A COMPARATIVE ANALYSIS OF BERNSTEIN TYPE ESTIMATES FOR THE DERIVATIVE OF MULTIVARIATE POLYNOMIALS

SZILÁRD GY. RÉVÉSZ

Abstract. We compare the yields of two methods to obtain Bernstein type pointwise estimates for the derivative of a multivariate polynomial in points of some domain, where the polynomial is assumed to have sup norm at most 1. One method, due to Sarantopoulos, relies on inscribing ellipses into the convex domain \( K \). The other, pluripotential theoretic approach, mainly due to Baran, works for even more general sets, and yields estimates through the use of the pluricomplex Green function (the Siciak-Zaharjuta extremal function). Using the inscribed ellipse method on non-symmetric convex domains, a key role was played by the generalized Minkowski functional \( \alpha(K,x) \). With the aid of this functional, our current knowledge is precise within a constant \( \sqrt{2} \) factor. Recently L. Milev and the author derived the exact yield of this method in the case of the simplex, and a number of numerical improvements were obtained compared to the general estimates known. Here we compare the yields of this real, geometric method and the results of the complex, pluripotential theoretical approaches on the case of the simplex. In conclusion we can observe a few remarkable facts, comment on the existing conjectures, and formulate a number of new hypothesis.

1. Introduction

If a univariate algebraic polynomial \( p \) is given with degree at most \( n \), then by the classical Bernstein-Szegő inequality ([36], [9], [1]) we have

\[
|p'(x)| \leq \frac{n \sqrt{||p||^2_{C[a,b]} - p^2(x)}}{\sqrt{(b-x)(x-a)}} \quad (a < x < b).
\]

This inequality is sharp for every \( n \) and every point \( x \in (a,b) \), as

\[
\sup \left\{ \frac{|p'(x)|}{\sqrt{||p||^2_{C[a,b]} - p^2(x)}} : \deg p \leq n, |p(x)| < ||p||_{C[a,b]} \right\} = \frac{n}{\sqrt{(b-x)(x-a)}}.
\]

We may say that the upper estimate (1) is exact, and the right hand side is just the "true Bernstein factor" of the problem.


Key words and phrases. convex body, generalized Minkowski functional, polynomials on normed spaces, gradient, convex hull, support functional, Bernstein-Szegő inequality, maximal chord, minimal width, pluripotential theory, Siciak-Zaharjuta extremal function, Monge-Ampère equation, complex equilibrium measure, Baran’s Conjecture, Révész-Sarantopoulos Conjecture.

Support of this work was rejected by Hungarian National Foundation for Scientific Research, project # T-047074.

This research was completed during the author’s stay in Sofia, Bulgaria under the exchange program of the Bulgarian and Hungarian Academies of Sciences.

1
Polynomials and continuous polynomials are defined even on topological vector spaces, see e.g. [10]. The set of continuous polynomials over $X$ will be denoted by $\mathcal{P} = \mathcal{P}(X)$ and polynomials in $\mathcal{P}$ with degree not exceeding $n$ by $\mathcal{P}_n = \mathcal{P}_n(X)$.

In the multivariate setting a number of extensions were proved for the classical result (1). However, due to the geometric variety of possible convex sets replacing intervals of $\mathbb{R}$, our present knowledge is still not final. The exact Bernstein inequality is known only for symmetric convex bodies, and we are within a bound of some constant factor in the general, nonsymmetric case.

For more precise notation we may define formally for any topological vector space $X$, a subset $K \subset X$, and a point $x \in K$ the $n^{th}$ "Bernstein factor" as

$$B_n(K, x) := \frac{1}{n} \sup \left\{ \frac{\|Dp(x)\|}{\sqrt{\|p\|^2_{C(K)} - p^2(x)}} : \deg p \leq n, \|p(x)\| < \|p\|_{C(K)} \right\},$$

where $Dp(x)$ is the derivative of $p$ at $x$, and even for an arbitrary unit vector $y \in X$

$$B_n(K, x, y) := \frac{1}{n} \sup \left\{ \frac{\langle Dp(x), y \rangle}{\sqrt{\|p\|^2_{C(K)} - p^2(x)}} : \deg p \leq n, \|p(x)\| < \|p\|_{C(K)} \right\},$$

where $\langle Dp(x), y \rangle$ is the directional derivative in direction $y$ (which equals the value attained by the gradient, as a linear functional, on $y$).

Our aim is to investigate these and related quantities, and to analyze the methods of estimating them.

2. The inscribed ellipse method of Sarantopoulos

Recall that a set $K \subset X$ is called convex body in a normed space (or in a topological vector space) $X$ if it is a bounded, closed convex set that has a non-empty interior. The convex body $K$ is symmetric, iff there exists a center of symmetry $x$ so that reflection of $K$ at $x$ leaves the set invariant, that is, $K = -(K - x) + x = -K + 2x$. In the following we will term $K$ to be centrally symmetric if it is symmetric with respect to the origin, i.e. if $K = -K$. This occurs iff $K$ can be considered the unit ball with respect to a norm $\| \cdot \|_{(K)}$, which is then equivalent to the original norm $\| \cdot \|$ of the space $X$ in view of $B_X, \| \|_{(0, r)} \subset K \subset B_X, \| \|_{(0, R)}$.

The "maximal chord" of $K$ in direction of $v \neq 0$ is

$$\tau(K, v) := \sup \{ \lambda \geq 0 : \exists y, z \in K \text{ s.t. } z = y + \lambda v \}$$

$$= \sup \{ \lambda \geq 0 : K \cap (K + \lambda v) \neq \emptyset \}$$

$$= \sup \{ \lambda \geq 0 : \lambda v \in K - K \}$$

Usually $\tau(K, v)$ is not a "maximal" chord length, but only a supremum, however we shall use the familiar finite dimensional terminology (see for example [38]).

