Total Positivity of the Spline Collocation Matrix

CARL de BOOR

0. Definitions. For given positive integers n and k, and a given real nondecreasing sequence $\mathbf{t} := (t_i)_1^{n+k}$ with

$$t_i < t_{i+k}\,, \quad \text{all } i,$$

denote by $\mathbf{S}_{k,\mathbf{t}}$ the linear span of the n normalized B-splines $N_{1,k}\,, \cdots, N_{n,k}\,,$ given by the rule that, for each t,

$$N_{i,k}(t) := ([t_{i+1}\,, \cdots, t_{i+k}] - [t_i\,, \cdots, t_{i+k-1}])(\cdot - t)_+^{k-1}.$$

Here,

$$[\rho_0\,, \cdots, \rho_r]f$$

denotes the rth divided difference of the function f at the points $\rho_0\,, \cdots, \rho_r$. The elements of $\mathbf{S}_{k,\mathbf{t}}$ are called *polynomial splines of order k with knot sequence* $\mathbf{t}$.

Let $\boldsymbol{\tau} := (\tau_i)_1^n$ be a strictly increasing real sequence. As is essentially shown by Schoenberg and Whitney [7], *there exists, for given f, exactly one $s \in \mathbf{S}_{k,\mathbf{t}}$ such that*

$$s(\tau_i) = f(\tau_i), \qquad i = 1, \cdots, n,$$

if and only if

$$N_{i,k}(\tau_i) \neq 0, \qquad i = 1, \cdots, n,$$

i.e., if

$$t_i < \tau_i < t_{i+k}\,, \qquad i = 1, \cdots, n.$$

Their proof is an appeal to much more general results in the same paper concerning the positivity of translation determinants. This theorem has been generalized by Karlin and Ziegler [3] to include the case of repeated or osculatory collocation, as stated in Theorem 1 below. The proof given here relies only on Rolle's Theorem and on the facts that

$$(1) \qquad (d/dt) \sum_i \alpha_i N_{i,k} = (k-1) \sum_i \frac{\alpha_i - \alpha_{i-1}}{t_{i+k-1} - t_i} N_{i,k-1}\,;$$

Indiana University Mathematics Journal, ©, Vol. 25, No. 6 (1976)

(2) $N_{i,k}$ is positive on (t_i, t_{i+k}) and zero off $[t_i, t_{i+k}]$;
(3) if $\alpha_n \neq 0$, then $\alpha_n \sum_i \alpha_i N_{i,k} > 0$ on $(t_{n+k} - \epsilon, t_{n+k})$ for some $\epsilon > 0$;
(4) if $t_i < t_{i+1}$, then every straight line on $[t_i, t_{i+1}]$ can be written as $\alpha_{i-1}N_{i-1,2} + \alpha_i N_{i,2}$ for appropriate α_{i-1}, α_i;

all of which follow directly from the definition of B-splines.

Karlin [1; Ch. 10, Theorem 4.1] proved that the $n \times n$ matrix $(N_{j,k}(\tau_i))$ is *totally positive, i.e.*, has all minors nonnegative. His proof applies to Chebyshev splines and not just polynomial splines but uses machinery which seems readily available only after reading [1]. In Theorem 2 below, we give a proof of the total positivity of $(N_{j,k}(\tau_i))$ based on the proof of Theorem 1 and on Karlin's elegant proof of the Fekete Lemma [1, Ch. 2, Theorem 3.2] making use of the determinant identity (0.20) of [1]. As a byproduct, we obtain in Theorem 2 the curious extension of the Schoenberg-Whitney Theorem that even for a *subsequence* $(p_1, \cdots, p_r)$ of $(1, \cdots, n)$ and for strictly increasing $\boldsymbol{\tau}$, $(N_{p_j,k}(\tau_i))_1^r$ is invertible iff $N_{p_i,k}(\tau_i) \neq 0$, all i.

1. Collocation. We consider collocation at the points of a given nondecreasing sequence $\boldsymbol{\tau} := (\tau_i)_1^n$, repeated points indicating repeated or osculatory interpolation in the usual way. Precisely, define

$$\lambda_i := [\tau_i]D^r \quad \text{with} \quad r := \max \{j \mid \tau_{i-j} = \tau_i\}.$$

Then we say that *s agrees with f at (the points of)* $\boldsymbol{\tau}$ provided

$$\lambda_i s = \lambda_i f, \qquad i = 1, \cdots, n.$$

If $\mathbf{t} = (t_i)_1^{n+k}$ with $t_i < t_{i+k}$, all i, and $s \in \mathbf{S}_{k,\mathbf{t}}$, then $[\tau]s^{(r)}$ makes sense as long as $\tau \notin \operatorname{ran} \mathbf{t}$. If, on the other hand, $\tau = t_j$ for some j, then $[\tau]D^r$ is defined on $\mathbf{S}_{k,\mathbf{t}}$ only if the multiplicity of t_j does not exceed $k - 1 - r$. Hence we must assume

$$\tau_{i+1} = \cdots = \tau_{i+r} = t_{j+1} = \cdots = t_{j+s} \text{ implies } r + s \leqq k, \tag{5}$$

if we want to consider collocation at $\boldsymbol{\tau}$ by elements of $\mathbf{S}_{k,\mathbf{t}}$.

