INTERNAL REPORT 99

RICHER: A DATA ENRICHMENT SUBROUTINE

E. R. Hamerly
University of Washington

ABSTRACT

RICHER is a sophisticated interpolation routine which allows the user to generate data for points he does not have. The following activities are at best marginally legitimate.

A. Regenerate a dataset at a different step size (monthly points instead of bimonthly)
B. Estimate missing data due to equipment failure
C. Estimate state vectors for times that are not increments of the simulation step.

INTRODUCTION

RICHER was written upon request for George Hendry and Dave Culver to allow better estimation of heat budgets in the four lakes study. In exchange for RICHER, the information bank received two additional useful tools:

(1) GRID -- a subroutine which squares data to a rectangular grid by estimating the grid value from the 6 closest points

(2) ISØPLET -- (author George Hendry) a program to draw temperature isopleths over the annual cycle

DIGDAT[1], GRID, and ISØPLET form an analysis chain for producing contour maps of lake bottoms from strip charts of transect runs allowing better estimation of the amount of water at different depths in the lakes. RICHER and ISØPLET furnish an analysis for producing more accurate isopleths of the annual temperature regimes at different depths.

THE ALGORITHM

RICHER selects three data points at a time $(\alpha_{i-2}, \beta_{i-2})$, $(\alpha_{i-1}, \beta_{i-1})$, (α_i, β_i), where $i=3, \ldots, n$ and $n=$ the number of data points. The three data points being considered will be referred to as (x_1, y_1), (x_2, y_2), and (x_3, y_3). The point $(\alpha_{i-3}, \beta_{i-3})$ is referred to as (x_0, y_0).

---

[1]See Internal Report 61.

RICHER checks that y_2 is strictly between y_1 and y_3. If it is not, then (x_3, y_3) is reset so that a parabola passing through the three points will have its extremum at (x_2, y_2).

$$(x_3, y_3) = (x_1 + 2(x_2 - x_1), y_1) \qquad (1)$$

This does not alter (α_i, β_i) because (x_3, y_3) are stored in separate working areas. RICHER passes the parabola $y = c_0 + c_1 x + c_2 x^2$ through the three points. c_0, c_1, and c_2 are calculated from the simultaneous equations

$$\begin{aligned} c_0 + x_1 c_1 + x_1^2 c &= y_1 \\ c_0 + x_2 c_1 + x_2^2 c_2 &= y_2 \\ c_0 + x_3 c_1 + x_3^2 c_3 &= y_3 \end{aligned} \qquad (2)$$

The parabola passes through (x_1, y_1), (x_2, y_2) and (x_3, y_3).

Let X be the matrix of coefficients for C

$$X = \begin{bmatrix} 1, & x_1, & x_1^2 \\ 1, & x_2, & x_2^2 \\ 1, & x_3, & x_3^2 \end{bmatrix}, \quad C = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} \text{ and}$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

Equation (2) is represented by XC = Y. C is found by $C = X^{-1}Y$.

A second parabola $y = b_0 + b_1 x + b_2 x^2$ is fitted through (x_1, y_1) and (x_2, y_2) with the constraint that y_1' has the same value as $y_2' = a_1 + 2a_2 x_2$ where a_1 and a_2 are the values found for c_1 and c_2 on points (x_0, y_0), (x_1, y_1), and (x_2, y_2). (y_1' is the slope of the line at x_1).

b_0, b_1, and b_2 are calculated from the equations

$$\begin{aligned} b_0 + x_1 b_1 + x_1^2 b_2 &= y_1 \\ b_0 + x_2 b_1 + x_2^2 b_2 &= y_3 \\ b_1 + 2x_1 b_2 &= a_1 + 2x_1 a_2 \end{aligned} \qquad (3)$$

Here

$$X = \begin{bmatrix} 1, & x_1, & x_1^2 \\ 1, & x_2, & x_2^2 \\ 0, & 1, & 2x_1 \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ a_1 + 2x_1 a_2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$$

Equation set (3) becomes XB = Y and B is found by $B = X^{-1}Y$.

The weighted average of the two parabolas is formed

$$y = [(x_2 - x)\ (b_0 + b_1 x + b_2 x^2) + (x - x_1)\ (c_0 + c_1 x + c_2 x^2)]\ /\ (x_2 - x_1) \qquad (4)$$

Close inspection will show equation (4) is not a parabola, but a third degree polynomial

$$y = d_0 + d_1 x + d_2 x^2 + d_3 x^3 \qquad (5)$$

where $d_0 = (x_2b_0 - x_1c_0) / (x_2 - x_1)$, $d_1 = (x_2b_1 - b_0 + c_0 - x_1c_1) / (x_2 - x_1)$, $d_2 = (x_2b_2 - b_1 + c_1 - x_1c_2) / (x_2 - x_1)$, $d_3 = (c_2 - b_2) / (x_2 - x_1)$.

