Title
UCRL LECTURES ON NUMERICAL ANALYSIS AND APPLIED MATHEMATICS, LECTURE II: NUMERICAL DIFFERENTIATION

Permalink
https://escholarship.org/uc/item/4qs1p48m

Author
Mitchell, Richard.

Publication Date
1952-09-30
UCRL LECTURES ON NUMERICAL ANALYSIS AND APPLIED MATHEMATICS

Lecture II

Richard Mitchell

September 30, 1952

Berkeley, California
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
I. NUMERICAL DIFFERENTIATION

1. Introduction

Formulas for numerical differentiation may be derived in a natural way from the interpolating polynomial formulas. The errors involved may be determined and the methods may then be used to practical advantage.

2. Relations between Differences and Derivatives

By means of induction arguments it may be shown that

\[ (a) \quad \lim_{\Delta x \to 0} \frac{\Delta^n f(x)}{(\Delta x)^n} = f^{(n)}(x) = \frac{d^n f(x)}{dx^n} \]

\[ (b) \quad \Delta^n f(x) = (\Delta x)^n f^{(n)}(x + \theta n \Delta x) \text{ where } 0 < \theta < 1. \]

Relation (b) shows that there is some point between \( x \) and \( x + n \Delta x \) such that

\[ (c) \quad f(x + n\theta \Delta x) = \frac{\Delta^n f(x)}{(\Delta x)^n} \quad \text{for } 0 < \theta < 1. \]

When the function \( f(x) \) is unknown, formulas (a) and (c) aid in the evaluation of the error term arising when the difference quotient is
substituted for the derivative.

3. Evaluation of the Error

It may be shown that if \( F(x) \) is a polynomial of degree \( m \) and is equal to the true function \( f(x) \) at \( m + 1 \) points, \( x_0, x_1, \ldots, x_m \), then

\[
f(x) = F(x) + \frac{(x - x_0)(x - x_1) \ldots (x - x_m)}{(m + 1)!} f^{(m+1)}(\xi),
\]

where \( \xi \) lies in the range \( x_0 \) to \( x_m \). Differentiating,

\[
\frac{df(x)}{dx} = \frac{dF(x)}{dx} + \frac{f^{(m+1)}(\xi)}{(m + 1)!} \frac{d}{dx} \left[ (x - x_0) \ldots (x - x_m) \right]
\]

\[
+ \frac{(x - x_0) \ldots (x - x_m)}{(m + 1)!} \frac{df^{(m+1)}}{dx}(\xi).
\]

From this equation it may be seen that the error can not be evaluated at every point if \( f(x) \) is unknown, since \( \frac{df^{(m+1)}}{dx}(\xi) \) cannot be evaluated. However, the error can be evaluated at the tabulated points \( x_0, x_1, \ldots, x_m \).

Let \( g(x) = (x - x_0)(x - x_1) \ldots (x - x_m) \)

then

\[
\frac{dg(x)}{dx} = \sum_{i=0}^{m} (x - x_0)(x - x_1) \ldots (x - x_{i-1})(x - x_{i+1}) \ldots (x - x_m),
\]

\[
\left. \frac{dg(x)}{dx} \right|_{x=x_j} = (x_j - x_0)(x_j - x_1) \ldots (x_j - x_{j-1})(x_j - x_{j+1}) \ldots (x_j - x_m).
\]

Also \( g(x) \bigg|_{x=x_j} = 0 \). Thus the error or remainder term at \( x = x_j \) is
\[ R(x) = \frac{f^{(m+1)}(\xi)}{(m+1)!} (x_j - x_0) \cdots (x_j - x_{j-1})(x_j - x_{j+1}) \cdots (x_j - x_m) \]

4. Differentiation of Newton's Interpolation Formula.

Newton's interpolation formula in terms of forward differences is

\[ F(x) = f(a) + n\Delta f(a) + \frac{n(n-1)}{2!} \Delta^2 f(a) + \ldots + \frac{n(n-1)\ldots(n-m+1)}{m!} \Delta^m f(a) \]

where \( x = a + nh \).