The n-term sequence $(\lambda_i)_1^n$ is linearly independent while the n-term sequence $(N_{i,k})_1^n$ generates $\mathbf{S}_{k,\mathbf{t}}$, by definition. Existence and uniqueness of $s \in \mathbf{S}_{k,\mathbf{t}}$ for given f for which

$$\lambda_i s = \lambda_i f, \qquad i = 1, \cdots, n,$$

is therefore equivalent to the invertibility of the *Gramian*

$$A := (\lambda_i N_{j,k})_{i,j=1}^n$$

or, equivalent to the statement that $s \in \mathbf{S}_{k,\mathbf{t}}$ vanishes at $\tau_1, \cdots, \tau_n$ (repeats indicating multiplicity in the usual way) iff s vanishes identically. We will use each of these equivalent formulations in the arguments to follow without belaboring their equivalence any further.

Theorem 1. *Let $(N_{i,k})_1^n$ be the sequence of B-splines of order k for the nondecreasing knot sequence $\mathbf{t} = (t_i)_1^{n+k}$ with $t_i < t_{i+k}$, all i, and, for $i = 1, \cdots, n$,* *let*

$$\lambda_i := [\tau_i]D^r \text{ with } r := \max\{j \mid \tau_{i-j} = \tau_i\}$$

for some given nondecreasing sequence $\boldsymbol{\tau} = (\tau_i)_1^n$ *satisfying* (5). *Then*

$$A := (\lambda_i N_{j,k}) \text{ is invertible iff } N_{i,k}(\tau_i) \neq 0, i = 1, \cdots, n,$$

i.e., $\tau_i \in (t_i, t_{i+k})$ *for all* i.

Proof. If, for some r, $N_{r,k}(\tau_r) = 0$, then $\tau_r \notin (t_r, t_{r+k})$ by (2), hence, without loss, $\tau_r \leqq t_r$. But then $\lambda_i N_{j,k} = 0$ for $1 \leqq i \leqq r \leqq j \leqq n$, showing columns $r, \cdots, n$ of A to be linearly dependent; hence A is then not invertible. Assume now, conversely, that $N_{r,k}(\tau_r) \neq 0$ for all r, *i.e.*,

$$\tau_r \in (t_r, t_{r+k}), \qquad r = 1, \cdots, n.$$

Then, *for* $n = 1$ *and all* k, the 1×1 matrix A is trivially invertible. Also, *for all* n *and* $k = 1$, A is invertible as then $A = 1$. As for the remaining *general case* $n > 1$ *and* $k > 1$, we assume without loss that both the first superdiagonal and the first subdiagonal of A is nonzero, *i.e.*,

(6) $\qquad t_{i+1} < \tau_i, i = 1, \cdots, n-1, \quad \text{and} \quad \tau_i < t_{i+k-1}, i = 2, \cdots, n.$

For, if, *e.g.*, $t_{r+1} \geqq \tau_r$ for some r then $\tau_r < \tau_{r+1}$ and $\lambda_i N_{j,k} = 0$ for $1 \leqq i \leqq r < j \leqq n$, and the invertibility of A is equivalent to the invertibility of the two *smaller* matrices $(\lambda_i N_{j,k})_{i,j=1}^r$ and $(\lambda_i N_{j,k})_{i,j=r+1}^n$ *of the same form*. We may further assume without loss that $k = 2$. For, if $k > 2$ and A is not invertible, then we can find a nonzero $\boldsymbol{\alpha} \in \mathbf{R}^n$ so that $A\boldsymbol{\alpha} = \mathbf{0}$. But then, the function

$$f := \sum_{j=1}^n \alpha_j N_{j,k}$$

is in $C^{(1)}$ and vanishes at the $n + 2$ points

$$t_1 =: \tau_0, \tau_1, \cdots, \tau_n, \tau_{n+1} := t_{n+k}$$

(repeats indicating multiplicity in the usual way). Its derivative is therefore by (1), of the form

$$f' = \sum_{j=1}^{n+1} \beta_j N_{j,k-1}$$

for some nonzero $\boldsymbol{\beta} \in \mathbf{R}^{n+1}$, and vanishes at certain $n + 1$ points $(\tau'_i)_1^{n+1}$ with

$$t_{i+1} \leqq \tau_i \leqq \tau'_{i+1} \leqq \tau_{i+1} \leqq t_{i+k}, \qquad i = 0, \cdots, n,$$

