

LEGENDRE

On Least Squares

[Translated from the French by Professor Henry A Ruger and Professor Helen M Walker, Teachers College, Columbia University, New York City.]

The great advances in mathematical astronomy made during the early years of the nineteenth century were due in no small part to the development of the method of least squares. The same method is the foundation for the calculus of errors of observation now occupying a place of great importance in the scientific study of social, economic, biological, and psychological problems. Gauss says in his work on the *Theory of Motions of the Heavenly Bodies* (1809) that he had made use of this principle since 1795 but that it was first published by Legendre. The first statement of the method appeared as an appendix entitled “Sur la Méthode des moindres carrés” in Legendre’s *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris 1805. The portion of the work translated here is found on pages 72–75.

Adrien-Marie Legendre (1752–1833) was for five years a professor of mathematics in the *École Militaire* at Paris, and his early studies on the paths of projectiles provided a background for later work on the paths of heavenly bodies. He wrote on astronomy, the theory of numbers, elliptic functions, the calculus, higher geometry, mechanics and physics. His work on geometry, in which he rearranged the propositions of Euclid, is one of the most successful textbooks ever written.

On the Method of Least Squares

In the majority of investigations in which the problem is to get from measures given by observation the most exact result which they can furnish, there almost always arises a system of equations of the form

$$a + bx + cy + fz + \&c.$$

in which $a, b, c, f, \&c.$ are the known coefficients which vary from one equation to another, and $x, y, z, \&c.$ are the unknowns which must be determined in accordance with the condition that the value of E shall for each equation reduce to a quantity which is either zero or very small.

If there are the same number of equations as unknowns $x, y, z, \&c.$, there is no difficulty in determining the unknowns, and the error E can be made absolutely zero. But more often the number of equations is greater than that of the unknowns, and it is impossible to do away with all the errors.

In a situation of this sort, which is the usual thing in physical and astronomical problems, where there is an attempt to determine certain important components, a degree of arbitrariness necessarily enters in the distribution of the errors, and it is not to be expected that all the hypotheses shall lead to exactly the same results; but it is particularly important to proceed in such a

way that extreme errors, whether positive or negative, shall be confined within as narrow limits as possible.

Of all the principles which can be proposed for that purpose, I think there is none more general, more exact, and more easy of application, that of which we made use in the preceding researches, and which consists of rendering the sum of squares of the errors a minimum. By this means, there is established among the errors a sort of equilibrium which, preventing the extremes from exerting an undue influence, is very well fitted to reveal that state of the system which most nearly approaches the truth.

The sum of the squares of the errors $E^2 + E'^2 + E''^2 + \&c.$ being

$$\begin{aligned} & (a + bx + cy + fz + \&c.)^2 \\ + & (a' + b'x + c'y + f'z + \&c.)^2 \\ + & (a'' + b''x + c''y + f''z + \&c.)^2 \\ + & \&c., \end{aligned}$$

if its *minimum* is desired, when x alone varies, the resulting equation will be

$$o = \int ab + x \int b^2 + y \int bc + z \int bf + \&c.,$$

in which by $\int ab$ we understand the sum of similar products, i.e., $ab + a'b' + a''b'' + \&c.$; by $\int b^2$ the sum of the squares of the coefficients of x , namely $b^2 + b'^2 + b''^2 + \&c.$, and similarly for the other terms.

Similarly the minimum with respect to y will be

$$o = \int ac + x \int bc + y \int c^2 + z \int fc + \&c.,$$

and the minimum with respect to z ,

$$o = \int af + x \int bf + y \int cf + z \int f^2 + \&c.,$$

in which it is apparent that the same coefficients $\int bc$, $\int bf$, $\&c.$ are common to two equations, a fact which facilitates the calculation.

In general, to form the equation of the minimum with respect to one of the unknowns, it is necessary to multiply all the terms of each given equation by the coefficient of the unknown in that equation, taken with regard to its sign, and to find the sum of these products.

The number of equations of minimum derived in this manner will be equal to the number of the unknowns, and these equations are then to be solved by the established methods. But it will be well to reduce the amount of computation both in multiplication and in solution, by retaining in each operation only so much signification of figures, integers or decimals, as are determined by the degree of approximation for which the inquiry calls.

Even if by a rare chance it were possible to satisfy all the equations at once by making all the errors zero, we could obtain the same result from the equations of minimum; for if after having found the values of x , y , z , $\&c.$ which make E , E' , $\&c.$ equal to zero, we let x , y , z vary by δx , δy , δz , $\&c.$, it is evident that E^2 , which was zero, will become by that variation $(a\delta x + b\delta y + c\delta z + \&c.)^2$. The

same will be true of E'^2 , E''^2 , &c. Thus we see that the sum of squares of the errors will by variation become a quantity of the second order with respect to δx , δy , &c., which is in accord with the nature of a minimum.

If after having determined all the unknowns x , y , z , &c., we substitute their values in the given equations, we will find the value of the different errors E , E' , E'' , &c., to which the system gives rise, and which cannot be reduced without increasing the sum of their squares. If among these error are some which appear too large to be admissible, then those equations which produced these errors will be rejected, as coming from too faulty experiments, and the unknowns will be determined by means of the other equations, which will then give much smaller errors. It is further to be noted that one will not then be obliged to begin the calculations anew, for since the equations of minimum are formed by the addition of the products made in each of the given equations, it will suffice to remove from the addition those products furnished by the equations which would have led to errors that were too large.

The rule by which one finds the mean among the results of difference observations is only a very simple consequence of our general method, which we will call the method of least squares.

Indeed, if experiments have given different values a , a' , a'' , &c. for a certain quantity x , the sum of squares of the errors will be $(a' - x)^2 + (a'' - y)^2 + (a''' - x)^2$, and on making that sum a minimum, we have

$$o = (a' - x) + (a'' - y) + (a''' - x),$$

from which it follows that

$$x = \frac{a' + a'' + a''' + \&c.}{n},$$

n being the number of the observations.

In the same way, if to determine the position of a point in space, a first experiment has given the coordinates a' , b' , c' ; a second the coordinates a'' , b'' , c'' ; and so on, and if the true coordinates of the point are denoted by x , y , z ; then the error in the first experiment will be the distance from the point (a', b', c') to the point (x, y, z) . The square of this distance is

$$(a' - x)^2 + (a'' - y)^2 + (a''' - x)^2,$$

If we make the sum of the squares of all such distances a minimum, we get three equations which give

$$x = \frac{\int a}{n}, \quad y = \frac{\int b}{n}, \quad z = \frac{\int c}{n},$$

n being the number of points given by the experiments. These formulas are precisely the ones by which one might find the common centre of gravity of several equal masses situated at the given points, whence it is evident that the centre of gravity of any body possesses this general property.

If we divide the mass of a body into particles which are equal and sufficiently small to be treated as points, the sum of the square of the distances from the particles to the centre of gravity will be a minimum.

We see then that the method of least squares reveals to us, in a fashion, the centre about which all the results furnished by experiments tend to distribute themselves, in such a manner as to make their deviations from it as small as possible. The application which we are now about to make of this method to the measurement of the meridian will display most clearly its simplicity and fertility.¹

From D E Smith, *A Source Book in Mathematics*, McGraw-Hill 1929 and Dover 1959, Volume II, pages 576–579.

¹An application of the method to an astronomical problem follows.