

STRUCTURED INVERSION OF THE BERNSTEIN–VANDERMONDE MATRIX*

LARRY ALLEN[†] AND ROBERT C. KIRBY[†]

Abstract. Bernstein polynomials, long a staple of approximation theory and computational geometry, have also increasingly become of interest in finite element methods. Many fundamental problems in interpolation and approximation give rise to interesting linear algebra questions. When attempting to find a polynomial approximation of boundary or initial data, one encounters the Bernstein–Vandermonde matrix, which is found to be highly ill conditioned. In [L. Allen and R. C. Kirby, *SIAM J. Matrix Anal. Appl.*, 41 (2020), pp. 413–431], we used the relationship between monomial Bézout matrices and the inverse of Hankel matrices to obtain a decomposition of the inverse of the Bernstein mass matrix in terms of Hankel, Toeplitz, and diagonal matrices. In this paper, we use properties of the Bernstein–Bézout matrix to factor the inverse of the Bernstein–Vandermonde matrix into a difference of products of Hankel, Toeplitz, and diagonal matrices. We also use the nonstandard matrix norm defined in [L. Allen and R. C. Kirby, *SIAM J. Matrix Anal. Appl.*, 41 (2020), pp. 413–431] to study the conditioning of the Bernstein–Vandermonde matrix, showing that the conditioning in this case is better than in the standard 2 norm. Additionally, we use properties of multivariate Bernstein polynomials to derive a block LU decomposition of the Bernstein–Vandermonde matrix corresponding to equispaced nodes on the d -simplex.

Key words. Bernstein polynomials, Lagrange polynomials, Legendre polynomials, Bernstein–Vandermonde matrix, Bernstein mass matrix, interpolation, matrix inverse, Bézout matrix, Hankel matrix, Toeplitz matrix, conditioning

AMS subject classifications. 65F05, 65F15, 65F35

DOI. 10.1137/20M1336606

1. Introduction. Given data $\{f_j\}_{j=0}^n$ and distinct nodes $\{x_j\}_{j=0}^n$, the interpolation problem consists of finding a polynomial p of degree n that satisfies

$$(1.1) \quad p(x_j) = f_j$$

for each $0 \leq j \leq n$. If a basis for the space of polynomials of degree at most n is chosen, then the interpolation problem can be expressed as a system of linear equations, where the coefficient matrix is a Vandermonde-like matrix [11]. Recently, Bernstein polynomials have been considered as a tool for high-order approximation of partial differential equations via the finite element method [1, 7], and the interpolant is often used as a polynomial approximation of initial or boundary data. The corresponding Bernstein–Vandermonde matrix is found to be highly ill conditioned [5] in the 2 norm, but the structure of the matrix has led to fast algorithms that avoid some of the issues that arise from the ill-conditioning. For example, Marco and Martínez [16] used the fact that the Bernstein–Vandermonde matrix is strictly totally positive [9] to obtain a bidiagonal factorization of the inverse, and Ainsworth and Sánchez [2] adapted the standard divided difference algorithm for the monomial Vandermonde matrix [4] to the Bernstein basis. In addition, it was observed in [6] that the collocation matrices of the Bernstein basis are optimally conditioned among all other totally positive bases when considering the conditioning in the infinity norm.

*Received by the editors May 7, 2020; accepted for publication (in revised form) by S. P. MacLachlan January 8, 2021; published electronically April 8, 2021.

<https://doi.org/10.1137/20M1336606>

[†]Department of Mathematics, Baylor University, Waco, TX 76798-7328 USA (larry_allen@baylor.edu, robert.kirby@baylor.edu).

In [3], we used the relationship between monomial Bézout matrices and the inverse of Hankel matrices described by Heinig and Rost [12] to obtain a decomposition of the inverse of the Bernstein mass matrix in terms of Hankel, Toeplitz, and diagonal matrices. In this paper, we generalize an argument made by Kaplan, Shapiro, and Teicher [13] to obtain a decomposition of the inverse of Vandermonde-like matrices in terms of their transpose, a diagonal matrix, and the corresponding Bézout matrix. When applied to the Bernstein basis, this gives a decomposition of the inverse of the Bernstein–Vandermonde matrix in terms of Hankel, Toeplitz, and diagonal matrices, which in turn leads to a fast algorithm for solving the interpolation problem. Additionally, we use the nonstandard matrix norm defined in [3] to give an explanation for the relatively good performance of the Bernstein–Vandermonde matrix despite its massive condition number. While the application of the arguments from [13] to the Bernstein basis leads to an algorithm that is less stable than the ones observed in [2, 16], the derivation is generic and can be applied to other bases and used to investigate the corresponding interpolation problems.

Bernstein polynomials also extend naturally to give a basis for multivariate polynomials of total degree n . Properties of Bernstein polynomials lead to special recursive blockwise-structure for finite element matrices [14, 15]. In this paper, we use this structure to obtain a block LU decomposition of the Bernstein–Vandermonde matrix associated with equispaced nodes on the d -simplex. Equispaced nodes on the d -simplex are commonly used in finite element methods, and so this decomposition (combined with fast, stable algorithms for the one-dimensional Bernstein–Vandermonde matrix) leads to recursive, block-structured fast algorithms for the related interpolation problems.

The rest of the paper is organized as follows. In section 2, we use the connection between Bézout and Vandermonde matrices to give a structured decomposition of the inverse of the Bernstein–Vandermonde matrix. Different algorithms for applying the inverse to a vector are discussed in section 3. In section 4, we use a nonstandard matrix norm to investigate the conditioning of the Bernstein–Vandermonde matrix. In section 5, we give a block LU decomposition of the Bernstein–Vandermonde matrix corresponding to equispaced nodes on the d -simplex. A comparison of the algorithms discussed in section 3 is given in section 6, and final remarks are given in section 7.

2. Inverse formulas.

2.1. General formulas. For integers $m, n \geq 0$, let $\{b_j^n(x)\}_{j=0}^n$ be a basis for the space of univariate polynomials of degree at most n , and let $\mathbf{x} \in \mathbb{R}^{m+1}$ with $\mathbf{x}_i < \mathbf{x}_{i+1}$ for each $0 \leq i < m$, where boldface is used to distinguish the vector of nodes \mathbf{x} from the indeterminate x (a similar convention will be used to distinguish a polynomial v from its vector of coefficients \mathbf{v}). The Vandermonde matrix $V^n(\mathbf{x})$ associated with $\{b_j^n(x)\}_{j=0}^n$ and \mathbf{x} is the $(m+1) \times (n+1)$ matrix given by

$$(2.1) \quad V_{ij}^n(\mathbf{x}) = b_j^n(\mathbf{x}_i).$$

We consider the case $m = n$ so that $V^n(\mathbf{x})$ is invertible.

If v and w are polynomials of degree at most $n+1$ expressed in the basis $\{b_j^{n+1}(x)\}_{j=0}^{n+1}$, then the Bézout matrix generated by v and w , denoted by $\text{Bez}(v, w)$, is the $(n+1) \times (n+1)$ matrix whose entries satisfy

$$(2.2) \quad \frac{v(s)w(t) - v(t)w(s)}{s - t} = \sum_{i,j=0}^n \text{Bez}_{ij}(v, w) b_i^n(s) b_j^n(t).$$

In [13], it was shown that if the monomial basis is used, then the inverse of a Bézout matrix is a Hankel matrix. As a consequence of the arguments involved, one obtains a formula for the inverse of the monomial Vandermonde matrix in terms of its transpose, a diagonal matrix, and a particular Bézout matrix. Since none of the arguments requires anything specific to the monomial basis, we can generalize the arguments to any basis, which we summarize here.

For $t \in \mathbb{R}$, let $\mathbf{b}^n(t)$ denote the column vector $(b_0^n(t), \dots, b_n^n(t))^T$. As a consequence of (2.2), if $s \neq t$, then

$$\begin{aligned} (\mathbf{b}^n(s))^T \text{Bez}(v, w) \mathbf{b}^n(t) &= \sum_{i,j=0}^n \text{Bez}_{ij}(v, w) b_i^n(s) b_j^n(t) \\ &= \frac{v(s)w(t) - v(t)w(s)}{s - t}. \end{aligned}$$

This implies that

$$\begin{aligned} (\mathbf{b}^n(t))^T \text{Bez}(v, w) \mathbf{b}^n(t) &= \lim_{\varepsilon \rightarrow 0} [(\mathbf{b}^n(t))^T \text{Bez}(v, w) \mathbf{b}^n(t + \varepsilon)] \\ &= \lim_{\varepsilon \rightarrow 0} \frac{v(t + \varepsilon)w(t) - v(t)w(t + \varepsilon)}{\varepsilon} \\ &= v'(t)w(t) - v(t)w'(t). \end{aligned}$$

Therefore, if v has simple zeros at \mathbf{x}_i and $w(\mathbf{x}_i) \neq 0$ for each $0 \leq i \leq n$, then

$$(2.3) \quad V^n(\mathbf{x}) \text{Bez}(v, w) (V^n(\mathbf{x}))^T = \text{diag}(v'(\mathbf{x}_j)w(\mathbf{x}_j))_{j=0}^n,$$

and so we have the following.

THEOREM 2.1. *The inverse of $V^n(\mathbf{x})$ is given by*

$$(2.4) \quad (V^n(\mathbf{x}))^{-1} = \text{Bez}(v, w) (V^n(\mathbf{x}))^T \text{diag} \left(\frac{1}{v'(\mathbf{x}_j)w(\mathbf{x}_j)} \right)_{j=0}^n,$$

where v and w are any polynomials of degree $n + 1$ satisfying $v(\mathbf{x}_i) = 0$, $v'(\mathbf{x}_i) \neq 0$, and $w(\mathbf{x}_i) \neq 0$ for each $0 \leq i \leq n$.

