Matrix Methods for the Bernstein Form and Their Application in Global Optimization

June, 10

Jihad Titi Jürgen Garloff

University of Konstanz
Department of Mathematics and Statistics
and
University of Applied Sciences / HTWG Konstanz
Faculty of Computer Science

Outline

- The Bernstein expansion for polynomials over a box and a simplex
- New method for the computation of the Bernstein coefficients of multivariate Bernstein polynomials over a simplex

Outline

- The Bernstein expansion for polynomials over a box and a simplex
- New method for the computation of the Bernstein coefficients of multivariate Bernstein polynomials over a simplex

Outline

- The Bernstein expansion for polynomials over a box and a simplex
- New method for the computation of the Bernstein coefficients of multivariate Bernstein polynomials over a simplex
- New method for the calculation of the Bernstein coefficients over sub-boxes generated by subdivision

Outline

- The Bernstein expansion for polynomials over a box and a simplex
- New method for the computation of the Bernstein coefficients of multivariate Bernstein polynomials over a simplex
- New method for the calculation of the Bernstein coefficients over sub-boxes generated by subdivision
- Test for the convexity of a polynomial

Outline

- The Bernstein expansion for polynomials over a box and a simplex
- New method for the computation of the Bernstein coefficients of multivariate Bernstein polynomials over a simplex
- New method for the calculation of the Bernstein coefficients over sub-boxes generated by subdivision
- Test for the convexity of a polynomial

Notations

- We will consider the unit box $\mathbf{u} := [0,1]^n$, since any compact nonempty box $\mathbf{x}$ of $\mathbb{R}^n$ can be mapped affinely upon $\mathbf{u}$.
- The multi-index $(i_1, \ldots, i_n)$ is abbreviated by i, where n is the number of variables.
- The multi-index k is defined as $k = (k_1, k_2, \ldots, k_n)$.
- Comparison between and arithmetic operations with multi-indices are defined entry-wise.
- For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, its monomials are defined as $x^i := \prod_{j=1}^n x_j^{i_j}$.
- The abbreviations $\sum_{i=0}^{k} := \sum_{i_1=0}^{k_1} \ldots \sum_{i_n=0}^{k_n}$ and $\binom{k}{i} := \prod_{\alpha=1}^{n} \binom{k_\alpha}{i_\alpha}$ are used.
- $i_{s,q} := (i_1, i_2, \ldots, i_{s-1}, i_s + q, i_{s+1}, \ldots, i_n)$ where $0 \leq i_s + q \leq k_s$.

Bernstein Polynomials

- Let p be an n-variate polynomial of degree l

$$p(x) = \sum_{i=0}^{l} a_i x^i. \tag{1}$$

- The i-th Bernstein polynomial of degree k, $k \geq l$, is the polynomial $(0 \leq i \leq k)$

$$B_i^{(k)}(x) = \binom{k}{i} x^i (1-x)^{k-i}. \tag{2}$$

- The Bernstein polynomials of degree k form a basis of the vector space of the polynomials of degree at most k. Therefore, p can represented by

$$p(x) = \sum_{i=0}^{k} b_i^{(k)} B_i^{(k)}(x), \quad k \geq l. \tag{3}$$

- The coefficients of this expansion are given by ($a_j := 0$ for $j \geq k$ and $j \neq k$)

$$b_i^{(k)} = \sum_{j=0}^{i} \frac{\binom{i}{j}}{\binom{k}{j}} a_j, \quad 0 \leq i \leq k. \tag{4}$$

(*Bernstein coefficients*).

- The Bernstein coefficients can be organized in a multi-dimensional array $B(\mathbf{u}) = (b_i^{(k)})_{0 \leq i \leq k}$, the so-called *Bernstein patch*.

