

Program Name: CURVEFIT

Language: Compiled BASIC

Microcomputer Use: IBM-PC and compatibles, PC-DOS 2.0, color monitor, memory capacity of 128K or more, and optional printer.

Purpose: CURVEFIT is used to fit equations to data.

Program Input/Output: CURVEFIT is self-contained. Data may be entered from the keyboard or from a sequential (ASCII) data file. Outputs include data storage in a diskette file; printed copies of data, equations, predicted values, and residuals; and graphs of data, predicted vs. observed values, and data compared to the regression equation.

Program Interfaces: CURVEFIT requires PC-DOS 2.0 or higher. Copies of CURVEFIT are supplied on 9-sector disks. Printing of CURVEFIT graphic plots requires DOS GRAPHICS.COM entry before calling BASICA.COM. Reading of disk files is facilitated with FASTDISK.COM. String variable arrays are passed from one program to the next using a 16K RAM-disk file CURVESTR.DAT established by CONFIG.SYS and VDISK.SYS. **Note: "Shift" + "Print Screen" to print graphs will not work on Windows machines in DOS.**

Program Description: CURVEFIT uses a Gaussian elimination scheme to calculate least-squares coefficients to linearized equation forms. Linearization is accomplished with single keystrokes which cause transformation of data. New variables may be created, for example, to combine primary variables into new forms, or to form polynomial equations from simple two-variable input data. Useless variables may be removed. Errors such as dividing by zero, or taking logarithms of negative numbers are detected and prevented. The user may remove offending data values by reconsideration of the data.

Transformation choices include:

1. X
2. AX
3. X-A
4.  $X^N$ , N an integer
5.  $X^A$ , A a non-integer
6. exp (X)
7. exp (-X)
8. log(X), natural logarithm
9.  $A^X$

All are available by single keystrokes.

Permutation (or variable creation) choices include:

1. X
2.  $X^N$ , N an integer
3.  $X^A$ , A a non-integer
4.  $X1/X2$
5.  $X1*X2$ , multiplication
6.  $X1-X2$
7.  $X1^{X2}$

Up to 9 new variables may be created in addition to the original variables. The program automatically tells the user how many additional variables can be created and gives the form for each current variable at important points in the program.

To assist in the linearization process, a dictionary of equation forms is available to use. Fifteen linearizable equations are presented in graphical form, and the data may be plotted to compare shapes. The data which is plotted is the currently transformed data, so that progress toward linearization can be seen.

Solution of equations can be accomplished with either a zero or non-zero intercept. There are instances where non-zero intercepts are theoretically impossible. The solution includes correlation coefficient for an indication of fit of equation to the data. A graphical plot of observed against predicted values also assists in determination of goodness-of-fit. A permanent record of equation, graph, and equation residuals can be made on the printer.

After the equation has been obtained, predicted values of the untransformed dependent variable can be obtained for any untransformed independent variable values. A plot of untransformed variables can be made, and original data included to assure that the transformation process did not produce an equation which does not fit the original data.

The use of color graphics requires a monitor which responds to color commands. Without a color monitor, most program operations can still be performed.

Program Components: CURVEFIT is supplied on a 5¼" diskette.

Required files are:

CURVEFIT.EXE	42862 bytes	8 Oct 1986
CURVEPLT.EXE	18046 bytes	8 Oct 1986
CURVESOL.EXE	7262 bytes	8 Oct 1986
CURVECHK.EXE	4311 bytes	8 Oct 1986
CURVEPRD.EXE	24394 bytes	8 Oct 1986
TEST.DAT	655 bytes	10 Oct 1984
CURVEFIT.ASC	39309 bytes	8 Oct 1986
CURVEPLT.ASC	15535 bytes	8 Oct 1986
CURVESOL.ASC	5478 bytes	8 Oct 1986
CURVECHK.ASC	4311 bytes	8 Oct 1986
CURVEPRD.ASC	24394 bytes	8 Oct 1986
VDISK.SYS	3080 bytes	14 Aug 1984
CONFIG.SYS	21 bytes	9 Oct 1986
AUTOEXEC.BAT	30 bytes	8 Oct 1986
BASRUN20.EXE	64150 bytes	4 Apr 1986
FASTDISK.COM	1024 bytes	4 Dec 1984

The first 5 files are compiled program files. BASIC source code for the program is found in the .ASC files. Other DOS files are included to establish a 16K resident RAM disk. Also included on the CURVEFIT diskette is a test data file, with 39 data points which generally fit the linearized equation:

$$\log(1/y-1) = a + bx$$

An AUTOEXEC.BAT file has been provided for automatic loading and running of the program.

