Example 177. Give a basis for the space of all polynomials.

Solution. $1, x, x^2, x^3, \ldots$

Indeed, every polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$ can be written uniquely as a sum of these basis elements. ("can be" = span; "uniquely" = independent)

Comment. The dimension is ∞. But we can make a list of basis elements, which is the "smallest kind of ∞" and is referred to as **countably infinite**. For the space of all functions, no such list can be made.

Just for fun. Let us indicate this difference in infiniteness in a slightly simpler situation: first, the natural numbers $0, 1, 2, 3\ldots$ are infinite but they are countable, because we can make a (infinite but complete) list starting with a first, then a second element and so on (hence, the name "countable"). On the other hand, consider the real numbers between 0 and 1. Clearly, there is infinitely many such numbers. The somewhat shocking fact (first realized by Georg Cantor in 1874) is that every attempt of making a complete list of these numbers must fail because every list will inevitably miss some numbers. Here's a brief indication of how the famous diagonal argument goes: suppose you can make a list, say:

$$\begin{array}{ll} \#1 & 0.111111\ldots \\ \#2 & 0.123456\ldots \\ \#3 & 0.750000\ldots \\ \vdots & \end{array}$$

Now, we are going to construct a new number $x = 0.x_1 x_2 x_3\ldots$ with decimal digits x_i in such a way that the digit x_i differs (by more than 1) from the ith digit of number $\#i$ on our list. For instance, $0.352\ldots$ in our case (for instance, $x_3 = 2$ differs from 0, the 3rd digit of sequence $\#3$). By construction, the number x is missing from the list.

Follow-up. What if we only consider rational numbers in the interval $[0, 1]$? Does the previous argument still apply? Or, can we now make a list?

Comment on fun. The statement "some infinities are bigger than others" nicely captures our observation. It appears in the book *The Fault in Our Stars* by John Green, where it is said by a cranky old author who attributes it to Cantor. Hazel, the main character, later reflects on that statement and compares $[0, 1]$ to $[0, 2]$. Can you explain why that is actually not what Cantor meant$\ldots$?

Comment.

Our choice of inner product $\langle f, g \rangle = \int_a^b f(t) g(t) \mathrm{d}t$ for (square-integrable) functions on $[a, b]$ gives rise to the norm $\| f \| = \left(\int_a^b f(t)^2 \mathrm{d}t \right)^{1/2}$. This is known as the L^2-norm (and often written as $\| f \|_2$).

It is the continuous analog of the usual Euclidean norm $\| \boldsymbol{v} \| = (v_1^2 + v_2^2 + \ldots)^{1/2}$ (known as ℓ^2-norm).

There do exist other norms to measure the magnitude of vectors, such as the ℓ_1-norm $\| \boldsymbol{v} \|_1 = |v_1| + |v_2| + \ldots$ or, more generally, for $p \geqslant 1$, the ℓ_p-norms $\| \boldsymbol{v} \|_p = (|v_1|^p + |v_2|^p + \ldots)^{1/p}$.

Likewise, for functions, we have the L^p-norms $\| f \|_p = \left(\int_a^b f(t)^p \mathrm{d}t \right)^{1/p}$.

Only in the case $p = 2$ do these norms come from an inner product. That's a mathematical (as opposed to geometric) reason why we especially care about that case.

Armin Straub
straub@southalabama.edu

Example 178. Find the best approximation of $f(x) = \sqrt{x}$ on the interval $[0, 1]$ using a function of the form $y = a + bx$.