The support function to $K$, where $K$ can be an arbitrary set, is defined for all $v^* \in X^*$ (sometimes only for $v^* \in S^* := \{ v^* \in X^* : \|v^*\| = 1 \}$) as

$$h(K, v^*) := \sup_{K} v^* = \sup \{ \langle v^*, x \rangle : x \in K \},$$
and the width of $K$ in direction $v^* \in X^*$ (or $v^* \in S^*$) is
\[
\begin{aligned}
w(K, v^*) := h(K, v^*) + h(K, -v^*) &= \sup_K v^* + \sup_K (-v^*) = \\
&= \sup \{ \langle v^*, x - y \rangle : x, y \in K \} = 2h(C, v^*) = w(C, v^*).
\end{aligned}
\]
Then the minimal width of $K$ is $\inf_{S^*} w(K, v^*)$ and the sharp inequalities
\[
w(K) \leq \tau(K, v) \leq d(K), \quad w(K) \leq w(K, v^*) \leq d(K),
\]
always hold, even in infinite dimensional spaces, compare [31, §2].

In $\mathbb{R}$ the position of a point $x \in \mathbb{R}$ with respect to the "convex body" $I$ can be expressed simply by $|x|$ (as $\pm x$ occupy symmetric positions). For this in the multivariate case the most frequent tool is the Minkowski functional. For any $x \in X$ the Minkowski functional or \textit{(Minkowski) distance function} [12, p. 57] or \textit{gauge} [27, p. 28] or \textit{Minkowski gauge functional} [26, §1.1(d)] is defined as
\[
\varphi_K(x) := \inf\{\lambda > 0 : x \in \lambda K\}.
\]
Clearly (8) is a norm on $X$ if and only if the convex body $K$ is centrally symmetric with respect to the origin. If $K \subset X$ is a centrally symmetric convex body, then the norm $\| \cdot \|_K := \varphi_K$ can be used successfully in approximation theoretic questions as well. As said above, for $\| \cdot \|_K$ the unit ball of $X$ will be $K$ itself. In case $K$ is nonsymmetric, the role of the so called \textit{generalized Minkowski functional} emerged in the above quantification problem. This generalized Minkowski functional $\alpha(K, x)$ also goes back to Minkowski [21] and Radon [28], see also [11], [31]. There are several ways to introduce it: perhaps the shortest is the following. First let
\[
\gamma(K, x) := \inf \left\{ \frac{2\sqrt{||x - a|| \cdot ||x - b||}}{||a - b||} : a, b \in \partial K, \text{s.t. } x \in [a, b] \right\}.
\]
Then we have
\[
\alpha(K, x) = \sqrt{1 - \gamma^2(K, x)}.
\]
In fact, usefulness of (10) and the possibility of the wide ranging applications stems from the fact that this geometric quantity incorporates quite nicely the geometric aspects of the configuration of $x$ with respect to $K$, which is mirrored by about a dozen (!), sometimes strikingly different-looking, equivalent formulations of it. For the above and many other equivalent formulations with full proofs, further geometric properties and some notes on the applications in approximation theory see [31] and the references therein; for the first appearance of it in approximation theoretical questions see [32].

The \textit{method of inscribed ellipses} was introduced into the subject by Y. Sarantopoulos [33]. This works for arbitrary interior points of any, possibly nonsymmetric convex body. The key of all of the method is the next

\textbf{Lemma 1 (Inscribed Ellipse Lemma, Sarantopoulos, 1991).} Let $K$ be any subset in a vector space $X$. Suppose that $x \in K$ and the ellipse
\[
\mathbf{r}(t) = \cos t a + b \sin t y + x - a \quad (t \in [-\pi, \pi])
\]
lies inside $K$. Then we have for any polynomial $p$ of degree at most $n$ the Bernstein type inequality
\[
|\langle Dp(x), y \rangle| \leq \frac{n}{b} \sqrt{||p||^2_{C(K)} - p^2(x)}.
\]

\[\begin{aligned}
\end{aligned}\]
Theorem 1 (Sarantopoulos, 1991). Let $p$ be any polynomial of degree at most $n$ over the normed space $X$. Then we have for any unit vector $y \in X$ the Bernstein type inequality
\begin{equation}
|\langle Dp(x), y \rangle| \leq \frac{n\sqrt{||p||_{C(K)}^2 - p^2(x)}}{\sqrt{1 - ||x||_{(K)}^2}}.
\end{equation}

Theorem 2 (Sarantopoulos, 1991). Let $K$ be a symmetric convex body and $y$ a unit vector in the normed space $X$. Let $p$ be any polynomial of degree at most $n$. We have
\begin{equation}
|\langle Dp(x), y \rangle| \leq \frac{2n\sqrt{||p||_{C(K)}^2 - p^2(x)}}{\tau(K,y)\sqrt{1 - \varphi^2(K,x)}}.
\end{equation}
In particular, we have
\begin{equation}
\|Dp(x)\| \leq \frac{2n\sqrt{||p||_{C(K)}^2 - p^2(x)}}{w(K)\sqrt{1 - \varphi^2(K,x)}},
\end{equation}
where $w(K)$ stands for the width of $K$.

The above solve the question for the case of a symmetric convex body $K$. However, in the general, non-symmetric case it can be rather difficult to determine, or even to estimate the $b$-parameter of the "best ellipse", what can be inscribed into a convex body $K$ through $x \in K$ and tangential to direction of $y$. Still, we can formalize what we want to find.

Definition 1 (Milev-Révész, 2003). For arbitrary $K \subset X$ and $x \in K$, $y \in X$ the corresponding "best ellipse constants" are the extremal quantities
\begin{equation}
E(K,x,y) := \sup\{b : r \subset K \text{ with } r \text{ as given in (11)}\}.
\end{equation}
Also, for the immediately resulting estimation of the gradient, we defined in [19]
\begin{equation}
E(K,x) := \inf\{E(K,x,y) : y \in X, ||y|| = 1\}.
\end{equation}

Clearly, the inscribed ellipse method yields Bernstein type estimates whenever we can derive some estimate of the ellipse constants. In case of symmetric convex bodies, Sarantopoulos’s Theorems 1 and 2 are sharp; for the nonsymmetric case we know only the following result.

