Robert Adrain and the Method of Least Squares

Author(s): Jacques Dutka

Source: *Archive for History of Exact Sciences*, 1990, Vol. 41, No. 2 (1990), pp. 171-184
Published by: Springer Nature
Stable URL: https://www.jstor.org/stable/41133885

Robert Adrain and the Method of Least Squares

JACQUES DUTKA

Communicated by H. J. M. BOS

1. Introduction

The method of least squares is a very important numerical technique of applied mathematics where it is used for the adjustment of observations, statistical estimation, curve fitting, *etc.* Publications on the method by A. M. LEGENDRE, ROBERT ADRAIN and C. F. GAUSS originally appeared in the first decade of the nineteenth century. The rival claims of LEGENDRE and GAUSS for priority of discovery generated considerable controversy in the years following.

For a long time the relatively unavailable publications of ROBERT ADRAIN on the method remained comparatively unknown, but in 1980 they were reprinted in STIGLER [1; Vol. 1]. The primary purpose of this paper is to present ADRAIN'S derivations and applications of the method of least squares in modern terminology.

A sketch of ADRAIN'S mathematical career is given in Section 2. A brief history of the adjustment of observations in the eighteenth century and of the method of least squares is given in Section 3. A surveying problem which was the stimulus for R. ADRAIN'S work on least squares is given in Section 4. ADRAIN'S derivations of the normal law and of the method of least squares are discussed in Section 5 and his applications of the method in Section 6. Finally the question of the originality of ADRAIN'S work is treated in Section 7.

2. Adrain's mathematical career

ROBERT ADRAIN (1775–1843) emigrated from Ireland to the United States in 1798. Although without much formal education, he taught in secondary schools in New Jersey, Pennsylvania, New York, and colleges including Queens (Rutgers), Columbia, and the University of Pennsylvania.

In their lifetimes, NATHANIEL BOWDITCH (1773–1838) and ADRAIN were regarded as the outstanding mathematicians in America. Like ADRAIN, BOWDITCH was also largely self-taught and well-versed in the contemporary developments of the French mathematical school, which was particularly active in the early nineteenth century.

Opportunities for mathematical publication in the United States were quite limited and such periodicals as existed were mainly devoted to the proposal and solution of problems, and both men displayed considerable ability in their solutions of difficult problems. But, in their more serious work, they exhibited marked contrasts. BOWDITCH applied mathematics principally to astronomy and related sciences. He translated and gave a remarkably detailed commentary on the first four volumes of P. S. LAPLACE's monumental work on celestial mechanics. (See LAPLACE [1].)

ADRAIN too was principally concerned with applied mathematics. He once wrote (HOGAN [1; 167]):

"The last and highest department of mathematical science consists in its applications to the laws and phenomena of the material world."

But he was much more varied and original in outlook than BOWDITCH in his mathematical publications. His mathematical work has been discussed in detail by COOLIDGE [1], STRUIK [1], HOGAN [1], and others. Here, only a sketch of this work, other than that concerned with the theory of errors, will be given. The latter will be described in Sections 5ff.

ADRAIN was a frequent contributor to the *Mathematical Correspondent* edited by GEORGE BARON and begun in 1804. In this, the first American mathematical periodical, ADRAIN's most important publication was on Diophantine analysis. He also proposed and partially solved the question of the form of the "skipping (or jump) rope" (*catenaria volvens*).* After the demise of this periodical, ADRAIN began a new publication — as he was to do several times in his career. He became the publisher and editor of *The Analyst* in 1808. He was also an active contributor, often giving additional solutions of problems which extended and generalized the original problems. Notable among these are his geometric developments of "isotomous curves" which stemmed from an analysis of the hygrometer of D. RITTENHOUSE, his ingenious application of techniques of the calculus of variations in solving a problem of R. PATTERSON, and his most important work relating to the theory of errors. He was also one of those who solved a remarkable astronomical problem originally posed by LAPLACE in 1796, concerning a "black hole."

The possibility that there were stars whose gravitational attractions were so great as to prevent any of their light (conceived of as luminous particles) from escaping had been considered by JOHN MICHELL, a British astronomer, in 1784 and later by LAPLACE. BOWDITCH proposed LAPLACE's problem (which LAPLACE had actually solved in 1799) as "Question 45" and gave a solution.**

The Analyst ceased publication after only a single volume had appeared. ADRAIN attempted to revive it some years later without success. Despite this,

* This was only completely solved many years later after the development of elliptic functions. (See *e.g.* BOWMAN [1; 29–31].)

** See *The Analyst*, 114, 223–224.

ADRAIN continued his professional activities, other than teaching, as an editor of mathematical books and periodicals, and as a contributor to various publications. The range of his interests was quite extensive, including geophysical problems, the history of mathematics and descriptive geometry in addition to the research areas mentioned above. Many unpublished manuscripts were left on his death. These have since been lost or destroyed.

