CHAPTER 4

Best approximation

1. Gauss and least squares

Mercury, Venus, Earth, Mars, Jupiter and Saturn were the only planets known in ancient times. After the invention of the telescope by Galileo in 1609, astronomers made rapid progress; with improved observational data, Kepler and Newton were able to formulate their famous laws and thus provided an essentially correct mathematical model of the solar system. The German astronomers Titius (1766) and Bode (1778) remarked an intriguing apparent pattern in the mean distances of the planets from the sun. *Bode's law* relates the distances to a simple number sequence obtained as follows.

- Begin with sequence

$$0, 3, 6, 12, 24, 48, 96, 192, \ldots$$

- Add 4 to each number in the sequence and then divide by 10.

It can be seen from Table 4.1 that the resulting sequence is very close to the distribution of the mean distances of the planets measured in *astronomical units*. A seventh planet, Uranus, was discovered in 1781 and, again, its mean distance agrees well with the mysterious law. Astronomers therefore began to search for a new planet between the orbits of Mars and Jupiter which would account for the fifth term in the Titius–Bode sequence.

For many years, Gauss had been pondering on a related problem, namely that of deducing the orbit of a celestial object from observations of its position "not embracing a great period of time" and, as is the case in practice, contaminated by errors. In the preface to his book on planetary motion [**5**], Gauss writes:

> "Some ideas occured to me in the month of September of the year 1801... For just about this time the report of a new planet , discovered on the first day of January of that year with the telescope at Palermo, was the subject of universal conversation; and soon afterwards the observations made by that distinguished astronomer PIAZZI from the above date to the eleventh of February were published. Nowhere in the annals of astronomy do we meet with so great an opportunity, and

Planet	Mercury	Venus	Earth	Mars	*Ceres*	Jupiter	Saturn	*Uranus*
Actual	0.39	0.72	1.00	1.52	2.77	5.20	9.54	19.19
Predicted	0.4	0.7	1.0	1.6	2.8	5.2	10.0	19.6

TABLE 4.1. Mean distance from the Sun (in astronomical units) and the prediction from Bode's law. Ceres and Uranus were found a few years after the law was formulated.

> a greater one could hardly be imagined, for showing most strikingly the value of this problem, than in the crisis and urgent necessity, when all hope of discovering in the heavens this planetary atom, among innumerable small stars after the lapse of nearly a year, rested solely upon a sufficiently approximate knowledge of its orbit to be based upon these very few observations. Could I ever have found a more seasonable opportunity to test the value of my conceptions, than now in employing them for the determination of the orbit of the planet Ceres, which during these forty-one days had described a geocentric arc of only 3°, and after the lapse of a year must be looked for in a region of the heavens very remote from that in which it was last seen? The first application of the method was made in the month of October, 1801, and the first clear night, when the planet was sought for[1] as directed by the numbers deduced from it, restored the fugitive to observation. Three other new planets, subsequently discovered, furnished new opportunities for examining and verifying the efficiency and generality of the method."

Gauss' solution proceeded roughly as follows. Firstly, since the path of Ceres must, by Kepler's laws, be elliptical, the problem reduces to finding a couple of parameters in the general equation for an elliptical orbit. For the sake of clarity, let us collect these unknown parameters into a vector

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$$

and write

$$R(t, x, \lambda) = 0$$

for the equation of the orbit. Here the variables t and x may be thought of as time and position respectively. It is assumed that some values

$$(t_1, x_1^\star), (t_2, x_2^\star), \ldots, (t_N, x_N^\star)$$

have been obtained by direct observation or measurement. In practice, N is usually much larger than k. Set

$$R_n(\lambda) = R(t_n, x_n^\star, \lambda), \quad 1 \leq n \leq N.$$

$R_n(\lambda)$ is called the *nth residual.* If the data were *exact* and the orbit model correct, then there would exist a value of λ such that all the residuals vanish, i.e.

$$\forall\ 1 \leq n \leq N, \quad R_n(\lambda) = 0.$$

Unfortunatly, due to errors in measurement, these data are inherently inexact and so it would be naive to hope that a value for the parameter vector λ could be found that precisely matches it. Instead, one has to be content with parameter values that *minimise*, in some sense, the discrepancy between the orbit model (an ellipse) and the data. Gauss used the quantity

$$\frac{1}{2} \sum_{n=1}^{N} R_n^2(\lambda) \tag{4.1}$$

as a measure of the discrepancy and computed the vector λ for which it is smallest. This is Gauss' *method of least squares.*

[1] By de ZACH, December 7, 1801

The least squares formulation is particularly simple when the residuals depend linearly on the parameter λ. We shall return to that important special case in §4. Gauss' method belongs to a set of techniques that use the important concept of *best approximation*.