Theorem 3 ((Króó–Révész, [16], 1998). Let $K$ be an arbitrary convex body, $x \in \text{int}K$ and $||y|| = 1$, where $X$ can be an arbitrary normed space. Then we have
\begin{equation}
|\langle Dp(x), y \rangle| \leq \frac{2n\sqrt{||p||_{C(K)}^2 - p^2(x)}}{\tau(K,y)\sqrt{1 - \alpha(K,x)}},
\end{equation}
for any polynomial $p$ of degree at most $n$. Moreover, we also have
\begin{equation}
\|Dp(x)\| \leq \frac{2n\sqrt{||p||_{C(K)}^2 - p^2(x)}}{w(K)\sqrt{1 - \alpha(K,x)}} \leq \frac{2\sqrt{2}n\sqrt{||p||_{C(K)}^2 - p^2(x)}}{w(K)\sqrt{1 - \alpha^2(K,x)}}.
\end{equation}
BERNSTEIN TYPE ESTIMATES FOR MULTIVARIATE POLYNOMIALS

Note that in [16] the best ellipse is not found; for most cases, the construction there gives only a good estimate, but not an exact value of (14) or (15). (In fact, here we quoted [16] in a strengthened form: the original paper contains a somewhat weaker formulation only.)

It is worthy to recall here that geometrically the proof of (16) follows the following idea. To construct an ellipse through \( x \), there parallel to \( y \), and inscribed into \( K \) it suffices to find the best (i.e., of maximal possible \( b \) parameter) such ellipse, which is inscribed within the quadrangle of the vertices of a maximal chord in direction of \( y \) (or, in infinite dimension, some chord in the direction and \( \epsilon \)-almost maximal), and the vertices of the parallel chord through the point \( x \). That ellipse is precisely calculated, and its \( b \)-parameter is estimated independently of the location of these chords (even if they degenerate into one line, in which case the ellipse becomes a line segment). (In general the best \( b \)-parameter could not be calculated, though.) We will remind this geometrical construction later.

One of the most intriguing questions of the topic is the following conjecture, formulated first in [31].

Conjecture A. (Révész–Sarantopoulos, 2001). Let \( X \) be a topological vector space, and \( K \) be a convex body in \( X \). For every point \( x \in \text{int} K \) and every (bounded) polynomial \( p \) of degree at most \( n \) over \( X \) we have

\[
\|Dp(x)\| \leq \frac{2n\sqrt{||p||^2_{C(K)} - p^2(x)}}{w(K)\sqrt{1 - \alpha^2(K, x)}},
\]

where \( w(K) \) stands for the width of \( K \).

3. SOME RESULTS ON THE SIMPLEX

We denote \( |x|_2 := \left( \sum_{i=1}^d x_i^2 \right)^{1/2} \) the Euclidean norm of \( x = (x_1, \ldots, x_d) \in \mathbb{R}^d \). Let

\[
\Delta := \Delta_d := \{(x_1, \ldots, x_d) : x_i \geq 0, i = 1, \ldots, d; \sum_{i=1}^d x_i \leq 1\}
\]

be the standard simplex in \( \mathbb{R}^d \). For fixed \( x \in \text{int} \Delta \), and \( y = (y_1, \ldots, y_d) \), \( |y|_2 = 1 \) the best ellipse constant of \( \Delta \) is, by Definition 1, \( E(\Delta, x, y) \). By a tedious calculation via the Kuhn-Tucker theorem and some geometry, the following was obtained in [19].

Theorem 4 (Milev-Révész, 2003). Let \( p \in \mathcal{P}^d_n \). Then for every \( x \in \text{int} \Delta \) and \( y \in S^{d-1} \) we have

\[
|D_y p(x)| \leq \frac{n\sqrt{||p||^2_{C(\Delta)} - p^2(x)}}{E(\Delta, x, y)},
\]

where \( E(\Delta, x, y) \) has the precise value

\[
E(\Delta, x, y) = \left\{ \frac{y_1^2}{x_1} + \cdots + \frac{y_d^2}{x_d} + \frac{(y_1 + \cdots + y_d)^2}{1 - x_1 - \cdots - x_d} \right\}^{-1/2}.
\]

Note that the inequality

\[
\frac{1}{E(\Delta, x, y)} \leq \frac{2}{\tau(\Delta, y)\sqrt{1 - \alpha(\Delta, x)}}
\]

remains valid even when the parallel chord degenerates to a line.
holds true for every $x \in \text{int} \Delta$ and $y \in S^1$, which is not by chance: the general estimation of (16) must be valid also for $\Delta$, and the precise value, calculated for $\Delta$, can only be better. But equality occurs for some directions, to what we return soon.

From now on let us restrict ourselves to the case $d = 2$. We denote the vertices of $\Delta$ by $O = (0, 0), A = (1, 0), B = (0, 1)$ and the centroid (i.e. mass point) of $\Delta$ by $M = (1/3, 1/3)$. It is calculated in [19] that

$$\alpha(\Delta, x) = 1 - 2r(x)$$

with $r(x) = \min\{x_1, x_2, 1 - x_1 - x_2\} = \begin{cases} x_1, & x \in \Delta OMB \\ x_2, & x \in \Delta OMA \\ 1 - x_1 - x_2, & x \in \Delta AMB \end{cases}$

and if $y = (\cos \varphi, \sin \varphi) (0 \leq \varphi \leq \pi)$ then

$$\tau(\Delta, y) = \begin{cases} 1/(y_1 + y_2), & \varphi \in [0, \pi/2] \\ 1/y_2, & \varphi \in (\pi/2, 3\pi/4] \\ -1/y_1, & \varphi \in (3\pi/4, \pi]. \end{cases}$$

Then it can be calculated that we have equality in (20) exactly for the directions $y = (\cos \varphi, \sin \varphi)$ with $\varphi = 0, \pi/2, 3\pi/4$ and its ± contemporaries (i.e. at $\varphi + \pi \mathbb{Z}$ with these values of $\varphi$), and for some corresponding values of the point $x$.