As pointed out above, in ADRAIN's lifetime, and particularly after the article by ABBE [1] in 1871 calling attention to the work of 1808 on the theory of errors, ADRAIN had a reputation as an able and original mathematician. But in the twentieth century, particularly in modern times, this view has been questioned by several critics. This will be discussed further in Section 7.

3. The adjustment of observations and the least squares controversy

The development of such observational sciences in the eighteenth century as astronomy and geodesy led to a concomitant development of numerical methods for reducing observational data. Frequently, there were problems in which the number of equations formed from the observations exceeded the number of unknown parameters in the equations — the problems were "overdetermined." Procedures were developed for obtaining a set of values for the parameters which, by some criterion, should fit the system of equations as well as possible.

From about mid-century on, increasing attention was given to such problems by LEONHARD EULER, TOBIAS MAYER, J. H. LAMBERT in Germany, THOMAS SIMPSON in England, R. J. BOSCOVICH in Italy, and J. L. LAGRANGE and P. S. LAPLACE in France to name only the most prominent. Details and references concerning this work can be found in a survey article by H. L. HARTER [1], and in a recent book by S. M. STIGLER [3].

The first published account of the method of least squares, MLS, now by far the most frequently used technique, was published by A. M. LEGENDRE [1] in 1805. The problems he discussed may be put as follows: An observation is assumed to have the theoretical form $\sum_{i=1}^{m} \alpha_i x_i$ where $(\alpha_1, \alpha_2, \ldots, \alpha_m)$ are parameters to be estimated and $(x_1, x_2, \ldots, x_m)$ are a given set of variables. To estimate the parameters, a series of observations, $y_1, y_2, \ldots, y_n$ are made and a quadratic form $\sum_{i=1}^{n} (y_j - a_1 x_{1j} - a_2 x_{2j} \ldots - a_m x_{mj})^2$ is considered in which x_{ij} is a known value of x_i, $m < n$ in general, and the quadratic form is a function of the coefficients $(a_1, a_2, \ldots, a_m)$. The set of values $(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_m)$ for which the value of the quadratic form is a minimum is taken as an estimate of the parameters $(\alpha_1, \alpha_2, \ldots, \alpha_m)$. LEGENDRE's method of least squares can be extended to cover a wide variety of applications. It is objective, general and (usually) facile from a computational standpoint. But there is another viewpoint involving the theory of probability which was not considered by LEGENDRE.

Suppose there is a series of measurements of a physical quantity, such as a length, whose true value is unknown. The measurements are assumed to be independent without systematic error and equally reliable. The differences between the measurements and the length are regarded as random errors subject to a law of error.

If the individual measurements are denoted by $x_1, x_2, \ldots, x_n$, and a is to be an estimate of the length, what is the "best" value to ascribe to a? An answer to this, stemming from GAUSS, which leads to the method of least squares can be obtained if it is assumed that $x_1 - a, x_2 - a, \ldots, x_n - a$, represent independent "normally" distributed random variables. *I.e.*, each random variable (r.v.) has a probability density

$$f(x) = \frac{h}{\sqrt{\pi}} \exp -(h(x-a))^2, \qquad -\infty < x < \infty \tag{1}$$

where h is associated with the reliability of a measurement. The "most probable value" of a is obtained when one maximizes the likelihood, a function proportional to the joint probability densities,

$$L = \exp - h^2[(x_1 - a)^2 + (x_2 - a)^2 + \ldots + (x_n - a)^2]. \tag{2}$$

This reduces to finding the value of a for which the exponent on the right is a minimum — essentially equivalent to the criterion for the method of least squares. If $\bar{x} = (1/n) \sum_{j=1}^{n} x_j$ denotes the arithmetic mean of $x_1, x_2, \ldots, x_n$, then

$$\sum_{j=1}^{n} (x_j - a)^2 = \sum_{j=1}^{n} [(x_j - \bar{x}) + (\bar{x} - a)]^2$$

$$= \sum_{j=1}^{n} (x_j - \bar{x})^2 + 2(\bar{x} - a) \sum_{j=1}^{n} (x_j - \bar{x}) + n(\bar{x} - a)^2.$$

From the definition of $\bar{x}$, $\sum_{j=1}^{n} (x_j = \bar{x}) = 0$, and $(\bar{x} - a)^2 = 0$ when a is set equal to $\bar{x}$. Thus, in this case, $\sum_{j=1}^{n} (x_j - a)^2$ is minimized when the most probable value of a is the arithmetic mean $\bar{x}$.*

The procedure described above in which measurement errors are assumed to be normally distributed and the most probable values are obtained by maximizing the likelihood will be called the probabilistic justification of the MLS in what follows.