2. Review of linear algebra

This section recalls some important concepts covered in the Core B syllabus.

DEFINITION 2.1. A real (or complex) *vector space* V is a set equipped with two rules, addition and multiplication by a scalar, that satisfy the following axioms:

Addition

Given any two elements x and y of V, the *sum* of x and y is denoted $x+y$. The addition rule has the properties:

- $\forall\, x, y \in V,\ x+y=y+x$.
- $\forall\, x, y, z \in V,\ (x+y)+z=x+(y+z)$.
- $\exists\, 0 \in V$ called the *zero vector* such that $\forall\, x \in V,\ x+0=x$.
- $\forall\, x \in V, \exists\, y \in V$ called the *negative of* x such that $x+y=0$.

Multiplication by a scalar

Given any $x \in V$ and $\lambda \in \mathbb{R}$ or $\mathbb{C}$, the *product* of the vector x and the scalar λ is denoted λx. The rule for multiplication by a scalar has the properties:

- $\forall\, x \in V,\ 1x=x$.
- $\forall\, \lambda, \mu \in \mathbb{R}$ (or $\mathbb{C}$) and $x \in V,\ \lambda(\mu x)=(\lambda\mu)x$.
- $\forall\, \lambda, \mu \in \mathbb{R}$ (or $\mathbb{C}$) and $x \in V,\ (\lambda+\mu)x=\lambda x+\mu x$.
- $\forall\, \lambda \in \mathbb{R}$ (or $\mathbb{C}$) and $x, y \in V,\ \lambda(x+y)=\lambda x+\lambda y$.

A *subspace* of a vector space V is a subset that is itself a vector space. In particular, V and $\{0\}$ are subspaces of V; we call them the *trivial subspaces*.

EXAMPLE 2.1. *Let $d \in \mathbb{N}$. The set $\mathbb{R}^d$ (or $\mathbb{C}^d$), equipped with the usual operations of addition and multiplication by a scalar, is a vector space.*

- *For $d=1$, there are no nontrivial subspaces.*
- *For $d=2$, the only nontrivial subspaces are, geometrically speaking, straight lines passing through the origin. For $d=3$, they are straight lines passing through the origin and planes passing through the origin.*

EXAMPLE 2.2. *The set of all real (or complex) sequences, equipped with the usual operations of addition and of multiplication by a scalar, is a vector space. The following are some of the subspaces:*

- *The set of all bounded sequences.*
- *The set of all convergent sequences.*
- *The set of all sequences that converge to 0.*

EXAMPLE 2.3. *Let $A \subset \mathbb{R}$. Let V denote the set of all real (or complex) valued functions defined on A. V, equipped with the usual operations of addition of two functions and multiplication of a function by a scalar, is a vector space. The following are particular subspaces of V:*

- *The set of all polynomials defined on A.*
- *The set of all trigonometric polynomials defined on A.*
- *The set of all continuous functions defined on A.*

- *The set of all differentiable functions defined on A.*
- *The set of all integrable functions defined on A.*

DEFINITION 2.2. Let V denote a vector space. We say that $x \in V$ is a *linear combination* of the vectors $e_1, \ldots, e_n \in V$ if there exist scalars $x_1, \ldots, x_n$ such that

$$x = x_1 e_1 + \ldots + x_n e_n \,.$$

Let $E \subset V$. We denote by $\operatorname{span} E$ the set of all the linear combinations of vectors in E. We say that the vectors in E are *linearly dependent* if one vector in E is a linear combination of other vectors in the set. Otherwise, we say that the vectors in E are *linearly independent.* We say that E is a *basis* for V if

- The vectors in E are linearly independent and
- $\operatorname{span} E = V$.

The *dimension* of V is the number of elements in a basis for V.

EXAMPLE 2.4. *In spite of the formalism, these concepts are quite practical.*

- *The dimension of $\mathbb{R}^d$ is d.*
- *There is no finite basis for any of the vector spaces in Examples 2.2 and 2.3. Such spaces are said to be* infinite dimensional.
- *Let P_N denote the space of all polynomials of degree less than N. Any $p \in P_N$ can be expressed as*

  $$p = p_0 + \ldots + p_{N-1}\, x^{N-1}$$

  where the scalars $p_0, \ldots, p_{N-1}$ are the coefficients of p. The vectors in the set $\{1, \ldots, x^{N-1}\}$ are clearly linearly independent. So the set is a basis for P_N and the dimension of P_N is N.