Properties of Bernstein Coefficients

- **Endpoint interpolation property**:

$$b_{0,0,\ldots,0} = a_{0,0,\ldots,0} = p(0,0,\ldots,0), \qquad b_k = \sum_{i=0}^{k} a_i = p(1,1,\ldots,1). \tag{5}$$

- The first partial derivative of the polynomial p with respect to x_s is given by

$$\frac{\partial p}{\partial x_s} = \sum_{i \le k_{s,-1}} b'_i B_{k_{s,-1},i}(x), \tag{6}$$

where

$$b'_i = k_s[b_{i_s,1} - b_i], \quad 1 \le s \le n, \quad x \in \mathbf{u}. \tag{7}$$

- **convex hull property**:The graph of p over $\mathbf{u}$ is contained in the convex hull of the control points.

$$\left\{ \begin{pmatrix} x \\ p(x) \end{pmatrix} \; : \; x \in \mathbf{u} \right\} \subseteq \operatorname{conv} \left\{ \begin{pmatrix} i/k \\ b_i \end{pmatrix} \; : \; 0 \leq i \leq k \right\}. \qquad (8)$$

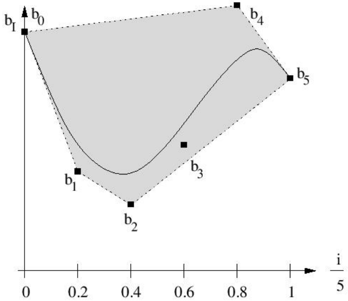

Figure 1:The graph of a degree 5 polynomial and the convex hull (shaded) of its control points (marked by squares).

- **range enclosing property**: For all $x \in \mathbf{u}$

$$\min b_i^{(k)} \leq p(x) \leq \max b_i^{(k)}. \tag{9}$$

Equality holds in the left or right inequality in (9) if and only if the minimum or the maximum, respectively, is attained at a vertex of $\mathbf{u}$, i.e., if $i_j \in \{0, k_j\}$, $j = 1, \ldots, n$.

Simplex

- Let $\mathbf{v}_0, \ldots, \mathbf{v}_n$ be $n+1$ points of $\mathbb{R}^n$. The ordered list $V = [\mathbf{v}_0, \ldots, \mathbf{v}_n]$ is called *simplex of vertices* $\mathbf{v}_0, \ldots, \mathbf{v}_n$.
- *The* realization $|V|$ of the simplex V is the set of $\mathbb{R}^n$ defined as the convex hull of the points $\mathbf{v}_0, \ldots, \mathbf{v}_n$.
- Any vector $x \in \mathbb{R}^n$ can be written as an affine combination of the vertices $\mathbf{v}_0, \ldots, \mathbf{v}_n$ with weights $\lambda_0, \ldots, \lambda_n$ called *barycentric coordinates.*
- *If* $x = (x_1, \ldots, x_n) \in \Delta$, *then* $\lambda = (\lambda_0, \ldots, \lambda_n) = (1 - \sum_{i=1}^n x_i, x_1, \ldots, x_n)$.

- For every multi-index $\alpha = (\alpha_0, \dots, \alpha_n) \in \mathbb{N}^{n+1}$ and $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_n) \in \mathbb{R}^{n+1}$ we write $|\alpha| := \alpha_0 + \ldots + \alpha_n$ and $\lambda^\alpha := \prod_{i=0}^n \lambda_i^{\alpha_i}$.
- Let k be a natural number. The Bernstein polynomials of degree k with respect to V are the polynomials

$$B_\alpha^k := \binom{k}{\alpha} \lambda^\alpha, |\alpha| = k. \tag{10}$$

Bernstein Polynomials

- The Bernstein polynomials of degree k form a basis of the vector space $\mathbb{R}_k[\mathbf{X}]$ of polynomials of degree at most k. Therefore, p can be uniquely represented as

$$p(x) = \sum_{|\alpha|=k} b_\alpha(p,k,V) B_\alpha^k, \quad l \leq k. \tag{11}$$

- The coefficients of this expansion are given by ($a_j := 0$ for $j \geq k$ and $j \neq k$)

$$b_\alpha(p,k,\Delta) = \sum_{\beta \leq \alpha} \frac{\binom{\alpha}{\beta}}{\binom{k}{\beta}} a_\beta \tag{12}$$

(*Bernstein coefficients*).

