Computer Aided Geometric Design 3 (1986) 205–215
North-Holland

Error analysis for piecewise quadratic curve fitting algorithms

Ronald A. DeVORE * and Zheng YAN

Department of Mathematics, University of South Carolina, Columbia SC 29208, U.S.A.

Received 16 December 1985
Revised 12 August 1986

Abstract. We analyze the convergence of piecewise quadratic curve fitting algorithms which preserve geometric properties of the data such as monotonicity and convexity. In the process, we suggest two new algorithms which improve the order of convergence of existing algorithms.

Keywords. Curve fitting, quadratic splines, monotonicity preservation.

1. Introduction

We are interested in numerically fitting a curve through a given finite set of points $P_i = (x_i, y_i)$, $i = 0, \dots, n$, in the plane. We shall assume that $0 = x_0 < x_1 < \cdots < x_n = 1$, so that these points can be thought of as coming from the graph of some function f defined on [0, 1]. We are particularly interested in algorithms which locally preserve geometric properties of the data or function such as monotonicity or convexity. For example if $\Delta y_i := y_{i+1} - y_i \geqslant 0$, $a \leqslant i < b$, then we may want the resulting interpolant to be monotone on (x_a, x_b).

The best known algorithms of this type were given by [McAllister and Roulier '81] and [Fritz and Carlson '80]. They generate a piecewise polynomial S (quadratic in the first paper and cubic in the latter) such that $y = S(x)$ is the desired interpolant. Here, we shall focus only on those algorithms which use C^1 piecewise quadratics S whose knots are at the x_i and at perhaps one additional point ξ_i in each interval (x_i, x_{i+1}). The McCallister–Roullier algorithm (subsequently denoted by MR) and the algorithms discussed in [Schumaker '83] are of this type.

The main purpose of the present paper is to systematically study the order of convergence of such curve fitting algorithms. In the process, we introduce two new algorithms which have improved convergence property. For our analysis, we assume that the data come from a smooth function: $y_i = f(x_i)$, $i = 0, 1, \dots, n$ and then analyze the error $f - S$.

It is well known that C^1 piecewise quadratics can approximate a three times continuously differentiable f to an order $O(h^3)$ with h the maximum spacing of the knots. We would therefore like our interpolation algorithms to also have this same accuracy. It turns out (as was noted in [Roulier and McAllister '80] that the MR algorithm is of order $O(h^3)$ except near the zeroes of f', where it is only of order $O(h^2)$. Basically, this and other algorithms have difficulty approximating f near the zeroes of f'. We propose in this paper two new algorithms which modify the interpolant when f' is small and thereby have better approximation properties while still preserving monotonicity and convexity.

* This author was supported by NSF grant DMS 8320562.

2. Piecewise quadratics

The knots of a piecewise polynomial S are the points of discontinuity of S or its derivatives. We are interested in piecewise quadratics S which are in C^1 [0, 1], which interpolate our data (x_i, y_i), $i = 0, \ldots, n$, and which have knots at the points x_i, $1 \leqslant i < n$ as well as at an additional point ξ_i in each of the intervals (x_i, x_{i+1}), $0 \leqslant i < n$. Such a piecewise quadratic is completely determined by its values $y_i = S(x_i)$ and its derivative values $s_i := S'(x_i)$, $i = 0, \ldots, n$. Since the values y_i are given by the data, the algorithms for generating a monotonicity preserving interpolant center around making a judicious choice for the derivative values s_i so that monotonicity is preserved. It is not difficult to see [Schumaker '83, Lemma 2.5] that S will be monotone on (x_i, x_{i+1}) if and only if $s_i s_{i+1} \geqslant 0$ and with $\delta_i := \Delta y_i/\Delta x_i$, $\lambda_i := (\xi_i - x_i)/\Delta x_i$ and $\rho_i := (x_{i+1} - \xi_i)/\Delta x_i$,

$$\begin{aligned} \lambda_i s_i + \rho_i s_{i+1} &\leqslant 2\delta_i, \quad \text{when } s_i, s_{i+1} \geqslant 0, \\ \lambda_i s_i + \rho_i s_{i+1} &\geqslant 2\delta_i, \quad \text{when } s_i, s_{i+1} \leqslant 0. \end{aligned} \tag{2.1}$$

Here, (2.1) guarantees that the slope s_{ξ_i} at ξ_i has the same sign as s_i and s_{i+1}. Namely, from the fact that the area under S' on $[x_i, x_{i+1}]$ must equal Δy_i, we find

$$s_{\xi_i} = 2\delta_i - \lambda_i s_i - \rho_i s_{i+1}. \tag{2.2}$$

Since S' is piecewise linear, S will be convex or concave on (x_i, x_{i+1}) if and only if

$$(s_{\xi_i} - s_i)(s_{i+1} - s_{\xi_i}) \geqslant 0. \tag{2.3}$$

In general, given s_i, s_{i+1} and δ_i of the same sign, there may not be any choice of ξ_i in (x_i, x_{i+1}) for which (2.1) holds. In fact, if $\delta_i = 0$, then there is a choice only when both $s_i = 0$ and $s_{i+1} = 0$. If $\delta_i \neq 0$, then there is a choice precisely when

$$0 \leqslant \min(s_i/\delta_i, s_{i+1}/\delta_i) < 2, \quad \text{or when} \quad s_i = s_{i+1} = 2\delta_i. \tag{2.4}$$

The collection of knots ξ_i for which (2.1) holds is a subinterval of (x_i, x_{i+1}) which we call the admissible interval for monotonicity. This interval contains points near x_{i+1} if s_i/δ_i is less than 2, similarly for x_i.

