use the formula

$$\underset{\sim}{p} = \frac{(\underset{\sim}{b}_1, \underset{\sim}{v})}{(\underset{\sim}{b}_1, \underset{\sim}{b}_1)} \underset{\sim}{b}_1 + \frac{(\underset{\sim}{b}_2, \underset{\sim}{v})}{(\underset{\sim}{b}_2, \underset{\sim}{b}_2)} \underset{\sim}{b}_2 + \cdots + \frac{(\underset{\sim}{b}_k, \underset{\sim}{v})}{(\underset{\sim}{b}_k, \underset{\sim}{b}_k)} \underset{\sim}{b}_k \tag{2}$$

to compute the projection of $\underset{\sim}{v}$ onto W. Remember that Eq. (2) is only valid when the basis $\{\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_k\}$ is othogonal!

Legendre polynomials

Let $\mathcal{C}$ be the space of continuous functions $[-1, 1] \to \mathbb{R}$, with inner product defined by

$$(f, g) = \int_{-1}^{1} f(x)g(x)\, dx.$$

Let f_0, f_1, f_2, …, $f_k \in \mathcal{C}$ be defined by $f_n(x) = x^n$ for all $x \in [-1, 1]$, and let $\mathcal{P}_k$ be the subspace of $\mathcal{C}$ spanned by f_0, f_1, …, f_k. That is, $\mathcal{P}_k$ is the subspace of $\mathcal{C}$ given by polynomial functions of degree at most k.

Given a continuous function $f\colon [-1, 1] \to \mathbb{R}$, the projection of f onto $\mathcal{P}_k$ is the polynomial function of degree at most k that is the best approximation to f on $[-1, 1]$, in the "least squares" sense: it gives the minimal value for $\int_{-1}^{1} (f(x) - p(x))^2\, dx$, subject to p being a polynomial of degree at most k.

As explained above, for the purpose of computing such projections conveniently, we need an orthogonal basis for $\mathcal{P}_k$. To obtain one, we apply the Gram-Schmidt process, starting with the basis $\{f_0, f_1, \ldots, f_k\}$. The formulas are as follows:

$$\begin{aligned}
g_0 &= f_0 \\
g_1 &= f_1 - \frac{(f_1, g_0)}{(g_0, g_0)} g_0 \\
g_2 &= f_2 - \frac{(f_2, g_0)}{(g_0, g_0)} g_0 - \frac{(f_2, g_1)}{(g_1, g_1)} g_1 \\
g_3 &= f_3 - \frac{(f_3, g_0)}{(g_0, g_0)} g_0 - \frac{(f_3, g_1)}{(g_1, g_1)} g_1 - \frac{(f_3, g_2)}{(g_2, g_2)} g_2 \\
&\ \vdots
\end{aligned}$$

and so on.

Let us compute these. Firstly, we have

$$(f_1, g_0) = \int_{-1}^{1} x\,1\,dx = \tfrac{1}{2}x^2\Big]_{-1}^{1} = 0,$$

and so it follows that $g_1 = f_1$. Next,

$$\begin{aligned}
(f_2, g_0) &= \int_{-1}^{1} x^2 1\,dx = \tfrac{1}{3}x^3\Big]_{-1}^{1} = \tfrac{2}{3} \\
(g_0, g_0) &= \int_{-1}^{1} 1\,dx = x\Big]_{-1}^{1} = 2 \\
(f_2, g_1) &= \int_{-1}^{1} x^2 x\,dx = \tfrac{1}{4}x^4\Big]_{-1}^{1} = 0 \\
(g_1, g_1) &= \int_{-1}^{1} x^2\,dx = \tfrac{2}{3},
\end{aligned}$$

and so

$$g_2 = f_2 - \tfrac{(2/3)}{2} g_0 - \tfrac{0}{(2/3)} g_1 = f_2 - \tfrac{1}{3} g_0.$$

Similarly, we find that

$$\begin{aligned}
(f_3, g_0) &= \int_{-1}^{1} x^3\,dx = 0 \\
(f_3, g_1) &= \int_{-1}^{1} x^4\,dx = \tfrac{2}{5} \\
(f_3, g_2) &= \int_{-1}^{1} x^3(x^2 - \tfrac{1}{3})\,dx = 0,
\end{aligned}$$

and this gives

$$g_3 = f_3 - 0g_0 - \tfrac{(2/5)}{(2/3)} g_1 - 0g_2 = f_3 - \tfrac{3}{5} g_1.$$

Students are invited to compute further terms for themselves.