Why is that so? For these, and only these vectors can we have a coincidence of the above geometrical figure, the quadrangle occurring in the proof of (16), and the exact domain into what we must really inscribe the ellipse through $x$ and there parallel to $y$: for all other directions the maximal chord in direction of $y$ lies strictly inside $\Delta$ and another ellipse, slightly stretched behind that chord, can also be inscribed. Therefore, it is geometrically natural that it occurs only for these directions that nothing better can be obtained, than the ellipse calculated in Theorem 3, while for other directions precise calculation of the best ellipse must always yield a better ellipse constant.

In [19] then the following estimates were deduced from Theorem 4.

**Proposition 5 (Milev-Révész, 2003).** Let $p \in P^2_n$. Then for every $x \in \text{int} \Delta$ we have

$$|Dp(x)|_2 \leq n \sqrt{||p||^2(\Delta) - p^2(x)} / E(\Delta, x),$$

where

$$E(\Delta, x) = \sqrt{\frac{2x_1x_2(1 - x_1 - x_2)}{x_1(1 - x_1) + x_2(1 - x_2) + D(x)}}$$

with

$$D(x) := \sqrt{(x_1(1 - x_1) + x_2(1 - x_2))^2 - 4x_1x_2(1 - x_1 - x_2)^2} = \sqrt{[x_1(1 - x_1) - x_2(1 - x_2)]^2 + 4x_1^2x_2^2} > 0 \quad (\forall x \in \text{int}\Delta).$$

From this the following improvements were achieved in Theorem 3 for the special case of $K = \Delta$. 

Proposition 6 (Milev-Révész, 2003). Let $p \in \mathcal{P}_n^2$ and $||p||_{C(\Delta)} = 1$. Then for every $x \in \text{int} \ \Delta$ we have
\begin{equation}
|Dp(x)|_2 \leq \frac{\sqrt{3}n\sqrt{||p||^2_{C(\Delta)} - p^2(x)}}{w(\Delta)\sqrt{1 - \alpha(\Delta, x)}}.
\end{equation}
Furthermore, using the conjectural quantity $\sqrt{1 - \alpha^2(\Delta, x)}$ on the right, we even have
\begin{equation}
|Dp(x)|_2 \leq \frac{\sqrt{3 + \sqrt{5}}n\sqrt{||p||^2_{C(\Delta)} - p^2(x)}}{w(\Delta)\sqrt{1 - \alpha^2(\Delta, x)}}.
\end{equation}

The result (27) improve the constant in Theorem 3 but falls short of Conjecture A, since $2\sqrt{2} = 2.8284\ldots > \sqrt{3 + \sqrt{5}} = 2.2882\ldots > 2$. On the way proving these, it was noted that no better constants follow from the inscribed ellipse method, interpreted so that $E(K, x)$ is considered the yield of the ellipse method. To this we shall return.

4. Baran’s pluripotential theoretic method

Another method of considerable success in proving Bernstein (and Markov) type inequalities is the pluripotential theoretical approach. Classically, all that was considered only in the finite dimensional case, but nowadays even the normed spaces setting is cultivated. To explain these, one needs an understanding of complexifications of real normed spaces, see e.g. [23, 5], as well as the Siciak-Zaharjuta extremal function $V(z)$. We start with a formulation which is perhaps easier to digest. That is very much like the Chebyshev problem, cf. [31, §8], except that we consider it all over the complexification $Y := X + iX$ of $X$, take logarithms, and after normalization by the degree, merge the information derived by all polynomials of any degree into one clustered quantity. Namely, for any bounded $E \subset Y$, $V_E$ vanishes on $E$, while outside $E$ we have the definition
\begin{equation}
V_E(z) := \sup \left\{ \frac{1}{n} \log |p(z)| : 0 \neq p \in \mathcal{P}_n(Y), \ |p||_E \leq 1, \ n \in \mathbb{N} \right\}, \quad (z \notin E)
\end{equation}
For $E \subset X$ one can easily restrict even to $p \in \mathcal{P}(X)$.

Note that $\log |p(z)|$ is a plurisubharmonic function (PSH, for short), as its one complex dimensional restrictions are just logarithms of univariate polynomials over $\mathbb{C}$, which are sums of zero factors each being the potential of a Dirac measure at the particular zero point. Therefore, after normalization by the degree, $(1/n)\log |p(z)|$ has very regular growth towards infinity: it is at most $\log |z| + O(1)$, in any direction. So it is reasonable to consider now the Lelong class of all such functions:
\begin{equation}
\mathcal{L}(E) := \{ u \ PSH : u|_E \leq 0, \ u(z) \leq \log |z| + O(1) \text{ in any direction} \}
\end{equation}
and to define
\begin{equation}
U_E(z) := \sup \{ u(z) : u \in \mathcal{L}(E) \}.
\end{equation}
This function may be named the pluricomplex Green function. The Siciak-Zaharjuta Theorem says that (30) and (28) are equal, at least as long as $E$ is compact, which we now assume together with $E$ being a non-pluripolar set. Then, as suprema of PSH functions (subharmonic functions on all complex "lines"), they are, modulo upper semicontinuous regularization, PSH themselves. They play a central role in the theory.
The extension of the Laplace- and Poisson equations is the so-called complex Monge-Ampere equation:

\[(\partial \partial u)^d := d!4^d \det \left[ \frac{\partial^2 u}{\partial z_j \partial z_k}(z) \right] dV(z),\]

where \(dV(z) = dx_1 \wedge dy_1 \wedge \cdots \wedge dx_d \wedge dy_d\) is just the usual volume element in \(\mathbb{C}^d\). At first this equation is applied only to smooth functions \(u \in \text{PSH} \cap C^2\), say, but due to the work of Bedford and Taylor \([6]\), the operator extends, in the appropriate sense, even to the whole set of locally bounded PSH functions (which covers the case of \(V_E\) for \(E\) non-pluripolar, see e.g. \([15]\)). Therefore, it makes sense to consider

\[(\partial \partial V_E)^d,\]

which is then a compactly supported measure \(\lambda_E\) and is called the complex equilibrium measure of the set \(E\). It is shown in the theory \([6]\) that in fact the measure is kind of normalized, as \(\lambda_E(\mathbb{C}^d) = \lambda_E(\hat{E}) = (2\pi)^d\).

