Algorithms for Computing Shape Preserving Spline Interpolations to Data

By David F. McAllister*, Eli Passow** and John A. Roulier*

Abstract. Algorithms are presented for computing a smooth piecewise polynomial interpolation which preserves the monotonicity and/or convexity of the data.

1. Introduction. Shape preserving polynomial or spline interpolation to monotonic and/or convex data has been investigated by several authors in recent years (see [1]–[5] and [7]–[16]). It is the purpose of this paper to develop and test algorithms for computing smooth monotonicity and/or convexity preserving spline interpolation. The algorithms presented herein are based on the theory developed by Passow and Roulier [11]. The algorithms are tested on several examples and are compared with ordinary Lagrange and cubic spline interpolation of increasing convex data. It should be noted that algorithms for convexity preserving interpolation by exponential splines are given in [14].

The authors thank the referee of the original version of this article for pointing out numerous related articles on computer aided geometric design. See [1], [3], and [4]. In particular, see the articles in [1] by Gordon and Riesenfeld p. 95, Bezier p. 127, Forrest p. 17, Wielinga p. 153, and Nielson p. 209. Many of these articles refer to Bezier curves which are Bernstein polynomials of parametrized polygonal segments. While the ideas are similar to those presented here and make use of the convexity preserving properties of Bernstein polynomials, they deal more with construction of some suitable shape rather than interpolation of given data with shape preservation. Here we assume that the data is given and that no additional data points can be easily obtained. Moreover, the alpha algorithm and its use in Theorem 3.2 allows one to minimize the degree of the piecewise polynomials without losing the properties of interpolation or convexity. This algorithm is automatic and requires no interaction on the part of the user as do many of the techniques in computer-aided geometric design.

2. Notation and Background. Let \( \Delta = \{x_0 < x_1 < \cdots < x_N\} \) be fixed real numbers and for \( j \leq n \) let \( S_n^j = S_n^j(\Delta) \) be the set of splines of degree \( n \) and deficiency \( n-j \) on \( \Delta \). Thus, \( f \in S_n^j(\Delta) \) if and only if \( f \in C^j[x_0, x_N] \) and \( f \) is an algebraic polynomial of degree \( n \) or less on \([x_{i-1}, x_i]\) for \( i = 1, 2, \ldots, N \).

Given corresponding real numbers \( y_0, \ldots, y_N \), define \( S_i = (y_i - y_{i-1})/(x_i - x_{i-1}) \) for \( i = 1, 2, \ldots, N \). We say that the data are nondecreasing if \( y_0 \leq y_1 \leq \cdots \leq y_N \), and the data are nonconcave if \( S_1 \leq S_2 \leq \cdots \leq S_N \). If

Received June 6, 1976; revised November 21, 1976.
*Supported in part by NSF Grant # MCS 76-04033.
**Supported in part by a Temple University Grant-in-Aid of Research.
no equality exists in either of these cases, we say that the data are increasing or convex, respectively.

**Definition 2.1.** Suppose that the data \((x_i, y_i), i = 0, 1, \ldots, N\) are nondecreasing (and/or nonconcave). Let \(\{\alpha_i\}_{i=1}^N\) be given with \(0 < \alpha_i < 1\). Let \(x_i = x_{i-1} + \alpha_i \Delta x_i\) where \(\Delta x_i = x_i - x_{i-1}\) for \(i = 1, 2, \ldots, N\). A set of numbers \(\{t_i\}_{i=1}^N\) is said to be increasing (and/or convex) \(\{\alpha_i\}\)-admissible if the piecewise linear function \(L(x)\) generated by the points

\[(x_0, y_0), (\bar{x}_1, t_1), (\bar{x}_2, t_2), \ldots, (\bar{x}_N, t_N), (x_N, y_N)\]

passes through the points \((x_i, y_i), i = 1, \ldots, N-1\), and is nondecreasing (and/or nonconcave). If \(\alpha_i = \alpha\) for all \(i\), we say that the numbers are \(\alpha\)-admissible.