The polynomials we have been calculating are known as *Legendre polynomials.* Expressed in terms of x, the first few are as follows:

$$\begin{aligned}
g_0(x) &= 1 \\
g_1(x) &= x \\
g_2(x) &= x^2 - \tfrac{1}{3} \\
g_3(x) &= x^3 - \tfrac{3}{5}x \\
g_4(x) &= x^4 - \tfrac{6}{7}x^2 + \tfrac{3}{35} \\
g_5(x) &= x^5 - \tfrac{10}{9}x^3 + \tfrac{5}{21}x \\
g_6(x) &= x^6 - \tfrac{15}{11}x^4 + \tfrac{5}{11}x^2 - \tfrac{5}{231}.
\end{aligned}$$

Strictly speaking, the Legendre polynomials are not these, but scalar multiples of these, the scalars being chosen so that the polynomials take the value 1 at $x = 1$. Performing this scaling gives

$$\begin{aligned}
P_0(x) &= 1\\
P_1(x) &= x\\
P_2(x) &= \tfrac{1}{2}(3x^2 - 1)\\
P_3(x) &= \tfrac{1}{2}(5x^3 - 3x)\\
P_4(x) &= \tfrac{1}{8}(35x^4 - 30x^2 + 3)\\
P_5(x) &= \tfrac{1}{8}(63x^5 - 70x^3 + 15x)\\
P_6(x) &= \tfrac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5).
\end{aligned}$$

In fact, the general formula is

$$P_n(x) = \sum_{k=0}^{[n/2]} \frac{(-1)^k(2n-k)!}{2^n k!(n-k)!(n-2k)!} x^{n-2k},$$

but we shall not prove this.

Here is some MAGMA code for calculating Legendre polynomials. It utilizes the MAGMA functions `Integral` and `Evaluate`: if `f` is a polynomial in `x` then `Integral(f)` is the polynomial with zero constant term whose derivative is `f`; if `f` is a polynomial in `x` and `t` a number then `Evaluate(f,t)` is the number obtained by evaluating `f` when `x` is given the value `t`. (It is `f(t)`, so to speak, although `f(t)` is not correct MAGMA syntax.)

```
> R := RealField();
> P<x> := PolynomialAlgebra(R);
> polyip := func< f,g | Evaluate(Integral(f*g),1) -
>     Evaluate(Integral(f*g),-1) >;
> //
> // So polyip(f,g) is the integral of f*g from -1 to 1,
> // which is exactly the inner product of f and g.
> //
> f := [];
> for i in [0..6] do
for> f[i+1] := x^i;
for> end for;
> f;
[
  1
  x
  x^2
  x^3
  x^4
  x^5
  x^6
]
```

```
> g := [ ];
> for i in [0..6] do
for> g[i+1]:=f[i+1];
for> for j in [1..i] do
for|for> g[i+1] := g[i+1] - (polyip(g[i+1],g[j])/
for|for>        polyip(g[j],g[j]))*g[j];
for|for> end for;
for> end for;
> g;
[
  1,
  x,
  x^2 - 1/3,
  x^3 - 3/5*x,
  x^4 - 6/7*x^2 + 3/35,
  x^5 - 10/9*x^3 + 5/21*x,
  x^6 - 15/11*x^4 + 5/11*x^2 - 5/231
]
> q:=[ ];
> for i in [1..7] do
for> q[i] := g[i]/Evaluate(g[i],1);
> end for;
>q;
[
  1,
  x,
  3/2*x^2 - 1/2,
  5/2*x^3 - 3/2*x,
  35/8*x^4 - 15/4*x^2 + 3/8,
  63/8*x^5 - 35/4*x^3 + 15/8*x,
  231/16*x^6 - 315/16*x^4 + 105/16*x^2 - 5/16
]
```