In Numerical Analysis, the idea of approximating vectors in an infinite dimensional space by elements of a finite dimensional subspace is fundamental. We need a precise mathematical concept for "size" of a vector which generalises the familiar notion of the absolute value of a number.

DEFINITION 2.3. Let V be a vector space. The mapping $x \mapsto \|x\|$ from V to $\mathbb{R}$ is called a *norm* on V if the following holds:

- $\forall\, x \in V,\ \|x\| \geq 0$ with equality iff $x = 0$.
- $\forall\, x \in V,\ \lambda \in \mathbb{R}$ (or $\mathbb{C}$), $\|\lambda x\| = |\lambda|\, \|x\|$.
- $\forall\, x, y \in V,\ \|x + y\| \leq \|x\| + \|y\|$. ($\triangle$ inequality)

We call $(V, \|\cdot\|)$ a *normed space.*

EXAMPLE 2.5. *Every vector in $\mathbb{R}^d$ can be represented as a d-tuple. Let us write*

$$x = (x_1, \ldots, x_d) \in \mathbb{R}^d, \text{ where } x_1, \ldots, x_d \in \mathbb{R}.$$

The following all define norms on $\mathbb{R}^d$:

$$\|x\| = \left(\sum_{i=1}^{d} |x_i|^p \right)^{1/p}, \; p \geq 1.$$

$$\|x\| = \max_{1 \leq i \leq d} |x_i|.$$

EXAMPLE 2.6. *Let $C[0,1]$ denote the space of all the continuous functions defined on the closed interval $[0,1]$. The following are norms on $C[0,1]$:*

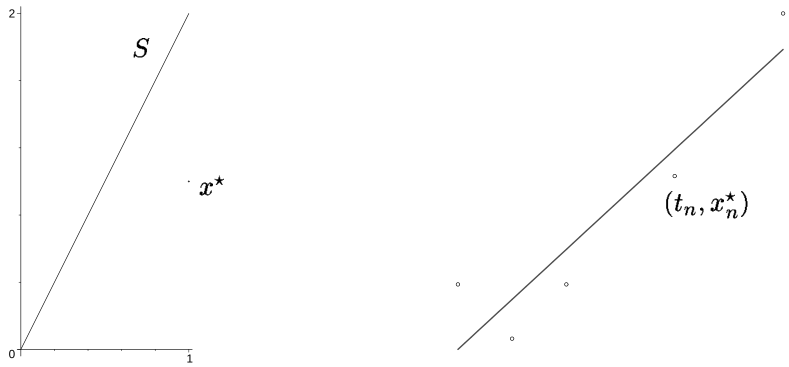
$$\|x\| = \left(\int_0^1 |x(t)|^p \, dt\right)^{1/p}, \; p \geq 1.$$

$$\|x\| = \max_{0 \leq t \leq 1} |x(t)|.$$

3. Best approximation in general normed spaces

Consider the following illustrative problems.

EXAMPLE 3.1. *Given a point $x^\star$ and a straight line S that passes through the origin, find the point x on the line that is nearest to $x^\star$ (see Figure 4.1).*



(a) (b)

FIGURE 4.1. (a) Approximating a point in the plane by a point on a straight line. (b) Fitting a straight line to a set of scattered points.

EXAMPLE 3.2. *Given a continuous function, find a polynomial that "best" approximates it.*

EXAMPLE 3.3. *Given a set $(t_n, x_n^\star)$, $1 \leq n \leq N$, of scattered data points, find a straight line that best fits it.*

In each case,

- a function, a point or some data are given;
- an "approximating set" is specified;
- the object is to find some element of the approximating set that is "optimally close".

In order to give full meaning to the notion of "closeness", we reformulate each problem in terms of a vector space in which a norm is used to measure the distance between vectors. The abstract form of the problem is

Problem (P) Let S be a subspace of the normed space $(V, \|\cdot\|)$. Given $x^\star \in V$, find $x \in S$ such that

$$\forall \; y \in S, \quad \|x - x^\star\| \leq \|y - x^\star\|. \tag{P}$$

This problem is solvable.