Regarding (2.3), we have a similar situation; (2.3) only holds for a (possibly empty) subinterval of (x_i, x_{i+1}) which we call the admissible interval for convexity. This interval is non-empty [Schumaker '83] if

$$s_i < \delta_i < s_{i+1} \quad \text{or} \quad s_i > \delta_i > s_{i+1} \quad \text{or} \quad s_i = \delta_i = s_{i+1}. \tag{2.5}$$

Again, one can easily check that this admissible interval contains points near an end point whenever it is non-empty. For example, if $|s_i - \delta_i| \geqslant |s_{i+1} - \delta_i|$ and (2.5) holds, then the admissible interval is $(x_i, x_i + 2(s_{i+1} - \delta_i)\, \Delta x_i/(s_{i+1} - s_i)]$. While, if $|s_i - \delta_i| \leqslant |s_{i+1} - \delta_i|$ and (2.5) holds, then the interval is $[x_{i+1} + 2(s_i - \delta_i)\, \Delta x_i/(s_{i+1} - s_i), x_{i+1})$.

Important for us the fact that the the admissible interval for convexity is always contained in the interval for monotonicity provided that $s_i s_{i+1} \geqslant 0$. In fact, if ξ is a point in the admissible interval for convexity and if s_i and s_{i+1} are not of opposite sign, then by (2.3), $S'(\xi)$ shares their sign and hence the piecewise linear function S' does not change sign on (x_i, x_{i+1}), so that S is monotone on this interval.

3. The MR algorithm

The MR algorithm was described by its authors using very involved geometric constructions. It turns out however to have a very simple description using the formulation of Section 2.

Namely, we let

$$h_i := \frac{2\delta_{i-1}\delta_i}{\delta_{i-1}+\delta_i}. \tag{3.1}$$

Then, h_i is the harmonic mean of δ_{i-1} and δ_i. The slopes in the MR algorithm are then defined as follows: for $i = 1,\ldots, n-1$,

$$s_i := \begin{cases} 0, & \text{if } \delta_{i-1}\delta_i \leqslant 0, \\ h_i, & \text{if } \delta_{i-1}\delta_i > 0, \end{cases} \tag{3.2}$$

while at the end points

$$\begin{aligned} s_0 &:= \begin{cases} 0, & \text{if } \delta_0(2\delta_0 - s_1) \leqslant 0, \\ 2\delta_0 - s_1, & \text{else}, \end{cases} \\ s_n &:= \begin{cases} 0, & \text{if } \delta_{n-1}(2\delta_{n-1} - s_{n-1}) \leqslant 0, \\ 2\delta_{n-1} - s_{n-1}, & \text{else}. \end{cases} \end{aligned} \tag{3.3}$$

With this definition for the slopes, it is easy to see that the admissible interval for monotonicity is always the entire interval (x_i, x_{i+1}). Indeed s_i and s_{i+1} are never of opposite signs and therefore this follows from (2.1) and the inequalities $|s_i|$, $|s_{i+1}| \leqslant 2|\delta_i|$. The latter inequalities follow because either $s_i = 0$ or else both δ_{i-1} and δ_i are of the same sign and hence $|s_i| = |h_i| \leqslant 2|\delta_i|$ by the definition of h_i.

In the MR algorithm, the knots ξ_i are chosen as the midpoint of the admissible interval for convexity if that interval is non-empty, otherwise as the midpoint of the interval for monotonicity which as we have already mentioned is all of (x_i, x_{i+1}).

The simplest of examples shows that the MR algorithm is only of order 2. For example the function $f(x) = x^2$ is only approximated with order $O(h^2)$ by this algorithm. The main problem is that the harmonic mean h_i is generally only a first-order approximation to $f'(x_i)$ when x_i is near a zero of f'. The algorithms of the next section attempt to remedy this situation.

4. Alternative algorithms

We want to circumvent the difficulties the MR algorithm has near the zeros of f'. We shall do this by changing the definition for the slopes s_i. We let

$$d_i := (\delta_{i-1}\,\Delta x_i + \delta_i\,\Delta x_{i-1})/(\Delta x_{i-1} + \Delta x_i), \quad 1 \leqslant i < n.$$

The expression for d_i is a three-point formula for calculating derivatives at x_i which is exact for quadratics.

We wish to propose two new curve fitting algorithms. The first will preserve monotonicity in the sense that when the data is monotone on an interval $[x_a, x_b]$, then the resulting interpolant is also monotone on that interval. We show in the next section that this algorithm is of order three for monotone functions. However, it still is generally only of order two near changes of monotonicity. Perhaps more interesting is the second algorithm since it is of order three for arbritrary three times continuously differentiable functions. This algorithm also preserves monotonicity but in a slightly different sense. Namely, if f is monotone on $[x_a, x_b]$, then the interpolant is also monotone on a slightly smaller interval. Both algorithms also have convexity preserving properties.

The slopes for the first algorithm are defined for $i = 1, 2,\ldots, n-2$ by

$$s_i := \begin{cases} 0, & \text{if } \delta_{i-1}\delta_i \leqslant 0, \\ h_i, & \text{if } \delta_{i-1}\delta_i > 0 \text{ and } \min(d_i/\delta_i,\ d_{i+1}/\delta_i) \geqslant 2, \\ d_i, & \text{else}. \end{cases} \tag{4.2}$$

For $i = n-1$, we simply let

$$s_{n-1} := \begin{cases} 0, & \text{if } \delta_{n-2}\delta_{n-1} \leqslant 0, \\ d_{n-1}, & \text{else}. \end{cases} \tag{4.2$'$}$$

At the end points, the slopes are defined as in (3.3). Thus, the difference between these slope assignments and those in MR are that the harmonic mean is not used when δ_{i-1} and δ_i are of comparable size. For example, if $0 \leqslant \delta_{i-1} < 2\delta_i$, then $d_i < 2\delta_i$ and so s_i is defined to be d_i in this case.

Once our slopes have been defined, we choose our knots ξ_i as follows: if the admissible interval for convexity on (x_i, x_{i+1}) is non-empty, we let ξ_i be the midpoint of this interval; if the admissible convexity interval is empty, we let ξ_i be the midpoint of the admissible interval for monotonicity, which as we shall now show is non-empty, irrespective of the original data.