Bivariate Case over a simplex

A bivariate polynomial of degree l in power form can be expressed as

$$
\begin{aligned}
p(x) &= \sum_{|\beta|\le l} a_\beta x^\beta \\
&= X_1 A X_2, \qquad (13)
\end{aligned}
$$

where

$$X_1 = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{l_1} \end{bmatrix}, \qquad (14)$$

$$X_2 = \begin{bmatrix} 1 & x_2 & x_2^2 & \dots & x_2^{l_2} \end{bmatrix}, \qquad (15)$$

$$A = \begin{bmatrix} a_{00} & a_{01} & \dots & a_{0l_1} \\ a_{10} & a_{11} & \dots & a_{1l_2} \\ \vdots & \vdots & \dots & \vdots \\ a_{l_10} & a_{l_11} & \dots & a_{l_1l_2} \end{bmatrix}. \qquad (16)$$

A bivariate polynomial in the simplicial Bernstein form can be expressed as

$$
\begin{aligned}
p(x) &= \sum_{|\alpha|=k} b_{\alpha_1,\alpha_2} \frac{x_1^{\alpha_1}}{\alpha_1!} \frac{x_2^{\alpha_2}}{\alpha_2!} \frac{(1-x_1-x_2)^{k-\alpha_1-\alpha_2}}{(k-\alpha_1-\alpha_2)!} \\
&= X_1 M X_2, \qquad (17)
\end{aligned}
$$

where

$$X_1 = \begin{bmatrix} 1 & x_1 & \frac{x_1^2}{2!} & \dots & \frac{x_1^{\alpha_1}}{\alpha_1!} \end{bmatrix}, \qquad (18)$$

$$X_2 = \begin{bmatrix} 1 & x_2 & \frac{x_2^2}{2!} & \dots & \frac{x_2^{\alpha_2}}{\alpha_2!} \end{bmatrix}, \qquad (19)$$

$$M = \begin{bmatrix} \frac{b_{00}(1-x_1-x_2)^k}{k!} & \frac{b_{01}(1-x_1-x_2)^{k-1}}{(k-1)!} & \dots & \frac{b_{0(k-1)}(1-x_1-x_2)}{1!} & b_{0k} \\ \frac{b_{10}(1-x_1-x_2)^{k-1}}{(k-1)!} & \frac{b_{11}(1-x_1-x_2)^{k-2}}{(k-2)!} & \dots & b_{1k}(1-x_1-x_2) & 0 \\ \vdots & \vdots & \vdots & \iddots & \vdots \\ b_{k0} & 0 & 0 & \dots & 0 \end{bmatrix} \qquad (20)$$

The 2-dimensional array for the Bernstein coefficients can be obtained as

$$B = \frac{1}{k!}(U_{x_2}(U_{x_1}W)^T)^T = \begin{bmatrix} b_{00} & b_{01} & \dots & b_{0l_1} \\ b_{10} & b_{11} & \dots & 0 \\ \vdots & \vdots & \iddots & \vdots \\ b_{l_1 0} & 0 & \dots & 0 \end{bmatrix}, \tag{21}$$

where

$$U_{x_1} \quad = \quad U_{x_2} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & \binom{2}{1}2! & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \binom{k}{1}1! & \binom{k}{2}2! & \dots & 1 \end{bmatrix}, \tag{22}$$

$$W \quad = \quad \begin{bmatrix} a_{00}k! & a_{01}(k-1)! & \dots & a_{0(k-1)}1! & a_{0k} \\ a_{10}(k-1)! & a_{11}(k-2)! & \dots & a_{1(k-1)}1! & 0 \\ \vdots & \vdots & \dots & \iddots & \vdots \\ a_{k0} & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{23}$$