Program Use: CURVEFIT is highly interactive and user-friendly. All questions are intended to be easily understood and the program structure is intended to allow looping and reconsideration should the user desire.

To begin CURVEFIT, just put the CURVEFIT disk in your computer and turn on the computer. The disk you are using must contain DOS, GRAPHICS, VDISK, CONFIG and AUTOEXEC files.

Data may be entered from the keyboard or from an ASCII sequential data file. The structure of each file record is  $X_1, X_2, \dots, X_n, Y$ . An ASCII file is used to allow file editing under DOS.

CURVEFIT may be used to iterate the best-fit solution to a non-linearizable equation. When the data is entered, it should be transformed and/or new variables created as far as possible. The new data can be saved in a data file called "INTERIM." A guess can be made for the value of one or more unknown parameters, and the data from "INTERIM" transformed into linearized form. The equation is then solved, and new guesses made. Data entry is made simpler by using the

"INTERIM" file. The optimum value of the unknown parameter will be that which gives the highest correlation coefficient value.

Comparison of the equation with zero intercept against that with a non-zero intercept is simple with CURVEFIT. Simply choose one or other of the choices the first time through. Then answer "Y" to the question which asks whether you wish to return to the variable manipulation point. You can then solve the best-fit equation with the remaining choice. If you had printed the first equation on the printer, no pencil-and-paper notation need to have been made. This is an illustration of the way in which program use can easily loop or return to previous points for reconsideration.

There is a non-stated limit of 300 data sets in each data file. This limit has been set to help detect errors occurring when data files are improperly named. This limit occurs in statement 2890 of CURVEFIT.ASC, and can be changed by the user.

Program Example: Follow program directions until you are asked:

HOW MANY INDEPENDENT VARIABLES (1-9)?

respond with "1"

IS DATA TO COME FROM A FILE (Y/N)?

respond with "Y"

WHAT NAME DOES THE DATA FILE HAVE?

respond with "TEST", "enter"

WHICH DISK DRIVE DO YOU WISH TO USE?

respond with the drive with the disk in, probably "A"

DO YOU WISH TO CHECK THE DATA (Y/N)?

you may respond either way

DO YOU WISH TO LIST THE DATA ON THE PRINTER (Y/N)?

you may respond either way. If you respond "Y" you'll obtain a listing similar to Listing 1.

DO YOU WISH TO SAVE THE DATA IN A FILE (Y/N)?

respond "N"

VARIABLE MANIPULATION:

respond "8"

respond "\". You will see a plot of the data similar to figure 1.

respond "/"

VARIABLE MANIPULATION:

respond "3"

WHICH OF THE TWO VARIABLES DO YOU WISH TO TRANSFORM?

respond "2"

TRANSFORMATION CHOICES:

respond "4"

ENTER THE VALUE FOR "N".

respond "-1", "enter"

then press any key

VARIABLE MANIPULATION:

respond "3"

WHICH OF THE 2 VARIABLES DO YOU WISH TO TRANSFORM?

respond "2"

TRANSFORMATION CHOICES:

respond "3"

ENTER THE VALUE FOR "A".

respond "1", "enter"

then press any key

VARIABLE MANIPULATION:

respond "3"

WHICH OF THE 2 VARIABLES DO YOU WISH TO TRANSFORM?

respond "2"

TRANSFORMATION CHOICES:

respond "8"

then press any key

VARIABLE MANIPULATION:

respond "8"

respond "\". You will see how the data has been linearized.

respond "/".