In 1808, some three years after LEGENDRE's publication, ROBERT ADRAIN developed the MLS from a probabilistic standpoint and gave some important applications, then and later. This will be discussed in detail in Sections 4 to 6. But ADRAIN's researches were either unknown or ignored in Europe and had no influence on subsequent developments of the MLS there.

In 1809 C. F. GAUSS [1] published a book principally concerned with his methods of determining orbits of celestial bodies. In Art. 175ff. he gave a demonstration of the MLS in which he derived (1) on the hypothesis that the most

* *Cf.* the discussion of Problem I in Section 6.

probable value of a set of equally reliable observations of an unknown parameter is the arithmetic mean of the observations. GAUSS acknowledged LEGENDRE's prior publication of the principle, but claimed that he had been using the MLS since 1795. LEGENDRE responded negatively, and then became increasingly bitter. The question of priority of discovery of the method soon became the most notorious controversy in the history of statistics.

A detailed discussion of the matter has been given by R. L. PLACKETT [1] who has given English versions of much of the relevant correspondence and history. Claims are sometimes made that there were additional discoverers of the MLS such as the Swiss mathematician DANIEL HUBER (1769–1829). In 1939, WALTER SPIESS [1] examined HUBER's geodetic work which included the triangulation of the Canton of Basel from 1813 to 1824. The claim for HUBER as a discoverer of the MLS was rejected by SPIESS.

The normal distribution (1) and the method of least squares soon gained an importance in the nineteenth century which went far beyond the original questions in astronomy and geodesy which gave rise to them. Their theoretical developments and applications in numerous other disciplines had a considerable, even dominant, influence on the growth of mathematical probability and statistics during the nineteenth century. This influence, though much reduced, persists until the present day, and accounts for the continuing interest in the history of these subjects.

4. Patterson's surveying problem and Bowditch's solution

Among the contributors to ADRAIN's *Analyst* of 1808 were some of the foremost American mathematicians of the time. One of these "ingenious correspondents" was ROBERT PATTERSON, of the University of Pennsylvania.

In an early issue (No. II, p. 42), he proposed a surveying problem which remained unsolved for months, although he had offered a prize of ten dollars "for the best satisfactory solution ... to be adjudged by the Editor." The problem was:

"In order to find the content of a piece of ground, having a plane level surface I measured with a common circumferentor and chain, the bearings and lengths of several sides, or boundary line which I found as follows:

1. N 45° E 40 perches
2. S 30° W 25 ditto
3. S 5° E 36 ditto
4. W 29.6 ditto
5. N 20° E 31 ditto to the place of beginning.

But upon casting up the difference of latitude and departure, I discovered what will perhaps always be the case in actual surveys, that some error has been contracted in taking the dimensions. Now it is required to compute the area of this enclosure, on the *most probable supposition* of this error."

The problem concerns the area of a polygon which does not close because of observational errors in the measurements of its angles and sides. The sides are given in terms of bearings and lengths and are expressed in rectangular coordinates of departures and latitudes. (The departure is the distance between two meridians of longitude at any given latitude.) The circumferentor (plain surveying compass) is used to take bearings from one vertex to another, and the error generated is principally due to the angular difference between the true direction and that indicated by compass readings. The surveyor's chain of wrought iron or steel contained a standard number of uniform links, and its length was measured in terms of perches (poles, rods) of 16.5 feet. Errors in measurements of length were due to irregularities in terrain, temperature variations, *etc.*

The prize was awarded to NATHANIEL BOWDITCH of Salem whose solution was published (No. IV, pp. 88–93). BOWDITCH, who had previously had surveying experience, was already at the time probably the most famous contributor to *The Analyst* because of his highly successful treatise on navigation.*

BOWDITCH begins by stating some principles for adjusting a survey and then explains how the positions of the vertices B, C, D, E of a field whose measured boundary is represented by $ABCDE$ can be sequentially adjusted to eliminate the closure error. He observes that

"in measuring the lengths of any lines the errors would probably be in proportion to their lengths.
... In observing the bearings of all the boundary lines, equal errors are liable to be committed."

BOWDITCH does not justify these assumptions. The first assumption, which was accepted by ADRAIN (Section 5), was criticized by GLAISHER [1; 78] who wrote "it seems scarcely likely that the error should be directly proportional to the distance measured." BOWDITCH summarizes his results in a geometric and an arithmetic procedure for correcting the survey.

In the geometric procedure, the boundary $ABCDE$ is drawn from the observed bearings and lengths and the magnitude and direction of the closure error EA are found. A quotient, $r = |AE|/(|AB| + |BC| + |CD| + |DE|)$, is obtained. Through the vertices B, C, D, E, lines BB'; CC'', ... are drawn, all parallel to and in the direction of EA. Here $|BB'| = r \cdot |AB|$, $|CC''| = r \cdot (|AB| + |BC|)$, $|DD''| = r \cdot (|AB| + |BC| + |CD|)$, ... Through the points A, B', C'', D''', ... the adjusted boundary lines are drawn. The area may be found by dividing the figure into triangles.

