Polynomial curve fitting when abscissas and ordinates are both subject to error

By M. O'Neill,* I. G. Sinclair† and Francis J. Smith

Departments of Computer Science and Applied Mathematics, School of Physics and Applied Mathematics, The Queen's University of Belfast, Northern Ireland

An iterative method is described for the least-square curve fitting of a polynomial to a set of points in two dimensions when both the abscissas and ordinates are subject to error and when the weights of all the readings are known. The process converges, in general, to a polynomial giving the exact minimum of the 'weighted' perpendicular distances onto the curve. It is shown that in practice Deming's method gives a solution close to this optimum polynomial.

(First received December 1967 and in revised form July 1968)

1. Introduction

The problem of the least square curve fitting of a function to a set of experimental data (x_i, y_i), $1 \leqslant i \leqslant n$, when x_i and y_i are both subject to error with weights ω_{xi} and ω_{yi} respectively, has been described and reviewed by Deming (1943). If $f(x, a_r)$ is the function and a_r, $1 \leqslant r \leqslant m$, is a set of variable parameters the problem reduces to the minimisation of the sum of the squares of the weighted perpendicular distances from the points (x_i, y_i) to the curve, that is, the minimisation of the sum:

$$S(a_r, x_i') = \sum_{i=1}^{n} \{\omega_{xi}(x_i - x_i')^2 + \omega_{yi}[y_i - f(x_i', a_r)]^2\} \quad (1)$$

with respect to the parameters a_r and with respect to the 'adjusted' values x_i' where $[x_i', f(x_i')]$ lies on the curve.

A method for finding the approximate minimum of $S(a_r, x_i')$ when $f(x, a_r)$ is the polynomial

$$f(x, a_r) = \sum_{r=0}^{n} a_r x^r \quad (2)$$

has been described by Deming, but this does not give the exact minimum of S. More recently a number of writers have studied the special case when $f(x, a_r)$ is a straight line. This work has been reviewed by York (1966) who has developed a method for the straight line based on the solution of a cubic equation and a given approximate value of the slope of the line.

In this paper we describe an iterative method which in general determines to any given accuracy the polynomial minimising S; the polynomial can have any specified degree so our solution includes the case of the straight line. We show also that Deming's method quickly gives a close approximation to the exact solution.

We use expansions of $f(x, a_r)$ in sets of orthogonal polynomials, following Forsythe (1957), because this reduces computing time and increases the accuracy of the computations. However, the theory we describe can be simply altered to fit a power series in cases when the coefficients of certain powers are constrained (for example, we might wish to fit a quadratic of the form $f(x) = a_0 + a_4x^4$ to the set of data).

* *Now with Independent Computer Services (N. Ireland) Ltd., Belfast.*

† *Now with Short Brothers & Harland Ltd., Belfast.*

2. Theory

The problem to be solved is basically the minimisation of the function $S(a_r, x_i')$ with respect to its independent variables, a_r and x_i'. A number of methods are available for minimising a function in many variables; these have been recently reviewed by Powell (1966). Using a library program for our I.C.T. 1905 computer based on the method for minimising a sum of squares due to Powell (1965) we were able to obtain the polynomials for which S is a minimum, but this and other common methods for minimising a function are inefficient compared with the methods we are going to describe because they are not able to take account of certain special features of the function S.

The minimisation process which we found to be most successful is based on the well known Newton–Raphson method for minimising a function of many variables. In the following we adopt the notation of Powell's review as far as possible.

