Interpolation of medians

Tim N.T. Goodman [a] and Thomas P.Y. Yu [b]

[a] Department of Mathematics, University of Dundee, Dundee DD1 4HN, UK
[b] Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180-3590, USA

Received September 1998; revised March 1999
Communicated by T. Lyche

We study the median of a continuous function on an interval and show that for certain spaces of functions there is a unique function in the space whose medians on given intervals take given values.

Keywords: median, interpolation, polynomial, spline

AMS subject classification: 41A05, 41A10, 41A15

1. Medians

Take $n \geqslant 0$ and $n+1$ disjoint, bounded intervals $I_0, \ldots, I_n$, possible single points. In [2], motivated by applications in signal processing, it was conjectured that for any numbers $a_0, \ldots, a_n$, there is a unique polynomial of degree n whose median value on I_j equals a_j for $j = 0, \ldots, n$, and this was confirmed for $n = 2$. We shall verify this conjecture in section 2. Indeed we prove a much more general result which includes median interpolation by spline functions. Firstly, in this section, we define precisely what is meant by the median value of any continuous function on a bounded interval and show that this varies continuously both with the function and with the interval. The concept of median is well known in statistics but we include elementary proofs of propositions 1.1 and 1.3 in order to make the paper self-contained. We also mention a connection between the median and best L_1 approximation. Finally, section 3 studies the derivative of the median for the case when the function is a polynomial.

We denote Lebesgue measure on $\mathbb{R}$ by m.

Proposition 1.1. Let f be a real-valued continuous function on a bounded interval I with $m(I) > 0$. Then there is a unique value c such that

$$m\big(\{x\colon\ f(x) < c\}\big) \leqslant \frac{1}{2}m(I) \quad \text{and} \quad m\big(\{x\colon\ f(x) > c\}\big) \leqslant \frac{1}{2}m(I).$$

Proof. Let

$$c = \sup\left\{\alpha\colon\ m\big(\{x\colon\ f(x) < \alpha\}\big) \leqslant \frac{1}{2}m(I)\right\}.$$

Then there is an increasing sequence α_n converging to c such that for all n,

$$m\big(\{x\colon\ f(x) < \alpha_n\}\big) \leqslant \frac{1}{2}m(I).$$

Since the sets $\{x\colon\ f(x) < \alpha_n\}$ are increasing and have union $\{x\colon\ f(x) < c\}$,

$$m\big(\{x\colon\ f(x) < c\}\big) = \lim_{n\to\infty} m\big(\{x\colon\ f(x) < \alpha_n\}\big) \leqslant \frac{1}{2}m(I).$$

For any $\beta > c$,

$$m\big(\{x\colon\ f(x) < \beta\}\big) > \frac{1}{2}m(I),$$

and so

$$m\big(\{x\colon\ f(x) \geqslant \beta\}\big) < \frac{1}{2}m(I).$$

Taking a decreasing sequence β_n converging to c, the sets $\{x\colon\ f(x) \geqslant \beta_n\}$ are increasing with union $\{x\colon\ f(x) > c\}$ and hence

$$m\big(\{x\colon\ f(x) > c\}\big) \leqslant \frac{1}{2}m(I).$$

So c satisfies the conditions of the proposition and it remains to prove that it is unique. Suppose there are two such values $c_1 < c_2$. Since $m(\{x\colon\ f(x) < c_2\}) < m(I)$, there is a point a in I with $f(a) \geqslant c_2$. Since $m(\{x\colon\ f(x) > c_1\}) < m(I)$, there is a point b in I with $f(b) \leqslant c_1$. Take any d, $c_1 < d < c_2$. Since f is continuous, there is a point z in I with $f(z) = d$ and there is an open neighbourhood of z on which $c_1 < f(x) < c_2$. So $m(\{x\colon\ c_1 < f(x) < c_2\}) > 0$ and hence

$$m\big(\{x\colon\ f(x) < c_2\}\big) > m\big(\{f(x) \leqslant c_1\}\big) = m(I) - m\big(\{x\colon\ f(x) > c_1\}\big) \geqslant \frac{1}{2}m(I),$$

which is a contradiction. □

Remark. Note that the existence part of the above proof does not depend on the continuity of f. However, if f is discontinuous, the value of c need not be unique. For example, if $I = (-1, 1)$, $f(x) = 0$, $-1 < x < 0$, and $f(x) = 1$, $0 \leqslant x < 1$, then c can take any value between 0 and 1.

