ON THE DISTRIBUTION OF COEFFICIENTS OF RESIDUE
POLYNOMIALS

LÁSZLÓ MAJOR

Abstract. For any nonconstant polynomial with positive coefficients and positive integer \(d\) if \(r(p^n)\) denotes the least degree residue of \(p^n\) modulo \(x^d - 1\), then the coefficients of \(r(p^n)\) are asymptotically equal as \(n\) tends to infinity.

When Euler (1765) was investigating the properties of the trinomial coefficients (see Andrews [1]), he obtained an unimodal distribution of the coefficients by expanding the \(n^{th}\) power of the polynomial \(1 + x + x^2\). In the analogous case of the polynomial \(1 + x\) the corresponding unimodal distribution is the \(n^{th}\) row of Pascal’s triangle. In general the distribution of the coefficients of the \(n^{th}\) power of the polynomial \(p = p_0 + p_1 x + \cdots + p_m x^m\) with nonnegative real coefficients is not necessarily unimodal, but a sufficient condition is given for unimodality by Boros and Moll [2]. We propose to examine the coefficient distribution of the powers of \(p\) in the polynomial residue class ring modulo \(x^d - 1\).

Let \(2 \leq d\) be a fixed positive integer. We consider the polynomial ring \(\mathbb{R}[x]\) and the quotient ring \(\mathbb{R}[x]/x^d - 1\). Let \(p\) denote both a polynomial of \(\mathbb{R}[x]\) with nonnegative coefficients and the residue class of \(\mathbb{R}[x]/x^d - 1\) represented by \(p\). The proper meaning will be determined by the context. We use the most common representative system formed by all polynomials of degree less than \(d\). Let \(r(p)\) denote that unique polynomial of degree less than \(d\) (denoted \(\deg(r(p)) < d\)) for which the polynomial \(p \in \mathbb{R}[x]\) is congruent to \(r(p)\) modulo \(x^d - 1\). In consequence the \(n^{th}\) power of \(p \in \mathbb{R}[x]/x^d - 1\) will mean the residue class of \(r(p^n)\). A polynomial with positive (nonnegative) coefficients is said to be positive (nonnegative) polynomial.

The main result is Theorem 1. The lemmas and corollaries provide background and an extension.

Lemma 1. Let \(p \in \mathbb{R}[x]\) be a nonnegative, nonzero polynomial. For any \(q \in \mathbb{R}[x]\) and \(i \in \mathbb{N}\) let \(q_i\) denote the coefficient of \(x^i\) in \(q\). If there exist \(k, l \in \{0, \ldots, d-1\}\) such that \(r(p)_k > 0\), \(r(p)_l > 0\) and \(\gcd(d, k - l) = 1\), then \(r(p^{d-1})\) is a positive polynomial.

Proof. It can be assumed that \(k < l\). Let us denote the difference \(l - k\) by \(h\). It is easy to see that if \(r((x^k + x^l)^{d-1})\) is positive then also \(r(p^{d-1})\) is positive. In addition, \(r((x^k + x^l)^{d-1}) = r(x^{k(d-1)}(1 + x^h)^{d-1})\) is positive if \(r((1 + x^h)^{d-1})\) is positive, so let us concentrate on the polynomial \(r((1 + x^h)^{d-1})\). We expand the
expression \((1 + x^h)^{d-1}\) in the polynomial ring \(\mathbb{R}[x]\) by using the Binomial Theorem:

\[(0.1) \quad (1 + x^h)^{d-1} = \sum_{i=0}^{d-1} \binom{d-1}{i} (x^h)^i.\]

Let us assume that the following congruence holds for some \(i, j \in \{0, \ldots, d-1\}:

\[(0.2) \quad \binom{d-1}{i} (x^h)^i \equiv \binom{d-1}{j} (x^h)^j \pmod{x^{d-1}}.\]

From the equation (0.2) follows that \(h^i - h^j\) is divisible by \(d\). But it was assumed that \(\gcd(h, d) = 1\), consequently \(i = j\). Therefore \(r(p^{d-1})\) has exactly \(d\) different nonzero coefficients, in other words it is positive. \(\Box\)

**Corollary 1.** If \(n \geq d-1\) is an integer and \(p\) satisfies the conditions of the Lemma 1, then \(r(p^n)\) is a positive polynomial.

In order to be able to deal more efficiently with powers of polynomials we need the concept and some useful properties of **circulant matrices**. For a classical reference see e.g. Davis [3]. Let \(v = (v_0, v_1, \ldots, v_{d-1})\) be a row vector in \(\mathbb{R}^d\). The permutation \(\rho : \mathbb{R}^d \rightarrow \mathbb{R}^d\) given by

\[\rho(v_0, v_1, \ldots, v_{d-1}) = (v_{d-1}, v_0, \ldots, v_{d-2})\]

is called cyclic permutation. The **circulant matrix** associated to the vector \(v\) is the \(d \times d\) matrix whose \(i^{th}\) row is \(\rho^{i-1}(v)\), \(i = 1, \ldots, d\) and it is denoted by

\[C = \text{circ}(v_0, v_1, \ldots, v_{d-1}) = \text{circ}(v).\]

The product of two circulant matrices is circulant, therefore any positive integer power of a circulant matrix is circulant. We shall use the connection between powers of circulant matrices and the powers of residue polynomials. If \(a = (a_0, a_1, \ldots, a_{d-1})\) is the coefficient vector of some polynomial \(r(p) \in \mathbb{R}[x]\), then the first row of \((\text{circ}(a))^n\) is the coefficient vector of \(r(p^n)\). Lally and Fitzpatrick [4] give a general overview in this subject.

