Check for updates

Numerical Mathematics
R.A. Willoughby Editor

Least Squares Piecewise Cubic Curve Fitting

J. Ferguson and P.A. Staley
Teledyne Ryan Aeronautical Company

The matrices involved in a linear least squares formulation are determined for the problem of fitting piecewise cubic functions, those possessing a continuous derivative, to arrays of planar data.

Key Words and Phrases: curve fitting, data reduction, function approximation, approximation splines

CR Categories: 5.13

Introduction

A class of functions—smooth (C^1) piecewise cubics—is used to find a fit to a finite set of approximate, planar data, using a least squares technique. (Recall that a C^1 function is one having a continuous first derivative.) The desirability of using such functions is linked to the advantages of using spline functions (piecewise cubic with a continuous second derivative), in that they are computationally simple and offer high degrees of flexibility in applications to rather general curve fitting situations. Here, we are interested only in the C^1 case since, for example, it provides a simple solution to producing visually smooth plotted curves needed in many applications. Additionally, we have a larger class of fitting functions than in the case requiring second derivative continuity, and we reach an interesting closed form solution with relative ease.

One quickly will recognize that a stringent constraint is subsequently imposed: the nodes of the fitting functions are fixed during the least squares process. A more satisfying approach would rather be to consider it otherwise; however, the problem becomes severely nonlinear, thereby presenting manifold complexities. So, we treat the results of this report as support computation for some forthcoming nonlinear, iterative, node-adjusting scheme.

Copyright © 1973, Association for Computing Machinery, Inc. General permission to republish, but not for profit, all or part of this material is granted provided that ACM's copyright notice is given and that reference is made to the publication, to its date of issue, and to the fact that reprinting privileges were granted by permission of the Association for Computing Machinery.

Authors' address: Teledyne Ryan Aeronautical Company, 2701 Harbor Drive, San Diego, CA 92112.

Since the time when this problem was initially considered [1], some related work has been accomplished using splines, and it would be profitable to consult some of the references [2–8, 11]. Apparently the treatment given here has not been developed elsewhere, even though it seems more natural to consider the C^1 case first.

Definitions and Problem Statement

We work in a planar rectangular coordinate system, denoting as usual, points as ordered pairs of numbers, such as (a, b). Derivatives may be written as $f'(x) \equiv df(x)/dx$.

Definition. Let $x_1 < x_2 < \cdots < x_{M+1}$ be $M + 1$ given numbers, called *joints* (or nodes or knots). A C^1 *piecewise cubic function* (PCF) is a C^1 smooth real function, $f(x)$, having domain $x_1 \leq x \leq x_{M+1}$ such that $f_j(x) \equiv f(x)$, $x_j \leq x \leq x_{j+1}$, is at most a cubic polynomial. (Note that this function $f(x)$ is completely determined if the values of $f(x_j), f'(x_j), j = 1, \ldots, M + 1$ are known.)

Let (t_i, y_i), $i = 1, \ldots, N$, be N points satisfying: (a) $t_1 < t_2 < \cdots < t_N$, and (b) $x_1 \leq t_1$, $x_{M+1} \geq t_N$.

Problem. Determine that C^1, PCF (assuming it is unique), $f(x)$, with joints at the x_j, which best fits (in a least squares sense) the points (t_i, y_i).