Definition. We call the value c in proposition 1.1 the *median value of f on I* and denote it by $M(f, I)$. If I comprises a single point a, we write $M(f, \{a\}) = f(a)$.

In the above definition, the interval I need not be closed and so the function f need not be integrable. However, in the case where f is integrable, we have the following characterisation of $M(f, I)$.

Proposition 1.2. If f is an integrable, continuous real-valued function on a bounded interval I, then $M(f, I)$ is the best approximation to f in the L_1 norm from the space of constant functions.

Proof. The best approximation to f in the L_1 norm from the space of constants is known to be unique [3]. Let us denote it by c. Let $m^- = m(\{x\colon\ f(x) < c\})$, $m^+ = m(\{x\colon\ f(x) > c\})$ and $Z_0 = \{x\colon\ f(x) = c\}$. From the characterisation of best L_1 approximation [5, p. 104],

$$\left| \int_I c \,\mathrm{sgn}(f - c) \right| \leqslant \int_{Z_0} |c|,$$

i.e.,

$$\left| \int_I \mathrm{sgn}(f - c) \right| \leqslant m(Z_0)$$

or

$$\left| m^+ - m^- \right| \leqslant m(Z_0). \tag{1.1}$$

Since $m^+ + m^- + m(Z_0) = m(I)$, we have

$$2m^+ + m(Z_0) - m(I) = m^+ - m^- \leqslant m(Z_0),$$

by (1.1). Hence, $m^+ \leqslant \frac{1}{2}m(I)$. Similarly from (1.1) we can deduce $m^- \leqslant \frac{1}{2}m(I)$. So, by definition, $c = M(f, I)$. □

It is known that if V is a finite dimensional subspace of a normed space E and if each f in E has a unique best approximation in V, denoted $P(f)$, then P is continuous. So from proposition 1.2, $M(f, I)$ is continuous in f on $C(I) \cap L_1(I)$ with the L_1 norm. For $f \in L_\infty(I)$, $||f||_1 \leqslant m(I)||f||_\infty$ and so $M(f, I)$ is also continuous in f on $C(I) \cap L_\infty(I)$ with the L_∞ norm. We now give an elementary proof of a result which is stronger than this last result.

Proposition 1.3. If f, g are continuous real-valued functions on a bounded interval I, then

$$\left| M(f, I) - M(g, I) \right| \leqslant \sup \left\{ \left| f(x) - g(x) \right| \colon\ x \in I \right\}.$$

Proof. Trivially the result is true if I comprises a single point. So we may assume $m(I) > 0$.

Take $\varepsilon > 0$ with $\sup\{|f(x) - g(x)|:\ x \in I\} < \varepsilon$. If $g(x) < M(f, I) - \varepsilon$, then $f(x) < M(f, I)$ and so

$$m\big(\{x:\ g(x) < M(f, I) - \varepsilon\}\big) \leqslant m\big(\{x:\ f(x) < M(f, I)\}\big) \leqslant \frac{1}{2} m(I).$$

But from the proof of proposition 1.1,

$$M(g, I) = \sup\left\{\alpha:\ m\big(\{x:\ g(x) < \alpha\}\big) \leqslant \frac{1}{2} m(I)\right\}$$

and so $M(f, I) - \varepsilon \leqslant M(g, I)$.