Multidimensional case

The polynomial coefficients given by

$$\begin{array}{cccc}
a_{000\ldots 0} & a_{100\ldots 0} & \cdots & a_{l_1 00\ldots 0} \\
a_{010\ldots 0} & a_{110\ldots 0} & \cdots & a_{l_1 10\ldots 0} \\
\vdots & \vdots & \vdots & \vdots \\
a_{0 l_2 0\ldots 0} & a_{1 l_2 0\ldots 0} & \cdots & a_{l_1 l_2 0\ldots 0} \\
\vdots & \vdots & \ldots & \vdots \\
a_{00 l_3\ldots 0} & a_{10 l_3\ldots 0} & \cdots & a_{l_1 0 l_3\ldots 0} \\
a_{01 l_3\ldots 0} & a_{11 l_3\ldots 0} & \cdots & a_{l_1 1 l_3\ldots 0} \\
a_{0 l_2 l_3\ldots 0} & a_{1 l_2 l_3\ldots 0} & \cdots & a_{l_1 l_2 l_3\ldots 0} \\
\vdots & \vdots & \ldots & \vdots \\
a_{0 l_2 l_3\ldots l_n} & a_{1 l_2 l_3\ldots l_n} & \cdots & a_{l_1 l_2 l_3\ldots l_n}
\end{array}.$$

The Bernstein coefficients given by

$$B = \frac{1}{k!}(U_{x_n}\ldots(U_{x_i}\ldots(U_{x_3}(U_{x_2}(U_{x_1}W)^T)^T)^{T\cdots})^{T\cdots})^T, \tag{24}$$

where W can be obtained by multiplying the entries $a_{i_1 i_2\ldots i_n}$ of A by $(k-\sum_{r=1}^{n} i_r)!$ and $U_{x_i} = U_{x_1}$ for all $i = 2, 3, \ldots, n$ (they are given in equation (22)).

The partial derivative with respect to x_s of p in the simplicial Bernstein form is

$$p_r'(x) = \sum_{|\alpha|=k-1} b_\alpha'(p, k-1, V) B_\alpha^{(k-1)}(x) = k \sum_{|\alpha|=k-1} (b_\alpha - b_{\alpha_{s,-1}}) B_{\alpha_{s,-1}}^{(k-1}(x) \quad (25$$

In the two-dimensional case, the Bernstein coefficients of p over the standard simplex Δ are given as

$$\begin{bmatrix} b_{00} & b_{01} & \dots & b_{0l_1} \\ b_{10} & b_{11} & \dots & 0 \\ \vdots & \vdots & \iddots & \vdots \\ b_{l_1 0} & 0 & \dots & 0 \end{bmatrix}. \tag{26}$$

The Bernstein coefficients of $\frac{\partial p}{\partial x_1}$ over Δ are given as

$$B' = \begin{bmatrix} b_{10} - b_{00} & b_{11} - b_{01} & \dots & b_{1(l_2-1)} - b_{0(l_2-1)} \\ b_{20} - b_{10} & b_{21} - b_{11} & \dots & 0 \\ \vdots & \vdots & \iddots & \vdots \\ b_{l_1 0} - b_{(l_1-1)0} & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} b_{00}' & b_{01}' & \dots & b_{0l_2}' \\ b_{10}' & b_{11}' & \dots & 0 \\ \vdots & \vdots & \iddots & \vdots \\ b_{(l_1-1)0}' & 0 & \dots & 0 \end{bmatrix}.$$

Subdivision

From the Bernstein coefficients $b_i^{(k)}$ of p over $\mathbf{u}$, we can compute by the de Casteljau algorithm the Bernstein coefficients over sub-boxes $\mathbf{u}_1$ and $\mathbf{u}_2$ resulting from subdividing $\mathbf{u}$ in the s-th direction, i.e.,

$$\begin{aligned} \mathbf{u}_1 &:= [0,1] \times \ldots \times [0,\lambda] \times \ldots \times [0,1], \\ \mathbf{u}_2 &:= [0,1] \times \ldots \times [\lambda,1] \times \ldots \times [0,1], \end{aligned} \tag{27}$$

for some $\lambda \in (0,1)$.

