A new bound for polynomials when all the roots are real [1]

RWD Nickalls [2]

The Mathematical Gazette (2011); vol. 95 (November, No 534), pp. 520–526
www.nickalls.org/dick/papers/maths/bound2011.pdf

1 Introduction

Motivated by the recent notes on polynomial bounds [1, 2], I would like to suggest that sometimes a polynomial application is only appropriate, or of interest, when all the roots are real, in which case a much tighter bound is generally possible. With this setting in mind this note describes an upper bound having the property of being exact in the case of $n-1$ multiple roots.

Definition 1. *Let $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$ be a polynomial of degree n with real coefficients ($a_n \neq 0$). If x is a complex root of $P(x) = 0$, then let $|x|$ denote the modulus of x.*

Definition 2. *Let $P(x)$ be differentiated $n-2$ times to yield a quadratic with roots given by $\dfrac{-a_{n-1}}{na_n} \pm \Omega$, where* [3]

$$\Omega^2 = \frac{a_{n-1}^2}{n^2 a_n^2} - \frac{2a_{n-2}}{n(n-1)a_n}. \tag{1}$$

Significance of Ω

Since Ω is a factor of the x^{n-2} coefficient of the reduced form of a polynomial, it is also a factor of the elementary symmetric function of the roots $\Sigma x_1 x_2$. For example, by rearranging (1) and letting $a_{n-1} = 0$, we obtain

$$\frac{a_{n-2}}{a_n} = \frac{-n(n-1)}{2}\Omega^2 \equiv -\binom{n}{2}\Omega^2 \equiv \Sigma x_1 x_2. \tag{2}$$

2 Polynomial bound

Some insight regarding the general case is afforded by the polynomials of degree less than five having all real roots, since the maximum absolute value of any root of the reduced linear, quadratic and cubic equations are respectively $0 \times \Omega$, 1Ω, and 2Ω. For example, the roots of a reduced cubic with three real roots can be expressed in the form $2\Omega\cos(\theta + 2k\pi/3)$, $(k = 0, 1, 2)$ [4], and hence their upper absolute bound is 2Ω. This heuristic therefore suggested the following theorem.

[1] This article will be available in the JSTOR archive in due course.

[2] Nottingham, UK. email: dick@nickalls.org

Theorem. *If all the roots x of $P(x)=0$ are real, then all the roots lie in the range bounded by $-a_{n-1}/(na_n) \pm (n-1)\Omega$. For a reduced equation ($a_{n-1}=0$) this is equivalent to $|x| \leq (n-1)|\Omega|$.*

Proof. With no loss of generality let $P(x)$ be a reduced polynomial ($a_{n-1}=0$) with roots x_i $(i=1,\ldots,n)$. We distinguish two cases as follows:

1. Roots all equal: If $P(x)$ has a root of multiplicity n, then $P(x)\equiv x^n=0$, and hence $x=\Omega=0$, which is consistent with the assertion $|x|\leq(n-1)|\Omega|$.

2. All other cases: Let x be any root, with multiplicity r $(1\leq r<n)$, of the reduced form $P(x)$. Let the remaining $n-r$ roots have arithmetic mean g and be denoted as $x_j=\{g+\delta_j : \Sigma\delta_j=0 : j=1,2,\ldots,n-r\}$. Since $P(x)$ is reduced, then

$$rx=-(n-r)g. \tag{3}$$

In order to express x in terms of Ω we now proceed to determine the elementary symmetric function Σx_1x_2. It is convenient to distinguish three sets of terms in Σx_1x_2 namely A, the $\binom{r}{2}$ terms containing x only, B, the $r(n-r)$ terms which contain both x and $(g+\delta_j)$ factors, and C, the remaining $\binom{n-r}{2}$ terms which contain only $(g+\delta_j)$ factors, as follows:

$$\begin{aligned}
A &= \binom{r}{2}x^2=\frac{r(r-1)}{2}x^2=\frac{(r-1)(n-r)^2g^2}{2r}.\\
B &= r\left\{x(g+\delta_1)+\ldots+x(g+\delta_{n-r})\right\},\\
&= rx\left\{(g+\delta_1)+\ldots+(g+\delta_{n-r})\right\},\\
&= \left\{-(n-r)g\right\}\left\{(n-r)g+\delta_1+\ldots+\delta_{n-r}\right\}.
\end{aligned}$$