Similarly,

$$M(g, I) = \inf\left\{\beta:\ m\big(\{x:\ g(x) > \beta\}\big) \leqslant \frac{1}{2} m(I)\right\}$$

and since

$$m\big(\{x:\ g(x) > M(f, I) + \varepsilon\}\big) \leqslant m\big(\{x:\ f(x) > M(f, I)\}\big) \leqslant \frac{1}{2} m(I),$$

$$M(f, I) + \varepsilon \geqslant M(g, I).$$

So $|M(f, I) - M(g, I)| \leqslant \varepsilon$ and the result follows. □

For any bounded interval I and continuous function f on $\bar{I}$, clearly $M(f, I) = M(f, \bar{I})$. Thus there is no loss of generality in restricting attention to closed intervals in the following result.

Proposition 1.4. For any real-valued continuous function on an interval I, $M(f, [x, y])$ is a continuous function on $\{(x, y):\ x \leqslant y,\ [x, y] \subseteq I\}$.

We shall need the following simple technical result.

Lemma 1.1. Take a continuous, real-valued function f on $[a, b]$ and for $\varepsilon > 0$ choose $\delta > 0$ so that $|f(x) - f(y)| < \varepsilon$ whenever x, y in $[a, b]$ satisfy $|x - y| < \delta$. Then for any $p < q$ in the range of f with $m(\{x:\ p < f(x) < q\}) < \delta$, we have $q - p < \varepsilon$.

Proof. Choose $c,\ d$ in $[a, b]$ with $f(c) = p,\ f(d) = q$. First suppose $c < d$. Let

$$\widetilde{c} = \sup\left\{x \in [a, d]:\ f(x) = p\right\} \quad \text{and} \quad \widetilde{d} = \inf\left\{c \in [\widetilde{c}, b]:\ f(x) = q\right\}.$$

Then $\widetilde{c} < \widetilde{d}$, $f(\widetilde{c}) = p$, $f(\widetilde{d}) = q$ and $f((\widetilde{c}, \widetilde{d})) = (p, q)$. Then

$$\widetilde{d} - \widetilde{c} \leqslant m\big(\{x:\ p < f(x) < q\}\big) < \delta$$

and so $q - p = f(\widetilde{d}) - f(\widetilde{c}) < \varepsilon$.

Similarly we can derive the result when $c > d$. □

Proof of proposition 1.4. We first show continuity at a point (a, a), $a \in I$. For $\varepsilon > 0$, choose $\delta > 0$ so that $|f(x) - f(a)| < \varepsilon$ for x in I with $|x - a| < \delta$. Take $\alpha \leqslant \beta$ with $|\alpha - a|$, $|\beta - a| < \delta$. Clearly,

$$\inf\left\{f(x)\colon\ \alpha \leqslant x \leqslant \beta\right\} \leqslant M\big(f, [\alpha, \beta]\big) \leqslant \sup\left\{f(x)\colon\ \alpha \leqslant x \leqslant \beta\right\}$$

and so $f(a) - \varepsilon < M(f, [\alpha, \beta]) < f(a) + \varepsilon$, i.e., $|M(f, [\alpha, \beta]) - M(f, \{a\})| < \varepsilon$. Thus $M(f, [x, y])$ is continuous at $(x, y) = (a, a)$.

We now show continuity at any point (a, b) for a, b in I, $a < b$. By symmetry it is sufficient to show that $M(f, [x, y])$ is continuous in x at $x = a$ uniformly in y in a neighbourhood of $y = b$. Let J be a bounded, closed neighbourhood of $[a, b]$ in I. For $\varepsilon > 0$, choose $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for x, y in J with $|x - y| < \delta$. Choose β in J with $\beta > a$. Take α in J with $\alpha < \beta$ and $|\alpha - a| < \delta$.