(the inner two inequalities being strict in case $\tau_i < \tau_{i+1}$). Consequently, the matrix $(\lambda_i' N_{j,k-1})_{i,j=1}^{n+1}$ involving splines of order $k - 1$ is then not invertible even though the λ_i' derive from $\boldsymbol{\tau}'$ satisfying

$$\tau'_i \in (t_i, t_{i+k-1}), \qquad i = 1, \cdots, n+1$$

and

$$\tau'_{i+1} = \cdots = \tau'_{i+r} = t_{j+1} = \cdots = t_{j+s} \quad \text{implies} \quad r + s \leqq k - 1.$$

This leaves the case $k = 2$. But then $t_{i+1} = t_{i+k-1}$, hence (6) allows only either $n = 1$, already settled above, or $n = 2$ and

$$t_2 < \tau_1 \leqq \tau_2 < t_3 ,$$

in which case A is trivially invertible by (4). □

Corollary. *The determinant of the Gramian matrix $(\lambda_i N_{j,k})$ of the preceding theorem is nonnegative.*

Proof. Assume initially that $\boldsymbol{\tau}$ is strictly increasing. By the theorem, it is sufficient to consider the case $\tau_i \in (t_i, t_{i+k})$, all i. Since then $\det (N_{j,k}(\tau_i))_1^n$ is trivially positive for $n = 1$, it is sufficient to prove that

$$\alpha_n := \det (N_{j,k}(\tau_i))_1^{n-1}/\det (N_{j,k}(\tau_i))_1^n > 0 \quad \text{for} \quad n > 1.$$

But, in this case, α_n is the n-th B-spline coefficient of the unique $f \in \mathbf{S}_{k,\mathbf{t}}$ which satisfies

$$f(\tau_i) = \delta_{in} , \quad \text{all } i,$$

while, by (3), there exists $\tau \in (\tau_n, t_{n+k})$ so that

$$\operatorname{sign} f(\tau) = \operatorname{sign} \alpha_n .$$

This allows the desired conclusion that sign $\alpha_n = 1$, since otherwise $f(\tau) < 0$ while $f(\tau_n) = 1 > 0$, hence f would vanish at some point $\tau_n' \in (t_n, t_{n+k})$ in addition to its zeroes $\tau_i \in (t_i, t_{i+k})$ for $i = 1, \cdots, n-1$, hence would have to vanish identically by the theorem, a contradiction.

Consider now the general case with $\boldsymbol{\tau}$ not necessarily strictly increasing. If, *e.g.*,

$$\tau_{r-s-1} < \tau_{r-s} = \cdots = \tau_r < \tau_{r+1}$$

then, for given $\epsilon > 0$, we can pick $\tau'_{r-s} < \cdots < \tau'_r$ all in $(\tau_{r-s-1}, \tau_{r+1})$ so that replacing $\lambda_{r-p} N_{j,k} = [\tau_r] N_{j,k}^{(s-p)}$ by $(s-p)!$ times the $(s-p)$th divided difference $[\tau'_{r-s}, \cdots, \tau'_{r-p}] N_{j,k}$ $(p = 0, \cdots, s; j = 1, \cdots, n)$ produces a new matrix whose determinant differs from that of $(\lambda_i N_{j,k})$ by no more than ϵ. But, after all coincidences in the τ_i's have been dealt with in this way, we obtain a matrix with a nonnegative determinant since it is of the form $L \cdot (N_{j,k}(\tau'_i))$ for some lower triangular matrix L with positive diagonal entries and for some strictly increasing sequence $\boldsymbol{\tau}'$. This shows $\det (\lambda_i N_{j,k})$ to be within ϵ of being nonnegative for arbitrary ϵ. □

2. Total positivity of the collocation matrix. The nonnegativity of the determinant of the collocation matrix $(\lambda_i N_{j,k})$, proved in the corollary to Theorem 1, is an indication of the remarkable and important fact that $(N_{j,k}(\tau_i))$ is *totally positive*, *i.e.*, all its minors are nonnegative. This fact can be inferred from the following generalization of Theorem 1.

