Inner Product Spaces

In $\mathbb{R}^n$, we defined an inner product $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T\boldsymbol{v} = u_1v_1 + ... + u_nv_n$. Another notation sometimes used is $\boldsymbol{u} \cdot \boldsymbol{v} = \ < \boldsymbol{u}, \boldsymbol{v} >$.

The inner product in $\mathbb{R}^n$ has several important properties (*see Theorem 1, p. 331*) that we have used over and over. Written with the $< \boldsymbol{u}, \boldsymbol{v} >$ notation, they are

a) $< \boldsymbol{u}, \boldsymbol{v} > \ = \ < \boldsymbol{v}, \boldsymbol{u} >$
b) $< \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} > \ = \ < \boldsymbol{u}, \boldsymbol{w} > + < \boldsymbol{v}, \boldsymbol{w} >$
c) $< c\boldsymbol{u}, \boldsymbol{v} > \ = c < \boldsymbol{u}, \boldsymbol{v} >$
d) $< \boldsymbol{u}, \boldsymbol{u} > \ \geq 0$ and $< \boldsymbol{u}, \boldsymbol{u} > \ = 0$ if and only if $\boldsymbol{u} = \boldsymbol{0}$.

Using the inner product, we then defined length $\|\boldsymbol{u}\| = \ < \boldsymbol{u}, \boldsymbol{u} >^{1/2}$ and distance between two vectors: $||\boldsymbol{u} - \boldsymbol{v}|| = \ < \boldsymbol{u} - \boldsymbol{v}, \boldsymbol{u} - \boldsymbol{v} >^{1/2}$. Finally we discussed the angle between vectors and defined orthogonality ($\boldsymbol{u} \perp \boldsymbol{v}$) by $< \boldsymbol{u}, \boldsymbol{v} > \ = 0$.

Earlier in the course, we took the essential properties of vectors in $\mathbb{R}^n$ for a starting point to define more general vector spaces V (*p. 190*). In the same spirit, we now use the essential properties a)-d) of the inner product in $\mathbb{R}^n$ as a guide for inner products in any vector space V. For a vector space V with real scalars, an inner product is a rule $< \boldsymbol{u}, \boldsymbol{v} >$ that produces a scalar for every pair of vectors $\boldsymbol{u}, \boldsymbol{v}$ in V in a way that a) - d) are true. We call such a rule an inner product because it acts like an inner product (in $\mathbb{R}^n$). (*Properties a) - d) are modified slightly when complex scalars are allowed.*) A vector space V with an inner product defined is called an inner product space. Because any inner product "acts just like" the inner product from $\mathbb{R}^n$, many of the theorems we proved about inner products for $\mathbb{R}^n$ are also true in any inner product space. You can look at an introduction to this material in Section 6.7 of the textbook.

Here is a little more detail using one specific example, $C[-\pi, \pi]$ = the vector space of all continuous real-valued functions defined on the interval $[-\pi, \pi]$. We'll call this vector space C for short.

Everything hereafter in these notes is happening in C.

For vectors (functions) f, g in C, define an inner product by

$$< f, g > \ = \int_{-\pi}^{\pi} f(x)g(x)\, dx \quad \text{(a scalar!)}$$

A purely heuristic comment: in Calculus I this integral is defined (roughly) as follows:

- *divide* $[-\pi, \pi]$ *into subintervals of length* $\frac{1}{n}$ *and pick a point* x_i *in each subinterval*
- *form a "Riemann sum"* $\sum f(x_i)g(x_i)\frac{1}{n} = \frac{1}{n}\sum f(x_i)g(x_i)$
- *let* $n \to \infty$*: the integral is the limit of the Riemann sums*

The Riemann sum $\sum f(x_i)g(x_i)$ *resembles the definition for the dot product in* $\mathbb{R}^n$ *: if you imagine* $f(x_i)$ *and* $g(x_i)$ *as "the* x_i *coordinates for the "vectors"* f *and* g*, then* $\sum f(x_i)g(x_i)$ *is analogous to an inner product in* $\mathbb{R}^n$ *: "add up the product of coordinates from* f *and* g*."*

Notice that $< f, g >$ does satisfy a) - d): that is, $< f, g >$ behaves like the inner product in $\mathbb{R}^n$:

a) $< f, g > \; = \; < g, f >$
because $\int_{-\pi}^{\pi} f(x)g(x)dx = \int_{-\pi}^{\pi} g(x)f(x)\,dx$

b) $< f + g, h > \; = \; < f, h > + < g, h >$
because $\int_{-\pi}^{\pi} (f(x) + g(x))h(x)\,dx = \int_{-\pi}^{\pi} f(x)h(x)\,dx + \int_{-\pi}^{\pi} g(x)h(x)\,dx$

c) $< \mathrm{c}f, g > \; = c < f, g >$
because ___?___

d) $< f, f > \; \geq 0$ because $\int_{-\pi}^{\pi} f^2(x)\,dx \geq 0$. And $< f, f > \; = 0$ if and only if $f = \mathbf{0}$ ($=$ the <u>constant function</u> $\mathbf{0}$) *You should try to convince yourself about d). Checking it needs the fact that function f in C are <u>continuous</u>. Why?*