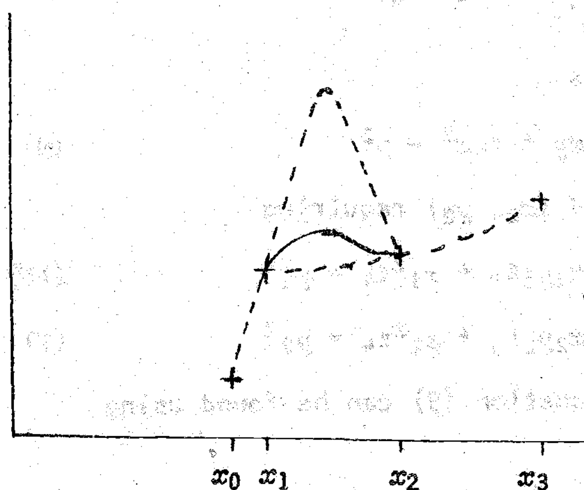
Equation (4) is not valid to use between (x_1, y_1) and (x_2, y_2) if a relative extremum occurs between x_1 and x_2. This happens if

$$d_1 + 2d_2x + 3d_3x^2 = 0 \qquad (6)$$

Equation (6) has at most two solutions

$$x = \{ [-e_1 - (e_1{}^2 - 4e_2e_0)^{\frac{1}{2}}] / 2e_2, \; [-e_1 + (e_1{}^2 - 4e_2e_0)^{\frac{1}{2}}] / 2e_2 \} \qquad (7)$$

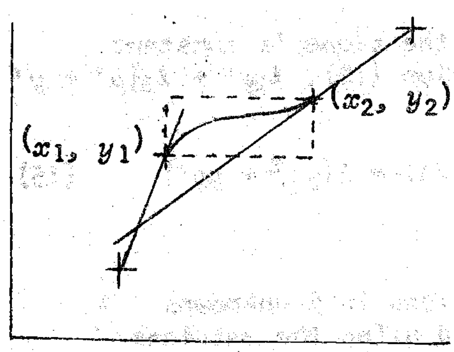
where $e_0 = d_1$, $e_1 = 2d_2$, and $e_2 = 3d_3$. Further refinement of the curve is needed if either value of x in equation (7) is strictly between x_1 and x_2. Equation (5) has inflection points when $2d_2 + 6d_3x = 0$ or $x = -d_2/2d_3$.



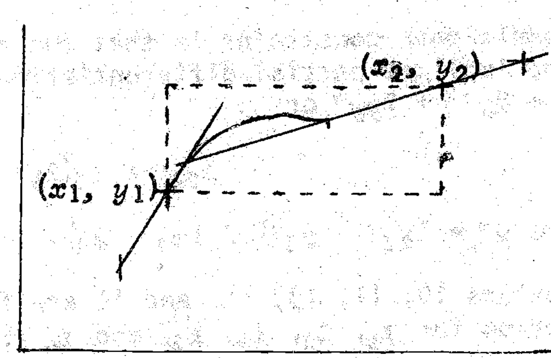
The dashed lines are the two parabolas.

The solid line is the weighted average which in this case has an undesired extremum between x_1 and x_2.

An inflection between x_1 and x_2 is only desired if the line tangent to equation (5) at (x_1, y_1) does not intersect the line tangent to equation (5) at (x_2, y_2) within the rectangle (x_1, y_1), (x_2, y_1), (x_2, y_2), (x_1, y_2).



The intersection is outside the rectangle. The inflection is desired.



The intersection is inside the rectangle. The inflection is not desired.

The line tangent to equation (5) at the point (ε, γ) will have the equation

$$y'_\varepsilon = (y - \gamma) / (x - \varepsilon) \tag{8}$$

where y'_ε is the slope of equation (5) at (ε, γ). Equation (8) can be written $g_1 x + y = g_0$, where $g_0 = \gamma + y'_\varepsilon$ and $g_1 = y'_\varepsilon$. Two equations are formed where (x_I, y_I) is the intersection. $h_1 x_I + y_I = h_0$ and $k_1 x_I + y_I = k_0$, or $X = \begin{bmatrix} h_1, & 1 \\ k_1, & 1 \end{bmatrix}$, $H = \begin{bmatrix} h_0 \\ k_0 \end{bmatrix}$, $P = \begin{bmatrix} x_I \\ y_I \end{bmatrix}$ or $XP = H$, $P = X^{-1}H$. Equation (5) is adequate if X_I is not strictly between x_1 and x_2 or if y_i is not strictly between y_1 and y_2. Equation (5) furnishes an adequate solution when y_2 is not strictly between y_1 and y_3. More refinement may be needed when y_1, y_2, and y_3 are strictly ordered. Equation (5) is such that $y'x_1 = y_2 - y_1 / x_2 - x_1$ and $y'x_2 = y_3 - y_2 / x_2 - x_1$. y_1, y_2, and y_3 being strictly ordered require that $y'x_1$ and $y'x_2$ be of the same sign. An elliptical segment can be passed through (x_1, y_1) and (x_2, y_2) such that the slope at x_1 is y_1' and the slope at x_2 is y_2'.