Theorem 2. *Let $(N_{i,k})_1^n$ be the sequence of B-splines of order k for the nondecreasing knot sequence $\mathbf{t} := (t_i)_1^{n+k}$ with $t_i < t_{i+k}$, all i, and, for $i = 1, \cdots, n$,* *let*

$$\lambda_i := [\tau_i]D^r \quad \textit{with} \quad r := \max\,\{j \mid \tau_{i-j} = \tau_i\}$$

for some nondecreasing sequence $\boldsymbol{\tau} := (\tau_i)_1^{\,n}$ *satisfying*

$$\tau_{i+1} = \cdots = \tau_{i+r} = t_{j+1} = \cdots = t_{j+s} \quad \textit{implies} \quad r + s \leqq k,$$

and set

$$U := (u_{ij}) := (\lambda_i N_{j,k})_1^{\,n}.$$

If $\mathbf{o} := (o_i)_1^{\,m}$ *is a subsequence of* $(1, \cdots, n)$ *such that (with* $o_0 = 0$)

$$o_{i-1} < o_i - 1 \quad \textit{implies} \quad \tau_{o_i - 1} < \tau_{o_i},$$

then, for every subsequence $\mathbf{p} := (p_i)_1^{\,m}$ *of* $(1, \cdots, n)$,

$$\det U\begin{pmatrix} o_1, & \cdots, & o_m \\ p_1, & \cdots, & p_m \end{pmatrix} \geqq 0$$

with equality iff, for some i, $N_{p_i,k}(\tau_{o_i}) = 0$.

Proof. The restriction on the selection $\mathbf{o}$ of rows for the submatrix

$$U\begin{pmatrix} o_1, & \cdots, & o_m \\ p_1, & \cdots, & p_m \end{pmatrix}$$

is designed to exclude the possibility that a row involving some derivative is included in the submatrix without also including all rows involving lower derivatives at the same point. This exclusion is essential, as the theorem is not true without it. In particular, U need not be totally positive in case of coincidences among the τ_i's.

$N_{j,k}(\tau_i)$ vanishes if supp $(\lambda_i) \cap$ supp $(N_{j,k}) = \varnothing$. But if this happens, then the underlying geometry implies that

$$\text{either } U\begin{pmatrix} 1, & \cdots, & i \\ j, & \cdots, & n \end{pmatrix} = 0 \quad \text{or else} \quad U\begin{pmatrix} i, & \cdots, & n \\ 1, & \cdots, & j \end{pmatrix} = 0.$$

Hence, the submatrix

$$U\begin{pmatrix} o_r, & \cdots, & o_s \\ p_r, & \cdots, & p_s \end{pmatrix}$$

is not invertible unless supp $(\lambda_{o_i}) \cap$ supp $(N_{p_i,k}) \neq \varnothing$ for $i = r, \cdots, s$. A submatrix satisfying this latter condition will be called "good" for the duration of this proof.

We assume that $\mathbf{o} = (1, \cdots, m)$ (by going over to a new nondecreasing sequence $\hat{\boldsymbol{\tau}}$ which starts $\tau_{o_1}, \cdots, \tau_{o_m}, \cdots$, if necessary). With this, we proceed to prove that

$$\text{(7)} \qquad \text{if} \quad U\begin{pmatrix} 1, & \cdots, & m \\ p_1, & \cdots, & p_m \end{pmatrix} \text{ is "good", then} \quad \det U\begin{pmatrix} 1, & \cdots, & m \\ p_1, & \cdots, & p_m \end{pmatrix} > 0$$

by induction on m and on

$$\operatorname{var} \mathbf{p} := p_m - p_1 ,$$

starting with the case $\operatorname{var} \mathbf{p} = m - 1$. In this case, there are no gaps in $\mathbf{p}$, hence Theorem 1 and its corollary imply (7). Note that this case includes the case $m = 1$.

Let now $m > 1$, and let $\operatorname{var} \mathbf{p} \geqq m$. For a given $m \times (m + 1)$ matrix $A = (a_{ij})$, the matrix

$$\tilde{A} := \begin{bmatrix} a_{1,1} & \cdots & a_{1,m+1} & a_{1,2} & \cdots & a_{1,m-1} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,m+1} & a_{m,2} & \cdots & a_{m,m-1} \\ a_{1,1} & \cdots & a_{1,m+1} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m-1,1} & \cdots & a_{m-1,m+1} & 0 & \cdots & 0 \end{bmatrix}$$

is not invertible since all $(m + 1)$-minors in the first $m + 1$ columns are zero. Hence, expanding by m-minors of the first m rows, we get

$$\begin{aligned} 0 = \det \tilde{A} &= \det A\begin{pmatrix} 1 & , \cdots , & m \\ m, m+1, 2, & \cdots & , m-1 \end{pmatrix} \det A\begin{pmatrix} 1, \cdots, m-1 \\ 1, \cdots, m-1 \end{pmatrix} \\ &+ (-1)^{m-1} \det A\begin{pmatrix} 1 & , \cdots , & m \\ 1, m+1, 2, & \cdots & , m-1 \end{pmatrix} \det A\begin{pmatrix} 1, \cdots, m-1 \\ 2, \cdots, m \end{pmatrix} \\ &+ (-1)^{m} \det A\begin{pmatrix} 1 & , \cdots , & m \\ 1, m, 2, & \cdots & , m-1 \end{pmatrix} \det A\begin{pmatrix} 1, \cdots, & m-1 \\ 2, \cdots, m-1, & m+1 \end{pmatrix}, \end{aligned}$$