Differentiating,

\[ F'(x) = \frac{1}{h} \frac{d}{dn} \left[ f(a) + n\Delta f(a) + \ldots \right] \]

\[ F'(a + nh) = \frac{1}{h} \left[ \Delta f(a) + \frac{2n-1}{2} \Delta^2 f(a) + \frac{3n^2-6n+2}{6} \Delta^3 f(a) + \ldots \right] \]

\[ F'(a) = \frac{1}{h} \left[ \Delta f(a) - \frac{1}{2} \Delta^2 f(a) + \frac{1}{3} \Delta^3 f(a) + \ldots + (-1)^{n-1} \frac{m!}{m} \Delta^m f(a) \right] \]

Evaluating the remainder term for the derivative of Newton's interpolation formula,

\[ R = (-1)^m \frac{m!}{h} f^{(m+1)}(\xi) \]

Now if the analytic form of \( f(x) \) is unknown, the best approximation to the error is obtained by approximately evaluating \( f^{(m+1)}(\xi) \) by means of relation (c).

\[ f^{(m+1)}(\xi) = \frac{\Delta^{m+1} f(x_0)}{(\Delta x)^{m+1}} = \frac{\Delta^{m+1} f(x_0)}{h^{m+1}} \]

since \( \xi = x_0 + nh \) and \( x_0 < \xi < x_m \).
Now \( R \sim \frac{(-1)^m}{h^m} \frac{\Delta^{m+1} f(x_0)}{(m+1)h^{m+1}} = \frac{(-1)^m}{m+1} \frac{\Delta^{m+1} f(x_0)}{(m+1)h^{m+1}} \)

Thus as \( h \) becomes smaller, the approximation becomes better by formula (a).

The second and higher derivatives are obtained the same manner.

For example, \( \delta \)

\[ F''(x) = \frac{1}{h} \frac{d}{dn} F'(a + nh). \]

5. Comparison of Standard Formulas.

Letting \( x = a_0 + nh \), the first three terms of the standard formulas are

(a) Newton's

\[ F'(a_0 + nh) = \frac{1}{h} \left[ \Delta f(a_0) + \frac{2n - 1}{2} \Delta^2 f(a_0) + \frac{3n^2 - 6n + 2}{6} \Delta^3 f(a_0) + \ldots \right] \]

(b) Stirling's

\[ F'(a_0 + nh) = \frac{1}{h} \left[ \frac{\Delta f(a_0) + \Delta f(a_0)}{2} + n \Delta^2 f(a_0) \right. \]

\[ + \left. \frac{3n^2 - 1}{3!} \frac{\Delta^3 f(a_0)}{2} + \Delta^3 f(a_0) + \ldots \right] \]

(c) Bessel's

\[ F'(a_0 + nh) = \frac{1}{h} \left[ \Delta f(a_0) + \frac{n}{2} \frac{\Delta^2 f(a_0) - \Delta^2 f(a_0)}{2} + \frac{3n^2 - 1}{3!} \frac{\Delta^3 f(a_0)}{2} + \ldots \right] \]

Newton's formula is a polynomial of order \( n \) and is generally used at the beginning or end of tables. Stirling's formula is a polynomial of
order \( 2n \) and is used most effectively close to tabulated values, for 
\( |n| \leq 1/4 \), while Bessel's formula of order \( 2n + 1 \) is used between 
\( 1/4 < n < 3/4 \).

For the tabulated values, \( n = 0 \), the formulas are

\[
\text{Newton's} \quad F'(a_0) = \frac{1}{h} \left[ \Delta f(a_0) - \frac{1}{2} \Delta^2 f(a_0) + \frac{1}{3} \Delta^3 f(a_0) + \ldots \right] \\
\text{Stirling's} \quad F'(a_0) = \frac{1}{h} \left[ \frac{\Delta f(a-1) + \Delta f(a_0)}{2} - \frac{1}{3!} \frac{\Delta^3 f(a-2) + \Delta^3 f(a-1)}{2} + \ldots \right] \\
\text{Bessel's} \quad F'(a_0) = \frac{1}{h} \left[ \Delta f(a_0) - \frac{1}{4 \times 3!} \Delta^3 f(a-1) + \ldots \right]
\]

By observation it is seen that Newton's formula for tabulated 
values is the easiest to compute with. However, if only first approximations 
are wanted, then the first term of Stirling's equation is the most satisfactory.

