Matrix factorization of multivariate Bernstein polynomials

Rune Dalmo

Faculty of Technology
Narvik University College
PO Box 385, 8505 Narvik, NORWAY

2015

Abstract

Ordinary univariate Bernstein polynomials can be represented in matrix form using factor matrices. In this paper we present the definition and basic properties of such factor matrices extended from the univariate case to the general case of arbitrary number of variables by using barycentric coordinates in the hyper-simplices of respective dimension. The main results in the paper are related to the design of an iterative algorithm for fast convex computation of multivariate Bernstein polynomials based on sparse-matrix factorization. In the process of derivation of this algorithm, we investigate some properties of the factorization, including symmetry, commutativity and differentiability of the factor matrices, and address the relevance of this factorization to the de Casteljau algorithm for evaluating curves and surfaces on Bézier form. A set of representative examples is provided, including a geometric interpretation of the de Casteljau algorithm, and representation by factor matrices of multivariate surfaces and their derivatives in Bézier form. Another new result is the observation that inverting the order of steps of a part of the new factorization algorithm provides a new, matrix-based, algebraic representation of a multivariate generalization of a special case of the de Boor-Cox computational algorithm.
AMS Subject Classification: 15A23, 15A27, 65D07, 65D17, 65D10, 65F30, 65F50
Keywords and phrases: Bernstein polynomials, matrix factorization, de Casteljau algorithm

1 Introduction

Univariate Bernstein polynomials were introduced by Sergei Bernstein [1] in his proof of the Weierstrass approximation theorem. They have been studied extensively in the literature since then, see, e.g. [2], or a textbook on approximation theory, e.g., [3].

Bernstein polynomials consitute the basis of Bézier curves and surfaces [4], and they are commonly used as basis functions in spline constructions [5, 6]. One way of representing Bernstein polynomials is by matrix form as factor matrices. Such matrices are related to de Casteljau's corner cutting algorithm [7]. The univariate case of Bernstein factor matrices was addressed in [8], however, univariate Bernstein factor matrices seem to have been exposed and used prior to that, see e.g. [9], where the de Casteljau algorithm is expressed on matrix form, or the method presented in [10, algorithm 12]. In addition, we mention that Tom Lyche and Knut Mørken have been using a slightly different matrix representation of B-splines in their lecture notes [11] at the university of Oslo. A particularly attractive property of the de Casteljau algorithm and its matrix-based version

is the convexity of computation, whereby the total accumulated error of computation stays within the convex hull of the errors made in the process of the algorithm's iterations.

The definition of ordinary Bernstein polynomials can be generalized with preservation of the convexity-of-computation property by using barycentric coordinates, see e.g. [12, 13]. In this paper we investigate symmetric and recursive properties of the factor matrices for higher dimensions of the polynomials' argument. Furthermore, we present an equivalent recursive definition which facilitates the construction of Bernstein factor matrices for arbitrary dimensions of the barycentric argument, and investigate some properties which follow from the layout and construction of the matrices. We are interested in these findings mainly because the factorization can be seen to correspond to the de Casteljau algorithm and to a special case of the de Boor-Cox recursion formula for B-splines. Parts of the results, including an outline of the recursive definition of the Bernstein factor matrices, were announced in [14].

The organization of the paper is, as follows. In subsection 1.1 we provide a brief description of Bernstein polynomials, the Bézier representation of polynomial curves, and univariate Bernstein factor matrices. Subsection 1.2 is concerning some properties of the algebra of square matrices. We generalize the factor matrices to the multivariate setting in section 2, via first considering the case of $\mathbb{R}^2$, and then the d-variate case. Then in section 3 we address the directional derivatives of Bernstein polynomials and Bernstein factor matrices in the multivariate setting. Section 4 covers the relevance of the factorization to the de Casteljau algorithm, followed by some examples of how the construction can be used. Proofs of some of the lemmas and theorems are provided in section 5. Finally, in section 6, we give our concluding remarks where we suggest some topics for future work.

1.1 Univariate Bernstein polynomials and curves on Bézier form

Computing the binomial expansion

$$1 = (t + (1-t))^n = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i},$$

where $t \in \mathbb{R}$, leads to the Bernstein polynomials [1] of degree n,

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i}, \quad i = 0, \ldots, n \ , \tag{1}$$

which satisfy a number of important properties, including linear independence, symmetry, roots at 0 and 1 only, partition of unity and positivity (i.e., convex partition of unity) in $(0, 1)$. They can be expressed through the following (convex) recursion formula:

$$B_i^n(t) = tB_{i-1}^{n-1}(t) + (1-t)B_i^{n-1}(t) \ , \tag{2}$$

where $B_{-1}^{n-1} = B_n^{n-1} = 0$ and $B_0^0 = 1$.

Since the $n+1$ linearly independent Bernstein polynomials B_i^n form a basis for all polynomials of degree $\leq n$, any polynomial curve $c(t)$ of degree $\leq n$ has a unique n-th degree Bernstein-Bézier representation

$$c(t) = \sum_{i=0}^{n} \mathbf{c}_i B_i^n(t), \tag{3}$$

where the coefficients $\mathbf{c}_i \in \mathbb{R}^d$, called Bézier points [13], are elements of an affine space, i.e., the elements of $\mathbb{R}^d$ are points with coordinates defined with respect to a frame, consisting of origin with coordinates $(\underbrace{0, \ldots, 0}_{d \text{ times}})$ and the canonical orthonormal basis with coordinates $(\underbrace{0, \ldots, 0}_{i-1 \text{ times}}, 1, \underbrace{0, \ldots, 0}_{n-i-1 \text{ times}})$.

We note that the Bernstein-Bézier representation is sometimes referred to as the B-form, as suggested by Carl de Boor in [15].

The curve in (3) can be evaluated using de Casteljau's corner cutting algorithm [7] by letting $\mathbf{c}_i^0 = \mathbf{c}_i$ and repeatedly applying the recursion formula for Bernstein polynomials in (2):

$$c(t) = \sum_{i=0}^{n} \mathbf{c}_i^0 B_i^n(t) = \sum_{i=0}^{n-1} \mathbf{c}_i^1 B_i^{n-1}(t) = \cdots = \sum_{i=0}^{0} \mathbf{c}_i^n B_i^0(t) = \mathbf{c}_0^n \ ,$$

where $\mathbf{c}_i^k = (1-t)\mathbf{c}_i^{k-1} + t\mathbf{c}_{i+1}^{k-1}$ are intermediate points of de Casteljau's algorithm. This recursive convex linear interpolation between two points can be expressed in matrix form as outlined in [8]. As an example, computing from right to left, for $n = 3$:

$$c(t) = \begin{pmatrix} 1-t & t \end{pmatrix} \begin{pmatrix} 1-t & t & 0 \\ 0 & 1-t & t \end{pmatrix} \begin{pmatrix} 1-t & t & 0 & 0 \\ 0 & 1-t & t & 0 \\ 0 & 0 & 1-t & t \end{pmatrix} \begin{pmatrix} \mathbf{c}_0 \\ \mathbf{c}_1 \\ \mathbf{c}_2 \\ \mathbf{c}_3 \end{pmatrix}. \tag{4}$$

By consecutively multiplying the matrices in (4) from right to left, the dimension of the vector on the RHS is reduced by 1 for each computation. The factor matrices are of dimension $n \times (n+1)$ given by their number of rows and columns respectively. It follows that (3) can be rewritten as

$$c(t) = \mathbf{T}_1(t)\mathbf{T}_2(t)\mathbf{T}_3(t)\mathbf{c}, \tag{5}$$

where $\mathbf{c} = (\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)^T$ and the factor matrices are denoted by $\mathbf{T}_n(t)$. As noted in [10], computing the three matrices from the left results in a vector containing the four Bernstein polynomials of degree three:

$$\begin{aligned} \mathbf{T}_1(t)\mathbf{T}_2(t)\mathbf{T}_3(t) &= \begin{pmatrix} (1-t)^3 & 3t(1-t)^2 & 3t^2(1-t) & t^3 \end{pmatrix} \\ &= \begin{pmatrix} B_0^3(t) & B_1^3(t) & B_2^3(t) & B_3^3(t) \end{pmatrix}. \end{aligned}$$

This method can be seen as a special case of the de Boor-Cox recursion formula [16, 17] for B-splines, since B-splines are the proper generalization of Bézier curves [18].

The curve $c(t)$ can be evaluated by multiplying the vector of Bernstein polynomials together with the vector of coefficients $\mathbf{c}$. We introduce the following matrix notation for the set of Bernstein polynomials of degree n:

$$\mathbf{B}^n(t) = \mathbf{T}_1(t)\mathbf{T}_2(t)\cdots\mathbf{T}_n(t). \tag{6}$$

Thus, by applying (6) to (5), the cubic Bézier curve in (4) can be expressed as

$$c(t) = \mathbf{B}^3(t)\mathbf{c}.$$

The binomial in (4) can be expressed using barycentric coordinates simply by substituting $1-t$ with v and t with w, so that (v, w) is a barycentric coordinate with respect to a line segment $L = \begin{pmatrix} \mathbf{l}_0 & \mathbf{l}_1 \end{pmatrix} \subset \mathbb{R}^1$, where $v + w = 1$. The first three matrices of (4) would then become

$$\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3 = \begin{pmatrix} v & w \end{pmatrix} \begin{pmatrix} v & w & 0 \\ 0 & v & w \end{pmatrix} \begin{pmatrix} v & w & 0 & 0 \\ 0 & v & w & 0 \\ 0 & 0 & v & w \end{pmatrix}. \tag{7}$$

Prior to introducing multivariate Bernstein polynomials, we provide a brief description of square matrices and some of their properties, which we shall use in some of the proofs and examples in the sequel of this paper.

1.2 The algebra of square matrices

N-th order square matrices form a vector space with respect to the ordinary summation of matrices and multiplication with a scalar [19, §93 (Example 1), §94]. Its dimension is N^2, and one basis in it is provided by the canonical matrices $E_{ij} = (\delta_{ik}\delta_{\ell j})_{k,\ell=1}^{N}$, $i, j = 1, \ldots, N$, where

$$\delta_{ij} = \begin{cases} 1, & \text{if} \quad i = j, \\ 0, & \text{otherwise}, \end{cases}$$

is the Kronecker delta. For the linear span of the so-defined canonical basis we shall use the notation $\mathbf{E} = \mathbf{E}_{N\times N} = \mathbf{E}_{N\times N}(\mathbb{F})$ where $\mathbb{F}$ denotes the field of scalars of the linear span and where, in our case, we shall always consider only the case $\mathbb{F} = \mathbb{R}$ (real scalars). It is well-known, that when endowed with matrix multiplication, $\mathbf{E}_{N\times N}$ becomes a non-commutative algebra [19, §93]. Since in the sequel we shall not be using all the properties of $\mathbf{E}_{N\times N}$ as an algebra (for example, we shall not be discussing inverse elements in $\mathbf{E}_{N\times N}$), our choice here is to provide a self-consistent presentation of only those of the properties of the algebra $\mathbf{E}_{N\times N}$ which are of relevance to this article through the next few lemmas.