**Theorem 2.1.** Let \(n_i, m_i\) be natural numbers, with \(m_i < n_i, i = 1, 2, \ldots, N\), and let \(\alpha_i = m_i/n_i, i = 1, 2, \ldots, N\). Let \(n = \max\{n_i : i = 1, \ldots, N\}\) and \(m = \min\{m_i - n_i : i = 1, \ldots, N\}\). Then there exist increasing (and/or convex) \(\{\alpha_i\}\)-admissible points for \(\Delta\) if and only if for all \(k > 1\) there exists \(f \in S_{mn}^{m}(\Delta)\) satisfying \(f\) is a polynomial of degree \(kn_i\) or less on \([x_{i-1}, x_i], i = 1, \ldots, N\), and

\[
\begin{align*}
(a) \quad & f(x_i) = y_i, i = 0, 1, \ldots, N; \\
(b) \quad & f^{(j)}(x_i^+) = 0, j = 2, 3, \ldots, m_i + 1; i = 1, 2, \ldots, N-1; \\
(c) \quad & f^{(j)}(x_i^-) = 0, j = 2, 3, \ldots, (n_i - m_i); i = 1, 2, \ldots, N-1; \\
(d) \quad & f'(x) > 0, x \in [x_0, x_N] \text{ (and/or } f''(x) > 0, x \in [x_0, x_N]).
\end{align*}
\]

(Thus, in particular, \(f^{(j)}(x_i) = 0, j = 2, 3, \ldots, m_i; i = 1, 2, \ldots, N-1.\)

Furthermore, if there exist increasing (and/or convex) \(\{\alpha_i\}\)-admissible points as above, then an \(f\) is constructed by setting \(f(x) = q_i(x)\) on \([x_{i-1}, x_i]\) where \(q_i\) is the Bernstein polynomial (see [6]) of degree \(kn_i\) of the restriction of the broken line segment function \(L(x)\) to \([x_{i-1}, x_i]\). That is,

\[
q_i(x) = \frac{1}{(x_i - x_{i-1})^{kn_i}}
\sum_{\nu=0}^{kn_i} \frac{(kn_i)_\nu}{\nu!} L(x_i - x_{i-1}^{\nu})(x - x_{i-1})^{\nu}(x_i - x)^{kn_i - \nu}.
\]

The function \(f\) is not in general unique, since the set \(\{t_i\}\) is not in general unique. Thus, any algorithm which will give increasing (and/or convex) \(\{\alpha_i\}\)-admissible points will result in an algorithm for computing such an interpolant \(f\). Furthermore, for computational purposes, it is generally not worthwhile to consider \(f \in S_m^n(\Delta)\) for \(m \geq 2\) because of (2.1)(b) and (c). However, the fact that higher order continuity is possible is significant theoretically and indicates that it may be possible to avoid the difficulties of \(f^{(k)}(x_i) = 0\) for \(k \geq 2\).

It is also shown in [11] for convex interpolation that if \(n\) is any fixed positive integer there exists a convex set of data points \((x_i, y_i), i = 0, 1, 2, 3, 4,\) such that no \(f \in S_n^4(x_0, x_1, x_2, x_3, x_4)\) satisfies \(f(x_i) = y_i\) for \(i = 0, 1, 2, 3, 4,\) and \(f\) is convex.
SHAPE PRESERVING SPLINE INTERPOLATIONS

This result has been used to produce examples of convex increasing data for which neither Lagrange nor cubic spline interpolation is convex and increasing. The example in Section 4 of this paper was constructed using this.