6. Maxima and Minima

These formulas can be used to calculate the value of \( x \) at a 
maximum or minimum. Using Newton's formula, where \( x = a_0 + nh \),

\[
F'(a_0 + nh) = 0 = \frac{1}{h} \left[ \Delta f(a_0) + \frac{2n - 1}{2} \Delta^2 f(a_0) \\
+ \frac{3n^2 - 6n + 2}{6} \Delta^3 f(a_0) + \ldots \right]
\]

Using the first three terms,

\[
0 = \frac{1}{2} n \Delta^3 f(a_0) + n(\Delta^2 f(a_0)) - \Delta^3 f(a) + \left[ \Delta f(a) - \frac{1}{2} \Delta^2 f(a) + \frac{1}{3} \Delta^3 f(a) \right]
\]
If, upon solving this formula, \( n \) is greater than one, then shifting position in the table will improve the solution. Also, taking more terms will improve the accuracy of \( x \).

7. Partial Derivatives

Using the double interpolation formula, \( \frac{\partial z}{\partial x} \) and \( \frac{\partial z}{\partial y} \) may be computed for \( z = f(x, y) \). Using arguments, \( x = a_0 + uh \) and \( y = b_0 + vk \)

where

\[
h = a_1 - a_0 = a_2 - a_1 = \ldots = a_m - a_{m-1}
\]

\[
k = b_1 - b_0 = b_2 - b_1 = \ldots = b_m - b_{m-1}
\]

\[
z = f(x, y) = z_{oo} + u \Delta^{1+o} z_{oo} + v \Delta^{o+1} z_{oo} + \frac{1}{2} \left[ u(u-1) \Delta^{2+o} z_{oo} + 2uv \Delta^{1+1} z_{oo} + v(v-1) \Delta^{o+2} z_{oo} \right]
\]

\[
+ \frac{1}{3!} \left[ u(u-1)(u-2) \Delta^{3+o} z_{oo} + 3u(u-1)v \Delta^{2+1} z_{oo} + 3uv(v-1) \Delta^{1+2} z_{oo} + v(v-1)(v-2) \Delta^{o+3} z_{oo} \right] + \ldots
\]

where

\[
\Delta^{m+n} z_{oo} = \Delta^{m+o} z_{oo} - n \Delta^{m+o} z_{o,n-1} + \binom{n-1}{2} \Delta^{m+o} z_{o,n-2} + \ldots + \Delta^{m+o} z_{oo}
\]

or

\[
\Delta^{m+n} z_{oo} = \Delta^{o+n} z_{mo} - m \Delta^{o+n} z_{m-1,o} + \frac{m(m-1)}{2} \Delta^{o+n} z_{m-2,o} + \ldots + \Delta^{o+n} z_{oo}
\]

and \( z_{ik} = a_i b_k \).
The error term is derived in Scarborough, first edition.

Differentiating,

\[
\frac{\partial z}{\partial x} = \frac{\partial u}{\partial x} \frac{\partial z}{\partial u} = \frac{1}{h} \frac{\partial z}{\partial u}
\]

\[
\frac{\partial z}{\partial y} = \frac{1}{k} \frac{\partial z}{\partial v}
\]

\[
\frac{\partial z}{\partial x} = \frac{1}{h} \left\{ \Delta^{1+0} z_{oo} + \frac{1}{2!} \left[ (2u - 1) \Delta^{2+0} z_{oo} + 2v \Delta^{1+1} z_{oo} \right] \right. \\
+ \frac{1}{3!} \left[ (3u^2 - 6u + 2) \Delta^{3+0} z_{oo} + (6u - 3)v \Delta^{2+1} z_{oo} \right. \\
+ 3v(v - 1) \Delta^{1+2} z_{oo} \right\} + \ldots
\]

\[
\frac{\partial z}{\partial y}
\]

is similarly obtained.
II. CURVE FITTING

1. Introduction

The subject of curve fitting is twofold. First a function must be found to fit the given data and secondly the constants of the function must be evaluated in order to obtain the best fit of the function. The first problem is concerned with finding an empirical function of which there may be many. In the second case the exact functional form of the curve may be known and the problem is to find the best fit to the observational data. Since there is such a large number of possible approximating functions, no systematic method of finding the best empirical function is possible.

