Example 163. What is the orthogonal projection of $f: [a, b] \rightarrow \mathbb{R}$ onto the space of constant functions (that is, $\operatorname{span}\{1\}$)?

Solution. The orthogonal projection of $f: [a, b] \rightarrow \mathbb{R}$ onto $\operatorname{span}\{1\}$ is

$$\frac{\langle f, 1\rangle}{\langle 1, 1\rangle}1 = \frac{\int_a^b f(t)1\mathrm{d}t}{\int_a^b 1^2\mathrm{d}t} = \frac{1}{b-a}\int_a^b f(t)\mathrm{d}t.$$

This is the average of $f(x)$ on $[a, b]$.

Comment. Makes perfect sense, doesn't it? Intuitively, the best approximation of a function by a constant should indeed be the one where the constant is the average.

Orthogonal polynomials

Example 164. Compute an orthogonal basis for the subspace $\operatorname{span}\{1, x, x^2, x^3\}$ (polynomials of degree at most 3) of functions on $[0, 1]$.

Solution. Our inner product is $\langle p_1, p_2\rangle = \int_0^1 p_1(t)p_2(t)\mathrm{d}t$. To find an orthogonal basis, we use Gram–Schmidt:

$$\begin{aligned}
q_1 &= 1\\
q_2 &= x - \frac{\langle x, q_1\rangle}{\langle q_1, q_1\rangle}q_1 = x - \frac{\langle x, 1\rangle}{\langle 1, 1\rangle}1 = x - \frac{1}{2}\\
q_3 &= x^2 - \frac{\langle x^2, q_1\rangle}{\langle q_1, q_1\rangle}q_1 - \frac{\langle x^2, q_2\rangle}{\langle q_2, q_2\rangle}q_2 = x^2 - \frac{\langle x^2, 1\rangle}{\langle 1, 1\rangle}1 - \frac{\left\langle x^2, x - \frac{1}{2}\right\rangle}{\left\langle x - \frac{1}{2}, x - \frac{1}{2}\right\rangle}\left(x - \frac{1}{2}\right)\\
&= x^2 - \frac{\frac{1}{3}}{1}1 - \frac{\frac{1}{12}}{\frac{1}{12}}\left(x - \frac{1}{2}\right) = x^2 - x + \frac{1}{6}\\
q_4 &= x^3 - \frac{\langle x^3, q_1\rangle}{\langle q_1, q_1\rangle}q_1 - \frac{\langle x^3, q_2\rangle}{\langle q_2, q_2\rangle}q_2 - \frac{\langle x^3, q_3\rangle}{\langle q_3, q_3\rangle}q_3 = \ldots = x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}
\end{aligned}$$

The polynomials 1, $x - \frac{1}{2}$, $x^2 - x + \frac{1}{6}$, $x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}$ form an orthogonal basis for the space of polynomials of degree at most 3.

Comment. Of course, we could keep going by next including $x^4, x^5, \ldots$ Up to scaling, the resulting polynomials are known as the **shifted Legendre polynomials** and they are an example of a family of **orthogonal polynomials**. They are important, for instance, in approximating more complicated functions using polynomials (see the previous example, for instance).

Homework. Fill in the details of the computation for q_4 (maybe using Sage for support). For instance, here is how to compute $\int_0^1 t^2\left(t - \frac{1}{2}\right)\mathrm{d}t$ using Sage:

```
>>> t = var('t')

>>> integral(t^2*(t-1/2), t, 0, 1)
```

$$\frac{1}{12}$$

In the literature, the interval $[0, 1]$ is often replaced with the interval $[-1, 1]$ (because of the symmetry). If we proceed as above, then the resulting orthogonal polynomials are known as the **Legendre polynomials**. In the case of the interval $[-1,1]$, we consider the space of all polynomials (with real coefficients) together with the dot product

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

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 165. 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 166. 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 $\tilde{\boldsymbol{p}}_n(x) = \boldsymbol{p}_n(2x-1)$. Compute the first few and compare with Example 164.

Comment on other norms. 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.

Example 167. 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 are 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, increasing digits by 2 (with 8 and 9 becoming 0 and 1) we get $0.342\ldots$ 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.

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...?