CONVERGENT BOUNDS FOR THE RANGE OF MULTIVARIATE POLYNOMIALS

J. Garloff

Institut für Angewandte Mathematik

Universität Freiburg i.Br.

Freiburg i.Br.

West Germany

1. Introduction

In this paper we consider the following problem: We are given a *bivariate polynomial* p, i.e., a polynomial in two variables

$$p(x,y) := \sum_{\mu,\nu=0}^{n} a_{\mu\nu} x^{\mu} y^{\nu} \tag{1.1}$$

having real coefficients $a_{\mu\nu}$ and a rectangle

$$Q = X \times Y \text{ with } X = [\underline{x},\overline{x}], \quad Y = [\underline{y},\overline{y}] \in \mathbb{I}(\mathbb{R}). \tag{1.2}$$

Here $\mathbb{I}(\mathbb{R})$ denotes the set of the compact, nonempty real intervals, henceforth referred to simply as intervals. We are seeking for the *range* of p over Q, i.e.,

$$P(Q) = \{p(x,y) \mid (x,y) \in Q\} = [\underline{m},\overline{m}],$$

$$\text{where } \underline{m} = \min_{(x,y)\in Q} p(x,y), \quad \overline{m} = \max_{(x,y)\in Q} p(x,y).$$

Knowledge of this range, or the equivalent global maximum $\overline{m}$ and global minimum $\underline{m}$ is relevant for numerous investigations and applications in numerical and functional analysis, optimization etc. It is therefore important to find easy and efficient methods for getting good approximations to this range. An exposition of available methods is given in the monograph [10].

In our paper we present two methods for finding convergent upper and lower bounds $\overline{m}_k$, $\underline{m}_k$ for the range $p(Q)$, i.e., $\overline{m}_k \geq \overline{m}$ and $\underline{m}_k \leq \underline{m}$ with $\overline{m}_k \to \overline{m}$ and $\underline{m}_k \to \underline{m}$ for $k \to \infty$. Both methods can easily be extended to the higher-dimensional case. For the sake of simplicity we will present our results only in the bivariate case since the generalization to the higher-dimensional case will be obvious.

The first method is based on the mean value theorem and is given in Section 2. The other method is based on the expansion of a bivariate polynomial in Bernstein polynomials and is discussed in detail in Section 3. Here we also address the problem of finding convergent bounds for the range of p over the unit triangle. It turns out that this can be handled in a similar way as for the rectangle. In Section 4 we consider the case that the coefficients of the polynomial p are not exactly known but can be located between upper and lower bounds

$$a_{\mu\nu} \in A_{\mu\nu} = [\underline{a}_{\mu\nu}, \overline{a}_{\mu\nu}] \in \mathbb{I}(\mathbb{R}), \quad \mu,\nu = 0(1)n. \tag{1.3}$$

Required is now to find the range of a set of bivariate polynomials over Q

$$\left\{ \sum_{\mu,\nu=0}^{n} a_{\mu\nu} x^{\mu} y^{\nu} \,\middle|\, (x,y) \in Q,\ a_{\mu\nu} \in A_{\mu\nu},\ \mu,\nu = 0(1)n \right\}.$$

We conclude our paper with a particular application to a problem in multidimensional system theory, namely testing a bivariate polynomial for positivity.
Each real interval can be mapped onto [0,1] by a linear function. So we will confine our discussion mainly to the *unit square* $I := [0,1]\times[0,1]$.
For an integer k we define $K := \{(i,j) \mid i,j = 0(1)k\}$.

2. Bounds using function values

The following is an extension of a method developed by Rivlin [12] for the univariate case. The bounds involve the function values of the polynomial on the grid on the unit square I given by $(\mu/k, \nu/k)$, $(\mu,\nu) \in K$.

Theorem 1 (without proof):
Let p be given by (1.1). Then

$$\max_{(x,y)\in I} p(x,y) = \overline{m} \leqq \max_{(\mu,\nu)\in K} p\left(\frac{\mu}{k},\frac{\nu}{k}\right) + \alpha_k$$
$$\min_{(x,y)\in I} p(x,y) = \underline{m} \geqq \min_{(\mu,\nu)\in K} p\left(\frac{\mu}{k},\frac{\nu}{k}\right) - \alpha_k, \tag{2.1}$$

where

$$\alpha_k := \frac{1}{8k^2} \sum_{i,j=0}^{n} (i+j)(i+j-1)|a_{ij}|.$$

Remark: If p has a large number of vanishing coefficients it might be advantageous to apply the algorithm given in [11] to evaluate p at the grid points since this algorithm takes account of the sparsity pattern of p.

3. Bounds using the Bernstein form

In this section we derive bounds for the range of a bivariate polynomial (1.1) on the unit square using the so called Bernstein form. This form is intimately related to Bernstein polynomials (a good reference for Bernstein polynomials is the monograph [9]). The first application to the range of univariate polynomials was given by Cargo and Shisha [5]; Rivlin [12] improved upon the bounds obtained by Cargo and Shisha. Grassmann and Rokne [6] and Rokne [13-16] applied the results of Rivlin to real and complex interval polynomials. Finally, Lane and Riesenfeld [8] discussed subdivision in the univariate case.

