

1.8: Simultaneous Linear Equations, $N > n$

Consider the following Equations

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1 = 0 \quad (1.8.1)$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2 = 0 \quad (1.8.2)$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + b_3 = 0 \quad (1.8.3)$$

$$a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + b_4 = 0 \quad (1.8.4)$$

$$a_{51}x_1 + a_{52}x_2 + a_{53}x_3 + b_5 = 0 \quad (1.8.5)$$

Here we have five Equations in only three unknowns, and there is no solution that will satisfy all five Equations exactly. We refer to these Equations as the *Equations of condition*. The problem is to find the set of values of x_1 , x_2 and x_3 that, while not satisfying any one of the Equations exactly, will come closest to satisfying all of them with as small an error as possible. The problem was well stated by Carl Friedrich Gauss in his famous *Theoria Motus*. In 1801 Gauss was faced with the problem of calculating the orbit of the newly discovered minor planet Ceres. The problem was to calculate the six elements of the planetary orbit, and he was faced with solving more than six Equations for six unknowns. In the course of this, he invented the method of least squares. It is hardly possible to describe the nature of the problem more clearly than did Gauss himself:

"...as all our observations, on account of the imperfection of the instruments and the senses, are only approximations to the truth, an orbit based only on the six absolutely necessary data may still be liable to considerable errors. In order to diminish these as much as possible, and thus to reach the greatest precision attainable, no other method will be given except to accumulate the greatest number of the most perfect observations, and to adjust the elements, not so as to satisfy this or that set of observations with absolute exactness, but so as to agree with all in the best possible manner."

If we can find some set of values of x_1 , x_2 and x_3 that satisfy our five Equations fairly closely, but without necessarily satisfying any one of them exactly, we shall find that, when these values are substituted into the left hand sides of the Equations, the right hand sides will not be exactly zero, but will be a small number known as the residual, R .

Thus:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1 = R_1 \quad (1.8.6)$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2 = R_2 \quad (1.8.7)$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + b_3 = R_3 \quad (1.8.8)$$

$$a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + b_4 = R_4 \quad (1.8.9)$$

$$a_{51}x_1 + a_{52}x_2 + a_{53}x_3 + b_5 = R_5 \quad (1.8.10)$$

Gauss proposed a "best" set of values such that, when substituted in the Equations, gives rise to a set of residuals such that the sum of the squares of the residuals is least. (It would in principle be possible to find a set of solutions that minimized the sum of the absolute values of the residuals, rather than their squares. It turns out that the analysis and the calculation involved is a good deal more difficult than minimizing the sum of the squares, with no very obvious advantage.) Let S be the sum of the squares of the residuals for a given set of values of x_1 , x_2 and x_3 . If any one of the x -values is changed, S will change - unless S is a minimum, in which case the derivative of S with respect to each variable is zero. The three Equations

$$\frac{\partial S}{\partial x_1} = 0, \quad \frac{\partial S}{\partial x_2} = 0, \quad \frac{\partial S}{\partial x_3} = 0 \quad (1.8.11)$$

express the conditions that the sum of the squares of the residuals is least with respect to each of the variables, and these three Equations are called the *normal Equations*. If the reader will write out the value of S in full in terms of the variables x_1 , x_2 and x_3 , he or she will find, by differentiation of S with respect to x_1 , x_2 and x_3 in turn, that the three normal Equations are

$$A_{11}x_1 + A_{12}x_2 + A_{13}x_3 + B_1 = 0 \quad (1.8.12)$$

$$A_{12}x_1 + A_{22}x_2 + A_{23}x_3 + B_2 = 0 \quad (1.8.13)$$

$$A_{13}x_1 + A_{23}x_2 + A_{33}x_3 + B_3 = 0 \quad (1.8.14)$$

where

$$A_{11} = \sum a_{i1}^2, \quad A_{12} = \sum a_{i1}a_{i2}, \quad A_{13} = \sum a_{i1}a_{i3}, \quad B_1 = \sum a_{i1}b_i, \quad (1.8.15)$$

$$A_{22} = \sum a_{i2}^2, \quad A_{23} = \sum a_{i2}a_{i3}, \quad B_2 = \sum a_{i2}b_i, \quad (1.8.16)$$

$$A_{33} = \sum a_{i3}^2, \quad B_3 = \sum a_{i3}b_i, \quad (1.8.17)$$

and where each sum is from $i = 1$ to $i = 5$.

These three normal Equations, when solved for the three unknowns x_1 , x_2 and x_3 , will give the three values that will result in the lowest sum of the squares of the residuals of the original five Equations of condition.

Let us look at a numerical example, in which we show the running checks that are made in order to detect mistakes as they are made. Suppose the Equations of condition to be solved are

$$7x_1 - 6x_2 + 8x_3 - 15 = 0 \quad -6$$

$$3x_1 + 5x_2 - 2x_3 - 27 = 0 \quad -21$$

$$2x_1 - 2x_2 + 7x_3 - 20 = 0 \quad -13$$

$$4x_1 + 2x_2 - 5x_3 - 2 = 0 \quad -1$$

$$9x_1 - 8x_2 + 7x_3 - 5 = 0 \quad 3$$

$$-108 \quad -69 \quad -71$$

The column of numbers to the right of the Equations is the sum of the coefficients (including the constant term). Let us call these numbers s_1 , s_2 , s_3 , s_4 , s_5 .

The three numbers below the Equations are $\sum a_{i1}s_i$, $\sum a_{i2}s_i$, $\sum a_{i3}s_i$

Set up the normal Equations:

$$159x_1 - 95x_2 + 107x_3 - 279 = 0 \quad -108$$

$$-95x_1 + 133x_2 - 138x_3 + 31 = 0 \quad -69$$

$$107x_1 - 138x_2 + 191x_3 - 231 = 0 \quad -71$$

The column of numbers to the right of the normal Equations is the sum of the coefficients (including the constant term). These numbers are equal to the row of numbers below the Equations of condition, and serve as a check that we have correctly set up the normal Equations. The solutions to the normal Equations are

$$x_1 = 2.474 \quad x_2 = 5.397 \quad x_3 = 3.723$$

and these are the numbers that satisfy the Equations of condition such that the sum of the squares of the residuals is a minimum.

I am going to suggest here that you write a computer program, in the language of your choice, for finding the least squares solutions for N Equations in n unknowns. You are going to need such a program over and over again in the future – not least when you come to Section 1.12 of this chapter!.

This page titled [1.8: Simultaneous Linear Equations, \$N > n\$](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Jeremy Tatum](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.