2. Graphical Methods

It is very often possible to use function scales which are so chosen that the graph of the function becomes approximately a straight line. Semi-log and log-log graph papers are commonly used.

If \( y = k a^x \), then \( \log y = \log k + (m \log a)x \) and \( \log y \) and \( x \) are linearly related.

If \( y = b x^n \) then \( \log y = \log b + n \log x \) and \( \log y \) and \( \log x \) are linearly related.

In general, if \( N(y) = m \cdot f(x) + b \), where \( f \) and \( N \) are function scales for \( x \) and \( y \), then \( f(x) \) and \( N(y) \) are linearly related.

3. Use of Differences

Using the observed pairs, \((x_i, y_i)\), a table of differences of the \( y_i \) is set up. If the \( x_i \) form an arithmetic progression, \( x_i - x_{i+1} = k \) for all \( i \), and if the \( r \)'th difference of the \( y_i \) is constant, all higher
differences being zero, then \( y \) is a polynomial in \( x \) of degree \( r \). If the 
\( r' \)th difference is almost constant, then a polynomial of degree \( r' \) will be 
a good approximation to \( y(x) \).

Modifying the conditions on the \( x \)'s and \( y \)'s, other possible 
representations of \( y(x) \) may be found. If the values of \( x_i \) form an 
arithmetic progression and if the \( r' \)th difference of \( y_1 \) is constant or 
almost constant, then 
\[
y_1^m = a_0 + a_1 x^n + a_2 (x^n)^2 + \cdots + a_r (x^n)^r.
\]

If \( m = n = -1 \) and \( r = 1 \), then \( y = \frac{x}{a_1 + a_0 x} \) and the 
function \( y \) has asymptotes \( x = -\frac{a_1}{a_0} \) and \( y = \frac{1}{a_0} \). In this case

\( 1/y \) versus \( 1/x \) is a straight line.

If the \( x_i \)'s form an arithmetic progression and the \( r' \)th differences 
of the \( y_i \) form a geometric progression, \( \Delta^r y_i = k \Delta^r y_{i+1} \) for all \( i \), then

\[
y = a_0 + a_1 x + \cdots + a_{r-1} x^{r-1} + k a x^r.
\]

If \( r = 1 \), then \( y = a_0 + k a x \).

If \( a > 1 \), then \( y \) increases indefinitely with \( x \).

If \( 0 < a < 1 \), then \( y \) decreases with \( x \) and is asymptotic to

\( y = a_0 \).

Now if the \( x_i \)'s form a geometric progression, and the \( y_i \)'s also 
form a geometric progression, then \( y = a x^n \).

4. Evaluation of Arbitrary Constants

4.1 Introduction

Substituting the observational values, \((x_i, y_i), i=1, \ldots, n,\) 
into the approximating function \( f(x) \), relationships which the constants 
must satisfy are obtained.

\[
y_i - f(x_i) = 0 \quad \text{for } i = 1, \ldots, n.
\]
If there are \( r \), less than \( n \), arbitrary constants, then \( r \) of the \( n \) equations involving the constants may be selected in hopes of obtaining the best approximation. An alternative is to stipulate that

\[
  f(x_i) - y_i = f(x_j) - y_j
\]

for certain points. However care must be exerted in order that a very large negative error does not cancel a large positive error.

In these two common procedures there is no systematic method, nor is there a reliable estimate of the probable error at every point, nor is there assurance that the best fit has been found. On the other hand the method of least squares has the advantage that it satisfies these three requirements.

4.2 Method of Least Squares.

If the function \( y(x) \) is to be approximated by

\[
  f(x) = \sum_{k=0}^{m} a_k g_k(x)
\]

where the \( g_k(x) \) are linearly independent and \( n > m \), then the method of least squares asserts that the residuals \( v_i = f(x_i) - y(x_i) \) should be such that \( \sum_{i=1}^{n} v_i^2 \) is a minimum.