COROLLARY 2.2. *The inverse of $V^n(\mathbf{x})$ is given by*

$$(2.5) \quad (V^n(\mathbf{x}))^{-1} = \text{Bez}(v, 1) (V^n(\mathbf{x}))^T \text{diag} \left(\frac{1}{v'(\mathbf{x}_j)} \right)_{j=0}^n,$$

where v is any polynomial of degree $n + 1$ that has simple zeros at \mathbf{x}_i for each $0 \leq i \leq n$.

2.2. Bernstein–Vandermonde. We now focus on the case where $0 \leq \mathbf{x}_i \leq 1$ for each $0 \leq i \leq n$ and the basis consists of the polynomials $\{B_j^n(x)\}_{j=0}^n$, where

$$(2.6) \quad B_j^n(x) = \binom{n}{j} x^j (1 - x)^{n-j}$$

are the Bernstein polynomials of degree n . In order to apply Theorem 2.1 or Corollary 2.2, we need a method for computing the entries of the Bernstein–Bézout matrix; in fact, they satisfy a recurrence relation involving the Bernstein coefficients of the polynomials.

THEOREM 2.3. If $v(t) = \sum_{i=0}^{n+1} \mathbf{v}_i B_i^{n+1}(t)$ and $w(t) = \sum_{i=0}^{n+1} \mathbf{w}_i B_i^{n+1}(t)$ are polynomials of degree at most $n+1$, then the entries b_{ij} of the Bernstein-Bézout matrix $\text{Bez}(v, w)$ generated by v and w satisfy

$$(2.7) \quad b_{ij} = \frac{1}{(i+1)(n-j+1)} [j(n-i)b_{i+1,j-1} + (n+1)^2 (\mathbf{v}_{i+1} \mathbf{w}_j - \mathbf{v}_j \mathbf{w}_{i+1})].$$

Proof. By (2.6), we have that

$$(2.8) \quad tB_j^n(t) = \frac{j+1}{n+1} B_{j+1}^{n+1}(t)$$

and

$$(2.9) \quad B_j^n(t) = \frac{n-j+1}{n+1} B_j^{n+1}(t) + \frac{j+1}{n+1} B_{j+1}^{n+1}(t).$$

Therefore,

$$\begin{aligned} s \sum_{i,j=0}^n b_{ij} B_i^n(s) B_j^n(t) &= \sum_{i,j=0}^n \frac{i+1}{n+1} b_{ij} B_{i+1}^{n+1}(s) B_j^n(t) \\ &= \sum_{i,j=0}^n \frac{(i+1)(n-j+1)}{(n+1)^2} b_{ij} B_{i+1}^{n+1}(s) B_j^{n+1}(t) \\ &\quad + \sum_{i,j=0}^n \frac{(i+1)(j+1)}{(n+1)^2} b_{ij} B_{i+1}^{n+1}(s) B_{j+1}^{n+1}(t). \end{aligned}$$

Similarly, we have that

$$\begin{aligned} t \sum_{i,j=0}^n b_{ij} B_i^n(s) B_j^n(t) &= \sum_{i,j=0}^n \frac{j+1}{n+1} b_{ij} B_i^n(s) B_{j+1}^{n+1}(t) \\ &= \sum_{i,j=0}^n \frac{(n-i+1)(j+1)}{(n+1)^2} b_{ij} B_i^{n+1}(s) B_{j+1}^{n+1}(t) \\ &\quad + \sum_{i,j=0}^n \frac{(i+1)(j+1)}{(n+1)^2} b_{ij} B_{i+1}^{n+1}(s) B_{j+1}^{n+1}(t). \end{aligned}$$

This implies that

$$\begin{aligned} (s-t) \sum_{i,j=0}^n b_{ij} B_i^n(s) B_j^n(t) &= \sum_{i,j=0}^n \frac{(i+1)(n-j+1)}{(n+1)^2} b_{ij} B_{i+1}^{n+1}(s) B_j^{n+1}(t) \\ &\quad - \sum_{i,j=0}^n \frac{(n-i+1)(j+1)}{(n+1)^2} b_{ij} B_i^{n+1}(s) B_{j+1}^{n+1}(t), \end{aligned}$$

and so

$$\begin{aligned} (s-t) \sum_{i,j=0}^n b_{ij} B_i^n(s) B_j^n(t) &= \sum_{i,j=0}^{n+1} \frac{i(n-j+1)}{(n+1)^2} b_{i-1,j} B_i^{n+1}(s) B_j^{n+1}(t) \\ &\quad - \sum_{i,j=0}^{n+1} \frac{j(n-i+1)}{(n+1)^2} b_{i,j-1} B_i^{n+1}(s) B_j^{n+1}(t). \end{aligned}$$

On the other hand,

$$(2.10) \quad v(s)w(t) - v(t)w(s) = \sum_{i,j=0}^{n+1} (\mathbf{v}_i \mathbf{w}_j - \mathbf{v}_j \mathbf{w}_i) B_i^{n+1}(s) B_j^{n+1}(t),$$

and so the result follows by comparing coefficients in (2.2). \square

We can form the first column and last row of $\text{Bez}(v, w)$ by using (2.7) with $j = 0$ and $i = n$. The rest of the columns can then be built by applying the recurrence relation. Therefore, we have the following corollary.

COROLLARY 2.4. *$\text{Bez}(v, w)$ can be constructed in $\mathcal{O}(n^2)$ operations.*

By repeatedly applying (2.7), we obtain a closed form for the entries of the Bernstein-Bézout matrix.

COROLLARY 2.5. *If $v(t) = \sum_{i=0}^{n+1} \mathbf{v}_i B_i^{n+1}(t)$ and $w(t) = \sum_{i=0}^{n+1} \mathbf{w}_i B_i^{n+1}(t)$ are polynomials of degree at most $n+1$, then the entries b_{ij} of the Bernstein-Bézout matrix $\text{Bez}(v, w)$ generated by v and w are given by*

$$(2.11) \quad b_{ij} = \frac{1}{\binom{n}{i} \binom{n}{j}} \sum_{k=0}^{m_{ij}} \binom{n+1}{i+k+1} \binom{n+1}{j-k} (\mathbf{v}_{i+k+1} \mathbf{w}_{j-k} - \mathbf{v}_{j-k} \mathbf{w}_{i+k+1}),$$

where $m_{ij} = \min\{j, n-i\}$.

Constructing $\text{Bez}(v, w)$ by using (2.11) would require $\mathcal{O}(n^3)$ operations, and so Corollary 2.4 implies that (2.11) is not optimal for building $\text{Bez}(v, w)$; however, since $\binom{n}{i} = 0$ whenever $i > n$, the upper limit of the sum in (2.11) can be replaced with $k = n$, and so we recognize that $\text{Bez}(v, w)$ is given by a difference of matrix products.

COROLLARY 2.6. *Let $v(t) = \sum_{i=0}^{n+1} \mathbf{v}_i B_i^{n+1}(t)$ and $w(t) = \sum_{i=0}^{n+1} \mathbf{w}_i B_i^{n+1}(t)$ be polynomials of degree at most $n+1$. Define the Toeplitz matrices $T^{v,n}$ and $T^{w,n}$ by*

$$T_{ij}^{v,n} = \binom{n+1}{j-i} \mathbf{v}_{j-i} \quad \text{and} \quad T_{ij}^{w,n} = \binom{n+1}{j-i} \mathbf{w}_{j-i},$$

define the Hankel matrices $H^{v,n}$ and $H^{w,n}$ by

$$H_{ij}^{v,n} = \binom{n+1}{i+j+1} \mathbf{v}_{i+j+1} \quad \text{and} \quad H_{ij}^{w,n} = \binom{n+1}{i+j+1} \mathbf{w}_{i+j+1},$$

and let $\Delta^n = \text{diag} \left(\binom{n}{j} \right)_{j=0}^n$. Then the Bernstein-Bézout matrix $\text{Bez}(v, w)$ generated by v and w is given by

$$(2.12) \quad \text{Bez}(v, w) = (\Delta^n)^{-1} [H^{v,n} T^{w,n} - H^{w,n} T^{v,n}] (\Delta^n)^{-1}.$$

Remark 2.7. In [3], it was shown that the inverse of the degree n Bernstein mass matrix M^n has the decomposition

$$(2.13) \quad (M^n)^{-1} = (\Delta^n)^{-1} [\tilde{T}^n H^n - T^n \tilde{H}^n] (\Delta^n)^{-1},$$

where \tilde{H}^n and H^n are Hankel matrices, \tilde{T}^n and T^n are Toeplitz matrices, and $\Delta^n = \text{diag} \left(\binom{n}{j} \right)_{j=0}^n$. This decomposition was obtained from the monomial Bézout matrix corresponding to polynomials whose coefficients are contained in the last column of $(M^n)^{-1}$. It is interesting to note the similarities between (2.12) and (2.13) despite the different matrices and bases being considered.