2.1. Newton–Raphson method

We expand $f(x, a_r)$ in the form

$$f(x, a_r) = \sum_{r=0}^{m} a_r p_r(x) \quad (3)$$

where $p_r(x)$ is a polynomial of degree r. The orthogonal properties of these polynomials will be discussed later. We now wish to minimise $S(a_r, x_i')$. We consider an approximate set of variables a_{r0} and x_{i0}' close to the minimum of S, then if the minimum is at $a_r = a_{r0} + \delta a_r$ and $x_i' = x_{i0}' + \delta x_i'$

$$\frac{\partial S}{\partial a_r} = \frac{\partial S}{\partial x_i} = 0 \quad (4)$$

at the minimum and using Taylor's series to first order in δa_r and $\delta x_i'$

$$\left(\frac{\partial S}{\partial a_r}\right)_0 + \sum_{s=0}^{m} \left(\frac{\partial^2 S}{\partial a_r \partial a_s}\right)_0 \delta a_s + \sum_{j=1}^{n} \left(\frac{\partial^2 S}{\partial a_r \partial x_j'}\right)_0 \delta x_j' = 0 \quad (5)$$

and

$$\left(\frac{\partial S}{\partial x_i'}\right)_0 + \sum_{s=0}^{m} \left(\frac{\partial^2 S}{\partial x_i' \partial a_s}\right)_0 \delta a_s + \sum_{j=1}^{n} \left(\frac{\partial^2 S}{\partial x_i' \partial x_j'}\right)_0 \delta x_j' = 0. \quad (6)$$

We have $m + n$ non-linear equations in the $m + n$ unknowns δa_s and $\delta x_i'$; these equations can be written in matrix form by defining the following matrices and vectors for $0 \leqslant r, s \leqslant m$ and $1 \leqslant i, j \leqslant n$:

$$\left.\begin{aligned} g_r^{(1)} &= \left(\frac{\partial S}{\partial a_r}\right)_0; \; g_i^{(2)} = \left(\frac{\partial S}{\partial x_i'}\right)_0; \; A_{sr} = \left(\frac{\partial^2 S}{\partial a_s \partial a_r}\right)_0; \\ B_{ji} &= \left(\frac{\partial^2 S}{\partial x_j' \partial x_i'}\right)_0; \; C_{rj} = C_{jr}' = \left(\frac{\partial^2 S}{\partial a_r x_j'}\right)_0. \end{aligned}\right\} \quad (7)$$

Then in the form of a partitioned matrix we have

$$\begin{bmatrix} A & C \\ C' & B \end{bmatrix} \begin{bmatrix} \delta a \\ \delta x' \end{bmatrix} = - \begin{bmatrix} g^{(1)} \\ g^{(2)} \end{bmatrix}. \quad (8)$$

If we write the partitioned matrix in this equation as G, then the corrections δa_r and $\delta x_i'$ are given by the matrix equation

$$\begin{bmatrix} \delta a \\ \delta x' \end{bmatrix} = - G^{-1} \begin{bmatrix} g^{(1)} \\ g^{(2)} \end{bmatrix}. \quad (9)$$

The position of the exact minimum can be found by the successive use of this equation (provided that the iteration converges).

This process appears impractical because at each iteration we have to invert a matrix of order $m + n$, and in many cases n may be large (it is not unusual to curve fit several hundred experimental points). The problem is made easier, however, because

$$B_{ij} = 0 \text{ if } i \neq j; \quad (10)$$

and, in addition, if the polynomials $p_r(x)$ are orthogonal over the set of points x_{i0}' with respect to the weights ω_{yi}, that is if

$$\sum_{i=1}^{n} p_r(x_{i0}') p_s(x_{i0}') \omega_{yi} = 0 \quad (11)$$

then $$A_{rs} = 0.$$

Therefore A and B are diagonal matrices. In addition the elements of the matrices C given by

$$C_{ri} = - 2\omega_{yi} \frac{\partial}{\partial x_{i0}'} \{[y_i - f(x_{i0}')] p_r(x_{i0}')\} \quad (12)$$

contain fewer terms and are usually smaller than the diagonal elements A_{rr} and B_{ii}. To a first approximation, therefore, we can approximate G^{-1} by the simple diagonal matrix

$$G^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix}. \quad (13)$$

Then

$$a_r = a_{r0} - [A^{-1} g^{(1)}]_r \quad (14)$$

which, after simplification, takes the form

$$a_r = \sum_{i=1}^{n} \omega_{yi} y_i p_r(x_{i0}') / \sum_{i=1}^{n} \omega_{yi} [p_r(x_{i0}')]^2. \quad (15)$$