For the theory related to this function and some recent developments concerning Bernstein and also Markov type inequalities for convex bodies or even more general sets, we refer to \([2, 3, 4, 5, 6, 15, 18, 17, 25]\).

There are further yields of the theory of PSH functions, when applied to the Bernstein problem: here we present a few results of Miroslav Baran. For more precise notation now we introduce (interpreting 0/0 as 0 here)

Definition 2.

\[G(E, x) := \left\{ \frac{\text{grad} \, p(x)}{n \sqrt{||p||^2 - p(x)^2}} : 0 \neq p \in \mathcal{P}_n, n \in \mathbb{N} \right\},\]

and following Baran we consider also

\[\bar{G}(E, x) := \text{con} \, G(E, x).\]

Clearly for any compact \(E \subset \mathbb{R}^d\) \(\sup_{n \in \mathbb{N}} B_n(E, x) = \sup_{u \in G(E, x)} ||u||\) holds.

Theorem 7 (Baran, 1995). Let \(E\) be a compact subset of \(\mathbb{R}^d\) with nonempty interior. Then the equilibrium measure \(\lambda_E\) is absolutely continuous in the interior of \(E\) with respect to the Lebesgue measure of \(\mathbb{R}^d\). Denote its density function by \(\lambda(x)\) for all \(x \in \text{int} \, E\). Then we have \(\frac{1}{d} \lambda(x) \geq \text{vol} \, \bar{G}(E, x)\) for a.a. \(x \in \text{int} \, E\). Moreover, if \(E\) is a symmetric convex domain of \(\mathbb{R}^d\), then here we have exact equality.

Conjecture B. (Baran, 1995). We have \(\frac{1}{d} \lambda(x) = \text{vol} \, \bar{G}(E, x)\) even if \(E\) is a non-symmetric convex body in \(\mathbb{R}^d\).

Now consider \(E = K \subset X\), where \(K\) is now a convex body. Our more precise results in \([30]\), see also \([31, \S 8]\), yield \(V_K(x) = \alpha(K, x) + \sqrt{\alpha(K, x)^2 - 1}\). However, in the Bernstein problem the values of \(V_K\) are much more of interest for complex points \(z = x + iy\), in particular for \(x \in K\) and \(y\) small and nonzero. More precisely, the important quantity is the normal (sub)derivative

\[D^+_y V_E(x) := \liminf_{\epsilon \to 0} \frac{V_E(x + i \epsilon y)}{\epsilon},\]
as this quantity occurs in the next estimation of the directional derivative and thus also in the gradient.

**Theorem 8 (Baran, 1994 & 2004).** Let \( E \subset X \) be any bounded, closed set, \( x \in \text{int} E \) and \( 0 \neq y \in X \). Then for all \( p \in \mathcal{P}_n(X) \) we have

\[
\langle (Dp(x), y) \rangle \leq n D^+_y V_E(x) \sqrt{\|p\|_E^2 - p(x)^2}.
\]

**Proof.** In fact, \([2]\) contains this only for \( \mathbb{R}^d \) and partial derivatives; the same estimate in case of infinite dimensional spaces are considered in \([5]\), but only for symmetric convex bodies. The same estimate occurs, also without proof but with reference to Baran, even in the recent publication \([7]\).

For a proof for arbitrary directions \( y \in \mathbb{R}^d \) one can consider a rotation (orthogonal transformation) \( A : \mathbb{R}^d \to \mathbb{R}^d \), taking \( e_1 \) to \( y \), and can take \( \tilde{E} := A^{-1}(E) \). For this set we have \( V_{\tilde{E}}(z) = V_{A^{-1}E}(z) = V_E(Az) \). Calculating the upper partial (sub)derivative \( \partial^+_i := D^+_e \) we get

\[
\partial^+_i V_{\tilde{E}}(x) = \lim_{\epsilon \to 0^+} V_E(Ax + i\epsilon e_1) = \lim_{\epsilon \to 0^+} V_E(Ax + i\epsilon y) = D^+_y V_E(Ax),
\]

(37) so rotating back (i.e. applying the same at a rotated point \( \tilde{x} := A^{-1}x \)) we find

\[
D^+_y V_E(x) = \partial^+_i V_{\tilde{E}}(A^{-1}x) = \partial^+_i V_{\tilde{E}}(\tilde{x}).
\]

Consider now the rotated polynomial \( q(x) := p(Ax) \), which is of the same bound on \( \tilde{E} \) as \( p \) on \( E \), and satisfies \( Dp = D(q \circ A^{-1}) = Dq \circ A^{-1} A^{-1} \) in view of the chain rule. Hence from (38) and applying the estimate (36) only for the first partial derivative of \( q \) on \( \tilde{E} \) and at an arbitrary point \( A^{-1}x = \tilde{x} \in \text{int} \tilde{E} \) (corresponding to \( x \in \text{int} E \)) we finally obtain

\[
\langle (Dq(x), A^{-1} e_1) \rangle = \langle (Dq(\tilde{x}), e_1) \rangle \\
\leq n D^+_e V_{\tilde{E}}(\tilde{x}) \sqrt{|q(\tilde{x})|^2 - q(\tilde{x})^2} = n D^+_y V_E(x) \sqrt{\|p\|_E^2 - p(x)^2}.
\]

It is not obvious, how such theoretical estimates can be applied to concrete cases. First, one has to find the precise value of \( V_E \), in such a precision, that even the derivative can be computed: then the derivatives must be obtained and only then do we really have something. However, even that is addressed by considering the Bedford-Taylor theory of the Monge-Ampere equation and the equilibrium measure \([6]\), as the density of the equilibrium measure gives the extremal function. In some concrete applications all that may be calculated, a particular example (see \([4, \text{Example 4.8}]\)) being the following.