VARIABLE MANIPULATION:

respond "4"

DO YOU WANT A ZERO OR NON-ZERO INTERCEPT (Z/N)?

respond "N". You will see the equation and correlation coefficient

DO YOU WISH TO SEE A PLOT OF THE PREDICTED OBSERVATIONS (Y/N)?

respond "Y"

then press any key. You will see the plot in figure 3.

then press any key

DO YOU WISH TO LIST OBSERVED AND PREDICTED VALUES OF THIS VARIABLE ON THE SCREEN (Y/N)?

respond either way

DO YOU WISH TO MAKE A COPY OF THE EQUATION AND PREDICTED VALUES ON THE PRINTER (Y/N)?

If you respond "Y", you will obtain a listing similar to Listing 2.

DO YOU WISH TO PREDICT VALUES OF 'Y' FROM SPECIFIED VALUES OF X'S (Y/N)?

Respond "N"

DO YOU WISH TO SEE A PLOT OF THE EQUATION (Y/N)?

respond "Y"

MINIMUM VALUE:

respond "-20", "enter"

MAXIMUM VALUE:

respond "30", "enter"

MINIMUM VALUE:

respond "0". "enter"

MAXIMUM VALUE:

respond "1". "enter"

push key F5. You will see a plot of the equation.

push any key

CHOICES:

respond "D". You will see a plot of equation and data similar to figure 4.

press any key

CHOICES:

respond "T"

DO YOU WISH TO RECONSIDER THE SAME DATA OR RETURN TO VARIABLE MANIPULATION (Y/N)?

respond "Y"

VARIABLE MANIPULATION:

respond "7"

Program Listing: Because the CURVEFIT Program includes about 3000 statements, providing a program listing is impractical. The user can easily generate his own listing because the program is supplied in interpreted BASIC.

Reference: Johnson, A. T., 1984, Multidimensional Curve Fitting Program for Biological Data, Computer Programs in Biomedicine 18:259-264.