Lemma 1. *The dimension of the space $\mathbf{E}_{N\times N}$ is determined by the largest dimension of the matrices involved in a given multiplication. For instance, let*

$$\mathbf{AB} = \mathbf{C},$$

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are matrices with dimensions $K \times L$, $L \times M$ and $K \times M$, respectively. Then $N = \max\{K, L, M\}$, and we can, for instance, express the matrix $\mathbf{A}$ as

$$\mathbf{E}_{N\times N} = \sum_{k=1}^{N}\sum_{l=1}^{N} a_{kl}E_{kl},$$

where a_{kl} are the entries in $\mathbf{A}$, which we define to be zero if $k > K$ or if $l > L$. The result is a space $\mathbf{E}_{N\times N}$ which contains $\mathbf{A}$ as element of the linear span of E_{ij} for the first K values of i and the first L values of j, i.e., corresponding to a $K \times L$-rectangular submatrix situated in the upper left corner of the $N \times N$ square matrix. Entries from $\mathbf{B}$ and $\mathbf{C}$ can be obtained analogously.

Lemma 1 allows us to simplify the considerations for multiplication of several algebrae [19, §94 (item 2)], to a consideration within one single algebra $\mathbf{E}_{N\times N}$ of sufficiently high order N [19, §93 (item 1)].

Lemma 2. *The product of two basis matrices is*

$$E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell}, \tag{8}$$

where $i, j, k, \ell = 1, \ldots, N$.

Formula (8) is a concise form of the respective formula in [19, §93 (item 1)] and a simplification of the respective formulae in [19, §94 (item 2)]. This will prove essential in the proofs of our main results, because there (8) will be used iteratively.

Lemma 3. *Let $\mathbf{A} = (a_{ij})_{i=1,j=1}^{N,N} \in \mathbf{E}_{N\times N}$, $\mathbf{B} = (b_{kl})_{k=1,l=1}^{N,N} \in \mathbf{E}_{N\times N}$, $\mathbf{C} = (c_{mn})_{m=1,n=1}^{N,N} \in \mathbf{E}_{N\times N}$ and $\mathbf{D} = (d_{pq})_{p=1,q=1}^{N,N} \in \mathbf{E}_{N\times N}$ be four matrices. Then $\mathbf{AB} - \mathbf{CD}$ is an element of $\mathbf{E}_{N\times N}$, defined by*

$$\mathbf{AB} - \mathbf{CD} = \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\left[\sum_{\lambda=1}^{N}(a_{\alpha\lambda}b_{\lambda\beta} - c_{\alpha\lambda}d_{\lambda\beta})\right]E_{\alpha\beta}. \tag{9}$$

For the particular case when $\mathbf{C} = \mathbf{B}$, $\mathbf{D} = \mathbf{A}$ in formula (9) we obtain the following:

Corollary 4. *The commutator of* $\mathbf{A} = (a_{ij})_{i=1,j=1}^{N,N} \in \mathbf{E}_{N\times N}$ *and* $\mathbf{B} = (b_{kl})_{k=1,l=1}^{N,N} \in \mathbf{E}_{N\times N}$ *is an element of* $\mathbf{E}_{N\times N}$*, defined by*

$$[\mathbf{A}, \mathbf{B}] = \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\left[\sum_{\lambda=1}^{N}(a_{\alpha\lambda}b_{\lambda\beta} - a_{\lambda\beta}b_{\alpha\lambda})\right]E_{\alpha\beta}.$$

2 Multivariate Bernstein polynomials

The symmetric properties of the Bernstein polynomials suggest that barycentric coordinates are a natural choice to represent multivariate scenarios. We proceed by considering factorization of the Bernstein polynomials in two variables expressed in barycentric coordinates with respect to a triangle in $\mathbb{R}^2$. Then we shall generalize the resulting factor matrices to the multivariate setting, with a recursive definition based on a particular decomposition of the matrices into submatrices.

2.1 Bivariate Bernstein polynomials

Proceeding as in (1), we compute the trinomial expansion

$$1 = (u + v + w)^n = \sum_{i,j,k}\binom{n}{i,j,k}u^i v^j w^k, \tag{10}$$

where $i, j, k \geq 0$ and $i + j + k = n$. This leads to the Bernstein polynomials of degree n,

$$B_{ijk}^n(u, v, w) = \binom{n}{i,j,k}u^i v^j w^k, \tag{11}$$

where (u, v, w) is the barycentric coordinate of a point $\mathbf{p}$ with respect to a triangle $A = (\mathbf{a_0}, \mathbf{a_1}, \mathbf{a_2}) \subset \mathbb{R}^2$, $(i, j, k) \in \{0, 1, \ldots, n\}^3$, and $i + j + k = n$. We note that (u, v, w) represents a local parameter with respect to A and $\mathbf{p}$ a global parameter, since $\mathbf{p} = A\mathbf{u} = \mathbf{a_0}u + \mathbf{a_1}v + \mathbf{a_2}w$, with $u + v + w = 1$.

The recurrence relation for Bernstein polynomials associated with a triangle $A \subset \mathbb{R}^2$ in homogeneous barycentric coordinates is defined (see e.g. [12, 13]) as

$$B_{ijk}^n = uB_{i-1,j,k}^{n-1} + vB_{i,j-1,k}^{n-1} + wB_{i,j,k-1}^{n-1}, \tag{12}$$

where we consider expressions with negative subscripts to be zero. It is well known [12] that the Bernstein polynomials

$$\mathbf{B}_2^n = \{B_{ijk}^n\}_{i+j+k=n} \tag{13}$$

form a basis for the space of polynomials of degree n. Thus, given any triangle A, every polynomial p of degree n has a unique Bernstein-Bézier representation

$$p = \sum_{i+j+k=n} c_{ijk}B_{ijk}^n, \tag{14}$$

where B_{ijk}^n are the Bernstein polynomials associated with A.

In the sequel we shall assume a lexicographic order of the $\binom{n+2}{2}$ Bernstein polynomials, as provided in (13) and (14), such that $B_{rst}^n > B_{ijk}^n$ when $r > i$, or if $r = i$, then $s > j$, or if $r = i$ and $s = j$, then $t > k$. As an example, the order of B_{ijk}^n for $i + j + k = n$ and $n = 3$ is the following:

$$B_{300}^3 > B_{210}^3 > B_{201}^3 > B_{120}^3 > B_{111}^3 > B_{102}^3 > B_{030}^3 > B_{021}^3 > B_{012}^3 > B_{003}^3.$$

We use the recurrence relation in (12) to define the triangular (bivariate) Bernstein factor matrix in homogeneous barycentric coordinates, as follows:

Definition 5. The Bernstein factor matrix $\mathbf{T}_{2,n}(u,v,w)$ is a $\binom{n+1}{2} \times \binom{n+2}{2}$ band-limited matrix with three non-zero elements on each row. The columns of $\mathbf{T}_{2,n}(u,v,w)$ correspond to the $\binom{n+2}{2}$ vectors of terms of the recurrence relation in (12) in lexicographic order, such that the non-zero elements in every column are

1. positioned according to the lexicographic index numbers of the B^{n-1}_{ijk} on the RHS, and
2. taking the values of their associated variables; u, v, or w.

It follows immediately from Definition 5 that the first three triangular (bivariate) factor matrices in homogeneous barycentric coordinates are as follows:

$$\mathbf{T}_{2,1} = \begin{pmatrix} u & v & w \end{pmatrix}, \tag{15}$$

$$\mathbf{T}_{2,2} = \left(\begin{array}{ccc|ccc} u & v & w & 0 & 0 & 0 \\ \hline 0 & u & 0 & v & w & 0 \\ 0 & 0 & u & 0 & v & w \end{array}\right), \tag{16}$$

$$\mathbf{T}_{2,3} = \left(\begin{array}{cccccc|cccc} u & v & w & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & u & 0 & v & w & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & u & 0 & v & w & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & u & 0 & 0 & v & w & 0 & 0 \\ 0 & 0 & 0 & 0 & u & 0 & 0 & v & w & 0 \\ 0 & 0 & 0 & 0 & 0 & u & 0 & 0 & v & w \end{array}\right), \tag{17}$$

where the horizontal and vertical lines are used to annotate sub-matrices, which we shall describe later.

We find it appropriate to include the following lemma, which states that the proposed factor matrices are a factorization of the set of Bernstein polynomials in three variables in barycentric coordinates.

Lemma 6. *The Bernstein factor matrices* $\{\mathbf{T}_{2,i}\}_{i=1}^{n}$ *factor the Bernstein polynomials* $\mathbf{B}_2^n$*:*

$$\mathbf{T}_{2,1}\mathbf{T}_{2,2}\cdots\mathbf{T}_{2,n} = \mathbf{B}_2^n. \tag{18}$$

By investigating the matrices in (7) and (15)-(17) we observe that a specific Bernstein factor matrix consists of four sub-matrices: The upper left sub-matrix is the factor matrix of the same dimension but of one degree lower. The lower right sub-matrix is the factor matrix of one dimension lower but of the same degree. The upper right sub-matrix is a zero matrix. In the lower left sub-matrix the only non-zero elements are on the diagonal from one of the entries on the top to the lower right element. We shall use the term right diagonal, since it can be seen as a main diagonal which is shifted as far as possible to the right in a square matrix.

It turns out that this relation holds for arbitrary degree of the factor matrices. We, therefore, summarize the findings in the following lemma.

Lemma 7. *The factor matrix* $\mathbf{T}_{2,n}(u,v,w)$*, for degree* $n > 1$*, consists of four sub-matrices, such that*

$$\mathbf{T}_{2,n}(u,v,w) = \left(\begin{array}{c|c} \mathbf{T}_{2,n-1}(u,v,w) & 0 \\ \hline \mathbf{T}^{\mathrm{diag}}_{2,n}(u) & \mathbf{T}_{1,n}(v,w) \end{array}\right),$$

where $\mathbf{T}^{\mathrm{diag}}_{2,n}(u)$ *is a matrix where the only non-zero elements are on the right diagonal and equal to* u*.*

2.2 The d-variate case, $d = 2, 3, \ldots$

Multivariate Bernstein polynomials over a d-dimensional simplex $A \subset \mathbb{R}^d$ are defined in [13] as follows:

$$B^n_{\mathbf{i}}(\mathbf{u}) = \binom{n}{\mathbf{i}}\mathbf{u}^{\mathbf{i}} = \binom{n}{i_0, \ldots, i_d} u_0^{i_0} \ldots u_d^{i_d}, \tag{19}$$

where $\mathbf{i} = (i_0, \ldots, i_d) \in \{0, 1, \ldots, n\}^{d+1}, |\mathbf{i}| = i_0 + \cdots + i_d = n$, $\mathbf{u} = (u_0, \ldots, u_d)$ is the barycentric coordinate of a point $\mathbf{p}$ with respect to A, $\mathbf{u}^{\mathbf{i}} = (u_0^{i_0} \ldots u_d^{i_d})$, and $u_0 + \cdots + u_d = 1$. There are $\binom{n+d}{d}$ Bernstein polynomials of degree n. They form a basis for all d-variate polynomials of total degree $\leq n$, they form a partition of unity and are positive for $\mathbf{u} > 0$, which is one reason to consider Bernstein polynomials over the simplex A.

