Good non-zeros of polynomials

Michael Moeller
FB Mathematik der Universität Dortmund
Vogelpothsweg 87, 44221 Dortmund, Germany
email: moeller@math.uni-dortmund.de

In numerical applications, one needs sometimes a point x^* with $F(x^*) \neq 0$, if $F \neq 0$. The usual procedure which helps in exact arithmetic (at most deg(F) unsuccessful trials) cannot be applied, since small values $F(x^*) \approx 0$ are numerically useless. Hence the problem is how to find an x^* with $|F(x^*)|$ in a reasonable order of magnitude compared to a norm of $F \neq 0$. In this exposition, I describe two ways of finding good points x^*.

The appropriate norm in this context is certainly the maximum norm over an interval or a compact region in the complex plane and x^* chosen from that region. However, the computation of this norm requires the solution of a maximum problem. Therefore it is more convenient to use a norm of the coefficient vector for fixing the norm of the polynomial and to use known results on the comparison of different polynomial norms like in the book [2]. Unfortunately, in [2] the maximum norm for polynomials is not considered. In the formulas (2) and (6) below, I compare a maximum norm with the euclidean norm. I hope, that the beauty of the following identity (1) and its elegant proof is also of some interest for the reader.

In Schönhage's article on quasi-gcd computations [4] x^* is constructed by means of the identity

$$\sum_{\nu=0}^{n} |a_\nu|^2 = \frac{1}{N} \sum_{j=0}^{N-1} |F(\omega^j)|^2$$

where

$$F(z) = \sum_{\nu=0}^{n} a_\nu z^\nu, \ \omega = exp(\frac{2\pi i}{N}), n < N.$$

This identity says, that the squared euclidean norm of the coefficient vector is the arithmetical mean of the values $|F(\omega^j)|^2$, $j = 0, \dots, N-1$. (1) can be proved by using properties of the primitive N-th root ω. I found a more direct proof of this formula, which goes as follows.

Proof. Let $N \in \mathbb{N}$ and $\omega = exp(\frac{2\pi i}{N})$. Then define the $N \times N$-matrix $\Omega := (\omega^{jk})_{j,k=0}^{N-1}$. It is well known that with $\overline{\Omega} = (\overline{\omega}^{jk})_{j,k=0}^{N-1}$ the identity

$$\overline{\Omega}\Omega = N \cdot I$$

holds, see for instance [1,p. 131]. Hence the symmetric matrix $\frac{1}{\sqrt{N}}\Omega$ is unitary. Let $a_\nu := 0$ for $\nu = n+1, \dots, N-1$ and define

$$M_1 := \begin{pmatrix} \overline{a_0} \\ \overline{a_1} \\ \vdots \\ \overline{a_{N-1}} \end{pmatrix} (a_0, \dots, a_{N-1}).$$

Then the matrices M_1 and

$$M_2 := \frac{1}{\sqrt{N}} \overline{\Omega} M_1 \Omega \frac{1}{\sqrt{N}} = \frac{1}{N} (\overline{F(\omega^j)} F(\omega^k))_{j,k=0}^{N-1}$$

are similar. Therefore

$$\sum_{\nu=0}^{n} |a_\nu|^2 = trace(M_1) = trace(M_2) = \frac{1}{N} \sum_{j=0}^{N-1} |F(\omega^j)|^2.$$

Corollary 1. Let $\|F\|_2 := \sqrt{\sum_{\nu=0}^{n} |a_\nu|^2}$, where $F(z) = \sum_{\nu=0}^{n} a_\nu z^\nu$, and $\|F\|_\infty := max\{ |F(z)| \ : \ |z| \leq 1\}$. Then

$$\frac{1}{\sqrt{n+1}} \|F\|_\infty \leq \|F\|_2 \leq \|F\|_\infty. \tag{1}$$

Proof. By the maximum principle for holomorphic functions, $\|F\|_\infty = |F(e^{it})|$ for a real t. Define F_t by $F_t(z) := F(ze^{it})$. Then $\|F\|_2 = \|F_t\|_2$ and $\|F\|_\infty = \|F_t\|_\infty = |F_t(1)|$. The application of (1) with $N = n+1$ on F_t gives

$$\frac{1}{n+1} |F_t(1)|^2 \leq \|F_t\|_2^2 = \frac{1}{n+1} \sum_{j=0}^{n} |F_t(\omega^j)|^2$$
$$\leq \frac{1}{n+1} \sum_{j=0}^{n} \|F_t\|_\infty^2 = \|F_t\|_\infty^2$$

whence the assertion follows.

Both bounds are sharp. The upper bound because of $F = 1$, the lower because of $F(z) = \sum_{\nu=0}^{n} z^\nu$, since then $\|F\|_\infty = F(1) = n+1$ and $\|F\|_2 = \sqrt{n+1}$.