The assumption of uniqueness can be discussed, at least heuristically. The task is to compute $2M + 2$ numbers, i.e. the slopes $s_j \equiv f'(x_j)$ and the ordinates $z_j \equiv f(x_j)$, $j = 1, \ldots, M + 1$, thereby fixing $f(x)$. So, there should be at at least $2M + 2$ data points, i.e. $N \geq 2M + 2$. However, this criterion alone is not enough to ensure uniqueness since all N of the points may be contained in some interval $x_j \leq x \leq x_{j+1}$. Here, finding the best $f(x)$ is equivalent to finding the best single cubic $f_j(x)$, leaving the determination of the remaining cubics almost arbitrary. Thus, other restrictions on the distribution of the N data points should be considered. One possible thought is to determine the smallest value for k such that forcing at least k-points over each interval $x_j \leq x \leq x_{j+1}$, augmented by a fixed number of otherwise randomly placed points (altogether satisfying (a), (b)), always gives a unique $f(x)$. We know $k = 4$ will work and $k = 0$ or 1 won't work. Intuitively and in practice it has been found that $k = 2$ is large enough. This alone accounts for $2M$ data points, so that we must augment these by at least two more points to satisfy $N \geq 2M + 2$. We assume in what follows that this does give a unique $f(x)$. (An interesting theorem would involve a proof of this observation.)

Problem Solution. Denoting $(x_{j+1} - x_j)$ by δx_j, we find that the conditions

$$f_j(x_j) = z_j, \quad f_j'(x_j) = s_j,$$
$$f_j(x_{j+1}) = z_{j+1}, \quad f_j'(x_{j+1}) = s_{j+1}, \tag{1}$$

give, for each $j = 1, \ldots, M$, the following equation for the jth cubic defined on $x_j \le x \le x_{j+1}$

$$\begin{aligned} f_j(x) = {} & [2u_j(x) + 1]v_j^2(x)z_j \\ & + u_j^2(x)[1 - 2v_j(x)]z_{j+1} \\ & + (\delta x_j)u_j(x)v_j^2(x)s_j \\ & + (\delta x_j)u_j^2(x)v_j(x)s_{j+1}, \end{aligned} \tag{2}$$

where $u_j(x) \equiv (x - x_j)/\delta x_j$ and $v_j(x) \equiv u_j(x) - 1$.

Now, for each $j = 1, \ldots, M$ assume there exists an L such that $x_j < t_L < t_{L+1} < x_{j+1}$ and define n_j equals the largest i satisfying $t_i < x_{j+1}$ ($\le$ when $j = M$). The problem is to minimize the following multivariable function

$$E(s_p, z_p; p = 1, \ldots, M+1) = \sum_{j=1}^{M} \sum_{i=m_j}^{n_j} [f_j(t_i) - y_i]^2 \tag{3}$$

with respect to the unknowns $s_1, \ldots, s_{M+1}, z_1, \ldots, z_{M+1}$. Note that we have written m_j instead of $n_{j-1} + 1$ ($n_0 \equiv 0$). The following further abbreviations and equations are used in the algebra for minimizing (3).

$$\begin{aligned} &(a)\ u_{ij} \equiv u_j(t_i) = (t_i - x_j)/\delta x_j, \\ &(b)\ v_{ij} \equiv v_j(t_i) = u_{ij} - 1, \\ &(c)\ \partial f_j(t_i)/\partial s_j = (\delta x_j)u_{ij}v_{ij}^2, \\ &(d)\ \partial f_j(t_i)/\partial s_{j+1} = (\delta x_j)u_{ij}^2 v_{ij}, \\ &(e)\ \partial f_j(t_i)/\partial z_j = (2u_{ij} + 1)v_{ij}^2, \\ &(f)\ \partial f_j(t_i)/\partial z_{j+1} = u_{ij}^2(1 - 2v_{ij}). \end{aligned} \tag{4}$$

Now, the $2M + 2$ equations

$$\begin{aligned} &\partial E/\partial s_p = 0 \\ &\partial E/\partial z_p = 0, \quad p = 1, \ldots, M+1, \end{aligned} \tag{5}$$

which must be satisfied by those values $s_1, \ldots, s_{M+1}, z_1 \ldots, z_{M+1}$, minimizing (3), can be written as a set of linear equations in the unknowns. The result obtained by expanding (5), using (2), (3), and (4), is summarized in the following matrix equation