all other terms in the expansion being trivially zero. On applying this determinant identity to the $m \times (m + 1)$ matrix

$$A = U\begin{pmatrix} 1, \cdots, m \\ p_1, \cdots, p_{m+1} \end{pmatrix},$$

we obtain (after appropriate interchanges) the fact that

$$\begin{aligned} -\det U\begin{pmatrix} 1, \cdots, m \\ p_2, \cdots, p_{m+1} \end{pmatrix} \det U\begin{pmatrix} 1, \cdots, m-1 \\ p_1, \cdots, p_{m-1} \end{pmatrix} + \det U\begin{pmatrix} 1, \cdots, & m \\ p_1, \cdots, p_{m-1}, & p_{m+1} \end{pmatrix} \\ \cdot \det U\begin{pmatrix} 1, \cdots, m-1 \\ p_2, \cdots, p_m \end{pmatrix} = \det U\begin{pmatrix} 1, \cdots, m \\ p_1, \cdots, p_m \end{pmatrix} \det U\begin{pmatrix} 1, \cdots, & m-1 \\ p_2, \cdots, p_{m-1}, & p_{m+1} \end{pmatrix}, \end{aligned}$$

with p_{m+1} an arbitrary index. Now pick p_{m+1} different from $p_1, \cdots, p_m$ but lying between p_1 and p_m, and let $\mathbf{q} := (q_i)_1^{m+1}$ be the strictly increasing sequence obtained from $p_1, \cdots, p_m$ and p_{m+1}. Then

$$\begin{aligned}(8)\qquad \det U\binom{1, \cdots, m}{q_2, \cdots, q_{m+1}} \det U\binom{1, \cdots, m-1}{p_1, \cdots, p_{m-1}} \\ + \det U\binom{1, \cdots, m}{q_1, \cdots, q_m} \det U\binom{1, \cdots, m-1}{p_2, \cdots, p_m} \\ = \det U\binom{1, \cdots, m}{p_1, \cdots, p_m} \det U\binom{1, \cdots, m-1}{q_2, \cdots, q_m}\end{aligned}$$

since it takes as many interchanges to order $p_1, \cdots, p_{m-1}, p_{m+1}$ as it does to order $p_2, \cdots, p_{m-1}, p_{m+1}$, but one less than does the ordering of $p_2, \cdots, p_m, p_{m+1}$.

If now, for some r,

$$\operatorname{supp}(\lambda_r) \cap \operatorname{supp}(N_{p_{r+1},k}) = \varnothing \quad \text{or} \quad \operatorname{supp}(\lambda_{r+1}) \cap \operatorname{supp}(N_{p_r,k}) = \varnothing,$$

then, much as in the proof of Theorem 1, $\tau_r < \tau_{r+1}$ and

$$\det U\binom{1, \cdots, m}{p_1, \cdots, p_m} = \det U\binom{1, \cdots, r}{p_1, \cdots, p_r} \det U\binom{r+1, \cdots, m}{p_{r+1}, \cdots, p_m} > 0,$$

since both factors on the right are "good", hence positive by induction hypothesis on m. Otherwise

$$\operatorname{supp}(\lambda_r) \cap \operatorname{supp}(N_{p_{r+1},k}) \neq \varnothing, \operatorname{supp}(\lambda_{r+1}) \cap \operatorname{supp}(N_{p_r,k}) \neq \varnothing, \qquad r = 1, \cdots, m-1,$$

hence each of the six submatrices in (8) is "good". This implies by induction hypothesis on m that each of the three $(m-1) \times (m-1)$ submatrices in (8) has positive determinant. Further, since also $q_{m+1} - q_2$ and $q_m - q_1$ are both less than var $\mathbf{p} = p_m - p_1$, we know by induction hypothesis that both $m \times m$ matrices on the left side of (8) have positive determinant. But then, also

$$\det U\binom{1, \cdots, m}{p_1, \cdots, p_m} > 0. \qquad \square$$

Corollary. *With $(N_{i,k})_1^n$ the sequence of B-splines of order k for some knot sequence $\mathbf{t} = (t_i)_1^{n+k}$, nondecreasing with $t_i < t_{i+k}$, all i, and $\boldsymbol{\tau} = (\tau_i)_1^n$ nondecreasing, the Gramian matrix $(N_{j,k}(\tau_i))_1^n$ is totally positive.*

Remark. The proof of Theorem 2 above differs from Karlin's proof of Theorem 3.2 in Chapter 2 of [1] only in that allowance was made here for the possibility that one of the minors appearing in (8) was actually zero. Karlin circumvented this possibility in his proof of the corollary to Theorem 2 [1; Theorem 4.1 of Ch. 10] by applying some smoothing to $(N_{j,k}(\tau_i))$. But such smoothing made it difficult to determine exactly which minors of $(N_{j,k}(\tau_i))$ are actually positive.