Similarly

$$x_i' = x_{i0}' + \left\{ \frac{\omega_{yi}[y_i - f(x_{i0}')] f'(x_{i0}') + \omega_{xi}(x_i - x_{i0}')}{\omega_{yi}[f'(x_{i0}')]^2 - \omega_{yi}[y_i - f(x_{i0}')] f''(x_{i0}') + \omega_{xi}} \right\}. \quad (16)$$

In this last equation more accurate values of x_i' are obtained if the new values of a_r given in (15) are used in the calculation of $f(x_{i0}')$, $f'(x_{i0}')$ and $f''(x_{i0}')$ in (16).

2.2. The algorithm

We suggest the following process, based on equations (15) and (16), for finding the least-square curve fit. We begin with the experimental points x_i as our first approximation to x_i'; that is we let $x_{i0}' = x_i$. The polynomials orthogonal over this set of points are generated by the recurrence relations (Forsythe, 1957).

$$p_0(x) = 1; \; p_1(x) = (x - \alpha_1);$$
$$p_r(x) = (x - \alpha_r) p_{r-1}(x) - \beta_r p_{r-2}(x), \quad (17)$$

in which

$$\alpha_r = \sum_i \{\omega_{yi}[p_{r-1}(x_{i0}')]^2 x_{i0}'\} / \sum_i \{\omega_{yi}[p_{r-1}(x_{i0}')]^2\}, \quad (18)$$

$$\beta_r = \sum_i \{\omega_{yi} p_{r-1}(x_{i0}') p_{r-2}(x_{i0}') x_{i0}'\} / \sum_i \{\omega_{yi}[p_{r-2}(x_{i0}')]^2\}. \quad (19)$$

The coefficients a_r are computed from equation (15). This gives us the polynomial which fits the experimental data when the x_i coordinates are error free. We use this as our first approximation.

We now iterate to the correct solution. We begin by computing from equation (16) approximate values of x_i'. In (16) the quantities $f(x_{i0}')$, $f'(x_{i0}')$, $f''(x_{i0}')$ are computed using the algorithm for summing orthogonal polynomial series and their derivatives due to Smith (1965). In this algorithm less than $6m$ multiplications and $9m$ additions or subtractions are needed to compute all three quantities with minimal rounding error. The summation in the numerator in (15) is evaluated in the same way.

When the new set of x_i' are calculated, a new set of orthogonal polynomials are generated by computing sets of α_r and β_r using (18) and (19), and a_r coefficients using (15). This completes the first iteration and gives us the second approximation to the polynomial. This iterative procedure is then repeated till convergence is reached. We call this the direct iterative method.

The total amount of computation is small and it increases linearly with the number of experimental points. Each iteration consists of approximately $(17\,nm)$ multiplications, $(16\,nm)$ additions and subtractions and $(2\,nm)$ divisions where n is the number of experimental points and m is the order of the polynomial. On our I.C.T. 1905 computer when $n = 10$ and $m = 2$ a complete iteration took approximately 2 seconds and cost about 6 pence. Comparable costs can be worked out for other problems from these figures.

Often a polynomial power series is required rather than an orthogonal polynomial at the end of the calculation. If we write

$$\sum_{r=0}^{m} a_r p_r(x) = \sum_{s=0}^{m} c_s x^s;$$

then the coefficients, c_s, can be calculated from the following recurrence relations:

$$\gamma_{k,i} = 0 \text{ if } k > i \text{ or if } k, i < 1;$$
$$= 1 \text{ if } k = i;$$
$$\gamma_{k+1,i+1} = \gamma_{k,i} - \alpha_i \gamma_{k+1,i} - \beta_i \gamma_{k+1,i-1} \text{ if } k < i;$$
$$c_s = \sum_{r=s}^{m} a_r \gamma_{s+1,r+1} \text{ for } 0 \leqslant s \leqslant m.$$