THEOREM 3.1. *If S is finite dimensional, then Problem (P) has a solution.*

PROOF. By the triangle inequality, the function

$$y \mapsto \|y - x^\star\|$$

is continuous. Indeed

$$|\,\|y - x^\star\| - \|z - x^\star\|\,| \le \|(y - x^\star) - (z - x^\star)\| = \|y - z\|\,.$$

Consider the set

$$B = \{y \in S : \ \|y\| \le 2\,\|x^\star\|\}\ .$$

This set is closed and bounded. The function $y \mapsto \|y - x^\star\|$ therefore attains a minimum for some $x \in B$. So

$$\forall\, y \in B, \quad \|x - x^\star\| \le \|y - x^\star\|\,.$$

Furthermore, if $y \notin B$, then

$$\|y - x^\star\| \ge |\,\|y\| - \|x^\star\|\,| \ge \|x^\star\| = \|x^\star - 0\| \ge \|x - x^\star\|$$

since $0 \in B$. □

Now that we know that there exists a solution, let us try to find it. We shall need the

DEFINITION 3.1. Let V be a vector space equipped with the norm $\|\cdot\|$. For $r > 0$ and $x^\star \in V$, the set

$$\{y \in V : \ \|y - x^\star\| \le r\}$$

is called the *ball of radius r centered at $x^\star$*. This ball is the set of vectors that are within a distance r of the vector $x^\star$.

For example, in $\mathbb{R}^2$, the unit balls centered at the origin for the norms $\|\cdot\|_2$, $\|\cdot\|_\infty$ and $\|\cdot\|_1$ defined respectively by

$$\|x\|_2 = \left(\sum_{i=1}^{2} |x_i|^2\right)^{1/2}, \quad \|x\|_\infty = \max_{1\le i\le 2} |x_i| \quad \text{and} \quad \|x\|_1 = \sum_{i=1}^{2} |x_i|,$$

are shown in Figure 4.2 (a).

EXAMPLE 3.1 **(revisited)**. In this case, $V = \mathbb{R}^2$ and $S = \text{span}\{a\}$. Our task is to minimise the function, call it $e : \mathbb{R} \to \mathbb{R}$, defined by

$$e(\lambda) = \|\lambda a - x^\star\|\,.$$

Of the many different norms that can be defined in $\mathbb{R}^2$, we shall consider three "popular" choices. As we shall see, the difficulty of constructing a solution depends crucially on the particular norm used.

- Consider the choice $\|\cdot\|_2$. To minimise e is equivalent to minimising its square

$$e^2(\lambda) = (\lambda\, a_1 - x_1^\star)^2 + (\lambda\, a_2 - x_2^\star)^2\,.$$

Using standard calculus techniques, we seek a value of λ for which the derivative vanishes:

$$2\,(\lambda\, a_1 - x_1^\star)\, a_1 + 2\,(\lambda\, a_2 - x_2^\star)\, a_2 = 0\,. \tag{4.2}$$

This is a *linear* equation with a *unique* solution. We find

$$x = \lambda\, a, \quad \text{where} \quad \lambda = \frac{x_1^\star\, a_1 + x_2^\star\, a_2}{a_1^2 + a_2^2}\,.$$

- Next, consider the choice $\|\cdot\|_\infty$. In this case, though e is continuous, it is *not differentiable*. Hence we cannot exploit the recipe that worked so well for the choice $\|\cdot\|_2$. We shall instead resort to a geometric construction. Consider a tiny ball centered at $x^\star$. Gradually increase the radius of that ball until it touches the straight line. The points of intersection of the ball and S solve Problem (P). This is illustrated in Figure 4.2 (b). Now, in the case of the $\|\cdot\|_\infty$ norm, the balls actually look like *squares* and we see that, if $x = \lambda\, a$ solves Problem (P), then it must be at a corner of the smallest square that touches S. This yields the necessary condition

$$|\lambda\, a_1 - x_1^\star| = |\lambda\, a_2 - x_2^\star|\,.$$

There are two possibilities:

$$\text{(i) } \lambda = \frac{x_1^\star - x_2^\star}{a_1 - a_2} \quad \text{or} \quad \text{(ii) } \lambda = \frac{x_1^\star + x_2^\star}{a_1 + a_2}\,.$$

Let us give a numerical example. Figure 4.2 was obtained by setting $x^\star = (1,1)$ and $a = (1,2)$. Equation (i) then gives $\lambda = 0$ and thus $\|\lambda\, a - x^\star\| = 1$. On the other hand, Equation (ii) yields $\lambda = 2/3$ and so $\|\lambda\, a - x^\star\| = 1/3$. Since this value of λ gives the smaller distance, we conclude that $x = (4/3, 2/3)$.