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^T & \mathcal{C} \end{bmatrix} \begin{bmatrix} \mathcal{S} \\ \mathcal{Z} \end{bmatrix} = \begin{bmatrix} \mathcal{K} \\ \mathcal{L} \end{bmatrix} \tag{6}$$

where $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are tridiagonal $M + 1$ square matrices ($\mathcal{B}^T$ is the transpose of $\mathcal{B}$), and

$$\begin{aligned} &\mathcal{S}^T = (s_1, \ldots, s_{M+1}), \quad \mathcal{Z}^T = (z_1, \ldots, z_{M+1}), \\ &\mathcal{K}^T = (\delta x_1 B_1, \delta x_1 A_1 + \delta x_2 B_2, \ldots, \delta x_{M-1} A_{M-1} \\ &\qquad + \delta x_M B_M, \delta x_M A_M), \\ &\mathcal{L}^T = (D_1, C_1 + D_2, \ldots, C_{M-1} + D_M, C_M). \end{aligned} \tag{6a}$$

Detailed information on $\mathcal{A}, \mathcal{B}, \mathcal{C}$ follows (u.d. $\equiv$ upper diagonal, etc.):

$$\begin{aligned} \text{u.d. } \mathcal{A} &= (\delta x_1^2 P_1, \delta x_2^2 P_2, \ldots, \delta x_M^2 P_M) = \text{l.d.} \\ \text{diag. } \mathcal{A} &= (\delta x_1^2 R_1, \delta x_1^2 Q_1 + \delta x_2^2 R_2, \ldots, \delta x_{M-1}^2 Q_{M-1} \\ &\qquad + \delta x_M^2 R_M, \delta x_M^2 Q_M) \\ \text{u.d. } \mathcal{B} &= (\delta x_1 H_1, \delta x_2 H_2, \ldots, \delta x_M H_M) \\ \text{l.d. } \mathcal{B} &= (\delta x_1 E_1, \delta x_2 E_2, \ldots, \delta x_M E_M) \\ \text{diag. } \mathcal{B} &= (\delta x_1 G_1, \delta x_1 F_1 + \delta x_2 G_2, \ldots, \delta x_{M-1} F_{M-1} \\ &\qquad + \delta x_M G_M, \delta x_M F_M) \\ \text{u.d. } \mathcal{C} &= (S_1, \ldots, S_M) = \text{l.d.} \\ \text{diag. } \mathcal{C} &= (V_1, T_1 + V_2, \ldots, T_{M-1} + V_M, T_M). \end{aligned} \tag{6b}$$

The subscripted upper case coefficients used in (6a) and (6b) above are given below for $k = 1, \ldots, M$, where all summations are from m_k to n_k:

$$\begin{aligned} A_k &= \textstyle\sum y_i u_{ik} e_{ik}, & H_k &= d_k - 2P_k, \\ B_k &= \textstyle\sum y_i v_{ik} e_{ik}, & S_k &= \textstyle\sum e_{ik}^2 + 2(H_k - a_k), \\ C_k &= \textstyle\sum y_i u_{ik}^2 - 2A_k, & T_k &= \textstyle\sum u_{ik}^4 + 4(Q_k - b_k), \\ D_k &= 2B_k + \textstyle\sum y_i v_{ik}^2, & V_k &= \textstyle\sum v_{ik}^4 + 4(c_k + R_k), \\ E_k &= 2P_k + a_k, \\ F_k &= b_k - 2Q_k, \\ G_k &= 2R_k + c_k, \end{aligned}$$

where, $w_{ik} \equiv e_{ik}^2$, $e_{ik} \equiv u_{ik} v_{ik}$ and

$$\begin{aligned} a_k &= \textstyle\sum w_{ik} v_{ik}, & P_k &= \textstyle\sum e_{ik} w_{ik}, \\ b_k &= \textstyle\sum u_{ik}^3 e_{ik}, & Q_k &= \textstyle\sum u_{ik}^2 w_{ik}, \\ c_k &= \textstyle\sum e_{ik} v_{ik}^3, & R_k &= \textstyle\sum v_{ik}^2 w_{ik}, \\ d_k &= \textstyle\sum u_{ik} w_{ik}. \end{aligned} \tag{6c}$$