2.3. Convergence

A question naturally arises about whether or not this direct iterative procedure converges, and if it does converge, is it to the minimum? We found no proof that the process always converges, but a large number of examples were tried, varying the weights and the positions of the abscissas, and convergence was found in all cases. Examples are given in a thesis by O'Neill (1967). However, in a few extreme cases, such as in an example we give at the end of this paper, convergence is slow and over 100 iterations are needed to give accurate results. It therefore seems likely that examples may be found of non-convergence or of convergence which is negligibly slow. If any such cases occur they might be solved by other methods such as Powell's method for the minimisation of a function or using the exact inverse of G in place of the approximate inverse in (13).

It is easily shown that if the iteration does converge it usually converges to the polynomial minimising S. From (14) and (16) it follows that for all r and i

$$0 = g_r^{(1)} = g_i^{(2)}. \tag{20}$$

Hence

$$\left(\frac{\partial S}{\partial a_r}\right) = \left(\frac{\partial S}{\partial x_i'}\right) = 0 \tag{21}$$

for all r and i. Thus, a minimum is obtained provided that all of the second derivatives

$$\left(\frac{\partial^2 S}{\partial a_r^2}\right) \text{ and } \left(\frac{\partial^2 S}{\partial x_i'^2}\right)$$

are positive, the off-diagonal elements being small. It is readily shown that the first of these second derivatives is always positive, but the second, given by the denominator in equation (16), is unfortunately not necessarily positive, though in practice it almost always is so.

A case when this failure might occur and when there are three curves making S stationary is illustrated schematically in **Fig. 1** (the x and y coordinates have been scaled so that the lines joining (x_i, y_i) to the points on the curve at which $\left(\frac{\partial S}{\partial x_i'}\right) = 0$ are perpendicular to the curve). If the curve passes through a sharp maximum and if the experimental point (x_i, y_i) lies below the maximum then as the point $[x_i', f(x_i')]$ passes along the curve, the partial sum

$$S_i(x_i') = \omega_{xi}(x_i - x_i')^2 + \omega_{yi}[y_i - f(x_i')]^2 \tag{22}$$

passes through a maximum at P_2 and two minima at P_1 and P_3. It is possible that our process might find the maximum point, but this possibility is easily eliminated (in a computer program) by checking that the denominator in (16) is not negative. If it is negative then it follows that we have found a maximum point and that there are two minima on either side of the maximum; approximate positions of these minima can be found by scanning S_i on either side of P_2 and more accurate values can be calculated using Newton's method in (16). The lowest of the two minimum points (in the example, P_3) then gives us our value of x_i'.

It might happen that we find the local minimum at P_1 initially rather than the lower minimum at P_3. This possibility can only be guarded against by scanning S_i

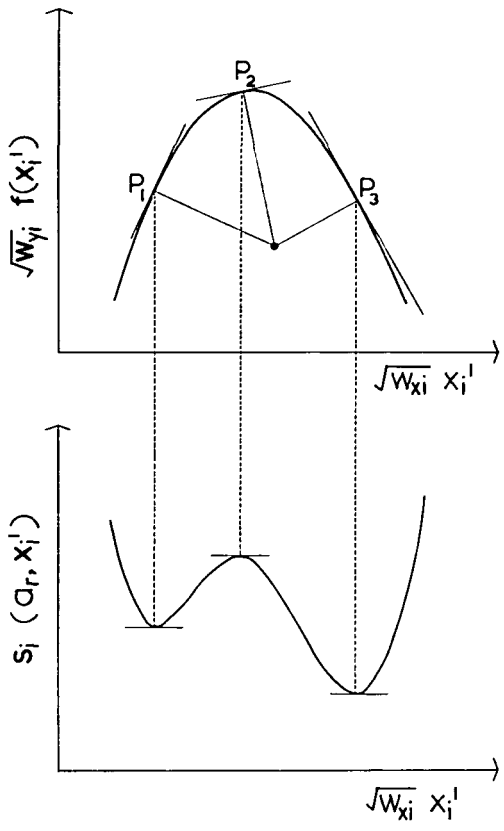

Fig. 1. A schematic illustration of an example in which S, the weighted sum of the squares of the errors, has a local maximum and two minima

when an experimental point is found inside a maximum or minimum of $f(x)$.