In the arithmetic procedure, the measured bearings and lengths in PATTERSON's problem are expressed in terms of latitude and departure components and the closure error is found (0.10 perches S, 0.08 perches W). The closure error is corrected by adjusting the positions of the vertices. (B, C, D, E, F are each adjusted 0.02 pN in the latitude. B, D, F, which terminate the longer sides, are adjusted by 0.02 pE, while C, D are adjusted by 0.01 pE to complete the departure corrections. The arithmetic procedure is not an exact equivalent of the geometric. BOWDITCH [1; 88] states that corrections should be made in cases where they affect the error of

* Updated editions of his *American Practical Navigator* are still published every few years.

the survey but not otherwise — presumably for practical convenience. Finally, the area of the closed polygon is computed (854.56 sq. p.).)

As formulated by ADRAIN [1; 108], BOWDITCH's rule for correcting a survey is as follows:

"Say as the sum of the distances is to each particular distance, so is the whole error in departure to the correction of the corresponding departure; each correction being so applied as to diminish the whole error in departure: proceed the same way for the corrections in latitude."

BOWDITCH's rule for the adjustment of traverses is still frequently mentioned in modern textbooks on surveying. It will be shown in Section 6 that the rule as stated here is equivalent to the solution obtained by the MLS.

5. Adrain's derivation of the method of least squares

Immediately following BOWDITCH's solution of the surveying problem is a much longer article by ADRAIN [1; 93–108] which also includes a solution. (As mentioned previously, ADRAIN frequently gave additional or general solutions to problems submitted by contributors.) In ADRAIN's article, the MLS is developed and applied to several problems including PATTERSON's. For convenience, ADRAIN's notation will be modernized and his article will be treated in two parts, here and in Section 6.

Suppose one measures a number of successive lengths on a line *AB*, *BC*, *etc.*, and obtains the values *Ab*, *bc*, *etc.*, the whole error being *Cc*. ADRAIN assumes with BOWDITCH that the most probable values of the measurement errors are proportional to their lengths. Let X and Y denote the errors involved in measuring the lengths $AB = a$, $BC = b$ X and Y have a joint probability density $f(x; a) \cdot f(y; b)$, *i.e.* X and Y are independent but the individual probability densities (p.d.'s) have the same form. If the joint p.d. is maximized subject to the condition that the total error $X + Y = E$ is fixed, the result obtained should yield $x/a = y/b$. ADRAIN proceeds to maximize $\ln f(x; a) \cdot f(y; b)$ when $x + y =$ constant. He thus obtains

$$\frac{d}{dx}\ln f(x;\, a) = \frac{d}{dy}\ln f(y;\, b) \tag{1}$$

and says that this ought to be equivalent to $x/a = y/b$. This, says ADRAIN, "is effected in the simplest manner possible" by assuming

$$\frac{f'(x;\, a)}{f(x;\, a)} = m\cdot\frac{x}{a}, \quad \frac{f'(y;\, b)}{f(y;\, b)} = m\cdot\frac{y}{b} \tag{2}$$

where m is a constant.* On integration, one finds the curve of probability in the form

$$f(x;\, a) = C\cdot\exp\frac{mx^2}{2a} \tag{3}$$

* GLAISHER [1; 81] criticizes this assumption.

where C is a (normalizing) constant and ADRAIN shows that m must be negative. When several (independent) errors, $X, Y, Z, \ldots$, each having a p.d. of this type are added together, their joint p.d. has a maximum when $(x^2/a + y^2/b + z^2/c + \ldots)$ is a minimum — this is ADRAIN's demonstration of the principle of least squares.

ADRAIN's procedure is based on *ad hoc* assumptions, that the most probable values of the measurement errors are proportional to their lengths and that (2) is supposed in order to obtain a particular solution of (1) which is then developed further. It is not persuasive and was criticized many years later in a fine review of the MLS by J. W. L. GLAISHER [1; 76–81].

Perhaps ADRAIN himself had reservations, for he attempts to give a second proof. After assuming that measurements of lengths and bearings are independent, he represents them as rectangular coordinates. The p.d.'s of the errors in length and bearing are each assumed to be symmetrically distributed about axes which are parallel to the respective coordinate axes. By a dubious argument, he infers that the horizontal (W–E) and vertical (N–S) components of error in measuring a distance are independent and similarly distributed. The joint probability density is $f(x) \cdot f(y)$ and is constant for all points lying on a given circle $x^2 + y^2 = r^2$ where r is the error of a distance measurement. He maximizes $\ln f(x) f(y) =$ const. with this condition. On differentiating these equations, he finds

$$\frac{1}{x}\frac{d(\ln f(x))}{dx} = \frac{1}{y}\frac{d(\ln f(y))}{dy} = n \tag{4}$$

where n is a constant. On integrating, he gets $f(x) = \exp(C + \frac{1}{2}nx^2)$, where n is negative, for the probability of x grows less as x grows greater.