NORMALIZED  
DATA PLOT

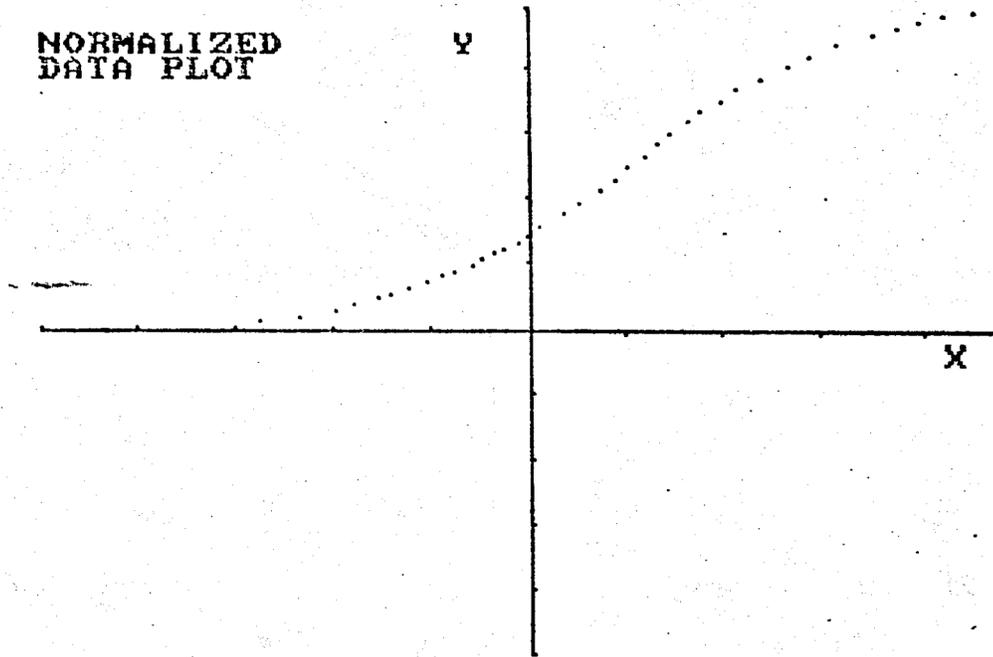


Figure 1

NORMALIZED  
DATA PLOT

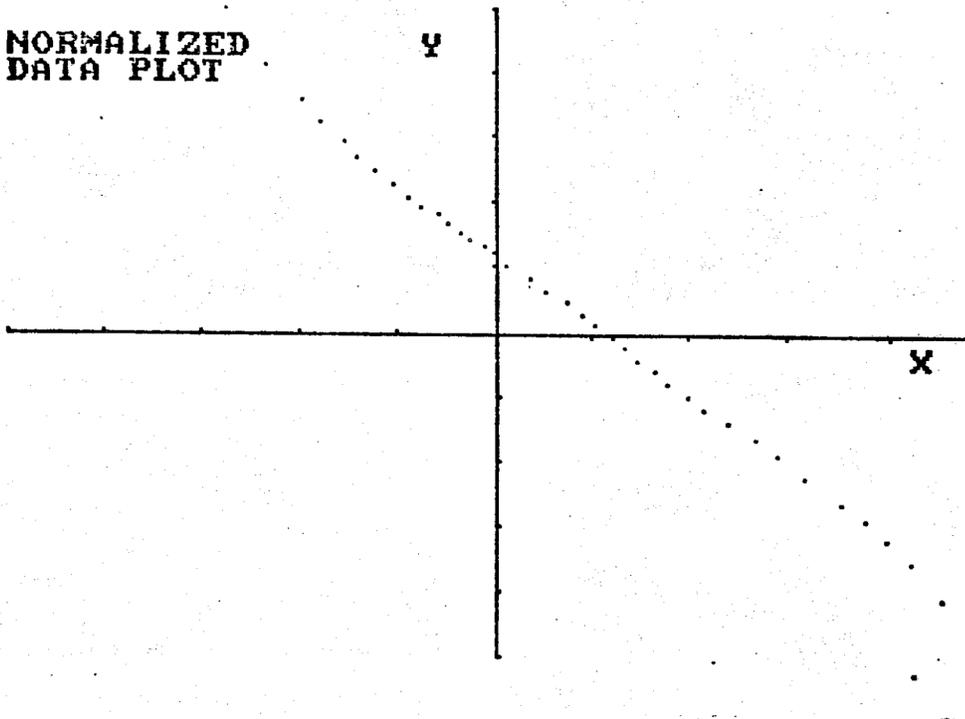


Figure 2

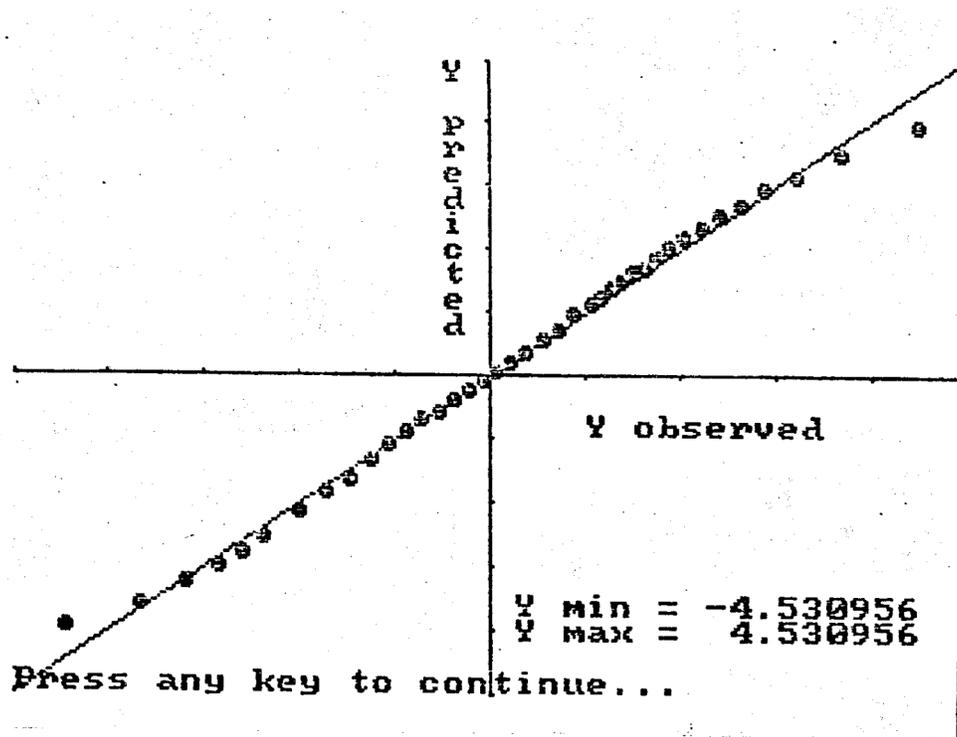


Figure 3

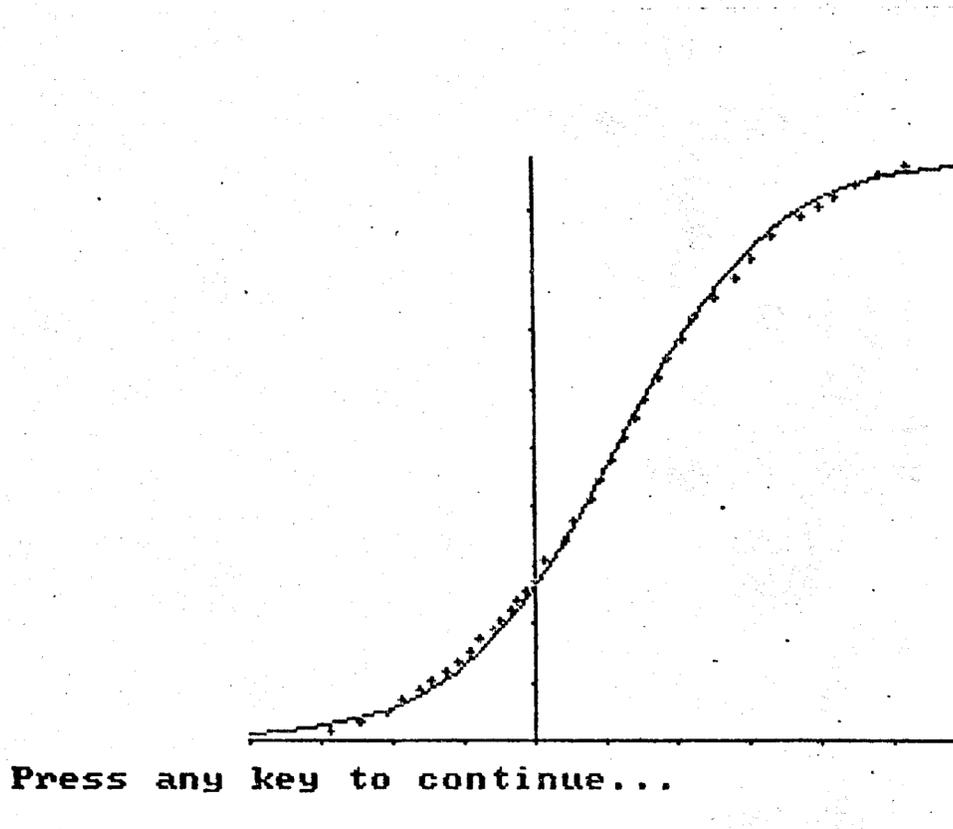


Figure 4

COMPLIANCE DATA FOR TEST  
DATA  
CURVEFIT PROGRAM  
03-26-1985

X(1)	Y
-14.4	.0172
-12.27	.0345
-10.33	.0517
-9.33	.069
-8.07	.0862
-7.33	.1034
-6.33	.1207
-5.4	.1379
-4.6	.1552
-4	.1724
-2.87	.1897
-2.53	.2069
-1.73	.2241
-1.27	.2414
-.53	.2586
0	.2759
.67	.3103
2	.3448
2.73	.3793
4	.4138
4.67	.4483
5.33	.4828
6.33	.5172
7.07	.5517
7.73	.5862
8.67	.6207
9.399999	.6552
10.4	.6897
11.33	.7241
12.67	.7586
14.07	.7931
15.2	.8276
16.67	.8621
18.67	.8966
20	.9138
21.07	.931
22.4	.9483
24.07	.9655
26	.9828

