Estimates for Positive Roots of Polynomials

Doru Ștefănescu\textsuperscript{1}  Vladimir Gerdt\textsuperscript{2}  Simeon Evlakhov\textsuperscript{2}

\textsuperscript{1}University of Bucharest, Romania
\textsuperscript{2}JINR Dubna, Russia

29 October 2009, CADE 2009, PAMPLONA
Contents

- Introduction
- Computation of Polynomial Roots
- Bounds for Complex Roots
- Bounds for Real Roots
- Conclusions
Introduction

- The computation of accurate bounds for univariate complex polynomials has many applications in various problems involving polynomials.
- We review some methods for computing polynomial bounds, emphasizing on the case of real zeros.
- For real roots we propose a general method that extends the theorems of Kioustelidis, Ştefănescu a. o. These bounds are expressed in function of the degree, the size of the coefficients and families of parameters that can be properly chosen.
- We use these bounds for the evaluation of the sizes of polynomial divisors.
One of the fundamental problems in Numerical Mathematics is the computation of polynomial zeros. The determination of the roots arises frequently in applications. Since the exact computation of the zeros in function of the coefficients of the polynomial is not possible for general polynomials, for all practical purposes it is useful to handle efficient methods for estimation. Bounds for complex roots were obtained, among other, by Cauchy, Kuniyeda, Fujiwara, Landau and Montel. Upper and lower bounds can be derived using polynomial sizes defined in function of the coefficients, such that the norm, the length, the height and the measure.
For computing real roots it is necessary to have efficient root isolation methods.

Root isolation means the computation of intervals containing exactly one real root.

The CF (continued fraction) algorithms use estimates of the lowest bound for positive root.

The computation of lowest bounds is equivalent to that of upper bounds.

There are few methods that computes only bounds for real roots.

We describe a new method for computing upper bounds for positive roots.

Finally we discuss bounds concerning roots of orthogonal polynomials.
• We remind some of the most used bounds for complex roots of univariate polynomials.
• There exist many bounds for absolute values of complex roots of polynomials.
• All the bounds for complex roots can be used for computing upper bounds real roots.
Cauchy

Theorem (A.–L. Cauchy, 1829)

All the roots of the nonconstant complex polynomial

\[ P(X) = a_0 + a_1 X + \cdots + a_d X^d \]

are contained in the disk \(|z| \leq \xi\), where \(\xi\) is the unique positive solution of the equation

\[
|a_d|X^d = |a_0| + |a_1|X + \cdots + |a_{d-1}|X^{d-1}. \tag{1}
\]
**Theorem (M. Kuniyeda, 1916)**

If and $p, q > 0$ are such that $\frac{1}{p} + \frac{1}{q} = 1$, then all the roots of the polynomial $P$ are contained in the disk $|z| \leq \xi$, where

\[
\xi = \left(1 + \left(\sum_{j=0}^{n-1} \frac{|a_j|}{|a_n|} \frac{p^j}{q^j}\right)^{\frac{1}{q}}\right). 
\]
If $\lambda_1, \ldots, \lambda_d \in (0, \infty)$ and

$$\frac{1}{\lambda_1} + \cdots + \frac{1}{\lambda_d} = 1,$$

then all the roots of the polynomial $P$ are contained in the disk $|z| \leq \xi$, where

$$\xi = \max_{1 \leq k \leq d} \left( \lambda_k \frac{|a_{n-d}|}{a_n} \right)^{\frac{1}{k}}.$$
Dominant Roots

- A root $\alpha \in \mathbb{C}$ of the polynomial $P \in \mathbb{C}[X]$ is called dominant if $|\alpha| > |\beta|$ for any other root $\beta$.
- The computation of dominant roots was considered by Newton (1707) in his *Arithmetica Universalis*, no. 133–137. His idea was developed by Daniel Bernoulli (1728), who used linear recurrent sequences for approaching the dominant roots.
Open Problems