Given a polynomial $v(t) = \sum_{k=0}^{n+1} \mathbf{v}_k B_k^{n+1}(t)$ of degree at most $n+1$, we can express v in the monomial basis as $v(t) = \sum_{k=0}^{n+1} \tilde{\mathbf{v}}_k t^k$. It was shown in [8] that

$$(2.14) \quad \mathbf{v}_k = \frac{1}{\binom{n+1}{k}} \sum_{\ell=0}^k \binom{n-\ell+1}{k-\ell} \tilde{\mathbf{v}}_\ell.$$

This implies that

$$\begin{aligned} H_{ij}^{v,n} &= \binom{n+1}{i+j+1} \mathbf{v}_{i+j+1} = \sum_{k=0}^{i+j+1} \binom{n-k+1}{i+j-k+1} \tilde{\mathbf{v}}_k \\ &= \sum_{k=0}^{i+j+1} \binom{n-k+1}{n-i-j} \tilde{\mathbf{v}}_k. \end{aligned}$$

Similarly,

$$T_{ij}^{v,n} = \sum_{k=0}^{j-i} \binom{n-k+1}{j-i-k} \tilde{\mathbf{v}}_k.$$

By Vieta's formula (see, for example, [18]),

$$(2.15) \quad \prod_{i=0}^n (x - \mathbf{x}_i) = \sum_{k=0}^n (-1)^{n-k+1} \sigma_{n-k+1}(\mathbf{x}) x^k,$$

where

$$(2.16) \quad \sigma_k(\mathbf{x}) = \begin{cases} 1 & \text{if } k = 0, \\ \sum_{0 \leq i_0 < \dots < i_{k-1} \leq n} \mathbf{x}_{i_0} \cdots \mathbf{x}_{i_{k-1}} & \text{otherwise} \end{cases}$$

is the k th elementary symmetric function in the $n+1$ variables $\mathbf{x}_0, \dots, \mathbf{x}_n$. In addition, since

$$(2.17) \quad 1 = \sum_{i=0}^{n+1} B_i^{n+1}(x)$$

and

$$(2.18) \quad \frac{d}{dx} \left[\prod_{i=0}^n (x - \mathbf{x}_i) \right]_{x=\mathbf{x}_j} = \prod_{i \in \{0, \dots, n\} \setminus \{j\}} (\mathbf{x}_j - \mathbf{x}_i),$$

we can combine the previous discussion with Corollaries 2.2 and 2.6 to obtain a decomposition of the inverse of the Bernstein–Vandermonde matrix.

THEOREM 2.8. *Define the Hankel matrices H^n and $\tilde{H}^n(\mathbf{x})$ by*

$$H_{ij}^n = \binom{n+1}{i+j+1} \quad \text{and} \quad \tilde{H}_{ij}^n(\mathbf{x}) = \sum_{k=0}^{i+j+1} (-1)^{n-k+1} \binom{n-k+1}{n-i-j} \sigma_{n-k+1}(\mathbf{x}),$$

define the Toeplitz matrices T^n and $\tilde{T}^n(\mathbf{x})$ by

$$T_{ij}^n = \binom{n+1}{j-i} \quad \text{and} \quad \tilde{T}_{ij}^n(\mathbf{x}) = \sum_{k=0}^{j-i} (-1)^{n-k+1} \binom{n-k+1}{j-i-k} \sigma_{n-k+1}(\mathbf{x}),$$

and define the diagonal matrices $D^n(\mathbf{x})$ and Δ^n by

$$D^n(\mathbf{x}) = \text{diag} \left(\prod_{i \in \{0, \dots, n\} \setminus \{j\}} (\mathbf{x}_j - \mathbf{x}_i) \right)_{j=0}^n \quad \text{and} \quad \Delta^n = \text{diag} \left(\binom{n}{j} \right)_{j=0}^n.$$

In addition, let $\tilde{V}^n(\mathbf{x})$ be the scaled Bernstein-Vandermonde matrix

$$\tilde{V}_{ij}^n(\mathbf{x}) = \mathbf{x}_i^j (1 - \mathbf{x}_i)^{n-j}.$$

Then

$$(2.19) \quad (V^n(\mathbf{x}))^{-1} = (\Delta^n)^{-1} \left[\tilde{H}^n(\mathbf{x}) T^n - H^n \tilde{T}^n(\mathbf{x}) \right] \left(\tilde{V}^n(\mathbf{x}) \right)^T (D^n(\mathbf{x}))^{-1}.$$

2.3. Equispaced nodes. In section 5, we will derive a block LU decomposition for the Bernstein-Vandermonde matrix associated with equispaced nodes on the d -simplex. This decomposition leads to a recursive, block-structured algorithm for the corresponding interpolation problem. The base for this algorithm is the one-dimensional Bernstein-Vandermonde matrix associated with equispaced nodes, and so we now briefly focus on the case where $\mathbf{x}_i = i/n$ for each $0 \leq i \leq n$. For these nodes, we observe that (2.15) becomes

$$\begin{aligned} \prod_{i=0}^n (x - i/n) &= \frac{1}{n^{n+1}} \prod_{i=0}^{n+1} (nx - i) = \frac{1}{n^{n+1}} \sum_{k=0}^{n+1} s(n+1, k) (nx)^k \\ &= \sum_{k=0}^{n+1} \frac{s(n+1, k)}{n^{n-k+1}} x^k, \end{aligned}$$

where $s(n, k)$ are the (signed) Stirling numbers of the first kind [19]. In addition, we have that

$$(2.20) \quad \prod_{i \in \{0, \dots, n\} \setminus \{j\}} (j/n - i/n) = (-1)^{n-j} \frac{j!(n-j)!}{n^n}.$$

COROLLARY 2.9. Define the Hankel matrices H^n and \tilde{H}^n by

$$H_{ij}^n = \binom{n+1}{i+j+1} \quad \text{and} \quad \tilde{H}_{ij}^n = \sum_{k=0}^{i+j+1} \binom{n-k+1}{n-i-j} \frac{s(n+1, k)}{n^{n-k+1}},$$

define the Toeplitz matrices T^n and \tilde{T}^n by

$$T_{ij}^n = \binom{n+1}{j-i} \quad \text{and} \quad \tilde{T}_{ij}^n = \sum_{k=0}^{j-i} \binom{n-k+1}{j-i-k} \frac{s(n+1, k)}{n^{n-k+1}},$$

and define the diagonal matrices D^n and Δ^n by

$$D^n = \text{diag} \left((-1)^{n-j} [j!(n-j)!] \right)_{j=0}^n \quad \text{and} \quad \Delta^n = \text{diag} \left(\binom{n}{j} \right)_{j=0}^n.$$

In addition, let \tilde{V}^n be the scaled Bernstein-Vandermonde matrix

$$\tilde{V}_{ij}^n = i^j (n-i)^{n-j}.$$

Then the inverse of $V^n = V^n(0, 1/n, \dots, 1)$ is given by

$$(2.21) \quad (V^n)^{-1} = (\Delta^n)^{-1} \left[\tilde{H}^n T^n - H^n \tilde{T}^n \right] \left(\tilde{V}^n \right)^T (D^n)^{-1}.$$

3. Applying the inverse. Now, we describe several approaches to applying $(V^n(\mathbf{x}))^{-1}$ to a vector.

LU factorization. We can decompose $V^n(\mathbf{x})$ as

$$(3.1) \quad V^n(\mathbf{x}) = L^n(\mathbf{x})U^n(\mathbf{x}),$$

where $L^n(\mathbf{x})$ and $U^n(\mathbf{x})$ are lower and upper triangular matrices, respectively. Widely available in libraries, computing $L^n(\mathbf{x})$ and $U^n(\mathbf{x})$ requires $\mathcal{O}(n^3)$ operations, and each of the subsequent triangular solves requires $\mathcal{O}(n^2)$ operations to perform.

Exact inverse. In light of Corollary 2.2, we can directly form $(V^n(\mathbf{x}))^{-1}$. By Theorem 2.3, we can form the inverse in $\mathcal{O}(n^2)$ operations. The inverse can then be applied to any vector in $\mathcal{O}(n^2)$ operations using the standard algorithm.

DFT-based application. By Theorem 2.8, we can express the inverse of $V^n(\mathbf{x})$ in terms of its (scaled) transpose and Hankel, Toeplitz, and diagonal matrices. The matrices $H^n(\mathbf{x})$ and $T^n(\mathbf{x})$ can be formed in $\mathcal{O}(n)$ operations, but the matrices $\tilde{H}^n(\mathbf{x})$ and $\tilde{T}^n(\mathbf{x})$ require $\mathcal{O}(n^2)$ operations to form, since each entry involves a sum. Even though the Hankel and Toeplitz matrices can be applied to a vector in $\mathcal{O}(n \log n)$ operations via circulant embedding and a couple of FFT/iFFT [20], the scaled transpose still requires $\mathcal{O}(n^2)$ operations to apply to a vector.

Newton algorithm. In [2], Ainsworth and Sánchez gave a fast algorithm for solving $V^n(\mathbf{x})\mathbf{c} = \mathbf{b}$ via a recursion relation for the Bézier control points of the Newton form of the Lagrangian interpolant. The numerical results given suggest that the algorithm is stable even for high polynomial degree, and so we include it for comparison.

Marco–Martínez algorithm. In [16], Marco and Martínez gave a bidiagonal factorization of $(V^n(\mathbf{x}))^{-1}$ in the case where the nodes are contained in the interior of $[0, 1]$, which can be extended to the case where one or both of the endpoints are contained in the list of nodes. The bidiagonal factors can be built in $\mathcal{O}(n^2)$ operations, and the inverse can subsequently be applied to a vector in $\mathcal{O}(n^2)$ operations. The startup and application costs are similar to the exact inverse method, and so we include it for comparison.

4. Conditioning and accuracy. In [3], we gave a partial explanation of the relatively high accuracy, compared to its large 2-norm condition number, of working with the Bernstein mass matrix. We will use similar reasoning to investigate the performance of the Bernstein–Vandermonde matrix.

Recall that for an integer $n \geq 0$, the Bernstein mass matrix M^n is the $(n+1) \times (n+1)$ matrix given by

$$(4.1) \quad M_{ij}^n = \int_0^1 B_i^n(x)B_j^n(x)dx.$$

The mass matrix (for Bernstein or any other family of polynomials) plays an important role in connecting the L^2 topology on the finite-dimensional space to linear algebra. To see this, we first define mappings connecting polynomials of degree n to \mathbb{R}^{n+1} . Given any $\mathbf{c} \in \mathbb{R}^{n+1}$, we let $\pi(\mathbf{c})$ be the polynomial expressed in the Bernstein basis with coefficients contained in \mathbf{c} :

$$(4.2) \quad \pi(\mathbf{c})(x) = \sum_{i=0}^n \mathbf{c}_i B_i^n(x).$$

We let $\mathbf{\Pi}$ be the inverse of this mapping, sending any polynomial of degree at most n to the vector of $n+1$ coefficients with respect to the Bernstein basis.