The idea which arises in ADRAIN's second proof of obtaining the independence of rectangular components of distance error from the independence of the length and bearing errors also occurred independently to Sir JOHN HERSCHEL, a noted British astronomer, in a book review of 1850 (HERSCHEL [1; 17]). Although the fallacy was soon pointed out by R. L. ELLIS [1; 325–327], HERSCHEL's derivation of the normal law continued to be accepted for some time thereafter by some of the most eminent British physicists of the time.

ADRAIN concludes his theoretical discussion by noting that if two independent normally distributed errors X and Y are scaled in the ratio 1 to p, their joint p.d. is proportional to $\exp - (x^2/a + y^2/p^2 a)$ and in this case "the curve of *equal probability* is an ellipsis."

As pointed out above, neither of ADRAIN's "proofs" of the normal law of error as a probabilistic basis for the MLS is now accepted as valid. Yet ADRAIN appears to have been *the first to publish* the relation between them — in particular, before GAUSS [1; Art. 175]. Moreover, it must be recognized that in the nineteenth century, a large number of unsuccessful attempts were made to derive the normal distribution from simple, readily acceptable assumptions by many of the most competent mathematicians and scientists.* Indeed near the end of the nineteenth century, G. LIPPMANN, a French physicist, once remarked, "Every-

* See GLAISHER [1], MERRIMAN [1], HARTER [1], KNOBLOCH [1], *et al.*

body believes in the law of errors [the normal distribution], the experimenters because they think it is a mathematical theorem, the mathematicians because they think it is an experimental fact."*

6. Adrain's illustrative problems

ADRAIN follows his theoretical derivation of the MLS by some examples which increase in order of complexity.

Problem I is concerned with finding the most probable value x of a set of measurements $a, b, c, \ldots$ The errors are $x - a, x - b, x - c, \ldots,$ and (assuming that they are independent and normally distributed) the logarithms of their p.d.'s are (essentially) given by $-(x-a)^2, -(x-b)^2, -(x-c)^2, \ldots$ Thus, one minimizes the sum $(x-a)^2 + (x-b)^2 + (x-c)^2 + \ldots$ ADRAIN differentiates and finds that the arithmetic mean $x = (a + b + c + \ldots)/n$ is the most probable value. He notes that this coincides with the practice of astronomers, navigators, *etc.*

Problem II is an extension of the foregoing and involves finding the most probable position of a point in space when a set of observed positions is given. ADRAIN proceeds in an individualistic and complex manner which will only be outlined here. The observed positions are projected perpendicularly onto an arbitrary plane. A fixed line is selected in the plane and a fixed point is selected on the line. Altitudes to the fixed line are constructed from the feet of the perpendiculars to the plane. An observed position is characterized by three parameters: the height of the perpendicular above the fixed plane, the length of an altitude to the line and the distance of the foot of the altitude from the fixed point. Let the corresponding coordinates of the most probable position be (x, y, z). ADRAIN minimizes the quadratic form $\sum_{i-1}^{n} (x - a_i)^2 + (y - b_i)^2 + (z - c_i)^2$ and reduces the problem to three one-dimensional problems like those in Problem I. He solves these as above and obtains the center of gravity of the observed positions. It is then shown how the process may be modified when the observed positions have unequal weights, *etc.*, and the results are applied to some subproblems.

There are similarities and differences between the publications on least squares by LEGENDRE (1805) and ADRAIN (1808). Both show that the arithmetic mean of a set of observations, and the center of gravity of a set of points (in space) are consequences of applying the MLS. But, while LEGENDRE only dicusses the advantages of the MLS as a numerical algorithm and does not attempt a further justification, ADRAIN gives a probabilistic demonstration for the method and notes that it yields rules which are in accord with those generally used.

Moreover, it is in Problems III and IV that ADRAIN treats considerably more sophisticated applications of the MLS than LEGENDRE. (See also ADRAIN [2] and [3].)

ADRAIN's Problem III is an application of the method to navigation in which the correction of the dead reckoning (of a ship's position) at sea by an observation

* H. POINCARÉ [1; 171]

of latitude is calculated. From an analytical standpoint, this is similar to the more complex surveying problem which follows, and the details of ADRAIN's application will be omitted here.