The following recurrence relation holds for the Bernstein polynomials in (19):

$$B^n_{\mathbf{i}}(\mathbf{u}) = u_0 B^{n-1}_{\mathbf{i}-\mathbf{e}_0} + \cdots + u_d B^{n-1}_{\mathbf{i}-\mathbf{e}_d}, \tag{20}$$

where $\mathbf{e}_0, \ldots, \mathbf{e}_d$ represents the columns of the $(d+1) \times (d+1)$ identity matrix, $B^0_{\mathbf{0}} = 1$, $B^0_{\mathbf{i}} = 0$ if $\mathbf{i}$ has negative components, and $|\mathbf{i} - \mathbf{e}_j| = n - 1$. We shall assume a lexicographic order of the multivariate Bernstein polynomials, similar to the $\mathbb{R}^2$ case, based on the index numbers $i_0, \ldots, i_d$. The Bernstein factor matrices for arbitrary number of variables are defined as follows:

Definition 8. The Bernstein factor matrix $\mathbf{T}_{d,n}(u_0, \ldots, u_d)$ is a $\binom{n+d-1}{d} \times \binom{n+d}{d}$ band-limited matrix with d non-zero elements on each row. The columns of $\mathbf{T}_{d,n}(u_0, \ldots, u_d)$ correspond to the $\binom{n+d}{d}$ vectors of terms in the recurrence relation (20) in lexicographic order, such that the non-zero elements in every column are

1. positioned according to the lexicographic index numbers of the $B^{n-1}_{i-\mathbf{e}_j}$ on the RHS, and
2. taking the values of their associated variables; $u_0, u_1, \cdots$, or u_d.

We recall Lemma 7, which can be extended to the multivariate case. An outline of a proof is to use the recurrence relation provided in (20) and Definition 8, which is similar to the proof of Lemma 7. We use this to propose an alternative equivalent definition of the multivariate Bernstein factor matrices of arbitrary degree, based on recursion of the sub-matrices, as summarized in the following theorem, which is a main result:

Theorem 9. *The Bernstein factor matrix $\mathbf{T}_{d,n}(\mathbf{u})$ is a $\binom{n+d-1}{d} \times \binom{n+d}{d}$ band-limited matrix with $d+1$ nonzero elements on each row. The matrix is defined recursively as follows:*

$$\mathbf{T}_{d,n}(u_0, \ldots, u_d) = \left(\begin{array}{c|c} \mathbf{T}_{d,n-1}(u_0, \ldots, u_d) & 0 \\ \hline \mathbf{T}^{\text{diag}}_{d,n}(u_0) & \mathbf{T}_{d-1,n}(u_1, \ldots, u_d) \end{array}\right), \tag{21}$$

where $d \geq 0$ is the dimension, $n \geq 1$ is the degree, $\mathbf{T}_{0,q} = (u_d)$ for $q \leq n$, $\mathbf{T}_{p,1} = (u_{d-p}, \ldots, u_d)$ for $p \leq d$, and $\mathbf{T}^{\text{diag}}_{q,n}(u_{d-q})$ is a matrix where the only non-zero elements are on the right diagonal and equal to u_{d-q}.

In the following we will in some cases omit specifying the parameters of the matrix $\mathbf{T}$ and its sub-matrices for simplification. We note that the sub-matrices in (21) are such that $\mathbf{T}_{d,n-1}$ is of size $\binom{n+d-2}{d} \times \binom{n+d-1}{d}$, $\mathbf{T}_{d-1,n}$ is $\binom{n+d-2}{d-1} \times \binom{n+d-1}{d-1}$ and $\mathbf{T}^{\text{diag}}_{d,n}$ is $\binom{d+n-2}{d} \times \binom{n+d-1}{d-1}$. The upper right sub-matrix is a $\binom{d+n-2}{d} \times \binom{n+d-1}{d-1}$ matrix where all the entries are zero.

Lemma 10. *The entries in every row in $\mathbf{T}_{d,n}(\mathbf{u})$ are either $0, u_0, \ldots,$ or u_d, such that each of the elements $u_0, \ldots, u_d$ occurs once, in increasing order, with optional zeros in-between.*

For the sake of clarity, we formalize that the proposed factor matrices are a factorization of the set of Bernstein polynomials in d variables, similar to Lemma 6 for the $\mathbb{R}^2$ case, in the following lemma.

Lemma 11. *The Bernstein factor matrices* $\{\mathbf{T}_{d,i}\}_{i=1}^{n}$ *factor the Bernstein polynomials* $\mathbf{B}_d^n$*:*

$$\mathbf{T}_{d,1}\mathbf{T}_{d,2}\cdots\mathbf{T}_{d,n} = \mathbf{B}_d^n. \tag{22}$$

We include the following lemma concerning a symmetry property of the Bernstein factor matrices:

Lemma 12. *Given the barycentric coordinates* $\mathbf{u} = (u_0, \ldots, u_d)$ *and* $\mathbf{z} = (z_0, \ldots, z_d)$ *with respect to a d-dimensional simplex A. Then*

$$\mathbf{T}_{d,k-1}(\mathbf{z})\mathbf{T}_{d,k}(\mathbf{u}) = \mathbf{T}_{d,k-1}(\mathbf{u})\mathbf{T}_{d,k}(\mathbf{z}). \tag{23}$$

3 Directional derivatives

Let us consider the line given by

$$l(t) = \mathbf{p} + t\mathbf{v},$$

where $\mathbf{p} = \mathbf{a}_0u_0 + \cdots + \mathbf{a}_du_d$ where $u_0 + \cdots + u_d = 1$ is a point in $\mathbb{R}^d$ with barycentric coordinates

$$\mathbf{u} = (u_0, u_1, \ldots, u_d), \tag{24}$$

and $\mathbf{v} = \mathbf{a}_0v_0 + \cdots + \mathbf{a}_dv_d$ where $v_0 + \cdots + v_d = 0$ is a vector with barycentric coordinates

$$\mathbf{v} = (v_0, v_1, \ldots, v_d), \tag{25}$$

and a polynomial $s(\mathbf{u})$ in Bernstein-Bézier form:

$$s(\mathbf{p}) = \sum_{|\mathbf{i}|=n} \mathbf{c}_\mathbf{i} B_\mathbf{i}^n(\mathbf{u}). \tag{26}$$

Consider $\mathbb{R}^{d+1}$ with Cartesian coordinates $u_0, u_1, \ldots, u_d$, a $(d+1)$-variate sufficiently smooth function $\sigma(u_0, u_1, \ldots, u_d)$ and a unit vector in $\mathbb{R}^{d+1}$ with Cartesian coordinates $(v_0, v_1, \ldots, v_d)$:

$$\sum_{k=0}^{d} v_k^2 = 1. \tag{27}$$

The 1-st directional derivative of σ at $\mathbf{u}^0 = (u_0^0, u_1^0, \ldots, u_d^0) \in \mathbb{R}^{d+1}$ in the direction of the vector $\mathbf{v} = (v_0, v_1, \ldots, v_d)$ satisfying (27) is, as habitual,

$$\lim_{t\to 0+} \frac{\mathrm{d}}{\mathrm{d}t}\sigma(t), \tag{28}$$

where $\sigma(t) = \sigma(\mathbf{u}^0 + t\mathbf{v})$, $t \in (0, t_0)$, $0 < t_0 < \infty$.

Definition 13. The 1-st directional derivative of $s(\mathbf{u}) = s(u_0, u_1, \ldots, u_d)$ at the point $\mathbf{p} = (u_0^0, u_1^0, \ldots, u_d^0) \in \mathbb{R}^d$ in the direction of the vector $\mathbf{v} = (v_0^0, v_1^0, \ldots, v_d^0) \in \mathbb{R}^d$ where u_k^0, v_k^0, $k = 0, 1, \ldots, d$ are barycentric coordinates in $\mathbb{R}^d$ satisfying (24) and (25), respectively, is the directional derivative in $\mathbb{R}^{d+1}$ in the particular case when $\mathbf{u}^0 = \mathbf{p}$ belongs to the affine manifold (hyperplane in $\mathbb{R}^{d+1}$) $\left\{(u_0, u_1, \ldots, u_d) \in \mathbb{R}^{d+1} : \sum_{k=0}^d u_k = 1\right\}$ and the directional vector $\mathbf{v}$ belongs to the intersection of the unit sphere $\left\{(v_0, v_1, \ldots, v_d) \in \mathbb{R}^{d+1} : \sum_{k=0}^d v_k^2 = 1\right\}$ and the d-dimensional subspace (hyperplane in $\mathbb{R}^{d+1}$) $\left\{(v_0, v_1, \ldots, v_d) \in \mathbb{R}^{d+1} : \sum_{k=0}^d v_k^2 = 0\right\}$.

Directional derivatives in barycentric coordinates of order higher than first are defined analogously.

From the chain rule, using also the properties (24) and (25) of the barycentric coordinates and Definition 13, we have that

$$D_{\mathbf{v}} s(\mathbf{p}) = v_0 D_{u_0} s(\mathbf{p}) + \cdots + v_d D_{u_d} s(\mathbf{p}).$$

Partial derivatives of a Bernstein polynomial with respect to each of its parameters can be obtained as

$$\frac{\partial}{\partial u_j} B_{\mathbf{i}}^n = n B_{\mathbf{i}-\mathbf{e}_j}^{n-1},$$

where $\mathbf{i}$ is a multiindex, $|\mathbf{i}| = n$, $|\mathbf{i} - \mathbf{e}_j| = n - 1$, and $B_{\mathbf{j}} = 0$ if $\mathbf{j}$ has a negative component. This gives rise to the following lemma, which can be found in a trivariate setting in [12]:

Lemma 14. *For arbitrary $i_0 + \cdots + i_d = n$, the directional derivative of a Bernstein polynomial $B_{\mathbf{i}}^n(\mathbf{p})$ in the direction of $\mathbf{v}$ is*

$$D_{\mathbf{v}} B_{\mathbf{i}}^n(\mathbf{p}) = n \left[v_0 B_{\mathbf{i}-\mathbf{e}_0}^{n-1}(\mathbf{u}) + \cdots + v_d B_{\mathbf{i}-\mathbf{e}_d}^{n-1}(\mathbf{u}) \right]. \tag{29}$$

Recall from Lemma 11 that

$$\mathbf{B}_d^n(\mathbf{u}) = \mathbf{T}_{d,1}(\mathbf{u}) \mathbf{T}_{d,2}(\mathbf{u}) \cdots \mathbf{T}_{d,n}(\mathbf{u}). \tag{30}$$

This product of matrices can be differentiated by the same formulae as if the factors were numbers. If $\mathbf{T}(\mathbf{u})$ is a matrix whose entries are functions of $\mathbf{u}$, then the derivative in the direction of $\mathbf{v}$, $D_{\mathbf{v}}\mathbf{T}$ of $\mathbf{T}$, is defined as the matrix obtained by differentiating each entry of $\mathbf{T}$ with respect to $\mathbf{v}$. The entries are obtained by differentiation of linear convex combinations, which yields constants independent of $\mathbf{u}$.