By partially differentiating \( \sum_{i=1}^{n} v_i^2 \) with respect to the \( m + 1 \) arbitrary constants and equating to zero, the "normal equations" are obtained which must be satisfied by the arbitrary constants in order that \( \sum_{i=1}^{n} v_i^2 \) shall be a minimum.
\[ a_0 \sum_{i=1}^{n} g_0(x_i) + a_1 \sum_{i=1}^{n} g_0(x_i) g_1(x_i) + \cdots + a_m \sum_{i=1}^{n} g_0(x_i) g_m(x_i) = \sum_{i=1}^{n} g_0(x_i) y(x_i) \]

\[ a_0 \sum_{i=1}^{n} g_1(x_i) g_0(x_i) + a_1 \sum_{i=1}^{n} g_1(x_i) + \cdots + a_m \sum_{i=1}^{n} g_1(x_i) g_m(x_i) = \sum_{i=1}^{n} g_1(x_i) y(x_i) \]

\[ \vdots \]

\[ a_0 \sum_{i=1}^{n} g_m(x_i) g_0(x_i) + a_1 \sum_{i=1}^{n} g_m(x_i) g_1(x_i) + \cdots + a_m \sum_{i=1}^{n} g_m(x_i) = \sum_{i=1}^{n} g_m(x_i) y(x_i) \]

There are \( m + 1 \) equations in the \( m + 1 \) unknowns, \( a_0, a_1, \ldots, a_m \). There will be a solution if the determinant of the coefficients is not zero.

4.3 Applications of Least Square Method

4.3.1 Linear Functions

If \( y \) is to be approximated by \( f(x) = ax + b \), then the normal equations are

\[ b \sum_{i=1}^{n} x_i + a \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \]
or
\[
a = \frac{\sum x_i \sum y_i - n \sum x_i y_i}{\sum x_i \sum x_i - n \sum x_i^2}
\]
and
\[
b = \frac{\sum x_i y_i \sum x_i - \sum y_i \sum x_i^2}{\sum x_i \sum x_i - n \sum x_i^2}
\]

4.32 Polynomials

Let \( f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \).

In order to simplify the problem, \( x \) will be transformed so that the interval \((a, b)\) in which we are interested becomes the interval \((-1, +1)\).

Then \( x = \frac{a + b - a - b}{2} t \) and \( f(t) = c_0 + c_1 t + c_2 t^2 + \ldots + c_n t^n \).

where \( g_k(t) = t^k \). Now \( \sum_{i=1}^{n} t_i^{2k+1} = 0 \) when \( y \) is tabulated in equal intervals of \( x \). The normal equations become

\[
nc_0 + c_2 \sum t_i^2 + c_4 \sum t_i^4 + \ldots = \sum y_i
\]
\[
c_1 \sum t_i^2 + c_3 \sum t_i^4 + \ldots = \sum t_i y_i
\]
\[
c_0 \sum t_i^2 + c_2 \sum t_i^4 + c_4 \sum t_i^6 + \ldots = \sum t_i^2 y_i
\]

The normal equations thus divide into two groups, one with even numbered coefficients and the other with odd numbered coefficients. The solution now is much easier to find.
4.33 Computing procedures

The following examples are given to illustrate how the work of computation with a calculator may be reduced. A tabulation form with positive checks is illustrated since most mistakes enter into the calculations at the stage of tabulation. The abridged form for solving the normal equations is simple but the computer must be careful of the loss of significant figures since the process may involve subtraction of numbers of the same order of magnitude.

Example 1: Applied least square method.

