Mathematica
Balkanica

New Series Vol. 22, 2008, Fasc. 3-4

Bounds for Polynomial Zeros Using the Companion Matrix

Rică Zamfir

Presented by V. Kirykova

Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in \mathbf{C}[X]$ with all $a_k \neq 0$. We apply several matrix inequalities to the Frobenius companion matrix of f to derive new bounds for the zeros of f.

Key Words: bounds for zeros, companion matrix, polynomial zeros

2000 Subject Mathematical Classification: 12D10

1. Introduction

It is well-known that matrix methods can be used to obtain certain root-location results for polynomials (see [4] or [5]). Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in \mathbf{C}[X]$ with $a_n \neq 0$. The Frobenius, or companion matrix is

$$(1) \qquad A = \begin{pmatrix} -\frac{a_{n-1}}{a_n} & -\frac{a_{n-2}}{a_n} & \cdots & -\frac{a_1}{a_n} & -\frac{a_0}{a_n} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} = (a_{ij})_{1 \le i,j \le n}$$

and the eigenvalues of A are the zeros of f. This simple result from Linear Algebra gives us the possibility to find bounds for zeros of f if we know bounds for the eigenvalues of its companion matrix.

If $p_1, p_2, ..., p_n > 0$, let $D = diag(p_1, p_2, ..., p_n)$. We have

$$D^{-1}AD = \left(\frac{p_j a_{ij}}{p_i}\right)_{1 \le i,j \le n},$$

therefore we can write

$$D^{-1}AD = \begin{pmatrix} -\frac{p_1 a_{n-1}}{p_1 a_n} & -\frac{p_2 a_{n-2}}{p_1 a_n} & \cdots & -\frac{p_{n-1} a_1}{p_1 a_n} & -\frac{p_n a_0}{p_1 a_n} \\ \frac{p_1}{p_2} & 0 & \cdots & 0 & 0 \\ 0 & \frac{p_2}{p_3} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \frac{p_{n-1}}{p_n} & 0 \end{pmatrix}. \tag{2}$$

We recall that if $\sigma(A)$ is the spectrum of A, then the spectral radius of A is

$$r(A) = max\{|\lambda|\,;\lambda \in \sigma(A)\} \tag{3}$$

and it is well known that $r(A) = r(D^{-1}AD)$. We are also using the fact that for every matrix norm $\|\cdot\|$ we have the inequality

$$r(A) \le \|A\|\,. \tag{4}$$

In [2], page 292 (see also [4], Chapter 6) we find a classical theorem due to Gershgorin.

Theorem 1. (Gershgorin) *Let $A = (a_{ij})$ be an $n \times n$ complex matrix, and let R_i be the sum of the moduli of the off-diagonal elements in the i-th row. Then each eigenvalue of A lies in the union of the circles*

$$|z - a_{ii}| \le R_i,$$

where $1 \le i \le n$. The analogous result holds, if columns of A are considered.

It is known (see [1], [2], or [4], Chapter 6) that for some choices of p_k we obtain some bounds for zeros of f. If, for example, we take

$$p_k = \left|\frac{a_k}{a_n}\right| \tag{5}$$

and we apply the Gershgorin theorem to $D^{-1}AD$, we find a bound due to Kojima (see [2], page 293)

$$|z| \le max\left\{\left|\frac{a_0}{a_1}\right|, 2\left|\frac{a_1}{a_2}\right|, ..., 2\left|\frac{a_{n-1}}{a_n}\right|\right\}.$$

2. Main results

In what follows we find new bounds for the zeros of f by choosing p_k in many different ways. These new bounds will be compared with some classical bounds.

Theorem 2. *Let* $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in \boldsymbol{C}[X]$ *with all* $a_k \neq 0$. *If* z *is a zero of* f, *then*

$$|z| \leq max\left\{\left|\frac{a_0}{a_1}\right|, \left|\frac{a_1}{a_2}\right|, ..., \left|\frac{a_{n-2}}{a_{n-1}}\right|, n\left|\frac{a_{n-1}}{a_n}\right|\right\}. \tag{6}$$

P r o o f. If we choose

$$p_k = \frac{1}{|a_{n-k}|},$$

then for all $1 \leq k \leq n$ we have

$$\left|\frac{p_k \cdot a_{n-k}}{p_1 \cdot a_n}\right| = \left|\frac{a_{n-1}}{a_n}\right| \tag{7}$$

and for $1 \leq k \leq n-1$ we obtain

$$\frac{p_k}{p_{k+1}} = \left|\frac{a_{n-k-1}}{a_{n-k}}\right|. \tag{8}$$

It is well known that

$$\left\|D^{-1}AD\right\|_\infty = max_{1\leq i\leq n}\left\{\sum_{j=1}^{n}|d_{ij}|\right\} \tag{9}$$

is a matrix norm and from (2.2) and (2.3) we obtain

$$\left\|D^{-1}AD\right\|_\infty = max\left\{\left|\frac{a_0}{a_1}\right|, \left|\frac{a_1}{a_2}\right|, ..., \left|\frac{a_{n-2}}{a_{n-1}}\right|, n\left|\frac{a_{n-1}}{a_n}\right|\right\}. \tag{10}$$