Using this new inner product, we make definitions in C parallel to definitions in $\mathbb{R}^n$:

Define the <u>norm</u> of f by $||f|| = \; < f, f >^{1/2} \; = (\int_{-\pi}^{\pi} f^2(x)\,dx)^{1/2}$

Define the <u>distance between</u> f <u>and</u> g as $||\, f - g \,|| = (\int_{-\pi}^{\pi} (f(x) - g(x))^2\,dx)^{1/2}$

There are certainly other ways to define "distance" between two functions f and g. This way is called the <u>mean square distance</u> between f and g. It's popular with mathematicians and statisticians because it resembles the definition of distance in $\mathbb{R}^n$ as a "square root of a sum of squares" $-$ where "sum" is replaced by "integral." Moreover, $||\, f - g \,||$ behaves nicely, with properties similar to ordinary distance in $\mathbb{R}^n$.

Notice that because of the squaring in the formula, "large differences" $|f(x) - g(x)|$ have more influence than "small differences" in calculating the distance $||f - g||$.

We say f <u>and</u> g <u>are</u> <u>orthogonal</u> if $< f, g > \; = \int_{-\pi}^{\pi} f(x)g(x)\,dx \; = 0$

For example: sin and cos are orthogonal on $[-\pi, \pi]$ because
$< \sin, \cos > \; = \int_{-\pi}^{\pi} (\sin x)(\cos x)\,dx$
$= \frac{1}{2}\int_{-\pi}^{\pi} \sin(2x)\,dx = -\frac{1}{2}\cos(2x)|_{-\pi}^{\pi} = 0$

Most of the tools we developed using inner products in $\mathbb{R}^n$ still work. For example:

For a subspace W of C : we define $W^{\perp} = \{f : \; < f, g > \; = 0$ for all g in $W\}$

<u>Orthogonal Decomposition Theorem</u>: Suppose $f \in C$ and that W is a subspace of C with an <u>orthogonal basis</u> $\{g_1, ..., g_n\}$. Then we can write $f = \widehat{f} + g$ where $\widehat{f} \in W$ and $g \in W^{\perp}$, and $\widehat{f}$ and g are unique.

$\widehat{f}$ is called the projection of f on W, also denoted $\mathrm{proj}_W f$, and

$$\widehat{f} = \frac{<f,g_1>}{<g_1,g_1>}\, g_1 + \ldots + \frac{<f,g_n>}{<g_n,g_n>}\, g_n$$

$\widehat{f}$ is the function in W closest to f, meaning that

$$\|f - \widehat{f}\,\| < \|f - g\| \text{ for all } g \text{ in } W,\ g \neq \widehat{f}$$

If $\{f_1, f_2, \ldots, f_n\}$ is a basis for a subspace W of C, we can convert the basis into an orthogonal basis $\{g_1, \ldots, g_n\}$ using the same Gram Schmidt formulas as in $\mathbb{R}^n$.

Note: Unlike $\mathbb{R}^n$, C is infinite dimensional but that doesn't matter for the results we just listed. What does matter is that the subspace W is finite dimensional: W has a finite (orthogonal) basis $\{g_1, \ldots, g_n\}$.

Some approximations in C: three examples

Example 1 Consider the subspace of C :

$$W = \text{Span}\,\{1,\ \cos x,\ \cos 2x,\ \ldots, \cos nx,\ \sin x,\ \sin 2x, \ldots,\ \sin nx\}$$

Functions g in W are the linear combinations of the functions that span W; these are sometimes called trigonometric polynomials:

$$(*) \quad \boldsymbol{g(x) = c_0 + a_1 \cos x + \ldots + a_n \cos nx + b_1 \sin x + \ldots + b_n \sin nx} \qquad (*)$$

$$\boldsymbol{= c_0 + \sum_{k=1}^{n} a_k \cos kx + \sum_{k=1}^{n} b_k \sin kx}$$

$\text{proj}_W f$ is one such function: it is in W and it is the best approximation to f from W – meaning that if g is in W, then the distance $||f - g||$ is smallest possible when $g = \text{proj}_W f$.