- There exist still open problems concerning the estimation of dominant roots.
- For example, let us suppose that a complex polynomial $P$ has exactly four dominant roots $\alpha_1, \ldots, \alpha_4$ such that $\alpha_1, \alpha_2$ are real and $\alpha_3, \alpha_4$ are complex conjugate.
- For example, the powerful methods of Bernoulli and Lobachevskii–Graeffe give inconvenient results.
Bounds for Real Roots

- Bounds for Real Roots can be derived from bounds for Complex Roots
- Lagrange obtained two upper bounds for positive real roots. One of them is surprisingly efficient.
- New bounds for positive roots were obtained by Kiostelidis (1986) and Ştefănescu (2005).
- We propose a general bound for positive roots.
Theorem (1)

[J.-L. Lagrange, 1769] Let

\[ P(X) = a_0 X^d + \cdots + a_m X^{d-m} - a_{m+1} X^{d-m-1} \pm \cdots \pm a_d \in \mathbb{R}[X], \]

with all \( a_i \geq 0, a_0, a_{m+1} > 0 \). Let

\[ A = \max \left\{ a_i ; \text{coeff} (X^{d-i}) < 0 \right\}. \]

The number

\[ 1 + \left( \frac{A}{a_0} \right)^{1/(m+1)} \]

is an upper bound for the positive roots of \( P \).
Lagrange L2

Note that in some special cases the following other bound of Lagrange can be useful:

**Theorem (2)**

If

\[
P(X) = X^d - \sum_{j \in J} a_j X^{d-j} + \text{positive terms} + \ldots,
\]

and

\[
\sqrt[\nu]{|a_u|} := R \geq \rho := \sqrt[\nu]{|a_j|} \geq \sqrt[i]{|a_i|} \quad \text{for all} \quad i \neq u, v.
\]

the number \( R + \rho \) is an upper bound for the positive roots.
The Bound of Fujiwara

One of the most efficient is the following

\[ Fw(P) = 2 \cdot \max \left| \frac{a_d - i}{a_d} \right|^{1/i} \]

In fact, the previous bound is a general one for complex roots, was obtained for the first time by Lagrange and rediscovered in 1916 by Fujiwara [3]. It was recently used by V. Sharma [10]. For complex roots it is one of the best.
Recent Bounds

Specific bounds for positive real roots were obtained by Kioustelidis (1986, [4]) and Ştefănescu (2005, [8]).

- They can be applied to polynomials having roots smaller than 1.
- They are more efficient than the classical bounds.
- They give methods for improving the bounds for particular classes of polynomials.
Theorem (3)

(J. B. Kioustelidis, 1986 [4]) Let

\[ P(X) = X^d - b_1 X^{d-m_1} - \ldots - b_k X^{d-m_k} + \sum_{j \neq m_1, \ldots, m_k} a_j X^{d-j}, \]

with \( b_1, \ldots, b_k > 0 \) and \( a_j \geq 0 \) for all \( j \notin \{m_1, \ldots, m_k\} \).

The number

\[ K(P) = 2 \cdot \{b_1^{1/m_1}, \ldots, b_k^{1/m_k}\} \]

is an upper bound for the positive roots of \( P \).
A result of Ştefănescu (2005)

Theorem (4)

(D. Ştefănescu, 2005 [8]) Let

\[ P(X) = X^d - b_1 X^{d-m_1} - \cdots - b_k X^{d-m_k} + \sum_{j \neq m_1, \ldots, m_k} a_j X^{d-j}, \]

with \( b_1, \ldots, b_k > 0 \) and \( a_j \geq 0 \) for all \( j \not\in \{m_1, \ldots, m_k\} \).

The number

\[ B_1(P) = \max\{ (kb_1)^{1/m_1}, \ldots, (kb_k)^{1/m_k} \} \]

is an upper bound for the positive roots of \( P \).
A Theorem of Ștefănescu (2005)

We remind our main bound from [8]:

**Theorem (5)**

D. Ștefănescu, 2005 [8] Let \( P(X) \in \mathbb{R}[X] \) be such that the number of sign variations of its coefficients is even. If

\[
P(X) = a_1 X^{d_1} - b_1 X^{m_1} + \cdots + a_s X^{d_s} - b_s X^{m_s} + g(X),
\]

where \( g(X) \in \mathbb{R}_+[X] \), \( a_j > 0 \), \( b_j > 0 \), \( d_j > m_j \) for all \( j \), the number

\[
St(P) = \max_{1 \leq j \leq s} \left\{ \left( \frac{b_j}{a_j} \right)^{1/(d_j-m_j)} \right\}
\]
Improved Methods