We formalize the definition of the derivative of the Bernstein factor matrix $\mathbf{T}_{d,n}(\mathbf{u})$:

Lemma 15. *The derivative of a Bernstein factorial matrix in the direction of $\mathbf{v} = (v_0, \ldots, v_d)$,*

$$D_{\mathbf{v}} \mathbf{T}_{d,n} = \mathbf{T}_{d,n}(\mathbf{v}),$$

is a $\binom{d+n-1}{n-1} \times \binom{d+n}{n}$ band-limited matrix with $d+1$ nonzero constant elements on each row (independent of $\mathbf{u}$).

We note here that the submatrix $D_{\mathbf{v}} \mathbf{T}_{d,n}^{\mathrm{diag}}$ is zero if $v_0 = 0$, since the only non-zero elements in $\mathbf{T}_{d,n}^{\mathrm{diag}}$ are equal to v_0. Furthermore, since $\mathbf{T}_{d-1,n}(\mathbf{v})$ does not contain v_0, the submatrix $D_{\mathbf{v}} \mathbf{T}_{d-1,n}$ is zero if $v_1 = \cdots = v_d = 0$, i.e. if $D_{\mathbf{v}}$ corresponds to the partial derivative D_{u_0}.

The remainder of this section is partially based on an analogous treatment of univariate B-splines on matrix form in [11], adjusted here to the setting of multivariate Bernstein polynomials. The following well-known rule for differentiating a product of two matrices will be used in order to find the derivative of the factorization.

Lemma 16. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two matrices with compatible dimensions for the matrix product* $\mathbf{AB}$ *and with entries that are functions of* $(u_0, \ldots, u_d)$*. Then*

$$D(\mathbf{AB}) = (D\mathbf{A})\mathbf{B} + \mathbf{A}(D\mathbf{B}).$$

By applying the rule from Lemma 16 to the product of (30), we get

$$D_{\mathbf{v}}\mathbf{B}_d^n(\mathbf{u}) = \sum_{k=1}^{n} \mathbf{T}_{d,1}(\mathbf{u}) \cdots \mathbf{T}_{d,k-1}(\mathbf{u}) D_{\mathbf{v}}\mathbf{T}_{d,k}\mathbf{T}_{d,k+1}(\mathbf{u}) \cdots \mathbf{T}_{d,n}(\mathbf{u}). \tag{31}$$

Formula (31) can be simplified by applying the following lemma:

Lemma 17. *The matrices* $\mathbf{T}_{d,k-1}$ *and* $\mathbf{T}_{d,k}$ *satisfy the relation*

$$D_{\mathbf{v}}\mathbf{T}_{d,k-1}\mathbf{T}_{d,k}(\mathbf{u}) = \mathbf{T}_{d,k-1}(\mathbf{u}) D_{\mathbf{v}}\mathbf{T}_{d,k}, \tag{32}$$

for $k \geq 2$*, any* $\mathbf{u} = (u_0, \ldots, u_d)$*,* $u_k \in \mathbb{R}$*,* $k = 0, 1, \ldots, d$*, and any given direction* $\mathbf{v} = (v_0, \ldots, v_d)$ *with* $\sum_{k=0}^{d} v_k^2 > 0$*.*

The differentiation operator $D_{\mathbf{v}}$ in (31) can be shifted between the matrices by using (32), and by making use of Lemma 15 we obtain

$$D_{\mathbf{v}}\mathbf{B}_d^n(\mathbf{u}) = n\mathbf{T}_{d,1}(\mathbf{u}) \cdots \mathbf{T}_{d,n-1}(\mathbf{u}) D_{\mathbf{v}}\mathbf{T}_{d,n} = n\mathbf{B}_d^{n-1}\mathbf{T}_{d,n}(\mathbf{v}). \tag{33}$$

We are now ready to provide one of the main results, a theorem which states that the r-th directional derivative of the set of Bernstein polynomials $\mathbf{B}_d^n$ can be obtained by differentiating r of the factor matrices.

Theorem 18. *Given a set of Bernstein polynomials* $\mathbf{B}_d^n(\mathbf{u})$ *and the directions* $\mathbf{v}_1, \ldots, \mathbf{v}_r$*, for* $1 \leq r \leq n$*. Then*

$$D_{\mathbf{v}_1} \cdots D_{\mathbf{v}_r}\mathbf{B}_d^n(\mathbf{u}) = \frac{n!}{(n-r)!}\mathbf{B}_d^{n-r}(\mathbf{u})\mathbf{T}_{d,n-r+1}(\mathbf{v}_1) \cdots \mathbf{T}_{d,n}(\mathbf{v}_r). \tag{34}$$

We recall from (26) that any d-variate polynomial of degree n has a Bernstein-Bézier representation,

$$s(\mathbf{u}) = \mathbf{B}_d^n(\mathbf{u})\mathbf{c},$$

where $\mathbf{c} = \begin{pmatrix} \mathbf{c}_0 & \cdots & \mathbf{c}_d \end{pmatrix}^T$. Next, we obtain the following via applying Theorem 18:

Corollary 19. *Given a polynomial* $s(\mathbf{u})$ *on Bernstein-Bézier form and the directions* $\mathbf{v}_1, \ldots, \mathbf{v}_r$*. Then*

$$D_{\mathbf{v}_1} \cdots D_{\mathbf{v}_r}s(\mathbf{p}) = \frac{n!}{(n-r)!}\mathbf{B}_d^{n-r}(\mathbf{u})\mathbf{T}_{d,n-r+1}(\mathbf{v}_1) \cdots \mathbf{T}_{d,n}(\mathbf{v}_r)\mathbf{c}. \tag{35}$$

4 Relevance to the de Casteljau algorithm

The Bernstein polynomials form a basis; thus, every d-variate polynomial $s(\mathbf{u})$ has a Bernstein-Bézier representation with respect to a reference simplex A:

$$s(\mathbf{u}) = \sum_{|\mathbf{i}|=n} \mathbf{c}_{\mathbf{i}}^0 B_{\mathbf{i}}^n(\mathbf{u}), \tag{36}$$

which can be evaluated with de Casteljau's algorithm by applying the recurrence relation in (20). Similar to the case of curves, we obtain

$$s(\mathbf{u}) = \sum_{|\mathbf{i}|=n-1} \mathbf{c}_{\mathbf{i}}^1 B_{\mathbf{i}}^{n-1}(\mathbf{u}),$$

and, after $n-2$ steps,

$$s(\mathbf{u}) = \sum_{|\mathbf{i}|=0} \mathbf{c}_{\mathbf{i}}^n B_{\mathbf{i}}^0(\mathbf{u}) = \mathbf{c}_{0\ldots0},$$

where $\mathbf{c}_{\mathbf{i}}^k = u_0 \mathbf{c}_{\mathbf{i}+\mathbf{e}_0}^{k-1} + \cdots + u_d \mathbf{c}_{\mathbf{i}+\mathbf{e}_d}^{k-1}$, with $|\mathbf{i}| = n, \ldots, 0$, are intermediate coefficients. These intermediate coefficients are polynomials of degree k.

The directional derivative of a polynomial in Bernstein-Bézier form is formulated, based on a description in [13], as follows.

Lemma 20. *The directional derivative of a d-variate polynomial $s(\mathbf{u})$ on Bernstein-Bézier form, at the point $\mathbf{p}$ in the direction of $\mathbf{v}$, is given by*

$$D_{\mathbf{v}} s(\mathbf{p}) = n \sum_{\mathbf{j}} \mathbf{d}_{\mathbf{j}} B_{\mathbf{j}}^{n-1}(\mathbf{u}),$$

where

$$\mathbf{d}_{\mathbf{j}} = v_0 \mathbf{c}_{\mathbf{j}+\mathbf{e}_0} + \cdots + v_d \mathbf{c}_{\mathbf{j}+\mathbf{e}_d}$$

are the intermediate points resulting from the first step of the de Casteljau algorithm based on the barycentric coordinates of $\mathbf{v}$.

We observe that the coefficients of the first derivative of a polynomial on Bernstein-Bézier form are as provided by Lemma 20 and verify that $D_{\mathbf{v}} s$ at the point $\mathbf{p}$ with barycentric coordinates $\mathbf{u}$ can be evaluated using the de Casteljau algorithm by applying one step of the algorithm using $\mathbf{v}$ and then $n-1$ steps using $\mathbf{u}$.

We abbreviate the intermediate points $\mathbf{d}_{\mathbf{j}}$, using the notation from [13], by $\mathbf{d}_{\mathbf{j}} = \Delta_{\mathbf{v}} \mathbf{c}_{\mathbf{j}}$. Higher-order derivatives are obtained in the same way, e.g., an r-th directional derivative $D_{\mathbf{v}_1} \cdots D_{\mathbf{v}_r} s$ has the Bézier coefficients $\Delta_{\mathbf{v}_1} \cdots \Delta_{\mathbf{v}_r} \mathbf{c}_{\mathbf{j}}$, where $|\mathbf{j}| = n - r$. The following result is obtained by applying Lemma 20 repeatedly.

Lemma 21. *Given the directions $\mathbf{v}_1, \ldots, \mathbf{v}_r$, for $1 \leq r \leq n$. Then,*

$$D_{\mathbf{v}_1} \cdots D_{\mathbf{v}_r} s(\mathbf{p}) = \frac{n!}{(n-r)!} \sum_{\mathbf{j}=n-r} \mathbf{d}_{\mathbf{j}}^{(r)} B_{\mathbf{j}}^{n-r}(\mathbf{u}),$$

where $\mathbf{d}_{\mathbf{j}}^{(r)}$ are the intermediate points obtained after performing r steps of the de Casteljau algorithm using $\mathbf{v}_1, \ldots, \mathbf{v}_r$ in this order.

One consequence of Lemma 21 is that $D_{\mathbf{v}_1} \cdots D_{\mathbf{v}_r} s$ can be evaluated at the point $\mathbf{p}$ with barycentric coordinates $\mathbf{u}$ by first applying r steps of the de Casteljau algorithm using $\mathbf{v}_1 \cdots \mathbf{v}_r$ in this order, and then by using $\mathbf{u}$ in the following $n - r$ steps.

The forward difference operator $\Delta_{\mathbf{v}}$ commutes with the steps of the de Casteljau algorithm [13] since the computation of affine combinations of affine combinations is commutative:

$$\sum_i \alpha_i \sum_j \beta_j \mathbf{p}_{ij} = \sum_j \beta_j \sum_i \alpha_i \mathbf{p}_{ij}.$$

Thus, an r-th directional derivative can also be obtained by first computing $n - r$ steps of the de Casteljau algorithm followed by r differentiation steps.