Lemma 1. *Let the data* (x_i, y_i), $i = 0,\dots, n$ *be arbitrary. For the slope assignments* (4.2) *and* (3.3), *the admissible intervals for monotonicity are always non-empty.*

Proof. If $\delta_i = 0$, then $s_i = s_{i+1} = 0$ by (4.2) or (3.3) and therefore the admissible interval for monotonicity is all of (x_i, x_{i+1}). If $\delta_i \neq 0$, it is enough to show (2.4). Hence, we can also assume that $s_i \neq 0$, $s_{i+1} \neq 0$. We consider the remaining possibilities:

Case $i = 0$. It follows from (3.3) that δ_0 and $2\delta_0 - s_1$ are of the same sign. If they are both positive, then $s_0 = 2\delta_0 - s_1 > 0$ and hence (2.4) holds. A similar argument holds when they are both negative.

Case $i = n-1$. This is similar to the above case.

Case $0 < i < n-1$. Since s_i, $s_{i+1} \neq 0$, it follows from (4.2) that δ_{i-1}, δ_i and δ_{i+1} are all of the same sign and hence so are s_i and s_{i+1}. We assume that they are positive; when they are negative the argument is similar. If s_i is defined to be h_i by (4.2), then $s_i = h_i = 2\delta_{i-1}\delta_i/(\delta_i + \delta_{i-1})$, and so $s_i < 2\delta_i$ because $\delta_{i-1}/(\delta_i + \delta_{i-1}) < 1$; similarly if $s_{i+1} := h_{i+1}$. This gives (2.4). In the remaining case $s_i := d_i$ and $s_{i+1} := d_{i+1}$, and therefore the criteria in (4.2) for defining s_i gives (2.4). □

We now show that S preserves the monotonicity and convexity of the data.

Lemma 2. *If* $\Delta y_i \geqslant 0$ ($\Delta y_i \leqslant 0$) *for* $a \leqslant i < b$, *then* S *is non-decreasing* (*non-increasing*) *on* $[x_a, x_b]$. *If* $\Delta\delta_i > 0$ ($\Delta\delta_i < 0$) *for* $a \leqslant i < b$, *then* S *is convex* (*concave*) *on* $[x_{a+1}, x_b]$. *In addition, if* $a = 0$, *then* f *is convex on* $[0, x_b]$. *Similarly, when* $b = n-1$, S *is convex on* $[x_{a+1}, 1]$.

Proof. We know that S is monotone on each interval $[x_i, x_{i+1}]$, $a \leqslant i \leqslant b$ because our knots are chosen from the admissible interval for montonicity. Now, if $\Delta y_i \geqslant 0$, $a \leqslant i < b$, then by (4.2), (3.3), $s_i \geqslant 0$, $a \leqslant i < b$. If s_i or s_{i+1} is positive, S must be non-decreasing on $[x_i, x_{i+1}]$. If both are 0, then from (2.2), $s_{\xi_i} = 2\delta_i$ and again S is non-decreasing on this interval. Since this applies for all $a \leqslant i < b$, S is non-decreasing on $[x_a, x_b]$.

Now suppose $\Delta\delta_i > 0$, $a \leqslant i < b$. In addition, we assume for the moment that $\delta_i \neq 0$, $a \leqslant i < b$. If $\delta_{i-1}\delta_i > 0$, then since s_i is either the harmonic mean or an average of δ_i and δ_{i-1}, it follows that $\delta_{i-1} < s_i < \delta_i$. This also holds if $\delta_{i-1}\delta_i < 0$, since in this case (4.2) gives $s_i = 0$. Hence, we have $\delta_a < s_{a+1} < \delta_{a+1} < \cdots < s_b < \delta_b$. This shows by (2.5) that the admissible interval for convexity in $[x_i, x_{i+1}]$ is non-empty for $i = a+1,\dots, b-1$ and since our knot ξ_i is chosen from this interval, S is convex on $[x_i, x_{i+1}]$, $i = a+1,\dots, b-1$. That is, $s_i \leqslant S'(x) \leqslant s_{i+1}$ on this interval. Hence, S is convex on $[x_{a+1}, x_b]$. If $\delta_j = 0$ for some $a \leqslant j < b$, then $s_j = s_{j+1} = 0$ and therefore the admissible interval for convexity on $[x_j, x_{j+1}]$ is again non-empty. Using this with the above argument gives that S is convex on $[x_{a+1}, x_b]$.

If $a=0$, we get in addition that the admissible interval for convexity in $[x_0, x_1]$ is non-empty. Indeed, if $\delta_0=0$, then this is obvious because $s_0=s_1=0$. If $\delta_0(2\delta_0-s_1)>0$, then (3.3) gives $s_0:=2\delta_0-s_1<\delta_0$ because $\delta_0<s_1$ which shows that (2.5) holds. On the other hand, if $\delta_0\neq 0$ and $\delta_0(2\delta_0-s_1)\leqslant 0$, then we have two possibilities. Either $\delta_0>0$ in which case, $s_0=0<\delta_0$ and therefore (2.5) holds, or $\delta_0\leqslant 0$ and $2\delta_0-s_1\geqslant 0$, in which case $s_1\leqslant 2\delta_0\leqslant\delta_0$, a contradiction. The new information that S is convex on $[x_0, x_1]$ together with what we have already shown gives that S is convex on $[0, x_b]$. In the same way, we can treat the case $b=n-1$. □

In Lemma 2, we get convexity for S only when we assume that the δ_i are strictly increasing. We could weaken this and allow merely non-decreasing of the δ_i if we would allow the knots ξ_i to be chosen as x_i or x_{i+1}. However, this would mean a loss of one degree of smoothness in S at these points.

