A New Method for Polynomial Real Root Isolation

Alkiviadis G. Akritas (North Carolina State University)

Abstract. A new method is presented for the isolation of the real roots of a given integral, univariate, square-free polynomial P. This method is based on Vincent's theorem and only uses: (i) Descartes' rule of signs, and (ii) transformations of the form $x = a_1 + \frac{1}{x'}$, $x' = a_2 + \frac{1}{x''}$, $x'' = a_3 + \frac{1}{x'''}$, ..., for positive, integral a_i's. The key element in this procedure is the calculation of the quantities $a_1, a_2, a_3, \ldots$. We compute them as "positive lower root bounds" of polynomials and the resulting algorithm has the best theoretical computing time achieved thus far. Empirical results also verify the superiority of our method over all others existing.

1. Introduction

In order to solve numerically an algebraic equation of degree higher than four, Fourier suggested to first isolate its roots and then approximate them to any desired degree of accuracy.

Isolation of the real roots of a polynomial equation is the process of finding real intervals, such that each contains exactly one real root and every real root is contained in some interval. Since 1830 the only method widely known and used for this purpose, was that of Sturm. Heindel [6] implemented it within a computer algebra system (using infinite-precision arithmetic) and showed that it is $O(n^{13}L(|P|_\infty)^3)$, where n is the degree of the integral, univariate, square-free polynomial P, and $L(|P|_\infty)$ the length, in bits, of its maximum coefficient in absolute value.

Quite recently, Collins and Loos [5] realized the drawbacks of Sturm's method and together they developed a new algorithm for the isolation of the real roots of an integral, univariate polynomial P. Their algorithm utilizes the sequence of derivatives of the given polynomial, relying on Rolle's theorem and a tangent construction to decide whether an interval contains two roots or none. They proved, [5] theorem 8, that their algorithm is $O(n^{10} + n^7L(|P|_\infty)^3)$.

Shortly thereafter, in Uspensky's Theory of Equations [8], the author came across Vincent's theorem [9]. This remarkable theorem asserts that if a univariate, square-free polynomial is successively transformed by successive substitutions of the form $x \leftarrow a_i + \frac{1}{x}$, for arbitrary positive a_i's, one eventually obtains a polynomial with at most one sign variation. Clearly, the calculation of the quantities $a_1, a_2, \ldots$ constitutes the real root isolation procedure. Two algorithms result--Vincent's and Akritas'--corresponding to the two different ways of computing these a_i's; the difference between these methods can be thought of as being analogous to the one between the integrals of Riemann and Lebesgue. Vincent's approach is by "brute force", in the sense that he computes a particular a_i by always incrementing it by one. As the reader can easily feel, this results in an exponential algorithm. Collins and Akritas [4] attempted to construct an algorithm with polynomial bound.

Our present method is quite different; instead of using "brute force", we easily compute these $a_1, a_2, \ldots$ as the "positive lower root bounds" of polynomials, and thus, we are able to prove that, for an integral, univariate, square-free polynomial P, our method is $O(n^5L(|P|_\infty)^3)$. This is the best theoretical computing time achieved thus far and empirical results verify the superiority of Akritas' method over all others existing.

2. Vincent's Theorem

Relevant to the statement of this theorem is Descartes' rule of signs. We begin with the following:

Definition. We say that a sign variation exists between two members c_p and c_q $(p<q)$ of a finite or infinite sequence of real numbers

$$c_1, c_2, \ldots, \qquad (1)$$

if c_p and c_q are not zero, have opposite signs, and in case $q \geq p + 2$ (that is, c_q does not immediately follow c_p) then, the numbers $c_{p+1}, \ldots, c_{q-1}$ are all zero.

The following facts are direct consequences of the definition.

(i) The sequence

(2) $$c_n, c_{n-1}, \ldots, c_0$$

has the same number of sign variations as the sequence

(3) $$c_0, c_1, \ldots, c_n .$$

(ii) If $c_0 \cdot c_n > 0$ then (3) has an even number of sign variations, whereas, if $c_0 \cdot c_n < 0$ then the number of sign variations in (3) is odd.

Theorem 1 (Descartes' rule of signs). The number of positive roots of a polynomial equation with real coefficients

(4) $$c_0 x^n + c_1 x^{n-1} + \ldots + c_{n-1} x + c_n = 0$$

is never greater than the number of sign variations in the sequence of coefficients

(5) $$c_0, c_1, \ldots, c_{n-1}, c_n ;$$

when it is less, then the difference is an even number.

We would like to remark that Descartes' rule of signs gives the exact number of roots only if there is one or no sign variation. In the first case there is one positive real root, whereas in the second, there is no root. This observation is of great importance in what follows.