We obtain the following result; by using Theorem 9, Lemma 11, and (20), formula (36) can be expressed in matrix form:

Corollary 22. *Any d-variate polynomial $s(\mathbf{u})$ in Bernstein-Bézier form,*

$$s(\mathbf{u}) = \sum_{|\mathbf{i}|=n} \mathbf{c}_{\mathbf{i}}^0 B_{\mathbf{i}}^n(\mathbf{u}),$$

can be expressed by using Bernstein factor matrices as:

$$s(\mathbf{u}) = \mathbf{T}_{d,1}(\mathbf{u})\mathbf{T}_{d,2}(\mathbf{u}) \cdots \mathbf{T}_{d,n}(\mathbf{u})\mathbf{c}, \tag{37}$$

where $\mathbf{c} = \begin{pmatrix} \mathbf{c}_0 & \cdots & \mathbf{c}_d \end{pmatrix}^T$. Computing the RHS in (37) from right to left corresponds to the steps of the de Casteljau algorithm.

We note that computing only the matrices $\mathbf{T}_{d,1} \cdots \mathbf{T}_{d,n}$, from left to right, yields a vector containing the $\binom{d+n}{d}$ Bernstein polynomials of degree n in $d+1$ variables with respect to the simplex $A \subset \mathbb{R}^d$:

$$\mathbf{B}_d^n(\mathbf{u}) = \mathbf{T}_{d,1}(\mathbf{u}) \cdots \mathbf{T}_{d,n}(\mathbf{u}).$$

One important new observation here is that, in a manner similar to the case of curves, this method can be seen to correspond to a special case of the de Boor-Cox recursion formula for B-splines in a multivariate setting.

We conclude this section by recalling Corollary 19 and Lemma 21 to make another important new observation: namely Corollary 22 is applicable also for computations of directional derivatives.

4.1 Examples

4.1.1 The de Casteljau algorithm as an iterative procedure

The de Casteljau algorithm can be described as an iterative process by using projection and a linear step. Let us recall the expression in (5) for a cubic parametric curve on Bézier form by factorization using matrices, where in general for degree n we obtain:

$$c(t) = \mathbf{T}_1(t)\mathbf{T}_2(t) \cdots \mathbf{T}_n(t)\mathbf{v}_0,$$

where $\mathbf{v}_0 = \begin{pmatrix} \mathbf{c}_0 & \cdots & \mathbf{c}_0 \end{pmatrix}^T = \begin{pmatrix} \mathbf{c}_0^0 & \mathbf{c}_1^0 & \cdots & \mathbf{c}_{n-1}^0 & \mathbf{c}_n^0 \end{pmatrix}^T$ is a vector containing the coefficients, or Bézier points, for the curve. Then we extend the matrices $\mathbf{T}_k(t)$ for $1 \leq k \leq n$, as described in Lemma 1, to obtain a product of n square matrices of size $(n+1)^2$, multiplied with $\mathbf{v}_0$. After

performing one step of the de Casteljau algorithm, which corresponds to multiplying together the two rightmost factors in the product, we find

$$c(t) = \underbrace{\mathbf{T}_1(t)\cdots\mathbf{T}_{n-1}(t)}_{\substack{\text{n-1 times matrices}\\ \text{of size } (n+1)^2}}\mathbf{v}_1,$$

where $\mathbf{v}_1 = \begin{pmatrix}\mathbf{c}_0^1(t) & \mathbf{c}_1^1(t) & \cdots & \mathbf{c}_{n-1}^1(t) & 0\end{pmatrix}^T$. We investigate the difference between the two steps by considering the difference between the vectors of coefficients:

$$\mathbf{v}_1 - \mathbf{v}_0 = \begin{pmatrix}\mathbf{c}_0^1(t) - \mathbf{c}_0^0\\ \mathbf{c}_1^1(t) - \mathbf{c}_1^0\\ \vdots\\ \mathbf{c}_{n-1}^1(t) - \mathbf{c}_{n-1}^0\\ 0 - \mathbf{c}_n^0\end{pmatrix} = \mathbf{d}_1.$$

Recall that the intermediate points $\mathbf{c}_i^k = (1-t)\mathbf{c}_i^{k-1} + t\mathbf{c}_{i+1}^{k-1}$. Then the first $n-1$ entries in $\mathbf{d}_1$ are

$$\mathbf{c}_i^1(t) - \mathbf{c}_i^0 = (1-t)\mathbf{c}_i^0 + t\mathbf{c}_{i+1}^0 - \mathbf{c}_i^0 = t(\mathbf{c}_{i+1}^0 - \mathbf{c}_i^0).$$

It follows that the difference vector $\mathbf{d}_1$ can be decomposed into an orthogonal projection from $\mathbb{R}^{n+1}$ onto $\mathbb{R}^n$, and a linear step:

$$\mathbf{d}_1 = \mathbf{p}_1 + \mathbf{l}_1,$$

where

$$\mathbf{p}_1 = \begin{pmatrix}0\\0\\\vdots\\0\\-\mathbf{c}_n^0\end{pmatrix}, \quad \text{and} \quad \mathbf{l}_1 = t\begin{pmatrix}\mathbf{c}_1^0 - \mathbf{c}_0^0\\ \mathbf{c}_2^0 - \mathbf{c}_1^0\\ \vdots\\ \mathbf{c}_n^0 - \mathbf{c}_{n-1}^0\\ 0\end{pmatrix}.$$

In general, we find that the k-th step of the iterative procedure, which yields

$$c(t) = \underbrace{\mathbf{T}_1(t)\cdots\mathbf{T}_{n-k}(t)}_{\substack{\text{n-k times matrices}\\ \text{of size } (n+1)^2}}\mathbf{v}_k,$$

consists of a similar projection from $\mathbb{R}^{n+2-k}$ onto $\mathbb{R}^{n+1-k}$ and a linear step, such that

$$\mathbf{d}_k = \mathbf{v}_k - \mathbf{v}_{k-1} = \mathbf{p}_k + \mathbf{l}_k,$$

where

$$\mathbf{p}_k = \begin{pmatrix}0\\\vdots\\0\\-\mathbf{c}_{n-k}^{k-1}(t)\\0\\\vdots\\0\end{pmatrix}, \quad \text{and} \quad \mathbf{l}_k = t\begin{pmatrix}\mathbf{c}_1^{k-1}(t) - \mathbf{c}_0^{k-1}(t)\\ \vdots\\ \mathbf{c}_{n-k}^{k-1}(t) - \mathbf{c}_{n-k-1}^{k-1}(t)\\ 0\\ \vdots\\ 0\end{pmatrix}.$$

The process terminates when

$$c(t) = \mathbf{v}_n = \begin{pmatrix}\mathbf{c}_0^n(t)\end{pmatrix} = \mathbf{c}_0^n(t).$$

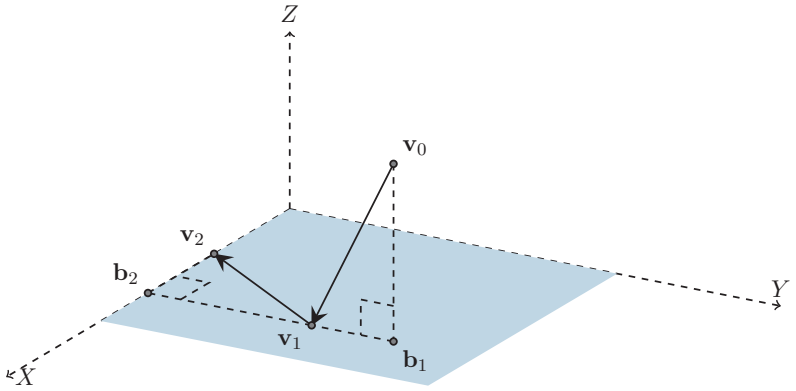

Figure 1: An illustration of the steps of the de Casteljau algorithm by iteratively performing a orthonormal projection followed by a linear step, for a parametric curve, with $n = 2$. The initial coefficients, or points, are contained in the vector $\mathbf{v}_0 \in \mathbb{R}^3$. $\mathbf{b}_1$ is the orthogonal projection of $\mathbf{v}_0$ onto $\mathbb{R}^2 \subset \mathbb{R}^3$, $\mathbf{v}_1$ constitutes the vector of intermediate points after the first step of the algorithm, and $\mathbf{b}_2$ is the orthogonal projection of $\mathbf{v}_1$ onto $\mathbb{R}^1 \subset \mathbb{R}^2$. The final step yields the vector $\mathbf{v}_2 \in \mathbb{R}^1$. It contains one point which constitutes the value of the curve at the parameter value t.

Figure 1 shows how the vectors of coefficients are repeatedly projected onto a subspace and translated by a linear step depending on t for the case of a parametric curve with $n = 2$. The number of coefficients, or points, is reduced by one for each projection until one single point remains.

In this example, $\mathbb{R}^{n+1-k} = W_k$ is a subset of the Euclidean space $\mathbb{R}^{n+2-k}$, and $\mathbf{b}_k = \mathbf{v}_{k-1} + \mathbf{p}_k$ is the orthogonal projection of $\mathbf{v}_{k-1}$ onto W_k. Then $\mathbf{b}_k$ is the closest point in W_k to $\mathbf{v}_{k-1}$ in the sense that $||\mathbf{v}_{k-1} - \mathbf{b}_k|| < ||\mathbf{v}_{k-1} - \mathbf{x}||$ for all $\mathbf{x}$ in W_k distinct from $\mathbf{b}_k$. It follows that $\mathbf{b}_k$ is the best approximation to $\mathbf{v}_{k-1}$ by elements of W_k.

We note that the space of square matrices used in the considered example is a complete inner product space (i.e., a Hilbert space) with the Euclidean norm $||\cdot||_{\ell^2}$. The method can be generalized to uniformly convex Banach spaces with ℓ^p-norm, with $1 < p < \infty$, since there will still exist a unique best approximation in such cases [3]. However, that will most likely yield non-polynomial special functions, in general, but they may share some properties with the polynomial case, where $p = 2$.