3.1 The Bernstein form of a bivariate polynomial on the unit square

Let $k \geqq n$ be an integer. We define for $(i,j) \in K$.

$$p_{ij}^{(k)}(x,y) := \binom{k}{i}\binom{k}{j}x^i(1-x)^{k-i}y^j(1-y)^{k-j},\ x,y \in I. \tag{3.1}$$

Then by some manipulations we get the identity

$$x^\mu y^\nu = \sum_{s=\mu,t=\nu}^{k} \binom{s}{\mu}\binom{t}{\nu}\rho_{\mu\nu}^{(k)} p_{st}^{(k)}(x,y), \tag{3.2}$$

$$\text{where } \rho_{\mu\nu}^{(k)} := \left[\binom{k}{\mu}\binom{k}{\nu}\right]^{-1},\ (\mu,\nu) \in K. \tag{3.3}$$

Substituting (3.2) into (1.1) gives

$$p(x,y) = \sum_{(i,j)\in K} b_{ij}^{(k)} p_{ij}^{(k)}(x,y), \tag{3.4}$$

$$\text{where } b_{ij}^{(k)} := \sum_{s=0}^{i}\sum_{t=0}^{j} \binom{i}{s}\binom{j}{t}\rho_{st}^{(k)} a_{st}, \tag{3.5}$$

with the convention that $a_{st} = 0$ for $s > n$ or $t > n$.

We call the $b_{ij}^{(k)}$ the *Bernstein coefficients* and (3.4) the *Bernstein form* of p (on the unit square).

Theorem 2:
If p is given by (1.1), then we have

$$\max_{(i,j)\in K} b_{ij}^{(k)} \geq \overline{m},\ \underline{m} \geq \min_{(i,j)\in K} b_{ij}^{(k)} \tag{3.6}$$

for each $k \geq n$; equality holds in the left (resp., right) inequality if and only if $\max_{(i,j)\in K} b_{ij}^{(k)}$ (resp., $\min_{(i,j)\in K} b_{ij}^{(k)}$) is one of $b_{00}^{(k)}$, $b_{k0}^{(k)}$, $b_{0k}^{(k)}$, $b_{kk}^{(k)}$.

Proof: Since $0 \leq p_{ij}^{(k)}(x,y)$ for all $(x,y) \in I$ and $(i,j) \in K$, and

$$\sum_{(i,j)\in K} p_{ij}^{(k)}(x,y) = 1 \quad \text{for all } (x,y) \in I \tag{3.7}$$

the inequalities (3.6) follow.
The "if" part is obvious from

$$b_{00}^{(k)} = a_{00} = p(0,0)\ , \quad b_{0k}^{(k)} = \sum_{t=0}^{k} a_{0t} = p(0,1),$$

$$b_{k0}^{(k)} = \sum_{s=0}^{k} a_{s0} = p(1,0), \quad b_{kk}^{(k)} = \sum_{(s,t)\in K} a_{st} = p(1,1).$$

We now assume that $\max b_{ij}^{(k)} = \bar{m} = p(\hat{x},\hat{y})$, $(\hat{x},\hat{y}) \in I$, and $\max b_{ij}^{(k)} > b_{00}^{(k)}, b_{0k}^{(k)}, b_{k0}^{(k)}, b_{kk}^{(k)}$. If $\hat{x},\hat{y} \in (0,1)$, then by (3.7)

$$p(\hat{x},\hat{y}) < \max b_{ij}^{(k)} \sum_{(i,j)\in K} p_{ij}(\hat{x},\hat{y}) = \max b_{ij}^{(k)},$$

a contradiction. The proof of the other cases and for $\min b_{ij}^{(k)}$ is analogous.

∎

We now show that the bounds given in (3.6) converge to $\underline{m}$ and $\bar{m}$, respectively.

<u>Theorem 3</u>:
If $k \geqq 2$, then

$$\max_{(i,j)\in K} b_{ij}^{(k)} - \bar{m}\ ,\ \underline{m} - \min_{(i,j)\in K} b_{ij}^{(k)} \leqq \gamma(k-1)k^{-2},$$

where

$$\gamma := \sum_{\mu,\nu=0}^{n} ((\mu-1)_+^2 + (\nu-1)_+^2)\,|a_{\mu\nu}| \tag{3.8}$$

and $(x)_+ = \max(0,x)$.

<u>Proof</u>: Since some of our considerations follow Rivlin's proof for the univariate case [12] we only give an outline of the proof.

For a function f defined on I let

$$B_k(f;x,y) := \sum_{(i,j)\in K} f(\tfrac{i}{k},\tfrac{j}{k})p_{ij}^{(k)}(x,y).$$

For $s,t \leqq n$, denote by $\delta_{ij}(s,t)$, $(i,j) \in K$, the Bernstein coefficients of the polynomial $B_k(x^s x^t;x,y) - x^s y^t$. Since $B_k(x^s y^t;x,y) = x^s y^t$ for $s,t \leqq 1$ we have $\delta_{ij}(s,t) = 0$ for $(i,j) \in K$ and $s,t \leqq 1$. Therefore, we assume that $s \geqq 2$ or $t \geqq 2$.
If $0 \leqq i < s$ we have by (3.2)

$$\delta_{ij}(s,t) \leqq (\tfrac{i}{k})^s(\tfrac{j}{k})^t \leqq (\tfrac{i}{k})^s \leqq (\tfrac{s-1}{k})^s \leqq \frac{(s-1)^2}{k}\,(1-\tfrac{1}{k}).$$