**Proposition 9 (Baran, 1995).** The extremal function of the standard simplex in \( \mathbb{R}^d \) is \( V_\Delta(z) = \log |h(|z_1| + |z_2| + \cdots + |z_n| + |1 - (z_1 + z_2 + \cdots + z_n)|)|. \) Here \( h(z) := z + \sqrt{z^2 - 1} \) is inverse to the Joukowski mapping \( \zeta \to (1/2)(\zeta + 1/\zeta) \), with that choice of the square-root that it is positive for positive \( z \) exceeding 1, so that \( h \) maps to the exterior of the unit disk.

From this and the calculation with the rotated directions above, we calculate\(^1\)

\(^1\)The same formula is mentioned in the recent publication \([7]\), see p. 145.
Hypothesis A. The pluripotential theoretical estimate (36) of Baran, calculated for the standard simplex of $\mathbb{R}^d$ in (40), gives the result exactly identical to (18), obtained from the inscribed ellipse method.

Much remains to explain in this striking coincidence, the first being the next.

Hypothesis A. Let $K \subset X$ be a convex body. Then for all points $x \in \text{int } K$ the inscribed ellipse method and the pluripotential theoretical method of Baran results in exactly the same estimate, i.e. for all $y \in S^*$ we have

$$D^+_y V_K(x) = \frac{1}{E(K, x, y)}.$$
5. Further geometric calculations

At this point it seems worthy to formulate a few naturally occurring assumptions.

**Hypothesis B.** Let $K \subset X$ be convex body. Then for all points $x \in \text{int} K$ the exact Bernstein factor is just what results from the pluripotential theoretical method of Baran:

\begin{equation}
B_n(K, x) = \sup_{y \in S^+} D_y^+ V_K(x) .
\end{equation}

**Hypothesis C.** Let $K \subset X$ be convex body. Then for all points $x \in \text{int} K$ the exact Bernstein factor is just what results from the inscribed ellipse method of Sarantopoulos:

\begin{equation}
B_n(K, x) = \frac{1}{E(K, x)} .
\end{equation}

These hypothesis are certainly not true for the directional derivatives of all directions $y \in S^*$, where both methods can be improved upon for some $y$, as is seen below. Care has to be exercised in formulating conjectures and hypothesis in these matters: the situation is more complex than one might like to have, and the simple heuristics of extending the results of the symmetric case do fail sometimes. In this respect see also [8, 17, 18] and also [7], where another case of deviation from symmetric case extension is observed for the so-called "Baran metric" on the simplex.

There is an important and immediate observation we did not utilize until here. Namely, we found methods (actually, two equivalently strong ones) to estimate $D_y p(x)$. However, if we are looking for the total derivative $\nabla p(x)$, then the estimate we used was only the trivial $||\nabla p(x)|| \leq \sup_{y \in S} |D_y p(x)|$. Can we do any better? Yes, depending on the estimating functions we have for $D_y p(x)$, we can.

Consider e.g. the estimates from Theorem 3, which was got back also for the simplex and thus the triangle $\Delta$. For the triangle we have an explicit computation of the maximal chords $\tau(\Delta, x)$, c.f. (22), and also of the generalized Minkowski functional $\alpha(\Delta, x)$, see (21), so everything is explicit and we can compute the estimating functions. As an example, consider e.g. the point $M := (\frac{1}{3}, \frac{1}{3})$ and compute all quantities involved in the normalization of the directional derivative estimates. As a result, we can exactly determine the arising domain $H(\Delta, M)$.

In fact, it turns out that the domain $H(\Delta, M)$ what the general estimates of Theorem 3 describe is a fleecy-cloud like domain which is symmetric with respect to the origin, and its upper half is (the part above the $x$-axis of) the union of three disks:

$D((\sqrt{\frac{3}{2}}, \sqrt{\frac{3}{2}}), \sqrt{3}) \cup D((0, \sqrt{\frac{3}{2}}), \sqrt{\frac{3}{2}}) \cup D((-\sqrt{\frac{3}{2}}, 0), \sqrt{\frac{3}{2}})$.

(Here the reader may wish to draw a figure for better visualization of all possible directional derivatives of the gradient.

We can conclude that if some domain

\begin{equation}
H := H(K, x) := \{v = ty : y = (y_1, \ldots, y_d), |t| \leq r(y)\}
\end{equation}
is given with \( r(y) \) being a valid estimation for the directional derivative in direction of \( y \), then to find or estimate \( G(K, x) \) an additional process of restricting to the "kernel" part

\[
\tilde{H} := \tilde{H}(K, x) := \bigcap_{\mathcal{S}} \{ v : |\langle v, y \rangle | \leq r(y) \}
\]

is available. That is, we always have \( \tilde{G}(K, x) \subset \tilde{H} \). Note that \( \tilde{H} \) is a convex, symmetric domain for whatever point set \( H \).

In order to illustrate this "kernel technique", let us come back to the above case of estimates from Theorem 3 for the triangle at point \( M \). After some standard considerations with Thales circles we find that \( \tilde{H} \) is the hexagonal domain

\[
\tilde{H}(\Delta, M) = \text{con}\{(\sqrt{6}, 0), (\sqrt{6}, \sqrt{6}), (0, \sqrt{6}), (-\sqrt{6}, 0), (-\sqrt{6}, -\sqrt{6}), (0, -\sqrt{6})\}.
\]

Observe that the area of the possible stretch of \( G \) is considerable reduced from the "fleecy-cloud" domain to the derived hexagonal domain as area \( H(\Delta, M) = 9 + \frac{\pi}{2} = 23.137... \), while area \( \tilde{H}(\Delta, x) = 18. \) For comparison recall that Baran’s Conjecture B would say that the area should be \( \frac{1}{2} \lambda_\Delta(M) = \frac{\pi}{\sqrt{3}} = 16.324... \).