Finding $\text{proj}_W f$ is relatively easy because the functions $1,\ \sin x,\ \sin 2x, \ldots,\ \sin nx,\ \cos x,$ $\cos 2x, \ldots,\ \cos nx$ form an orthogonal basis for W. However, we should check that fact. Doing so involves some integrations that use some trig identities:

$$\sin A \sin B = \tfrac{1}{2}[\cos(A-B) - \cos(A+B)]$$

$$\cos A \cos B = \tfrac{1}{2}[\cos(A-B) + \cos(A+B)]$$

$$\sin A \cos B = \tfrac{1}{2}[\sin(A-B) + \sin(A+B)]$$

$$\cos^2 A = \frac{1 + \cos 2A}{2}$$

$$\sin^2 A = \frac{1 - \cos 2A}{2}$$

- 1 is orthogonal to any of the other functions $\cos kx$ to $\sin kx$:

$$\int_{-\pi}^{\pi} 1 \cdot \cos(kx)\, dx = \tfrac{1}{k}\sin(kx)\Big|_{-\pi}^{\pi} = 0$$
$$\int_{-\pi}^{\pi} 1 \cdot \sin(kx)\, dx = -\tfrac{1}{k}\cos(kx)\Big|_{-\pi}^{\pi} = 0$$

- $\sin kx$ and $\cos kx$ are orthogonal:

$$\int_{-\pi}^{\pi} \sin(kx)\cos(kx)\, dx = \frac{\sin^2(kx)}{2k}\Big|_{-\pi}^{\pi} = 0$$

- finally, for $k \neq m$

$$\int_{-\pi}^{\pi} \sin(mx)\sin(kx)\, dx = \int_{-\pi}^{\pi} \tfrac{1}{2}\left[\cos(m-k)x - \cos(m+k)x\right] dx$$
$$= \left(\frac{\sin(m-k)x}{2(m-k)} - \frac{\sin(m+k)x}{2(m+k)}\right)\Big|_{-\pi}^{\pi} = 0$$

$$\int_{-\pi}^{\pi} \cos(mx)\cos(kx)\, dx = \int_{-\pi}^{\pi} \tfrac{1}{2}\left[\cos(m-k)x + \cos(m+k)x\right] dx$$
$$= \left(\frac{\sin(m-k)x}{2(m-k)} + \frac{\sin(m+k)x}{2(m+k)}\right)\Big|_{-\pi}^{\pi} = 0$$

$$\int_{-\pi}^{\pi} \sin(mx)\cos(kx)\, dx = \int_{-\pi}^{\pi} \tfrac{1}{2}\left[\sin(m-k)x + \sin(m+k)x\right] dx$$
$$= \left(-\frac{\cos(m-k)x}{2(m-k)} - \frac{\cos(m+k)x}{2(m+k)}\right)\Big|_{-\pi}^{\pi} = 0$$

We can also compute two other integrals that we will need:

$$\int_{-\pi}^{\pi} \cos^2 kx \cos kx\, dx = \int_{-\pi}^{\pi} \frac{1+\cos 2kx}{2}\, dx = \left(\frac{x}{2} + \frac{\sin 2kx}{4k}\right)\Big|_{-i}^{\pi} = \pi$$

$$\int_{-\pi}^{\pi} \sin^2 kx \cos kx\, dx = \int_{-\pi}^{\pi} \frac{1-\cos 2kx}{2}\, dx = \left(\frac{x}{2} - \frac{\sin 2kx}{4k}\right)\Big|_{-i}^{\pi} = \pi$$

Notation Alert: If f is any function in C, we can compute $proj_W f$ = the function in W closest to f. As you might recognize, these calculations are closely related to a topic called Fourier analysis. In Fourier analysis, some functions f have what's called a "Fourier transform," denoted by $\widehat{f}$. This is something very different from $\widehat{f} = proj_W f$; so in this example we will use only the notation $proj_W f$ instead of $\widehat{f}$ to avoid possible confusion with the Fourier transform by those who know something about it.)

$\text{proj}_W f$ is a function that looks like (*). We can use the projection formula to determine its coefficients.

$$\text{proj}_W f = \frac{<f(x),1>}{<1,1>}1 + \frac{<f(x),\cos x>}{<\cos x,\cos x>}\cos x + \ldots + \frac{<f(x),\cos nx>}{<\cos nx,\cos nx>}\cos nx$$

$$+ \frac{<f(x),\sin x>}{<\sin x,\sin x>}\sin x + \ldots + \frac{<f(x),\sin nx>}{<\sin nx,\sin nx>}\sin nx$$

The denominators don't depend on f :

$$< 1, 1 > = \int_{-\pi}^{\pi} 1 \cdot 1\, dx = 2\pi$$

$$< \cos kx, \cos kx> = \int_{-\pi}^{\pi} \cos^2 kx\, dx = \pi \qquad (\textit{see calculation above})$$

$$< \sin kx, \sin kx> = \int_{-\pi}^{\pi} \sin^2 kx\, dx = \pi \qquad (\textit{see calculation above})$$