But $\Sigma\delta_j=0$, so this reduces to $B=-(n-r)^2g^2$.

$$\begin{aligned}
C =\ & (g+\delta_1)(g+\delta_2)+(g+\delta_1)(g+\delta_3)\\
& +\ldots+(g+\delta_{n-r-1})(g+\delta_{n-r}),\\
=\ & \left\{g^2+g(\delta_1+\delta_2)+\delta_1\delta_2\right\}+\left\{g^2+g(\delta_1+\delta_3)+\delta_1\delta_3\right\}\\
& +\ldots+\left\{g^2+g(\delta_{n-r-1}+\delta_{n-r})+\delta_{n-r-1}\delta_{n-r}\right\},\\
=\ & \binom{n-r}{2}g^2+g\left\{(\delta_1+\delta_2)+(\delta_1+\delta_3)\right.\\
& \left.+\ldots+(\delta_{n-r-1}+\delta_{n-r})\right\}+\Sigma\delta_1\delta_2,\\
=\ & \binom{n-r}{2}g^2+g(n-r-1)\Sigma\delta_j+\Sigma\delta_1\delta_2.
\end{aligned}$$

But $\Sigma\delta_j=0$, so C reduces to

$$C=\binom{n-r}{2}g^2+\Sigma\delta_1\delta_2.$$

Summing these expressions for A, B and C we have

$$\begin{aligned}
\Sigma x_1x_2 &= A+B+C\\
&= \frac{(r-1)(n-r)^2g^2}{2r}-(n-r)^2g^2\\
&\quad+\frac{(n-r)(n-r-1)g^2}{2}+\Sigma\delta_1\delta_2,
\end{aligned}$$

which simplifies to

$$\Sigma x_1 x_2 = \frac{-n(n-r)g^2}{2r} + \Sigma\delta_1\delta_2. \tag{4}$$

Now from (3) we have $x = -(n-r)g/r$, and so (4) is equivalent to

$$\Sigma x_1 x_2 = -x^2 \frac{rn}{2(n-r)} + \Sigma\delta_1\delta_2.$$

But, from (2) we have $\Sigma x_1 x_2 = -\binom{n}{2}\Omega^2$, and hence we can write

$$x^2 \frac{rn}{2(n-r)} - \Sigma\delta_1\delta_2 = \binom{n}{2}\Omega^2,$$

which is equivalent to

$$x^2 = \frac{(n-1)(n-r)}{r}\Omega^2 + \frac{2(n-r)}{rn}\Sigma\delta_1\delta_2. \tag{5}$$

If the roots of $P(x)$ are all real then, since the maximum absolute value of x, say x_{max}, is associated with the condition $r = 1$ and the remaining $n-1$ roots all being equal (that is, $\Sigma\delta_1\delta_2 = 0$), (5) reduces to

$$(x_{max})^2 = (n-1)^2\Omega^2. \tag{6}$$

If $r > 1$ then $(n-r)/r < n-1$ and also $\Sigma\delta_1\delta_2 \leq 0$ (see Lemma), and hence it follows from (5) that

$$x^2 < (n-1)^2\Omega^2, \tag{7}$$

irrespective of whether the remaining $n-r$ roots are equal ($\Sigma\delta_1\delta_2 = 0$) or not ($\Sigma\delta_1\delta_2 < 0$). It follows from (6) and (7), therefore, that if x is any root of a reduced univariate polynomial having all real roots then

$$|x| \leq (n-1)|\Omega|, \tag{8}$$

as required. □

Lemma. *If s_i $\{\in \mathbb{R} : i = 1, 2, \ldots, n\}$ are independent rational values such that $\Sigma s_i = 0$, then $\Sigma s_1 s_2 \leq 0$.*

Proof: Since $(\Sigma s_i)^2 \equiv \Sigma s_i^2 + 2\Sigma s_1 s_2$, if $\Sigma s_i = 0$ then $2\Sigma s_1 s_2 = -\Sigma s_i^2$. Since $\Sigma s_i^2 \geq 0$ it follows that $\Sigma s_1 s_2 \leq 0$. □

