

## MODULE 5

---

### Topics: Best approximation and orthogonal projections

---

It is a common process in the application of mathematics to approximate a given vector by a vector in a specified subspace. For example, a linear system like

$$Ax = b$$

which does not have a solution, may be approximated by a linear system

$$Ax = b'$$

for which there does exist a solution and where the vector  $b'$  is chosen so that it is “close” to  $b$ . This situation will be discussed later in connection with the least squares solution of linear systems. Another common example is the approximation of a given function  $f$  defined on some interval, say  $(-\pi, \pi)$ , in terms of a trigonometric sum like

$$f(t) \sim \alpha_0 + \sum_{j=1}^N \alpha_j \cos jt + \sum_{j=1}^N \beta_j \sin jt$$

which leads to the concept of Fourier series. As we shall discover,  $b'$  and the trigonometric sum will be a “best” approximation to  $b$  and  $f$ , respectively.

**Definition:** Given a vector space  $V$  with norm  $\| \cdot \|$  and a subspace  $M \subset V$  the best approximation of a given vector  $x \in V$  in the subspace  $M$  is a vector  $\hat{m}$  which satisfies

$$\|x - \hat{m}\| \leq \|x - m\| \quad \text{for all } m \in M.$$

Note that the best approximation in this definition is tied to a norm. Changing the norm will usually change the element  $\hat{m}$ . The choice of norm is dictated by the application or the desire to compute easily the best approximation.

To illustrate that the best approximation can be easy or hard to find depending on the choice of norm consider the following simply stated problem:

Find the best approximation to the vector  $(1, 2, 3)$  in the subspace

$$M = \text{span}\{(3, 2, 1)\} \subset \mathbb{R}_3.$$

i) when  $\|x\| = \|x\|_\infty$

- ii) when  $\|x\| = \|x\|_1$
- iii) when  $\|x\| = \|x\|_2$ .

**Answer:**

- i) Since  $m = \alpha(3, 2, 1)$  for  $\alpha \in (-\infty, \infty)$  the problem is to find an  $\alpha$  which minimizes the expression

$$\|(1, 2, 3) - \alpha(3, 2, 1)\|_\infty = \max\{|1 - 3\alpha|, |2 - 2\alpha|, |3 - \alpha|\} \equiv f(\alpha).$$

If one plots  $f(\alpha)$  vs.  $\alpha$  one finds that it has a minimum at  $\hat{\alpha} = 1$  with  $f(1) = 2$ . Note that this  $\hat{\alpha}$  cannot be found with calculus because the function is not differentiable. Hence the best approximation in  $M$  in this norm is  $\hat{m} = (3, 2, 1)$ .

- ii)  $\|(1, 2, 3) - \alpha(3, 2, 1)\|_1 = |1 - 3\alpha| + |2 - 2\alpha| + |3 - \alpha| \equiv f(\alpha)$ . The function  $f$  is piecewise linear and constant on the interval  $[1/3, 1]$  where it assumes its minimum of  $f(1/3) = 4$ . Hence the best approximation is not unique.  $\hat{m}$  may be chosen to be  $\alpha(3, 2, 1)$  for any  $\alpha \in [1/3, 1]$ .

- iii) Since the Euclidean norm involves square roots it is usually advantageous to minimize the square of the norm rather than the norm itself. Thus,

$$\|(1, 2, 3) - \alpha(3, 2, 1)\|_2^2 = (1 - 3\alpha)^2 + (2 - 2\alpha)^2 + (3 - \alpha)^2 \equiv f(\alpha)$$

is minimized where  $f'(\alpha) = 0$ . A simple calculation shows that  $\alpha = 5/7$  so that  $\hat{m} = 5/7(3, 2, 1)$ . If we set  $v = x - \hat{m}$  then we find by direct calculation that  $v \cdot (3, 2, 1) = 0$ , so that  $v$  is orthogonal to  $M$ . Hence  $\hat{m}$  is the orthogonal projection  $Px$  of  $x$  onto  $M$ . According to Module 4 we can calculate the projection as

$$Px = \alpha(3, 2, 1)$$

where

$$\alpha = \frac{(1, 2, 3) \cdot (3, 2, 1)}{(3, 2, 1) \cdot (3, 2, 1)} = \frac{5}{7}$$

which shows that the best approximation is obtainable also without calculus in this case.

The next theorem shows that if the norm is derived from an inner product then the best approximation always is the orthogonal projection.

**Theorem:** Let  $V$  be a vector space (real or complex) with inner product  $\langle \cdot, \cdot \rangle$  and norm  $\| \cdot \| = (\langle \cdot, \cdot \rangle)^{1/2}$ . Let  $M = \text{span}\{x_1, \dots, x_n\} \subset V$  be a subspace of dimension  $n$ . Given  $y \in V$  then  $\hat{m} \in M$  is the best approximation to  $y$  in  $M$  if and only if  $\hat{m}$  is the orthogonal projection  $Py$  of  $y$  onto  $M$ .