3. Schoenberg's variation diminishing spline approximation. The total positivity of $(N_{j,k}(\tau_i))$ has many ramifications, one of which is the remarkable fact that for any nondecreasing sequence $\boldsymbol{\tau} := (\tau_i)_1^n$ and any f, the spline approximation process

$$V : f \mapsto \sum_{i=1}^{n} f(\tau_i) N_{i,k}$$

to f is *variation diminishing, i.e.,* Vf has no more sign changes than does f itself. This follows immediately from the statement that for a totally positive matrix A and any vector $\mathbf{x}$ (of appropriate length), the vector $A\mathbf{x}$ has no more sign changes than does $\mathbf{x}$. This statement is part of Theorem 5.1.4 in Karlin [1]. But its proof there is more involved than is necessary if one only deals with the specific matrix $A = (N_{j,k}(\tau_i))$, as the arguments below hopefully show.

We say that the function f *has at least* p *strong sign changes* if f *alternates on some sequence* $(\tau_i)_0^p$, *i.e.,*

$$f(\tau_0) \neq 0, \quad \text{and} \quad f(\tau_{i-1}) f(\tau_i) < 0, \quad i = 1, \cdots, p$$

for some nondecreasing sequence $(\tau_i)_0^p$ (in the domain of f). The definition is phrased so as to allow the cases $p < 0$ (vacuously satisfied by every real function) and $p = 0$ (satisfied by every function which does not vanish identically). It is customary to denote by

$$S^-(f)$$

the total number of strong sign changes of f on its domain.

A very simple, but typical connection between total positivity and sign changes is contained in the following lemma and corollary.

Lemma 1. *Let* $f := \sum_{j=0}^{p} \alpha_j v_j$ *be a linear combination of certain functions* $v_0, \cdots, v_p$, *and suppose that for some*

$$\tau_0 < \cdots < \tau_{p+1}$$

all $(p+1)$*-minors of* $U := (v_j(\tau_i))_{i,j=0}^{p+1,p}$ *are nonnegative. If* $(v_j(\tau_i))_{i,j=0}^{p}$ *is invertible and*

$$(-1)^i f(\tau_i) \geqq 0, \qquad i = 0, \cdots, p$$

then

$$(-1)^p f(\tau_{p+1}) \geqq 0.$$

Proof. The matrix

$$\left(\begin{array}{c:c} & f(\tau_0) \\ U & \vdots \\ & f(\tau_{p+1}) \end{array}\right)$$

is not invertible since its last column is a linear combination of its first $p + 1$ columns. Hence, expanding its determinant by elements of its last column, we obtain that

$$0 = \sum_{i=0}^{p} (-1)^i f(\tau_i) \det U\begin{pmatrix}0, & \cdots, & i-1, i+1, & \cdots, & p+1\\ 0 & & , \cdots, & & p\end{pmatrix}$$
$$+ (-1)^{p+1} f(\tau_{p+1}) \det (v_j(\tau_i))_{i,j=0}^{p}.$$

But, by assumption, the sum is nonnegative, while the coefficient of $(-1)^{p+1}f(\tau_{p+1})$ is positive. □

Corollary. *Let* $f := \sum_{i=0}^{p} \alpha_i v_i$. *If* $\det (v_j(\tau_i))_{i,j=0}^{p} \geqq 0$ *for all nondecreasing* $(\tau_i)_0^p$, *with strict inequality whenever* f *alternates on* $(\tau_i)_0^p$, *then*

$$S^-(f) \geqq p \quad \textit{implies} \quad S^-(f) = p.$$

Theorem 3. *Let* $(N_{i,k})_1^n$ *be the B-spline sequence of order* k *for the nondecreasing knot sequence* $\mathbf{t} := (t_i)_1^{n+k}$ *with* $t_i < t_{i+k}$, *all* i. *If* $f = \sum_{i=1}^{n} \alpha_i N_{i,k}$, *then* $S^-(f) \leqq S^-(\boldsymbol{\alpha})$.