Theorem (6)

Let

\[ P(X) = a_1 X^{d_1} + a_2 X^{d_2} + \cdots + a_s X^{d_s} - b_1 X^{e_1} - b_2 X^{e_2} - \cdots - b_t X^{e_t} \in \mathbb{R}[X], \]

where \( a_i > 0 \), \( b_j > 0 \), \( d_1 = \text{deg}(P) \) and \( d_1 > d_2 > \cdots > d_s \).

An upper bound for the positive roots of \( P \) is given by

\[
\max_{\substack{1 \leq i \leq s \\ 1 \leq j \leq t \\ \beta_j \neq 0}} \left( \frac{\gamma_{ji} b_j}{\beta_j a_i} \right) \frac{1}{d_i - e_j}
\]
for any $\beta_j \geq 0$, $\gamma_{jk} \geq 0$ such that

$$\sum_{j=1}^{t} \beta_j \leq 1,$$

$$\sum_{i=1}^{s} \gamma_{ji} \geq 1 \quad \text{with} \quad \gamma_{ji} = 0 \quad \text{if} \quad d_i < e_j.$$

From the previous Theorem it is easy to obtain the bound $K$ of Kioustelidis (Theorem 3) and the bound $B_1$ of Ţşefănescu (Theorem 5), that apply to polynomials of the form

$$X^d - b_1 X^{d-m_1} - \cdots - b_k X^{d-m_k} + \text{positive terms}$$
A bound for \( s \geq t \)

**Proposition (7)**

If \( s \geq t \) and \( d_i > e_j \) for all \( i \) and \( j \), the number

\[
B_5(P) = \max_{1 \leq i \leq s, 1 \leq j \leq t} \left( \frac{b_j}{a_i} \right)^{1/(d_i-e_j)}
\]

is an upper bound for the positive roots of the polynomial

\[
P(X) = a_1 X^{d_1} + a_2 X^{d_2} + \cdots + a_s X^{d_s} - b_1 X^{e_1} - b_2 X^{e_2} - \cdots - b_t X^{e_t}.
\]
A New General Result

Theorem (8)

Let

\[ P(X) = a_1 X^{d_1} + a_2 X^{d_2} + \cdots + a_s X^{d_s} - b_1 X^{e_1} - b_2 X^{e_2} - \cdots - b_t X^{e_t} \in \mathbb{R}[X], \]

where \( a_i > 0 \), \( b_j > 0 \), \( d_1 = \deg(P) \) and \( d_1 > d_2 > \cdots > d_s \). An upper bound for the positive roots of \( P \) is given by

\[
B_6(P) = \max_{1 \leq i \leq s, 1 \leq j \leq t} \left( \frac{b_j}{\beta_j a_i} \right)^{\frac{1}{d_i - e_j}}
\]

for any \( \beta_j > 0 \) such that

\[
\beta_1 + \cdots + \beta_t \leq 1.
\]
The following cases will be considered:

\[ \beta_1 = \beta_2 = \ldots = \beta_t = \frac{1}{t}. \]  \tag{2} 

and

\[ \beta_1 = \frac{1}{2}, \quad \beta_2 = \ldots = \beta_t = \frac{1}{2(t - 1)}. \]  \tag{3} 

For (3) we obtained the following
B6 pseudocode

Input: $P(X) \in \mathbb{Q}[X]$. Output: A real number

$p := 0$
for $i = 1$ to $s$
  for $j = 1$ to $t$
    $r := d_i - e_j$
    if $r > 0$
      $q := \left( \frac{tb_j}{a_i} \right) \frac{1}{d_i - e_j}$
    fi
    if $q > p$
      $p = q$
    fi
  od
od
return $p$
Complexity of the algorithm