Let us calculate the "kernel set" \( \tilde{H}(\Delta, x) \) in case of the standard triangle \( \Delta \) from the exact estimates (18), (36), (40) which we obtain from the ellipse (and hence also of Baran’s) method! We obtain the following².

**Proposition 12.** With the above notations, \( \tilde{H}(\Delta, x) \) is an ellipse domain. Moreover, if the major axis of this ellipse domain is denoted by \( \mu := \mu(x) \) and the minor axis is denoted by \( \nu := \nu(x) \), then we have

\[
\mu = \sqrt{\frac{2}{x_1(1-x_1)+x_2(1-x_2)+D(x)}} \quad \text{and} \quad \nu = \sqrt{\frac{2}{x_1(1-x_1)+x_2(1-x_2)-D(x)}},
\]

where \( D(x) \) is the quantity already defined in (25).

**Proof.** For any given \( x \in \Delta \) our task is to solve the equation (47) for the case of \( K = \Delta \), the triangle, with the estimating function \( r(y) \) being the quantity (43) of \( \Delta \) defined by (19). That is, for given, fixed \( x \in \Delta \) we want to determine all those vectors \( u = (u_1, u_2) \in \mathbb{R}^2 \), which satisfy \( |\langle u, y \rangle | \leq 1/E(\Delta, x, y) \) for all directional vectors \( y = (\cos \varphi, \sin \varphi) \). Writing in (19) and squaring, the defining equations present themselves as

\[
\{ u : (u_1 \cos \varphi + u_2 \sin \varphi)^2 \leq \frac{\cos^2 \varphi}{x_1} + \frac{\sin^2 \varphi}{x_2} + \frac{(\cos \varphi + \sin \varphi)^2}{1-x_1-x_2} \quad (\forall \varphi \in \mathbb{R}) \}.
\]

If \( \cos \varphi = 0 \), then \( |\sin \varphi| = 1 \) and the equation reduces to

\[
u_2^2 \leq \frac{1}{x_2} + \frac{1}{1-x_2-x_3} = \frac{1-x_1}{x_2(1-x_1-x_2)} \}
\]

which we need to check besides the case when \( \cos^2 \varphi > 0 \). Put \( x_3 := 1-x_1-x_2 \). Division by \( \cos^2 \varphi > 0 \) yields

\[
(u_1 + u_2 t)^2 \leq \frac{1}{x_1} + \frac{t^2}{x_2} + \frac{(1+t)^2}{x_3} \quad (\forall t := \tan \varphi \in \mathbb{R}),
\]

²These computations were executed jointly with Nikola Naidenov from the University of Sofia during the author’s stay in Sofia in October 2004. The author regrets that in spite of his inevitable contribution [24] to the work, Nikola Naidenov chose not to be named as a coauthor.
that is, writing now \( z_j := 1/x_j \) for \( j = 1, 2, 3 \) and ordering,
\[
0 \leq (z_2 + z_3 - u_2^2) + 2(z_4 - u_1 u_2) t + (z_1 + z_3 - u_1^2) \quad (\forall t := \tan \varphi \in \mathbb{R}),
\]
which is a second degree equation in \( t \). Thus the point \( u = (u_1, u_2) \) is a solution iff the
discriminant of this quadratic equation is not positive, and either the quadratic coefficient
is positive, or it is zero and then not only the discriminant is (nonnegative hence) zero,
but also the constant term is nonnegative, too. Note that nonnegativity of the quadratic
coefficient is just (50), while the discriminant condition becomes
\[
0 \leq d(x) := (z_3 - u_1 u_2)^2 - (z_2 + z_3 - u_2^2)(z_1 + z_3 - u_1^2)
\]
which is again a quadratic equation, but now in the coordinates of the point \( u \). From this
form a calculation leads to \( (z_2 + z_3)u_1^2 + (z_1 + z_3)u_2^2 - 2z_3 u_1 u_2 \leq z_1 z_2 + z_1 z_3 + z_2 z_3 \), hence
multiplying by \( x_1 x_2 x_3 \) and using \( x_1 + x_2 + x_3 = 1 \) a few times we arrive at
\[
au_1^2 + bu_2^2 - cu_1 u_2 \leq 1,
\]
where here the coefficients
\[
a := a(x) := x_1(1 - x_1), \quad b := b(x) := x_2(1 - x_2), \quad c := c(x) := 2x_1 x_2
\]
are all strictly positive. Thus (54) determines an ellipse domain indeed. Among points
satisfying (54) it is not difficult to determine the maximal values of \( u_2 \). These will occur
at points where the function \( F(u_1, u_2) \) on the left of (54) have value 1 and a vertical
gradient, i.e. \( \partial_1 F = 0 \), from which conditions a standard calculation derives that \( u_1 = \sqrt{x_2}/\sqrt{(1 - x_1)x_3} \) and \( u_2 = \sqrt{1 - x_1}/\sqrt{x_2 x_3} \), showing that the maximal possible value of
\( u_2 \) satisfies (50) with equality. Moreover, in this extremal case we find that also (52)
is satisfied, the right hand side reducing to constant \( 1/[x_1(1 - x_1)] > 0 \) identically for all \( t \).

