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
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

Test for the convexity of a polynomial
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 $u := [0, 1]^n$, since any compact nonempty box $x$ of $\mathbb{R}^n$ can be mapped affinely upon $u$.
- 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 $\sum_{i=0}^k := \sum_{i_1=0}^{k_1} \cdots \sum_{i_n=0}^{k_n}$ 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 $0 \leq i_s + q \leq k_s$. 
Let \( p \) be an \( n \)-variate polynomial of degree \( l \)
\[
p(x) = \sum_{i=0}^{l} a_{i} x^{i}.
\] (1)

The \( i \)-th Bernstein polynomial of degree \( k \), \( k \geq l \), is the polynomial
\[
B_{i}^{(k)}(x) = \binom{k}{i} x^{i} (1 - x)^{k-i}.
\] (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 be represented by

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

The coefficients of this expansion are given by

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

(Bernstein coefficients).

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

- **Endpoint interpolation property:**

\[
b_0,0,...,0 = a_0,0,...,0 = p(0,0,...,0), \quad b_k = \sum_{i=0}^{k} a_i = p(1,1,...,1). \quad (5)
\]

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

\[
\frac{\partial p}{\partial x_s} = \sum_{i \leq k_s-1} b'_i B_{k_s,-1,i}(x), \quad (6)
\]

where

\[
b'_i = k_s [b_{i,s,1} - b_i], \quad 1 \leq s \leq n, \quad x \in u. \quad (7)
\]
convex hull property: The graph of $p$ over $u$ is contained in the convex hull of the control points.

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

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 $x \in u$

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

(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 $u$, i.e., if $i_j \in \{0, k_j\}$, $j = 1, \ldots, n$. 
Let $v_0, \ldots, v_n$ be $n + 1$ points of $\mathbb{R}^n$. The ordered list $V = [v_0, \ldots, v_n]$ is called *simplex of vertices* $v_0, \ldots, v_n$.

The realization $|V|$ of the simplex $V$ is the set of $\mathbb{R}^n$ defined as the convex hull of the points $v_0, \ldots, v_n$.

Any vector $x \in \mathbb{R}^n$ can be written as an affine combination of the vertices $v_0, \ldots, v_n$ 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 $\alpha = (\alpha_0, \ldots, \alpha_n) \in \mathbb{N}^{n+1}$ and $\lambda = (\lambda_0, \ldots, \lambda_n) \in \mathbb{R}^{n+1}$ 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

$$B^k_\alpha := \binom{k}{\alpha} \lambda^\alpha, |\alpha| = k. \quad (10)$$
Bernstein Polynomials

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

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

- The coefficients of this expansion are given by

$$(a_j := 0 \text{ for } j \geq k \text{ and } j \neq k)$$

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

(Bernstein coefficients).
Bivariate Case over a simplex

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

\[
p(x) = \sum_{|\beta| \leq l} a_\beta x^\beta
\]

\[
= X_1 A X_2,
\]

where

\[
X_1 = \begin{bmatrix}
1 & x_1 & x_1^2 & \cdots & x_1^l_1
\end{bmatrix},
\]

\[
X_2 = \begin{bmatrix}
1 & x_2 & x_2^2 & \cdots & x_2^l_2
\end{bmatrix},
\]

\[
A = \begin{bmatrix}
a_{00} & a_{01} & \cdots & a_{0l_1}
a_{10} & a_{11} & \cdots & a_{1l_2}
\vdots & \vdots & \ddots & \vdots
a_{l_10} & a_{l_11} & \cdots & a_{l_1l_2}
\end{bmatrix}.
\]
A bivariate polynomial in the simplicial Bernstein form can be expressed as

\[
p(x) = \sum_{|\alpha|=k} b_{\alpha_1,\alpha_2} \frac{x_1^{\alpha_1} x_2^{\alpha_2}}{\alpha_1! \alpha_2!} \frac{(1 - x_1 - x_2)^{k-\alpha_1-\alpha_2}}{(k - \alpha_1 - \alpha_2)!}
\]

\[
= X_1 MX_2,
\]

where

\[
X_1 = \begin{bmatrix}
1 & x_1 & \frac{x_1^2}{2!} & \ldots & \frac{x_1^{\alpha_1}}{\alpha_1!}
\end{bmatrix},
\]

\[
X_2 = \begin{bmatrix}
1 & x_2 & \frac{x_2^2}{2!} & \ldots & \frac{x_2^{\alpha_2}}{\alpha_2!}
\end{bmatrix},
\]

\[
M = \begin{bmatrix}
\frac{b_{00}(1-x_1-x_2)^k}{k!} & \frac{b_{01}(1-x_1-x_2)^{k-1}}{(k-1)!} & \ldots & \frac{b_{0(k-1)}(1-x_1-x_2)}{1!} & b_{0k} \\
\frac{b_{10}(1-x_1-x_2)^{k-1}}{(k-1)!} & \frac{b_{11}(1-x_1-x_2)^{k-2}}{(k-2)!} & \ldots & b_{1k}(1-x_1-x_2) & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
b_{k0} & 0 & 0 & \ldots & 0
\end{bmatrix}
\]
The 2-dimensional array for the Bernstein coefficients can be obtained as

$$B = \frac{1}{k!} (U_{x_2} (U_{x_1} W)^T)^T = \begin{bmatrix} b_{00} & b_{01} & \ldots & b_{0l_1} \\ b_{10} & b_{11} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{l_10} & 0 & \ldots & 0 \end{bmatrix}, \quad (21)$$

where

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

and

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

The polynomial coefficients given by

\[
\begin{array}{cccc}
  a_{000} & a_{100} & \cdots & a_{l_100} \\
  a_{010} & a_{110} & \cdots & a_{l_110} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{0l_20} & a_{1l_20} & \cdots & a_{l_1l_20} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{00l_3} & a_{10l_3} & \cdots & a_{l_10l_3} \\
  a_{01l_3} & a_{11l_3} & \cdots & a_{l_11l_3} \\
  a_{0l_2l_3} & a_{1l_2l_3} & \cdots & a_{l_1l_2l_3} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{0l_2l_3\cdots l_n} & a_{1l_2l_3\cdots l_n} & \cdots & a_{l_1l_2l_3\cdots l_n}
\end{array}
\]