If $2 \leqq s \leqq i$ and $2 \leqq t \leqq j$ we get after some algebraic manipulations

$$\delta_{ij}(s,t) = (\tfrac{i}{k})^s(\tfrac{j}{k})^t - \rho_{st}^{(k)}\binom{i}{s}\binom{j}{t}$$

$$\leqq (\tfrac{i}{k})^s(\tfrac{j}{k})^t\left[1-\left(1-\frac{(s-1)}{i}\right)^{s-1}\left(1-\frac{(t-1)}{j}\right)^{t-1}\right].$$

Now we apply the generalized Bernoulli inequality, see e.g. [7, p. 60], to obtain

$$\delta_{ij}(s,t) \leqq (\tfrac{i}{k})^s(\tfrac{j}{k})^t\left(\frac{(s-1)^2}{i}+\frac{(t-1)^2}{j}\right)$$

$$= (\tfrac{i}{k})^{s-1}(\tfrac{j}{k})^t\,\frac{(s-1)^2}{k} + (\tfrac{i}{k})^s(\tfrac{j}{k})^{t-1}\,\frac{(t-1)^2}{k}.$$

It follows that (note that $\delta_{kk}(s,t) = 0$)

$$\delta_{ij}(s,t) \leqq \frac{k-1}{k^2}\left((s-1)^2+(t-1)^2\right).$$

It is easy to see that this formula is also true in the remaining cases.
As in the univariate case now one shows that

$$\left|p(\tfrac{i}{k},\tfrac{j}{k}) - b_{ij}^{(k)}\right| \leqq \gamma\,\frac{k-1}{k^2}$$

from which the assertion follows. ∎

Because of Theorem 3 one expects that when increasing k the bounds become better. Before we discuss this in more detail we note another improvement, namely a correction of already calculated bounds.
We assume that

$$\max_{(i,j)\in K} b_{ij}^{(k)} = b_{\mu\nu},\quad (\mu,\nu) \notin \{(0,0),\ (0,k),\ (k,0),\ (k,k)\},$$

$$\max\{b_{ij}^{(k)} \mid (i,j) \neq (\mu,\nu)\} = b_{\hat{\mu},\hat{\nu}} < b_{\mu\nu}.$$

By a similar argument as in [12] one shows using

$$\max_{(x,y)\in I} p_{ij}(x,y) = p_{ij}(\tfrac{i}{k},\tfrac{j}{k})$$

that

$$\bar{m} \leqq p_{\mu\nu}(\tfrac{\mu}{k},\tfrac{\nu}{k})b_{\mu\nu} + (1 - p_{\mu\nu}(\tfrac{\mu}{k},\tfrac{\nu}{k}))b_{\hat{\mu}\hat{\nu}} < b_{\mu\nu} \ . \tag{3.9}$$

If the maximum of the Bernstein coefficients is assumed more than once then a similar bound holds involving the maxima of the corresponding Bernstein polynomials.
An analogous bound is valid for $\underline{m}$.

3.2 Calculation of the Bernstein coefficients

The calculation of the Bernstein coefficients by (3.5) is not economic since, e.g., the number of additions needed is $\frac{1}{4}n^4 + O(n^3)$ $(k = n)$. We present now a method for calculating the Bernstein coefficients which requires fewer arithmetical operations.

Proposition 1:
For $\mu,\nu = O(1)n$ we have

$$a_{\mu\nu} = (\rho_{\mu\nu}^{(k)})^{-1}\Delta_{\mu\nu}b_{OO}^{(k)} \ ,$$

where $\Delta_{\mu\nu}$ is a twodimensional forward difference operator defined by

$$\Delta_{\mu\nu}b_{ij}^{(k)} := \sum_{\sigma=0}^{\mu}\sum_{\tau=0}^{\nu}(-1)^{\mu+\nu-\sigma-\tau}\binom{\mu}{\sigma}\binom{\nu}{\tau}b_{i+\sigma,j+\tau}^{(k)} \ ,$$

$$\mu \leqq k-i, \ \nu \leqq k-j \ .$$

Proof: Straightforward calculation using

$$a_{st} = \frac{1}{s!t!}\frac{\partial^{s+t}p}{\partial x^s \partial y^t}(O,O) . \quad \blacksquare$$

To calculate the Bernstein coefficients one may proceed in two steps:

First, one computes

$$\Delta_{\mu\nu}b_{OO}^{(k)} = \rho_{\mu\nu}^{(k)}a_{\mu\nu} .$$

Then one computes the Bernstein coefficients from $\Delta_{\mu\nu}b_{OO}^{(k)}$ by using the following recurrence relations:

$$\left.\begin{aligned}
\Delta_{00} b_{ij}^{(k)} &= b_{ij}^{(k)} \;, \\
\Delta_{10} b_{ij}^{(k)} &= b_{i+1,j}^{(k)} - b_{ij}^{(k)} \;, \\
\Delta_{01} b_{ij}^{(k)} &= b_{i,j+1}^{(k)} - b_{ij}^{(k)} \;, \\
\Delta_{\mu+1,\nu} b_{ij}^{(k)} &= \Delta_{\mu\nu} b_{i+1,j}^{(k)} - \Delta_{\mu\nu} b_{ij}^{(k)} \;, \\
\Delta_{\mu,\nu+1} b_{ij}^{(k)} &= \Delta_{\mu\nu} b_{i,j+1}^{(k)} - \Delta_{\mu\nu} b_{ij}^{(k)} \;;
\end{aligned}\right\} \tag{3.10}$$

furthermore, we have

$$\Delta_{\mu\nu} b_{00}^{(k)} = 0 \quad \text{if} \quad \mu > n \quad \text{or} \quad \nu > n \;.$$

To clarify the second step we give the explicit calculations in the case $k = n = 2$ (for simplicity we suppress the upper index k).
We start with the table of the differences $\Delta_{\mu\nu} b_{00}$

$$\begin{array}{ccc}
b_{00} & \Delta_{01} b_{00} & \Delta_{02} b_{00} \\
\Delta_{10} b_{00} & \Delta_{11} b_{00} & \Delta_{12} b_{00} \\
\Delta_{20} b_{00} & \Delta_{21} b_{00} & \Delta_{22} b_{00}
\end{array} \;;$$

in the upper left corner we have $b_{00} = \Delta_{00} b_{00}$.