Let \( z(x, y) = ax + by + c \) where \( (x_1, y_1, z_1) \) are observed values.

\[
\sum_{i=1}^{n} v_i^2 = \sum_{i=1}^{n} (a x_1 + b y_1 + c - z_1)^2
\]

\[
\frac{\partial \sum v_i^2}{\partial a} = 0 ; \quad \frac{\partial \sum v_i^2}{\partial b} = 0 ; \quad \frac{\partial \sum v_i^2}{\partial c} = 0 .
\]

\[
n c + a \sum x_1 + b \sum y_1 = \sum z_1
\]

\[
c \sum x_1 + a \sum x_1^2 + b \sum x_1 y_1 = \sum x_1 z_1
\]

\[
c \sum y_1 + a \sum y_1 x_1 + b \sum y_1^2 = \sum y_1 z_1 .
\]
Tabulation with positive checks

<table>
<thead>
<tr>
<th>$x_1$</th>
<th>$y_1$</th>
<th>$z_1$</th>
<th>$x_1^2$</th>
<th>$x_1y_1$</th>
<th>$y_1^2$</th>
<th>$x_1z_1$</th>
<th>$y_1z_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_2$</td>
<td>$y_2$</td>
<td>$z_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_3$</td>
<td>$y_2$</td>
<td>$z_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_n$</td>
<td>$y_n$</td>
<td>$z_n$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\sum x_i \quad \sum y_i \quad \sum z_i \quad \sum x_i^2 \quad \sum x_1y_1 \quad \sum y_1^2 \quad \sum x_1z_1 \quad \sum y_1z_1
\]

\[
1 + x_1 + y_i \quad (1 + x_1 + y_i)^2 \quad z_1(1 + x_1 + y_i)
\]

\[
1 + x_1 + y_2
\]

\[
1 + x_2 + y_2
\]

\[
1 + x_3 + y_3 \quad \text{accumulate on machine}
\]

\[
\ldots
\]

\[
1 + x_n + y_n
\]

\[
\sum (1 + x_i + y_i) \quad \sum (1 + x_i + y_i)^2 \quad \sum z_i(1 + x_i + y_i)
\]
Checks:

1. \[ \sum (1 + x_i + y_i) = \sum x_i + \sum y_i + n \]
2. \[ \sum (1 + x_i + y_i)^2 = \sum x_i^2 + \sum y_i^2 + 2 \sum x_iy_i + \sum x_i + \sum y_i + n \]
3. \[ \sum z_i (1 + x_i + y_i) = \sum z_i + \sum z_i x_i + \sum z_i y_i \]

Example 2: Doolittle's Abridged Method.

Let the normal equations be

\[ a_{00} k + a_{10} l + a_{20} m = b_0 \]
\[ a_{01} k + a_{11} l + a_{21} m = b_1 \]
\[ a_{02} k + a_{12} l + a_{22} m = b_2 \]

where \( a_k l = a_{kl} \) 

<table>
<thead>
<tr>
<th>I</th>
<th>( a_{00} )</th>
<th>( a_{10} )</th>
<th>( a_{20} )</th>
<th>( b_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2)</td>
<td>( a_{11} )</td>
<td>( a_{21} )</td>
<td>( b_1 )</td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>( a_{22} )</td>
<td>( b_2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>( b_0 ) + ( a_{10} + a_{20} )</td>
<td>( b_1 + a_{11} + a_{21} )</td>
<td>( b_2 + a_{20} + a_{21} )</td>
<td>( b_1 + a_{10} + a_{11} + a_{21} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(4)</th>
<th>( \frac{a_{10}}{a_{00}} )</th>
<th>( \frac{-a_{20} - a_{10}}{a_{00}} )</th>
<th>( \frac{-b - a_{10}}{a_{00}} )</th>
<th>( \frac{-a_{10}}{a_{00}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_{00} )</td>
<td>( \frac{-a_{10} a_{10}}{a_{00}} )</td>
<td>( \frac{-a_{20} a_{10}}{a_{00}} )</td>
<td>( \frac{-b_{0} a_{10}}{a_{00}} )</td>
<td>( \frac{-a_{10}}{a_{00}} )</td>
</tr>
</tbody>
</table>
| \( b_{-1} \) | \( \left[ b_{0} + a_{00} + a_{10} + a_{20} \right] \)
II. (2) + (4) \[ a_{11} a_{21} b_1 \quad b_1 + a_{11} + a_{21} \quad \text{(check)} \]

(3) \[ a_{22} b_2 \quad b_2 + a_{20} + a_{21} + a_{22} \]