Example 2. Let us use the orthogonal basis $\{g_0, g_1, g_2\}$ to compute the projection onto the space W_2 of the function $f\colon [-1,1] \to \mathbb{R}$ given by $f(x) = e^x$ (for all $x \in [-1,1]$). By the formula, the projection p is given by

$$p = \frac{(f, g_0)}{(g_0, g_0)} g_0 + \frac{(f, g_1)}{(g_1, g_1)} g_1 + \frac{(f, g_2)}{(g_2, g_2)} g_2.$$

This formula, of course, is only applicable since the g_i form an orthogonal set.

We need to calculate a few integrals:

$$\begin{aligned}
(f, g_0) &= \int_{-1}^{1} e^x 1\, dx = e^x \Big]_{-1}^{1} = e - e^{-1}, \\
(f, g_1) &= \int_{-1}^{1} e^x x\, dx = xe^x - e^x \Big]_{-1}^{1} = 2e^{-1}, \\
(f, g_2) &= \int_{-1}^{1} e^x (x^2 - \tfrac{1}{3})\, dx = x^2 e^x - 2xe^x + \tfrac{5}{3} e^x \Big]_{-1}^{1} = \tfrac{2}{3} e - \tfrac{14}{3} e^{-1}.
\end{aligned}$$

We found above that $(g_0, g_0) = 2$ and $(g_1, g_1) = \frac{2}{3}$, and we also need

$$(g_2, g_2) = \int_{-1}^{1} (x^2 - \tfrac{1}{3})^2 \, dx = \tfrac{8}{45}.$$

The final answer is

$$\begin{aligned} p(x) &= \frac{e - e^{-1}}{2} + \frac{2e^{-1}}{(2/3)} x + \frac{(2/3)e - (14/3)e^{-1}}{(8/45)} (x^2 - \tfrac{1}{3}) \\ &\approx 0.996 + 1.104x + 0.537x^2. \end{aligned}$$

Note that the Taylor series for e^x about $x = 0$ is $1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots$, and in particular the polynomial we have found is not much different from the degree 2 Taylor polynomial, $1 + x + \frac{1}{2}x^2$. The difference is due to the fact that the accuracy of the Taylor polynomial as an approximation to e^x improves the closer x is to zero, whereas the least squares approximation gives equal weight to all values of x in the interval $[-1, 1]$.

Fourier series

Let $\mathcal{C}[-\pi, \pi]$ be the space of continuous functions $[-\pi, \pi] \to \mathbb{R}$, with inner product $(f, g) = \int_{-\pi}^{\pi} f(x)g(x)\, dx$. Let c_0, c_1, c_2, … and s_1, s_2, … be the elements of $\mathcal{C}[-\pi, \pi]$ defined by

$$\begin{aligned} c_n(x) &= \cos(nx) \\ s_n(x) &= \sin(nx) \end{aligned}$$

for x in the interval $[-\pi, \pi]$. Note that s_n is not defined for $n = 0$, while c_0 is the constant function $c_0(x) = 1$ (since $\cos(0x) = \cos 0 = 1$ for all x). To compute inner products involving these, we need to use some standard trigonometric identities. For all nonnegative integers n and m,

$$\begin{aligned} (s_n, c_m) &= \int_{-\pi}^{\pi} \sin(nx)\, \cos(mx)\, dx \\ &= \int_{-\pi}^{\pi} \tfrac{1}{2}(\sin(n+m)x - \sin(n-m)x)\, dx \\ &= \begin{cases} \tfrac{1}{2}\left(\tfrac{-1}{n+m}\cos(n+m)x - \tfrac{-1}{n-m}\cos(n-m)x\right)\Big]_{-\pi}^{\pi} & \text{if } n \neq m \\ \tfrac{1}{2}\left(\tfrac{-1}{2m}\cos(2m\, x)\right)\Big]_{-\pi}^{\pi} & \text{if } n = m \end{cases} \end{aligned}$$