First suppose that $\alpha < a$. Now

$$m\left(\left\{x \in [a, \beta]\colon\ f(x) < M\big(f, [a, \beta]\big)\right\}\right) \leqslant \frac{1}{2}(\beta - a)$$

and so

$$m\left(\left\{x \in [\alpha, \beta]\colon\ f(x) < M\big(f, [a, \beta]\big)\right\}\right) \leqslant \frac{1}{2}(\beta - a) + \delta.$$

Suppose $M(f, [\alpha, \beta]) < M(f, [a, \beta])$. Now

$$m\left(\left\{x \in [\alpha, \beta]\colon\ f(x) \leqslant M\big(f, [\alpha, \beta]\big)\right\}\right) \geqslant \frac{1}{2}(\beta - \alpha)$$

and so

$$\begin{aligned} &m\left(\left\{x \in [\alpha, \beta]\colon\ M\big(f, [\alpha, \beta]\big) < f(x) < M\big(f, [a, \beta]\big)\right\}\right) \\ &\quad \leqslant \frac{1}{2}(\beta - a) + \delta - \frac{1}{2}(\beta - \alpha) < \delta. \end{aligned}$$

Then by lemma 1.1, $M(f, [a, \beta]) - M(f, [\alpha, \beta]) < \varepsilon$.

Similarly,

$$m\left(\left\{x \in [\alpha, \beta]\colon\ f(x) > M\big(f, [a, \beta]\big)\right\}\right) \leqslant \frac{1}{2}(\beta - a) + \delta$$

and it follows that if $M(f, [\alpha, \beta]) > M(f, [a, \beta])$, then

$$M\big(f, [\alpha, \beta]\big) - M\big(f, [a, \beta]\big) < \varepsilon.$$

Similarly we obtain $|M(f, [\alpha, \beta]) - M(f, [a, \beta])| < \varepsilon$ when $\alpha > a$. Thus $M(f, [x, y])$ is continuous at $(x, y) = (a, b)$. □

2. Interpolation

Take $n \geqslant 0$ and let $I_0, \ldots, I_n$ be disjoint, bounded intervals in a closed, bounded interval I. Let $C(I)$ denote all continuous real-valued functions on I and let V_n be a vector subspace of $C(I)$ of dimension $n+1$. We shall assume that if any function f in V_n has a zero in all the intervals $I_0, \dots, I_n$, then $f = 0$.

We can now state our main result.

Theorem 2.1. In the above situation, for any real numbers $a_0, \dots, a_n$, there is a unique function f in V_n with

$$M(f, I_j) = a_j, \quad j = 0, \dots, n.$$

Now define $F: V_n \to \mathbb{R}^{n+1}$ by

$$F(f) = \big(M(f, I_0), \dots, M(f, I_n)\big). \tag{2.1}$$

Then theorem 2.1 is equivalent to showing that F is a bijection.

Lemma 2.1. Suppose $I = (a, b)$ and f, g in $C(I)$ satisfy $M(f, I) = M(g, I)$. Then there is a point α in (a, b) with $f(\alpha) = g(\alpha)$.

Proof. We suppose that there is no such point and reach a contradiction. Without loss of generality we may suppose $f(x) > g(x)$ for all x in (a, b). Let $M = M(f, I) = M(g, I)$ and

$$U = \{x \in I\colon\ f(x) > M\}, \qquad E = \{x \in I\colon\ g(x) \geqslant M\}.$$

Then $E \subseteq U$. By definition of $M(f, I)$, $m(U) \leqslant \frac{1}{2}(b-a)$ and by definition of $M(g, I)$, $m(E) \geqslant \frac{1}{2}(b-a)$. Thus $m(E) = m(U) = \frac{1}{2}(b-a)$. Now E is closed (in (a, b)) and U is open, $U \neq (a, b)$. So $E = U$ would contradict (a, b) being connected. Thus $U \setminus E$ is non-empty and open and so $m(U \setminus E) > 0$, which contradicts $m(E) = m(U)$. $\square$

Lemma 2.2. For any norm on V_n and F as in (2.1), $F(f) \to \infty$ as $||f|| \to \infty$.

Proof. Suppose that (f_k) is a sequence in V_n with $||f_k|| \to \infty$. We shall assume $F(f_k)$ is bounded and reach a contradiction. Write

$$g_k = \frac{f_k}{||f_k||}, \quad k = 1, 2, \dots.$$

Since the closed unit ball in V_n is compact, we can choose a subsequence (h_k) with $h_k \to h$ in V_n. Noting that for any $\lambda \in \mathbb{R}$, $f \in C(I)$, $F(\lambda f) = \lambda F(f)$, we have $F(h_k) \to 0$. By proposition 1.3, $F(h) = 0$. Thus h has a zero in all the intervals $I_0, \dots, I_n$ and, by assumption, $h = 0$. This contradicts $||h|| = 1$. $\square$

We shall need the following theorem, which follows from a result in [1, p. 93], which is a special case of a result in [4, p. 136].