The conclusion of the theorem follows now from (1.4) and (2.5). ∎

Theorem 3. *Let* $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in \boldsymbol{C}[X]$ *with all* $a_k \neq 0$. *If* z *is a zero of* f, *then*

$$|z| \leq \left|\frac{a_{n-1}}{a_n}\right| + max\left\{\left|\frac{a_0}{a_1}\right|, \left|\frac{a_1}{a_2}\right|, ..., \left|\frac{a_{n-2}}{a_{n-1}}\right|\right\}. \tag{11}$$

P r o o f.

If we choose

$$p_k = p_1 \cdot \left|\frac{a_{n-1}}{a_{n-k}}\right|,$$

then for $1 \le k \le n-1$ we have

$$\frac{p_k}{p_1} \cdot \left|\frac{a_{n-k}}{a_n}\right| + \frac{p_k}{p_{k+1}} = \left|\frac{a_{n-1}}{a_n}\right| + \left|\frac{a_{n-k-1}}{a_{n-k}}\right|, \tag{12}$$

and

$$\frac{p_n}{p_1} \cdot \left|\frac{a_0}{a_n}\right| = \left|\frac{a_{n-1}}{a_n}\right|. \tag{13}$$

It is well known that

$$\left\|D^{-1}AD\right\|_1 = max_{1\le j\le n}\left\{\sum_{i=1}^{n} |d_{ij}|\right\} \tag{14}$$

is a matrix norm and from (2.7) and (2.8) we obtain:

$$\left\|D^{-1}AD\right\| = \left|\frac{a_{n-1}}{a_n}\right| + max\left\{\left|\frac{a_0}{a_1}\right|, \left|\frac{a_1}{a_2}\right|, ..., \left|\frac{a_{n-2}}{a_{n-1}}\right|\right\}. \tag{15}$$

The conclusion of the theorem follows now from (1.4) and (2.10). ■

Corollary 1. *If $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in \boldsymbol{R}[X]$ have all $a_k \neq 0$ and $a_0, a_1, ..., a_n$ are in arithmetical progression, then for any zero z of f we have*

$$|z| \le \left|\frac{a_{n-1}}{a_n}\right| + max\left\{\left|\frac{a_0}{a_1}\right|, \left|\frac{a_{n-2}}{a_{n-1}}\right|\right\}. \tag{16}$$

P r o o f. Let $g(x) = \frac{x}{x+r}$, where r is the ratio of progression. We have $g'(x) = \frac{r}{(x+r)^2}$ so if $r \ge 0$, then g is an increasing function, therefore for every $x \in \{a_0, a_1, ..., a_{n-2}\}$ we will have

$$g(a_0) \le g(x) \le g(a_{n-2}),$$

so:

$$|g(x)| \leq max\left\{|g(a_0)|, |g(a_{n-2}|\right\}. \tag{17}$$

Applying Theorem 3, we find

$$\begin{aligned} |z| &\leq \left|\frac{a_{n-1}}{a_n}\right| + max\left\{|g(a_0)|, |g(a_1)|, ..., |g(a_{n-2})|\right\} \\ &= \left|\frac{a_{n-1}}{a_n}\right| + max\left\{|g(a_0)|, |g(a_{n-2}|)\right\} \\ &= \left|\frac{a_{n-1}}{a_n}\right| + max\left\{\left|\frac{a_0}{a_1}\right|, \left|\frac{a_{n-2}}{a_{n-1}}\right|\right\}. \end{aligned}$$

Inequality (2.12) holds even if $r < 0$, which concludes the proof. ■

In [3] J.L. Diaz-Barrero has established the following theorem.

Theorem 4. *Let $z_1, z_2, ..., z_n$ be the zeros of the monic complex polynomial*

$$f(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 + a_0$$

with $a_k \neq 0$ for $k = 0, 1, ..., n-1$. Then, z_i^2 lies in the disk $|z| \leq 2r$, where

$$r = max\left\{\frac{|a_k|^2 + \sum_{j=1}^{n-k+1}|a_{k-j}\cdot a_{k+j}|}{|a_{k+1}|^2}; k = 0, 1, ..., n-1\right\} \tag{18}$$

with $a_l = 0$ if $l < 0$ or $l > n$.