since the term $\sin(n - m)x$ vanishes when $n = m$. The crucial point to observe now is that, by the periodicity of the cosine function, $\cos(kx)$ takes the same value at $x = -\pi$ as at $x = \pi$, whenever k is an integer. So in both cases the above integral vanishes, and we deduce that $(s_n, c_m) = 0$. Similarly, if n, m are positive integers,

$$
\begin{aligned}
(s_n, s_m) &= \int_{-\pi}^{\pi} \sin(nx)\,\sin(mx)\,dx \\
&= \int_{-\pi}^{\pi} \tfrac{1}{2}(\cos(n-m)x - \cos(n+m)x)\,dx \\
&= \begin{cases} \tfrac{1}{2}\big(\tfrac{1}{n-m}\sin(n-m)x - \tfrac{1}{n+m}\sin(n+m)x\big)\Big]_{-\pi}^{\pi} & \text{if } n \neq m \\ \tfrac{1}{2}\big(x - \tfrac{1}{2m}\sin(2m\,x)\big)\Big]_{-\pi}^{\pi} & \text{if } n = m \end{cases} \\
&= \begin{cases} 0 & \text{if } n \neq m \\ \pi & \text{if } n = m, \end{cases}
\end{aligned}
$$

where again the periodicity of the sin function simplifies the calculations. (Note that in the case $n = m$ above the term $\cos(n-m)x$ becomes simply $\cos 0 = 1$, and integrating this gives x. The formula $\frac{1}{n-m}\sin(n-m)x$ for the integral of $\cos(n-m)x$ is not valid when $n = m$, but it is valid for all other values of n and m.)

The calculation of (c_n, c_m), when n and m are positive integers, is analogous to the calculation of (s_n, s_m):

$$
\begin{aligned}
(c_n, c_m) &= \int_{-\pi}^{\pi} \cos(nx)\,\cos(mx)\,dx \\
&= \int_{-\pi}^{\pi} \tfrac{1}{2}(\cos(n-m)x + \cos(n+m)x)\,dx \\
&= \begin{cases} \tfrac{1}{2}\big(\tfrac{1}{n-m}\sin(n-m)x + \tfrac{1}{n+m}\sin(n+m)x\big)\Big]_{-\pi}^{\pi} & \text{if } n \neq m \\ \tfrac{1}{2}\big(x + \tfrac{1}{2m}\sin(2m\,x)\big)\Big]_{-\pi}^{\pi} & \text{if } n = m \end{cases} \\
&= \begin{cases} 0 & \text{if } n \neq m \\ \pi & \text{if } n = m. \end{cases}
\end{aligned}
$$

When n and m are both zero the above calculation does not apply (because of the $\frac{1}{n+m}$), and in fact $(c_0, c_0) = \int_{-\pi}^{\pi} 1\,dx = 2\pi$.

These calculations have shown, in particular, that $\{c_0, s_1, c_1, s_2, c_2, s_3, c_3, \ldots, s_k, c_k\}$ is an orthogonal set in $\mathcal{C}[-\pi,\pi]$ (for any value of k). If we define W_k to be the subspace spanned by this set, then we can calculate the projection of any $f \in \mathcal{C}[-\pi,\pi]$ onto this subspace by means of the general formulas we have obtained. Specifically, if p is the projection of f then p is given by the formula

$$
p = \frac{(f, c_0)}{(c_0, c_0)}c_0 + \frac{(f, s_1)}{(s_1, s_1)}s_1 + \frac{(f, c_1)}{(c_1, c_1)}c_1 + \cdots + \frac{(f, c_k)}{(c_k, c_k)}c_k.
$$

In view of the formulas for (c_n, c_n) and (s_n, s_n) this yields, for all $x \in [-\pi, \pi]$,

$$
\begin{aligned}
p(x) = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt\Big) &+ \Big(\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin t\,dt\Big)\sin x + \Big(\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos t\,dt\Big)\cos x \\
&+ \Big(\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(2t)\,dt\Big)\sin(2x) + \cdots + \Big(\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(kt)\,dt\Big)\cos(kx).
\end{aligned}
$$

This is the best approximation—in the least squares sense—to the function f on the interval $[-\pi, \pi]$ by a function in the subspace W_k. The larger k is, the better the approximation. Letting k tend to ∞ yields an infinite series known as the *Fourier series* of f.