Theorem A. Take $n \geqslant 1$ and a continuous, injective function G from $\mathbb{R}^n$ to $\mathbb{R}^n$. If $G(x) \to \infty$ as $x \to \infty$, then G maps *onto* $\mathbb{R}^n$.

Proof of theorem 2.1. We first show that F as in (2.1) is injective. Suppose that for f, g in V_n, $F(f) = F(g)$. Then by lemma 2.1, there are points $\alpha_j \in I_j$, $j = 0, \ldots, n$, such that $f(\alpha_j) = g(\alpha_j)$, $j = 0, \ldots, n$. Thus $f - g$ has zeros in all the intervals $I_0, \ldots, I_n$ and, by our assumption on V_n, $f = g$.

It follows from proposition 1.3, lemma 2.2 and theorem A, that F is bijective. □

By putting V_n equal to the space of polynomials of degree n, we have the following immediate consequence of theorem 2.1.

Corollary 2.1. Take $n \geqslant 0$ and let $I_0, \ldots, I_n$ be disjoint, bounded intervals. Then for any real numbers $a_0, \ldots, a_n$, there is a unique real polynomial p of degree n with

$$M(p, I_j) = a_j, \quad j = 0, \ldots, n.$$

Clearly if all the intervals $I_0, \ldots, I_n$ comprise single points, then this result reduces to well-known Lagrange interpolation. Next we give a result which partially extends the Schoenberg–Whitney theorem [6] for interpolation by spline functions. For intervals I, J we write $I < J$ if $x < y$ for all $x \in I$, $y \in J$.

Corollary 2.2. Take $n \geqslant 0$, $m \geqslant 1$ and a sequence $t_0 \leqslant \cdots \leqslant t_{n+m+1}$ with $t_i < t_{i+m}$, $i = 1, \ldots, n$. Let $I_0 < \cdots < I_n$ be intervals such that for $i = 0, \ldots, n$, I_i lies in (t_i, t_{i+m+1}). Then for any real numbers $a_0, \ldots, a_n$, there is a unique spline function f of degree m with support in $[t_0, t_{n+m+1}]$ and knots (with multiplicity) in $(t_0, \ldots, t_{n+m+1})$ with

$$M(f, I_j) = a_j, \quad j = 0, \ldots, n.$$

Proof. Let V_n denote the $(n+1)$-dimensional space of all spline functions of degree m with support in $[t_0, t_{n+m+1}]$ and knots in (t_0, t_{n+m+1}). By our assumption on the knots, the elements of V_n are continuous. If f in V_n has a zero in all the intervals $I_0, \ldots, I_n$, then by the Schoenberg–Whitney theorem [6], $f = 0$. Thus the result follows from theorem 2.1. □

3. Derivatives

In proposition 1.3 we showed that the median $M(f, [a, b])$ is continuous in f in the L_∞ norm. A natural next step is to study the differentiability of this median. This is done in this section for the case when f is a polynomial.

For $n \geqslant 1$, let Π_n denote the space of real polynomials of degree n. For any function $h : \Pi_n \to \mathbb{R}$ and any q in Π_n, we define the directional derivative

$$D_q h(p) = \lim_{\varepsilon \to 0_+} \frac{h(p + \varepsilon q) - h(p)}{\varepsilon},$$

if this limit exists. Now fix an interval $[a,b]$ and define $m:\Pi_n \to \mathbb{R}$ by $m(p) = M(p,[a,b])$. We shall calculate directional derivatives of m. For clarity, we deal first with the simplest case in which m is differentiable.