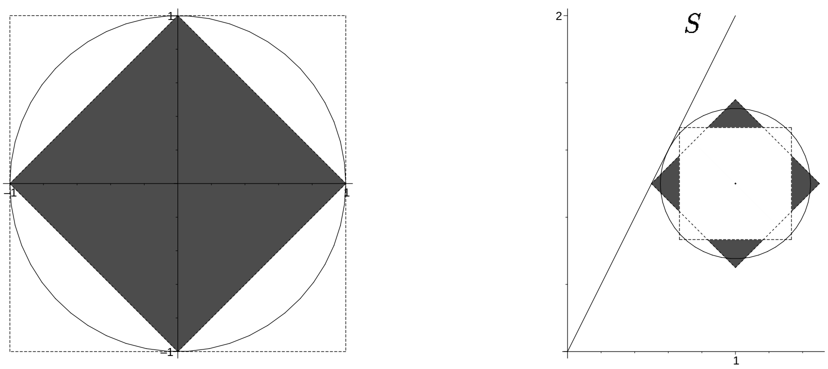
- Finally, we consider the choice $\|\cdot\|_1$ briefly. In this case,

$$e(\lambda) = |\lambda\, a_1 - x_1^\star| + |\lambda\, a_2 - x_2^\star|$$

is, again, not differentiable. Geometrically, the balls are diamond shaped. If x solves Problem (P), then x is a vertex of the smallest diamond that touches S and so either

$$\text{(i) } |\lambda\, a_1 - x_1^\star| = 0 \quad \text{or} \quad \text{(ii) } |\lambda\, a_2 - x_2^\star| = 0\,.$$

For the particular case where $x^\star = (1,1)$ and $a = (1,2)$, it is equation (i) that yields the optimal value of λ, namely $\lambda = 1/2$.



(a) (b)

FIGURE 4.2. (a) Unit balls for the norms $\|\cdot\|_2$, $\|\cdot\|_\infty$ and $\|\cdot\|_1$.
(b) Geometric solution of the problem posed in Example 4.1.

To summarise the discussion, we may say that the choice $\|\cdot\|_2$ leads to the most straightforward form of the Problem (P); the solution is obtained by solving a linear problem and it is unique. On the other hand, if one chooses the norm $\|\cdot\|_\infty$, one needs to resort to a geometrical approach and systematically test a number of possible candidates for the minimiser. (In $\mathbb{R}^d$, there are $d!$ possibilities to examine.) The solution is not always unique. Indeed, if the line S were aligned with one of the axes, then there would be an infinity of "nearest" points. This is due to the fact that the balls for this choice of norm are not strictly convex. The situation in the case of the norm $\|\cdot\|_1$ is analogous.

What is so special about the norm $\|\cdot\|_2$? It is a norm associated with an *inner product*.

4. Best approximation in euclidean spaces

A *euclidean space* is a vector space V equipped with an inner product— that is a mapping $(\cdot,\cdot) : V \times V \to \mathbb{R}$ satisfying the following properties: $\forall\, x,\, y,\, z \in V$ and $\forall\, \lambda \in \mathbb{R}$,

$$(x,\, y) = (y,\, x), \tag{4.3}$$

$$(x + y,\, z) = (x,\, z) + (y,\, z), \tag{4.4}$$

$$(\lambda\, x,\, y) = \lambda\, (x,\, y), \tag{4.5}$$

$$(x,\, x) \geq 0 \;\text{ and }\; (x,\, x) = 0 \implies x = 0\,. \tag{4.6}$$

EXAMPLE 4.1. $\mathbb{R}^d$ *equipped with the inner product*

$$(x,\, y) = \sum_{i=1}^{d} x_i\, y_i$$

is a euclidean space.

EXAMPLE 4.2. *Let us denote by* $C[a, b]$ *the space of all the continuous functions defined on the interval* $[a, b]$. *This is a euclidean space when equipped with the inner product*

$$(x,\, y) = \int_a^b x(t)\, y(t)\, dt\,.$$

In euclidean spaces, a norm may be defined very naturally by

$$\|x\| = \sqrt{(x,\, x)}\,.$$

In order to verify this assertion, we need to show that the triangle inequality holds. This is easily done by using the *Cauchy–Schwarz inequality*:

$$\forall\, x,\, y \in V, \;\; |(x,\, y)| \leq \|x\|\, \|y\|\,. \tag{4.7}$$

PROOF. Since $(\cdot,\, \cdot)$ is an inner product, we have, $\forall\, x,\, y$ and $\forall\, \lambda \in \mathbb{R}$,

$$0 \leq (x - \lambda\, y,\, x - \lambda\, y) = (x,\, x) - 2\,\lambda\,(x,\, y) + \lambda^2\,(y,\, y)\,.$$