Lemma 2 does not guarantee convexity on $[x_a, x_{b+1}]$, only on the smaller interval $[x_{a+1}, x_b]$. As was pointed out to us by C. de Boor, the following simple example shows that in general this is the best that we can expect. We consider $f(x):=(x-\tau)_+$ and let x_i, $i=0,\dots,n$ be such that $n\geqslant 4$ and $x_2=\tau$. Then any C^1 piecewise quadratic S which interpolates f at the x_i and which is co-convex with f must be 0 on $[0, \tau]$ (because f is both convex and concave on this interval and $S(x)=x-\tau$ on $[\tau, 1]$ (for the same reason.) But then S is not in C^1.

5. Convergence order

We begin by showing that the algorithm described in Section 4 gives a third-order approximation to a monotone function. We discuss only the case when f is nondecreasing. We let $\|\cdot\|[a, b]$ denote the sup norm on the interval $[a, b]$. When no dependence on the interval is indicated the norm is understood to be on $[0, 1]$. We consider the linear functionals $d_i(f):=d_i$ where d_i is defined as in (4.1) with $y_j:=f(x_j)$, $j=0,\dots,n$. It is easy to check that when f is a quadratic polynomial, we have $d_i(f)=f'(x_i)$, $0<i<n$. Now let f be three times continuously differentiable with $\|f^{(3)}\|=M$. Calculating the divided difference of f, we find

$$|f'(x_i)-d_i|=\Delta x_{i-1}\,\Delta x_i\,|[x_{i-1}, x_i, x_i, x_{i+1}]f|\leqslant Mh^2/6,\quad 0<i<n. \tag{5.1}$$

We can prove a similar inequality for $f'(x_i)-h_i$, $i=1,\dots,n-2$ provided $\min(d_i, d_{i+1})\geqslant 2\delta_i$. Indeed, d_i and h_i both lie between δ_{i-1} and δ_i and therefore by (5.1),

$$|f'(x_i)-h_i|\leqslant|f'(x_i)-d_i|+|d_i-h_i|\leqslant Mh^2/6+|\delta_i-\delta_{i-1}|. \tag{5.2}$$

In order to estimate $|d_i-\delta_{i-1}|$, we calculate the divided difference of f at the points x_j, $j=i-1,\dots,i+2$ and find

$$[x_{i-1}, x_i, x_{i+1}, x_{i+2}]f=\frac{\dfrac{\delta_{i+1}-\delta_i}{x_{i+2}-x_i}-\dfrac{\delta_i-\delta_{i-1}}{x_{i+1}-x_{i-1}}}{x_{i+2}-x_{i-1}}. \tag{5.3}$$

Since d_i is a convex combination of δ_{i-1} and δ_i and by our assumption $d_i\geqslant 2\delta_i$, we have $\delta_{i-1}\geqslant 2\delta_i$. Similarly, $\delta_{i+1}\geqslant 2\delta_i$. So the numerator on the right side of (5.3) is the sum of two non-negative terms and therefore each term in the numerator is smaller than $(x_{i+2}-x_{i-1})\|f^{(3)}\|/6$. This gives $|\delta_i-\delta_{i-1}|\leqslant Mh^2$. Using this in (5.2) shows that

$$|f'(x_i)-h_i|\leqslant 2Mh^2,\quad 0<i\leqslant n-2,\quad \text{provided}\ \min(d_i, d_{i+1})\geqslant 2\delta_i. \tag{5.4}$$

Theorem 1. *If f is three times continuously differentiable and monotone on* [0, 1], *then the algorithm of Section* 4 *generates a piecewise quadratic S satisfying*

$$\| f - S \| \leqslant 3 \| f^{(3)} \| h^3. \tag{5.5}$$

Proof. We first observe that

$$|f'(x_i) - s_i| \leqslant \begin{cases} 2Mh^2, & 0 < i < n, \\ 3Mh^2, & i = 0, n. \end{cases} \tag{5.6}$$

Consider first the case $0 < i < n$. If $s_i = d_i$ or $s_i = h_i$, then this is (5.1) or (5.4) respectively. If $s_i = 0$, then, from (4.2), either $\delta_{i-1} = 0$ or $\delta_i = 0$ and hence f' vanishes on an interval which contains x_i and (5.6) follows.

We now check the case $i = 0$; the case $i = n$ is handled similarly. We have

$$|2\delta_0(f) - f'(x_0) - f'(x_1)| = (x_1 - x_0)^2 |[x_0, x_0, x_1, x_1] f| \leqslant Mh^2/6. \tag{5.7}$$

Using this with (5.6) for $i = 1$, we have that $|2\delta_0 - s_1 - f'(x_0)| \leqslant 3Mh^2$. This is the desired estimate when s_0 is defined as $2\delta_0 - s_1$. On the other hand if $s_0 := 0$, then by (3.3), we have two possibilities. If $\delta_0 \leqslant 0$, then $\delta_0 = 0$ because f is monotone and $f'(x_0) = 0$ so that (5.6) is obvious. If $2\delta_0 - s_1$ is non-positive, then by (5.7) and the first part of (5.6), we find $f'(x_0) \leqslant 3Mh^2$ and (5.6) (for $i = 0$) follows in this case as well.