Let coefficients of a polynomial and its degree be bounded by $L$ bits, and let $n$ be an integer such that

$$n \gg \max(L, \log t)$$

**Theorem (9)**

*The bit-complexity for the $n$–bit evaluation of a function which can be exponential function, or a trigonometric function, or an elementary algebraic function, or their superposition, or their inverse, or a superposition of the inverses is given by

$$O(M(n) \log^2 n)$$

where $M(n)$ is the complexity function for multiplication of $n$–bit integers.*
Complexity of the algorithm (contd.)
Thus, the algorithm has the complexity of the bound estimation with accuracy up to \( n \) digits is

\[
O \left( s \cdot t \cdot M(n) \log^2 n \right)
\]

Here \( s \) and \( t \) are respectively the numbers of positive and negative coefficients in the input polynomial.

The best (w.r.t. complexity) known algorithms for integer multiplication are

- Karatsuba & Ofman (1962)
  \[
  M(n) = O \left( n^{\log_2 3} \approx n^{1.5849...} \right)
  \]
- Schönhage and Strassen (1971)
  \[
  M(n) = O(n \cdot \log n \cdot \log \log n)
  \]
Numerical Results

We compare various results on upper bounds for positive polynomials. The following notation will be used:

\[ L_1(P) = 1 + \left( \frac{A}{a_0} \right)^{1/(m+1)} \]
\[ L_2(P) = R + \rho \]
\[ Fw(P) = S(P) = 2 \cdot \max \left| \frac{a_{d-i}}{a_d} \right|^{1/i} \]
\[ K(P) = 2 \cdot \max \{ b_1^{1/m_1}, \ldots, b_k^{1/m_k} \} \]
Numerical Results (contd.)

\[ B_1(P) = \max \{ (kb_1)^{1/m_1}, \ldots, (kb_k)^{1/m_k} \} \]
\[ St(P) = \max_{1 \leq j \leq s} \left\{ \left( \frac{b_j}{a_j} \right)^{1/(d_j-m_j)} \right\} \]
\[ B_5(P) = \max_{1 \leq i \leq s, 1 \leq j \leq t} \left( \frac{b_j}{a_i} \right)^{1/(d_i-e_j)} \]
\[ B_6(P) = \max_{1 \leq i \leq s, 1 \leq j \leq t} \left( \frac{b_j}{\beta_j a_i} \right) \frac{1}{d_i-e_j} \]

We denote by TUB the true upper bound for positive roots.
The case all \( d_i > e_i \)

We consider the polynomials

\[
P_1(X) = X^{11} + 1.9X^2 - 3
\]

\[
P_2(X) = X^{11} + 3X^2 + 7X - 11.1
\]

\[
P_3(X) = X^{11} + 32X^3 + 2X^2 - 37
\]

<table>
<thead>
<tr>
<th>( P )</th>
<th>( L_1 )</th>
<th>( L_2 )</th>
<th>( Fw )</th>
<th>( K )</th>
<th>( B_1 )</th>
<th>( B_5 )</th>
<th>TUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 )</td>
<td>2.105</td>
<td>2.178</td>
<td>2.147</td>
<td>2.147</td>
<td>1.105</td>
<td>1.105</td>
<td>1.006</td>
</tr>
<tr>
<td>( P_2 )</td>
<td>2.244</td>
<td>2.489</td>
<td>2.489</td>
<td>2.489</td>
<td>1.244</td>
<td>1.923</td>
<td>1.004</td>
</tr>
<tr>
<td>( P_3 )</td>
<td>2.388</td>
<td>2.777</td>
<td>3.084</td>
<td>2.777</td>
<td>1.388</td>
<td>1.388</td>
<td>1.017</td>
</tr>
</tbody>
</table>
The case $s = t$

We consider the polynomials from [8] and $Q_5$.

\[ Q_1(X) = 3X^4 - X^3 + 7X^2 - 3X + 0.001 \]
\[ Q_2(X) = X^5 - 1.01X^4 + X^3 - 1.1X + 0.1 \]
\[ Q_3(X) = 3X^7 - X^6 + 7X^5 - 3X^2 + 0.001 \]
\[ Q_4(X) = 10X^9 - 17X^5 + 10X^4 - 13X + 1 \]
\[ Q_5(X) = 2X^{13} - 3X^5 + 12X^4 + 3X - 15.7 \]
The case $s = t$ (contd.)