The right-hand side of this inequality is a quadratic polynomial in λ whose graph lies above the λ axis. It follows that its discriminant is less than or equal to zero. This is precisely (4.7). □

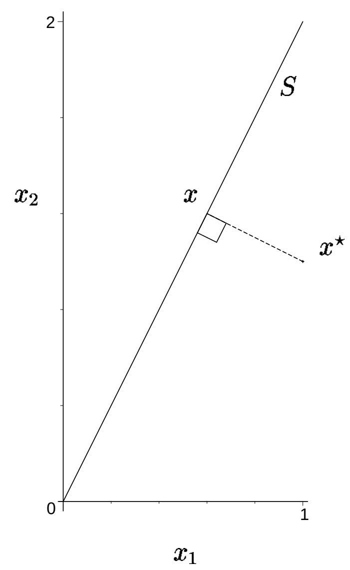


FIGURE 4.3. $x^\star - x$ is orthogonal to S.

The triangle inequality follows easily:

$$||x + y||^2 = ||x||^2 + 2\,(x,\, y) + ||y||^2 \leq (||x|| + ||y||)^2 \ .$$

So $||\cdot|| = \sqrt{(\cdot,\, \cdot)}$ is indeed a norm for V.

A key feature of euclidean spaces is that the concept of *angle*, say θ, between two vectors x and $y \in V$ can be given meaning via

$$\cos\theta := \frac{(x,\, y)}{||x||\,||y||}, \quad -\frac{\pi}{2} < \theta \leq \frac{\pi}{2}\,. \tag{4.8}$$

This definition for θ makes sense because, by virtue of the Cauchy–Schwarz inequality, the absolute value of the right-hand side does not exceed 1. In $\mathbb{R}^2$, this of course reduces to the familiar concept of angle and it is instructive to reexamine Example 4.1 with the choice $||\cdot|| = ||\cdot||_2$. Geometrically, we see from Figure 4.2 (b) that that the line S is tangent to the ball at the point x. An equivalent way of expressing this is to say that the vector $x - x^\star$ is *orthogonal* to S (see Figure 4.3). In terms of the inner product,

$$\forall\, y \in S, \quad (x - x^\star,\, y) = 0\,. \tag{4.9}$$

Set $y = a$ and $x = \lambda\, a$ to obtain

$$\lambda\,(a,\, a) = (a,\, x^\star)\,.$$

This is just another way of expressing (4.2)! This geometrical argument is correct and yields the solution of Problem (P) in any euclidean space.

THEOREM 4.1. *Let V be a euclidean space with inner product $(\cdot,\, \cdot)$ and corresponding norm $||\cdot|| = \sqrt{(\cdot,\, \cdot)}$. If S is finite dimensional, then the solution $x \in S$*

of Problem (P) is unique and is characterised by the condition

$$\forall\, y \in S, \quad (x,\, y) = (x^\star,\, y)\,. \tag{4.10}$$

PROOF. To begin, let us verify that equation (4.10) defines x uniquely. Let

$$k = \dim S$$

and let

$$\{e_1,\, e_2,\, \dots,\, e_k\} \subset S$$

be a basis for S. We seek

$$x = \sum_{i=1}^{k} \lambda_i\, e_i$$

such that (4.10) holds. This can be reformulated as a linear system of k equations for the unknown coefficients $\lambda_1,\, \lambda_2,\, \dots,\, \lambda_k$:

$$\sum_{i=1}^{k} \underbrace{(e_i,\, e_j)}_{\mathcal{A}_{ij}}\, \lambda_i = \underbrace{(x^\star,\, e_j)}_{b_j}, \quad 1 \le j \le k\,. \tag{4.11}$$

The $k \times k$ matrix $\mathcal{A}$ is *symmetric* since $\mathcal{A}_{ji} = \mathcal{A}_{ij}$ by virtue of (4.3). Furthermore, it is *positive definite*. Indeed, Let $0 \ne \mu \in \mathbb{R}^k$ and write $y = \sum_{i=1}^{k} \mu_i\, e_i \ne 0$. Then

$$\mu^T \mathcal{A} \mu = (y,\, y) > 0$$

by (4.6). Such a matrix is clearly invertible. This proves that x is uniquely defined by (4.10).

There remains to show that x solves Problem (P) if and only if (4.10) holds.