De Casteljau algorithm

By starting with $B^0(\mathbf{u}) = B(\mathbf{u})$ we set for $k = 1, 2, \ldots, n_s$,

$$b_i^{(k)} = \begin{cases} b_i^{(k-1)}, & i_s \leq k, \\ (1-\lambda) b_{i_{s,-1}}^{(k-1)} + \lambda b_i^{(k-1)}, & k \leq i_s. \end{cases} \quad (28)$$

To obtain the new coefficients, the above formula is applied for $i_j = 0, 1, \ldots, j = 1, 2, \ldots, s-1, s+1, \ldots, l$. Then,

$$b_i(\mathbf{u}_1) = b_i^{(n_s)}, \quad (29)$$

$$b_i(\mathbf{u}_2) = b_{i_1, i_2, \ldots, l_s, \ldots, i_n}^{(n_s - i_s)} \quad (30)$$

The Bernstein patch over $\mathbf{u}_1$ is given by

$$B(\mathbf{u}_1) = B^{(n_s)}(\mathbf{u}),$$

and Bernstein patche $B(\mathbf{u}_2)$ over the sub-box u_2 are obtained as intermediate values in this computation.

Subdivision Direction Selection

- Subdivision can be performed in any coordinate direction. It may be advantageous to subdivide in a particular direction to increase the probability of finding a sharp range enclosure.
- The merit function for the subdivision in coordinate direction

$$K = \min\left\{j : j \in \{1, 2, \ldots, l\}, y(j) = \max\left\{y(s), s = 1, 2, \ldots, l\right\}\right\}. \quad (31)$$

  - **Rule A**: $y(s) = \mathrm{wid}(\mathbf{u}_s)$, where $\mathrm{wid}(\mathbf{u}_s)$ is the width(edge length) of the box in the direction s.
  - **Rule B**: $y(s) = \max\left|\frac{\partial p}{\partial x_s}\right| = \max_{i \leq k_{s,-1}} |b_{i_{s,1}} - b_i|$.
  - **Rule C**: $y(s) = [\max_{i \leq k_{s,-1}}(b_{i_{s,1}} - b_i) - \min_{i \leq k_{s,-1}}(b_{i_{s,1}} - b_i)]\ \mathrm{wid}\ \mathbf{u}_s$.

The Bernstein coefficients can be calculated over a sub-box by premultiplying the matrix representing the Bernstein patch $B(\mathbf{u})$ by matrices which depend on the subdivision parameter point λ.

E.g., when the subdivision is applied in the first coordinate direction, then the Bernstein patches over $\mathbf{u}_1$ and $\mathbf{u}_2$ are given as

$$B(\mathbf{u}_1) = L_m L_{m-1} \dots L_1 B(\mathbf{u}), \quad B(\mathbf{u}_2) = L^*_m L^*_{m-1} \dots L^*_1 B(\mathbf{u}), \tag{32}$$

where for $t = 1, 2, \dots, m$

$$L_t = \begin{bmatrix} I_t & 0 \\ (1-\lambda)E_{1,t} & M_{m+1-t} \end{bmatrix}, \qquad L^*_t = \begin{bmatrix} M^*_{m+1-t} & \lambda E_{m+1-k,1} \\ 0 & I_t \end{bmatrix}. \tag{33}$$

where I_t is the txt identity matrix, $E_{1,t}$, $E_{m+1-k,1} \in \mathbb{R}^{m-1-t,t}$ with all of their entries are zero except the $(1,t)$ and $(m+1-t,1)$ entry is 1, respectively, and $M_{m+1-t} = (mij)$, $M^*_{m+1-t} = (m^*_{ij}) \in \mathbb{R}^{m+1-t,m+1-t}$,

$$m_{ij} := \begin{cases} \lambda & \text{if } i = j, \\ 1-\lambda, & \text{if } i = j+1, \\ 0, & \text{if } otherwise, \end{cases} \qquad m^*_{ij} := \begin{cases} 1-\lambda, & \text{if } i = j, \\ \lambda, & \text{if } i = j-1, \\ 0, & \text{if } otherwise. \end{cases} \tag{34}$$

The matrix method has the following advantages over the de Casteljau algorithm:

- Elegant.
- Easier to handle.
- The Bernstein coefficients over each sub-box appear directly.
- The matrix method of computation of the Bernstein coefficients over each sub-box and the matrix method proposed by Ray and Nataraj [6](for computation of the Bernstein coefficients over the entire box) are complement each other. Thus, the Bernstein coefficients can be calculated by using matrix methods only.