We obtain

<table>
<thead>
<tr>
<th>$P$</th>
<th>$F_w$</th>
<th>$K$</th>
<th>$B_1$</th>
<th>$St$</th>
<th>TUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_1$</td>
<td>3.055</td>
<td>2.0</td>
<td>0.857</td>
<td>0.428</td>
<td>0.421</td>
</tr>
<tr>
<td>$Q_2$</td>
<td>2.048</td>
<td>2.048</td>
<td>1.483</td>
<td>1.048</td>
<td>1.003</td>
</tr>
<tr>
<td>$Q_3$</td>
<td>3.055</td>
<td>2.0</td>
<td>0.949</td>
<td>0.753</td>
<td>0.725</td>
</tr>
<tr>
<td>$Q_4$</td>
<td>2.283</td>
<td>2.283</td>
<td>1.375</td>
<td>1.141</td>
<td>1.12</td>
</tr>
<tr>
<td>$Q_5$</td>
<td>2.471</td>
<td>2.343</td>
<td>1.303</td>
<td>1.069</td>
<td>1.025</td>
</tr>
</tbody>
</table>
The General Case

In the general case we consider the bounds $L_1, L_2, Fw, K, B_1$ and $B_6$. For convenience, in $B_6$, we consider

$$
\beta_1 = \beta_2 = \ldots = \beta_t = \frac{1}{t}.
$$

(4)

We consider the following polynomials

$$
R_1 = X^{21} + 7X^9 - 8X^8 - 9X^6 + 5X^5 - 11
$$

$$
R_2 = X^{21} - 5X^9 - 8X^8 - 9X^6 + 5X^5 - 11
$$

$$
R_3 = X^7 + 2X^6 - 2X^4 - 3X^2 - X + 1
$$

and we have

<table>
<thead>
<tr>
<th>$P$</th>
<th>$L_1$</th>
<th>$L_2$</th>
<th>$Fw$</th>
<th>$K$</th>
<th>$B_1$</th>
<th>$B_6$</th>
<th>TUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1$</td>
<td>12.000</td>
<td>2.331</td>
<td>2.352</td>
<td>2.346</td>
<td>1.279</td>
<td>3.428</td>
<td>1158</td>
</tr>
<tr>
<td>$R_2$</td>
<td>12.000</td>
<td>2.316</td>
<td>2.346</td>
<td>2.346</td>
<td>1.305</td>
<td>1.544</td>
<td>1.255</td>
</tr>
<tr>
<td>$R_3$</td>
<td>2.732</td>
<td>2.505</td>
<td>4.000</td>
<td>2.519</td>
<td>1.817</td>
<td>1.817</td>
<td>1.165</td>
</tr>
</tbody>
</table>
The General Case (contd.)

Note that for the choice (4) we always have $B_6(P) \leq B_1(P)$. But for other choices of $\alpha$ we can have $B_6(P) < B_1(P)$. For example, if we consider

$$\beta_1 = \frac{1}{2}, \quad \beta_2 = \ldots = \beta_t = \frac{1}{2(t-1)}.$$ (5)

we obtain

<table>
<thead>
<tr>
<th>$P$</th>
<th>$B_1$</th>
<th>$B_6$</th>
<th>TUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1$</td>
<td>1.279</td>
<td>2.285</td>
<td>1.158</td>
</tr>
<tr>
<td>$R_2$</td>
<td>1.305</td>
<td>1.675</td>
<td>1.255</td>
</tr>
<tr>
<td>$R_3$</td>
<td>1.817</td>
<td>1.643</td>
<td>1.165</td>
</tr>
</tbody>
</table>