Now, the problem we started with can be solved by selecting as x^* an ω^k such that $|F(\omega^k)|^2$ is not less than the arithmetic mean of all $|F(\omega^j)|^2$, $j = 0, \dots, n$. Then

$$\frac{1}{\sqrt{n+1}} \|F\|_\infty \leq |F(x^*)| \leq \|F\|_\infty \tag{2}$$

and obviously $\|F\|_2 \leq |F(x^*)|$.

An identity using only real points similar to (1) can be obtained by the discrete orthogonality of Chebyshev polynomials, as given for instance in [3, p.50],

$$\sum_{\nu=0}^{n}{}'' T_k(x_\nu)T_m(x_\nu) = \begin{cases} n & \text{for } k=m\in\{0,n\} \\ \frac{n}{2} & \text{for } k=m\in\{1,\dots,n-1\} \\ 0 & \text{for } k\neq m,\ 0\le k,m\le n \end{cases} \tag{3}$$

Here, $\sum_{\nu=0}^{n}{}'' u_\nu := \frac{1}{2}u_0 + u_1 + \ldots + u_{n-1} + \frac{1}{2}u_n$ and $x_\nu := cos(\frac{\nu\pi}{n})$ and $T_k(x) := cos(k\ arccos(x))$, the k-th Chebyshev polynomial (first kind). Using the $(n+1)\times(n+1)$ matrices $U := (T_k(x_\ell))_{k,\ell=0}^{n}$ and $D := diag(\frac{1}{\sqrt{2}}, 1, \ldots, 1, \frac{1}{\sqrt{2}})$, the discrete orthogonality (4) reads

$$UD(UD)^T = \frac{n}{2}D^{-2}.$$

U is symmetric because of $T_k(x_\ell) = cos(\frac{k\ell\pi}{n}) = T_\ell(x_k)$. Hence $V := \sqrt{\frac{2}{n}}DUD$ is a real-symmetric orthogonal matrix.

Proposition. Let $p := \sum_{\nu=0}^{n}{}'' b_\nu T_\nu$ and $x_\nu := cos(\frac{\nu\pi}{n})$ for $\nu = 0, \ldots, n$. Then

$$\sum_{\nu=0}^{n}{}'' b_\nu^2 = \frac{2}{n}\sum_{\nu=0}^{n}{}'' p(x_\nu)^2. \tag{4}$$

Proof. In analogy to the proof of (1), we define

$$M_1 := \begin{pmatrix} b_0/\sqrt{2} \\ b_1 \\ \vdots \\ b_{n-1} \\ b_n/\sqrt{2} \end{pmatrix} (\frac{b_0}{\sqrt{2}}, b_1, \ldots, b_{n-1}, \frac{b_n}{\sqrt{2}})$$

and $M_2 := VM_1V$. Then by some matrix calculations using the symmetry of V, the comparison of the traces gives the assertion.

If I define here $||p||_\infty := max\{|p(x)|\ :\ -1 \le x \le 1\}$, the Cauchy-Schwarz inequality gives using $|T_\nu(\xi)| \le 1$ for all $\xi \in [-1,1]$ and for the first and last summand of $\sum''$ using $\frac{ab}{2} = \frac{a}{\sqrt{2}}\frac{b}{\sqrt{2}}$,

$$\begin{aligned} ||p||_\infty^2 = |p(\xi)|^2 &= |\sum_{\nu=0}^{n}{}'' b_\nu T_\nu(\xi)|^2 \\ &\le \sum_{\nu=0}^{n}{}'' b_\nu^2 \sum_{\nu=0}^{n}{}'' T_\nu(\xi)^2 \le n\sum_{\nu=0}^{n}{}'' b_\nu^2. \end{aligned}$$

Hence $\frac{1}{\sqrt{n}}||p||_\infty \le ||p||_2 := \sqrt{\sum_{\nu=0}^{n}{}'' b_\nu^2}$. This bound is sharp because of $p := \sum_{\nu=0}^{n}{}'' T_\nu$. An immediate consequence of the proposition is then for $p := \sum_{\nu=0}^{n}{}'' b_\nu T_\nu$ and $x^* = cos(\frac{k\pi}{n})$ such that $|p(x^*)| = max_{\nu=0}^{n}|p(cos(\frac{\nu\pi}{n}))|$

$$\frac{1}{\sqrt{2n}}||p||_\infty \le \frac{1}{\sqrt{2}}||p||_2 \le |p(x^*)| \le ||p||_\infty. \tag{5}$$

References

[1] Geddes, K.A., Czapor, S.R., and Labahn, G.: Algorithms for Computer Algebra. Kluwer Academic Publisher, 1992.

[2] Mignotte, M. and Stefanescu, D.: Polynomials, an Algorithmic Approach. Springer, 1999.

[3] Rivlin, Th. J.: The Chebyshev-Polynomials, John Wiley&Sons, 1974.

[4] Schönhage, A. : Quasi-GCD computations, J. of Complexity, **1**, 118 – 137 (1985).