**Proof.** Let us show first that  $Py$  is a best approximation, i.e., that

$$\|y - Py\| \leq \|y - m\| \quad \text{for all } m \in M.$$

Let  $m$  be arbitrary in  $M$ . Then  $m = Py + (m - Py)$  and

$$\begin{aligned} \|y - m\|^2 &= \langle y - Py - (m - Py), y - Py - (m - Py) \rangle \\ &= \langle y - Py, y - Py \rangle - \langle m - Py, y - Py \rangle \\ &\quad - \langle y - Py, m - Py \rangle + \langle m - Py, m - Py \rangle. \end{aligned}$$

But by definition of the orthogonal project  $y - Py$  is orthogonal to  $M$ , while  $m - Py \in M$ . Hence the two middle terms on the right drop out and thus  $\|y - m\|^2 = \|y - Py\|^2 + \|m - Py\|^2$  so that

$$\|y - Py\| \leq \|y - m\| \quad \text{for all } m \in M.$$

Conversely, suppose that  $\hat{m}$  is the best approximation to  $y$  in  $M$ . For arbitrary but fixed  $m \in M$  and real  $t$  define

$$g(t) = \langle y - (\hat{m} + tm), y - (\hat{m} + tm) \rangle.$$

Since  $\hat{m} + tm$  is an element of  $M$  and  $\hat{m}$  is the best approximation it follows that  $g$  has a minimum at  $t = 0$ . Hence necessarily  $g'(0) = 0$ . Differentiation shows that

$$g'(0) = -[\langle m, y - \hat{m} \rangle + \langle y - \hat{m}, m \rangle] = -2\text{Re}\langle m, y - \hat{m} \rangle = 0.$$

Since  $m$  is arbitrary in  $M$  this implies that  $\langle m, y - \hat{m} \rangle = 0$  so that  $\hat{m}$  satisfies the definition of an orthogonal projection. Its uniqueness guarantees that  $\hat{m} = Py$ . Looking back at the examples of Module 4 we see that the function  $Pt^3 \equiv \frac{3}{5}t$  is the best approximation to the function  $f(t) \equiv t^3$  in the sense that

$$\|t^3 - \frac{3}{5}t\|_2^2 \equiv \int_{-1}^1 (t^3 - \frac{3}{5}t)^2 dt \leq \int_{-1}^1 (t^3 - P_2(t))^2 dt$$

for any other polynomial  $P_2(t)$  of degree  $\leq 2$ . In other words,  $Pt^3 \equiv \frac{3}{5}t$  is the best polynomial approximation of degree  $\leq 2$  to the function  $f(t) \equiv t^3$  in the mean square sense.

## Module 5 - Homework

- 1) Let  $y = (1, 2, 3, 4)$  and  $M = \text{span}\{(4, 3, 2, 1)\}$ . Find the best approximation  $\hat{m} \in M$  to  $y$
- in the  $\|\cdot\|_1$  norm
  - in the  $\|\cdot\|_2$  norm
  - in the  $\|\cdot\|_\infty$  norm.
- 2) Let  $M$  be the plane in  $\mathbb{R}_3$  given by

$$3x - 2y + z = 0$$

Find the orthogonal projection of the unit vector  $\hat{e}_1$  onto  $M$  when

- the inner product is the dot product
- the inner product is

$$\langle x, y \rangle = Ax \cdot y$$

where

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

(You may assume at this point without further checking that  $\langle x, y \rangle$  is indeed an inner product).

- 3) Compute the best approximation in the  $L_2(-\pi, \pi)$  sense of the function

$$H(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

in terms of the functions  $\{\sin nt\}_{n=1}^N$  and  $\{\cos nt\}_{n=0}^N$  where  $N > 0$  is some integer.

- 4) Let  $0 = t_0 < t_1 < \dots < t_N = 1$  where  $t_i = i\Delta t$  and  $\Delta t = 1/N$ . For each  $i = 0, 1, \dots, N$  define on  $[0, 1]$  the function

$$\phi_i(t) = \begin{cases} \frac{(t-t_{i-1})}{\Delta t} & t \in [t_{i-1}, t_i) \\ \frac{(t_{i+1}-t)}{\Delta t} & t \in [t_i, t_{i+1}) \\ 0 & \text{otherwise.} \end{cases}$$

- For  $N = 4$  plot

$$\phi_0(t) \quad \text{and} \quad \phi_3(t)$$

Let

$$M = \text{span}\{\phi_0(t), \phi_1(t), \phi_2(t), \phi_3(t), \phi_4(t)\}$$

We shall consider  $M$  as a subspace of  $L_2[0, 1]$  with inner product

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt$$

- ii) Compute and plot the orthogonal projection  $Pf$  of the function  $f(t) \equiv 1 + t$  onto  $M$ .
- iii) Compute and plot the orthogonal projection  $Pf$  of the function  $f(t) \equiv t^2$  onto  $M$ .