It remains to determine the major and minor axes of the ellipse domain of \( u \) described
by (54). In fact, the equation is almost in a canonical form. We need only to rotate the
coordinates by
\[
v_1 := \cos \alpha u_1 - \sin \alpha u_2 \quad u_1 = \cos \alpha v_1 + \sin \alpha v_2
\]
\[
v_2 := \sin \alpha u_1 + \cos \alpha u_2 \quad u_2 = \cos \alpha v_2 - \sin \alpha v_1
\]
to obtain
\[
A v_1^2 + B v_2^2 - C v_1 v_2 \leq 1
\]
with
\[
A := A(x) := a \cos^2 \alpha + b \sin^2 \alpha + c \cos \alpha \sin \alpha,
\]
\[
B := B(x) := a \sin^2 \alpha + b \cos^2 \alpha - c \cos \alpha \sin \alpha,
\]
\[
C := C(x) := a^2 \cos \alpha \sin \alpha - b^2 \cos \alpha \sin \alpha + c(\sin^2 \alpha - \cos^2 \alpha).
\]
In case \( a = b \) the rotational angle \( \alpha = \pi/4 \) will be proper, as then \( C \) vanishes and we get
\( 2A = a + b + c \), \( 2B = a + b - c \), and
\[
A v_1^2 + B v_2^2 \leq 1.
\]
In view of the formula \( D(x) = \sqrt{(a - b)^2 + c^2} \), easily seen from (25) and (55), (48) obtains
for \( a = b \).

Let now \( a \neq b \) and choose \( \alpha = \frac{1}{2} \arctan \frac{c}{a - b} \in (-\pi/4, \pi/4) \), which is chosen again to
annihilate \( C(x) \) and thus to reduce the equation (57) to (59); depending on the sign of \( a - b \),
the major and minor axis of the ellipse are to be \( 1/\sqrt{A} \) and \( 1/\sqrt{B} \) or conversely. The sum
of these two axes is \(a + b\), and a calculation also leads \(A - B = \text{sign} (a - b) \sqrt{(a - b)^2 + c^2} = \pm D(x)\), hence the values (48) obtain once again.

So we are led to the following result.

**Theorem 13.** With the above notations, we have \(\text{area} \tilde{H}(\Delta, x) = \frac{\pi}{\sqrt{x_1 x_2 (1 - x_1 - x_2)}}\).

**Proof.** As is well-known, the area of an ellipse domain \(E\) having major and minor axes \(\mu\) and \(\nu\) is \(\text{area} E = \pi \mu \nu\), hence Proposition 12 leads to the asserted value.

**Corollary 14.** We have \(G(x) \subseteq \text{con} G(x) \subseteq \text{area} \tilde{H}(x) = \frac{1}{2} \lambda(x)\). Hence either \(\text{con} G(x) = \tilde{H}(x)\), for all points \(x \in \Delta\), or Baran’s Conjecture B fails.

**Proof.** One must compute the density function \(\lambda(x)\) of the equilibrium measure. This has already been done by Baran, [4, Example 4.8]: we have \(\lambda(x) = \frac{2\pi}{\sqrt{x_1 x_2 (1 - x_1 - x_2)}}\). On comparing to Theorem 13 we find the asserted identity. Since \(\tilde{H}\) is an ellipse domain and also con \(G\) is a convex domain, con \(G(x) \subset \tilde{H}(x)\) and equality of their area entails that con \(G(x) = \tilde{H}(x)\). On the other hand if at some point \(x \in \Delta\) the respective areas differ, then area con \(G(x) < \text{area} \tilde{H}(x) = \frac{1}{2} \lambda(x)\), hence the conjectured identity of Baran fails.

**Remark 1.** While using the information on the support functional from \(H(\Delta, x)\) improves upon the known area estimates, it does not improve the maximal gradient norm estimate of [19].

Indeed, as \(\tilde{H}(\Delta, x)\) is an ellipse domain, we have to consider the major axis of this ellipse. It turns out that in the case of the standard triangle, this calculation yields \(\max_{v \in \tilde{H}} \|v\| = \max_{v \in H} \|v\| = 1/E(\Delta, x)\).

Note that \(\max_{v \in V} \|v\| = \max_{v \in \text{con} V} \|v\|\) for any set \(V\), hence regarding the maximal gradient norm estimate it makes no difference if we consider con \(G(x)\) or \(G(x)\) only. Also note that starting out from a set \(H \supset G\) and considering the ”kernel” \(\tilde{H}\), we necessarily obtain a convex set, so from \(G \subset \tilde{H}\) it follows that even taking convex hull we still have con \(G \subset \tilde{H}\).

**Corollary 15.** Conjecture A and Conjecture B can not hold simultaneously.

**Proof.** According to Corollary 14, Baran’s Conjecture B holds if only there can be no improvement on the estimates of the ellipse (or Baran’s) method on the simplex. But than Conjecture A fails. Conversely, if Conjecture A holds, then there is an improvement at least at certain points and in certain directions compared to the estimates of the ellipse (or Baran’s) method, hence the estimates of Corollary 14 are strictly exceeding the right value and Baran’s Conjecture B fails.

6. Concluding remarks

Also, another real, geometric method, of obtaining Bernstein type inequalities, due to Skalyga [34, 35], is to be mentioned here: the difficulty with that is that to the best of our knowledge, no one has ever been able to compute, neither for the seemingly least complicated case of the standard triangle of \(\mathbb{R}^2\), nor in any other particular non-symmetric case the yield of that abstract method. Hence in spite of some remarks that the method
is sharp in some sense, it is unclear how close these estimates are to the right answer and what use of them we can obtain in any concrete cases.

Given the above findings, it seemed to be plausible that Conjecture A, if not true, can be disproved by some explicit example. To construct a polynomial with large gradient, as compared to the norm, means to construct a highly oscillating polynomial. For that, various natural and more intricate ideas were tried by Nikola Naidenov [24] in Sofia during the Fall of 2004. We hope he will report on his experiences in the near future.

The author would like to thank for the enlightening comments and corrections of Norm Levenberg.

REFERENCES

[17] Levenberg, N., Remarks on pluripotential theory for convex bodies in $R^N$, manuscript, 11 pages.


A. Rényi Institute of Mathematics
Hungarian Academy of Sciences,
Budapest, P.O.B. 127, 1364
Hungary

E-mail address: revesz@renyi.hu