[$\Longrightarrow$]: Suppose that x solves Problem (P). Let $y \in S$ and consider the function $\mathcal{E} : \mathbb{R} \to \mathbb{R}$ defined by

$$\mathcal{E}(\epsilon) = \|x + \epsilon\, y - x^\star\|^2\,.$$

By definition of the norm,

$$\begin{aligned}\mathcal{E}(\epsilon) &= (x + \epsilon\, y - x^\star,\, x + \epsilon\, y - x^\star) \\ &= \|x - x^\star\|^2 + 2\,\epsilon(x - x^\star,\, y) + \epsilon^2\, \|y\|^2\,.\end{aligned}$$

Since $x,\, y \in S$ and S is a subspace of V, we note that

$$\forall\, \epsilon \in \mathbb{R}, \quad x + \epsilon\, y \in S\,.$$

Now, x solves Problem (P). Therefore, $\mathcal{E}$ has a minimum at $\epsilon = 0$. It follows that

$$0 = \frac{d\mathcal{E}}{d\epsilon}(0) = (x - x^\star,\, y)\,.$$

$y \in S$ is arbitrary. This proves that x satisfies (4.10).

[$\Longleftarrow$]: Suppose that (4.10) holds. Let $y \in S$. We have $y - x \in S$ and so

$$(x - x^\star,\, y - x) = 0\,.$$

Hence

$$\begin{aligned}\|x - x^\star\|^2 &= (x - x^\star,\, x - x^\star) = (x - x^\star,\, x - x^\star) + (x - x^\star,\, y - x) \\ &= (x - x^\star,\, x - x^\star + y - x) = (x - x^\star,\, y - x^\star) \le \|x - x^\star\|\, \|y - x^\star\|\end{aligned}$$

by the Cauchy–Schwarz inequality. After simplification this yields (P). $\square$

REMARK 4.1. The equations that result from (4.10) when y ranges among the vectors in a basis for S are called *the normal equations.*

The proof of Theorem 4.1 is *constructive.* The solution of Problem (P) can be calculated by solving the normal equations. All that is needed is to select a basis $\{e_1, e_2, \dots, e_k\}$ for the subspace S, compute the inner products (e_i, e_j), $1 \le i, j \le k$ and $(x^\star, e_j)$, $1 \le j \le k$, and solve the linear system

$$\mathcal{A}\lambda = b$$

in Equation (4.11).

5. Linear least squares

In this section, we return to the problem considered in Example 3.3: A set of data points $(t_n, x_n^\star)$, $1 \le n \le N$, is given and we seek the straight line that best fits the data. This problem can be recast in the form of Problem (P) as follows. Let $V = \mathbb{R}^N$, equipped with the familiar inner product

$$(x, y) = \sum_{n=1}^{N} x_n \, y_n \, .$$

The subspace S consists of the vectors in V of the form

$$x = \lambda_1 \underbrace{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}}_{e_1} + \lambda_2 \underbrace{\begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix}}_{e_2} .$$

Thus $S = \operatorname{span}\{e_1, e_2\}$, $k = 2$ and the vector λ is obtained by solving the 2×2 matrix system (4.11) with

$$\mathcal{A} = \begin{pmatrix} N & \sum_{n=1}^{N} t_n \\ \\ \sum_{n=1}^{N} t_n & \sum_{n=1}^{N} t_n^2 \end{pmatrix}, \quad b = \begin{pmatrix} \sum_{n=1}^{N} x_n^\star \\ \\ \sum_{n=1}^{N} t_n \, x_n^\star \end{pmatrix} .$$

For instance, the line in Figure 4.1(b) best fits the data points

$$(1, 0), \ (0, 1), \ (2, 1), \ (4, 3), \ (6, 6) .$$

In this case

$$\mathcal{A} = \begin{pmatrix} 5 & 13 \\ 13 & 57 \end{pmatrix}, \quad b = \begin{pmatrix} 11 \\ 50 \end{pmatrix}$$

and so the equation of the line is

$$x = \lambda_1 + \lambda_2 \, t = -\frac{23}{116} + \frac{107}{116} \, t .$$

A more general problem of data fitting is the following: We are given data points $(t_n, x_n^\star)$, $1 \le n \le N$ and we conjecture that the data would be represented well by a function of the form

$$\lambda_1 \, f_1(t) + \lambda_2 \, f_2(t) + \ldots + \lambda_k \, f_k(t)$$

where the f_i's are known functions of t. The object is to find the parameter vector λ that provides the best fit to the data in the sense that the quantity

$$\frac{1}{2}\sum_{n=1}^{N}\Bigg(\underbrace{\lambda_1 f_1(t_n)+\ldots+\lambda_k f_k(t_n)-x_n^\star}_{R_n(\lambda)}\Bigg)^2$$

is minimised. This is the linear version of Gauss' method of least squares discussed in §1. It is again a particular case of Problem (P) where $V=\mathbb{R}^N$, equipped with the usual inner product, and the subspace S consists of vectors of the form

$$x=\lambda_1\underbrace{\begin{pmatrix}f_1(t_1)\\ \vdots\\ f_1(t_N)\end{pmatrix}}_{e_1}+\lambda_2\underbrace{\begin{pmatrix}f_2(t_1)\\ \vdots\\ f_2(t_N)\end{pmatrix}}_{e_2}+\ldots+\lambda_k\underbrace{\begin{pmatrix}f_k(t_1)\\ \vdots\\ f_k(t_N)\end{pmatrix}}_{e_k}.$$