Proposition 3.1. Take $p_0 \in \Pi_n$ such that the only roots in $[a,b]$ of the equation $p_0(x) = m(p_0)$ are simple roots $\alpha_1 < \cdots < \alpha_r$ in (a,b). Then for any q in Π_n,

$$D_q m(p_0) = \frac{\sum_{j=1}^{r} (q(\alpha_j)/|p_0'(\alpha_j)|)}{\sum_{j=1}^{r} (1/|p_0'(\alpha_j)|)}. \tag{3.1}$$

Proof. For p in a small enough neighbourhood of p_0, the equation $p(x) = m(p)$ will have simple roots $\alpha_1(p), \ldots, \alpha_r(p)$ in (a,b). Then for $j = 1, \ldots, r$,

$$D_q m(p_0) = D_q p\big(\alpha_j(p)\big)|_{p=p_0} = q(\alpha_j) + p_0'(\alpha_j) D_q \alpha_j(p_0). \tag{3.2}$$

Also $\sum_{j=1}^{r} (-1)^j \alpha_j(p)$ is constant and so

$$0 = \sum_{j=1}^{r} (-1)^j D_q \alpha_j(p_0) = \sum_{j=1}^{r} \frac{(-1)^j}{p_0'(\alpha_j)} \big(D_q m(p_0) - q(\alpha_j)\big)$$

by (3.2). Now $(-1)^j p_0'(\alpha_j)$ has the same sign for $j = 1, \ldots, r$, and so

$$0 = D_q m(p_0) \sum_{j=1}^{r} \frac{1}{|p_0'(\alpha_j)|} - \sum_{j=1}^{r} \frac{q(\alpha_j)}{|p_0'(\alpha_j)|},$$

which gives (3.1). □

At a general polynomial p_0 in Π_n, the function m is not differentiable and so we consider a manifold $S \subset \Pi_n$ at p_0 such that m restricted to S is differentiable.

Theorem 3.1. Take $p_0 \in \Pi_n$ such that the equation $p_0(x) = m(p_0)$ has roots $\alpha_1 < \cdots < \alpha_r$ in (a,b) with multiplicities $m_j \geqslant 1$, $j = 1, \ldots, r$, and roots at a and b of multiplicities $\alpha, \beta \geqslant 0$, respectively. Let $S = S(p_0,[a,b]) \subset \Pi_n$ denote the manifold of all polynomials p in Π_n such that the roots of $p(x) = m(p)$ have the same multiplicities at a, b and in (a,b) with the same order as for $p = p_0$. Then the elements of the tangent hyperplane to S at p_0 are of the form $p_0 + q$, where q satisfies the following $\alpha + \beta + \sum_{j=1}^{r}(m_j - 1)$ conditions: for some number $x(q)$,

$$q(\alpha_j) = x(q) \quad \text{if } m_j \geqslant 2, \tag{3.3}$$

$$q^{(i)}(\alpha_j) = 0, \qquad i = 1, \ldots, m_j - 2, \tag{3.4}$$

$$q(a) = x(q) \quad \text{if } \alpha \geqslant 1, \tag{3.5}$$

$$q(b) = x(q) \quad \text{if } \beta \geqslant 1, \tag{3.6}$$

$$q^{(i)}(a) = 0, \qquad i = 1, \ldots, \alpha - 1, \tag{3.7}$$

$$q^{(i)}(b) = 0, \qquad i = 1, \ldots, \beta - 1, \tag{3.8}$$

$$\sum_{m_j \text{ odd}} \frac{q^{(m_j-1)}(\alpha_j)}{|p^{(m_j)}(\alpha_j)|} = x(q) \sum_{m_j=1} \frac{1}{|p'(\alpha_j)|}, \tag{3.9}$$

where the final summation is zero if $m_j \geqslant 2$, $j = 1, \ldots, r$. Moreover, for $p_0 + q$ in this tangent hyperplane,

$$D_q m(p_0) = x(q). \tag{3.10}$$

Proof. Suppose that for λ in a neighbourhood of 0, p_λ is a curve in S, differentiable in λ, and coinciding with p_0 when $\lambda = 0$. Suppose that $p_\lambda(x) = m(p_\lambda)$ has roots $\alpha_1(\lambda), \ldots, \alpha_r(\lambda)$ in (a, b) with multiplicities $m_1, \ldots, m_r$. Thus for $j = 1, \ldots, r$, $\alpha_j(0) = \alpha_j$ and

$$p_\lambda\big(\alpha_j(\lambda)\big) = m(p_\lambda). \tag{3.11}$$

We write $q = (\partial/\partial\lambda) p_\lambda|_{\lambda=0}$.