Important observation. The orthogonal projection of f: $[0, 1] \to \mathbb{R}$ onto $\operatorname{span}\{1, x\}$ is not simply the projection onto 1 plus the projection onto x. That's because 1 and x are not orthogonal:

$$\langle 1, x\rangle = \int_0^1 t\mathrm{d}t = \frac{1}{2} \neq 0.$$

Solution. To find an orthogonal basis for $\operatorname{span}\{1, x\}$, following Gram–Schmidt, we compute

$$x - \begin{pmatrix} \text{projection of} \\ x \text{ onto } 1 \end{pmatrix} = x - \frac{\langle x, 1\rangle}{\langle 1, 1\rangle} 1 = x - \frac{1}{2}.$$

Hence, $1, x - \frac{1}{2}$ is an orthogonal basis for $\operatorname{span}\{1, x\}$.

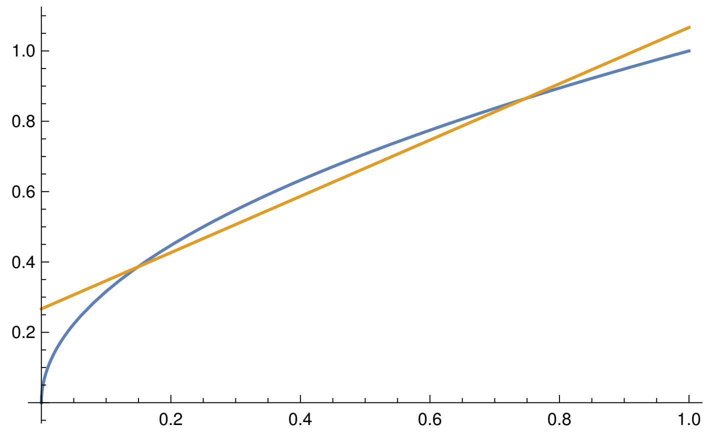
The orthogonal projection of $f\colon [0, 1] \to \mathbb{R}$ onto $\operatorname{span}\{1, x\} = \operatorname{span}\left\{1, x - \frac{1}{2}\right\}$ therefore is

$$\begin{aligned} \frac{\langle f, 1\rangle}{\langle 1, 1\rangle} 1 + \frac{\left\langle f, x - \frac{1}{2}\right\rangle}{\left\langle x - \frac{1}{2}, x - \frac{1}{2}\right\rangle}\left(x - \frac{1}{2}\right) &= \int_0^1 f(t)\mathrm{d}t + \frac{\int_0^1 f(t)\left(t - \frac{1}{2}\right)\mathrm{d}t}{\int_0^1 \left(t - \frac{1}{2}\right)^2 \mathrm{d}t}\left(x - \frac{1}{2}\right) \\ &= \int_0^1 f(t)\mathrm{d}t + (12x - 6)\int_0^1 f(t)\left(t - \frac{1}{2}\right)\mathrm{d}t. \end{aligned}$$

In our case, this best approximation is

$$\begin{aligned} \int_0^1 \sqrt{t}\mathrm{d}t + (12x - 6)\int_0^1 \sqrt{t}\left(t - \frac{1}{2}\right)\mathrm{d}t &= \left[\frac{1}{3/2} t^{3/2}\right]_0^1 + (12x - 6)\left[\frac{1}{5/2} t^{5/2} - \frac{1}{2}\frac{1}{3/2} t^{3/2}\right]_0^1 \\ &= \frac{2}{3} + (12x - 6)\left(\frac{2}{5} - \frac{1}{3}\right) \\ &= \frac{4}{5}\left(x + \frac{1}{3}\right). \end{aligned}$$

The plot below confirms how good this linear approximation is (compare with the previous example):



Armin Straub
straub@southalabama.edu

Orthogonal polynomials

Let us think about the space of all polynomials (with real coefficients). On that space, we consider the dot product

$$\langle p_1, p_2 \rangle = \int_{-1}^{1} p_1(t) p_2(t) \mathrm{d}t. \quad (1)$$

Comment. That dot product is useful if we are thinking about the polynomials as functions on $[-1, 1]$ only. You can, of course, consider any other interval and you will obtain a shifted version of what we get here.

Example 179. Are $1, x, x^2, \ldots$ orthogonal (with respect to the inner product (1))?