So $\text{proj}_W f = \frac{1}{2\pi} < f(x), 1 > 1 + \frac{1}{\pi} < f(x), \cos x > + \ldots + \frac{1}{\pi} < f(x), \cos nx>$

$$+ \frac{1}{\pi} < f(x), \sin x > + \ldots + \frac{1}{\pi} < f(x), \sin nx>$$

$$= \frac{1}{2}\left(\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cdot 1\, dx\right)1 + \frac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot\cos x\ dx\right)\cos x + \ldots + \frac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot\cos nx\, dx\right)\cos nx$$

$$+ \frac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot\sin x\ dx\right)\sin x + \ldots + \frac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot\sin nx\, dx\right)\sin nx$$

so

$$\text{proj}_W f = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + \sum_{k=1}^{n} b_k \sin kx \qquad (**)$$

where $a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cdot 1\, dx$

and $\begin{cases} a_k = \frac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot\cos kx\, dx\right) & \text{for } 1 \le k \le n \\ b_k = \frac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot\sin nx\, dx\right) & \text{for } 1 \le k \le n \end{cases}$

The coefficients a_k and b_k used here to write $\text{proj}_W f$ are called the Fourier coefficients of f and the "trigonometric series" (**) is the n^{th} Fourier approximation for f.

Fact: If $n \to \infty$, then the approximation error (as measured by our distance function) $\| f - \text{proj}_W f \|) \to 0$. This is called mean square convergence.

*Mean square convergence is **not** equivalent to saying:*

for each $x \in [-\pi, \pi]$, $\lim_{n\to\infty} \text{proj}_W f(x) \to f(x)$ (this is called pointwise convergence)

It's a much harder problem to characterize for which f's and x's this is true.

Concrete Example: Find the n^{th} Fourier approximation for $f(x) = x$ on the interval $[-\pi, \pi]$ and compare the graphs.

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} x \cdot 1\, dx \ = 0$$

$a_k = \frac{1}{\pi}(\int_{-\pi}^{\pi} x\cos kx\, dx = 0$ because $x\cos(kx)$ is an odd function (for an odd function over $[-\pi, \pi]$ the "positive" and "negative" areas between the graph and the x-axis cancel out)

$$b_k = \frac{1}{\pi}(\int_{-\pi}^{\pi} x \cdot \sin kx\, dx) = \frac{1}{\pi}(-x\frac{\cos kx}{k} + \frac{\sin kx}{k^2})\Big|_{-\pi}^{\pi} = \frac{1}{\pi}[-\pi\frac{(-1)^k}{k} - (\pi\frac{(-1)^k}{k})]$$

↑ integration by parts

$$= \frac{-2(-1)^k}{k} = \frac{2(-1)^{k+1}}{k}$$

So for $f(x) = x$ on $[-\pi, \pi]$, we get the approximation

$$x \approx \ b_1\sin x + \ldots + b_n \sin nx \qquad = 2\sin x - 1\sin 2x + \frac{2}{3}\sin 3x + \ldots + \frac{2(-1)^{n+1}}{n}\sin nx$$

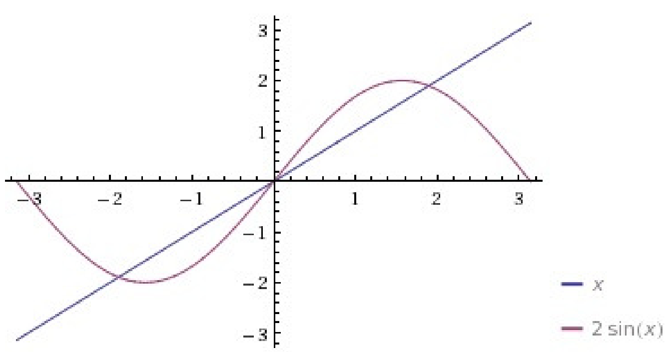
$$= 2\Big(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \ldots + (-1)^{n+1}\frac{\sin nx}{n}\Big)$$

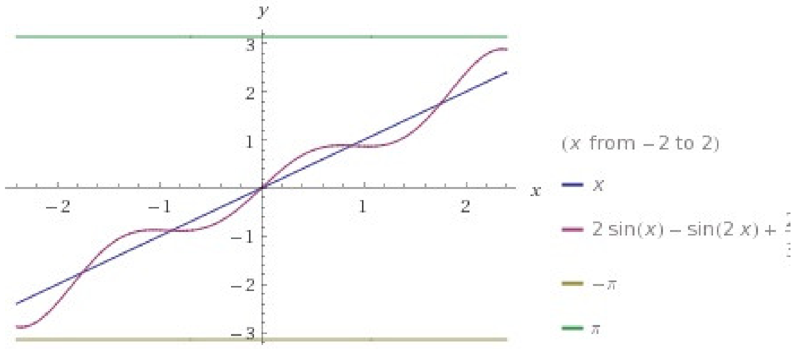
Here are three graphs for comparison, using $n = 1$, 3, and 10:

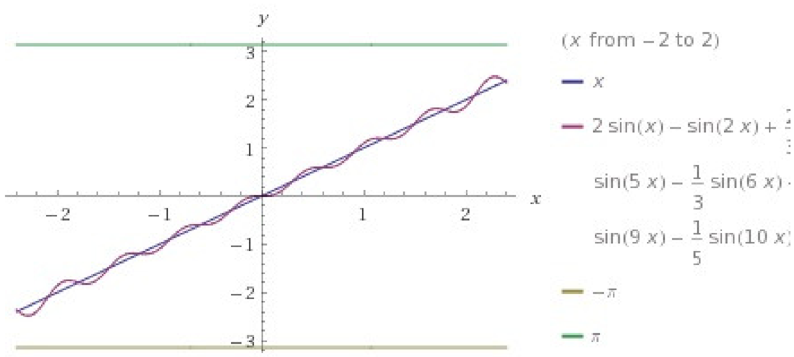
Notice (not clear in the pictures) that every Fourier approximation

$$2\Big(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \ldots + (-1)^{n+1}\frac{\sin nx}{x}\Big)$$

has value 0 at the endpoints $\pm\pi$, but the function $f(x) = x$ is not 0 at the endpoints: at π and $-\pi$, the value of Fourier polynomials do have a limit, 0, but $0 \neq f(\pi) = f(-\pi)$.







Example 2 We are still working in $C = C[-\pi, \pi]$. Let W be the subspace of C containing the polynomials of degree ≤ 5 :

$$W = \text{Span}\,\{1, x, x^2, x^3, x^4, x^5\}$$

Find the polynomial in W closest to the function sin. MATLAB will handle the details for us.

The polynomial we want is $q = \text{proj}_W$ sin; q is the "best approximation from W" to the function sin, meaning that the approximation error

$$||\, q - \sin || = \left(\int_{-\pi}^{\pi} |\, q(x) - \sin x|^2\, dx\right)^{1/2}$$ <u>is smaller than</u>

$$||\, p - \sin || = \left(\int_{-\pi}^{\pi} |\, p(x) - \sin x|^2\, dx\right)^{1/2}$$ for any $p \in W,\ p \neq q$

It's easy to compute $\text{proj}_W \sin$ if we have orthonormal basis for W, so we convert

the standard basis for W $\quad \{v_1, v_2, ..., v_6\} = \{1, x, x^2, ..., x^5\}$
into an orthonormal basis that we'll call $\quad \{e_1, e_2, e_3, ..., e_6\}$

using the Gram-Schmidt Process. Since MATLAB is going to do the work, we will normalize at each step in the process. (*For hand calculations, the arithmetic would be simpler to just use Gram Schmidt to get an orthogonal basis and after Gram Schmidt is completed, then normalize each of those vectors.*)

> *Note: the integrations below were done using MATLAB. Notice that every integration used to find the e_i's is very easy, but that the constants that arise are messy and pile up fast; they can easily lead to errors when the computation is done by hand. Try to compute at least e_1, e_2, e_3 for yourself (with or without computer assistance) to be sure you understand what's going on. The steps are the same as for the usual Gram Schmidt process in $\mathbb{R}^n$.*

We start the process with $v_1 = 1$. But v_1 is not a unit vector in C because $||v_1||^2 = ||1||^2 = \int_{-\pi}^{\pi} 1 \cdot 1 \, dx = 2\pi$. So we normalize and let

$$e_1 = \frac{v_1}{||v_1||} = \frac{1}{||1||} = \frac{1}{\sqrt{2\pi}}$$

For $j = 2, ..., 6$ in turn we use the Gram Schmidt formula, normalizing at each step to get a unit vector:

$$e_j = \frac{v_j - \langle v_j, e_1 \rangle e_1 - \langle v_j, e_2 \rangle e_2 - ... - \langle v_j, e_{j-1} \rangle e_{j-1}}{\| v_j - \langle v_j, e_1 \rangle e_1 - \langle v_j, e_2 \rangle e_2 - ... - \langle v_j, e_{j-1} \rangle e_{j-1} \|}$$

So $$e_2 = \frac{v_2 - \langle v_2, e_1 \rangle e_1}{||v_2 - \langle v_2, e_1 \rangle e_1||} = \frac{x - \left(\int_{-\pi}^{\pi} x \cdot \frac{1}{\sqrt{2\pi}} \, dx\right) \cdot \frac{1}{\sqrt{2\pi}}}{||x - \left(\int_{-\pi}^{\pi} x \cdot \frac{1}{\sqrt{2\pi}} \, dx\right) \cdot \frac{1}{\sqrt{2\pi}}||}$$

$$= \frac{x}{||x||} \qquad (\text{since } \int_{-\pi}^{\pi} x \, dx = 0)$$

$$= \frac{x}{\left(\int_{-\pi}^{\pi} x \cdot x \, dx\right)^{1/2}} = \frac{x}{\frac{1}{3}\pi(6\pi)^{1/2}} = \frac{\sqrt{6}\, x}{2\pi\sqrt{\pi}}$$