Example 3. Let us find the Fourier series for the function $f(x) = x$ on the interval $[-\pi, \pi]$.

It is useful to remember that if a function g has the property that $g(-x) = -g(x)$ for all $x \in [-a, a]$, then $\int_{-a}^{a} g(x)\,dx = 0$. In particular, the function g defined by $g(x) = x\cos(nx)$ has this property (for any value of n), and so

$$\int_{-\pi}^{\pi} x\cos(nx)\,dx = 0.$$

Thus the coefficient of $\cos nx$ in the Fourier series of x is zero for all n. The coefficient of $\sin(nx)$ is

$$\begin{aligned}\frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(nx)\,dx &= \frac{1}{\pi}\left(-\frac{x}{n}\cos(nx)\Big]_{-\pi}^{\pi} - \int_{-\pi}^{\pi}\frac{-1}{n}\cos(nx)\,dx\right)\\ &= \left(-\frac{x}{\pi n}\cos(nx) + \frac{1}{\pi n^2}\sin(nx)\right)\Big]_{-\pi}^{\pi}\\ &= \frac{-\pi}{\pi n}\cos(n\pi) + \frac{-\pi}{\pi n}\cos(n(-\pi))\\ &= \frac{2(-1)^{n+1}}{n}\end{aligned}$$

So the Fourier series of x on $[-\pi, \pi]$ is

$$2(\sin x - \tfrac{1}{2}\sin(2x) + \tfrac{1}{3}\sin(3x) - \tfrac{1}{4}\sin(4x) + \tfrac{1}{4}\sin(5x) - \cdots \quad).$$

It can be shown that this series converges to x when $x \in [-\pi, \pi]$. When $x \notin [-\pi, \pi]$ the series still converges, but, rather than x, the limit is $x - 2k\pi$, where k is defined by the requirement that $x - 2k\pi \in [-\pi, \pi]$.

We conclude the section of the course on inner product spaces with two more examples of calculations with spaces of continuous functions. The first of these was done incorrectly in lectures: I inadvertently omitted a couple of square root signs, so that the quantities which I said were the length of f and the length of g were in fact the squares of these lengths.

Example 4. We verify the Cauchy-Schwarz inequality for the functions $f(x) = x$ and $g(x) = -e^x$ in $\mathcal{C}[0, 2]$.

Recall that the Cauchy-Schwarz inequality says that

$$|(f.g)| \leq \|f\|\,\|g\|.$$

Verifying this for f and g as given is simply a matter of evaluating some integrals:

$$\begin{aligned}
\|f\| &= \sqrt{(f,f)} = \sqrt{\int_0^2 x^2\,dx} = \sqrt{\tfrac{1}{3}x^3\Big]_0^2} = \sqrt{\tfrac{8}{3}},\\
\|g\| &= \sqrt{(g,g)} = \sqrt{\int_0^2 e^{2x}\,dx} = \sqrt{\tfrac{1}{2}e^{2x}\Big]_0^2} = \sqrt{\tfrac{1}{2}(e^4-1)},\\
(f,g) &= \int_0^2 (-xe^x)\,dx = -xe^x + e^x\Big]_0^2 = (e^2 - 2e^2) - e^0 = -1 - e^2.
\end{aligned}$$

Thus $|(f,g)| = 1 + e^2 \approx 8.39$, which is less than $\|f\|\,\|g\| = \sqrt{\frac{4}{3}(e^4-1)} \approx 8.45$.

Note that, by definition, the angle between f and g is $\cos^{-1}\bigl(\frac{(f,g)}{\|f\|\,\|g\|}\bigr)$, although this quantity has no geometrical interpretation. In this case $\frac{(f,g)}{\|f\|\,\|g\|}$ is fairly close to 1, and so the angle is fairly close to zero. In fact it is approximately 0.127 radians, or 7.09 degrees.

