DECUS

PROGRAM LIBRARY

DECUS NO.	FOCAL8-61
TITLE	LEAST SQUARE FIT TO A POLYNOMIAL
AUTHOR	Adrian Demayo
COMPANY	Department of Energy, Mines and Resources Inland Water Branch, Water Quality Division Ontario, Canada
DATE	November 11, 1969
SOURCE LANGUAGE	FOCAL

LEAST SQUARE FIT TO A POLYNOMIAL

DECUS Program Library Write-up — DECUS NO. FOCAL8-61

Given L1 pairs of points X_j(obs), Y_j(obs) (j=1...L1) this program finds the coefficients B_i of expression:

$$Y_j\,(\text{calc}) = \sum_{i=1}^{L} B_i X_j^{\,i-1}\,(\text{obs}) \qquad L = NA\ldots NB \qquad j = 1\ldots L1 \qquad (1)$$

which minimizes the sum:

$$\sum_{j=1}^{L1} \left(Y(\text{obs}) - Y(\text{calc})\right)^2$$

The coefficients B_i are found by solving a linear system of L equations given in a general form by:

$$\sum_{j=1}^{L1} \left(Y_j\,(\text{Obs}) * X_j^{\,m}\,(\text{obs})\right) = \sum_{i=1}^{L1} \left[B_i * \sum_{j=1}^{L1} X_j^{\,i+m-1}\,(\text{obs}) \right] \qquad (2)$$

$$m = 0,1,\ldots,L-1.$$

In an expanded form the equations are shown in 3.

$$\sum_{j=1}^{L1} Y_i(obs)*X_i^{0}(obs) = B_1 * \sum_{j=1}^{L1} X_i^{0}(obs) + B_2 * \sum_{j=1}^{L1} X_i(obs) + \ldots + B_i * \sum_{j=1}^{L1} X_i^{i-1}(obs) + \ldots + B_L * \sum_{j=1}^{L1} X_i^{L-1}(obs)$$

$$\sum_{j=1}^{L1} Y_i(obs)*X_i(obs) = B_1 * \sum_{j=1}^{L1} X_i(obs) + B_2 * \sum_{j=1}^{L1} X_i^{2}(obs) + \ldots + B_i * \sum_{j=1}^{L1} X_i^{i}(obs) + \ldots + B_L * \sum_{j=1}^{L1} X_i^{L}(obs)$$

3

. . .

$$\sum_{j=1}^{L1} Y_i(obs)*X_i^{m}(obs) = B_1 * \sum_{j=1}^{L1} X_i^{m}(obs) + B_2 * \sum_{j=1}^{L1} X_i^{m+1}(obs) + \ldots + B_i * \sum_{j=1}^{L1} X_i^{i+m-1}(obs) + \ldots + B_L * \sum_{j=1}^{L1} X_i^{L+m-1}(obs)$$

. . .

$$\sum_{j=1}^{L1} Y_i(obs)*X_i^{L-1}(obs) = B_1 * \sum_{j=1}^{L1} X_i^{L-1}(obs) + B_2 * \sum_{j=1}^{L1} X_i^{L+m}(obs) + \ldots + B_i * \sum_{j=1}^{L1} X_i^{i+L-2}(obs) + \ldots + B_L * \sum_{j=1}^{L1} X_i^{2(L-1)}(obs)$$

L, the number of adjustable coefficients B_i, can vary between two limits, NA and NB (NA can also be equal to NB) which are specified in the input of the program. Thus:

$$L = NA, NA+1, \ldots, NB$$

The maximum values for NB and L1 are given in table 1.

Table 1

NB		Maximum value for L1	
	Memory size:	4K	8K
2		22	64
3		19	61
4		16	58
5		12	54
6		7	49
7		-	44
8		-	38
9		-	31
10		-	24
11		-	15