Problem IV is the correction of a survey and is a generalization, in analytical form, of PATTERSON's problem. In particular, BOWDITCH's rule is also obtained. A modernized version of ADRAIN's discussion (pp. 106–109), which is expressed in NEWTON's fluxional notation, will be given.*

Let $P_1P_2, P_2P_3, \ldots, P_{n-1}P_n$ denote the sides measured in a survey whose respective lengths and bearings are $a_1, a_2, \ldots, a_{n-1}$ and $\theta_1, \theta_2, \ldots, \theta_{n-1}$. Let $\Delta a_1, \Delta a_2, \ldots, \Delta a_{n-1}$ and $\Delta\theta_1, \Delta\theta_2, \ldots, \Delta\theta_{n-1}$ be the corrections required to obtain closure and adjust for the error P_nP_1. The departure and latitude components of P_mP_{m+1} are $a_m \sin\theta_m$ and $a_m \cos\theta_m$ respectively, $1 \leqq m \leqq n-1$. Thus one must have

$$\sum_{m=1}^{n-1} (a_m + \Delta a_m) \sin(\theta_m + \Delta\theta_m) = 0, \qquad \sum_{m=1}^{n-1} (a_m + \Delta a_m) \cos(\theta_m + \Delta\theta_m) = 0. \tag{1}$$

To a first order approximation, on setting $\Delta a_m = x_m$ and $a_m \Delta\theta = y_m$, where the radial and transverse components are regarded as independently distributed errors, one gets

$$\begin{aligned} D_m &= x_m \sin\theta_m + y_m \cos\theta_m, \\ L_m &= x_m \cos\theta_m - y_m \sin\theta_m \end{aligned} \tag{2}$$

for the respective components of error of P_mP_{m+1} in departure and latitude. ADRAIN (in effect) assigns the weights $1/a_m$ to x_m and $1/p^2 a_m$ to y_m and proceeds to apply the principle of least squares. Thus one has

$$\sum_{m=1}^{n-1} \left(\frac{x_m^2}{a_m} + \frac{y_m^2}{p^2 a_m} \right) = \text{minimum} \tag{3}$$

subject to the closure conditions of (1) and (2),

$$\begin{aligned} \sum_{m=1}^{n-1} D_m &= -\sum_{m=1}^{n-1} a_m \sin\theta_m = D, \\ \sum_{m=1}^{n-1} L_m &= -\sum_{m=1}^{n-1} a_m \cos\theta_m = L. \end{aligned} \tag{4}$$

On solving this minimization problem by introducing the LAGRANGE multipliers $2\lambda_D$, $2\lambda_L$** in (4), one has

$$x_m = a_m(\lambda_D \sin\theta_m + \lambda_L \cos\theta_m), \qquad y_m = a_m p^2 (\lambda_D \cos\theta_m - \lambda_L \sin\theta_m) \tag{5}$$

---

* See also the expositions of T. W. WRIGHT [1; 222–223] and E. HAMMER [1; 621–625].

** In ADRAIN's solutions the analogous multipliers are left unnamed.

whence it follows from (2) that

$$\begin{aligned} D_m &= \lambda_D a_m(\sin^2\theta_m + p^2\cos^2\theta_m) + \lambda_L a_m(1-p^2)\sin\theta_m\cos\theta_m, \\ L_m &= \lambda_D a_m(1-p^2)\sin\theta_m\cos\theta_m + \lambda_L a_m(\cos^2\theta_m + p^2\sin^2\theta_m) \end{aligned} \tag{6}$$

for $1 \leqq m \leqq n-1$.

ADRAIN continues:

"The simplest case of the problem is, when $p = 1, \ldots$, besides this seems to agree best with the imperfections of the common instruments used in surveying."

Let

$$D = -\sum_{m=1}^{n-1} a_m \sin\theta_m, \quad L = -\sum_{m=1}^{n-1} a_m \cos\theta_m.$$

In this case one has

$$\lambda_D = D\Big/\sum_{m=1}^{n-1} a_m, \quad \lambda_L = L\Big/\sum_{m=1}^{n-1} a_m \tag{7}$$

since

$$D_m = \lambda_D a_m, \quad L_m = \lambda_L a_m. \tag{8}$$

ADRAIN's development, which depends on an application of the MLS, has thus yielded a result equivalent to that obtained by BOWDITCH in Section 4.

The geometrical interpretations of (6) in some particular cases when $p \neq 1$ are also considered by ADRAIN. He concluded the article by stating that he had applied the principle of least squares to the determination of the most probable value of the earth's ellipticity but lacked the space to present it.*

In his Problem IV, ADRAIN anticipated a method later developed extensively by GAUSS and his successors for geodetic and other applications. It represented an important advance in the methodology of surveying. (See WRIGHT [1; 223] and HAMMER [1; 629].)

His work was mentioned by BOWDITCH in the latter's treatise on navigation and in some American books on surveying, but otherwise received little attention — perhaps because it was too far in advance of the level of mathematical education in America at the time.

The following quotation from a book by ROBERT GIBSON [1; f.n. 192–194] on surveying of 1828 is interesting in this regard. GIBSON refers to BOWDITCH's rule (Section 4) and states:

"This arithmetical rule was given by Mr. Bowditch in his solution of Mr. Patterson's question of correcting a survey in No. 4 of *The Analyst*. Also the Editor, Dr. Adrain has given precisely the same practical rule in his elegant solution to the said question, analytically demonstrated."