Now add the first column to the second and the second to the third to obtain

$$\begin{array}{ccc}
b_{00} & b_{01} & \Delta_{01} b_{01} \\
\Delta_{10} b_{00} & \Delta_{10} b_{01} & \Delta_{11} b_{01} \\
\Delta_{20} b_{00} & \Delta_{20} b_{01} & \Delta_{21} b_{01}
\end{array} \;;$$

as the second Bernstein coefficient we now know $b_{01} = \Delta_{00} b_{01}$. Now add row 1 to row 2 and row 2 to row 3 which gives

$$\begin{array}{ccc}
b_{00} & b_{01} & \Delta_{01} b_{01} \\
b_{10} & b_{11} & \Delta_{01} b_{11} \\
\Delta_{10} b_{10} & \Delta_{10} b_{11} & \Delta_{11} b_{11}
\end{array} \;.$$

In the last but one step add the second column to the third to obtain

$$\begin{array}{ccc} b_{00} & b_{01} & b_{02} \\ b_{10} & b_{11} & b_{12} \\ \hdashline \Delta_{10}b_{10} & \Delta_{10}b_{11} & \Delta_{10}b_{12} \end{array}$$

and in the last step the second row to the third which yields the table of the Bernstein coefficients b_{ij}, $i,j = 0(1)2$.

The number of additions required for this method is $n(n+1)^2$ $(k = n)$. Also the number of calculations of binomial coefficients is considerably smaller than for the direct calculation by (3.5).

When the Bernstein coefficients are computed by the difference table for several k then for each k all the Bernstein coefficients have to be calculated once again. Hence the difference table is unfavourable when it is used more than once. A better way is to calculate the Bernstein coefficients for fixed k-1 and then to make use of the following recurrence relations ($(i,j) \in K$):

$$\begin{aligned} b_{ij}^{(k)} &= k^{-2}[ij\, b_{i-1,j-1}^{(k-1)} + j(k-i)b_{i,j-1}^{(k-1)} \\ &\quad + i(k-j)b_{i-1,j}^{(k-1)} + (k-i)(k-j)b_{ij}^{(k-1)}] \\ &\quad \text{with } b_{-1,-1}^{(k-1)} = b_{-1,j}^{(k-1)} = b_{i,-1}^{(k-1)} = b_{ik}^{(k-1)} = b_{kj}^{(k-1)} = 0, \\ &\quad i,j = 0(1)k. \end{aligned} \tag{3.11}$$

We see that the Bernstein coefficients of order k are convex linear combinations of Bernstein coefficients of order k-1 and we may conclude that the convergence of the bounds is monotone:

$$\max b_{ij}^{(k-1)} \geqq \max b_{ij}^{(k)} ,$$

$$\min b_{ij}^{(k-1)} \leqq \min b_{ij}^{(k)} .$$

3.3 Subdivision

In this section we discuss subdivision, i.e., dividing the unit square I into four subsquares of edge length 1/2 (see Fig. 1)

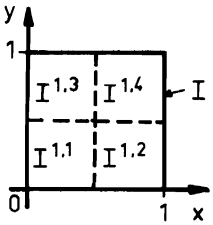

Fig. 1. Subdivision

and calculating the Bernstein coefficients of the given polynomial (1.1) on each subsquare. Here we mean by the Bernstein coefficients $b_{ij}(Q)$ of p on a rectangle Q given by (1.2) the Bernstein coefficients of the shifted polynomial

$$p(x,y) = \sum_{\mu,\nu=0}^{n} c_{\mu\nu}\xi^{\mu}\eta^{\nu} , \quad (\xi,\eta) \in I \tag{3.12}$$

with

$$c_{\mu\nu} = (\bar{x}-\underline{x})^{\mu}(\bar{y}-\underline{y})^{\nu} \sum_{s=\mu}^{n} \sum_{t=\nu}^{n} \binom{s}{\mu}\binom{t}{\nu}\underline{x}^{s-\mu}\underline{y}^{t-\nu}a_{st} . \tag{3.13}$$

The coefficients $c_{\mu\nu}$ of the shifted polynomial may be calculated by a twodimensional complete Horner scheme. The process may be continued by subdividing again each of the four subsquares into four subsquares of edge length 1/4 and calculating the Bernstein coefficients on each of the 16 subsquares and so on. Then the maximum (resp., minimum) taken over the Bernstein coefficients of p on all subsquares is an upper (resp., lower) bound for p over I.

We first give the explicit formulas for the Bernstein coefficients on the four subsquares of an arbitrary square. By iterated use of these formulas one sees that the Bernstein coefficients on the subsquares generated by subdivision may be calculated successively from the Bernstein coefficients on I. In particular, transformation of the subsquares onto I can be avoided. Then we show that the bounds converge quadratically when subdivision is applied iteratively.