For each continuously differentiable function g, we let

$$d'(g) := 2\delta_i(g) - \lambda_i g'(x_i) - \rho_i g'(x_{i+1})$$

with $\delta_i(g) := (g(x_{i+1}) - g(x_i))/\Delta x_i$ and λ_i, ρ_i as in (2.1). It then follows that $d'(Q) = Q'(\xi_i)$ for any quadratic polynomial Q. If we take Q as the quadratic Taylor polynomial at ξ_i, then $\| f' - Q' \| [x_i, x_{i+1}] \leqslant Mh^2/2$. Hence, using the mean value theorem, we have $|\delta_i(f - Q)| = |f'(\xi) - Q'(\xi)| \leqslant Mh^2/2$, with $\xi \in (x_i, x_{i+1})$. Hence

$$|f'(\xi_i) - d'(f)| = |d'(f - Q)| \leqslant 2Mh^2/2 + Mh^2/2 \leqslant 2Mh^2.$$

Using this with (5.6), and the fact that $S'(\xi_i) = s_{\xi_i}$ is given by (2.2), we find

$$|f'(\xi_i) - S'(\xi_i)| \leqslant 5Mh^2. \tag{5.8}$$

This shows that the piecewise linear function S' approximates f' to an error at most $5Mh^2$ at each of its beam points. Hence, if T is the piecewise linear interpolant to f' at these breakpoints, then $\| T - S' \| \leqslant 5Mh^2$. As is well known [2], $\| f' - T \| \leqslant Mh^2/8$. This gives $\| f' - S' \| \leqslant 6Mh^2$. Since $f - S$ vanishes at each x_i, $i = 0, \ldots, n$, we obtain (5.5) by integration from the nearest x_i to x. □

6. A second algorithm

We have shown third-order convergence for the algorithm in Section 5 only when f is monotone on all of [0, 1]. When f changes monotonicity on [0, 1], this same argument shows only third-order convergence on intervals $[x_a, x_b]$ contained *strictly* inside the intervals of montonicity. On intervals where the montonicity changes, there will generally only be second-order convergence. The reason for this is that if $\delta_i = 0$, then S is identically constant on $[x_i, x_{i+1}]$ while in general the most we can say about the error is that $\| f - S \| [x_i, x_{i+1}] \leqslant$ const. $\| f'' \| h^2$.

The first algorithm assigns slope zero at x_i if the data change monotonicity, i.e. $\delta_{i-1}\delta_i \leqslant 0$. On the other hand, if the data come from a function f, we know only that f changes monotonicity on (x_{i-1}, x_{i+1}). It is for this reason that the algorithm will generally only be of order two near such changes of monotonicity. On the other hand, there seems little point to requiring that S' change sign exactly at x_i. We therefore propose a second algorithm, where (4.2) is replaced by

$$s_i := \begin{cases} 0, & \text{if either } \delta_i = 0 \text{ and } \delta_{i-1}\delta_{i+1} \geqslant 0 \text{ or } \delta_{i-1} = 0 \text{ and} \delta_{i-2}\delta_i \geqslant 0, \\ h_i, & \text{if } \delta_{i-1}\delta_i > 0 \text{ and } \min(d_i/\delta_i,\ d_{i+1}/\delta_i) \geqslant 2, \\ d_i, & \text{otherwise.} \end{cases} \tag{6.1}$$

Thus, we have only changed s_i when $\delta_{i-1}\delta_i \leqslant 0$. For the end points, we define s_0 and s_n by

$$\begin{aligned} s_0 &:= 2\delta_0 - s_1, \\ s_n &:= 2\delta_{n-1} - s_{n-1}. \end{aligned} \tag{6.2}$$

Our knots ξ_i are defined as follows. If the admissible interval for convexity is not empty we take ξ_i as the midpoint of this interval; if the admissible interval for convexity is empty we take ξ_i as the midpoint of the admissible interval for monotonicity if it is not empty; otherwise we take ξ_i as the midpoint of $[x_i,\ x_{i+1}\}$. Because of our interpolation conditions, $S'(\xi_i)$ is again given by (2.2).

As before, for a function $f \in C[0, 1]$, we let $y_i := f(x_i)$, $i = 0,\dots, n$,

Theorem 2. *If f is three times continuously differentiable on* [0, 1], *then the algorithm of this section generates from f a piecewise quadratic S satisfying*

$$\| f - S \| \leqslant 3 \| f^{(3)} \| h^3. \tag{6.3}$$

Proof. We first show that

$$| f'(x_i) - s_i | \leqslant 2Mh^2, \quad i = 1,\dots, n-1. \tag{6.4}$$

We consider the three possible slope assignments of (6.1). If $s_i := d_i$, then, since (5.1) holds for any three times continuously differentiable f, we have (6.4). If $s_i := h_i$, then $\min(d_i/\delta_i,\ d_{i+1}/\delta_i) \geqslant 2$ and $\delta_{i-1}\delta_i > 0$ and so (5.4) holds, which gives (6.4) in this case. If s_i is defined by the first line of (6.1), then we have two cases. One possibility is that $\delta_i = 0$ and $\delta_{i-1}\delta_{i+1} \geqslant 0$. When this is the case, we compute the divided difference of f at $x_{i-1},\ x_i,\ x_{i+1},\ x_{i+2}$ as in (5.3). Since the numerator in (5.3) is the sum of two terms of the same sign, we get that $| \delta_i - \delta_{i-1} | \leqslant Mh^2$. Now, $\delta_i = 0$ and d_i lies between δ_{i-1} and δ_i. Hence, from (5.1) we find

$$| f'(x_i) - 0 | \leqslant | f'(x_i) - d_i | + | \delta_i - \delta_{i-1} | \leqslant Mh^2 + Mh^2 \leqslant 2Mh^2,$$

as desired. The remaining possibility when $\delta_{i-1} = 0$ and $\delta_{i-2}\delta_i \geqslant 0$ is proved in the same manner by using the divided difference of f at the points $x_{i-2},\ x_{i-1},\ x_i,\ x_{i+1}$. We have thus established (6.4).

