Title
AN OPTIMAL CUBIC SPLINE

Author
Young, Jonathan D.

Publication Date
1970-04-01

Submitted to The Logistics Review

UCRL-19763
Preprint c.2

AN OPTIMAL CUBIC SPLINE

Jonathan D. Young

April 1970

AEC Contract No. W-7405-eng-48

LAWRENCE RADIATION LABORATORY
UNIVERSITY of CALIFORNIA BERKELEY

UCRL-19763

AN OPTIMAL CUBIC SPLINE

Jonathan D. Young

Lawrence Radiation Laboratory
University of California
Berkeley, California

April 1970

ABSTRACT

We seek specifications for second derivatives at initial and terminal points which will minimize the third derivative discontinuities of a cubic spline fit passing through a given set of points.

INTRODUCTION

For a given set of data points, (x_i, t_i); $i=1, n$; $n \geq 3$ with the t_i distinct and increasing the cubic spline fit, $s(t)$ for $x(t)$ has the following Properties:

(1) $s(t_i) = x_i$; $i=1, n$

(2) $s \quad C_2 [x_1, x_n]$

(3) s''' is constanton (t_i, t_{i+1}); $i=1, n-1$

The above properties are not sufficient to determine a unique s. Two additional independent conditions must be imposed. Traditionally [1], this was done by stipulating:

$$s''(t_1) = 0 \qquad\qquad s''(t_n) = 0.$$

However [2], any numerical specification of

$$s'(t_1) \text{ or } s''(t_1) \quad \text{and} \quad s'(t_n) \text{ or } s''(t_n)$$

will suffice.

When the imposed conditions are based on known fact about $x(t)$, then the above approach is valid, and may in fact be necessary. On the other hand if specifications are made simply to make the system determinate, this freedom ought to be employed in some useful manner. We note that property (3) enables us to define s_i''' (left) and s_i''' (right) at each interior" point, t_i; $i=2, n-1$. The difference between these two values is the discontinuity of s''' at t_i. We now define an optimal cubic spline fit to be that one for which these discontinuities are minimized in the least square sense. For n=3 or n=4 the optimal spline is immediate since the data can be fitted exactly by a parabola or cubic respectively.

Small n Optimization. Jackson Laslett [3] proposes that conditions be imposed at t_1 and t_2 which will effect this minimization. His procedure has the attractive computational simplicity that the cubic spline fits may be constructed by means of a recursion formula without recourse to the solution of a linear system. The method is plagued by some difficulty unless n is restricted to relatively small values. He suggests $n \leq 10$.

The difficulties for large n seem to be due to

(1) the accumulation of error in successively applying the recursion formula

and

(2) the attenuated response at t_i (for large i) to conditions imposed at t_1 and t_2.

For large n ($n > 10$) it seems necessary to abandon the recursion scheme and return the former methods of imposing conditions at t_1 and t_n

General Optimization. We assume $n \geq 5$ and define

$$e_i \equiv s_i'''(\text{right}) - s_i'''(\text{left}) \quad ; \; i=2, n-1$$

and (somewhat) arbitrarily elect to impose conditions $s''(t_1)$ and $s''(t_n)$ of such values as to minimize

$$\sigma \equiv \sum_{2}^{n-1} e_i^2$$

As stated in [3], three "basic" spline fits are required to find the optimum fit. We elect (again somewhat arbitrarily):

$$s^{(0)} \quad \text{from} \quad s''(t_1) = 0 \quad \text{and} \quad s''(t_n) = 0$$

$$s^{(1)} \quad \text{from} \quad s''(t_1) = 1 \quad \text{and} \quad s''(t_n) = 0$$

$$s^{(2)} \quad \text{from} \quad s''(t_1) = 0 \quad \text{and} \quad s''(t_n) = 1$$

We now express our optimal spline, s, by (1)

$$s = as^{(0)} + bs^{(1)} + cs^{(2)}$$

We then have:

$$s(t_i) = as^{(0)}(t_i) + bs^{(1)}(t_i) + cs^{(2)}(t_i);\ i=1,\ n$$

and by property (1)

$$x_i = ax_i + bx_i + cx_i\ ;\ i=1,\ n$$

hence

$$1 = a + b + c$$

or

$$a = 1 - b - c$$

By differentiating (1)

$$s_i''' \text{ (right)} = as_i^{(0)'''} \text{ (right)} + bs_i^{(1)'''} \text{ (right)} + cs^{(2)'''} \text{ (right)}$$

$$s_i''' \text{ (left)} = as_i^{(0)'''} \text{ (left)} + bs^{(1)'''} \text{ (left)} + cs^{(2)'''} \text{ (left)};\ i=2,\ n-1$$