(5) \[ I \times \frac{a_{20}}{a_{00}} \quad a_{20} \quad \frac{-a_{20}}{a_{00}} \quad b_0 \quad \frac{-a_{20}}{a_{00}} \quad a_{20} \quad \left[ b_0 + a_{00} + a_{10} + a_{20} \right] \]

(6) \[ II \times \frac{a_{21}}{a_{11}} \quad a_{21} \quad \frac{-a_{21}}{a_{11}} \quad b_2 \quad \frac{-a_{21}}{a_{11}} \quad a_{21} \quad \left[ b_1 + a_{11} + a_{21} \right] \]

III. (3) + (5) + (6) \[ a_{22} b_2 \quad b_2 \quad a_{22} \quad \text{(check)} \]

Now the normal equations can be solved easily.

\[ a_{22} m = b_2 \]
\[ a_{11} \ell + a_{21} \cdot m = b_1 \]
\[ a_{00} k + a_{10} \ell + a_{20} \cdot m = b_0 \]

4.4 Integral Cases

If continuous curves rather than discrete observations are being considered, then \[ \int_a^b \left[ f(x) - y(x) \right]^2 \, dx \] is minimized. In the normal equations is substituted for \[ \sum_{i=1}^{n} \]. Transforming the interval.
(a, b) to (-1, +1), there are terms like \( \int_{-1}^{+1} g_m(t)g_n(t)\,dt \) where

\[
f(t) = \sum_{m=0}^{n} a_m t^m \quad \text{and} \quad g_m(x) = x^m
\]

and \( g_m(t) = t^m \).

For the general term,

\[
\int_{-1}^{+1} g_m(t)g_{\xi}(t)\,dt = \int_{-1}^{+1} g_{m+\xi}\,dt = \begin{cases} \\
\frac{2}{m+\xi+1} & m+n \text{ odd} \\
0 & m+n \text{ even}
\end{cases}
\]

The other terms are

\[
\int_{-1}^{+1} g_k(t) y\,dt = \int_{-1}^{+1} t^k y\,dt
\]

\[
2 J_0 = \int_{-1}^{+1} y\,dt
\]

\[
2 J_1 = \int_{-1}^{+1} y t\,dt
\]

\[
2 J_2 = \int_{-1}^{+1} y t^2\,dt, \quad \text{etc.}
\]

The normal equations are now
This set of normal equations divides into odd and even groups and the a's can be written in terms of the J's.

The disadvantage of the power series approximation is that the constants must be reevaluated if additional terms are added for greater accuracy.

If orthogonal polynomials are used, then only the diagonal terms in the normal equations remain and, as a consequence, the previous calculations do not change with the addition of more terms. Then

$$a_k = \frac{1}{b_k} \int_{-1}^{+1} g_k(t) y \, dt$$

where

$$b_k = \int_{-1}^{+1} \left[ g_k(t) \right]^2 \, dt$$

4.5 Fundamental Theorem.

The fundamental theorem states that the approximating function determined by the least square method is closer to the true function than the observed function.

Let \( z \) be a continuous function in \((a, b)\) which is to be approximated by
\[ y(x) = \sum_{i=1}^{m} a_i v_i(x) \] where
\[ v_1(x), v_2(x), \ldots, v_m(x) \] are linearly independent continuous functions of \( x \) in the interval \( (a, b) \). The \( a \)'s are determined so as to minimize the integral
\[ I = \int_{a}^{b} w(x - \sum_{i=1}^{m} a_i v_i)^2 \ dx \] where \( w(x) \), a weight function, is a continuous non-negative function of \( x \) in \( (a, b) \).

Now \( z(x) \) has been obtained from observations subject to random errors. There is a true function \( u(x) \) which would have been found if there were no errors in observation.

**Theorem:** If the true function \( u(x) \) can be represented by \( v \)'s, which are linearly independent, so that
\[ u(x) = b_1 v_1 + b_2 v_2 + \ldots + b_n v_n \],

and if \( a \)'s are determined so that
\[ y(x) = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n \]
is the least square approximation to the observed function \( z(x) \), then
\[ \int_{a}^{b} w(z - u)^2 \ dx = \int_{a}^{b} w(z - y)^2 \ dx + \int_{a}^{b} w(y - u)^2 \ dx \].