Then

$$e_3 = \frac{v_3 - \langle v_3, e_1 \rangle e_1 - \langle v_3, e_2 \rangle e_2}{|| v_3 - \langle v_3, e_1 \rangle e_1 - \langle v_3, e_2 \rangle e_2 ||} = \frac{x^2 - \left(\int_{-\pi}^{\pi} x^2 \cdot e_1 \, dx\right) e_1 - \left(\int_{-\pi}^{\pi} x^2 \cdot e_2 \, dx\right) e_2}{|| x^2 - \left(\int_{-\pi}^{\pi} x^2 \cdot e_1 \, dx\right) e_1 - \left(\int_{-\pi}^{\pi} x^2 \cdot e_2 \, dx\right) e_2 ||}$$

$$= \frac{x^2 - \left(\int_{-\pi}^{\pi} x^2 \cdot \frac{1}{\sqrt{2\pi}} \, dx\right) \cdot \frac{1}{\sqrt{2\pi}} - \left(\int_{-\pi}^{\pi} x^2 \cdot \frac{\sqrt{6}\, x}{2\pi\sqrt{\pi}} \, dx\right) \cdot \frac{\sqrt{6}\, x}{2\pi\sqrt{\pi}}}{|| x^2 - \left(\int_{-\pi}^{\pi} x^2 \cdot \frac{1}{\sqrt{2\pi}} \, dx\right) \cdot \frac{1}{\sqrt{2\pi}} - \left(\int_{-\pi}^{\pi} x^2 \cdot \frac{\sqrt{6}\, x}{2\pi\sqrt{\pi}} \, dx\right) \cdot \frac{\sqrt{6}\, x}{2\pi\sqrt{\pi}} ||}$$

> *Notice that each integration requires no calculus harder than $\int x^n \, dx$.*
> *But already the constants are becoming a headache.*

$$= \quad \ldots \quad = \frac{1}{8}\frac{\sqrt{8}\,\sqrt{45}\,(x^2 - \frac{1}{3}\pi^2)}{\sqrt{\pi^5}}$$

Continuing in this way gives:

$$e_4 = \quad \ldots \quad = \frac{1}{8}\frac{\sqrt{175}\,\sqrt{8}\,(x^3 - \frac{3}{5}\pi^2\,x)}{\sqrt{\pi^7}}$$

$$e_5 = \quad \ldots \quad = \frac{1}{128}\frac{\sqrt{11025}\,\sqrt{128}\left(x^4 - \frac{1}{5}\pi^4 - \frac{6}{7}\pi^2(x^2 - \frac{1}{3}\pi^2)\right)}{\sqrt{\pi^9}},$$ and finally

$$e_6 = \quad \ldots \quad = \frac{1}{128}\frac{\sqrt{128}\,\sqrt{43659}\left(x^5 - \frac{3}{7}\pi^4 x - \frac{10}{9}\pi^2\,(x^3 - \frac{3}{5}\pi^2\,x)\right)}{\sqrt{\pi^{11}}}$$

Then $e_1, \ldots, e_6$ form an orthonormal basis for W. With them, we can use the projection formula to compute

$$q(x) = \text{proj}_W \sin x = \ < \sin x, e_1 > e_1 + \ < \sin x, e_2 > e_2 + \ldots + \ < \sin x, e_6 > e_6$$

$$= (\int_{-\pi}^{\pi} (\sin x) e_1\, dx) e_1 + (\int_{-\pi}^{\pi} (\sin x) e_2\, dx) e_2 + \quad \ldots \quad + (\int_{-\pi}^{\pi} (\sin x) e_6\, dx) e_6$$

$$= (\int_{-\pi}^{\pi} (\sin x) \frac{1}{\sqrt{2\pi}}\, dx) \frac{1}{\sqrt{2\pi}} + \left(\int_{-\pi}^{\pi} (\sin x) \frac{\sqrt{6}\,x}{2\pi\sqrt{\pi}}\, dx\right) \cdot \frac{\sqrt{6}\,x}{2\pi\sqrt{\pi}}$$

$$+ \left(\int_{-\pi}^{\pi} (\sin x) \frac{1}{8}\frac{\sqrt{8}\,\sqrt{45}\,(x^2 - \frac{1}{3}\pi^2)}{\sqrt{\pi^5}}\, dx\right) \cdot \frac{1}{8}\frac{\sqrt{8}\,\sqrt{45}\,(x^2 - \frac{1}{3}\pi^2)}{\sqrt{\pi^5}}$$

$+ \quad \ldots \quad +$ **<u>three more integral terms</u>** corresponding to $e_4,\ e_5,$ and e_6

(*Notice that the integrations needed are now a bit more challenging because they involve terms like* $\int x^n \sin x\; dx$. *But they are manageable, with enough patience, using integration by parts.*)