Now, let $p(x)$ and $q(x)$ be polynomials of degree n with expansion coefficients $\mathbf{\Pi}(p) = \mathbf{p}$ and $\mathbf{\Pi}(q) = \mathbf{q}$. Then the L^2 inner product of p and q is given by the M^n -weighted inner product of \mathbf{p} and \mathbf{q} for

$$(4.3) \quad \int_0^1 p(x)q(x)dx = \sum_{i,j=0}^n \mathbf{p}_i \mathbf{q}_j \int_0^1 B_i^n(x)B_j^n(x)dx = \mathbf{p}^T M^n \mathbf{q}.$$

Similarly, if

$$(4.4) \quad \|\mathbf{p}\|_{M^n} = \sqrt{\mathbf{p}^T M^n \mathbf{p}}$$

is the M^n -weighted vector norm, then we have for $p = \pi(\mathbf{p})$,

$$(4.5) \quad \|p\|_{L^2} = \|\mathbf{p}\|_{M^n}.$$

We can interpret the Bernstein–Vandermonde matrix $V^n(\mathbf{x})$ as an operator mapping polynomials of degree at most n to \mathbb{R}^{n+1} via $p \mapsto V^n(\mathbf{x})\mathbf{\Pi}(p)$. Therefore, it makes sense to measure p in the L^2 norm and $V^n(\mathbf{x})\mathbf{\Pi}(p)$ in the 2 norm. By (4.5), we can also consider $\mathbf{\Pi}(p)$ in the M^n norm, which leads us to define the operator norm

$$(4.6) \quad \|A\|_{M^n \rightarrow 2} = \max_{\mathbf{y} \neq \mathbf{0}} \frac{\|A\mathbf{y}\|_2}{\|\mathbf{y}\|_{M^n}},$$

and going in the opposite direction,

$$(4.7) \quad \|A\|_{2 \rightarrow M^n} = \max_{\mathbf{y} \neq \mathbf{0}} \frac{\|A\mathbf{y}\|_{M^n}}{\|\mathbf{y}\|_2}.$$

These two norms naturally combine to define a new condition number

$$(4.8) \quad \kappa_{M^n \rightarrow 2}(A) = \|A\|_{M^n \rightarrow 2} \|A^{-1}\|_{2 \rightarrow M^n}.$$

Since M^n is symmetric and positive definite, it has a well-defined positive square root via the spectral decomposition.

LEMMA 4.1.

$$(4.9) \quad \kappa_{M^n \rightarrow 2}(V^n(\mathbf{x})) = \kappa_2 \left(V^n(\mathbf{x}) (M^n)^{-1/2} \right).$$

Proof. Observe that if $\mathbf{y} \neq \mathbf{0}$, then

$$\begin{aligned} \frac{\|(V^n(\mathbf{x}))^{-1} \mathbf{y}\|_{M^n}^2}{\|\mathbf{y}\|_2^2} &= \frac{\mathbf{y}^T (V^n(\mathbf{x}))^{-T} M^n (V^n(\mathbf{x}))^{-1} \mathbf{y}}{\|\mathbf{y}\|_2^2} \\ &= \frac{\left((M^n)^{1/2} (V^n(\mathbf{x}))^{-1} \mathbf{y} \right)^T \left((M^n)^{1/2} (V^n(\mathbf{x}))^{-1} \mathbf{y} \right)}{\|\mathbf{y}\|_2^2} \\ &= \frac{\left\| (M^n)^{1/2} (V^n(\mathbf{x}))^{-1} \mathbf{y} \right\|_2^2}{\|\mathbf{y}\|_2^2}, \end{aligned}$$

and so

$$(4.10) \quad \left\| (V^n(\mathbf{x}))^{-1} \right\|_{2 \rightarrow M^n} = \left\| (M^n)^{1/2} (V^n(\mathbf{x}))^{-1} \right\|_2.$$

Let $\mathbf{z} = (M^n)^{1/2} \mathbf{y}$. Then $\mathbf{z} \neq \mathbf{0}$ and

$$\begin{aligned} \frac{\|V^n(\mathbf{x})\mathbf{y}\|_2^2}{\|\mathbf{y}\|_{M^n}^2} &= \frac{\mathbf{y}^T (V^n(\mathbf{x}))^T V^n(\mathbf{x}) \mathbf{y}}{\mathbf{y}^T M^n \mathbf{y}} \\ &= \frac{\left(V^n(\mathbf{x}) (M^n)^{-1/2} \mathbf{z}\right)^T \left(V^n(\mathbf{x}) (M^n)^{-1/2} \mathbf{z}\right)}{\mathbf{z}^T \mathbf{z}} \\ &= \frac{\left\|V^n(\mathbf{x}) (M^n)^{-1/2} \mathbf{z}\right\|_2^2}{\|\mathbf{z}\|_2^2}, \end{aligned}$$

and so

$$(4.11) \quad \|V^n(\mathbf{x})\|_{M^n \rightarrow 2} = \left\|V^n(\mathbf{x}) (M^n)^{-1/2}\right\|_2.$$

It follows that

$$\begin{aligned} \kappa_{M^n \rightarrow 2}(V^n(\mathbf{x})) &= \|V^n(\mathbf{x})\|_{M^n \rightarrow 2} \left\|(V^n(\mathbf{x}))^{-1}\right\|_{2 \rightarrow M^n} \\ &= \left\|V^n(\mathbf{x}) (M^n)^{-1/2}\right\|_2 \left\|(M^n)^{1/2} (V^n(\mathbf{x}))^{-1}\right\|_2 \\ &= \left\|V^n(\mathbf{x}) (M^n)^{-1/2}\right\|_2 \left\|\left(V^n(\mathbf{x}) (M^n)^{-1/2}\right)^{-1}\right\|_2 \\ &= \kappa_2 \left(V^n(\mathbf{x}) (M^n)^{-1/2}\right). \quad \square \end{aligned}$$

To analyze $\kappa_2(V^n(\mathbf{x})(M^n)^{-1/2})$, we will need reference to the Legendre polynomials, mapped from their typical home on $[-1, 1]$ to $[0, 1]$. For $0 \leq j \leq n$, let $L^j(x)$ denote the Legendre polynomial of degree n over $[0, 1]$, scaled so that $L^n(1) = 1$ and

$$(4.12) \quad \|L^j\|_{L^2}^2 = \frac{1}{2j+1}.$$

The Legendre–Vandermonde matrix is the $(n+1) \times (n+1)$ matrix $\widehat{V}^n(\mathbf{x})$ given by

$$(4.13) \quad \widehat{V}_{ij}^n(\mathbf{x}) = L^j(\mathbf{x}_i).$$

Given any polynomial p of degree at most n , let $\Theta(p)$ denote the vector of $n+1$ coefficients with respect to the Legendre basis. Define T^n to be the $(n+1) \times (n+1)$ matrix satisfying $T^n \Theta(p) = \Pi(p)$ for all polynomials p of degree at most n . In particular, we have the relationship

$$(4.14) \quad \widehat{V}^n(\mathbf{x}) = V^n(\mathbf{x}) T^n.$$

Let $\{\lambda_j^n\}_{j=0}^n$ be the set of eigenvalues of M^n . In [3], it was shown that M^n admits the spectral decomposition

$$(4.15) \quad M^n = Q^n \Lambda^n (Q^n)^T,$$

where

$$(4.16) \quad Q^n = T^n \operatorname{diag} \left(\sqrt{(2j+1)\lambda_j^n} \right)_{j=0}^n$$

is an orthogonal matrix and $\Lambda^n = \text{diag}(\lambda_j^n)_{j=0}^n$ contains the eigenvalues. This decomposition combined with Lemma 4.1 and the transformation given in (4.14) suggest a relationship between the $M^n \rightarrow 2$ condition number of $V^n(x)$ and the 2-norm condition number of $\widehat{V}^n(\mathbf{x})$ scaled by a diagonal matrix.

THEOREM 4.2.

$$(4.17) \quad \kappa_{M^n \rightarrow 2}(V^n(\mathbf{x})) = \kappa_2 \left(\widehat{V}^n(\mathbf{x}) \text{diag} \left(\sqrt{2j+1} \right)_{j=0}^n \right).$$

Proof. We have that

$$\begin{aligned} V^n(\mathbf{x}) (M^n)^{-1/2} &= V^n(\mathbf{x}) Q^n (\Lambda^n)^{-1/2} (Q^n)^T \\ &= V^n(\mathbf{x}) T^n \text{diag} \left(\sqrt{(2j+1)\lambda_j^n} \right)_{j=0}^n \text{diag} \left(\frac{1}{\sqrt{\lambda_j^n}} \right)_{j=0}^n (Q^n)^T \\ &= \widehat{V}^n(\mathbf{x}) \text{diag} \left(\sqrt{2j+1} \right)_{j=0}^n (Q^n)^T. \end{aligned}$$

Since Q^n is orthogonal, the result follows from Lemma 4.1. \square

In [10], Gautschi computed the condition number in the Frobenius norm $\|\cdot\|_F$ for Vandermonde matrices associated with families of orthonormal polynomials. Since the diagonal matrix in Theorem 4.2 contains the reciprocal of the L^2 norms of the Legendre polynomials, we can adapt his arguments to bound the Frobenius norm of $(\widehat{V}^n(\mathbf{x}) \text{diag}(\sqrt{2j+1})_{j=0}^n)^{-1}$. In addition, the Legendre polynomials have the special property that $|L^j(x)| \leq 1$ for all $0 \leq x \leq 1$, and so we can bound the Frobenius norm of $\widehat{V}^n(\mathbf{x}) \text{diag}(\sqrt{2j+1})_{j=0}^n$.