The program contains a subroutine for the solution of a system of linear equations (Statements nos. 15.0 + 14.0). The system is solved by Gauss-Jordan reduction method. The steps of this method which the program follows very closely are as follows. Given the systems of equations:

$$A_{11}X_1 + A_{12}X_2 + \ldots + A_{1i}X_i + \ldots + A_{1L}X_L = C_1$$

$$A_{21}X_1 + A_{22}X_2 + \ldots + A_{2i}X_i + \ldots + A_{2L}X_L = C_2$$

$$\ldots \qquad (4)$$

$$A_{m1}X_1 + A_{m2}X_2 + \ldots + A_{mi}X_i + \ldots + A_{mL}X_L = C_m$$

$$\ldots$$

$$A_{L1}X_1 + A_{L2}X_2 + \ldots + A_{Li}X_i + \ldots + A_{LL}X_L = C_L$$

Consider the $L*(L+1)$ matrix formed by the coefficients A and terms C:

$$\begin{matrix} A_{11} & A_{12} & \dots A_{1i} & \dots A_{1L} & & C_{1} \\ A_{21} & A_{22} & \dots A_{2i} & \dots A_{2L} & & C_{2} \\ \dots & & & & & \\ A_{m1} & A_{m2} & \dots A_{mi} & \dots A_{mL} & & C_{m} \\ \dots & & & & & \\ A_{L1} & A_{L2} & \dots A_{Li} & \dots A_{LL} & & C_{L} \end{matrix} \qquad 5$$

a. Divide the 1st row by A_{11}. This makes element $A_{11}=1$.

b. Perform the subtraction: 2nd row - 1st row $*A_{21}$. This operation makes $A_{21}=0$.

c. Perform the subtraction: jth row - 1st row $*\ A_{j1}$. This operation makes $A_{j1}=0$.

d. Repeat c until $j=L$.

Now the matrix 5 has the following aspect:

$$\begin{matrix} 1 & A_{12} & A_{13} & \dots A_{1i} & \dots A_{1L} & & C_{1} \\ 0 & A_{22} & A_{23} & \dots A_{2i} & \dots A_{2L} & & C_{2} \\ \dots & & & & & & \\ 0 & A_{m2} & A_{m3} & \dots A_{mi} & \dots A_{mL} & & C_{m} \\ \dots & & & & & & \\ 0 & A_{L2} & A_{L3} & \dots A_{Li} & \dots A_{LL} & & C_{L} \end{matrix} \qquad 6$$

It must be pointed out that, though we used the same notation the elements of matrix 6 have different values from those of matrix 5.

e. Operations a – d are repeated starting this time with element A_{22} and row 2 and leaving 1st row alone. At the end of this series of calculation matrix 6 has become:

$$
\begin{array}{cccccccccc}
1 & A_{12} & A_{13} & \ldots & A_{1i} & \ldots & A_{1L} & & C_1 \\
0 & 1 & A_{23} & \ldots & A_{2i} & \ldots & A_{2L} & & C_2 \\
0 & 0 & A_{33} & \ldots & A_{3i} & \ldots & A_{3L} & & C_3 \\
\ldots & & & & & & & & \\
0 & 0 & A_{m3} & \ldots & A_{mi} & \ldots & A_{mL} & & C_m \\
\ldots & & & & & & & & \\
0 & 0 & A_{L3} & \ldots & A_{Li} & \ldots & A_{LL} & & C_L
\end{array}
\qquad (7)
$$

The same operations are repeated for each diagonal element and for each row. At the end the matrix of coefficients has become:

$$
\begin{array}{cccccccc}
1 & A_{12} & A_{13} & \ldots A_{1i} & \ldots A_{iL} & & C_1 \\
0 & 1 & A_{23} & \ldots A_{2i} & \ldots A_{2L} & & C_2 \\
\ldots & & & & & & \\
0 & 0 & 0 & \ldots 0 \ldots 1 & A_{L-1,L} & & C_{L-1} \\
0 & 0 & 0 & \ldots 0 \ldots 0 & A_{LL} & & C_L
\end{array}
$$

Thus, from the last row X_L can be obtained immediately. This value is substituted in row L-1 and X_{L-1} obtained, and so on, until all the X's are obtained.

LOADING THE PROGRAM TAPE AND INPUT/OUTPUT FORMAT.

Assume that the FOCAL-69 tape is already in the computer (no extended functions). If they were present delete them manually by using the patch given in ADVANCED FOCAL MANUAL, Digital Equipment Corp., Publ. No. DEC-08-AJBB-DL, p. 4-9. Set SR at 0200. Press STOP, LOAD, START. Enter via keyboard: ERASE ALL and press the key labelled "Return." Enter via tape reader the program tape. Enter via keyboard: Return-GO-Return. Now the program is ready to accept the data in the following order:

NA, NB, L1, and X-Y pairs. (Refer to the above write-up for the meaning of these symbols).