4.1.2 Representing a parametric surface

In this example we show how directional derivatives of a multivariate cubic Bézier representation of a parametric surface $s(\mathbf{u})$ can be expressed in matrix form by factorization. First of all, we consider the surface

$$s(\mathbf{u}) = \mathbf{B}_d^3(\mathbf{u})\mathbf{c},$$

where $\mathbf{B}_d^3(\mathbf{u})$ are the d-variate Bernstein polynomials of degree 3, and $\mathbf{c}$ are its Bézier coefficients. The derivative of $s(\mathbf{u})$ at the point $\mathbf{p}$ in the direction of $\mathbf{v}_0$,

$$D_{\mathbf{v}_0} s(\mathbf{p}) = D_{\mathbf{v}_0} \mathbf{B}_d^3(\mathbf{u})\mathbf{c},$$

can be expressed as a factorization of the Bernstein polynomials by matrices:

$$D_{\mathbf{v}_0} s(\mathbf{p}) = D_{\mathbf{v}_0} \left[\mathbf{T}_1(\mathbf{u})\mathbf{T}_2(\mathbf{u})\mathbf{T}_3(\mathbf{u})\right] \mathbf{c},$$

which computes to

$$D_{\mathbf{v}_0} s(\mathbf{p}) = \left[D_{\mathbf{v}_0} \mathbf{T}_1 \mathbf{T}_2(\mathbf{u}) \mathbf{T}_3(\mathbf{u}) + \mathbf{T}_1(\mathbf{u}) D_{\mathbf{v}_0} \mathbf{T}_2 \mathbf{T}_3(\mathbf{u}) + \mathbf{T}_1(\mathbf{u}) \mathbf{T}_2(\mathbf{u}) D_{\mathbf{v}_0} \mathbf{T}_3 \right] \mathbf{c},$$

and by applying Lemma 17 we obtain

$$D_{\mathbf{v}_0} s(\mathbf{p}) = 3\mathbf{B}_d^2(\mathbf{u}) D_{\mathbf{v}_0} \mathbf{T}_3 \mathbf{c}.$$

We proceed by applying one more differentiation step, this time in the direction of $\mathbf{v}_1$:

$$D_{\mathbf{v}_1} D_{\mathbf{v}_0} s(\mathbf{p}) = D_{\mathbf{v}_1} \left[3\mathbf{B}_d^2(\mathbf{u}) D_{\mathbf{v}_0} \mathbf{T}_3 \mathbf{c} \right] = 3 D_{\mathbf{v}_1} \mathbf{B}_d^2(\mathbf{u}) D_{\mathbf{v}_0} \mathbf{T}_3 \mathbf{c},$$

which computes to

$$D_{\mathbf{v}_1} D_{\mathbf{v}_0} s(\mathbf{p}) = 3 \left[D_{\mathbf{v}_1} \mathbf{T}_1 \mathbf{T}_2(\mathbf{u}) + \mathbf{T}_1(\mathbf{u}) D_{\mathbf{v}_1} \mathbf{T}_2 \right] D_{\mathbf{v}_0} \mathbf{T}_3 \mathbf{c},$$

and by applying Lemma 17 again and re-ordering the terms we obtain

$$D_{\mathbf{v}_1} D_{\mathbf{v}_0} s(\mathbf{p}) = 6\mathbf{B}_d^1(\mathbf{u}) D_{\mathbf{v}_1} \mathbf{T}_2 D_{\mathbf{v}_0} \mathbf{T}_3 \mathbf{c}.$$

With the use of Lemma 15 we write

$$D_{\mathbf{v}_1} D_{\mathbf{v}_0} s(\mathbf{p}) = 6\mathbf{B}_d^1(\mathbf{u}) \mathbf{T}_2(\mathbf{v}_1) \mathbf{T}_3(\mathbf{v}_0) \mathbf{c}.$$

4.1.3 Commutativity of multiplication

We shall now look at how a special property of the multivariate Bernstein factor matrix $\mathbf{T}_{d,n}(\mathbf{u})$ can be used to provide an alternative proof of Lemma 12 by using induction.

Let us recall from (23) that:

$$\mathbf{T}_{d,k-1}(\mathbf{z}) \mathbf{T}_{d,k}(\mathbf{u}) = \mathbf{T}_{d,k-1}(\mathbf{u}) \mathbf{T}_{d,k}(\mathbf{z}),$$

for $d \geq 0$ and $n \geq 1$.

Proof. The proof is based on induction, on the variables d and n, of (23) of the fully-ordered (well-ordered) set of all (d, n). We equip $\mathbb{N} \times \mathbb{N}$ with the lexicographic order relation $\leq$ defined by $(d, n) \leq (s, t)$ if ($(d < s)$ or $(d = s$ and $n \leq t)$). This is a full-order relation on $\mathbb{N} \times \mathbb{N}$, which means that we may use the induction theorem. Details on full-ordered sets and induction can be found in [20]. We proceed as follows to prove a property P for all (d, n) in $\mathbb{N} \times \mathbb{N}$:

1. Show that P holds for $(d, n) = (0, 1)$.
2. Suppose that P holds for all elements of $\mathbb{N} \times \mathbb{N}$ which are less than some arbitrary (d, n). Show that P holds for (d, n).
3. Conclude that P holds for all (d, n) in $\mathbb{N} \times \mathbb{N}$ by the induction theorem for the well-ordered sets.

1. Using $d = 0, n = 1$ in (23) gives the following relation:

$$\mathbf{T}_{0,1}(u) \mathbf{T}_{0,2}(z) = \mathbf{T}_{0,1}(z) \mathbf{T}_{0,2}(u),$$

which yields

$$\begin{pmatrix} u \end{pmatrix} \begin{pmatrix} z \end{pmatrix} = \begin{pmatrix} z \end{pmatrix} \begin{pmatrix} u \end{pmatrix}, \tag{38}$$

and it follows that both the LHS and the RHS of (38) compute to uz, since the product of two numbers commutes.

2. We want to prove $P(d,n)$ using any $(p,q) < (d,n)$. When applying (21) and its derivative from the definitions above, the left-hand side of (23) becomes

$$\left(\begin{array}{c|c} \mathbf{T}_{d,n-1}(\mathbf{u}) & 0 \\ \hline \mathbf{T}^{\text{diag}}_{d,n}(\mathbf{u}) & \mathbf{T}_{d-1,n}(\mathbf{u}) \end{array}\right)\left(\begin{array}{c|c} \mathbf{T}_{d,n}(\mathbf{z}) & 0 \\ \hline \mathbf{T}^{\text{diag}}_{d,n+1}(\mathbf{z}) & \mathbf{T}_{d-1,n+1}(\mathbf{z}) \end{array}\right),$$

which can be written as

$$\left(\begin{array}{c|c} \mathbf{T}_{d,n-1}(\mathbf{u})\mathbf{T}_{d,n}(\mathbf{z}) & 0 \\ \hline \mathbf{T}^{\text{diag}}_{d,n}(\mathbf{u})\mathbf{T}_{d,n}(\mathbf{z}) + \mathbf{T}_{d-1,n}(\mathbf{u})\mathbf{T}^{\text{diag}}_{d,n+1}(\mathbf{z}) & \mathbf{T}_{d-1,n}(\mathbf{u})\mathbf{T}_{d-1,n+1}(\mathbf{z}) \end{array}\right). \tag{39}$$

A similar computation for the RHS of (23) yields

$$\left(\begin{array}{c|c} \mathbf{T}_{d,n-1}(\mathbf{z})\mathbf{T}_{d,n}(\mathbf{u}) & 0 \\ \hline \mathbf{T}^{\text{diag}}_{d,n}(\mathbf{z})\mathbf{T}_{d,n}(\mathbf{u}) + \mathbf{T}_{d-1,n}(\mathbf{z})\mathbf{T}^{\text{diag}}_{d,n+1}(\mathbf{u}) & \mathbf{T}_{d-1,n}(\mathbf{z})\mathbf{T}_{d-1,n+1}(\mathbf{u}) \end{array}\right). \tag{40}$$

It follows that both the upper left and the lower right sub-matrices of (39) are equivalent to the corresponding sub-matrices of (40) by definition of the induction step.

Next, we need to check that the lower left sub-matrices of (39) and (40) are equivalent:

$$\mathbf{T}^{\text{diag}}_{d,n}(\mathbf{u})\mathbf{T}_{d,n}(\mathbf{z}) + \mathbf{T}_{d-1,n}(\mathbf{u})\mathbf{T}^{\text{diag}}_{d,n+1}(\mathbf{z}) = \mathbf{T}^{\text{diag}}_{d,n}(\mathbf{z})\mathbf{T}_{d,n}(\mathbf{u}) + \mathbf{T}_{d-1,n}(\mathbf{z})\mathbf{T}^{\text{diag}}_{d,n+1}(\mathbf{u}). \tag{41}$$

The non-zero elements of $\mathbf{T}^{\text{diag}}_{d,n}(\mathbf{u})$ and $\mathbf{T}^{\text{diag}}_{d,n}(\mathbf{z})$ are equal to u_0 and v_0, respectively. Thus, the LHS of (41) becomes

$$u_0\left(\begin{array}{c|c} \mathbf{T}^{\text{diag}}_{d,n}(\mathbf{z}) & \mathbf{T}_{d-1,n}(\mathbf{z}) \end{array}\right) + z_0\left(\begin{array}{c|c} 0 & \mathbf{T}_{d-1,n}(\mathbf{u}) \end{array}\right).$$

Similarly, we obtain for the RHS of (41):

$$z_0\left(\begin{array}{c|c} \mathbf{T}^{\text{diag}}_{d,n}(\mathbf{u}) & \mathbf{T}_{d-1,n}(\mathbf{u}) \end{array}\right) + u_0\left(\begin{array}{c|c} 0 & \mathbf{T}_{d-1,n}(\mathbf{z}) \end{array}\right).$$

Re-ordering the terms and inserting them into (41) yields

$$\left(\begin{array}{c|c} u_0\mathbf{T}^{\text{diag}}_{d,n}(\mathbf{z}) & u_0\mathbf{T}_{d-1,n}(\mathbf{z}) + z_0\mathbf{T}_{d-1,n}(\mathbf{u}) \end{array}\right) = \left(\begin{array}{c|c} z_0\mathbf{T}^{\text{diag}}_{d,n}(\mathbf{u}) & u_0\mathbf{T}_{d-1,n}(\mathbf{z}) + z_0\mathbf{T}_{d-1,n}(\mathbf{u}) \end{array}\right),$$

and it follows that LHS = RHS in (41), since $\mathbf{T}^{\text{diag}}_{d,n}(\mathbf{u})$ contains u_0s and $\mathbf{T}^{\text{diag}}_{d,n}(\mathbf{z})$ contains v_0s, and since the product of two real scalars is commutative.

3. Finally, we conclude that (23) holds for all (d,n) in $\mathbb{N}\times\mathbb{N}$ provided $d \geq 0, n \geq 1$. □

5 Proofs

This section contains proofs of some of the lemmas and theorems presented earlier in this article.

