

**Problem 1.** As you know, one way to approximate a function  $f$  of one variable is to replace it by its tangent line at some point of interest, or by the “best fitting” parabola at this point (these approximations correspond to using the first- or second-order Taylor polynomial of the function  $f$  at this point). This type of approximation, however, works very well only near this point, and can be very inaccurate over an entire *interval*.

One way to approximate a function  $f$  (of one variable) on an entire interval is the following. Choose some class of functions  $\mathcal{H}$ , say all linear functions. Then look for a function  $h$  from this class  $\mathcal{H}$  for which the “distance” between  $f$  and  $h$  is the smallest possible. The “distance” – which is usually called “error” – can be defined in many different ways. If we want to approximate  $f$  by a function  $h \in \mathcal{H}$  on the interval  $[a, b]$ , and we want  $|f(x) - h(x)|$  to be small for all  $x \in [a, b]$ , then an appropriate definition for the “error” would be  $E_\infty := \max_{x \in [a, b]} |f(x) - h(x)|$ . Another choice is to

minimize  $E_1 := \int_a^b |f(x) - h(x)| dx$ , but the expressions for  $E_\infty$  and  $E_1$  cause technical difficulties if one tries to use them in practice. The most convenient for numerical purposes expression for the error is

$$E_2 := \int_a^b [f(x) - h(x)]^2 dx ,$$

which we will use below.

In this problem you will find the best approximation of the function  $f(x) = x^3$  by a linear function,  $h_{\mu, \nu}(x) := \mu x + \nu$ , over the interval  $[0, 1]$  if the “error” is given by the integral

$$E_f(\mu, \nu) := \int_0^1 [f(x) - h_{\mu, \nu}(x)]^2 dx . \quad (1)$$

In other words, you have to choose the values of the constants  $\mu$  and  $\nu$  that minimize the error  $E_f(\mu, \nu)$  given by (1).

*Hint:* Here is a useful fact:  $\int_0^1 [x^3 - (\mu x + \nu)]^2 dx = \frac{1}{7} - \frac{2}{5} \mu + \frac{1}{3} \mu^2 - \frac{1}{2} \nu + \mu \nu + \nu^2$  .

**Problem 2.** The result of Problem 1 can be obtained in a different way. Think of functions defined on the interval  $[a, b] \subseteq \mathbb{R}$  as elements of a vector space; if all functions are allowed, this vector space is infinite-dimensional. Let us endow this vector space with the inner product

$$\langle f, g \rangle := \int_a^b f(x) g(x) dx . \quad (2)$$

Let  $V_n$  be the  $(n + 1)$ -dimensional space of polynomials defined on the interval  $[0, 1]$ , of degree no more than  $n$ , endowed with the inner product (2) (with  $a = 0$ ,  $b = 1$ ). Then the problem of best approximation of the polynomial  $f(x) = x^3$  by a linear function can be thought of as the problem of finding the vector  $h \in V_1$  (this is the desired linear function, and  $V_1$  is the set that was denoted by  $\mathcal{H}$  in Problem 1) that is closest to the “vector”  $P$  in the sense that  $\|f - h\|_2^2 := \langle f - h, f - h \rangle$  has the smallest possible value. (Think about the connection between  $\|f - h\|_2^2$  and the “error” (1)).

Another way of thinking about this is to visualize  $V_1$  (all linear polynomials of degree no more than 1) as a subspace of the space  $V_1$  where the function  $f(x) = x^3$  “lives”. Then  $\|f - h\|_2$  is exactly the distance from the “end” of  $f$  to  $V_1$ , and  $h \in V_1$  is the orthogonal projection of  $f$  onto  $V_1$ . Let us think like this and build an orthogonal (with respect to the inner product (2)) basis of  $V_3$  in which the first two vectors span the subspace  $V_1$ . Let  $\phi_0, \phi_1, \dots, \phi_n$  be a basis of the vector space  $V_n$  such that:

- $\phi_k$  is a polynomial of degree  $k$ ;
- the polynomials  $\phi_j$  and  $\phi_k$  are orthogonal (with respect to (2)), i.e.,  $\langle \phi_j, \phi_k \rangle = 0$  if  $j \neq k$ .

Then  $\{\phi_0, \phi_1, \phi_2, \phi_3\}$  form an orthogonal basis of  $V_3$ , while  $\{\phi_0, \phi_1\}$  is a basis of the subspace  $V_1$ . The basis of  $V_3$  it is easy to construct by starting from  $\phi_0$  and then recursively constructing the other  $\phi_j$ ’s using the Gram-Schmidt process. I did this and obtained the following expressions:

$$\begin{aligned}\phi_0(x) &:= 1, \\ \phi_1(x) &:= x - \frac{1}{2}, \\ \phi_2(x) &:= x^2 - x + \frac{1}{6}, \\ \phi_3(x) &:= x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}.\end{aligned}\tag{3}$$

Note that the basis  $\phi_j$  is not normalized.

- Write the polynomial  $f$  (defined by  $f(x) = x^3$ ) as a linear combination of the functions  $\phi_j$  (3).
- Use geometric arguments to find the vector  $h \in V_1$  that minimizes the error (1). Please explain your reasoning in detail.
- Demonstrate that your answer in part (b) is the same as your result in Problem 1.