The output consists, first of all, of a table of X-Y values. Then the adjustable coefficients B_i as they were calculated by the program are typed out. Finally a table of Y (obs), Y(calc) and the difference between them is typed out. The last line of the output contains the value of the mean square deviation calculated from expression:

$$SD=\left[\frac{\sum_{i=1}^{L1}\left(Y(obs)_i-Y(calc)_i\right)^2}{L1-L}\right]^{1/2}$$

Note that this treatment assumes that Y coordinates are subject to error while X coordinates are free of error.

To restart the program with a new series of values enter via the keyboard Return-GO-Return and proceed as before.

REFERENCES

1. W. Edwards Deming, Statistical Adjustment of Data, Dover Publications Inc., New York, N. Y., 1964.

```
*
*C-FOCAL, 1969
*
*Ø1.Ø6  E
*Ø1.Ø7  A    NA,NB,L1
*Ø1.Ø9  F    J1=1,L1;A X(J1),Y(J1)
*Ø1.1Ø  T    T ! !"DATA POINTS :",!!,"  NO.          X                    Y",!
*Ø1.11  F    J1=1,L1;T %2,J1,"          ",%,X(J1),"        ",Y(J1),!
*Ø1.15  F    L=NA,NB;D 2
*Ø1.2Ø  Q
*
*Ø2.17  S    N2=2*L-1
*Ø2.2Ø  F    J1=1,N2;S SX(J1)=Ø
*Ø2.22  F    J1=1,L;S YX(J1)=Ø
*Ø2.25  F    J2=1,N2;F J1=1,L1;S SX(J2)=SX(J2)+X(J1)↑(J2-1)
*Ø2.27  F    J2=1,L;F J1=1,L1;S YX(J2)=YX(J2)+Y(J1)*(X(J1)↑(J2-1))
*Ø2.28  T    !!"NO. OF ADJUSTABLE COEFFICIENTS",%2,L
*Ø2.3Ø  F    J2=1,L;F J1=1,L;S A(J2+J1*L)=SX(J1+(J2-1))
*Ø2.35  D    15.Ø
*Ø2.78  F    K=1,L; T !"B(",%2,K,")      ",%,B(K)
*Ø2.8Ø  F    J1=1,L1;S YX(J1)=Ø
*Ø2.82  S    SD=Ø
*Ø2.84  T    !!" NO.        Y OBS.                Y CALC.                DIF",!
*Ø2.85  F    J1=1,L1;S YX=Ø;D 8.Ø
*Ø2.86  S    SD=FSQT(SD/(L1-L))
*Ø2.88  T    !"MEAN SQUARE DEVIATION",SD;R
*
*Ø8.1Ø  F    J2=1,L;S YX=YX+B(J2)*(X(J1) ↑ (J2-1))
*Ø8.15  S    D=Y(J1)=YX;S SD=SD+D*D
*Ø8.2Ø  T    %2,J1,"     ",%,Y(J1),"     ",YX,"     ",D,!;R
*
*14.Ø5  S    N=K+1; S DD=A(N+II*L)/A(II+II*L)
*14.1Ø  F    J=II,L; S A(N+J*L)=A(N+J*L)-A(II+J*L)*DD
*14.15  S    YX(N)=YX(N)-YX(II)*DD; R
*
*15.Ø5  S    MM=L-1
*15.1Ø  F    II=1,MM; F K=II,MM; D 14.Ø
*15.15  S    B(L)=YX(L)/A(L+L*L)
*15.2Ø  F    M=2,L;S N=L+1-M;S KK=N+1;S B(N)=YX(N)/A(N+N*L);D 15.25
*15.21  G    15.3Ø
*15.25  F    K=KK,L; S B(N)=B(N)-A(N+K*L)*B(K)/A(N+N*L)
*15.3Ø  R
**
```

*G
11 12 15
1 2 3 4 5 6 7 8 9 1Ø 11 12 13 14 15 16 17 18 19 2Ø 21 22 23 24
25 26 27 28 29 30

DATA POINTS :