Another difficulty occurs if x_{i0}' is close to either of the two points of inflection, one of which is between P_1 and P_2 and the other between P_2 and P_3. Then the denominator in (16) will be small and the correction to x_{i0}' might be unrealistically large. This can be guarded against in a computer program by checking that this correction is never larger than some specified value, say the larger of $\frac{1}{2}|x_{i+1} - x_i|$ and $\frac{1}{2}|x_i - x_{i-1}|$. If it is larger then S_i must be scanned and the minimum found as before.

2.4. A possible improvement

In practice we have found that at least two iterations and sometimes more than 100 iterations were needed to obtain convergence in our algorithm. We thought this might be improved by obtaining a better approximation to the inverse G^{-1} than that given in (13). We obtained a better inverse by noting that if X_1 is an approximate inverse to the matrix G, then X_2 is a better approximation if

$$X_2 = X_1[2I - GX_1]. \tag{23}$$

Putting X_1 equal to the expression in (13) our improved inverse is given by

$$G^{-1} = \begin{bmatrix} A^{-1} & Q \\ Q' & B^{-1} \end{bmatrix} \tag{24}$$

where

$$Q = -A^{-1}CB^{-1}. \tag{25}$$

The matrices A, B and C are defined in (7). Then

$$a_r = a_{r0} - [A^{-1}g^{(1)}]_r - [Qg^{(2)}]_r \tag{26}$$

and

$$x_i' = x_{i0}' - [B^{-1}g^{(2)}]_r - [Q'g^{(1)}]_r. \tag{27}$$

These are identical with the equations we used previously except for the additional terms containing the matrices Q and Q'.

In practice we found that this gave hardly any improvement although it involved considerably more computation. If in the calculation of x_i' using equation (27) we use the values a_{r0} in the computation of $f(x)$ and its derivatives, then from its definition it is easily shown that $g^{(1)} = 0$. So the correction $Qg^{(1)}$ in (27) also is zero. Similarly it can be shown that if equation (27) is used successively until x_i' is the minimum point of S_i then in the subsequent calculation of a_r in (26) $g^{(2)} = 0$ and the extra correction in (26) is zero. Thus no improvement is obtained. If, however, we use the new values of a_r in the calculation of x_i' in equation (27), then $g^{(1)}$ will not be zero; nor will the correction be zero, but it will still be small. Similarly the correction in (26) will not be zero, but it will be small except in the initial step. In practice we found that the improvement was slight and rarely reduced the number of iterations although it increased the total amount of computing in each iteration. We conclude that this improvement is not worth the extra computing effort. We think that other approximations to the inverse of G are unlikely to be more efficient.

2.5. Deming's method

In a method due to Deming (1943) an approximate solution of considerable simplicity is obtained. The process we describe below is the same in principle as Deming's method but it is slightly different in practice.

At the maximum since $\dfrac{\partial S}{\partial x_i'} = 0$, we can write

$$x_i - x_i' = -\frac{\omega_{yi}}{\omega_{xi}} f'(x_i')[y_i - f(x_i')]. \tag{28}$$

Therefore, substituting in S we have, after simplification, that

$$S = \sum_i W_i[y_i - f(x_i')]^2 \tag{29}$$

where $$W_i = \omega_{yi}\left\{1 + \frac{\omega_{yi}}{\omega_{xi}}[f'(x_i')]^2\right\}.$$