Since (5.7) hold for any f, we have $| f'(x_0) - s_0 | \leqslant 3Mh^2$. The same is true for $i = n$, hence

$$| f'(x_i) - s_i | \leqslant 3Mh^2, \quad i = 0 \text{ and } i = n. \tag{6.5}$$

From (6.4) and (6.5), we can complete the proof exactly as in Theorem 1. □

We now discuss the monotonicity preserving properties of this algorithm,

Theorem 3. *The spline S generated by the algorithm of this section has no more monotonicity changes on $[x_1,\ x_{n-1}]$ than the sequence δ: $\delta_0,\dots, \delta_{n-1}$ has sign changes.*

Proof. We will show that S' changes sign on $[x_1,\ x_{n-1}]$ no more often than δ. In view of the definition of s_i, $i = 1,\dots, n-1$, we see that s_i is non-negative whenever both δ_{i-1} and δ_i are non-negative and s_i is non-positive if both δ_{i-1} and δ_i are non-positive. Hence, the sequence λ: $\delta_0, s_1, \delta_1, s_2, \ldots, \delta_{n-2}, s_{n-1}, \delta_{n-1}$ changes sign no more often than δ. We now want to insert s_{ξ_i}, into the sequence λ between s_i and s_{i+1}, $i = 1, \ldots, n-2$ in such a way that the resulting sequence λ^* has no more sign changes than λ. The only case which is not trivial is when s_i, δ_i, s_{i+1} are all of the same sign (or zero). We now show that in this case, the admissible interval for monotonicity on $[x_i, x_{i+1}]$ is non-empty and hence s_{ξ_i} agrees in sign with s_i and s_{i+1}. Thus it is possible to insert s_{ξ_i} in this case as well.

We shall only consider the case when s_i, δ_i, and s_{i+1} are all non-negative; the case when they are non-positive is similar. We shall check the various possibilities:

Case $\delta_i > 0$.

Subcase: $s_i = 0$, *or* $s_{i+1} = 0$. Here, (2.4) obviously holds.

Subcase: $s_i := h_i$ *or* $s_{i+1} := h_{i+1}$ *by* (6.1). Then either $s_i < 2\delta_i$ or $s_{i+1} < 2\delta_i$ and so (2.4) holds.

Subcase: $s_i := d_i$ *and* $s_{i+1} := d_{i+1}$ *by* (6.1). Then either $\delta_{i-1} > 0$ in which case the criteria in (6.1) gives (2.4) or $\delta_{i-1} \leqslant 0$, in which case $s_i \leqslant \delta_i$ (because it is a convex combination of δ_i and δ_{i-1}) and (2.4) holds.

Case $\delta_i = 0$.

Subcase: $s_i = s_{i+1} = 0$. The admissible interval for monotonicity is all of (x_i, x_{i+1}) in this case.

Subcase: $s_i > 0$ *and* $s_{i+1} = 0$. Then, we must have $s_i := d_i$ by (6.1) and hence $\delta_{i-1} > 0$. The criteria in (6.1) also give that $\delta_{i+1} < 0$ (else s_i is zero). But then s_{i+1} must be defined to be d_{i+1}

Table 1

n	MR	Algorithm of Section 4	Algorithm of Section 6
$y = x^2$			
16	0.488281250000d − 03	0.138777878078d − 16	
32	0.122070312500d − 03	0.277555756156d − 16	same as
64	0.305175781250d − 04	0.138777878078d − 16	algorithm of
128	0.762939453125d − 05	0.277555756156d − 16	Section 4
256	0.190734864281d − 05	0.277555756156d − 16	
	$O(h^2)$	zero	
$y = \cos(x)$			
16	0.277829405296d − 03	0.126783470478d − 04	
32	0.560392383724d − 04	0.161480136285d − 05	same as
64	0.157472067168d − 04	0.203664441756d − 06	algorithm of
128	0.387501773202d − 05	0.255695074003d − 07	Section 4
256	0.961169510830d − 06	0.320309312407d − 08	
	$O(h^2)$	$O(h^3)$	
$y = x \sin x$			
32	0.130084133320d − 03	0.5913541372l4d − 05	
64	0.314947442995d − 04	0.743824330129d − 06	same as
128	0.775005548855d − 05	0.932565455969d − 07	algorithm of
256	0.19223425350d − 05	0.116741301071d − 07	Section 4
512	0.47870542026d − 06	0.146032175241d − 08	
	$O(h^2)$	$O(h^3)$	
$y = \cos(6x)$			
32	0.277455600555d − 02	0.371189149842d − 02	0.294413496052d − 03
64	0.567109080345d − 03	0.104923798966d − 02	0.363051687600d − 04
128	0.277341531626d − 03	0.276519887848d − 03	0.448985110779d − 05
256	0.667778913560d − 04	0.655773692415d − 04	0.802927047516d − 06
512	0.152910747667d − 04	0.143150564327d − 04	0.979241505661d − 07
	$O(h^2)$	$O(h^2)$	$O(h^3)$

by (6.1) because the other two criteria are not met. Since d_{i+1} is a convex combination of δ_i and δ_{i+1}, we have $s_{i+1} < 0$ which is a contradiction and shows that this case does not occur.

Subcase: $s_i = 0$ *and* $s_{i+1} > 0$. Similar to the previous subcase.

We have therefore shown that it is possible to create the sequence λ^* as described above to have no more sign changes than λ. Now, τ: $s_1, s_{\xi_1}, s_2, s_{\xi_2}, \ldots, s_{n-1}$ is a subsequence of λ^* and hence changes sign no more often than λ^*. Since S' changes sign on $[x_1, x_{n-1}]$ no more often than the sequence τ, which in turn changes sign no more often than δ, we have the desired result. □

We note that the spline S may have an additional monotonicity change on the end intervals $[x_0, x_1]$ and $[x_{n-1}, x_n]$. This seems to be unavoidable (apparently because only forward differences can be used in the definition of s_0), if (6.3) is to hold for every f in $C^{(3)}$ [0, 1]. Indeed, the algorithm must then preserve quadratics. We can choose a quadratic Q with $Q'(x_0) < 0$ and $\delta_0, \delta_1, \delta_2 > 0$. Then, any function f with the same δ_0, δ_1 and δ_2 as Q will be assigned a negative slope s_0 at x_0. Since such a function can be monotone on [0, 1], we have an undesired change of monotonicity. Since S' approximates f' to an error $\leqslant cMh^2$, it follows that $|f'(x)| \leqslant cMh^2$ on $[x_0, x_1]$ with $M := \| f^{(3)} \|$. Therefore f actually varies only slightly on this interval relative to the size of M and h.