	X	Y
= 1	= Ø.1ØØØØØE+Ø1	= Ø.2ØØØØØE+Ø1
= 2	= Ø.3ØØØØØE+Ø1	= Ø.4ØØØØØE+Ø1
= 3	= Ø.5ØØØØØE+Ø1	= Ø.6ØØØØØE+Ø1
= 4	= Ø.7ØØØØØE+Ø1	= Ø.8ØØØØØE+Ø1
= 5	= Ø.9ØØØØØE+Ø1	= Ø.1ØØØØØE+Ø2
= 6	= Ø.11ØØØØE+Ø2	= Ø.12ØØØØE+Ø2
= 7	= Ø.13ØØØØE+Ø2	= Ø.14ØØØØE+Ø2
= 8	= Ø.15ØØØØE+Ø2	= Ø.16ØØØØE+Ø2
= 9	= Ø.17ØØØØE+Ø2	= Ø.18ØØØØE+Ø2
= 10	= Ø.19ØØØØE+Ø2	= Ø.2ØØØØØE+Ø2
= 11	= Ø.21ØØØØE+Ø2	= Ø.22ØØØØE+Ø2
= 12	= Ø.23ØØØØE+Ø2	= Ø.24ØØØØE+Ø2
= 13	= Ø.25ØØØØE+Ø2	= Ø.26ØØØØE+Ø2
= 14	= Ø.27ØØØØE+Ø2	= Ø.28ØØØØE+Ø2
= 15	= Ø.29ØØØØE+Ø2	= Ø.3ØØØØØE+Ø2

NO. OF ADJUSTABLE COEFFICIENTS= 11

B(= 1)	= Ø.1ØØØ91E+Ø1
B(= 2)	= Ø.998Ø14E+ØØ
B(= 3)	= Ø.92Ø187E-Ø3
B(= 4)	= -Ø.185Ø82E-Ø3
B(= 5)	= Ø.191286E-Ø4
B(= 6)	= -Ø.871372E-Ø6
B(= 7)	= -Ø.896964E-Ø8
B(= 8)	= Ø.263Ø42E-Ø8
B(= 9)	= -Ø.1Ø2Ø17E-Ø9
B(=10)	= Ø.145919E-11
B(11)	= -Ø.483212E-14

NO.	Y OBS.	Y CALC.	DIF
= 1	= Ø.2ØØØØØE+Ø1	= Ø.199968E+Ø1	= Ø.321388E-Ø3
= 2	= Ø.4ØØØØØE+Ø1	= Ø.399957E+Ø1	= Ø.427246E-Ø3
= 3	= Ø.6ØØØØØE+Ø1	= Ø.6ØØØ11E+Ø1	=- Ø.1Ø7765E-Ø3
= 4	= Ø.8ØØØØØE+Ø1	= Ø.8ØØØ47E+Ø1	=- Ø.471115E-Ø3
= 5	= Ø.1ØØØØØE+Ø2	= Ø.1ØØØØ7E+Ø2	=- Ø.658Ø35E-Ø3
= 6	= Ø.12ØØØØE+Ø2	= Ø.12ØØØ6E+Ø2	=- Ø.593186E-Ø3
= 7	= Ø.14ØØØØE+Ø2	= Ø.14ØØØ1E+Ø2	=- Ø.114441E-Ø3
= 8	= Ø.16ØØØØE+Ø2	= Ø.159993E+Ø2	= Ø.7Ø3812E-Ø3
= 9	= Ø.18ØØØØE+Ø2	= Ø.179987E+Ø2	= Ø.131989E-Ø2
= 10	= Ø.2ØØØØØE+Ø2	= Ø.19999ØE+Ø2	= Ø.984192E-Ø3

= 11	= Ø.22ØØØØE+Ø2	= Ø.22ØØØ5E+Ø2	=- Ø.473Ø23E-Ø3
= 12	= Ø.24ØØØØE+Ø2	= Ø.24ØØ19E+Ø2	=- Ø.189972E-Ø2
= 13	= Ø.26ØØØØE+Ø2	= Ø.26ØØ1ØE+Ø2	=- Ø.988ØØ7E-Ø3
= 14	= Ø.28ØØØØE+Ø2	= Ø.279974E+Ø2	= Ø.261688E-Ø2
= 15	= Ø.3ØØØØØE+Ø2	= Ø.3ØØØØ8E+Ø2	=- Ø.839233E-Ø3

MEAN SQUARE DEVIATION= Ø.2Ø5455E-Ø2

ADDENDUM TO DECUS NO. FOCAL8-61

Since the FCCAL exponentiation routine evaluates zero to the power zero as zero instead of one, there will be an error where J2=1 and X(J1)=Ø. The following patch should correct the problem:

```
Ø2.23  F  J1=1,L1;D 16

16.1Ø  I  (X(J1))16.3,16.2,16.3
16.2Ø  S  SX(1)=1;S YX(1)=Y(J1)
16.3Ø  R
*
```