Consequently

$$e_i = (1-b-c)\ e_i^{(0)} + be_i^{(1)} + ce_i^{(2)}$$

$$\sigma = \sum_2^{n-1} [(1-b-c)\ e_i^{(0)} + be_i^{(1)} + ce_i^{(2)}]^2$$

To minimize σ over (b,c) we must have

$$\frac{\partial\sigma}{\partial b} = 2 \sum_2^{n-1} [(1-b-c)\ e_i^{(0)} + be_i^{(1)} + ce_i^{(2)}]\ [-e^{(0)} + e^{(1)}] = 0 \qquad (2)$$

and

$$\frac{\partial\sigma}{\partial c} = 2 \sum_2^{n-1} [(1-b-c)\ e_i^{(0)} + be_i^{(1)} + ce_i^{(2)}]\ [-e_i^{(0)} + e_i^{(2)}] = 0. \qquad (3)$$

Let $E_i^{(1)} = e_i^{(1)} - e_i^{(0)}$ and $E_i^{(2)} = e_i^{(2)} - e_i^{(0)}$

and let $\sigma_{01} = \Sigma\, e_i^{(0)}\, E_i^{(1)}$ $\qquad \sigma_{02} = \Sigma\, e_i^{(0)}\, E_i^{(2)}$

$\sigma_{11} = \Sigma\, e_i^{(1)}\, E_i^{(1)}$ $\qquad \sigma_{12} = \Sigma\, e_i^{(1)}\, E_i^{(2)}$

$\sigma_{21} = \Sigma\, e_i^{(2)}\, E_i^{(1)}$ $\qquad \sigma_{22} = \Sigma\, e_i^{(2)}\, E_i^{(2)}$

Then from (2) and (3)

$$(1-b-c)\ \sigma_{01} + b\sigma_{11} + c\sigma_{21} = 0 \tag{4}$$

$$(1-b-c)\ \sigma_{02} + b\sigma_{12} + c\sigma_{22} = 0 \tag{5}$$

or

$$(\sigma_{01} - \sigma_{11})\ b + (\sigma_{01} - \sigma_{21})\ c = \sigma_{01} \tag{6}$$

$$(\sigma_{02} - \sigma_{12})\ b + (\sigma_{02} - \sigma_{22})\ c = \sigma_{02} \tag{7}$$

Having computed the coefficients and constant terms of Equations (6) and (7) from the "basic" spline fits we solve these equations for the optimum choice of (b, c).

Again by differentiation of (1)

$$s'' = as^{(0)''} + bs^{(1)''} + cs^{(2)''}$$

hence

$$s''(t_1) = 0 + b + 0$$

$$s''(t_n) = 0 + 0 + c$$

We therefore construct the optimum cubic spline

$$s \text{ from } s''(t_1) = b \quad \text{and} \quad s''(t_n) = c$$

CONCLUSION

If all the data points (t_1, x_n) do in fact lie on a cubic over $[t_1, t_n]$, the resulting optimum cubic spline will be that cubic. A computer code, OPSPLY was written in Fortran to perform the computation described in the previous section. For various sets of data, there was, in all cases, a significant reduction of the third derivative discontinuities for s as computed by the code with those of the natural (or clamped) spline $s^{(0)}$ ($s''(t_1) = 0$, $s''(t_n) = 0$). Listing and description of the computer code are available from the author.

REFERENCES

1. Birkhoff, G. and C. R. deBoor, "Piecewise Polynomial Interpolation and Approximation", Approximation of Functions, Henry L. Garabedian, Ed., (Etsevier Publishing Co., Amsterdam, 1965), 164-168.

2. Young, J., "Numerical Applications of Cubic Spline Functions", The Logistics Review, 3 (1967), 9-14.

3. Laslett, J., "Concerning A Cubic Spline Fit for Interpolation and Differentiation", University of California Lawrence Radiation Laboratory ERAN-68 of March 1970. (Unpublished).