Solution. Since $\langle x^r, x^s \rangle = \int_{-1}^{1} t^r t^s \mathrm{d}t = \int_{-1}^{1} t^{r+s} \mathrm{d}t$, we find that $\langle x^r, x^s \rangle = \begin{cases} \frac{2}{r+s+1}, & \text{if } r+s \text{ is even,} \\ 0, & \text{otherwise.} \end{cases}$

Hence, if $r+s$ is odd, then the monomials x^r and x^s are orthogonal. On the other hand, if $r+s$ is even, then x^r and x^s are not orthogonal.

Example 180. Use Gram-Schmidt to produce an orthogonal basis $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots$ for the space of polynomials with the dot product (1). Compute $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4$.

Instead of normalizing these polynomials, **standardize** them so that $\boldsymbol{p}_n(1) = 1$.

Solution. We construct an orthogonal basis $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots$ from $1, x, x^2, \ldots$ as follows:

- Starting with 1, we find $\boldsymbol{p}_0(x) = 1$.

  For future reference, let us note that $\|\boldsymbol{p}_0\|^2 = \int_{-1}^{1} 1 \mathrm{d}x = 2$.

- Starting with x, Gram–Schmidt produces $x - \begin{pmatrix} \text{projection of} \\ x \text{ onto } \boldsymbol{p}_0 \end{pmatrix} = x - \frac{\langle x, \boldsymbol{p}_0 \rangle}{\langle \boldsymbol{p}_0, \boldsymbol{p}_0 \rangle} \boldsymbol{p}_0 = x - \int_{-1}^{1} t \mathrm{d}t = x.$

  Again, that's already standardized, so that $\boldsymbol{p}_1(x) = x$.

  Comment. The previous problem already told us that x is orthogonal to 1.

  For future reference, let us note that $\|\boldsymbol{p}_1\|^2 = \int_{-1}^{1} t^2 \mathrm{d}t = \frac{2}{3}$.

- Starting with x, Gram–Schmidt produces $x^2 - \begin{pmatrix} \text{projection of } x^2 \\ \text{onto span}\{\boldsymbol{p}_0, \boldsymbol{p}_1\} \end{pmatrix} = x^2 - \frac{\langle x^2, \boldsymbol{p}_0 \rangle}{\langle \boldsymbol{p}_0, \boldsymbol{p}_0 \rangle} \boldsymbol{p}_0 - \frac{\langle x^2, \boldsymbol{p}_1 \rangle}{\langle \boldsymbol{p}_1, \boldsymbol{p}_1 \rangle} \boldsymbol{p}_1$
  $= x^2 - \frac{1}{2} \int_{-1}^{1} t^2 \mathrm{d}t - \frac{x}{2/3} \int_{-1}^{1} t^3 \mathrm{d}t = x^2 - \frac{1}{3}.$

  Hence, standardizing, $\boldsymbol{p}_2(x) = \frac{1}{2}(3x^2 - 1)$.

  Comment. The previous problem told us that x^2 is orthogonal to x (but not to 1).

- Continuing, we find $\boldsymbol{p}_3(x) = \frac{1}{2}(5x^3 - 3x)$ and $\boldsymbol{p}_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$.

Comment. These famous polynomials are known as the **Legendre polynomials**. The Legendre polynomial $\boldsymbol{p}_n$ is an even function if n is even, and an odd function if n is odd (can you explain why?!).

An explicit formula is $\boldsymbol{p}_n(x) = 2^{-n} \sum_{k=0}^{n} \binom{n}{k}^2 (x+1)^k (x-1)^{n-k}$.