GIBSON then proceeds to give not ADRAIN's "elegant solution", but BOWDITCH's in detail!

* This was actually published later. See ADRAIN [2], [3] and Section 7.

7. The question of the originality of Adrain's work

ADRAIN's article deriving the MLS appears to have been virtually unknown in Europe until the publication by C. ABBE [1] in 1871 which included a reproduction of ADRAIN's first derivation of the normal law. Following this and GLAISHER's critique of ADRAIN's proof, the general consensus until modern times was that GAUSS (1794–1795), LEGENDRE (1805), and ADRAIN (1808), were independent discoverers of the MLS. The priority dispute which arose between LEGENDRE and GAUSS, discussed in Section 3, is not dissimilar to controversies of the seventeenth century concerning the discoveries of logarithms, the calculus, *etc.*, by different investigators working independently.

In retrospect, what is remarkable, given the simple, almost mechanical procedure for *applying* the least squares principle to the adjustment of observations, is that it had not been discovered before 1794. Indeed, in a letter of 1812 to OLBERS, (PLACKETT [1; 244]), GAUSS says

"... The only thing which is surprising is that this principle ... was not already applied 50 or 100 years earlier by others, e.g., Euler or Lambert or Halley or Tobias Mayer ...".

The consensus mentioned above has been challenged recently by a number of critics, most persistently by S. M. STIGLER [1; 243–244], [2], and elsewhere. STIGLER [2] has examined some of the evidence for GAUSS's discovery of least squares before LEGENDRE's publication of 1805 and found it insufficient, yet conjectures with some evidence that GAUSS's claim of prior discovery was justified. Elsewhere STIGLER [3; 143] is more definite — GAUSS's demonstration of 1809 of the MLS is characterized as "nonsense."

But STIGLER's sharpest attacks are reserved for ADRAIN. Contrary to the opinion of numerous competent mathematicians such as J. L. COOLIDGE [1] and D. J. STRUIK [1] who have examined ADRAIN's mathematical work in detail, STIGLER [1; 243–245], says:

"His sole contribution to our field (and his sole original contribution to mathematics) appeared in a mathematical magazine he started in 1808. Adrain began his solution by presenting two derivations of the normal distribution both of which were more wishful thinking than proofs."

He notes that (in an article of 1926), M. J. BABB stated that ADRAIN's library also contained LEGENDRE's work of 1805 in the original paper cover. As STIGLER [2; 465] puts it: "Robert Adrain may have "discovered" it [the MLS] in Legendre's book." STIGLER [3; 374] also assigns a probable publication date of 1809 for ADRAIN [1].

Two questions arise here:

(a) When did ADRAIN first learn about LEGENDRE's prior work on the method of least squares?

(b) Was ADRAIN's probabilistic justification of the MLS and of its applications original with him?

The better part of two centuries have passed since ADRAIN wrote his initial article on least squares. It appears impossible to determine when he obtained a copy of LEGENDRE [1] and learned of the latter's priority. A summary of the available evidence which points to ADRAIN's independent discovery of the MLS follows:

1. ADRAIN refers several times in relevant publications to having found the method without mention of any predecessors. *E.g.*, he says in 1818 (ADRAIN [2]):

"Having in the year 1808 discovered a general method of resolving several useful problems by ascertaining the highest degree of probability where certainty cannot be found; I shall here apply that method to the determination of the earth's ellipticity ..."

ADRAIN proceeds to determine the ellipticity by aplying the MLS to a series of fifteen observations of the lengths of the seconds pendulum in different latitudes taken from LAPLACE [1].

2. ABBE [1] says that manuscripts left after ADRAIN's death indicate that the investigations published as ADRAIN [2] and [3] were completed in 1808.

3. A comparison of BOWDITCH's solution of PATTERSON's problem in Section 3 and ADRAIN's solution of Problem IV in Section 5, shows a natural (and unforced) relation. ADRAIN, in accordance with his editorial custom, was attempting to obtain a much more extensive and generalized development. Instead of BOWDITCH's rather vague intuitive use of probability, ADRAIN substitutes a principle based on a law of error. The application of the least squares principle is given in a general form, which is shown to include BOWDITCH's solution as a particular case.

So far as Question (b) above is concerned, there appears to be general unanimity. ADRAIN is accepted as an independent discoverer of the normal distribution, from which the least squares principle follows. Moreover there is no doubt that ADRAIN's application of the MLS in Problem IV represented an important development in the history of surveying.

Bibliography

ABBE, CLEVELAND [1], "Historical Note on the method of least squares", *American Journal of Science*, Vol. **1** (1871), 411–415. Repr. in STIGLER [1].