When all the smoke clears and terms are combined, MATLAB has produced

$$q(x) = \frac{21}{8\pi^{10}}\Big((5\pi^8 - 765\pi^6 + 7425\pi^4)x + (3750\pi^4 - 30\pi^6 - 34650\pi^2)x^3 + (33\pi^4 - 3465\pi^2 + 31185)x^5\Big) \quad (***)$$

If we convert the exact coefficients in (*) to approximate decimal coefficients we have

$$q(x) \approx 0.98786213557467x - 0.15527141063343\, x^3 + 0.00564311797635\, x^5 \qquad (***)$$

Remember that everything in these examples is happening on the interval $[-\pi, \pi]$: on that interval, the polynomial q is the "best fit to the function sin" from among the polynomials in W (the polynomials with degree ≤ 5). Best fit means that

> if p is any polynomial with degree ≤ 5, $p \neq q$.
>
> $||\, q - \sin\, || = (\int_{-\pi}^{\pi} (q(x) - \sin x)^2\, dx)^{1/2} < (\int_{-\pi}^{\pi} (p(x) - \sin x)^2\, dx)^{1/2} = ||\, p - \sin ||$

Example 3 For comparison with $q(x)$, here is a 5th degree polynomial approximation for $\sin x$ that you should already know from Calculus II: the 5th degree Taylor polynomial,

$$T_5(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(v)}(x)}{5!}x^5$$

For $f(x) = \sin x$, this gives

$$\sin x \approx T_5(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}$$

With approximate decimal coefficients,

$$T_5(x) = x - 0.16666666666667x^3 + 0.00833333333333x^5$$ (*compare coefficients with those in* $q(x)$)

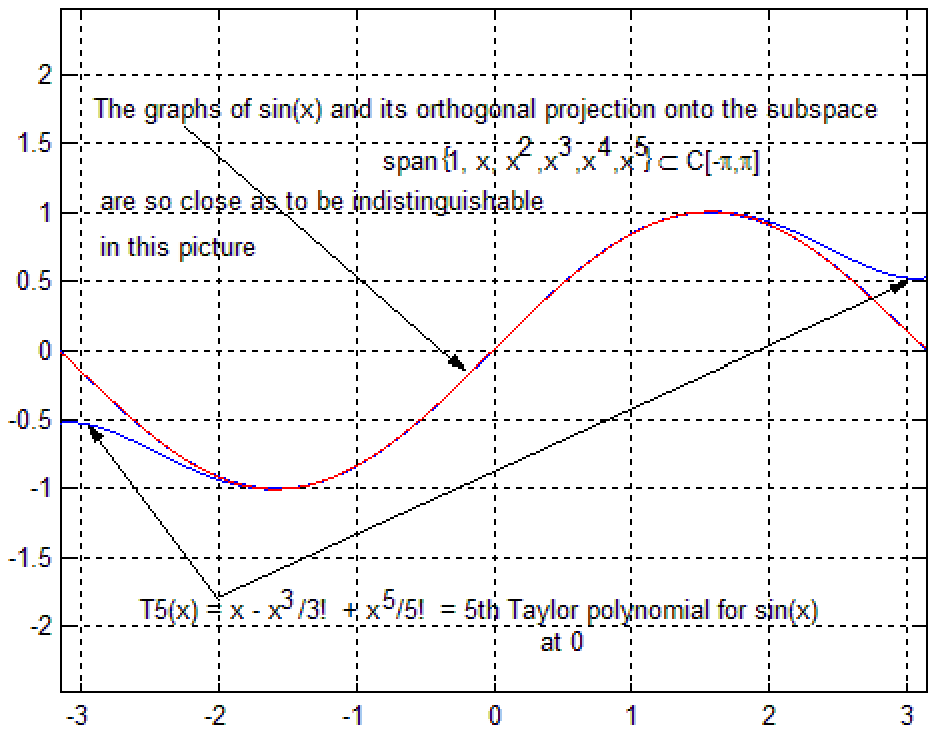
Because $T_5(x)$ is constructed using derivatives of f at 0, it gives a better approximation for $\sin x$ near 0, but the approximation error gets larger and larger error as x moves further from 0.

We can see this in the figure on the next page.

The figure below shows the graphs of $\sin x$, $T_5(x)$ and $q(x)$ on the interval $[-\pi, \pi]$.

Across the whole interval $[-\pi, \pi]$: the graphs of $\sin x$ and $q(x)$ are so close that you can't see a difference between them (at the scale of this graph).

Near 0, you also can't see the difference between $T_5(x)$ and $\sin x$, but that difference becomes clear in the picture as you move closer to the endpoints $\pm\pi$.