A matrix is called **positive (nonnegative)** if all its entries are positive (nonnegative). A \(d \times d\) nonnegative matrix \(M\) is said to be **doubly stochastic** if the sum of the entries in each row and in each column equals 1. The product of doubly stochastic matrices is doubly stochastic. The following lemma follows from a general result on eigenvectors of positive matrices, Theorem 8.2.8 of Horn and Johnson [5]. Here we provide a direct proof. A matrix whose entries are all ones will be denoted by \(U\).

**Lemma 2.** If \(M\) is a \(d \times d\) positive doubly stochastic matrix, then

\[(0.3) \quad \lim_{n \to \infty} M^n = \frac{1}{d} U.\]

**Proof.** Let us denote the doubly stochastic matrix \(\frac{1}{d} U\) by \(J\). If \(M = J\), then the statement is trivial. Otherwise the matrix \(M\) can be given in the following form:

\[(0.4) \quad M = \lambda J + (1 - \lambda) M_0,\]

where \(0 < \lambda < 1\) and \(M_0\) is a doubly stochastic matrix. Let us write the \(n^{th}\) power of \(M\) by using the Binomial Theorem and the fact that \(JD = DJ = J\) for all
doubly stochastic matrix $D$, 
\[
M^n = \sum_{i=0}^{n} \binom{n}{i} (\lambda J)^{n-i}((1 - \lambda)M_0)^i \\
= (\lambda J)^n + J \sum_{i=1}^{n} \binom{n}{i} \lambda^{n-i}(1 - \lambda)^i = (\lambda)^n J + J.
\]

From the above expression follows our statement, because $0 < \lambda < 1$ and consequently $\lim_{n \to \infty} (\lambda)^n = 0$. □

In order that the Lemma 2 could be applied for the powers of polynomials, we shall restrict our attention to polynomials $p$ whose coefficients sum to 1 ($p(1) = 1$). We may do this without loss of generality, because all polynomials can be written in the form $p = p(1) \cdot p'$, where $p'$ has the mentioned property. It can also be said that the coefficients of $p'$ form a stochastic vector. Clearly the sum of the coefficients of $(p')^n$ is also 1 and the sum of the coefficients of $p^n$ is equal to $p(1)^n$. By using Lemma 1, Lemma 2 and the connection between powers of circulant matrices and the powers of residue polynomials we obtain the following theorem:

**Theorem 1.** For any polynomial $q$ and nonnegative integer $i$ let $q_i$ denote the coefficient of $x^i$ in $q$. Let $p$ be a nonconstant polynomial with positive real coefficients summing to 1 and $d \geq 2$ a fixed integer. Then for all $0 \leq j \leq d - 1$

\[(0.5) \quad \lim_{n \to \infty} \sum_{k \equiv j \text{ mod } d} (p^n)_k = \frac{1}{d} \]

The restriction, that $p$ is positive, is quite strong. Theorem 1 can be extended to larger classes of polynomials. For instance the following corollary applies Lemma 1 so as to provide such an extension.

**Corollary 2.** If the nonnegative polynomial $p$ satisfies the conditions of Lemma 1 and its coefficients sum to 1, then the convergence (0.5) holds for $p$ or equivalently for all $0 \leq j \leq d - 1$

\[\lim_{n \to \infty} r(p^n)_j = \frac{1}{d}.\]

**Proof.** Let $p$ be a polynomial satisfying the conditions of Lemma 1 and let $p(1) = 1$. Let us denote the coefficient vector of $r(p)$ by $c$ and the circulant $\text{circ}(c)$ by $C$. By Corollary 1 if $m \geq d - 1$ then $C^m$ is a positive matrix. We use $J$ to denote the $d \times d$ matrix in which every element is $\frac{1}{d}$. By Lemma 2 $\lim_{n \to \infty} C^{(d-1)n} = J$. We may also express this fact by using the max norm of matrices:

\[\lim_{n \to \infty} \|C^{(d-1)n} - J\| = 0.\]

We show that for all $0 \leq i \leq d - 1$, $\lim_{n \to \infty} C^{(d-1)n+i} = J$. We recall that $JC^i = J$, so we have

\[(0.6) \quad \|C^{(d-1)n}C^i - J\| = \|C^{(d-1)n}C^i - JC^i\| \leq \|C^{(d-1)n} - J\| \cdot \|C^i\|.
\]

Because $\|C^i\|$ is a constant, we obtain that $\lim_{n \to \infty} C^{(d-1)n+i} = J$ for all $0 \leq i \leq d - 1$, consequently $C^m \to J$ as $m$ tends to infinity. □
Example Theorem \( \text{II} \) applies to the polynomial \( p = 1 + x \) and \( d = 3 \). The \( n^{th} \) row of Pascal’s triangle displays the coefficients of \( p^n \). In this case Theorem \( \text{II} \) says that the sum of every third element of the \( n^{th} \) row of Pascal’s triangle is asymptotically equal to \( \frac{1}{3} \cdot 2^n \).

References

[1] G.E. Andrews Euler’s "Exemplum Memorabile Inductionis Fallacis" and q-Trinomial Coefficients, J. American Math. Soc., 3 (1990), 653-669.
[2] G. Boros, V. Moll, : A Criterion for Unimodality, The Electronic Journal of Combinatorics, 6, (1999), 1-6.
[3] P.J. Davis: Circulant Matrices, John Wiley and Sons, 1979.
[4] K. Lally, P. Fitzpatrick: Algebraic Structure of Cuasicyclic Codes, Discrete Applied Mathematics, 111, (2001), 157-175.
[5] R.A. Horn, C.R. Johnson: Matrix Analysis, Cambridge University Press, 1991.