Proof. With $p := S^-(\boldsymbol{\alpha})$, it suffices to prove that

$$S^-(f) \geqq p \quad \text{implies} \quad S^-(f) = p$$

which we do by induction on p, it being obviously true for $p \leqq 0$. Hence, with $p > 0$, assume without loss of generality that

$$\alpha_1 > 0$$

and let $J_0, \cdots, J_p$ be the partition of $(1, \cdots, n)$ into the $p+1$ "intervals" on which $\boldsymbol{\alpha}$ has constant sign; *i.e.*, for some strictly increasing sequence $0 = \nu_{-1} < \cdots < \nu_p = n$ and for $i = 0, \cdots, p$, $J_i = (\nu_{i-1}+1, \cdots, \nu_i)$ and $(-1)^i \alpha_j \geqq 0$ for all $j \in J_i$ with at least one strict inequality. Correspondingly, let

$$v_i := \sum_{j \in J_i} |\alpha_j| \, N_{j,k} , \qquad i = 0, \cdots, p.$$

Then

$$f = \sum_{i=0}^{p} (-1)^i v_i$$

while, for any nondecreasing sequence $(\tau_i)_0^p$,

$$\det (v_j(\tau_i))_{i,j=0}^{p} = \sum_{j_0 \in J_0} \cdots \sum_{j_p \in J_p} |\alpha_{j_0}| \cdots |\alpha_{j_p}| \det (N_{j_r,k}(\tau_i))_{i,r=0}^{p}$$

is nonnegative by Theorem 2. Hence, by the corollary to Lemma 1, it suffices to prove that

$$\det (v_j(\tau_i))_{i,j=0}^{p} > 0 \tag{9}$$

whenever f alternates on $(\tau_i)_0^p$. Let $(\tau_i)_0^p$ be such a sequence. Since, by Theorem 2,

$$\det (N_{j_r,k}(\tau_i))_{i,r=0}^{p} > 0 \quad \text{iff} \quad N_{j_r,k}(\tau_r) \neq 0 \quad \text{for} \quad r = 0, \cdots, p,$$

assertion (9) is equivalent to the statement that, for $r = 0, \cdots, p$, there exists $j \in J_r$ such that $\alpha_j N_{j,k}(\tau_r) \neq 0$, hence equivalent to the statement that

$$v_r(\tau_r) \neq 0, \qquad \text{for} \quad r = 0, \cdots, p.$$

For its proof, observe that
(10) $v_i \neq 0$, $i = 0, \cdots, p$; and
(11) $v_i(s) = 0$ but $v_i(s')v_i(s'') \neq 0$ for some $s' < s < s''$ implies that $v_j(s) = 0$ for all j.
Hence, if $v_r(\tau_r) = 0$ for some r, then, by (10), $v_r(t) \neq 0$ for some $t > \tau_r$ (without loss of generality), hence, by (11), $v_r(s) = 0$ for $s \leqq \tau_r$, therefore $v_r(\tau_i) = 0$ for $i = 0, \cdots, r$, and $\tau_r < t_j$ for some $j \in J_r$. But then also $v_\mu(\tau_i) = 0$ for $i = 0, \cdots, r$ and $\mu = r + 1, \cdots, p$, hence $\hat{f} := \sum_{\mu=0}^{r-1} (-1)^\mu v_\mu$ would alternate on $(\tau_i)_0^r$ while having only $r - 1$ sign changes in its B-spline coefficients, contradicting the induction hypothesis. □

The argument can be sharpened slightly to give the

Corollary. *If $f := \sum_{i=1}^n \alpha_i N_{i,k}$ alternates on $(\tau_i)_0^q$, then*

$$f(\tau_i)\alpha_{j_i} N_{j_i,k}(\tau_i) > 0$$

for some subsequence $\mathbf{j}$ *of* $(1, \cdots, n)$.

Proof. Assuming again that $\alpha_1 > 0$ and writing

$$f = \sum_{i=0}^{p} (-1)^i v_i$$

as in the proof of the theorem, with $p = S^-(\boldsymbol{\alpha})$, and $J_0, \cdots, J_p$ the partition of $(1, \cdots, n)$ into intervals of constant sign for $\boldsymbol{\alpha}$, *etc.*, the corollary asserts the existence of some subsequence $\mathbf{j}$ of $(0, \cdots, p)$ so that

$$f(\tau_i)v_{j_i}(\tau_i) > 0, \qquad i = 0, \cdots, q.$$

This we prove by induction on p, it being obviously true for $p \leqq 0$. Let $p > 0$. We can assume that $v_0(\tau_0) \neq 0$, since $f(\tau_0) \neq 0$, hence $t_1 \leqq \tau_0$ while, by (3),

$$(12) \qquad f(t)v_0(t) > 0 \quad \text{for all} \quad t > t_1 \quad \text{and "near"} \ t_1,$$

therefore $v_0(\tau_0) = 0$ would imply that $\hat{f} := \sum_{i=1}^p (-1)^i v_i$ alternates on $(\tau_i)_0^q$ and the induction hypothesis would then furnish the desired subsequence $\mathbf{j}$ as a subsequence of $(1, \cdots, p)$. We can further assume that $f(\tau_0)v_0(\tau_0) > 0$, since the contrary case can always be reduced to it: if not $f(\tau_0)v_0(\tau_0) > 0$, then $f(\tau_0)v_0(\tau_0) < 0$, hence, by (12), f alternates on some $(\tau_{-1}, \cdots, \tau_q)$ for some additional point τ_{-1}.