Proof of Lemma 2

Proof. We consider the product

$$E_{ij}E_{k\ell} = E_{\alpha\beta},$$

where the entries in E_{ij} are

$$c_{\mu\nu} = \begin{cases} 1, & \text{if} \quad \mu = i, \nu = j, \\ 0, & \text{otherwise,} \end{cases}$$

and the elements of $E_{k\ell}$ are

$$d_{\lambda\eta} = \begin{cases} 1, & \text{if} \quad \lambda = k, \eta = \ell, \\ 0, & \text{otherwise.} \end{cases}$$

Since

$$E_{ij} = (\delta_{i\mu}\delta_{\nu j})_{\mu=1,\nu=1}^{N,N},$$

and

$$E_{k\ell} = (\delta_{k\lambda}\delta_{\eta\ell})_{\lambda=1,\eta=1}^{N,N},$$

the elements of $E_{\alpha\beta}$ can be expressed as

$$e_{\alpha\beta} = \sum_{a=1}^{N} c_{\alpha a}d_{a\beta} = \sum_{a=1}^{N} \delta_{i\alpha}\delta_{aj}\delta_{ka}\delta_{\beta\ell}.$$

Fixing $\alpha = i$ and $\beta = \ell$ yields

$$e_{\alpha\beta} = \sum_{a=1}^{N} \delta_{aj}\delta_{ka},$$

and for $a = j$ and $k = a$ we obtain

$$e_{\alpha\beta} = \sum_{a=j=k}^{N} 1 = 1.$$

Thus,

$$e_{\alpha\beta} = \delta_{kj}\delta_{i\alpha}\delta_{\ell\beta},$$

which means that

$$e_{i\ell} = \delta_{kj}.$$

Finally, we obtain

$$E_{ij}E_{kl} = \delta_{kj}E_{i\ell},$$

and the result follows. □

Proof of Lemma 3

Proof.

$$\mathbf{AB} - \mathbf{CD} = \left(\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}E_{ij}\right)\left(\sum_{k=1}^{N}\sum_{\ell=1}^{N} b_{k\ell}E_{k\ell}\right) - \left(\sum_{m=1}^{N}\sum_{n=1}^{N} c_{mn}E_{mn}\right)\left(\sum_{p=1}^{N}\sum_{q=1}^{N} d_{pq}E_{pq}\right),$$

where the RHS computes to

$$\mathbf{AB} - \mathbf{CD} = \sum_{i}\sum_{j}\sum_{k}\sum_{\ell} a_{ij}b_{k\ell}E_{ij}E_{k\ell} - \sum_{m}\sum_{n}\sum_{p}\sum_{q} c_{mn}d_{pq}E_{mn}E_{pq}.$$

By using the multiplier in (8) for $E_{ij}E_{k\ell}$ and $E_{mn}E_{pq}$ and re-ordering the sums we obtain

$$\mathbf{AB}-\mathbf{CD}=\sum_{i}\sum_{j}\sum_{k}\sum_{\ell}a_{ij}b_{k\ell}\delta_{jk}E_{i\ell}-\sum_{m}\sum_{n}\sum_{p}\sum_{q}c_{mn}d_{pq}\delta_{np}E_{mq},$$

where setting $j=k=\mu$ and $n=p=\nu$ yields

$$\mathbf{AB}-\mathbf{CD}=\sum_{i}\sum_{\ell}\left(\sum_{\mu}a_{i\mu}b_{\mu\ell}\right)E_{i\ell}-\sum_{m}\sum_{q}\left(\sum_{\nu}c_{m\nu}d_{\nu q}\right)E_{mq}.$$

Then we set $i=\alpha$, $\ell=\beta$, $m=\alpha$ and $q=\beta$, and we get

$$\begin{aligned}\mathbf{AB}-\mathbf{CD}&=\sum_{\alpha}\sum_{\beta}\left(\sum_{\mu}a_{\alpha\mu}b_{\mu\beta}\right)E_{\alpha\beta}-\sum_{\alpha}\sum_{\beta}\left(\sum_{\nu}c_{\alpha\nu}d_{\nu\beta}\right)E_{\alpha\beta}\\&=\sum_{\alpha}\sum_{\beta}\left(\sum_{\mu}a_{\alpha\mu}b_{\mu\beta}-\sum_{\nu}c_{\alpha\nu}d_{\nu\beta}\right)E_{\alpha\beta}.\end{aligned}$$

Finally, we fix $\nu=\mu=\lambda$, and the result follows. □

Proof of Corollary 4

Proof. The result follows by setting $\mathbf{C}=\mathbf{B}$ and $\mathbf{D}=\mathbf{A}$ in Lemma 3 and re-ordering the terms, since $[\mathbf{A},\mathbf{B}]=\mathbf{AB}-\mathbf{BA}$. □

Proof of Lemma 6

Proof. We use induction on the variable n of (18). It is trivial that $\mathbf{T}_{2,1}=\mathbf{B}_2^1=(u,v,w)$. Assuming that it holds for $n=r-1$ and setting $n=r$ in (18) yields

$$\mathbf{B}_2^{r-1}\mathbf{T}_{2,r}=\mathbf{B}_2^r.$$

But then the columns of $\mathbf{T}_{2,r}$ must correspond to the variables u, v and w in the RHS of the recurrence relation (12) for $n=r$, where the values u, v and w are indexed within every column according to the position of the indexes of the B_{ijk}^{r-1} in lexicographic order, which is exactly $\mathbf{T}_{2,r}$ by Definition 5. We conclude that (18) holds for arbitrary n by induction.

□

Proof of Lemma 7

Proof. The proof is based on using the recurrence relation (12) and Definition 5.
$\mathbf{T}_{2,n-1}(u,v,w)$ corresponds by Definition 5 to the set of all $\binom{n+1}{2}$ RHS terms of B_{ijk}^{n-1} in (12) for $i+j+k=n-1$. By Definition 5, a similar collection of terms from the first $\binom{n+1}{2}$ vectors obtained from (12), for B_{ijk}^{n}, which corresponds to the elements where $i>0$ since $i+j+k=n$, results in a matrix whose upper part contains $\mathbf{T}_{2,n-1}(u,v,w)$ and its lower part contains $\mathbf{T}_{2,n}^{\text{diag}}(u)$. Here, the upper part consists of the first $\binom{n}{2}$ rows, whereas the remaining $\binom{n+1}{2}-\binom{n}{2}$ rows consitute the lower part. Similarly, collecting the remaining $\binom{n+2}{2}-\binom{n+1}{2}$ vectors results in a matrix where the upper part is a zero matrix and the non-zero elements in the lower part are either v or w, since $i=0$, and they follow the layout of $\mathbf{T}_{1,n}(v,w)$. □

Proof of Lemma 10

Proof. According to Definition 8, the columns in $\mathbf{T}_{d,n}(\mathbf{u})$ are arranged such that

1. The order of the columns are based on the lexicographic order of the permutations of $\mathbf{i}$ for $|\mathbf{i}| = n$, and
2. u_i is placed on the $(\mathbf{i} - \mathbf{e}_i)$-th row within the column, if $0 \leq (\mathbf{i} - \mathbf{e}_i) < \binom{d+n-1}{d}$.

The possible results of computing $\mathbf{i} - \mathbf{e}_i$, for $i = 0, \ldots, d$ leads to the $\binom{n+d-1}{d}$ permutations of $\mathbf{j} = (j_0, \ldots, j_d)$, where $|\mathbf{j}| = n - 1$. There are $d + 1$ possible ways to obtain the position j_k, for $k = 1, \ldots, \binom{n+d-1}{d}$ in the lexicographic order of the values of $\mathbf{j}$, by taking $\mathbf{i} - \mathbf{e}_i$. Since the values of $\mathbf{i}$ are ordered, the values of j_k will appear in order such that the associated values of u_i are in the order $u_0, \ldots, u_d$. Since j_k corresponds to a row in $\mathbf{T}_{d,n}(\mathbf{u})$, the result follows. □

Proof of Lemma 11

Proof. We use use induction in the variable n of (22) for arbitrary degree d.
1. By setting $n = 2$ in the LHS of (22) and applying Theorem 9 we obtain

$$\mathbf{T}_{d,1}\mathbf{T}_{d,2} = \begin{pmatrix} u_0 & u_1 & \cdots & u_d \end{pmatrix} \left(\begin{array}{c|c} \mathbf{T}_{d,1}(u_0, \ldots, u_d) & 0 \\ \hline \mathbf{T}^{\text{diag}}_{d,n}(u_0) & \mathbf{T}_{1,n}(u_1, \ldots, u_d) \end{array} \right),$$

which computes to

$$\begin{pmatrix} u_0^2 & 2u_1u_0 & 2u_2u_0 & \cdots & u_1^2 & 2u_2u_1 & \cdots & 2u_du_1 & u_2^2 & \cdots & u_d^2 \end{pmatrix} = \mathbf{B}^2_d,$$

thus, (22) holds for $n = 2$.

2. Setting $n = r$ in (22) gives

$$\mathbf{T}_{d,1}\mathbf{T}_{d,2} \cdots \mathbf{T}_{d,r} = \mathbf{B}^r_d. \tag{42}$$

Assuming that $\mathbf{T}_{d,1}\mathbf{T}_{d,2} \cdots \mathbf{T}_{d,r-1} = \mathbf{B}^{r-1}_d$, we obtain

$$\mathbf{B}^{r-1}_d \mathbf{T}_{d,r} = \mathbf{B}^r_d. \tag{43}$$

Using the recurrence formula in (20), formula (43) holds if, and only if, the columns of $\mathbf{T}_{d,n}$ correspond to the variables $u_0, \ldots, u_d$ in the RHS of the recurrence relation in (20) for $n = r$, where the values $u_0, \ldots, u_d$ are indexed within every column according to the position of the indices of the associated $B^{r-1}_{i-\mathbf{e}_j}$ in lexicographic order. But, this is true according to Definition 8; hence, we conclude that (42) holds and that $\mathbf{T}_{d,1}\mathbf{T}_{d,2} \cdots \mathbf{T}_{d,n}$ is a factorization of the set of Bernstein polynomials $\mathbf{B}^n_d$. □

Proof of Lemma 12

Proof. Consider the following relation:

$$\mathbf{T}_{d,k-1}(\mathbf{z})\mathbf{T}_{d,k}(\mathbf{u}) - \mathbf{T}_{d,k-1}(\mathbf{u})\mathbf{T}_{d,k}(\mathbf{z}),$$

where $\mathbf{u} = (u_0, \ldots, u_d)$ and $\mathbf{z} = (z_0, \ldots, z_d)$ are barycentric coordinates with respect to a d-dimensional simplex A. Using Lemma 3, this translates to

$$\sum_{\alpha=1}^{N} \sum_{\beta=1}^{N} \left[\sum_{\lambda=1}^{N} (a_{\alpha\lambda} b_{\lambda\beta} - c_{\alpha\lambda} d_{\lambda\beta}) \right] E_{\alpha\beta}, \tag{44}$$

where $a_{\alpha\lambda}, b_{\lambda\beta}, c_{\alpha\lambda}$ and $d_{\lambda\beta}$ are elements of $\mathbf{T}_{d,k-1}(\mathbf{z})$, $\mathbf{T}_{d,k}(\mathbf{u})$, $\mathbf{T}_{d,k-1}(\mathbf{u})$ and $\mathbf{T}_{d,k}(\mathbf{z})$, respectively. Next, we change the order of the sums to

$$\sum_{\alpha=1}^{N}\sum_{\lambda=1}^{N}\sum_{\beta=1}^{N}\left(a_{\alpha\lambda}b_{\lambda\beta}-d_{\lambda\beta}c_{\alpha\lambda}\right)E_{\alpha\beta}.$$