3. Tests for $\{\alpha_i\}$-Admissibility. In this section we will restrict ourselves to increasing and convex $\{\alpha_i\}$-admissibility. For brevity we will use the term $\{\alpha_i\}$-admissibility with no prefix. We present in this section two algorithms for $\{\alpha_i\}$-admissibility. That is, we will construct interpolating, shape preserving polygonal segments whose vertices interlace the abscissas of the given data. The first of these is constructive in nature and produces for given nonconcave nondecreasing data, numbers $0 < \beta_i < 1, i = 1, \ldots, N$, such that there exists $\{\alpha_i\}$-admissible points for each collection $\{\alpha_i\}_{i=1}^N$ with $0 < \alpha_i \leq \beta_i$ for $i = 1, \ldots, N$. Once a set of such $\{\alpha_i\}$ is determined, the calculation of the set $\{t_i\}_{i=1}^N$ is given. This algorithm, however, does not necessarily specify all possible $\{\alpha_i\}_{i=1}^N$ for which there are $\{\alpha_i\}$-admissible points. The following theorem is the basis for this algorithm.

**Theorem 3.1.** Let the data $(x_i, y_i)$, $i = 0, 1, \ldots, N$, be convex and increasing and define $S_i$, $i = 1, \ldots, N$, as above. Let $S_0 = 0$ and $\tilde{x}_N = x_N$. Now define

$$
\beta_i = \frac{x_i - x_{i-1}}{\tilde{x}_i - \tilde{x}_{i-1}} \quad \text{for } i = 1, \ldots, N,
$$

where

$$
\tilde{x}_i = \frac{y_{i-1} - y_i + S_{i+1}x_{i+1} - S_{i-1}x_{i-1}}{S_{i+1} - S_{i-1}} \quad \text{for } i = 1, 2, \ldots, N-1.
$$

Then, given any $\{\alpha_i\}_{i=1}^N$ with

$$0 < \alpha_i \leq \beta_i \quad \text{for } i = 1, \ldots, N,$

there exist $\{\alpha_i\}$-admissible points for the given data.

**Proof.** Note that $\tilde{x}_i$ as defined above is the x-coordinate of the point of intersection of the lines determined by the points $(x_{i-2}, y_{i-2})$ and $(x_{i-1}, y_{i-1})$ and by $(x_i, y_i)$ and $(x_{i+1}, y_{i+1})$, for $i = 2, 3, \ldots, N-1$. Also, $\tilde{x}_1$ is the x-coordinate of the point of intersection of the horizontal line through $(x_0, y_0)$ and the line determined by $(x_1, y_1)$ and $(x_2, y_2)$. Finally, $\tilde{x}_N$ is the x-coordinate of the point of intersection of the vertical line through $(x_N, y_N)$ and the line determined by $(x_{N-2}, y_{N-2})$ and $(x_{N-1}, y_{N-1})$.

Now consider the triangle $T_i$ formed by the vertical line through $\tilde{x}_i$, the line determined by $(x_{i-1}, y_{i-1})$ and $(x_i, y_i)$ and the line determined by $(x_{i-2}, y_{i-2})$ and $(x_{i-1}, y_{i-1})$ for $i = 2, \ldots, N$. For $i = 1$, we replace the last line with the horizontal line through $(x_0, y_0)$.

From Figure 1 we see that if $(\tilde{x}_i, \tilde{y}_i)$ is any point in $T_i$ then the line formed by $(\tilde{x}_i, \tilde{y}_i)$ and $(x_i, y_i)$ lies in the triangle $T_{i+1}$ for all values of $x$ between $x_i$ and $\tilde{x}_{i+1}$. Thus given $\{\alpha_i\}_{i=1}^N$, $0 < \alpha_i \leq \beta_i$, the points $\xi_i = x_{i-1} + \alpha_i(x_i - x_{i-1})$ all satisfy $x_{i-1} < \tilde{x}_i \leq \xi_i$. Hence, if $\tilde{y}_i$ is chosen so that $(\tilde{x}_i, \tilde{y}_i)$ lies in $T_i$ then this uniquely determines $\tilde{y}_2$ so that $(\tilde{x}_2, \tilde{y}_2)$ lies in $T_2$ and so forth. It follows that $\{\tilde{y}_i\} = \{\alpha_i\}$-admissible.
The second algorithm gives necessary and sufficient conditions for the existence of an \( \{ \alpha_j \} \)-admissible set for given data. If \( \{ \alpha_j \}_{i=1}^N \) are determined for which there exists a set \( \{ t_j \}_{i=1}^N \) of \( \{ \alpha_j \} \)-admissible points, then the calculation of such a set \( \{ t_j \}_{i=1}^N \) is given. The alpha-algorithm presented below and Theorem 3.2 which refers to it are the basis for this second algorithm.