In the sequel we assume that the Bernstein coefficients of p are computed for fixed k, and we suppress the upper index k.

Proposition 2:
Let the rectangle Q be given by (1.2) and let $\xi := (\underline{x}+\overline{x})/2$, $\eta := (\underline{y}+\overline{y})/2$. Then the Bernstein coefficients on the four subrectangles are given by $((i,j) \in K)$:

$$b_{ij}(Q^{\ell}) = 2^{-i-j} \sum_{s=0}^{i} \sum_{t=0}^{j} \binom{i}{s}\binom{j}{t} \beta_{st}^{(\ell)} , \qquad \ell = 1,2,3,4,$$

where

$$\beta_{st}^{(1)} = b_{st}(Q) \qquad \text{on} \quad Q^1 := [\underline{x},\xi]\times[\underline{y},\eta] ,$$

$$\beta_{st}^{(2)} = b_{k-s,t}(Q) \qquad \text{on} \quad Q^2 := [\xi,\overline{x}]\times[\underline{y},\eta] ,$$

$$\beta_{st}^{(3)} = b_{s,k-t}(Q) \qquad \text{on} \quad Q^3 := [\underline{x},\xi]\times[\eta,\overline{y}] ,$$

$$\beta_{st}^{(4)} = b_{k-s,k-t}(Q) \qquad \text{on} \quad Q^4 := [\xi,\overline{x}]\times[\eta,\overline{y}] .$$

Proof: Similar as for subdivision in the univariate case [8]. ∎

For practical calculation of the Bernstein coefficients $b_{ij}(Q^{\ell})$, $\ell = 1,2,3,4$, one writes down the following four tables

$$(b_{ij}), \quad (b_{k-i,j}), \quad (b_{i,k-j}), \quad (b_{k-i,k-j})$$

and then proceeds for each table similarly as in the calculation of the Bernstein coefficients, cf. Section 3.2. The entries of the final tables have to be divided by 2^{-i-j}. We see that the $b_{ij}(Q^{\ell})$ are convex linear combinations of the $b_{ij}(Q)$. Therefore, we conclude that the bounds calculated from the Bernstein coefficients on the smaller rectangles are at least as good as those obtained by using the Bernstein coefficients of p on Q and that the bounds are monotone when subdivision is applied iteratively.

We denote the subsquares of edge length 2^{-m} generated by subdivision of I (arranged in any specified order) by $I^{m,\ell}$, $\ell = 1(1)4^m$.

Theorem 4:
If $k \geqq 2$ the following relation hold for all m = 3,4,5,...

$$\left.\begin{array}{l} \max\limits_{\substack{(i,j)\in K \\ \ell=1(1)4^m}} b_{ij}(I^{m,\ell}) - \bar{m} \\ \\ \bar{m} - \min\limits_{\substack{(i,j)\in K \\ \ell=1(1)4^m}} b_{ij}(I^{m,\ell}) \end{array}\right\} \leqq \varepsilon(k-1)k^{-2}m^{-2},$$

where

$$\varepsilon := \max_{\mu,\nu=0(1)n} \sum_{s=\mu}^{n} \sum_{t=\nu}^{n} \binom{s}{\mu}\binom{t}{\nu} |a_{st}| .$$

Proof: Let m be fixed and

$$\max_{\substack{(i,j)\in K \\ \ell=1(1)4^m}} b_{ij}(I^{m,\ell}) = \max_{(i,j)\in K} b_{ij}(I^{m,\ell_0}) .$$

Then by Theorem 3

$$\max_{(i,j)\in K} b_{ij}(I^{m,\ell_0}) - \max_{(x,y)\in I^{m,\ell_0}} p(x,y) \leqq \gamma(k-1)k^{-2} ,$$

where

$$\gamma = \sum_{\mu,\nu=1}^{n-1} (\mu^2 + \nu^2)|c_{\mu+1,\nu+1}|$$
$$+ \sum_{\mu=2}^{n} (\mu-1)^2 (|c_{\mu 0}|+|c_{\mu 1}|+|c_{0\mu}|+|c_{1\mu}|) ,$$

and the $c_{\mu\nu}$'s are the coefficients of the shifted polynomial p (3.12). Since by (3.13)

$$|c_{\mu\nu}| \leqq \varepsilon\cdot 2^{-m(\mu+\nu)} , \quad (\mu,\nu) \in K$$

we get (we assume w.l.o.g. $p \neq 0$)

$$\gamma\varepsilon^{-1} \leqq 2^{-2m} \sum_{\mu,\nu=1}^{n-1} (\mu^2+\nu^2)\,2^{-m(\mu+\nu)} + 2^{-m+1}(1+2^{-m}) \sum_{\mu=1}^{n-1} \mu^2 2^{-m\mu} .$$

Because all summands on the right hand side are positive we may estimate the sums by the respective infinite series. Making use of the relation

$$\sum_{\mu=1}^{\infty} \mu^2 x^\mu = x(1+x)(1-x)^{-3} \quad \text{for} \quad |x| < 1$$

we obtain

$$\begin{aligned} \gamma\varepsilon^{-1} &\leqq 2^{-4m+1}(1+2^{-m})(1-2^{-m})^{-4} + 2^{-2m+1}(1+2^{-m})^2(1-2^{-m})^{-3} \\ &= 2^{-2m+1}(1+2^{-m})(1-2^{-m})^{-4} \\ &= 2^{m+1}(2^m+1)(2^m-1)^{-4}. \end{aligned}$$

The last term is less than m^{-2} for $m \geqq 3$.
The proof for the minimum is entirely analogous. ∎

Remarks: i) One shows similarly that subdivision in the univariate case converges quadratically. This extends the results in [8].

ii) For the proof of quadratic convergence in the r-variate case with $r \geqq 3$ it is easier to use the rougher estimate

$$\gamma \leqq \varepsilon \sum_{\mu_1,\ldots,\mu_r=0}^{\infty} (\mu_1^2 + \ldots + \mu_r^2)\,2^{-m(\mu_1+\ldots+\mu_r)} .$$

For the question of which subsquares can be discarded from the list of further examination see Section 5.