Example 1. Akima data

x	0	2	3	5	6	8	9	11	12	14	15
y	10	10	10	10	10	10	10.5	15	50	60	85

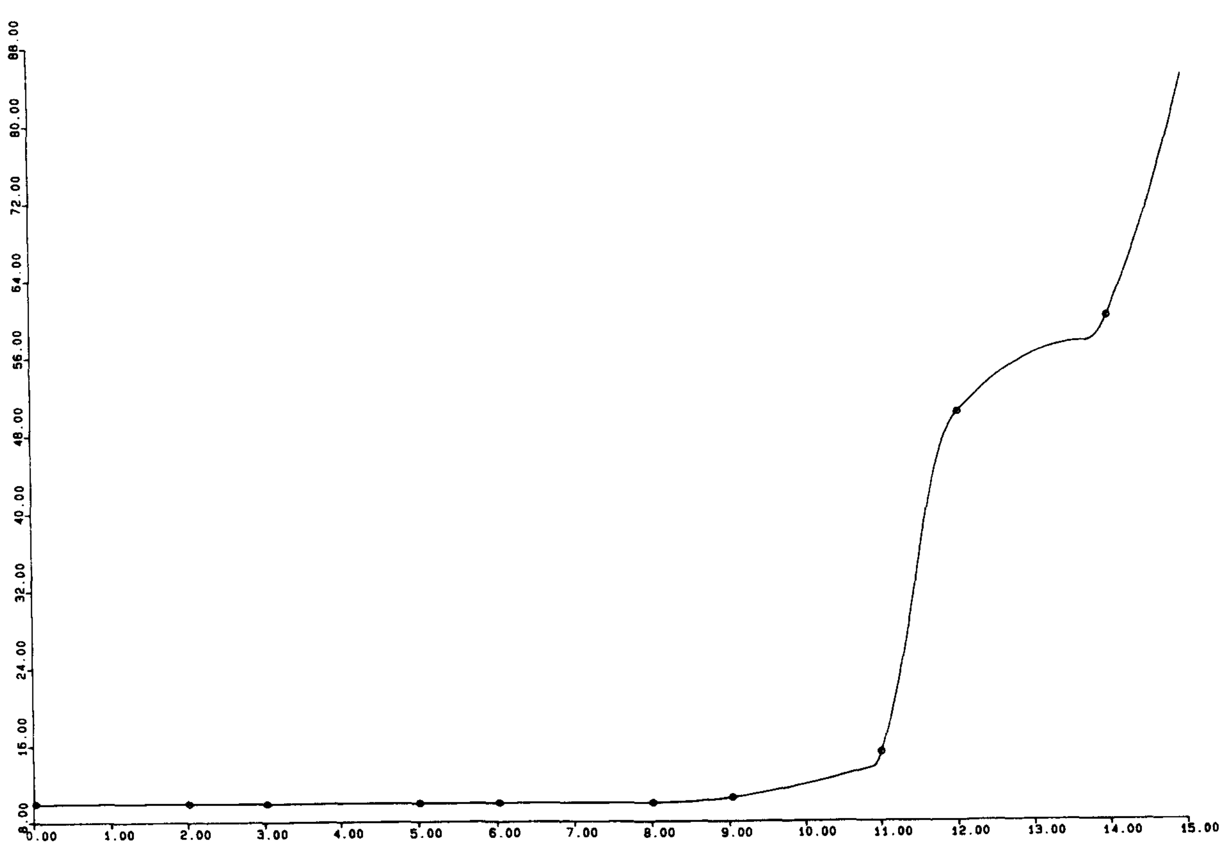

Fig. 1. Akima data.

7. Numerical examples

We give some examples of smooth functions and analyze the error in approximation by the MR algorithm and the algorithms of Section 4 and Section 6. In these examples, we take $n+1$ equally spaced points on [0, 1] and estimate the error using chopping algorithms for finding the maximum of a function. See Table 1.

8. Sample graphs

We now give some example of graphs produced by the Algorithms of Section 4 and Section 6. The examples are for three data sets which are standard examples in the literature. The first example was introduced by [Akima '70]; the second is RNP 14 which is acutal data from LLL radiochemical calculations and was introduced by [Fritsch and Carlson '80]; the third is an

Example 2. RNP 14 data

x	7.99	8.09	8.19	8.7	9.2	10	12	15	20
y	0	2.76429e−5	4.37498e−2	0.169183	0.469428	0.943740	0.998636	0.999919	0.999994