The Alpha-Algorithm. Let \( \{ \alpha_j \}_{i=1}^N \) with \( 0 < \alpha_i < 1 \) be given.

Define \( m_0 = 0 \) and \( M_0 = S_1 \). Now for \( i = 1, 2, 3, \ldots, N-1 \), define

\[
m_i = \frac{1}{1 - \alpha_i} [S_i - \alpha_i M_{i-1}]
\]

and

\[
M_i = \min \left( S_{i+1}, \frac{1}{1 - \alpha_i} [S_i - \alpha_i M_{i-1}] \right).
\]

Theorem 3.2. Given nondecreasing nonconcave data \( (x_i, y_i) \), \( i = 0, 1, \ldots, N \), and \( \{ \alpha_i \}_{i=1}^N \) with \( 0 < \alpha_i < 1 \), \( i = 1, \ldots, N \). Then

(3.1) there exist \( \{ \alpha_j \} \)-admissible points for this data, if and only if

(3.2) the alpha algorithm can be completed with \( m_i \leq S_{i+1} \) for \( i = 1, 2, 3, \ldots, N-1 \).

Proof. Let \( \bar{x}_0 = x_0, \bar{x}_{N+1} = x_N \), and as above, let \( \bar{x}_i = x_{i-1} + \alpha_i (x_i - x_{i-1}) \) for \( i = 1, 2, \ldots, N \).
Assume that a line of slope $\tilde{S}_i$ passes through $(x_{i-1}, y_{i-1})$. The point of intersection of this line with the vertical line through $x_i$ determines a unique line through $(x_i, y_i)$ of slope

$$\tilde{S}_{i+1} = \frac{1}{1 - \alpha_i} (S_i - \tilde{S}_i).$$

Similarly, a line of slope $\tilde{S}_{i+1}$ through $(x_i, y_i)$ determines a unique line through $(x_{i-1}, y_{i-1})$ of slope

$$\tilde{S}_i = \frac{1}{\alpha_i} (S_i - (1 - \alpha_i)\tilde{S}_{i+1}).$$

Also, note that

$$\tilde{S}_{i+1} - \tilde{S}_i = \frac{1}{1 - \alpha_i} (S_i - \tilde{S}_i).$$

Thus $\tilde{S}_i \leq S_i$ if, and only if, $\tilde{S}_i \leq \tilde{S}_{i+1}$. Moreover,

$$\tilde{S}_{i+1} - \tilde{S}_i = \frac{\alpha_i}{1 - \alpha_i} (S_i - \tilde{S}_i).$$

Hence

$$S_i \leq \tilde{S}_{i+1} \quad \text{if, and only if} \quad \tilde{S}_i \leq S_i.\quad (3.7)$$

Now suppose that $S_i^*$ and $S_{i+1}^*$ are two other pairs of slopes related as above. Then

$$\tilde{S}_{i+1} - S_{i+1}^* = \frac{\alpha_i}{1 - \alpha_i} (S_i^* - \tilde{S}_i).\quad (3.9)$$

That is,

$$\tilde{S}_{i+1} \geq S_{i+1}^* \quad \text{if, and only if} \quad S_i^* \geq \tilde{S}_i.\quad (3.10)$$

It is now clear from this discussion that $m_i$ is the slope of the line segment through $(x_i, y_i)$ determined by the line of slope $M_{i-1}$ through $(x_{i-1}, y_{i-1})$, and $M_i$ is the minimum of $S_{i+1}$ and the slope of the line through $(x_i, y_i)$ determined by the line of slope $m_{i-1}$ through $(x_{i-1}, y_{i-1})$.