The Bernstein coefficients given by

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

where \( W \) can be obtained by multiplying the entries \( a_{i_1i_2\cdots i_n} \) 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

$$p'_r(x) = \sum_{|\alpha|=k-1} b'_\alpha(p, k-1, V) B^{(k-1)}_{\alpha}(x) = k \sum_{|\alpha|=k-1} (b_\alpha - b_{\alpha,s,-1}) B^{(k-1)}_{\alpha,s,-1}(x) \quad (25)$$

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

$$\begin{bmatrix} b_{00} & b_{01} & \cdots & b_{0l_1} \\ b_{10} & b_{11} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{l_10} & 0 & \cdots & 0 \end{bmatrix} \quad (26)$$

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

$$B' = \begin{bmatrix} b_{10} - b_{00} & b_{11} - b_{01} & \cdots & b_{1(l_2-1)} - b_{0(l_2-1)} \\ b_{20} - b_{10} & b_{21} - b_{11} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{l_10} - b_{(l_1-1)0} & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} b'_{00} & b'_{01} & \cdots & b'_{0l_2} \\ b'_{10} & b'_{11} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b'_{(l_1-1)0} & 0 & \cdots & 0 \end{bmatrix}.$$
Subdivision

From the Bernstein coefficients $b_i^{(k)}$ of $p$ over $u$, we can compute by the de Casteljau algorithm the Bernstein coefficients over sub-boxes $u_1$ and $u_2$ resulting from subdividing $u$ in the $s$-th direction, i.e.,

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

(27)

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

By starting with $B^0(u) = B(u)$ 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,

$$b_i(u_1) = b_i^{(n_s)}, \quad (29)$$

$$b_i(u_2) = b_i^{(n_s-i_s)}(u_1, u_2, \ldots, u_s, \ldots, u_n) \quad (30)$$

The Bernstein patch over $u_1$ is given by

$$B(u_1) = B^{(n_s)}(u),$$

and Bernstein patche $B(u_2)$ 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

\[ K = \min \{ j : j \in \{1, 2, \ldots, l\}, y(j) = \max \{ y(s), s = 1, 2, \ldots, l \} \} \]  

(31)

- **Rule A**: \( y(s) = \text{wid}(u_s) \), where \( \text{wid}(u_s) \) is the width (edge length) of the box in the direction \( s \).
- **Rule B**: \( y(s) = \max |\frac{\partial p}{\partial x_s}| = \max_{i \leq k_s, -1} |b_{i_s,1} - b_i| \).
- **Rule C**: \( y(s) = [\max_{i \leq k_s, -1} (b_{i_s,1} - b_i) - \min_{i \leq k_s, -1} (b_{i_s,1} - b_i)] \text{wid } u_s \).
The Bernstein coefficients can be calculated over a sub-box by premultiplying the matrix representing the Bernstein patch \( B(u) \) by matrices which depend on the subdivision parameter point \( \lambda \).

E.g., when the subdivision is applied in the first coordinate direction, then the Bernstein patches over \( u_1 \) and \( u_2 \) are given as

\[
B(u_1) = L_m L_{m-1} \ldots L_1 B(u), \quad B(u_2) = L^*_m L^*_{m-1} \ldots L^*_1 B(u),
\]

where for \( t = 1, 2, \ldots, m \)

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

where \( I_t \) is the \( t \times t \) 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 \( M_{m+1-t} = (m_{ij}), \ M^*_{m+1-t} = (m^*_{ij}) \in \mathbb{R}^{m+1-t,m+1-t} \),

\[
m_{ij} := \begin{cases} 
\lambda & \text{if } i = j, \\
1 - \lambda & \text{if } i = j + 1, \\
0 & \text{if } \text{otherwise},
\end{cases}
\quad \quad m^*_{ij} := \begin{cases}
1 - \lambda & \text{if } i = j, \\
\lambda & \text{if } i = j - 1, \\
0 & \text{if } \text{otherwise}.
\end{cases}
\]
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 $[a] = [a, \bar{a}], a \leq \bar{a}$.
- $\mathbb{IR}^n$: 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.,

  $$[A] = ([a_{ij}])_{i,j=1}^n = ([a_{ij}, \bar{a}_{ij}])_{i,j=1}^n$$
  $$= [A, \bar{A}], \text{ where } A = (a_{ij})_{i,j=1}^n, \bar{A} = (\bar{a}_{ij})_{i,j=1}^n.$$  

- A vertex matrix of $[A]$ is a matrix $A = (a_{ij})_{i,j=1}^n$ with $a_{ij} \in \{ a_{ij}, \bar{a}_{ij} \}$, $i, j = 1, \ldots, n$. 
An interval matrix \([A] \in \mathbb{R}^{n,n}\) can be represented as

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

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

We introduce an auxiliary index set

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

For each \(z \in Y\) define the matrix

\[
A_z := A_c - T_z \Delta T_z,
\] (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 $\text{dom} f$. Then $f$ is convex if and only if the $\text{dom} f$ is convex and its Hessian matrix is positive semidefinite for all $x \in \text{dom} f$, i.e.,

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


Thank you for your attention!