Deming now uses Taylor's theorem to eliminate $f(x_i')$ and $f'(x_i')$ from these equations by assuming that

$$f(x_i') = f(x_i) + (x_i' - x_i)f'(x_i),$$
$$f'(x_i') = f'(x_i). \tag{30}$$

This is his first approximation when $f(x)$ is not a straight line. Using (28) again, S takes the form

$$S = \sum_i \omega_i[y_i - f(x_i)]^2 \tag{31}$$

where $$\omega_i = \omega_{yi}\Big/\left\{1 + \frac{\omega_{yi}}{\omega_{xi}}[f'(x_i)]^2\right\}.$$

In his second approximation Deming assumes that ω_i does not vary quickly with respect to the parameters a_r; thus ω_i is kept constant when S is differentiated with respect to the parameters a_r. Hence for $0 \leqslant r \leqslant m$ we have

$$\sum_i \omega_i[y_i - f(x_i)]p_r(x_i) = 0. \tag{32}$$

This gives us $m + 1$ linear equations in the $m + 1$ variables a_0 to a_m if we substitute an approximate value of $f'(x_i)$ into the expression for ω_i.

In practice we first solve equations (32) with ω_i replaced by ω_{yi} (equivalent to assuming that there are no errors in the x readings), then using the values of $f'(x_i)$ obtained from this solution we calculate ω_i and solve equations (32) again. The process can be repeated till convergence is reached. In each iteration we need approximately (15 nm) multiplications, (13 nm) additions and (2 nm) divisions, less than in each iteration in the previous method. In addition many fewer iterations are needed to reach convergence as we will show later.

2.6. An improvement on Deming's method

Deming's method works well in practice but it does not give the exact minimum of S because of the two approximations introduced. The first approximation in equation (30) can be eliminated by using equation (29) directly, that is by assuming W_i is constant and finding the position of x_i' by iteration. We begin with $W_i = \omega_{yi}$, calculate an approximate curve fit $f(x)$, find x_i' from (28), calculate $f'(x_i')$ and W_i, recompute a new polynomial, assuming W_i is constant and repeat till convergence is reached.

This should eliminate errors due to the first approximation and should give more accurate answers, but in practice we found that it made little difference to the polynomial fit and in some cases gave values further from the minimum than Deming's method. Obviously most of the error in Deming's method comes from keeping ω_i constant when differentiating with respect to a_r.

We conclude, therefore, that this apparent improvement is not worth the considerable extra computing effort.

3. Results

The direct iterative method in §2.2 and Deming's method have been tried on a number of real examples taken from scientific and industrial data and on some artificial examples chosen specifically to test the methods. Details of these are given in two theses by O'Neill (1967) and by Sinclair (1967). Some of the results obtained by examining these examples have already been mentioned.

Our main conclusions are that the direct iterative process does give a polynomial curve fit minimising S, the sum of the square of the errors, and that it usually does this in only a few iterations and with little computer time or computer storage. Deming's approximation is fast and so accurate that a better result will seldom be needed. Thus we suggest that Deming's method should be used to give us our first approximation to $f(x)$ in the direct iterative procedure. Usually only one iteration will then be necessary as a check.

In the rare event when the direct iterative process does not converge quickly, Aitken's process for accelerating a slowly converging sequence has been found to be useful.

We suggest that this procedure should be adopted if convergence is not reached after about 10 iterations.

We illustrate these main conclusions with three examples. The first example, due to Pearson (1901), is the fitting of a straight line to the set of data given in **Table 1**, with all of the ω_{xi} and ω_{yi} weights equal to unity. Pearson obtained, as the slope of the line,

Table 1

Data due to Pearson

i	x_i	y_i	i	x_i	y_i
1	0·0	5·9	6	4·4	3·7
2	0·9	5·4	7	5·2	2·8
3	1·8	4·4	8	6·1	2·8
4	2·6	4·6	9	6·5	2·4
5	3·3	3·5	10	7·4	1·5

−0·54556, York (1966) obtained −0·546, Deming's method gives −0·540 and the iterative approach gives Pearson's correct result, −0·54556 after 6 iterations.