It is also clear that

$$M_i - m_i = \min \left( S_{i+1} - m_i, \frac{\alpha_{i-1}}{1 - \alpha_{i-1}} (M_{i-1} - m_{i-1}) \right).\quad (3.12)$$

Thus, we have

$$M_i - m_i \geq 0 \quad \text{for } i = 1, 2, \ldots, N - 1,\quad (3.13)$$

if, and only if,

$$m_i \leq S_{i+1} \quad \text{for } i = 1, 2, \ldots, N - 1.\quad (3.14)$$

Moreover, in this case we have
\[(3.15) \quad S_i \leq m_i \leq M_i \leq S_{i+1} \quad \text{for } i = 0, 1, 2, \ldots, N-1,\]
since
\[m_i - S_i = \frac{\alpha_i}{1 - \alpha_i} (S_i - M_{i-1}) \quad \text{and} \quad M_i \leq S_{i+1} \quad \text{for } i = 1, \ldots, N-1.\]

Now, let \( L \) be any piecewise linear function with knots at \( \bar{x}_i, i = 0, 1, \ldots, N, \) and such that \( L(x_j) = y_j, j = 0, 1, \ldots, N. \) Let \( v_j \) be the slope of the line segment of the graph of \( L \) passing through \( (x_j, y_j), j = 0, 1, \ldots, N. \) If we define \( M_N \) and \( m_N \) to be the slopes of the lines through \( (x_N, y_N) \) induced by the lines through \( (x_{N-1}, y_{N-1}) \) of slopes \( m_{N-1} \) and \( M_{N-1}, \) respectively, then it follows from \( (3.10) \) and \( (3.15) \) that
\[(3.16) \quad m_j \leq v_j \leq M_j \quad \text{for } j = 0, 1, \ldots, N,
\]
if, and only if
\[(3.17) \quad m_N \leq v_N \leq M_N.\]

Furthermore, if \( v_N < m_N \) or \( v_N > M_N, \) then either \( v_j > S_{j+1} \) for some \( j \) or \( v_0 < 0. \) Thus, \( L \) is either not nonconcave or not nondecreasing. On the other hand, by \( (3.15), \)
any piecewise linear function \( L \) satisfying \( (3.16) \) is nondecreasing and nonconcave.
Thus, the only nondecreasing nonconcave piecewise linear functions \( L \) as above are those satisfying \( (3.17). \) But \( (3.17) \) can be satisfied if and only if \( (3.14) \) is satisfied.
This proves the theorem.

The following corollaries are immediate consequences of Theorem 3.2 and its proof.

**Corollary 1.** (a) If \( (3.2) \) holds as above then a suitable piecewise linear function, \( L \) can be constructed by choosing any \( v_N \) between \( m_N \) and \( M_N. \)
(b) This function \( L \) is unique if, and only if \( m_N = M_N. \)

**Corollary 2.** If \( S_{i+1} - 2S_i + S_{i-1} \geq 0 \) for \( i = 1, 2, \ldots, N-1, \) then the alpha-algorithm satisfies \( (3.2) \) with \( \alpha_i = \frac{1}{2} \) for \( i = 1, \ldots, N. \)

**Proof.** Observe that
\[S_{i+1} - m_i = S_{i+1} - 2S_i + M_{i-1}.\]

This gives two cases:

(i) \( M_{i-1} = S_i. \) In this case
\[S_{i+1} - m_i = S_{i+1} - S_i \geq 0.\]

(ii) \( M_{i-1} = 2S_{i-1} - m_{i-2}. \) In this case
\[S_{i+1} - m_i = S_{i+1} - 2S_i + 2S_{i-1} - m_{i-2} = S_{i+1} - 2S_i + S_{i-1} + S_{i-1} - m_{i-2} \geq S_{i-1} - m_{i-2}.\]

This holds for \( i = 2, 3, \ldots, N-1. \) Applied recursively, this gives
\[S_{i+1} - m_i \geq \min(S_1 - m_0, S_2 - m_1).\]

But
Thus

\[ S_1 - m_0 = S_1 \geq 0 \quad \text{and} \quad S_2 - m_1 = S_2 - S_1 \geq 0. \]

Thus

\[ S_{i+1} - m_i \geq 0 \quad \text{for } i = 0, 1, \ldots, N - 1. \]

This last corollary shows that the result of Passow [9] is contained in this theory.