ADRAIN, ROBERT [1], "Article XIV. Research concerning the probabilities of the errors ...," *The Analyst or Mathematical Museum*, Vol. **1** (1808), 93–108. Repr. in STIGLER [1].

ADRAIN, ROBERT [2], "Investigation of the Figure of the Earth, and of the Gravity in Different Latitudes," *Transactions of the American Philosophical Society*, Vol. **1**, N.S. (1818), 119–135. Repr. in STIGLER [1].

ADRAIN, ROBERT [3], "Research Concerning the Mean Diameter of the Earth," *Ibid.*, 353–366, (1818). Repr. in STIGLER [1].

BOWDITCH, N. [1], "Article XIII. Solution of Mr. Patterson's Prize Question for correcting a survey ... ", *The Analyst or Mathematical Museum*, Vol. **I** (1808), 88–93.

BOWMAN, F. [1], *Introduction to Elliptic Functions witt applications*, New York, 1953.

COOLIDGE, J. L. [1], "Robert Adrain and the Beginnings of American mathematics, *American Mathematical Monthly*, Vol. **33** (1926), 61–76.

ELLIS, R. L. [1], 'Remarks on an alleged Proof of the "Method of Least Squares," contained in a late number of the *Edinburgh Review*', *London Edinburgh and Dublin Philosophical Magazine*, Ser. 3, Vol. **57** (1850), 321–328.

GAUSS, K. F. [1], *Theory of the Motion of the Heavenly Bodies* ... (Trans. from original edition published in Latin in Hamburg, 1809). Reprinted, New York, 1963.

GIBSON, ROBERT [1], *The Theory and Practice of Surveying*, New York, 1828.

GLAISHER, J. W. L. [1], "On the Law of Facility of Errors of Observations, and on the Method of Least Squares," *Memoirs of the Royal Astronomical Society*, Vol. **39** (1872), 75–124.

HAMMER, E. [1], „Beitrag zur Geschichte der Ausgleichungsrechnung", *Zeitschrift für Vermessungswesen*, Bd. XXIX (1900), 613–628.

HARTER, H. L. [1], "The method of least squares and some alternatives — Part I", *International Statistical Review*, Vol. **42** (1974), 147–174.

HERSCHEL, J. F. W. [1], Review of Quetelet's "Lettres á S.A.R. le Duc regnant de Saxe Cobourg et Gotha", *Edinburgh Review*, Vol. **83** (1850), 1–57.

HOGAN, E. R. [1], "Robert Adrain: American mathematician", *Historia Mathematica*, Vol. **4** (1977), 157–172.

KNOBLOCH, E. [1], „Zur Grundlagenproblematik der Fehlertheorie", *Festschrift für Helmut Gericke*, Ed. by M. FOLKERTS, & U. LINDGREN, Wiesbaden, 1985, 561–590.

LAPLACE, P. S. [1], *Traité de Mécanique Céleste*, 5 vols. Paris, 1799–1825. Vols. I–IV transl. by N. BOWDITCH. Repr., New York, 1966–1969.

LEGENDRE, A. M. [1], *Nouvelles méthodes pour la détermination des orbites des comètes.* Paris, 1805. An abridged translation into English of the Appendix on least squares was published in D. E. SMITH, *A Source Book of Mathematics*, repr., in 2 vols., New York, NY, 1959.

MERRIMAN, M. [1], "A list of writings relating to the method of least squares ...", *Trans. of the Connecticut Academy of Arts and Sciences*, Vol. **4** (1877), 151–232. Repr. in STIGLER [1; Vol. I].

PLACKETT, R. L. [1], "The discovery of the method of least squares," ... *Biometrika*, Vol. **59** (1972), 239–251.

POINCARÉ, H. [1], *Calcul des Probabilités*, Paris, 1912.

SPIESS, W. [1], „Kann man für Daniel Huber ... Ansprüche als Erfinder, der „Methode der Kleinsten Quadrate geltend machen?" *Schweizerische Zeitschrift für Vermessungswesen und Kulturtechnik*, 37. Jahrgang (1939), 11–17 u. 21–23.

STIGLER, S. M. [1], *American Contributions to Mathematical Statistics in the Nineteenth Century*, 2 vols., New York, 1980.

STIGLER, S. M. [2], "Gauss and the invention of least squares," *Annals of Mathematical Statistics*, Vol. **9** (1981), 465–474.

STIGLER, S. M. [3], *The History of Statistics. The Measurement of Uncertainty before* 1900, Cambridge, Mass., 1986.

STRUIK, D. J. [1], "Adrain, Robert", *Dictionary of Scientific Biography*, New York, 1970, Vol. **1**, 65–66.

WRIGHT, T. W. [1], *A Treatise on the Adjustment of Observations* ..., New York, 1884.

Audits and Surveys, Inc.
650 Avenue of the Americas
New York, NY 10011

(Received May 1, 1990)