Then the terms $a_{\alpha\lambda}b_{\lambda\beta}$ are obtained in an order such that, for each row α, every entry in the λ-th column of $\mathbf{T}_{d,k-1}(\mathbf{z})$ is multiplied with all entries on the λ-th row of $\mathbf{T}_{d,k}(\mathbf{u})$. Furthermore, the order of the terms $d_{\lambda\beta}c_{\alpha\lambda}$ yields that all entries on the λ-th row of $\mathbf{T}_{d,k}(\mathbf{z})$ are multiplied with every entry in the λ-th column of $\mathbf{T}_{d,k-1}(\mathbf{u})$. We recall Lemma 10, and conclude that for each α we obtain

$$\begin{aligned}\sum_{\lambda}\sum_{\beta}\left([a_{\alpha\lambda}b_{\lambda\beta}]-[d_{\lambda\beta}c_{\alpha\lambda}]\right) &= [z_0(u_0+\cdots+u_d)+\cdots+z_d(u_0+\cdots+u_d)]\\ &\quad -[(z_0+\cdots+z_d)u_0+\cdots+(z_0+\cdots+z_d)u_d],\end{aligned}$$

where the RHS can be re-arranged to

$$(z_0+\cdots+z_d)(u_0+\cdots+u_d)-(u_0+\cdots+u_d)(z_0+\cdots+z_d),$$

which clearly is equal to zero since scalar multiplication is commutative. But then (44) is equal to zero, since changing the order of the sums commutes. Finally, since two matrices $\mathbf{A}$ and $\mathbf{B}$ commute if $\mathbf{AB}-\mathbf{BA}=0$ (see Corollary 4), it follows that

$$\mathbf{T}_{d,k-1}(\mathbf{z})\mathbf{T}_{d,k}(\mathbf{u})-\mathbf{T}_{d,k-1}(\mathbf{u})\mathbf{T}_{d,k}(\mathbf{z})=0.$$

□

Proof of Lemma 14

Proof. The barycentric coordinates of the point $\mathbf{p}+t\mathbf{v}$ are

$$(u_0+tv_0,\ldots,u_d+tv_d).$$

Using this with (19) yields

$$B^n_{\mathbf{i}}(\mathbf{p}+t\mathbf{v})=\binom{n}{\mathbf{i}}(u_0+tv_0)^{i_0}\cdots(u_d+tv_d)^{i_d}.$$

By differentiating with respect to t and evaluating at $t=0$ we obtain

$$D_{\mathbf{v}}B^n_{\mathbf{i}}(\mathbf{p})=\binom{n}{\mathbf{i}}\left[i_0u_0^{i_0-1}v_0u_1^{i_1}\ \cdots u_d^{i_d}+u_0^{i_0}i_1u_1^{i_1-1}v_1\cdots u_d^{i_d}\ +\cdots+u_0^{i_0}\cdots u_{d-1}^{i_{d-1}}i_du_d^{i_d-1}v_d\right],$$

where the RHS computes to (29). □

Proof of Lemma 15

Proof. By applying the chain rule we obtain

$$D_{\mathbf{v}}\mathbf{T}_{d,n}=v_0D_{u_0}\mathbf{T}_{d,n}+\cdots+v_dD_{u_d}\mathbf{T}_{d,n}.$$

The RHS corresponds to setting $\mathbf{T}_{d,n}(\mathbf{v})$, and the result follows. □

Proof of Lemma 17

Proof. The result follows by differentiating both sides of (23) with respect to $\mathbf{z}$ in the direction of $\mathbf{v}$. □

Proof of Theorem 18

Proof. We differentiate (33) to find the second derivative. Since $D_{\mathbf{v}_k}\left(\mathbf{T}_{d,n}(\mathbf{v}_1)\right) = 0$, we get

$$D_{\mathbf{v}_2} D_{\mathbf{v}_1} \mathbf{B}_d^n(\mathbf{u}) = n D_{\mathbf{v}_2} \mathbf{B}_d^{n-1}(\mathbf{u}) \mathbf{T}_{d,n}(\mathbf{v}_1),$$

and by applying (33) to $D_{\mathbf{v}_2}\mathbf{B}_d^{n-1}$, we obtain

$$D_{\mathbf{v}_2} D_{\mathbf{v}_1} \mathbf{B}_d^n(\mathbf{u}) = n(n-1) \mathbf{B}_d^{n-2}(\mathbf{u}) \mathbf{T}_{d,n-1}(\mathbf{v}_2) \mathbf{T}_{d,n}(\mathbf{v}_1).$$

Similarly, for the r-th derivative we find

$$D_{\mathbf{v}_r} \cdots D_{\mathbf{v}_1} \mathbf{B}_d^n(\mathbf{u}) = \frac{n!}{(n-r)!} \mathbf{B}_d^{n-r}(\mathbf{u}) \mathbf{T}_{d,n-r+1}(\mathbf{v}_r) \cdots \mathbf{T}_{d,n}(\mathbf{v}_1).$$

Since, in addition, $\mathbf{B}_d^{n-r}(\mathbf{u}) = \mathbf{T}_{d,1}(\mathbf{u}), \cdots, \mathbf{T}_{d,n-r}(\mathbf{u})$ holds, and since it does not matter which of the n matrices $\mathbf{T}_{d,k}$ is being differentiated (it only matters that we differentiate r of them), due to the symmetry property of (32), we conclude that (34) is true. □

Proof of Lemma 20

Proof. Since $l(t)$ is a line in the domain, it can be seen to correspond to a curve $s \circ l(t)$ on a given parametric object,

$$s \circ l(t) = s(\mathbf{p}) = \sum_{\mathbf{i}} \mathbf{c}_{\mathbf{i}} B_{\mathbf{i}}^n(\mathbf{u}).$$

Then the derivative of $s(\mathbf{u})$ at the point $\mathbf{p}$ in the direction of $\mathbf{v}$ is given by its derivative with respect to t at $t = 0+$:

$$\begin{aligned} D_{\mathbf{v}} s(\mathbf{p}) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(s \circ l(t)\right)\Big|_{t=0+} \\ &= v_0 \frac{\partial}{\partial u_0} s + \cdots + v_d \frac{\partial}{\partial u_d} s \\ &= n \sum_{\mathbf{j}} \mathbf{d}_{\mathbf{j}} B_{\mathbf{j}}^{n-1}(\mathbf{u}). \end{aligned}$$

□

6 Concluding remarks

We have proposed a matrix representation of multivariate Bernstein polynomials of arbitrary dimension by factorization. The factor matrix is defined recursively via a particular decomposition of the matrix into sub-matrices.

The multiplication of the factor matrices from right to left was shown to correspond to the steps of the de Casteljau algorithm. A similar relation to the de Boor-Cox recursion formula, specialized for Bézier representation, can be found by computing the matrices from left to right,

which yields a vector containing the Bernstein polynomials for the given degree in the specified number of variables. These polynomials can then be multiplied together with the associated vector of coefficients in order to evaluate the parametric object in question.

The matrix representation illustrates clearly one of the differences between the two methods; in the case of de Casteljau, every step yields intermediate points, whereas the intermediate values with the de Boor-Cox method are numbers. The total number of arithmetic operations when multiplying the matrices from right to left (de Casteljau) are higher than in the case of multiplication from left to right (de Boor-Cox), provided that the coefficients are points in $\mathbb{R}^d, d > 1$.

References

[1] S. Bernstein, *Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités.*, Communications de la Société Mathématique de Kharkow. 2-ée série **13**, No. 1 (1912), 1–2.

[2] G. G. Lorentz, *Bernstein Polynomials*, 2nd ed., AMS Chelsea Publishing, Providence, Rhode Island (1986), reprinted, 2012.

[3] P. J. Davis, *Interpolation and approximation*, Dover, New York (1975).

[4] G. Farin, *Curves and surfaces for CAGD: a practical guide*, 4th ed., Computer Science and Scientific Computing, Academic Press, San Diego, CA, USA (1997).

[5] C. de Boor, *A Practical Guide to Splines*, revised ed., vol. 27 of *Applied mathematical sciences*, Springer-Verlag, New York (2001).

[6] L. L. Schumaker, *Spline Functions*, 3rd ed., Cambridge University Press, Cambridge CB2 8RU, UK (2007).

[7] P. de Faget de Casteljau, *Shape mathematics and CAD*, English language ed., vol. 2 of *Mathematics and CAD*, Kogan Page, 120 Pentonville Road, London, N1 9JN (1986).

[8] A. Lakså, *A method for sparse-matrix computation of B-spline curves and surfaces*, in I. Lirkov, S. Margenov, J. Waśniewski (eds.), *Large-Scale Scientific Computing 2009*, Springer (2010), vol. 5910 of *Lecture Notes in Computer Science*, 796–804. DOI: http://dx.doi.org/10.1007/978-3-642-12535-5_95.

[9] P. Costantini, C. Manni, *Geometric Construction of Generalized Cubic Splines*, Rendiconti di Matematica e delle sue Applicazioni **26** (2006), 327–338.

[10] A. Lakså, *Basic properties of Expo-Rational B-splines and practical use in Computer Aided Geometric Design*, Ph.D. thesis, University of Oslo (2007), (Dr.philos.).

[11] T. Lyche, K. Mørken, *Spline Methods* (2011), lecture notes at the University of Oslo (Draft).

[12] M.-J. Lai, L. L. Schumaker, *Spline Functions on Triangulations*, vol. 110 of *Encyclopedia of Mathematics and Its Applications*, Cambridge University Press, Cambridge CB2 8RU, UK (2007).

[13] H. Prautzsch, W. Boehm, M. Paluszny, *Bézier and B-spline Techniques*, Mathematics and visualization, Springer-Verlag, Berlin Heidelberg (2002).

[14] R. Dalmo, J. Bratlie, A. Lakså, *Recursive construction of Bernstein factor matrices* (2014), communication at the 8th International Conference on Curves and Surfaces, Paris, France.

[15] C. de Boor, *B-form basics*, in G. E. Farin (ed.), *Geometric Modeling: Algorithms and New Trends*, SIAM, Philadelphia (1987), 131–148.

[16] C. de Boor, *On calculating with B-splines*, Journal of Approximation Theory **6**, No. 1 (1972), 50–62. DOI: http://dx.doi.org/10.1016/0021-9045(72)90080-9.

[17] M. G. Cox, *The Numerical Evaluation of B-splines*, J. Inst. Maths. Applics. **10** (1972), 134–149. DOI: http://dx.doi.org/10.1093/imamat/10.2.134.

[18] W. J. Gordon, R. Riesenfeld, *B-spline curves and surfaces*, in R. Barnhill, R. Riesenfeld (eds.), *Computer Aided Geometric Design*, Academic Press, New York (1974), 95–126.

[19] B. L. van der Waerden, *Algebra II*, Fünfte Auflage der Modernen Algebra ed., Springer-Verlag, Berlin Heidelberg New York (1967).

[20] K. Ciesielski, *Set Theory for the Working Mathematician*, London Mathematical Society Student Texts (Book 39), Cambridge University Press, Cambridge, UK (1997).