If now $f(\tau_i)v_i(\tau_i) > 0$ for $i = 0, \cdots, q$, then we are done. Otherwise,

$$f(\tau_i)v_i(\tau_i) > 0 \quad \text{for} \quad i = 0, \cdots, r-1 \quad \text{and} \quad v_r(\tau_r) = 0$$

for some $r \in (0, q]$, giving $j_i = i$, $i = 0, \cdots, r - 1$. But then, since $v_r \neq 0$, *either* $v_r(t) \neq 0$ for some $t > \tau_r$, hence, as $f(\tau_r) \neq 0$, $\tau_r < t_s$ for some $s \in J_r$, therefore $v_\mu(\tau_i) = 0$ for $i = 0, \cdots, r$ and $\mu = r + 1, \cdots, q$, and so $\hat{f} := \sum_{\mu=0}{}^{r-1} (-1)^\mu v_\mu$ would alternate on $\tau_0, \cdots, \tau_r$ while having only $r - 1$ sign changes in its B-spline coefficients, an impossibility; *or* $v_r(t) \neq 0$ for some $t < \tau_r$, hence $t_s < \tau_r$ for some $s \in J_r$, therefore $v_\mu(\tau_i) = 0$ for $\mu = 0, \cdots, r - 1$ and $i = r, \cdots, q$, and so $\hat{f} := \sum_{\mu=r}{}^{p} (-1)^\mu v_\mu$ agrees with f on $\tau_r, \cdots, \tau_q$ hence alternates there and the induction hypothesis supplies the missing part $j_r, \cdots, j_q$ of the subsequence $\mathbf{j}$ as a subsequence of $(r, \cdots, p)$. $\square$

It now follows that, for any nondecreasing sequence $(\tau_i)_1{}^n$, $Vf := \sum_i f(\tau_i)N_{i,k}$ has no more sign changes than does f. In fact

$$S^-(Vf - g) \leqq S^-(f - g) \quad \text{on} \quad [t_k, t_{n+1}] \tag{13}$$

for every constant function g since $\sum_i N_{i,k} = 1$ on that interval. Further, since for every straight line ℓ on $[t_k, t_{n+1}]$,

$$\ell = \sum_{i=1}^{n} \left\{\ell(\tau_i) + \ell^{(1)}(\tau_i)\left(\sum_{j=1}^{k-1} t_{i+j} - (k-1)\tau_i\right)/(k-1)\right\}N_{i,k}$$

(see [4]), we even have (13) whenever g is a straight line provided we choose

$$\tau_i = (t_{i+1} + \cdots + t_{i+k-1})/(k-1), \qquad i = 1, \cdots, n. \tag{14}$$

The resulting approximation is known as *Schoenberg's variation diminishing spline approximation*. Schoenberg introduced this scheme in [5], promising to give a complete proof of its variation diminishing property elsewhere. Marsden and Schoenberg discuss the scheme further in [4], referring for a proof of Theorem 3 above to [6]. Schoenberg's scheme was generalized to Chebyshev splines by Karlin and Karon [2].

References

1. S. Karlin, *Total positivity, Vol. I*, Stanford University Press, Stanford, California, 1968.
2. S. Karlin, & J. M. Karon, *A variation-diminishing generalized spline approximation method*, J. Approximation Theory **1** (1968) 255–268.
3. S. Karlin & Z. Ziegler, *Tchebysheffian spline functions*, SIAM J. Numer. Anal., Series B, **3** (1966) 514–543.
4. M. Marsden & I. J. Schoenberg, *On variation diminishing spline approximation methods*, Mathematica (Cluj), **8** (**31**) (1966) 61–82.
5. I. J. Schoenberg, *On spline functions*, in *Inequalities*, O. Shisha, ed., Academic Press, New York, 1967, 255–291.
6. I. J. Schoenberg, *On the variation diminishing properties of spline functions*, to appear.
7. I. J. Schoenberg & A. Whitney, *On Pólya frequency functions. III*, Trans. Amer. Math. Soc., **74** (1953) 246–259.

This work was sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

University of Wisconsin

Date communicated: November 8, 1974