For instance, $\boldsymbol{p}_2(x) = \frac{1}{4}((x-1)^2 + 2^2(x-1)(x+1) + (x+1)^2) = \frac{1}{2}(3x^2 - 1)$

https://en.wikipedia.org/wiki/Legendre_polynomials

Comment. Legendre polynomials are an example of **orthogonal polynomials**. Each choice of dot product gives rise to a family of such orthogonal polynomials.

https://en.wikipedia.org/wiki/Orthogonal_polynomials

Comment. It is also particularly natural to consider the dot product (1), where the integral is from 0 to 1. In that case, we obtain what's known as the shifted Legendre polynomials $\tilde{\boldsymbol{p}}_n(x) = \boldsymbol{p}_n(2x - 1)$.

A **Fourier series** for a function $f(x)$ is a series of the form

$$f(x) = a_0 + a_1\cos(x) + b_1\sin(x) + a_2\cos(2x) + b_2\sin(2x) + \cdots$$

You may have seen Fourier series in other classes before. Our goal here is to tie them in with what we have learned about orthogonality.

In these other classes, you would have seen formulas for the coefficients a_k and b_k. We will see where those come from.

Observe that the right-hand side combination of cosines and sines is 2π-periodic.

Let us consider (nice) functions on $[0, 2\pi]$.

Or, equivalently, functions that are 2π-periodic.

We know that a natural inner product for that space of functions is

$$\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)\mathrm{d}t.$$

Example 181. Show that $\cos(x)$ and $\sin(x)$ are orthogonal (in that sense).

Solution. $\langle \cos(x), \sin(x) \rangle = \int_0^{2\pi} \cos(t)\sin(t)\mathrm{d}t = \left[\frac{1}{2}(\sin(t))^2\right]_0^{2\pi} = 0$

In fact:

All the functions $1, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots$ are orthogonal to each other!

Moreover, they form a basis in the sense that every other (nice) function can be written as a (infinite) linear combination of these basis functions.

Example 182. What is the norm of $\cos(x)$?

Solution. $\langle \cos(x), \cos(x) \rangle = \int_0^{2\pi} \cos(t)\cos(t)\mathrm{d}t = \pi$

Why? There's many ways to evaluate this integral. For instance:

- integration by parts
- using a trig identity
- here's a simple way:
  - $\int_0^{2\pi} \cos^2(t)\mathrm{d}t = \int_0^{2\pi} \sin^2(t)\mathrm{d}t$ (cos and sin are just a shift apart)
  - $\cos^2(t) + \sin^2(t) = 1$
  - So: $\int_0^{2\pi} \cos^2(t)\mathrm{d}t = \frac{1}{2}\int_0^{2\pi} 1\,\mathrm{d}x = \pi$

Hence, $\cos(x)$ is not normalized. It has norm $\|\cos(x)\| = \sqrt{\pi}$.

Similarly. The same calculation shows that $\cos(kx)$ and $\sin(kx)$ have norm $\sqrt{\pi}$ as well.

Armin Straub
straub@southalabama.edu

Example 183. How do we find, say, b_2?

Solution. Since the functions $1, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots$, the term $b_2 \sin(2x)$ is the orthogonal projection of $f(x)$ onto $\sin(2x)$.

In particular, $b_2 = \frac{\langle f(x), \sin(2x) \rangle}{\langle \sin(2x), \sin(2x) \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(2t) \mathrm{d}x$.

In conclusion:

A (nice) $f(x)$ on $[0, 2\pi]$ has the Fourier series

$$f(x) = a_0 + a_1 \cos(x) + b_1 \sin(x) + a_2 \cos(2x) + b_2 \sin(2x) + \cdots$$

where

$$a_k = \frac{\langle f(x), \cos(kx) \rangle}{\langle \cos(kx), \cos(kx) \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(kt) \mathrm{d}t,$$

$$b_k = \frac{\langle f(x), \sin(kx) \rangle}{\langle \sin(kx), \sin(kx) \rangle} = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(kt) \mathrm{d}t,$$

$$a_0 = \frac{\langle f(x), 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{2\pi} \int_0^{2\pi} f(t) \mathrm{d}t.$$