Remark. If the polynomial formed by the remaining $n-r$ roots of $P(x)$ be denoted as $Q(x)$, then the δ_j can be regarded as the roots of the reduced form of $Q(x)$ and hence it follows from (2) that

$$\Sigma\delta_1\delta_2 = -\binom{n-r}{2}\Omega_Q^2.$$

Consequently (5) can be expressed more generally as

$$\left(x + \frac{a_{n-1}}{na_n}\right)^2 = \frac{(n-1)(n-r)}{r}\Omega_P^2 - \frac{(n-r)^2(n-r-1)}{rn}\Omega_Q^2, \tag{9}$$

where x is any root (real or complex) of $P(x)$ and the subscripts denote the parent polynomial (see Example).

Example

Consider $P(x) \equiv x^5 - 7x^3 - 2x^2 + 12x + 8 = 0$ (which has roots: $-2, -1, -1, +2, +2$) for which (1) gives $\Omega^2 = 7/10$, and (8) gives $|x| \leq 3{\cdot}4$.

Note that three of the roots (-2, $+2$, $+2$) have the absolute maximum value of 2. If we allocate $x = -2$ then we have $n = 5$, $r = 1$, $g = 0{\cdot}5$, which are consistent with (3). The δ_j are therefore $3/2$, $3/2$, $-3/2$, $-3/2$, giving $\Sigma\delta_j = 0$ and $\Sigma\delta_1\delta_2 = -9/2$, consistent with (5). From the perspective of (9) we have $a_{n-1} = 0$, $Q(x) \equiv x^4 - 2x^3 - 3x^2 + 4x + 4 = 0$, $\Omega_Q^2 = 3/4$ and $\Omega_P^2 = 7/10$, which are consistent with $x = -2$.

Alternatively, if we allocate $x = +2$ (double root), then $n = 5$, $r = 2$, $g = -4/3$, which are consistent with (3) and lead to values consistent with (5). As regards (9) we have $a_{n-1} = 0$, $Q(x) \equiv x^3 + 4x^2 + 5x + 2 = 0$, $\Omega_Q^2 = 1/9$ and $\Omega_P^2 = 7/10$, which are consistent with $x = +2$.

3 Interpretation

When all the roots of $P(x)$ are real, we can envisage the position vectors of the n roots of the reduced equation as radiating symmetrically in $\mathbb{R}^{n-1}$ from the origin, each with radius $(n-1)|\Omega|$. For example, we regard the roots of a three-real-root reduced cubic as lying symmetrically in the complex plane on the circumference of a circle of radius $2|\Omega|$ [4]. For the higher polynomials we imagine the position vectors directed symmetrically towards the surface of a hypersphere of radius $(n-1)|\Omega|$.

If there are $n-1$ equal roots, then symmetry considerations require the position vector of the remaining root to coincide with the real axis, and hence its absolute value equals the radius, and the bound is therefore exact [3].

A consequence of this is that the bound is exact for all quadratics, irrespective of whether the roots are real or complex, since for equations of degree 2 then $n - 1 = 1$. In other words, whichever root of a reduced quadratic is allocated as x the remaining root is always an instance of $n-1$ equal roots.

4 Comparison with other methods

Useful overviews of other methods are [5, 6, 7]. As might be expected, methods which give tighter bounds (Term Grouping bound, Newton bound) are significantly more computationally intensive than those generating wider bounds (Cauchy bound, modified Cauchy bound, Maclaurin bound, Negative Inverse Sum bound). The last two bounds, which are generally better than the Cauchy bounds, are well detailed in [8]. These last four bounds are defined as follows:

> If $f(x)$ is a monic polynomial with coefficients $a_0, \ldots, a_n$, and N is the absolute value of the most negative coefficient, then the upper bounds for all the real roots of $f(x)$ are given by [5]:
>
> **Cauchy bound**. $|x| < 1 + max\,(|a_i|)$
>
> **Modified Cauchy bound**. $|x| \leq 1 + N$