In demonstration of Theorem 4, which used the Gershgorin theorem, the key result is the following lemma.

Lemma 1. *Let $g(x) = x^n + b_{n-1}x^{n-1} + ... + b_1 + b_0$ be the monic polynomial whose zeros are $z_1^2, z_2^2, ..., z_n^2$. Then, the coefficients of g are related to coefficients of f by the expressions*

$$b_k = (-1)^{n-k}\cdot\left[a_k^2 + 2\sum_{j=1}^{n-k+1}(-1)^j\cdot a_{k-j}\cdot a_{k+j}\right] \tag{19}$$

$0 \leq k \leq n-1$ with $a_l = 0$ if $l < 0$ or $l > n$.

If we apply several times the theorem of Diaz-Barrero, we can obtain better bounds.

Let $P_0, P_1, ...P_m$ be a sequence of monic polynomials thus that the roots of P_i are

$$z_1^{2^i}, z_2^{2^i}, ..., z_n^{2^i}$$

for $i = 0, 1, ..., m$. If

$$P_i(x) = x^n + a_{n-1}^{(i)} x^{n-1} + ... + a_1^{(i)} + a_0^{(i)}$$

then, using Lemma 1, we have

$$a_k^{(i+1)} = (-1)^{n-k} \cdot \left[(a_k^{(i)})^2 + 2 \sum_{j=1}^{n-k+1} (-1)^j \cdot a_{k-j}^{(i)} \cdot a_{k+j}^{(i)} \right], \tag{20}$$

$0 \leq k \leq n-1$ with $a_l^{(i)} = 0$ if $l < 0$ or $l > n$, therefore we can calculate the polynomials $P_1, P_2, ..., P_m$ by recurrence (obviously, $P_0 = f$).

If in (1.1) we make some permutations of lines and rows, we deduce another form for companion matrix of P_m

$$A_m = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0^{(m)} \\ 1 & 0 & 0 & \cdots & 0 & -a_1^{(m)} \\ 0 & 1 & 0 & \cdots & 0 & -a_2^{(m)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1}^{(m)} \end{pmatrix}, \tag{21}$$

and if $D = diag(p_1, p_2, ..., p_n)$ with $p_i > 0$, then

$$D^{-1}AD = \begin{pmatrix} 0 & 0 & & 0 \cdots & 0 & -\frac{p_n}{p_1} \cdot a_0^{(m)} \\ \frac{p_1}{p_2} & 0 & & 0 \cdots & 0 & -\frac{p_n}{p_2} \cdot a_1^{(m)} \\ 0 & \frac{p_2}{p_3} & 0 & \cdots & 0 & -\frac{p_n}{p_3} \cdot a_2^{(m)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \frac{p_{n-1}}{p_n} & -\frac{p_n}{p_n} \cdot a_{n-1}^{(m)} \end{pmatrix}. \tag{22}$$

Now, we apply the Gershgorin theorem, and obtain the following theorem.

Theorem 5. *Let $P_m(x) = x^n + a_{n-1}^{(m)} x^{n-1} + ... + a_1^{(m)} + a_0^{(m)}$ be the monic polynomial whose roots are $z_1^{2^m}, z_2^{2^m}, ..., z_n^{2^m}$. Then, for $1 \leq i \leq n$ we have*

$$|z_i|^{2^m} \leq max_{1 \leq k \leq n-1} \left\{ \frac{p_n}{p_1} \cdot \left| a_0^{(m)} \right|, \frac{p_k + p_n \cdot \left| a_k^{(m)} \right|}{p_{k+1}} \right\}. \tag{23}$$

Corollary 2. *If* $f(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 + a_0 \in \boldsymbol{C}[X]$ *with all* $a_k \neq 0$, *then for every* $m \geq 1$ *all roots of* f *are inside the disk*

$$|z| \leq (2r)^{\frac{1}{2^m}},$$

where:

$$r = max_{0\leq k\leq n-1}\left\{\frac{\left|a_k^{(m-1)}\right|^2 + \sum_{j=1}^{n-k+1}\left|a_{k-j}^{(m-1)}\right|\cdot\left|a_{k+j}^{(m-1)}\right|}{\left|a_{k+1}^{(m-1)}\right|^2}\right\} \tag{24}$$

with $a_l^{(m-1)} = 0$ *if* $l < 0$ *or* $l > n$.