Notations

- $\mathbb{IR}$: set of the compact, nonempty real intervals $[a] = [\underline{a}, \overline{a}]$, $\underline{a} \leq \overline{a}$.
- $\mathbb{IR}^n$: set of n-vectors with components from $\mathbb{IR}$, *interval vectors*.
- $\mathbb{IR}^{n,n}$: set of n-by-n matrices with entries from $\mathbb{IR}$, *interval matrices*.
- Elements from $\mathbb{IR}^n$ and $\mathbb{IR}^{n,n}$ may be regarded as vector intervals and matrix intervals, respectively, w.r.t. the usual entrywise partial ordering, e.g.,

$$\begin{aligned} [A] &= ([a_{ij}])_{i,j=1}^{n} = \left([\underline{a}_{ij}, \overline{a}_{ij}]\right)_{i,j=1}^{n} \\ &= [\underline{A}, \bar{A}], \quad \text{where } \underline{A} = \left(\underline{a}_{ij}\right)_{i,j=1}^{n}, \ \overline{A} = (\overline{a}_{ij})_{i,j=1}^{n}. \end{aligned}$$

- A *vertex matrix* of $[A]$ is a matrix $A = (a_{ij})_{i,j=1}^{n}$ with $a_{ij} \in \{\underline{a}_{ij}, \overline{a}_{ij}\}$, $i, j = 1, \ldots, n$.

- An interval matrix $[A] \in \mathbb{R}^{n,n}$ can be represented as

$$[A] = [A_c - \Delta, A_c + \Delta] = \{A : A_c - \Delta \leq A \leq A_c + \Delta\} \quad (35)$$

with $A_c, \Delta \in \mathbb{R}^{n,n}$ and symmetric, $\Delta \geq 0$.

- We introduce an auxiliary index set

$$Y := \{z \in \mathbb{R}^n; |z_j| = 1 \text{ for } j = 1, 2, \dots n\}, \text{with cardinality } 2^n.$$

- For each $z \in Y$ define the matrix

$$A_z := A_c - T_z \Delta T_z, \quad (36)$$

where T_z is an *nxn* diagonal matrix with diagonal vector z.

- $A_z \in [A]$ for each $z \in Y$. The number of mutually different matrices A_z is at most 2^{n-1}.

Test for the convexity of a polynomial p

Theorem (Bialas and Garloff, 1984; Rohn, 1994)

Let $[A]$ be a square interval matrix. Then, $[A]$ is positive semidefinite if and only if A_z is positive semidefinite for each $z \in Y$.

Second order convexity condition

Let the function $f : \mathbb{R}^n \to \mathbb{R}$ be twice differentiable, that is, its Hessian matrix $\nabla^2 f$ exists at each point in the dom$f$. Then f is convex if and only if the dom$f$ is convex and its Hessian matrix is positive semidefinite for all $x \in \text{dom} f$, i.e.,

$$\nabla^2 f \succeq 0.$$

References

[1] G.T. Cargo and O. Shisha, The Bernstein Form of a Polynomial, *J. Res. Nat. Bur. Standards* 70(B):79–81, 1966.

[2] J. Garloff, Convergent Bounds for the Range of Multivariate Polynomials, *Interval Mathematics 1985*, K. Nickel, Ed., *Lecture Notes in Computer Science*, vol. 212, Springer, Berlin, Heidelberg, New York, 37–56, 1986.

[3] J. Garloff, The Bernstein Algorithm, *Interval Comput.* 2:154–168, 1993.

[4] P.S.V Nataraj and M. Arounassalame, A New Subdivision Algorithm for the Bernstein Polynomial Approach to Global Optimization, *Int. J. Automat. Comput.* 4(4):342–352, 2007.

[5] S. Ray and P.S.V. Nataraj, An Efficient Algorithm for Range Computation of Polynomials Using the Bernstein Form, *J. Global Optim.* 45:403–426, 2009.

[6] S. Ray and P.S.V. Nataraj, A Matrix Method for Efficient Computation of Bernstein Coefficients, *Reliab. Comput.* 17(1):40–71,

Thank you for your attention!