The general equation for the ellipse is

$$\ell_0 + \ell_1 x + \ell_2 y + \ell_3 xy + \ell_4 x^2 = y^2 \tag{9}$$

The ellipse passes through (x_1, y_1) and (x_2, y_2) requiring

$$\ell_0 + x_1\ell_1 + y_1\ell_2 + x_1y_1\ell_3 + x_1^2\ell_4 = y_1^2 \tag{10}$$

$$\ell_0 + x_2\ell_1 + y_2\ell_2 + x_2y_2\ell_3 + x_2^2\ell_4 = y_2^2 \tag{11}$$

A relation involving the slope y' of equation (9) can be found using implicit partial differentiation.

$$\ell_1 + \ell_2 y' + \ell_3(y + xy') + 2\ell_4 x = 2yy' \tag{12}$$

The two additional equations are

$$\ell_1 + y_1'\ell_2 + (y_1 + x_1y_1')\ell_3 + 2x_1\ell_4 = 2y_1y_1' \tag{13}$$

$$\ell_1 + y_2'\ell_2 + (y_2 + x_2y_2')\ell_3 + 2x_2\ell_4 = 2y_2y_2' \tag{14}$$

An additional constraint is that the change in the slope is constant. Using implicit partial differentiation on equation (12), $\ell_2 y'' + \ell_3(y' + y' + xy'') + 2\ell_4 = 2y'^2 + 2yy''$ or

$$\ell_2 y'' + \ell_3(2y' + xy'') + 2\ell_4 = 2(y'^2 + yy'') \tag{15}$$

where $y'' = (y_2' - y_1') / (x_2 - x_1)$.

Equations 10, 11, 13, 14, and 15 are five equations in 5 unknowns. A solution for ℓ_0, ℓ_1, ℓ_2, ℓ_3, and ℓ_4 can be found using the matrices

$$X = \begin{bmatrix} 1, & x_1, & y_1, & x_1y_1 & , x_1^2 \\ 1, & x_2, & y_2, & x_2y_2 & , x_2^2 \\ 0, & 1, & y_1, & (y_1 + x_1y_1') & , 2x_1 \\ 0, & 1, & y_2, & (y_2 + x_2y_2') & , 2x_2 \\ 0, & 0, & y'', & (2y_1' + x_1y'') & , 2 \end{bmatrix}, \quad Y = \begin{bmatrix} y_1^2 \\ y_2^2 \\ 2y_1y_1' \\ 2y_2y_2' \\ 2\,(y_2'y_2' + y_2y_2'') \end{bmatrix}$$

and $L = \begin{bmatrix} \ell_0 \\ \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix}$. The relation is XL = Y.

ℓ_0, ℓ_1, ℓ_2, ℓ_3, and ℓ_4 are found by $L = X^{-1}Y$. The ellipse has been fitted. Now a solution for y in terms of x must be found.

Starting from equation (9)

$$\begin{aligned} \ell_0 + \ell_1x + \ell_2y + \ell_3xy + \ell_4x^2 &= y^2 \\ \ell_0 + \ell_1x + \ell_4x^2 &= y^2 - \ell_2y - \ell_3xy \\ \ell_0 + \ell_1x + \ell_4x^2 &= y^2 - y(\ell_2 + \ell_3x) \end{aligned} \qquad (16)$$

Add $(\ell_2 + \ell_3x)^2/4$ to both sides completing the square. $y - \frac{1}{2}(\ell_2 + \ell_3x) = \pm[\ell_0 + \ell_1x + \ell_4x^2 + (\ell_2 + \ell_3x)^2/4]^{\frac{1}{2}}$, $y = \frac{1}{2}(\ell_2 + \ell_3x) \pm [\ell_0 + \ell_1x + \ell_4x^2 + (\ell_2 + \ell_3x)^2/4]^{\frac{1}{2}}$

The choice of the function depends upon the second derivative of the ellipse. $y'' = (y_2' - y_1') / (x_2 - x_1)$. The larger value of y must be taken if the ellipse is curving downward and the smaller value of y must be taken if the ellipse is curving upward.

USAGE

The call to subroutine RICHER is

```
CALLRICHER(XI,YI,NDI,XØ,YØ,NDØ,XMINØ,XSTEP)
```

where (XI(I), YI(I)), I = 1, ..., NDI are the NDI ordered pairs to be enriched.

(XØ(J), YØ(J)), J = 1, ..., NDØ are the NDØ ordered pairs generated by RICHER such that XØ(J) = (J - 1) * XSTEP + XMINØ, J = 1, ..., NDØ.

The deck structure using RICHER in a FØRTRAN program:

```
JØBCARD,CM40000,T60,PO.
REQUEST,RICHER,VRN=P894,DI,IN,FILES=41.
FØRTRAN
LØAD(RICHER)
LGO
(End of record card)
  FØRTRAN program
(End of record card)
  Data cards
(End of job card)
```

RICHER was used to generate 1000 points from the 14 points shown on the following graph. The solid line was formed by connecting the generated points by straight line segments.