For each integer $0 \leq j \leq n$, let $\ell^{j,n}$ be the j th Lagrange polynomial with respect to \mathbf{x} ; that is,

$$(4.18) \quad \ell^{j,n}(x) = \prod_{i \in \{0, \dots, n\} \setminus \{j\}} \frac{x - \mathbf{x}_i}{\mathbf{x}_j - \mathbf{x}_i}.$$

Define $\mathbf{w}^n \in \mathbb{R}^{n+1}$ by

$$(4.19) \quad \mathbf{w}_j^n = \|\ell^{j,n}\|_{L^2}.$$

COROLLARY 4.3.

$$(4.20) \quad \kappa_{M^n \rightarrow 2}(V^n(\mathbf{x})) \leq (n+1)^{3/2} \|\mathbf{w}^n\|_2.$$

Proof. Since $|L^j(x)| \leq 1$ for all $0 \leq x \leq 1$, we have that

$$\begin{aligned} \left\| \widehat{V}^n(\mathbf{x}) \text{diag} \left(\sqrt{2j+1} \right)_{j=0}^n \right\|_F &= \left(\sum_{i,j=0}^n (2j+1) (L^j(\mathbf{x}_i))^2 \right)^{1/2} \\ &\leq \left(\sum_{i,j=0}^n (2j+1) \right)^{1/2} \\ &= (n+1)^{3/2}. \end{aligned}$$

Since

$$(4.21) \quad \ell^{j,n}(\mathbf{x}_i) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

we have that $(\widehat{V}^n(\mathbf{x}))_{ij}^{-1} = \Theta(\ell^{j,n})_i$. Therefore, by (4.12), we have that

$$\begin{aligned} \|\mathbf{w}^n\|_2 &= \left(\sum_{j=0}^n (\mathbf{w}_j^n)^2 \right)^{1/2} \\ &= \left(\sum_{j=0}^n \int_0^1 (\ell^{j,n}(x))^2 dx \right)^{1/2} \\ &= \left(\sum_{j=0}^n \int_0^1 \sum_{i,k=0}^n \Theta(\ell^{j,n})_i \Theta(\ell^{j,n})_k L^i(x) L^k(x) dx \right)^{1/2} \\ &= \left(\sum_{j=0}^n \sum_{i,k=0}^n \Theta(\ell^{j,n})_i \Theta(\ell^{j,n})_k \int_0^1 L^i(x) L^k(x) dx \right)^{1/2} \\ &= \left(\sum_{i,j=0}^n \Theta(\ell^{j,n})_i^2 \int_0^1 (L^i(x))^2 dx \right)^{1/2} \\ &= \left(\sum_{i,j=0}^n \frac{\Theta(\ell^{j,n})_i^2}{2i+1} \right)^{1/2} \\ &= \left\| \text{diag} \left(\frac{1}{\sqrt{2i+1}} \right)_{i=0}^n \left(\widehat{V}^n(\mathbf{x}) \right)^{-1} \right\|_F \\ &= \left\| \left(\widehat{V}^n(\mathbf{x}) \text{diag} \left(\sqrt{2j+1} \right)_{j=0}^n \right)^{-1} \right\|_F. \end{aligned}$$

Since the condition number in the 2 norm is bounded above by the condition number in the Frobenius norm, the result follows from Theorem 4.2. \square

Theorem 4.2 relates the condition number of the Bernstein–Vandermonde matrix in the $M^n \rightarrow 2$ norm to the 2-norm condition number of the (normalized) Legendre–Vandermonde matrix. Since the 2-norm conditioning of the Vandermonde matrix is better in the Legendre basis than the Bernstein basis [10], this result indicates that, when measuring the relevant L^2 norm rather than the Euclidean norm of the solution process, we can expect much better results than the ill-conditioning of the Bernstein–Vandermonde matrix suggests. Corollary 4.3 gives an estimate of this condition number that can be numerically evaluated. A plot of the condition numbers (and the estimate) up to degree 20 is given in Figure 1.

5. Higher dimensions. Now, in the case of equispaced nodes on the d -simplex, we can use block-recursive structure as discussed in [15] to give low-complexity algorithms for the inversion of simplicial Bernstein–Vandermonde matrices. This gives, at least for the equispaced lattice, an algorithm for solving simplicial Bernstein interpolation problems that is an alternate to the one worked out in [1].

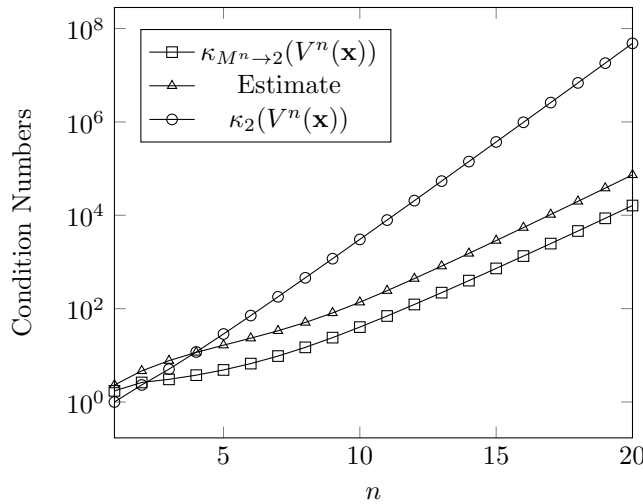


FIG. 1. Comparison of the condition number of V^n in the $M^n \rightarrow 2$ norm and the 2 norm for $1 \leq n \leq 20$, where V^n is the matrix described in subsection 2.3. We also include the estimate given in Corollary 4.3. We use Theorem 4.2 to compute $\kappa_{M^n \rightarrow 2}(V^n(\mathbf{x}))$.

5.1. Notation and preliminaries. For an integer $d \geq 1$, let S_d be a nondegenerate simplex in \mathbb{R}^d . Let $\{\mathbf{v}_i\}_{i=0}^d \subset \mathbb{R}^d$ be the vertices of S_d , and let $\{\mathbf{b}_i\}_{i=0}^d$ denote the barycentric coordinates of S_d . Each \mathbf{b}_i is an affine map from \mathbb{R}^d to \mathbb{R} such that

$$(5.1) \quad \mathbf{b}_i(\mathbf{v}_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j \end{cases}$$

for each vertex \mathbf{v}_j . Each \mathbf{b}_i is nonnegative on S_d , and

$$(5.2) \quad \sum_{i=0}^d \mathbf{b}_i = 1.$$

A multi-index $\boldsymbol{\beta}$ of length $d+1$ is a $(d+1)$ -tuple of nonnegative integers, written as

$$(5.3) \quad \boldsymbol{\beta} = (\beta_0, \dots, \beta_d).$$

The order of $\boldsymbol{\beta}$, denoted by $|\boldsymbol{\beta}|$, is given by

$$(5.4) \quad |\boldsymbol{\beta}| = \sum_{i=0}^d \beta_i.$$

The factorial $\boldsymbol{\beta}!$ of a multi-index $\boldsymbol{\beta}$ is defined by

$$(5.5) \quad \boldsymbol{\beta}! = \prod_{i=0}^d \beta_i.$$

For a multi-index $\boldsymbol{\beta}$ of length $d+1$, denote by $\boldsymbol{\beta}'$ the multi-index of length d given by

$$(5.6) \quad \boldsymbol{\beta}' = (\beta_1, \dots, \beta_d).$$

Given a nonnegative integer b and a multi-index $\beta' = (\beta_1, \dots, \beta_d)$ of length d , define a new multi-index $b \vdash \beta$ of length $d + 1$ by

$$(5.7) \quad b \vdash \beta' = (b, \beta_1, \dots, \beta_d).$$

In particular,

$$(5.8) \quad \beta = \beta_0 \vdash \beta'.$$

Multi-indices have a natural partial ordering given by

$$(5.9) \quad \beta \leq \tilde{\beta} \quad \text{if and only if} \quad \beta_i \leq \tilde{\beta}_i \quad \text{for all} \quad 0 \leq i \leq d.$$

The Bernstein polynomials of degree n on the d -simplex S_d are defined by

$$(5.10) \quad B_{\beta}^n = \frac{n!}{\beta!} \prod_{i=0}^d \mathbf{b}_i^{\beta_i}.$$

The complete set of Bernstein polynomials $\{B_{\beta}^n\}_{|\beta|=n}$ forms a basis for polynomials in d variables of complete degree at most n .

If $n_0 \leq n$, then any polynomial expressed in the basis $\{B_{\beta}^{n_0}\}_{|\beta|=n_0}$ can also be expressed in the basis $\{B_{\beta}^n\}_{|\beta|=n}$. We denote by $E^{d,n_0,n}$ the $\binom{n+d}{d} \times \binom{n_0+d}{d}$ matrix that maps the coefficients of the degree n_0 representation to the coefficients of the degree n representation. The matrix $E^{d,n_0,n}$ is sparse and can be applied matrix-free [14], if desired.