3.4 Bounds for the range of a multivariate polynomial on the unit triangle

In this section we consider the *unit triangle*

$$S := \{(x,y) \in \mathbb{R}^2 \mid x,y \geqq 0 \wedge x+y \leqq 1\}$$

instead of the unit square and address the question of finding bounds for the range $[\underline{m},\overline{m}]$ of the bivariate polynomial p given by (1.1) on S, i.e.,

$$\underline{m} = \min_{(x,y)\in S} p(x,y), \quad \overline{m} = \max_{(x,y)\in S} p(x,y).$$

We set

$$d := \max\{\mu+\nu \mid a_{\mu\nu} \neq 0\}, \tag{3.14}$$

e.g., $d = 2n$ if $a_{nn} \neq 0$, $d = n$ if $a_{\mu\nu} = 0$ for $\nu > n-\mu$. Let $k \geqq d$. The index set K is now defined by

$$K := \{(i,j) \mid i,j = 0(1)k,\ i+j \leqq k\}.$$

Appropriate bivariate Bernstein polynomials are now given by $((i,j)\in K)$:

$$p_{ij}^{(k)}(x,y) = \binom{k}{i\ j} x^i y^j (1-x-y)^{k-i-j}$$

with the generalized binomial coefficients

$$\binom{k}{i\ j} := \frac{k!}{i!j!(k-i-j)!} = \binom{k}{i}\binom{k-i}{j}.$$

Then on the unit triangle S we have again the Bernstein form (3.4) with the Bernstein coefficients (3.5) but now with the convention that

$$\rho_{\mu\nu}^{(k)} := \binom{k}{\mu\ \nu}^{-1} \quad \text{for all} \quad (\mu,\nu) \in K \tag{3.3'}$$

instead of (3.3).
For these Bernstein polynomials Theorem 2 remains true with the modification that in the case of equality $\max b_{ij}^{(k)}$, $\min b_{ij}^{(k)}$ $\in \{b_{00}^{(k)}, b_{0k}^{(k)}, b_{k0}^{(k)}\}$. Theorem 3 reads now (due to the fact that $\delta_{\mu\nu}(1,1) = -\mu\nu\, k^{-2}(k-1)^{-1}$, cf. the proof of Theorem 3):

<u>Theorem 3'</u>:
The following bound holds for all $k \geqq 2$

$$\max_{(i,j)\in K} b_{ij}^{(k)} - \overline{m}, \quad \underline{m} - \min_{(i,j)\in K} b_{ij}^{(k)} \leqq (\gamma + |a_{11}|)(k-1)k^{-2},$$

where γ is given by (3.8).

Also, already calculated bounds may be improved similarly as for the case of the unit square, cf. (3.9). The calculation of the Bernstein coefficients on I carries over to the Bernstein coefficients on S but now only the upper left half of the difference table is needed. The number of additions required is $(k=n)\ d(d+1)(d+2)/3$, where d is given by (3.14). Formula (3.11) reads now ($(i,j) \in K$):

$$b_{ij}^{(k)} = k^{-1}[i\, b_{i-1,j}^{(k-1)} + j\, b_{i,j-1}^{(k-1)} + (k-i-j) b_{ij}^{(k-1)}] \tag{3.11'}$$

$$\text{with } b_{-1,j}^{(k-1)} = b_{i,-1}^{(k-1)} = 0,\ b_{ij}^{(k-1)} = 0 \text{ if } i+j = k.$$

Results on subdivision will be given elsewhere.

3.5 Symmetric coefficients

It is advantageous if the coefficients of p given by (1.1) are *symmetric*, i.e., $a_{\mu\nu} = a_{\nu\mu}$, $\mu,\nu = 0(1)n$, because then the Bernstein coefficients (on the unit square and the unit triangle) are also symmetric and therefore the number of operations required for the calculation of the Bernstein coefficients can about be halved. The symmetry of the Bernstein coefficients on I carries over to the Bernstein coefficients on the subsquares generated by subdivision but in different forms. So we have for all $(i,j) \in K$ (cf. Figure 1)

$$b_{ij}^{(k)}(I^{1,\ell}) = b_{ji}^{(k)}(I^{1,\ell}), \quad \ell = 1,4$$

$$b_{ij}^{(k)}(I^{1,2}) = b_{ji}^{(k)}(I^{1,3})$$

and again the number of operations required can about be halved.

4. Bounds for the range of a multivariate interval polynomial

To guarantee the bounds obtained in Sections 2 and 3 in the presence of rounding errors (entailed by computing with fixed length floating point arithmetic) interval arithmetic, see, e.g., [1], should be applied. In this section we address another question involving interval arithmetic.