Let E be the $N\times k$ matrix with columns $e_1, e_2, \ldots\ e_k$. The matrix $\mathcal{A}$ may be written

$$\mathcal{A}=E^TE$$

and the normal equations take the form

$$E^TE\,\lambda=E^Tx^\star. \tag{4.12}$$

EXAMPLE 5.1. *Fit a parabola to the data points used in Figure 4.1(b), i.e*

$$(1,\,0),\ (0,\,1),\ (2,\,1),\ (4,\,3),\ (6,\,6).$$

In this case, the model function is

$$x=\lambda_1 f_1(t)+\lambda_2 f_2(t)+\lambda_3 f_3(t),$$

where $f_1=1$, $f_2=t$ and $f_3=t^2$. Therefore

$$E=\begin{pmatrix}1&0&0\\1&1&1\\1&2&4\\1&4&16\\1&6&36\end{pmatrix},\quad E^T=\begin{pmatrix}1&1&1&1&1\\0&1&2&4&6\\0&1&4&16&36\end{pmatrix}$$

and so the normal equations are

$$E^TE\,\lambda=\begin{pmatrix}5&13&57\\13&57&289\\57&289&1569\end{pmatrix}\lambda=\begin{pmatrix}11\\50\\268\end{pmatrix}=E^Tx^\star.$$

The solution is

$$\lambda_1=\frac{311}{469},\quad \lambda_2=-\frac{509}{1876}\quad and\quad \lambda_3=\frac{369}{1876}.$$

The corresponding parabola is show in Figure 4.4.

EXAMPLE 5.2. *Consider the data points $(t_n,\,y_n^\star)$, $1\le n\le 5$, given by*

$$(1,\,5.10),\ (1.25,\,5.79),\ (1.5,\,6.53),\ (1.75,\,7.45)\ \text{and}\ (2,\,8.46).$$

$y_n^\star$ increases rapidly with t_n. We therefore conjecture that the data can be represented adequately by a formula of the type

$$y(t)=b\,e^{at},$$

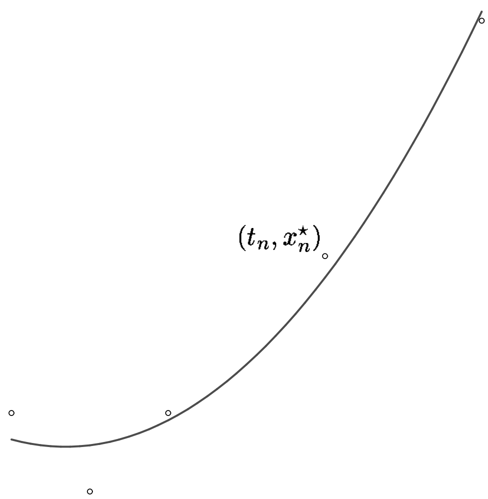


FIGURE 4.4. The parabola that best fits a set of points $(t_n, x_n^\star)$, $1 \le n \le N$.

where a and b are parameters yet to be determined. This is not a linear model in a and b. However, after taking logarithms, we obtain

$$x(t) := \ln y(t) = \underbrace{\ln b}_{\lambda_1} + \underbrace{a}_{\lambda_2} t$$

and we can use linear least squares. We have

$$E = \begin{pmatrix} 1 & 1 \\ 1 & 1.25 \\ 1 & 1.50 \\ 1 & 1.75 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad x^\star = \begin{pmatrix} \ln 5.10 \\ \ln 5.79 \\ \ln 6.53 \\ \ln 7.45 \\ \ln 8.46 \end{pmatrix}.$$

Hence, the normal equations are

$$\begin{pmatrix} 5 & 7.5 \\ 7.5 & 11.875 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 9.405 \\ 14.424 \end{pmatrix}.$$

This yields $\lambda_1 = 1.12$ (equivalently $b = 3.06$) and $\lambda_2 = a = 0.51$.