For a nonnegative integer m and a set of distinct nodes $\{\mathbf{x}_{\alpha}^m\}_{|\alpha|=m} \subset S_d$, define the Bernstein–Vandermonde matrix $V^{d,m,n}$ to be the $\binom{m+d}{d} \times \binom{n+d}{d}$ matrix given by

$$(5.11) \quad V_{\alpha\beta}^{d,m,n} = B_{\beta}^n(\mathbf{x}_{\alpha}^m).$$

THEOREM 5.1. *Let d, m, n_0 , and n be nonnegative integers with $d \geq 1$ and $n_0 \leq n$, and let $\{\mathbf{x}_{\alpha}^m\}_{|\alpha|=m}$ be distinct nodes contained in the d -simplex S_d . Then*

$$(5.12) \quad V^{d,m,n} E^{d,n_0,n} = V^{d,m,n_0}.$$

Proof. Given a polynomial $p \in \text{span}\{B_{\beta}^{n_0}\}_{|\beta|=n_0}$, the matrix V^{d,m,n_0} evaluates p at each of the nodes \mathbf{x}_{α}^m . On the other hand, the matrix $V^{d,m,n} E^{d,n_0,n}$ evaluates the degree n representation of p at each of the nodes \mathbf{x}_{α}^m . Since we are evaluating at the same nodes and the polynomial has not been modified, both evaluations must give the same result. \square

The partial ordering (5.9) implies a natural way to order the entries of $V^{d,m,n}$ and also imposes a block structure on $V^{d,m,n}$ by dividing the matrix into sections where α_0 and β_0 are constant. For integers $0 \leq \alpha_0 \leq m$ and $0 \leq \beta_0 \leq n$, let $V_{\alpha_0\beta_0}^{d,m,n}$ denote the $\binom{m-\alpha_0+d}{d} \times \binom{n-\beta_0+d}{d}$ submatrix of $V^{d,m,n}$ whose entries satisfy

$$(5.13) \quad \left(V_{\alpha_0\beta_0}^{d,m,n}\right)_{\alpha'\beta'} = V_{(\alpha_0\vdash\alpha')(\beta_0\vdash\beta')}^{d,m,n}.$$

5.2. Degree reduction. We now consider the case of equispaced nodes on the d -simplex S^d . For this case, the entries of the Bernstein–Vandermonde matrix are given by

$$(5.14) \quad V_{\alpha\beta}^{d,m,n} = \frac{n!}{\beta!} \frac{1}{m^n} \prod_{i=0}^d \alpha_i^{\beta_i}.$$

This representation of the entries allows us to represent the submatrix $V_{\alpha_0\beta_0}^{d,m,n}$ in terms of a lower-dimensional Bernstein–Vandermonde matrix.

THEOREM 5.2. *Let $d > 1$ and $m, n \geq 0$ be integers, and let α_0, β_0 be integers with $0 \leq \alpha_0 \leq m$ and $0 \leq \beta_0 \leq n$. Then*

$$(5.15) \quad V_{\alpha_0\beta_0}^{d,m,n} = m_{\alpha_0\beta_0} V^{d-1,m-\alpha_0,n-\beta_0},$$

where

$$(5.16) \quad m_{\alpha_0\beta_0} = \binom{n}{\beta_0} (\alpha_0/m)^{\beta_0} (1 - \alpha_0/m)^{n-\beta_0}.$$

Proof. If $\alpha_0 = m$ and $\beta_0 \neq n$, then both sides equal zero and we are done. Otherwise, by (5.14), we have that

$$(5.17) \quad V_{\alpha'\beta'}^{d-1,m-\alpha_0,n-\beta_0} = \frac{(n-\beta_0)!}{\beta'!} \frac{1}{(m-\alpha_0)^{n-\beta_0}} \prod_{i=1}^d \alpha_i^{\beta_i}.$$

On the other hand,

$$\begin{aligned} \left(V_{\alpha_0\beta_0}^{d,m,n}\right)_{\alpha'\beta'} &= V_{\alpha\beta}^{d,m,n} \\ &= \frac{n!}{\beta!} \frac{1}{m^n} \prod_{i=0}^d \alpha_i^{\beta_i} \\ &= \frac{n!}{\beta_0!(n-\beta_0)!} \frac{(n-\beta_0)!}{\beta'!} \frac{\alpha_0^{\beta_0}}{m^{\beta_0}} \frac{(m-\alpha_0)^{n-\beta_0}}{m^{n-\beta_0}} \frac{1}{(m-\alpha_0)^{n-\beta_0}} \prod_{i=1}^d \alpha_i^{\beta_i}, \end{aligned}$$

where we have separated terms containing α_0 and β_0 and multiplied and divided by $(n-\beta_0)!$ and $(m-\alpha_0)^{n-\beta_0}$. Since

$$(5.18) \quad \frac{n!}{\beta_0!(n-\beta_0)!} \frac{\alpha_0^{\beta_0}}{m^{\beta_0}} \frac{(m-\alpha_0)^{n-\beta_0}}{m^{n-\beta_0}} = \binom{n}{\beta_0} (\alpha_0/m)^{\beta_0} (1 - \alpha_0/m)^{n-\beta_0},$$

we have the desired result. \square

The fact that the blocks are multiples of lower-dimensional matrices is analogous to what has been observed for other matrices related to Bernstein polynomials. For example, the Bernstein–Vandermonde matrix associated with the zeros of Legendre polynomials has a similar structure [15], and so does the Bernstein mass matrix [14].

We recognize that $m_{\alpha_0\beta_0}$ describes a univariate Bernstein polynomial of degree n being evaluated at equispaced points. Therefore, if V^n is the one-dimensional Bernstein–Vandermonde matrix associated with equispaced nodes as described in subsection 2.3, then we have that $V^{d,n,n}$ admits the block structure

$$(5.19) \quad V^{d,n,n} = \begin{pmatrix} V_{00}^n V^{d-1,n,n} & V_{01}^n V^{d-1,n,n-1} & \cdots & V_{0n}^n V^{d-1,n,0} \\ V_{10}^n V^{d-1,n-1,n} & V_{11}^n V^{d-1,n-1,n-1} & \cdots & V_{1n}^n V^{d-1,n-1,0} \\ \vdots & \vdots & \ddots & \vdots \\ V_{n0}^n V^{d-1,0,n} & V_{n1}^n V^{d-1,0,n-1} & \cdots & V_{nn}^n V^{d-1,0,0} \end{pmatrix}.$$

In [17], Marco, Martínez, and Viaña considered the bivariate interpolation problem with Lagrange-type data. They showed that when using the bivariate tensor

product Bernstein basis, the corresponding coefficient matrix can be expressed as the Kronecker product of univariate Bernstein–Vandermonde matrices. While the case of equispaced nodes on the d -simplex does not fit in this framework, it is interesting to observe the similarity in the structure of (5.19) and the Kronecker product given in [17].

We can use Theorem 5.1 and the block structure given in (5.19) to perform block Gaussian elimination on $V^{d,n,n}$. This method was used in [14] to obtain the block LU decomposition of the Bernstein mass matrix. In the same way, we have the following.

THEOREM 5.3. *Let V^n be the one-dimensional Bernstein–Vandermonde matrix associated with equispaced nodes as described in subsection 2.3. Suppose $V^n = L^n U^n$ is the LU decomposition of V^n . Then*

$$(5.20) \quad V^{d,n,n} = L^{d,n} U^{d,n},$$

where $L^{d,n}$ is the block lower triangular matrix with blocks given by

$$(5.21) \quad L_{\alpha_0 \beta_0}^{d,n} = L_{\alpha_0 \beta_0}^n V^{d-1,n-\alpha_0,n-\beta_0}$$

and $U^{d,n}$ is the block upper triangular matrix with blocks given by

$$(5.22) \quad U_{\alpha_0 \beta_0}^{d,n} = U_{\alpha_0 \beta_0}^n E^{d-1,n-\beta_0,n-\alpha_0}.$$

Proof. By Theorem 5.1, if $0 \leq \alpha_0, \beta_0 \leq n$, and $\gamma_0 \leq \beta_0$, then

$$\begin{aligned} L_{\alpha_0 \gamma_0}^{d,n} U_{\gamma_0 \beta_0}^{d,n} &= L_{\alpha_0 \gamma_0}^n U_{\gamma_0 \beta_0}^n V^{d-1,n-\alpha_0,n-\gamma_0} E^{d-1,n-\beta_0,n-\gamma_0} \\ &= L_{\alpha_0 \gamma_0}^n U_{\gamma_0 \beta_0}^n V^{d-1,n-\alpha_0,n-\beta_0}. \end{aligned}$$

Therefore,

$$\begin{aligned} (L^{d,n} U^{d,n})_{\alpha_0 \beta_0} &= \sum_{\gamma_0=0}^{\beta_0} L_{\alpha_0 \gamma_0}^{d,n} U_{\gamma_0 \beta_0}^{d,n} \\ &= \left(\sum_{\gamma_0=0}^{\beta_0} L_{\alpha_0 \gamma_0}^n U_{\gamma_0 \beta_0}^n \right) V^{d-1,n-\alpha_0,n-\beta_0} \\ &= V_{\alpha_0 \beta_0}^n V^{d-1,n-\alpha_0,n-\beta_0} \\ &= V_{\alpha_0 \beta_0}^{d,n,n}. \end{aligned} \quad \square$$

6. Numerical results. Now, we consider the accuracy of the methods described in section 3 and subsection 5.2 on several problems. For all of the problems, we chose random solution vectors, computed the right-hand side by matrix multiplication, and then attempted to recover the solution. The numerical results are run in double precision arithmetic on a 2014 Macbook Air running macOS 10.13 and using Python 2.7.13. LU factorization and FFTs are performed using the `numpy` (v1.12.0) function calls. Also, because our code is a mix of pure Python and low-level compiled libraries, timings are not terribly informative. Consequently, we focus on assessing the stability and accuracy of our methods. If future work leads to more stable fast algorithms, greater care will be afforded to tuning our implementations for performance.