Our second example is due to York who introduced the extreme set of weights given in **Table 2** for Pearson's

Table 2

Weights given by York for the data in Table 1

i	ω_{xi}	ω_{yi}	i	ω_{xi}	ω_{yi}
1	1,000	1·0	6	80	20
2	1,000	1·8	7	60	70
3	500	4·0	8	20	70
4	800	8·0	9	1·8	100
5	200	20·0	10	1·0	500

data and obtained a slope −0·477. If the x-coordinates are assumed to be free from error, the slope is −0·611. Using this as our first approximation Deming's method gives a slope, −0·463 and a value of $S = 11{\cdot}95653$ in two iterations. The direct iterative method converges slowly: it obtains a result as accurate as Deming's method only after 40 iterations, but after 125 iterations it obtains a slope, −0·481 and a value of S down to $S = 11{\cdot}86646$. Using Deming's method as our first approximation and Aitken's accelerative process beginning after six iterations we obtained a slope, −0·4805 and the lowest value of S, $S = 11{\cdot}86635$. This involved a total of 13 iterations.

In our third example we fit a polynomial of degree 3 to Pearson's data with unit weights. Neglecting errors in the x-coordinates we find that $f(x) = 5{\cdot}982 - 0{\cdot}9936x + 0{\cdot}1563x^2 - 0{\cdot}0138x^3$ with $S = 0{\cdot}60991$. Using Deming's method we obtain after 4 iterations that

$$f(x) = 5{\cdot}9988 - 1{\cdot}0050x + 0{\cdot}1570x^2 - 0{\cdot}0137x^3$$

and that S is reduced to $S = 0{\cdot}48677$. In the mixed method of Deming's approximation and Aitken's accelerative process we find after 9 iterations that $f(x)$ converges to

$$f(x) = 6{\cdot}0152 - 0{\cdot}998x + 0{\cdot}1525x^2 - 0{\cdot}0132x^3$$

with $S = 0{\cdot}48515$. This again illustrated that Deming's method gives a good first approximation, being correct to one or two units in the second decimal place.

Acknowledgements

One of us (F.J.S.) has pleasure in thanking Mr. C. D. Kemp, Professor J. Meinguet and Mr. M. J. D. Powell for some helpful suggestions.

This research has been sponsored in part by the Physics Branch, Office of Naval Research, Washington, D.C., under Contract No. F61052-68-C-0012, and in part by the Ministry of Education, Northern Ireland.

References

DEMING, W. E. (1943). *Statistical Adjustment of Data*, New York: Wiley.

FORSYTHE, G. E. (1957). Generation and use of orthogonal polynomials for data fitting with a digital computer, *J. Soc. Indust. Appl. Math.*, Vol. 5, pp. 74–88.

O'NEILL, M. (1967). Least squares fitting of a polynomial when both sets of data are subject to errors. Thesis, Queen's University, Belfast.

PEARSON, K. (1901). On lines and planes of closest fit to systems of points in space, *Phil. Mag.*, Vol. 2, pp. 559–572.

POWELL, M. J. D. (1965). A method for minimising the sum of squares of non-linear functions without calculating derivatives, *Computer Journal*, Vol. 7, pp. 303–307.

POWELL, M. J. D. (1966). *Numerical Analysis: an Introduction*, Ed. J. Walsh, London: Academic Press, pp. 143–147.

SINCLAIR, I. G. (1967). Polynomial least squares fitting when both abscissae and ordinates are subject to errors. Thesis, Queen's University, Belfast.

SMITH, F. J. (1965). An algorithm for summing orthogonal polynomial series and their derivatives with applications to curve fitting and interpolation, *Math. Comp.*, Vol. 19, pp. 33–36.

YORK, D. (1966). Least-squares fitting of a straight line, *Can. J. of Phys.*, Vol. 44, pp. 1079–1086.