Assume that the coefficients $a_{\mu\nu}$ of the polynomial p given by (1.1) are not exactly known but can be located between upper and lower bounds (1.3). Then we have to consider the *interval polynomial*

$$P(x,y) := \sum_{\mu,\nu=0}^{n} A_{\mu\nu} x^{\mu} y^{\nu}$$

and it is required to find the range of P over the unit square I, $P(I) = \{P(x,y) \mid (x,y) \in I\}$. Clearly, $p(I) \subseteq P(I)$ holds. Bounds for $P(I)$ can be obtained from (2.1), (3.6) simply by replacing the real coefficients $a_{\mu\nu}$ by the respective interval coefficients $A_{\mu\nu}$ and the real arithmetical operations by the respective interval arithmetical operations. E.g., an enclosure of the Bernstein coefficients $b_{ij}^{(k)}$ is given by

$$B_{ij}^{(k)} = [\underline{b}_{ij}^{(k)}, \overline{b}_{ij}^{(k)}] := \sum_{s=0}^{i} \sum_{t=0}^{j} \binom{i}{s}\binom{j}{t} \rho_{st}^{(k)} A_{st}, \quad (i,j) \in K; \quad (4.1)$$

then $\max \overline{b}_{ij}^{(k)}$ (resp., $\min \underline{b}_{ij}^{(k)}$) is an upper (resp., lower) bound for $\max P(I)$ (resp., $\min P(I)$). We do not go into the details here. However, special attention has to be paid to the overestimation entailed by replacing real numbers by intervals. This concerns two problems which we will discuss in the sequel.

1) In (4.1) each coefficient $A_{\mu\nu}$ occurs only once in the calculation of each $B_{ij}^{(k)}$. It follows that the direct calculation by (4.1) is *optimal*, i.e., there is a real polynomial $\tilde{p}(x,y) = \sum_{\mu,\nu=0}^{n} \tilde{a}_{\mu\nu} x^{\mu} y^{\nu}$ with $\tilde{a}_{\mu\nu} \in A_{\mu\nu}$ and Bernstein coefficients $\tilde{b}_{ij}^{(k)}$ such that

$$\max \tilde{b}_{ij}^{(k)} = \max \overline{b}_{ij}^{(k)} \geq \max_{(x,y)\in I} P(x,y)$$

and analogously for $\min \underline{b}_{ij}^{(k)}$.

An enclosure of the real Bernstein coefficients $b_{ij}^{(k)}$ may also be obtained by interval performance of the procedure using the difference table, cf. Section 3.2, starting with $\rho_{\mu\nu}^{(k)} A_{\mu\nu}$. In the calculation of each entry of the resulting table some $A_{\mu\nu}$'s occur more than once. But this causes no overestimation because the real numbers by which the intervals under consideration are multiplied are positive and therefore the distributive law holds, see, e.g., [1, p. 3]. It follows that also the difference table produces the coefficients $B_{ij}^{(k)}$.

2) If the range of P over an arbitrary rectangle Q given by (1.2) is wanted the interval polynomial has to be shifted to I. Then when replacing in (3.13) all real coefficients $a_{\mu\nu}$ by the intervals $A_{\mu\nu}$ the width of the $B_{ij}^{(k)}(Q)$ could increase since each of the intervals $A_{\mu\nu}$ occurs more than once in the calculation of each $B_{ij}^{(k)}(Q)$. But if one plugs (3.13) into (3.5) then the resulting double sum can be rearranged as follows (for the univariate case see [16]):

$$b_{ij}^{(k)} = \sum_{s,t=0}^{n} a_{st} \sum_{\mu=0}^{\min\{s,n,i\}} \binom{s}{\mu} \underline{x}^{s-\mu} (\overline{x}-\underline{x})^{\mu} \frac{\binom{i}{\mu}}{\binom{k}{\mu}} \times \tag{4.2}$$

$$\times \sum_{\nu=0}^{\min\{t,n,j\}} \binom{t}{\nu} \underline{y}^{t-\nu} (\overline{y}-\underline{y})^{\nu} \frac{\binom{j}{\nu}}{\binom{k}{\nu}} .$$

Now each real number $a_{\mu\nu}$ occurs only once in the calculation of each $b_{ij}^{(k)}$ and replacing a_{st} by A_{st} in (4.2) gives an optimal formula, i.e., the endpoints of the resulting intervals are Bernstein coefficients of real polynomials with coefficients taken from the interval coefficients. However, this formula requires more calculations compared to (4.1).

Another way to avoid the overestimation entailed by the shift of the original polynomial is to divide the given rectangle into (at most four) rectangles lying in the four quadrants of $\mathbb{R}^2$. On each of these rectangles the two *corner polynomials* of P, i.e.,

$$\sum_{\mu,\nu=0}^{n} a_{\mu\nu}^{(1)} x^{\mu} y^{\nu} \leqq \sum_{\mu,\nu=0}^{n} \tilde{a}_{\mu\nu} x^{\mu} y^{\nu} \leqq \sum_{\mu,\nu=0}^{n} a_{\mu\nu}^{(2)} x^{\mu} y^{\nu}$$

$$\text{for all } \tilde{a}_{\mu\nu} \in A_{\mu\nu} \quad \text{with} \quad a_{\mu\nu}^{(1)}, a_{\mu\nu}^{(2)} \in A_{\mu\nu}$$

can be given explicitly, e.g. on $[0,\infty)\times(-\infty,0]$ we have

$$a_{\mu\nu}^{(1)} = \underline{a}_{\mu\nu} \quad \text{if } \nu \text{ is even}, \; = \overline{a}_{\mu\nu} \quad \text{if } \nu \text{ is odd}$$

$$a_{\mu\nu}^{(2)} = \overline{a}_{\mu\nu} \quad \text{if } \nu \text{ is even}, \; = \underline{a}_{\mu\nu} \quad \text{if } \nu \text{ is odd.}$$

Then the problem of finding bounds for the range of the interval polynomial P reduces to the problem of finding bounds for the range of at most 8 real polynomials for which the shift to I can be done exactly (except for small intervals in the coefficients of the shifted polynomial due to the use of machine interval arithmetic).