Proof. Let

$$p_k = \left|a_k^{(m-1)}\right|^2$$

Then:

$$\frac{p_n}{p_1}\left|a_0^{(m)}\right| = \frac{\left|a_0^{(m)}\right|}{\left|a_1^{(m-1)}\right|^2} = \frac{\left|a_0^{(m-1)}\right|^2}{\left|a_1^{(m-1)}\right|^2},$$

and:

$$\frac{p_k + p_n\cdot\left|a_k^{(m)}\right|}{p_{k+1}} = \frac{\left|a_k^{(m-1)}\right|^2 + \left|a_k^{(m)}\right|}{\left|a_{k+1}^{(m-1)}\right|^2}$$

$$= \frac{\left|a_k^{(m-1)}\right|^2 + \left|(-1)^{n-k}\left[(a_k^{(m-1)})^2 + 2\sum_{j=1}^{n-k+1}(-1)^j\,a_{k-j}^{(m-1)}a_{k+j}^{(m-1)}\right]\right|}{\left|a_{k+1}^{(m-1)}\right|^2}$$

$$\leq \frac{2\left[\left|a_k^{(m-1)}\right|^2 + \sum_{j=1}^{n-k+1}\left|a_{k-j}^{(m-1)}\right|\cdot\left|a_{k+j}^{(m-1)}\right|\right]}{\left|a_{k+1}^{(m-1)}\right|^2},$$

where $a_l^{(m-1)} = 0$ if $l < 0$ or $l > n$. Applying Theorem 5, we obtain the desired result. ■

Remark 1. For $m = 1$, we obtain Theorem 4.

3. Applications

1. Let $f(x) = 61x^6 + 12x^5 - 11x^4 - 10x^3 + 9x^2 + 8x + 7$.

1. If we apply Theorem 2, we find $|z| \leq M_1 = 1.180$.

2. The classical bound of Cauchy gives us

$$|z| \leq M_2 = 1 + max_{0 \leq k \leq 5} \left\{ \left| \frac{a_k}{a_6} \right| \right\} = 1.196.$$

3. Kojima's bound is $M_3 = 1.833$.

2. Let$f(x) = 79x^6 + 12x^5 + 11x^4 + 10x^3 + 9x^2 + 8x + 7$.

1. Cauchy's bound is $M_4 = 1.151$.

2. Carmichael-Mason's bounds is $M_5 = \left(1 + \frac{|a_0|^2 + |a_1|^2 + \ldots + |a_5|^2}{|a_6|^2}\right)^{\frac{1}{2}} = 1.043$.

3. The Theorem of Enestrom-Kakeya gives $M_6 \leq 1$.

4. Theorem 2 gives a better bound $M_7 = 0.911$.

3. For some polynomials the bound from Theorem 3 is better than the bound from Theorem 2.

For example, if $f(x) = 61x^6 + 12x^5 - 11x^4 - 10x^3 + 9x^2 + 8x + 7$, we saw that Theorem 2 gives the bound $M_1 = 1.180$, while Theorem 3 gives us the better bound: $M_8 = 1.068$.

4. If $f(x) = x^3 - 1.1x^2 - 1.2x + 1.3$, Diaz-Barrero found in [3] that all the zeros lie in the disk $|z| \leq r_1$, where $r_1 = \sqrt{4.82} \approx 2.195$.

If we apply Corollary 3 for $m = 2$, we find that all the zeros lie in the disk $|z| \leq r_2$, where $r_2 = \sqrt[4]{14.3122} \approx 1.945$, and if we apply for $m = 3$ we find that $|z| \leq r_3$, where $r_3 = \sqrt[8]{28.891014} \approx 1.522$. Solving the equation $f(x) = 0$, we observe that the root of maximal modulus is $x \approx 1.191$.

Remark 2. The bound from Theorem 2 is sharp, if the first coefficient a_n is much bigger than the others coefficients.

References

[1] A. Aziz, Q.G. Muhammad, Zero free-region for polynomials and some generalizations of Enestrom-Kakeya theorem, *Can. J. Math.* **27** (1984), 265-272.

[2] H. Bell, Gershgorin theorem and the zeros of polynomials, *AMM*, **72**, No 3 (1965), 292-295.

[3] J.L.Diaz-Barrero, A bound for the square of the zeros, *Missouri Journal of Mathematical Sciences,* **16**, No 1 (2004), 3-7.

[4] R.A. Horn, C.R. Johnson, *Matrix Analysis,* Cambridge University Press, NY, 1990.

[5] M. Marden, *Geometry of Polynomials,* AMS, Providence, Rhode Island, 1989.

Bd. Iuliu Maniu 55A, Bloc 17A,
Sc A, Apart 5, Sect 6
Bucuresti, ROMANIA
Received: 29 April 2008
E-mail: rzamfir62@gmail.com