Now from (3.11), for $j = 1, \ldots, r$,

$$x(q) := \frac{\partial}{\partial\lambda} m(p_\lambda)|_{\lambda=0} = q(\alpha_j) + p_0'(\alpha_j)\alpha_j'(0) \tag{3.12}$$

(as in (3.2)). Next suppose that $m_j \geqslant 2$. Then $p_0'(\alpha_j) = 0$ and (3.12) gives (3.3). Since $p_\lambda'(\alpha_j(\lambda)) = 0$,

$$\frac{\partial}{\partial\lambda} p_\lambda'\big(\alpha_j(\lambda)\big) + p_\lambda''\big(\alpha_j(\lambda)\big)\alpha_j'(\lambda) = 0.$$

Continuing to differentiate with respect to λ and putting $\lambda = 0$ gives (3.4) and

$$q^{(m_j-1)}(\alpha_j) + p_0^{(m_j)}(\alpha_j)\alpha_j'(0) = 0. \tag{3.13}$$

Now let $\beta_1(\lambda), \ldots, \beta_s(\lambda)$ comprise $\{\alpha_j(\lambda)\colon\ m_j \text{ odd}\}$, which must be non-empty. Then $\sum_{j=1}^{s} (-1)^j \beta_j(\lambda)$ is constant and so from (3.12) and (3.13),

$$0 = \sum_{m_j=1} \frac{q(\alpha_j) - x(q)}{|p_0'(\alpha_j)|} + \sum_{m_j \text{ odd} \geqslant 3} \frac{q^{(m_j-1)}(\alpha_j)}{|p_0^{(m_j)}(\alpha_j)|},$$

which gives (3.9).

If $\alpha \geqslant 1$, we have

$$p_\lambda(a) = m(p_\lambda), \qquad p_\lambda^{(i)}(a) = 0, \quad i = 1, \ldots, \alpha - 1.$$

Then differentiating with respect to λ and putting $\lambda = 0$ gives (3.5) and (3.7). Similarly we derive (3.6) and (3.8).

Finally, we note that

$$D_q m(p_0) = \frac{\partial}{\partial \lambda} m(p_\lambda)|_{\lambda=0}$$

and (3.10) follows from the definition of $x(q)$ in (3.12). □

Acknowledgements

Much of this work was undertaken by the first-named author during a visit to the National University of Singapore supported by the Wavelets Strategic Research Programme. We are grateful to the help of people there, in particular Prof. Lee Seng Luan and Dr. Wayne Lawton. The second-named author would like to thank Joyce McLaughlin and Stephen Weingram for discussions on topics related to the content of this paper. We are also grateful to the referee who suggested proposition 1.2 and the remark thereafter.

References

[1] D. Braess, *Nonlinear Approximation Theory* (Springer, Berlin, 1986).
[2] D.L. Donoho and T.P.-Y. Yu, Nonlinear “wavelet transforms” based on median-interpolation, Preprint.
[3] I. Jackson, A general class of problems in approximation, Amer. J. Math. 46 (1924) 215–234.
[4] J.M.O. Ortega amd W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Severable Variables* (Academic Press, New York, 1970).
[5] J.R. Rice, *The Approximation of Functions* (Addison-Wesley, Reading, MA, 1964).
[6] I.J. Schoenberg and A. Whitney, On Pólya frequency functions III: The positivity of translation determinants with application to the interpolation problem by spline curves, Trans. Amer. Math. Soc. 74 (1953) 246–259.