5. An application

This paper was stimulated by a problem often arising in multidimensional system theory, namely to test a multivariate polynomial p for positivity on a multidimensional rectangle, cf. [2 ; 3, Chapter 2, 17]. Such tests are often referred to as *local positivity tests.*
Computing the minimum of the Bernstein coefficients provides an alternative local positivity test. If $\min_{(i,j)\in K} b_{ij}^{(k)} > 0$ for an integer k then the positivity of p is guaranteed. On the other hand, if

$$\min_{(i,j)\in K} b_{ij}^{(k)} + \gamma(k-1)k^{-2} \leqq 0$$

then by Theorem 3, p assumes also nonpositive values.
If subdivision is applied iteratively one may reduce the computational effort by the following two observations. Let $I^* = I^{m,\ell}$ be a subsquare generated by subdivision.

i) If the polynomial p assumes its minimum over I* at one of the four vertices of I*, then by Theorem 2

$$\min_{(x,y)\in I^*} p(x,y) = \min_{(i,j)\in K} b_{ij}(I^*) \in \{b_{00}(I^*), b_{k0}(I^*), b_{0k}(I^*), b_{kk}(I^*)\}$$

and there is no need to subdivide I* further. If furthermore

$$\min_{(i,j)\in K} b_{ij}(I^*) = \min_{\substack{(i,j)\in K \\ \ell=1(1)4^m}} b_{ij}(I^{m,\ell})$$

then one already knows that

$$\min_{(x,y)\in I^*} p(x,y) = \min_{(x,y)\in I} p(x,y) .$$

ii) If

$$\min_{(i,j)\in K} b_{ij}(I^*) > \min_{\substack{(i,j)\in K \\ \ell=1(1)4^m}} b_{ij}(I^{m,\ell}) + \varepsilon(k-1)k^{-2}m^{-2} \qquad (5.1)$$

then I* may be discarded from the list for further examination since p can not assume its minimum over I on I*. If equality holds in (5.1) I* may also be discarded since the range of p over I* makes no additional contribution to the minimum of p over I.

Multivariate polynomials can be very sensitive to small perturbations of their coefficients. It is therefore useful to know the allowable intervals centered around the respective unperturbed values within the coefficients of the polynomial might fluctuate without losing the property of being positive for all $x \in \mathbb{R}^r$. This problem was solved in [4 ; see also 3, Section 2.5].

References

[1] Alefeld, G. and J. Herzberger, Introduction to Interval Computations, Academic Press, New York (1983)

[2] Bickard, T.A. and E.I. Jury, Real polynomials: a test for non-global non-negativity and non-global positivity, J. Math. Anal. Appl. 78, 17-32 (1980).

[3] Bose, N.K., Applied Multidimensional Systems Theory, van Nostrand Reinhold Comp., New York (1982).

[4] Bose, N.K. and J.P. Guiver, Multivariate polynomial positivity invariance under coefficient perturbation, IEEE Trans. Acoust. Speech Signal Process ASSP - 28, 660-665 (1980).

[5] Cargo, G.T. and O Shisha, The Bernstein form of a polynomial, J. Res. Nat. Bur. Standards 70B, 79-81 (1966).

[6] Grassmann, E. and J. Rokne, The range of values of a circular complex polynomial over a circular complex interval, Computing 23, 139-169 (1979).

[7] Hardy, G.H., J.E. Littlewood, and G. Pólya, Inequalities, 2nd ed., Cambridge Univ. Press, London and New York (1959).

[8] Lane, J.M. and R.F. Riesenfeld, Bounds on a polynomial, BIT 21, 112-117 (1981).

[9] Lorentz, G.G., Bernstein Polynomials, Univ. Toronto Press, Toronto (1953)

[10] Ratschek, H. and J. Rokne, Computer Methods for the Range of Functions, Ellis Horwood Ltd., Chichester (1984).

[11] Rheinboldt, W.C., C.K. Mesztenyi, and J.M. Fitzgerald, On the evaluation of multivariate polynomials and their derivatives, BIT 17, 437-450 (1977).

[12] Rivlin, T.J., Bounds on a polynomial, J. Res. Nat. Bur. Standards 74B, 47-54 (1970).

[13] Rokne, J., Bounds for an interval polynomial, Computing 18, 225-240 (1977).

[14] Rokne, J., A note on the Bernstein algorithm for bounds for interval polynomials, Computing 21, 159-170 (1979).

[15] Rokne, J., The range of values of a complex polynomial over a complex interval, Computing 22, 153-169 (1979).

[16] Rokne, J., Optimal computation of the Bernstein algorithm for the bound of an interval polynomial, Computing 28, 239-246 (1982).

[17] Walach, E. and E. Zeheb, Sign test of multivariable real polynomials, IEEE Trans. Circuits and Systems CAS-27, 619-625 (1980).