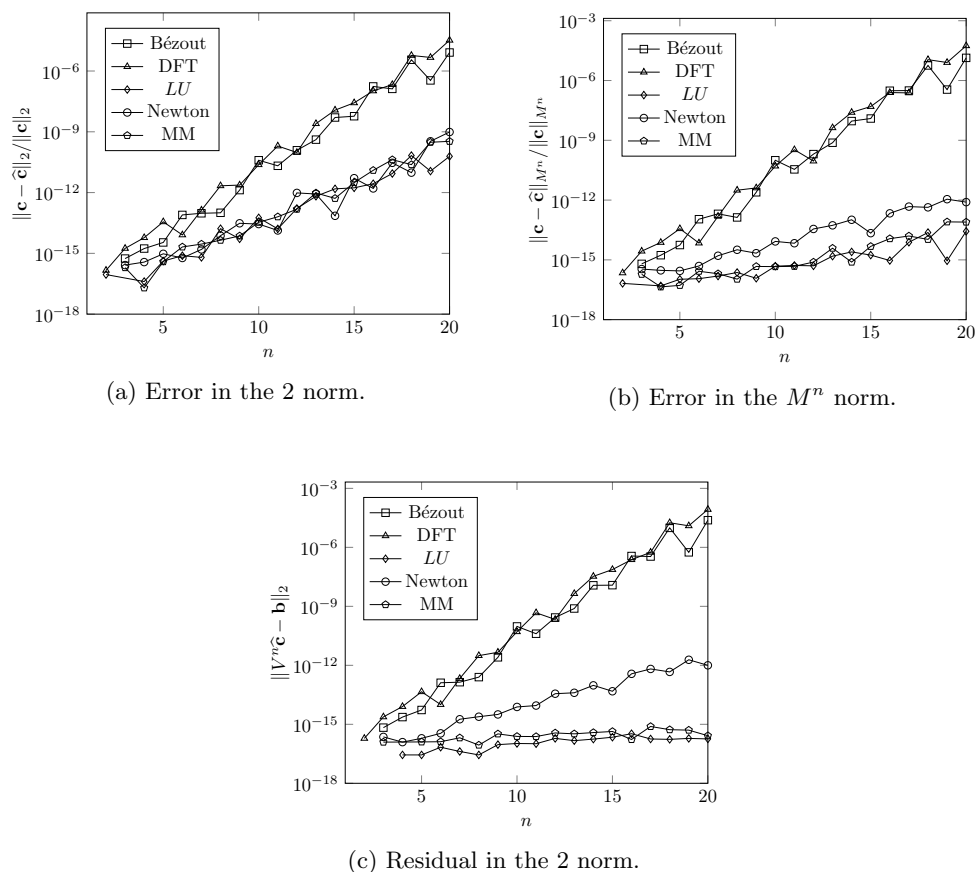


FIG. 2. Error/residual in using the methods described in section 3 to solve $V^n \mathbf{c} = \mathbf{b}$ for $1 \leq n \leq 20$, where V^n is the matrix described in subsection 2.3 and \mathbf{b} is a random vector in $[-1, 1]^{n+1}$. Bézout refers to Corollary 2.2, DFT refers to Corollary 2.9, LU refers to LU decomposition of V^n , Newton refers to the Ainsworth–Sanchez algorithm, and MM refers to the Marco–Martínez algorithm. We use $\hat{\mathbf{c}}$ to denote the computed solution.

In Figure 2, we consider the case of equispaced nodes. When considering the Euclidean norm of the error (Figure 2a), the LU, Newton, and MM algorithms have the best performance, followed by the DFT-based application and multiplying by the inverse. We observe that the LU, Newton, and MM methods have comparable performance, as do the DFT-based algorithm and multiplying by the inverse. This is in contrast to what was observed in [3], where a similar DFT-based algorithm quickly became unstable. The same behavior can be observed when comparing the Euclidean norm of the residual (Figure 2c), although some separation does occur between the Newton method and the LU and MM algorithms. Since the solution vector can be viewed as the Bernstein coefficients of the interpolation polynomial, we also measure the L^2 difference between the exact and computed solutions (Figure 2b); equivalently, we measure the relative M^n error, where M^n is the Bernstein mass matrix given in (4.1). All five solution methods have very similar behavior in the M^n norm, as they do when considering the Euclidean norm of the residual.

To ensure that the behavior of the solution methods does not depend on the

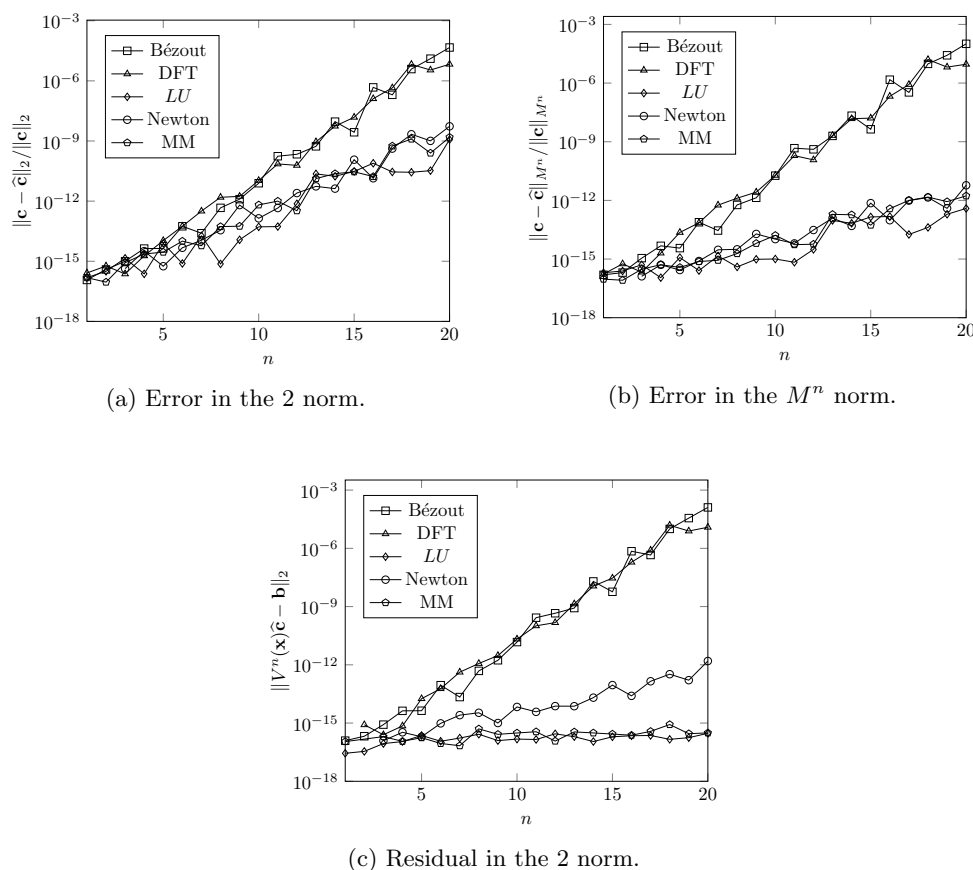


FIG. 3. Error/residual in using the methods described in section 3 to solve $V^n(\mathbf{x})\mathbf{c} = \mathbf{b}$ for $1 \leq n \leq 20$, where $V^n(\mathbf{x})$ is the Bernstein–Vandermonde matrix associated with \mathbf{x} , the nodes \mathbf{x}_j are randomly selected from $[j/(n+1), (j+1)/(n+1))$ for $0 \leq j \leq n$, and \mathbf{b} is a random vector in $[-1, 1]^{n+1}$. Bézout refers to Corollary 2.2, DFT refers to Theorem 2.8, LU refers to LU decomposition of V^n , Newton refers to the Ainsworth–Sanchez algorithm, and MM refers to the Marco–Martínez algorithm. We use $\hat{\mathbf{c}}$ to denote the computed solution.

choice of nodes, we also considered the case where the nodes \mathbf{x}_j are randomly selected from $[j/(n+1), (j+1)/(n+1))$ for each $0 \leq j \leq n$ (Figure 3); however, there was no significant difference in the quantities measured between this case and the equispaced case. In addition, we consider a case where the values of the right-hand side are all nonnegative (Figure 4). We see some slight improvement for the DFT and Bézout methods when considering the error in the Euclidean and M^n norms, but overall, the methods have performance comparable to the previous case.

We also used the block LU decomposition given in Theorem 5.3 combined with the one-dimensional LU algorithm to solve the interpolation problem for equispaced nodes on the d -simplex for $d = 2$ and $d = 3$ (Figure 5). Due to the recursive nature of the algorithm, the quantities measured are very similar to the ones observed for equispaced nodes.

7. Conclusion. We have studied several algorithms for the inversion of the univariate Bernstein–Vandermonde matrix. These algorithms, while less stable than