Listing 1

COMPLIANCE DATA FOR TEST

CURVE FITTING RESULTS

03-26-1985

$$\text{LOG}((1/(Y) - 1)) = 1.007533 - .1768786 * X(1)$$

WHERE:

X(1) = PRESSURE

Y = LUNG VOLUME/VITAL CAPACITY

CORRELATION (R) = .997283

OBSERVED VALUES	PREDICTED VALUES	RESIDUALS	RESIDUALS SQUARED	RESIDUAL PER CENT
4.045497	3.554585	-.490912	.2409946	-12.13478
3.331687	3.177834	-.1538532	.0236708	-4.617877
2.909213	2.834689	-.0745244	5.553887E-03	-2.561669
2.602153	2.65781	5.565739E-02	3.097745E-03	2.138898
2.360942	2.434943	7.400155E-02	5.476229E-03	3.134408
2.160005	2.304053	.1440482	2.074989E-02	6.668883
1.985818	2.127174	.1413565	1.998165E-02	7.1183
1.832843	1.962677	.1298349	.0168571	7.083799
1.694385	1.821175	.1267893	1.607553E-02	7.48291
1.568713	1.715047	.1463349	.0214139	9.328343
1.451961	1.515175	6.321406E-02	3.996018E-03	4.353704
1.343714	1.455036	.1113222	1.239262E-02	8.284663
1.241931	1.313533	7.160175E-02	5.12681E-03	5.765355
1.145019	1.232169	8.714962E-02	7.595056E-03	7.611192
1.053258	1.101279	4.802096E-02	2.306012E-03	4.559279
.964891	1.007533	4.264224E-02	1.81836E-03	4.419384
.7987171	.8890246	9.030742E-02	8.155429E-03	11.30656
.641976	.6537761	1.180005E-02	1.392412E-04	1.838083
.4925204	.5246547	.0321343	1.032613E-03	6.52446
.3482783	.3000189	-4.825941E-02	2.328971E-03	-13.85656
.2075417	.1815103	-2.603139E-02	6.776333E-04	-12.54273
6.882705E-02	.0647704	-4.056647E-03	1.645639E-05	-5.893973
-6.882716E-02	-.1121081	-4.328095E-02	1.873241E-03	62.88354
-.2075418	-.2429984	-3.545657E-02	1.257168E-03	17.08406
-.3482782	-.3597381	-1.145989E-02	1.31329E-04	3.29044
-.4925205	-.5260041	-.0334836	1.121151E-03	6.798416
-.6419759	-.6551251	-.0131492	1.729015E-04	2.048239
-.7987171	-.8320038	-3.328669E-02	1.108004E-03	4.167519
-.9648911	-.9965011	-3.161001E-02	9.991928E-04	3.276019
-1.145019	-1.233518	-8.849883E-02	7.832043E-03	7.729024
-1.343714	-1.481148	-.1374348	1.888834E-02	10.22799
-1.568712	-1.681021	-.1123086	1.261323E-02	7.159287
-1.832842	-1.941033	-.1081901	1.170509E-02	5.902857
-2.160005	-2.29479	-.1347845	1.816685E-02	6.240006
-2.360941	-2.530038	-.1690969	2.859378E-02	7.162268
-2.602153	-2.719298	-.1171455	1.372308E-02	4.50187
-2.909214	-2.954547	-4.533315E-02	2.055094E-03	1.558261
-3.331688	-3.249934	8.175349E-02	6.683634E-03	-2.453816
-4.045497	-3.59131	.4541879	.2062866	-11.227

Listing 2