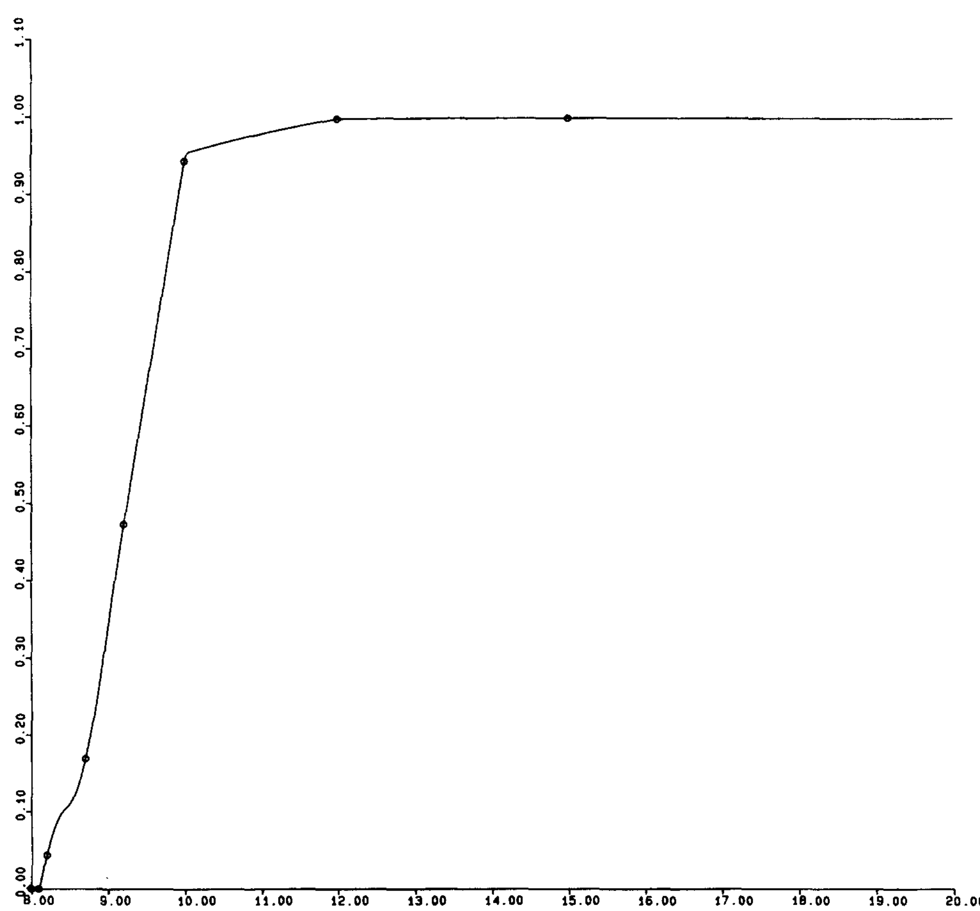

Fig. 2. RNP 14 data.

Example 3. Titanium heat data

$x_i = 585 + 10i, \quad i = 1, \dots, 49$

y_i = 0.644, 0.622, 0.638, 0.649, 0.652, 0.639, 0.646, 0.657, 0.652, 0.655, 0.644, 0.663, 0.663, 0.668, 0.676, 0.676, 0.686, 0.679, 0.678, 0.683, 0.694, 0.699, 0.710, 0.730, 0.763, 0.812, 0.987, 1.044, 1.336, 1.881, 2.169, 2.075, 1.598, 1.211, 0.916, 0.746, 0.672, 0.627, 0.615, 0.607, 0.606, 0.609, 0.603, 0.601, 0.603, 0.601, 0.611, 0.601, 0.608.

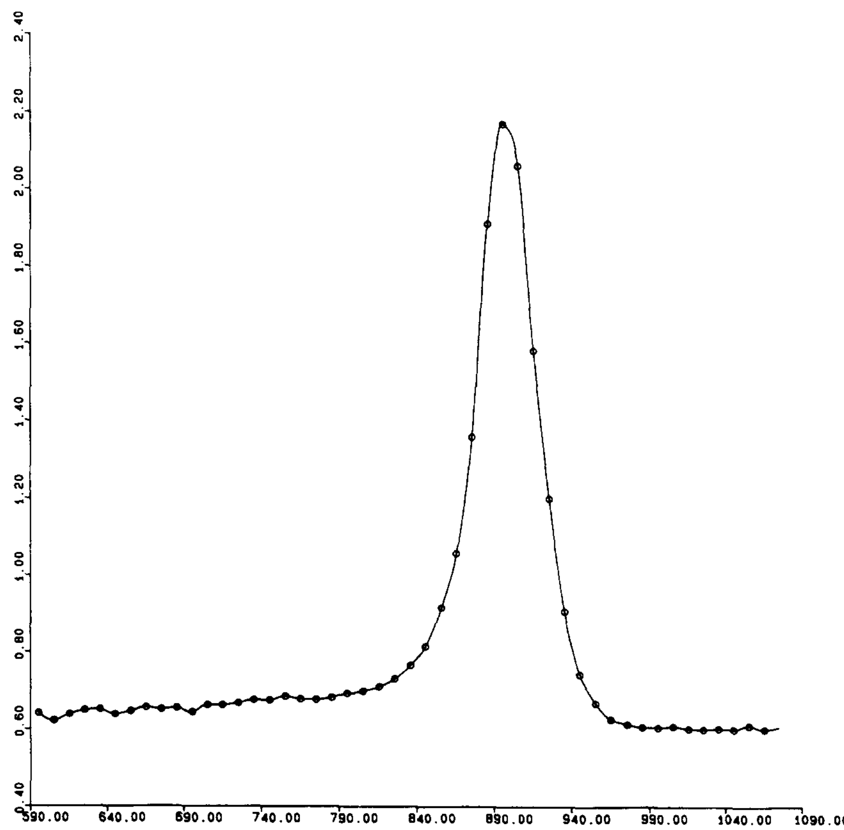

Fig. 3. Titanium heat data.

example of Titanium heat data which was introduced by [de Boor '78]. For this data, the algorithms of Sections 4 and 6 produce the exact same graphs.

Acknowledgement

The authors would like to thank Professor W. Boehm for his careful reading of the original manuscript especially for pointing out an error in the original version of the algorithm of Section 6

References

Akima, H. (1970), A new method of interpolation and smooth curve fitting based on local procedures, J. ACM 17, 589–602.

de Boor, C. (1978), *A Practical Guide to Splines*, Applied Mathematical Sciences, Vol. 27, Springer, New York.

Fritsch, F.N. and Carlson, R.E. (1980), Monotone piecewise cubic interpolation, SIAM J. Numer. Anal. 17, 238–246.

McAllister, D.F. and Roulier, J.A. (1981), An algorithm for computing a shape-preserving osculatory quadratic spline, ACM Trans. Math. Software 7, 381–347.

Schumaker, L.L. (1983), On shape preserving quadratic spline interpolation, SIAM J. Numer. Anal. 20, 854–864.

Roulier, J.A. and McAllister, D.F. (1980), Approximation by convex quadratic splines, in: E.W. Cheney, ed., *Approximation Theory* III, Academic Press, New York.