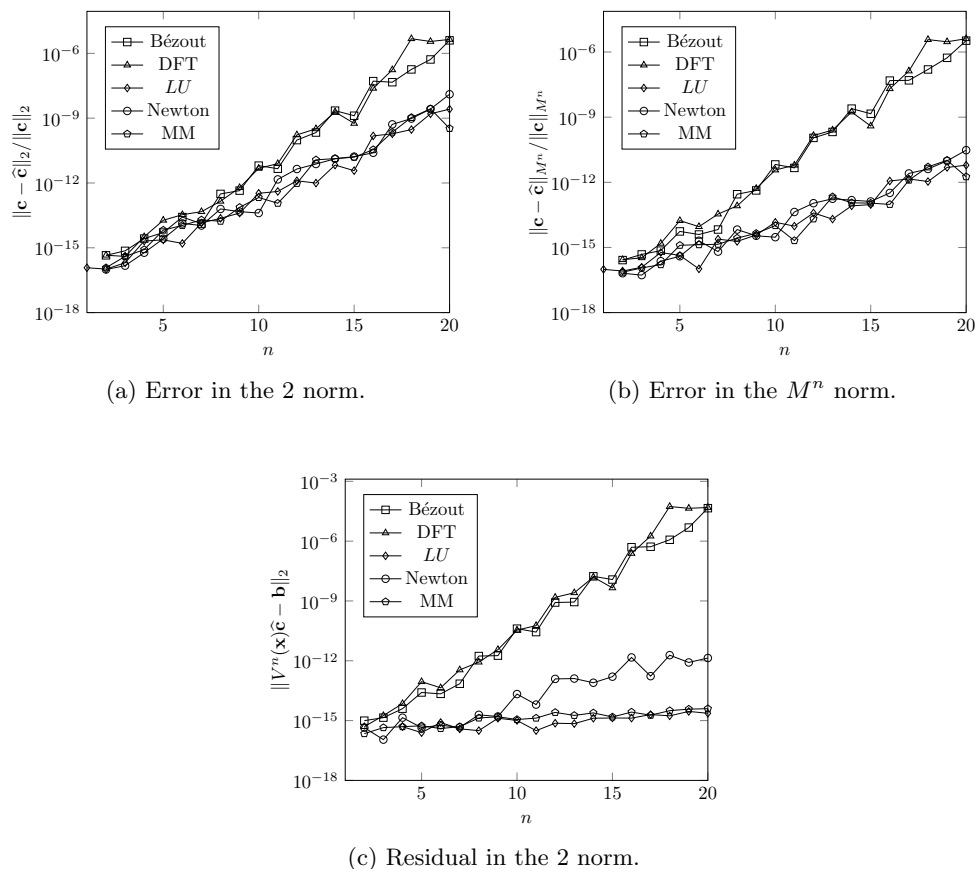


FIG. 4. Error/residual in using the methods described in section 3 to solve $V^n(\mathbf{x})\mathbf{c} = \mathbf{b}$ for $1 \leq n \leq 20$, where $V^n(\mathbf{x})$ is the Bernstein–Vandermonde matrix associated with \mathbf{x} , the nodes \mathbf{x}_j are randomly selected from $[j/(n+1), (j+1)/(n+1))$ for $0 \leq j \leq n$, and \mathbf{b} is a random vector in $[0, 4]^{n+1}$. Bézout refers to Corollary 2.2, DFT refers to Theorem 2.8, LU refers to LU decomposition of V^n , Newton refers to the Ainsworth–Sanchez algorithm, and MM refers to the Marco–Martínez algorithm. We use $\hat{\mathbf{c}}$ to denote the computed solution.

algorithms discovered previously, provide insight into the structure of the Bernstein–Vandermonde matrix and are remarkably similar to algorithms derived for the Bernstein mass matrix. Moreover, the derivation of the algorithm is not restricted to the Bernstein basis and can be applied to other polynomial bases and used to investigate algorithms for the corresponding interpolation problems. In addition, we have used a block LU decomposition of the Bernstein–Vandermonde matrix corresponding to equispaced nodes on the d -simplex to give a recursive, block-structured algorithm with accuracy comparable to the one-dimensional algorithm. We have also given a new perspective on the conditioning of the Bernstein–Vandermonde matrix, indicating that the interpolation problem is better-conditioned with respect to the L^2 norm than the Euclidean norm. In the future, we hope to expand this perspective to other polynomial problems and continue the development of fast and accurate methods for problems involving Bernstein polynomials.

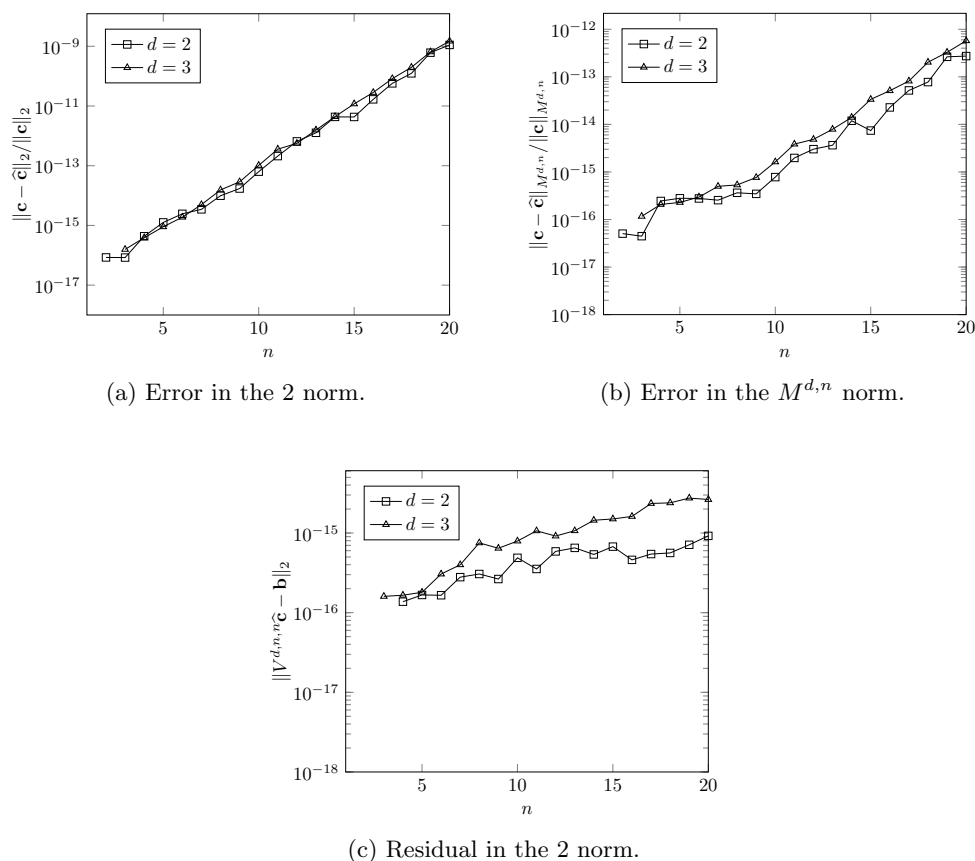


FIG. 5. Error/residual in using the block LU decomposition given in Theorem 5.3 to solve $V^{d,n,n}\mathbf{c} = \mathbf{b}$ for $1 \leq n \leq 20$ and $d = 2, 3$, where $V^{d,n,n}$ is the Bernstein–Vandermonde matrix given in (5.14) and \mathbf{b} is a random vector in $[-1, 1]^{\binom{n+d}{d}}$. We use $\hat{\mathbf{c}}$ to denote the computed solution.

REFERENCES

- [1] M. AINSWORTH, G. ANDRIAMARO, AND O. DAVYDOV, *Bernstein–Bézier finite elements of arbitrary order and optimal assembly procedures*, SIAM J. Sci. Comput., 33 (2011), pp. 3087–3109, <https://doi.org/10.1137/11082539X>.
- [2] M. AINSWORTH AND M. A. SÁNCHEZ, *Computing the Bézier control points of the Lagrangian interpolant in arbitrary dimension*, SIAM J. Sci. Comput., 38 (2016), pp. A1682–A1700, <https://doi.org/10.1137/15M1046113>.
- [3] L. ALLEN AND R. C. KIRBY, *Structured inversion of the Bernstein mass matrix*, SIAM J. Matrix Anal. Appl., 41 (2020), pp. 413–431, <https://doi.org/10.1137/19M1284166>.
- [4] R. L. BURDEN AND J. D. FAIRES, *Numerical Methods*, 4th ed., Cengage Learning, Boston, MA, 2012.
- [5] P. J. DAVIS, *Interpolation and Approximation*, Dover, New York, 1975.
- [6] J. DELGADO AND J. M. PEÑA, *Optimal conditioning of Bernstein collocation matrices*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 990–996, <https://doi.org/10.1137/080737976>.
- [7] M. G. DUFFY, *Quadrature over a pyramid or cube of integrands with a singularity at a vertex*, SIAM J. Numer. Anal., 19 (1982), pp. 1260–1262, <https://doi.org/10.1137/0719090>.
- [8] R. T. FAROUKI, *Legendre–Bernstein basis transformations*, J. Comput. Appl. Math., 119 (2000), pp. 145–160.
- [9] M. GASCA AND J. M. PEÑA, *Total positivity and Neville elimination*, Linear Algebra Appl., 165 (1992), pp. 25–44.

- [10] W. GAUTSCHI, *The condition of Vandermonde-like matrices involving orthogonal polynomials*, Linear Algebra Appl., 52/53 (1983), pp. 293–300.
- [11] W. GAUTSCHI, G. H. GOLUB, AND G. OPFER, EDS., *Applications and Computation of Orthogonal Polynomials*, Springer, Basel, 1999.
- [12] G. HEINIG AND K. ROST, *Algebraic Methods for Toeplitz-Like Matrices and Operators*, Oper. Theory Adv. Appl. 13, Birkhäuser, Basel, 1984.
- [13] S. KAPLAN, A. SHAPIRO, AND M. TEICHER, *Several Applications of Bézout Matrices*, preprint, <https://arxiv.org/abs/math/0601047>, 2006.
- [14] R. C. KIRBY, *Fast inversion of the simplicial Bernstein mass matrix*, Numer. Math., 135 (2017), pp. 73–95.
- [15] R. C. KIRBY AND K. T. THINH, *Fast simplicial quadrature-based finite element operators using Bernstein polynomials*, Numer. Math., 121 (2012), pp. 261–279.
- [16] A. MARCO AND J.-J. MARTÍNEZ, *A fast and accurate algorithm for solving Bernstein–Vandermonde linear systems*, Linear Algebra Appl., 422 (2007), pp. 616–628.
- [17] A. MARCO, J.-J. MARTÍNEZ, AND R. VIAÑA, *Accurate polynomial interpolation by using the Bernstein basis*, Numer. Algorithms, 75 (2017), pp. 655–674.
- [18] J. J. ROTMAN, *Advanced Modern Algebra*, 3rd ed., American Mathematical Society, Providence, RI, 2015.
- [19] I. A. STEGUN, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, U.S. Government Printing Office, Washington, D.C., 1965.
- [20] C. R. VOGEL, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002, <https://doi.org/10.1137/1.9780898717570>.