[3] For example, in the case of the quartic see Figure 3 in: Nickalls RWD (2012): The quartic equation: alignment with an equivalent tetrahedron *The Mathematical Gazette*, 96, 49–55. http://www.nickalls.org/dick/papers/maths/tetrahedron2012.pdf

Maclaurin bound. *If k is the index of the last negative coefficient in the sequence $a_0, \ldots, a_n$, then*

$$|x| \leq 1 + N^{1/(n-k)}$$

Negative Inverse Sum bound. *For each negative coefficient a_m of $f(x)$, let S_m be the sum of all positive coefficients following a_m in the sequence $a_0, \ldots, a_n$. Then*

$$|x| \leq 1 + max\left(\left|\frac{a_m}{S_m}\right|\right)$$

Examples

We now give some examples and compare the various bounds. Since the tightness of the bound is influenced by the number of multiple roots (see (5)), we present a series of reduced equations having all-real roots, with increasing numbers of multiple roots up to $n-1$. The bounds are denoted as Nic (author), Mac (Maclaurin), NIS (Negative Inverse Sum), Cau (Cauchy) and mCau (modified Cauchy).

(a) $x^5 - 20x^3 - 30x^2 + 19x + 30 = 0$ (roots: -3, -2, -1, 1, 5),
$|x| \leq 5{\cdot}7$ (Nic), 6·5 (Mac), 31 (NIS), 31 (Cau), 31 (mCau).

(b) $x^5 - 24x^3 - 26x^2 + 87x + 90 = 0$ (roots: -3, -3, -1, 2, 5),
$|x| \leq 6{\cdot}2$ (Nic), 6·1 (Mac), 27 (NIS), 91 (Cau), 27 (mCau).

(c) $x^5 - 40x^3 - 90x^2 + 135x + 378 = 0$ (roots: -3, -3, -3, 2, 7),
$|x| \leq 8$ (Nic), 10·5 (Mac), 91 (NIS), 379 (Cau), 91 (mCau).

(d) $x^5 - 90x^3 - 540x^2 - 1215x - 972 = 0$ (roots: -3, -3, -3, -3, 12),
$|x| \leq 12$ (Nic), 35·9 (Mac), 1216 (NIS), 1216 (Cau), 1216 (mCau).

Thus, although the Nic bound is exact for polynomials having all-real roots and $n-1$ equal roots (example d), the Maclaurin bound will occasionally just out-perform it when there are fewer than $n-1$ equal roots, as shown in example b.

Acknowledgements

I would like to acknowledge the useful comments of the anonymous referee.

5 References

1. A. M. Cohen, Bounds for the roots of polynomial equations. *Math. Gaz.* **93** (March 2009), pp. 87–88.
2. R. Larham, Feedback on 93:02. *Math. Gaz.* **94** (March 2010), pp. 151–152.
3. R. W. D. Nickalls, The quadratic formula is a particular case. *Math. Gaz.* **84** (July 2000), pp. 276–280.
   http://www.nickalls.org/dick/papers/maths/quadratic2000.pdf
   http://www.jstor.org/stable/3621662

4. R. W. D. Nickalls, A new approach to solving the cubic: Cardan's solution revealed. *Math. Gaz.* **77** (November 1993), pp. 354–359.
http://www.nickalls.org/dick/papers/maths/cubic1993.pdf
http://www.jstor.org/stable/3619777

5. M. R. Spencer, Polynomial real root finding in Bernstein form. PhD thesis, Brigham Young University, Provo, Utah, USA (August 1994).
http://hdl.lib.byu.edu/1877/2395

6. H. P. Hirst and W. T. Macey, Bounding the roots of polynomials. *The College Mathematics Journal*; **28** (Sept 1997), pp. 292–295.

7. M. Marden, Geometry of polynomials. (2nd ed.), *American Mathematical Society* (Mathematical Surveys, No. 3) (1966).

8. W. S. Burnside and A. W. Panton, *The theory of equations: with an introduction to the theory of binary algebraic forms.* (7th ed.), 2 vols; Longmans, Green and Company, London, (1912); vol 1, pp. 180–188 .

◊ ◊ ◊