Example 5. For our final example we compute the best approximation to $\cos x$ in $\mathcal{P}_2[-\pi,\pi]$ (the space of polynomial functions of degree at most 2 on the interval $[-\pi,\pi]$.

The necessary first step is to find an orthogonal basis for $\mathcal{P}_2[-\pi,\pi]$. We do this by applying the Gram-Schmidt process to the basis $\{f_0, f_1, f_2\}$, where $f_i(x) = x^2$ (for all $x \in [0,2]$). This is very similar to the calculation of the Legendre polynomials, but the numbers come out differently since we are working over a different interval now.

The new basis is

$$\begin{aligned}
g_0 &= f_0\\
g_1 &= f_1 - \frac{(f_1,g_0)}{(g_0,g_0)}g_0\\
g_2 &= f_2 - \frac{(f_2,g_0)}{(g_0,g_0)}g_0 - \frac{(f_2,g_1)}{(g_1,g_1)}g_1.
\end{aligned}$$

We have $(f_1,g_0) = \int_{-\pi}^{\pi} x\,1\,dx = 0$, and so $g_1 = f_1$. Now

$$\begin{aligned}
(f_2,g_0) &= \int_{-\pi}^{\pi} x^2\,dx = \tfrac{2}{3}\pi^3,\\
(f_2,g_1) &= \int_{-\pi}^{\pi} x^3\,dx = 0,\\
(g_0,g_0) &= \int_{-\pi}^{\pi} 1\,dx = 2\pi.
\end{aligned}$$

Thus

$$g_2 = f_2 - \frac{(2/3)\pi^3}{2\pi}g_0 = f_2 - \frac{\pi^2}{3}g_0,$$

so that $g_0(x) = 1$, $g_1(x) = x$ and $g_2(x) = x^2 - \frac{\pi^2}{3}$, for all $x \in [-\pi,\pi]$.

Since $\{g_0, g_1, g_2\}$ is an orthogonal basis for the space, the projection of cos onto this space is given by

$$\frac{(\cos,g_0)}{(g_0,g_0)}g_0 + \frac{(\cos,g_1)}{(g_1,g_1)}g_1 + \frac{(\cos,g_2)}{(g_2,g_2)}g_2.$$

We find that $(\cos, g_0) = \int_{-\pi}^{\pi} \cos x\, dx = \sin x\,\Big]_{-\pi}^{\pi} = 0$. And $(\cos, g_1) = \int_{-\pi}^{\pi} x\cos x\, dx = 0$, since the function $f(x) = x\cos x$ satisfies $f(-x) = f(x)$ for all $x \in [-\pi, \pi]$. Now

$$
\begin{aligned}
(\cos, g_2) &= \int_{-\pi}^{\pi} (\cos x)(x^2 - \tfrac{\pi^2}{3})\, dx \\
&= \int_{-\pi}^{\pi} x^2(\cos x)\, dx \\
&= x^2(\sin x) - \int 2x(\sin x)\, dx\,\Big]_{-\pi}^{\pi} \\
&= x^2(\sin x) - \big(-2x(\cos x) + \int 2(\cos x)\, dx\,\big)\Big]_{-\pi}^{\pi} \\
&= x^2(\sin x) + 2x(\cos x) - 2(\sin x)\,\Big]_{-\pi}^{\pi} \\
&= 2\pi(-1) - 2(-\pi)(-1) \\
&= -4\pi,
\end{aligned}
$$

and we also have that

$$
(g_2, g_2) = \int_{-\pi}^{\pi} x^4 - \tfrac{2\pi^2}{3}x^2 + \tfrac{\pi^4}{9}\; dx = \tfrac{2}{5}\pi^5 - \tfrac{4}{9}\pi^5 + \tfrac{2}{9}\pi^5 = \tfrac{8}{45}\pi^5.
$$

So the projection of cos onto $\mathcal{P}_2[-\pi, \pi]$ is $\frac{-4\pi}{(8/45)\pi^5}(x^2 - \frac{\pi^2}{3}) = \frac{-45}{2\pi^4}(x^2 - \frac{\pi^2}{3})$.

The diagram below is a fairly accurate graph of cos and its projection.