The assumptions used in the solution (6) imply that the coefficient matrix is nonsingular. If we can show that $\mathcal{A}$ is nonsingular, then it follows that $(\mathcal{C} - \mathcal{B}^T\mathcal{A}^{-1}\mathcal{B})$ is nonsingular. This is true since

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^T & \mathcal{C} \end{bmatrix} \cdot \begin{bmatrix} I & -\mathcal{A}^{-1}\mathcal{B} \\ 0 & I \end{bmatrix} = \begin{bmatrix} \mathcal{A} & 0 \\ \mathcal{B}^T & \mathcal{C} - \mathcal{B}^T\mathcal{A}^{-1}\mathcal{B} \end{bmatrix}$$

from which

$$\det \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^T & \mathcal{A} \end{bmatrix} \cdot \det(\mathcal{A}^{-1}) = \det(\mathcal{C} - \mathcal{B}^T\mathcal{A}^{-1}\mathcal{B}).$$

Therefore we can reduce the size of the matrix inversion problem by computing (using block decomposition)

$$\mathcal{Z} = (\mathcal{C} - \mathcal{B}^T\mathcal{A}^{-1}\mathcal{B})^{-1} * (\mathcal{L} - \mathcal{B}^T\mathcal{A}^{-1}\mathcal{K}),$$

and then setting $\mathcal{S} = \mathcal{A}^{-1} * (\mathcal{K} - \mathcal{B}\mathcal{Z})$.

The fact that $\mathcal{A}^{-1}$ exists follows from the next statements.

LEMMA. *For each $k = 1, \ldots, M$ we have*

$$R_k Q_k - P_k^2 > 0.$$

PROOF. First drop the subscript k temporarily and note that

$$\begin{aligned} P &= \textstyle\sum u_i^3(u_i - 1)^3, \\ Q &= \textstyle\sum u_i^4(u_i - 1)^2, \\ R &= \textstyle\sum u_i^2(u_i - 1)^4. \end{aligned}$$

Then, $(RQ - P^2)$

$$= \sum_j \sum_i u_i^2 u_j^2 (u_i - 1)^2 (u_j - 1)^2$$

$$\left[u_i^2(u_j - 1)^2 - u_i u_j (u_i - 1)(u_j - 1)\right]$$

$$\sum_{i<j}\sum u_i^2 u_j^2 (u - 1)^2 (u - 1)^2 \left[u_j - u_i\right]^2 > 0$$

since by assumption there exist at least two distinct values for u in each subinterval.

THEOREM. $\mathcal{A}$ *is positive definite.*

PROOF. We are going to establish that the principal minors of $\mathcal{A}$, $1 < k < m$,

$$\psi_k = \det \begin{bmatrix} \delta x_1^2 R_1 & \delta x_1^2 R_1 & & \\ \delta x_1^2 P_1 & (\delta x_1^2 Q_1 + \delta x_2^2 R_2) & \ddots & \delta x_k^2 P_k \\ & \ddots & \ddots & (\delta x_{k-1} Q_{k-1} \\ & & \delta x_k^2 P_k & + \delta x_k^2 R_k) \end{bmatrix}$$

$\psi_1 = \delta x_1^2 R_1$, $\psi_M = \det \mathcal{A}$, are all positive, and hence $\mathcal{A}$ satisfies Sylvester's criterion for positive definiteness. $\delta x_k^2 > 0$ and $R_k > 0$ are obvious. $\psi_1 = (\delta x_1)^2 R_1 > 0$. Assume $\psi_{k+1} \geq \psi_k (\delta x_{k+1})^2 R_{k+1} > 0$ for $1 \leq k < n$. For $k = n$ we expand ψ_{n+1} using row $n + 1$.

