PROGRAM ABSTRACTS/ALGORITHMS

Curve fitting by estimation and iteration

LARRY HOCHHAUS

Oklahoma State University, Stillwater, Oklahoma 74078
and University of Oklahoma Health Science Center
Oklahoma City, Oklahoma 73190

Finding the best-fitting equation for a set of data points is a common operation in evaluating scientific experiments. Straight lines are frequently established by estimating the position of the best-fitting line by eye. Using the method of linear regression, an exact function can be obtained that minimizes the vertical (dependent variable) squared deviations between the actual data points and their corresponding theoretical estimates. Often, these two techniques for fitting straight lines show close agreement.

Curvilinear functions present a more difficult problem, but at times these may be converted to the linear case by choice of a suitable transformation on the dependent variable. This may not always be an easy operation. The widespread availability of small computer systems that include options for graphics presentation provides an added method for obtaining solutions to problems involving complex curves.

The two BASIC programs developed to meet these needs were created using an Apple microprocessor with 48K RAM, but the programs can be adapted to work on other systems. The GRAPH program uses video graphics to display raw scores and the curve produced by an equation assumed to represent the data. Parameter values in the underlying equation may be changed to allow selection of the best-fitting curve by visual inspection. The ITERATION program may be used to find the equation that minimizes the sum of squared deviations between raw scores and estimates generated by the equation.

A speed-accuracy tradeoff curve is used to illustrate the programs. The raw data in this example are six pairs of accuracy (d') and response speed (RT) scores for a single subject. While other mathematical models of the function relating speed and accuracy are available (e.g., McClelland, 1979), Wickelgren's (1977) equation was selected for use as a first approximation. This equation requires the estimation of three parameters, X-intercept (t), rate (r), and asymptote (a). The equation is $d' = a [1 - e^{-r(RT - t)}]$ for values of RT greater than t ; $d' = 0$ otherwise. The computation of the equation is

```
10 REM ESTIMATION PROGRAM
20 DIM D(6), RT(6)
30 REM D' & RT VALUES
40 FOR I = 1 TO 6: READ D(I): NEXT I
50 DATA .3,.08,.1,.35,.72,1.76
60 FOR I = 1 TO 6: READ RT(I): NEXT I
70 DATA 169,209,258,304,410,1229
80 ITC = .177:ASY = 2.98:RTE = 3.4:E = 2.7183
90 HOME:VTAB 10
100 PRINT "DO YOU WANT TO SEE THE AXES?"
110 POKE -16368,0: GET AS
120 HGR: HCOLOR=3
130 IF AS <> "Y" THEN 310
140 HPLLOT 3,3: HPLLOT TO 3,157
150 HPLLOT TO 257,157: HPLLOT TO 257,3
160 HPLLOT TO 3,3
170 FOR X = 3 TO 257:T = X - 3
180 IF T / 10 <> INT (T / 10) THEN 240
190 HPLLOT X,158
200 IF T / 20 <> INT (T / 20) THEN 240
210 HPLLOT X,159: HPLLOT X,156: HPLLOT X,155
220 IF T / 200 <> INT (T / 200) THEN 240
230 HPLLOT X,159: HPLLOT TO X,145
240 NEXT X
250 FOR Y = 3 TO 157:T = Y - 7
260 IF T / 15 <> INT (T / 15) THEN 300
270 HPLLOT 2,Y
280 IF T / 30 <> INT (T / 30) THEN 300
290 HPLLOT 1,Y: HPLLOT TO 5,Y
300 NEXT Y
310 FOR K = 1 TO 6
320 X = INT (RT(K) / 5) + 3
330 Y = 157 - D(K) * 30
340 HPLLOT X,Y: NEXT K
350 HOME:VTAB 21
360 PRINT "STUDY TIME-----"
370 GOTO 500
380 REM PLOTTING SUBROUTINE
390 FOR T = 0 TO 1200 STEP 20
400 IF T <= ITC * 1000 THEN HPLLOT (INT (T / 5) + 3), 156
410 IF T <= ITC * 1000 THEN 480
420 M = T / 1000
430 POW = -RTE * (M - ITC)
440 D = ASY * (1 - E + POW)
450 X = T / 5 + 3
460 Y = 157 - D * 30
470 HPLLOT X,Y
480 NEXT T
490 RETURN
500 REM
510 GOSUB 380
520 PRINT "ASYMPTOTE=";ASY;" INTERCEPT=";ITC;" RATE=";RTE
530 PRINT "CHANGE A PARAMETER? (Y/N)"
540 POKE -16368,0: GET AS
550 IF AS = "N" THEN 670
560 HCOLOR = 2
570 GOSUB 380
580 HCOLOR = 1
590 PRINT "WHICH? (A, I, OR R)"
600 POKE -16368,0: GET AS
610 PRINT "ASYMPTOTE=";ASY;" INTERCEPT=";ITC;" RATE=";RTE
620 IF AS = "A" THEN INPUT "ENTER NEW ASYMPTOTE:";ASY
630 IF AS = "I" THEN INPUT "ENTER NEW INTERCEPT:";ITC
640 IF AS = "R" THEN INPUT "ENTER NEW RATE:";RTE
650 GOSUB 380
660 GOTO 520
670 HOME:VTAB 21
680 PRINT "ASYMPTOTE=";ASY;" INTERCEPT=";ITC;" RATE=";RTE
690 PRINT "PRESS <SPACE BAR> TO CLEAR GRAPHICS."
700 POKE -16368,0: GET AS: TEXT: HOME:VTAB 3
710 PRINT "A=";ASY;" I=";ITC;" R=";RTE
```