4. Computations and Examples. Given convex increasing data \((x_i, y_i), i = 0, 1, \ldots, N,\) and numbers \(\{\alpha_i\}_{i=1}^N\) as above, Theorem 3.2 gives a method of determining if there exist \(\{\alpha_i\}\)-admissible points for this data. Corollary 1(a) and Theorem 3.2 show how to construct a suitable piecewise linear function \(L\) if \(\{\alpha_i\}\)-admissible points exist. Once \(L\) and the desired continuity class \(m\) have been determined, a convex increasing spline in \(S_n^m\) can then be calculated by (2.2). It is clear from Theorem
that the most desirable and efficient choice of \( \{ \alpha_i \}_{i=1}^N \) is \( \alpha_i = \frac{1}{2} \) for \( i = 1, \ldots, N \). Unfortunately, this is not always possible. The problem of how to actually choose a set \( \{ \alpha_i \}_{i=1}^N \) for which there exist \( \{ \alpha_i \} \)-admissible points can be handled by Theorem 3.1 either directly or implicitly in conjunction with Theorem 3.2. That is, Theorem 3.1 certainly gives a direct method of locating some sequences \( \{ \alpha_i \}_{i=1}^N \) for which there exist \( \{ \alpha_i \} \)-admissible points. On the other hand, Theorem 3.1 shows that by choosing a sequence of sets \( A_k = \{ a_{1,k}, a_{2,k}, \ldots, a_{N,k} \} \) with \( \lim_{k \to \infty} a_{i,k} = 0 \) for \( i = 1, \ldots, N \), then for \( k \) sufficiently large all sets \( A_k = \{ a_{1,k}, \ldots, a_{N,k} \} \) will do. Thus, a repeated application of Theorem 3.2 to such a sequence of sets will eventually be successful. This latter effort is probably more desirable since Theorem 3.1 does not in general give all possible sets \( \{ \alpha_i \}_{i=1}^N \) for which there exist \( \{ \alpha_i \} \)-admissible points.

For computational examples in this paper we have used the latter approach with \( \alpha_{i,k} = 1/(k + 2) \) for \( k = 0, 1, 2, \ldots \) and all \( i \). The alpha-algorithm was run for successively larger \( k \) until condition (3.2) was satisfied. The function \( L \) was determined by choosing \( \nu_{N-1} = \frac{1}{2}(m_{N-1} + M_{N-1}) \).

The algorithm was tested on several examples and compared with Lagrange and standard cubic spline interpolation.

The following example is for \( f(x) = 1/x^2 \) on \([-2, -0.2]\) interpolating at \( x = -2, -1, -3, \) and \(-0.2\). Figure 2 shows that interpolation to these points by natural cubic splines \( S''(-2) = S''(-0.2) = 0 \) and by ordinary polynomials yields functions which are neither convex nor increasing. However, there exist \( \frac{1}{2} \)-admissible points for this data and our technique presents a convex increasing \( f \in S_3^3(-2, -3, -0.3, -0.2) \) for this case. As mentioned in [11], the data can be constructed so that no spline in \( S_3^3 \) (or even \( S_3^2 \)) can interpolate the data and preserve convexity and monotonicity. Other examples were tested and run on the computer with similar results. In one case the algorithm obtained \( \alpha = 1/10 \) and in another case \( \alpha = 1/21 \), while Lagrange and cubic spline interpolation failed.

Acknowledgement. The authors are grateful to George Bishop, who wrote some of the computer programs during the preliminary testing of our algorithms, and to Stephen Bento, who painstakingly drew the figures.

Department of Computer Science
North Carolina State University
Raleigh, North Carolina 27607

Department of Mathematics
Temple University
Philadelphia, Pennsylvania 19122

Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27607