$$\psi_{n+1} = ((\delta x_n)^2 Q_n + (\delta x_{n+1})^2 R_{n+1})\psi_n - (\delta x_n)^4 P_n^2 \psi_{n-1}$$

$$\geq ((\delta x_n)^2 Q_n + (\delta x_{n+1})^2 R_{n+1})\psi_n - (\delta x_n)^2 P_n^2 \frac{\psi_n}{R_n}$$

from the above assumption,

$$\geq \psi_n ((\delta x_{n+1})^2 R_{n+1} + (\delta x_n)^2 (Q_n - \frac{P_n^2}{R_n}))$$

$$\geq \psi_n (\delta x_{n+1})^2 R_{n+1} > 0,$$

by the above lemma. It follows that $\mathcal{A}$ is positive definite.

Received August 1971, revised March 1972.

References

1. Ferguson, J. Least squares 'spline' curve fitting. TRW Applied Math. Report 3122-11, Feb. 18, 1966.
2. Barrodale, I., and Young A. A note on numerical procedures for approximation by spline functions. *Computer J. 9* (Nov. 1966), 318–320.
3. Reinsch, C.H. Smoothing by spline functions. *Num. Math. 10* (1967), 177–183.
4. Curtis, A.R., and Powell, M.J.D. Using cubic splines to approximate functions of one variable to prescribed accuracy. AERE-R 5602. Harwell, Berk, England, Nov. 1967.
5. Anselone, P.M., and Laurent, P.J. A general method for the construction of interpolating or smoothing spline functions. *Num. Math. 12* (1968), 66–82.
6. deBoor, C., and Rice, J.R. Least squares cubic spline approximation, I—Fixed Knots, II—Variable Knots. CSD TR20 and CSD TR21, Purdue U. Lafayette, Ind., Apr. 1968.
7. Greville, T.N.E. (Ed.) *Theory and Applications of Spline Functions*. Academic Press, New York, 1969.
8. Rivlin, T.J. *An Introduction to the Approximation of Functions*. Blaisdell Pub. Co., Waltham, Mass. 1969.
9 Harrison Jr, J.O. Piecewise polynomial approximation for large scale digital computers. *Math. Tables Aids to Comp. 3* (1949), 400–407.
10. Stone, H. Approximation of curves by line segments. *Math. 15* (1961), 40–47.
11. Hanson, R.J. Computing quadratic programming problems: linear inequality and equality constraints. Jet Propulsion Lab. Sec. 314, Tech. Memo No. 240, Feb. 9, 1970.

Numerical Mathematics

R.A. Willoughby Editor

Cubic Spline Solutions to Fourth-order Boundary Value Problems

W.D. Hoskins
University of Manitoba

The cubic spline approximation to the fourth-order differential equation $y^{iv} + p(x)y'' + q(x)y' + r(x)y = t(x)$ is shown to reduce to the solution of a five-term recurrence relationship. For some special cases the approximation is shown to be simply related to a finite difference representation with a local truncation error of order $(1/720)\delta^8 y$.

Key Words and Phrases: cubic spline, differential equations, boundary value problem

CR Categories: 5.13, 5.17

1. Introduction

The differential equation arising in the case of small deflections of a loaded beam is a fourth-order equation of boundary value type (cf. Hildebrand, 1965 [10]).

Common methods of solution involve finite differences or Ritz Galerkin methods. However, in a recent paper Fyfe, 1970 [8], has demonstrated the application of cubic splines to the solution of such equations, with the final feature of his analysis being the solution of a set of linear equations whose coefficient matrix is of order $n + 3$ and essentially lower triangular with two super diagonals.

Copyright © 1973, Association for Computing Machinery, Inc. General permission to republish, but not for profit, all or part of this material is granted provided that ACM's copyright notice is given and that reference is made to the publication, to its date of issue, and to the fact that reprinting privileges were granted by permission of the Association for Computing Machinery.

Author's address: Computer Science Department, University of Manitoba, Winnipeg, Manitoba, Canada.