Figure 1. Program listing of the GRAPH routine.

embodied in Lines 420-440 of the GRAPH program listed in Figure 1.

Once the three parameters are approximated using the GRAPH program, these values are recorded and used as the starting point for the ITERATION program listed in Figure 2. The ITERATION program then adds six parameter values by successively adding and subtracting a small amount to each of the original three parameters. The resulting 27 equations are then evaluated to determine the equation that provides the minimum sum of

```

10  REM ITERATION PROGRAM
20  DIM D (6), RT(6), DEST(6)
30  REM VECTORS FOR VALUES OF D', RT, AND D' ESTIMATES
40  E = 2.7183: SM = 100
50  FOR I = 1 TO 6: READ D(I): NEXT I
60  DATA .3,.08,.1,.35,.72,1.76
70  FOR I = 1 TO 6: READ RT (I): NEXT I
80  DATA 169,209,258,304,410,1229
90  INPUT "ASYMPTOTE ESTIMATE?";P2
100 INPUT "INTERCEPT ESTIMATE?";P1
110 INPUT "RATE ESTIMATE?";P3
120 REM TRY P1=.24: P2=1.85: P3=3.8
130 INPUT "ASYMPTOTE INCREMENT? (.01 OR .1)";AI
140 INPUT "INTERCEPT INCREMENT? (.001 OR .01)";II
150 INPUT "RATE INCREMENT? (.1 OR 1)";RI
160 AI = AI * 100:II = II * 1000:RI = RI * 10
170 HOME: VTAB 12: PRINT,"COMPUTING"
180 REM ENTRY POINT FOR RE-CYCLE
190 FOR A = (P2 * 100 - AI) TO (P2 * 100 + AI) STEP AI
200 FOR I = (P1 * 1000 - II) TO (P1 * 1000 + II) STEP II
210 FOR R = (P3 * 10 - RI) TO (P3 * 10 + RI) STEP RI
220 FOR J = 1 TO 6
230 IF RT(J) < P1 * 1000 THEN DEST (J) = 0
240 IF RT(J) < P1 * 1000 THEN 270
250 POW = -R / 10 * (RT(J) / 1000 - I / 1000)
260 DEST (J) = A / 100 * (1 - E * POW)
270 SS = SS + (D(J) - DEST (J)) + 2
280 NEXT J
290 IF SS > SM THEN 340
300 IF SS < SM THEN AM = A
310 IF SS < SM THEN IM = I
320 IF SS < SM THEN RM = R
330 IF SS < SM THEN SM = SS
340 SS = 0
350 NEXT R,I,A
360 IF AM = 01 AND IM = 02 AND RM = 03 THEN 410
370 O1 = AM:O2 = IM:O3 = RM
380 PRINT "A=";AM / 100: SPC( 5);"I=";IM / 1000: SPC( 5); "R=";RM / 10
390 P1 = IM / 1000: P2 = AM / 100:P3 = RM / 10
400 GOTO 180
410 PRINT : PRINT: PRINT "PROGRAM HAS CONVERGED."
420 AM = AM / 100: IM = IM / 1000: RM = RM / 10
430 PRINT "MINIMUMS: ASY=";AM;" INT=";IM;" RATE=";RM

```

Figure 2. Program listing of the ITERATION routine.

squared deviations between raw scores and equation estimates. The minimizing parameters are printed on the screen. The program inserts these values into the beginning of the routine. The procedures continue to improve the parameter estimates until the same values are found for each parameter on two successive cycles through the program.

REFERENCES

- McCLELLAND, J. L. On the time relations of mental processes: An examination of system of processes in cascade. *Psychological Review*, 1979, 86, 287-330.
- WICKELGREN, W. A. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 1977, 41, 67-85.

(Accepted for publication August 6, 1981.)