The table on the following page illustrates three things (the third is not clear from the graphs above):

i) As we move away from 0 toward $\pm\pi$, the approximation to $\sin x$ using $T_5(x)$ is not as good as the $q(x)$ approximation

ii) Linear algebra gives us the $q(x)$ approximation. It has the advantage that it gives us a good approximation for $\sin x$ over the whole interval $[-\pi, \pi]$.

iii) Near 0, the Taylor polynomial $T_5(x)$ is actually a better approximation than $q(x)$ to $\sin x$ (*in some sense, $T_5(x)$ is actually the best of all 5th degree polynomial approximations to* $\sin x$ *near* 0).

All table values are rounded to 4 significant digits : for example, 0.0000 is not exactly 0

x	$\sin x$	$T_5(x)$	$q(x)$	$\lvert\text{Error}\rvert =$ $\lvert \sin x - T_5(x) \rvert$		$\lvert\text{Error}\rvert$*** $=$ $\lvert \sin x - q(x) \rvert$	
-3.1416	-0.0000	-0.5240	-0.0160	0.5240	← **large T_5**	0.0160	← **smallish q**
-2.9416	-0.1987	-0.5347	-0.1966	0.3361	← **|error|**	0.0021	← **|error|**
-2.7416	-0.3894	-0.5979	-0.3827	0.2084	← **near $-\pi$**	0.0067	← **over**
-2.5416	-0.5646	-0.6891	-0.5600	0.1244	⋮	0.0047	← **the whole**
-2.3416	-0.7174	-0.7884	-0.7169	0.0710		0.0005	← **interval**
-2.1416	-0.8415	-0.8800	-0.8447	0.0385		0.0032	⋮
-1.9416	-0.9320	-0.9516	-0.9372	0.0196		0.0052	
-1.7416	-0.9854	-0.9947	-0.9906	0.0092	↕	0.0052	↕
-1.5416	-0.9996	-1.0035	-1.0032	0.0040		0.0036	
-1.3416	-0.9738	-0.9754	-0.9749	0.0015		0.0011	
-1.1416	-0.9093	-0.9098	-0.9077	0.0005		0.0016	
-0.9416	-0.8085	-0.8086	-0.8047	0.0001		0.0038	
-0.5416	-0.5155	-0.5155	-0.5106	0.0000	← **very**	0.0049	← **but |error|**
-0.3416	-0.3350	-0.3350	-0.3313	0.0000	← **small**	0.0037	← **for q**
-0.1416	-0.1411	-0.1411	-0.1394	0.0000	← **|error|**	0.0017	← **near 0 is small**
0.0584	0.0584	0.0584	0.0577	0.0000	← **near 0**	0.0007	← **but larger**
0.2584	0.2555	0.2555	0.2526	0.0000	← **using**	0.0029	← **than |error|**
0.4584	0.4425	0.4425	0.4380	0.0000	← **T_5**	0.0045	← **for the T_5**
0.6584	0.6119	0.6119	0.6068	0.0000	← **approx**	0.0051	← **approx**
0.8584	0.7568	0.7569	0.7524	0.0001	⋮	0.0044	⋮
1.0584	0.8716	0.8719	0.8690	0.0003		0.0026	
1.2584	0.9516	0.9526	0.9515	0.0010		0.0001	
1.4584	0.9937	0.9964	0.9963	0.0027		0.0026	
1.6584	0.9962	1.0028	1.0009	0.0066		0.0047	
1.8584	0.9589	0.9734	0.9644	0.0145	↕	0.0054	↕
2.0584	0.8835	0.9128	0.8877	0.0293		0.0043	
2.2584	0.7728	0.8282	0.7740	0.0554		0.0012	
2.4584	0.6313	0.7304	0.6283	0.0991		0.0030	← **smallish q**
2.6584	0.4646	0.6336	0.4583	0.1690	⋮	0.0063	← **|error|**
2.8584	0.2794	0.5561	0.2742	0.2767	← **large T_5**	0.0052	← **over**
3.0584	0.0831	0.5204	0.0894	0.4373	← **|error|**	0.0063	← **the whole**
3.1416	9.0000	0.5240	0.0160	0.5240	← **near π**	0.0160	← **interval**

***Note: To be honest, when we picked q as a "best approximation" for sin on $[-\pi, \pi]$, we did so to make $\|\sin - q\| = (\int_{-\pi}^{\pi} (q(x) - \sin x)^2 \, dx)^{1/2}$ as a small as possible; we didn't actually look at "point-by-point" of $\lvert q(x) - \sin x \rvert$, as we do in the table.

For a Fourier approximation $F_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + \sum_{k=1}^{n} b_k \sin kx$ to a function f, we specifically said that having $\| F_N - f \|$ get small might not force $\| F_N(x) - f(x) \|$ get small for particular values of x.

Nevertheless, the table helps to give a feel of how the trigonometric polynomial $q(x)$ is a better approximation to the sin function over the whole interval $[-\pi, \pi]$ than some other polynomial of degree 5 such as $T_5(x)$.