We observe that in this case $B_6$ is smaller for $R_2$ and $R_3$. On the other hand we have

$$B_6(R_3) < B_1(R_3).$$
Chebychev Polynomials of First Kind $T_d(X)$

$$T_d(X) = \frac{d}{2} \sum_{k=0}^{\lfloor d/2 \rfloor} \frac{(-1)^k 2^{d-2k}}{d-k} \binom{d-k}{k} X^{d-2k}$$

<table>
<thead>
<tr>
<th>$d$</th>
<th>$L_2$</th>
<th>$K$</th>
<th>$B_1$</th>
<th>$B_2$</th>
<th>$B_3$</th>
<th>TUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.7</td>
<td>1.732</td>
<td>0.866</td>
<td>0.866</td>
<td>1.225</td>
<td>0.866</td>
</tr>
<tr>
<td>16</td>
<td>3.3</td>
<td>4.000</td>
<td>4.000</td>
<td>4.000</td>
<td>2.828</td>
<td>0.995</td>
</tr>
<tr>
<td>63</td>
<td>6.8</td>
<td>7.937</td>
<td>15.875</td>
<td>15.875</td>
<td>11.941</td>
<td>0.999</td>
</tr>
<tr>
<td>251</td>
<td>13.7</td>
<td>15.843</td>
<td>62.875</td>
<td>62.875</td>
<td>50.316</td>
<td>0.999</td>
</tr>
<tr>
<td>356</td>
<td>16.4</td>
<td>18.868</td>
<td>89.000</td>
<td>89.000</td>
<td>71.648</td>
<td>0.999</td>
</tr>
<tr>
<td>500</td>
<td>19.4</td>
<td>22.361</td>
<td>125.000</td>
<td>125.000</td>
<td>101.042</td>
<td>0.999</td>
</tr>
</tbody>
</table>
Chebychev Polynomials of First Kind $T_d(X)$ (contd.)
Laguerre Polynomials $L_d(X)$

\[ L_d(X) = \sum_{k=0}^{d} (-1)^k \frac{1}{k!} \binom{n}{k} X^k. \]

<table>
<thead>
<tr>
<th>$d$</th>
<th>K</th>
<th>$B_1$</th>
<th>$B_6$</th>
<th>Mth-1</th>
<th>Mth-2</th>
<th>TUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>18.0</td>
<td>18.0</td>
<td>9.0</td>
<td>13.08</td>
<td>27.9</td>
<td>6.28</td>
</tr>
<tr>
<td>16</td>
<td>512.0</td>
<td>2048.0</td>
<td>256.0</td>
<td>$35 \cdot 10^2$</td>
<td>489.5</td>
<td>51.7</td>
</tr>
<tr>
<td>63</td>
<td>7938.0</td>
<td>$13 \cdot 10^4$</td>
<td>3970.0</td>
<td>$23 \cdot 10^2$</td>
<td>$29 \cdot 10^2$</td>
<td>230.9</td>
</tr>
<tr>
<td>300</td>
<td>$18 \cdot 10^4$</td>
<td>$13 \cdot 10^6$</td>
<td>$9 \cdot 10^4$</td>
<td>$3 \cdot 10^4$</td>
<td>$7 \cdot 10^5$</td>
<td>972.4</td>
</tr>
<tr>
<td>500</td>
<td>$5 \cdot 10^5$</td>
<td>$63 \cdot 10^6$</td>
<td>$25 \cdot 10^4$</td>
<td>$2 \cdot 10^6$</td>
<td>$6 \cdot 10^5$</td>
<td>7050.4</td>
</tr>
</tbody>
</table>
Conclusions

- There exist still open problems for estimating the dominant roots.
- The general bounds for complex roots can give good estimates, in particular the bound of Fujiwara.
- The bound $L_1$ can be used only in particular cases, especially when the largest positive root is greater than 1.
- The bound of Kioustelidis can be used for any polynomials having positive roots.
- The bound $R + \rho$ of Lagrange is better than that of Fujiwara.
- For polynomials with an even number of sign variations the bound Ştefănescu ST gives the best result.
- For polynomials with an arbitrary number of sign variations suitable applications of Theorem 6 can be used.
- Theorem 8 gives good estimates for all cases.
REFERENCES


THANK YOU VERY MUCH FOR YOUR ATTENTION!