This states that except for \( y = z \), a trivial case,
\[ \int_{a}^{b} w(y - u)^2 \ dx < \int_{a}^{b} (z - u)^2 \ dx \]
or that \( y \) is closer to the true function than \( z \), the observed function.

4.6 Non-linear Cases

It was previously assumed that the arbitrary constants appeared linearly in the equations. This is not the case in general. Consequently if this method is to be applied to non-linear equations, they must be transformed by some means into a linear form.

If \( N(y) = a \xi(x) + b \) and if \( (x_i, y_i) \) are the observational values, then

\[
\Delta N_i = N_i - a \xi_i - b \approx N'(y_i) \Delta y_i
\]

\[
\Delta y_i = \frac{\Delta N_i}{N'(y_i)} = \frac{N_i}{N'(y_i)} - a \frac{\xi_i}{N'(y_i)} - b \frac{1}{N'(y_i)}
\]

The normal equations are

\[
\frac{\partial}{\partial \theta} \sum \Delta y_i^2 = 0 = \sum \frac{N_i}{[N'(y_i)]^2} - a \sum \frac{\xi_i}{[N'(y_i)]^2} - b \sum \frac{1}{[N'(y_i)]^2}
\]

\[
\frac{\partial}{\partial \alpha} \sum \Delta y_i^2 = 0 = \sum \frac{N_i \xi_i}{[N'(y_i)]^2} - a \sum \frac{\xi_i^2}{[N'(y_i)]^2} - b \sum \frac{\xi_i}{[N'(y_i)]^2}
\]

If it is not possible to transform the equations into a linear form, but if an approximate solution to the problem is known, then by means of Taylor's expansion about the approximate solution, a linear equation for correction terms can be written. This is sometimes called a differential correction method.
Let \( y = f(x, a, b, c) \) be the equation involving the constants in a non-linear way and let \( a_0, b_0, c_0 \) be a known approximate solution. We wish
\[
\begin{align*}
\Delta a &= a - a_0 \\
\Delta b &= b - b_0 \\
\Delta c &= c - c_0
\end{align*}
\]
where \( \Delta a, \Delta b, \Delta c \) are small.

\[
v_i = f(x_i, a, b, c) - y_i
\]

\[
v_i = \left[ f(x_i, a_0 + \Delta a, b_0 + \Delta b, c_0 + \Delta c) - y_i \right]
\]

\[
= \left[ f(x_i, a_0, b_0, c_0) + \frac{\partial f}{\partial a}(x_i, a_0, b_0, c_0) \Delta a + \frac{\partial f}{\partial b}(x_i, a_0, b_0, c_0) \Delta b + \frac{\partial f}{\partial c}(x_i, a_0, b_0, c_0) \Delta c \right. \\
\left. + \text{higher order terms} \right] - y_i
\]

If the higher order terms are neglected, then \( \Delta a, \Delta b, \Delta c \) will be in a linear form and the terms for the normal equations can be computed. This gives a solution for \( \Delta a, \Delta b, \Delta c \): Repetition of this process will provide a check on the computation and will also improve the evaluation of \( a, b, \) and \( c \) to within the limits of the experimental error.

5. Criticism

It must be recognized that the argument was assumed all along to be exact. This is not true in general for most experimental data. It is hoped that the error made in assuming the argument exact is small and unimportant in the desired answer.
If the error in the argument is of the same order of magnitude as that in the function, then two curves could be drawn through a given set of points, one for exact argument and one for exact function.

It is presumed that the best curve lies between these curves. However, the real advantage of the least square fit is now lost since there is no way of judging the best fit or the error involved.

The best discussion of the estimation of errors is found in Birge's notes.

6. Bibliography

6.1 Milne, "Numerical Calculus".

6.2 Sokolinkoff, "Higher Mathematics for Engineers and Physicists".

6.3 Willer, "Practical Analysis".

6.4 Scarborough, "Numerical Mathematical Analysis".

6.5 Birge, R. T., Chairman of Physics Department, University of California, Lecture notes for Physics 204.