Theorem 2 (Vincent). If in a polynomial equation with rational coefficients and without multiple roots, one makes the successive transformations of the form

(6) $$x = a_1 + \frac{1}{x'}, \; x' = a_2 + \frac{1}{x''}, \; x'' = a_3 + \frac{1}{x'''}, \ldots ,$$

where each $a_1, a_2, a_3, \ldots$ is any positive integer, then the resulting, transformed equation has either zero or one sign variation. In the latter, the equation has a single positive real root represented by the continued fraction

(7) $$a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}} ,$$

whereas, in the former case there is no root.

This theorem appeared in 1836 but, as Uspensky notes [8] p. 298, even such a capital work as the Enzyclopädie der mathematischen Wissenschaften ignores it. It was totally forgotten for over a century, despite the fact that it is the basis of the most efficient algorithm for polynomial real root isolation. The author came across Vincent's theorem, while reviewing methods for the isolation of real roots of equations as presented by Uspensky [8]; a historical survey and some speculations as to the reasons why it was consigned to oblivion, is the subject of another paper [3]. (Notice that the negative roots are investigated by replacing x by -x in the original polynomial and by investigating the positive roots of the transformed polynomial.)

Uspensky, [8] pp. 298-303, extended Vincent's theorem, in the sense that he computed the upper bound m, on the number of transformations required, so that the resulting polynomial has not more than one sign variation. Uspensky's theorem though, contains an error, both in its statement and the proof. The author corrected it in [1], and [2], and proved that m is computed as the index of the smallest number in the Fibonacci sequence (F_k, $k \geq 0$)

(8) 1, 1, 2, 3, 5, 8, 13, 21, ...

such that both of the following inequalities hold simultaneously:

(9) $$F_{m-1} \cdot \frac{\Delta}{2} > 1 \text{ and } F_{m-1} \cdot F_m \cdot \Delta > 1 + \frac{1}{\varepsilon} ,$$

where

(10) $$\varepsilon = (1 + \frac{1}{n})^{\frac{1}{n-1}} - 1 ,$$

and $\Delta > 0$ is the smallest distance between any two roots of the given equation (of degree $n > 1$).

It is known from the analysis of complex functions, [7] pp. 127-153, that the transformation $x = a + \frac{1}{y}$ is equivalent to a general translation, $x = a + y_1$, followed by an inversion, $y_1 = \frac{1}{y}$. As a result of the Vincent-Uspensky-Akritas theorem the number of general translations as well as that of inversions that have to be performed for each positive real root in Theorem 2, are bounded by m, which in turn is [1]

(11) $$m \leq n \cdot L(|P|_\infty) .$$

Notice, moreover, that the general translation $x = a + y$, is equivalent to a unit translations, $x = 1 + y_i$.

We have already seen that Descartes' rule of signs is a rather weak proposition and, applied to a polynomial equation, does not give the exact number of positive (or negative) real roots except when the number of sign variations is zero or one. But exactly these particular cases, when combined with Vincent's theorem, supply two real root isolation methods, corresponding to the two different ways in which the positive, integral quantities $a_1, a_2, \ldots$ can be computed.

The first method is due to Vincent himself (1836) and will be briefly described here (see Uspensky's book [8] pp. 127-137 for more details, and certain improvements to Vincent's original presentation). Each a_i is computed separately; it is first initialized to zero and then, by always performing the pair of substitutions $x \leftarrow 1 + x$ and $x \leftarrow \frac{1}{1+x}$, a decision is made, with the help of Descartes' rule of signs, as to whether a_i has already been computed, or has to be incremented by one, $a_i \leftarrow a_i + 1$. For a particular root, a_i has to be determined before the computation of a_{i+1} can be initiated.

Notice that the quantities $a_1, a_2, \ldots$ are all of finite magnitude; despite this fact though, since their incrementation is only by one unit, we may have a prohibitively large number of executions of $x \leftarrow 1 + x$ and $x \leftarrow \frac{1}{1+x}$ in order to compute a specific a_i. This intuitively proves that Vincent's algorithm has an exponential computing time. An example makes the point clearer.

Example. Consider the polynomial equation $P(x) = (x-\alpha)\cdot(x-\beta) = 0$, where $\alpha = 5\cdot10^9+\varepsilon$, $\varepsilon>0$, and $\beta = \alpha + 1$. Following Uspensky, [8] p. 132, we see that the computations in Vincent's algorithm can be represented in the form of a binary tree, where a polynomial and a transformation are associated with each node. Each path from a node to a successor corresponds to a substitution $x \leftarrow 1 + x$ or $x \leftarrow \frac{1}{1+x}$, which is applied to both the polynomial and the transformation. Polynomials at terminal nodes have at most one sign variation; the associated transformations of those with one sign variation, when applied to $(0,\infty)$, produce isolating intervals.

In our case, we will have the following tree:

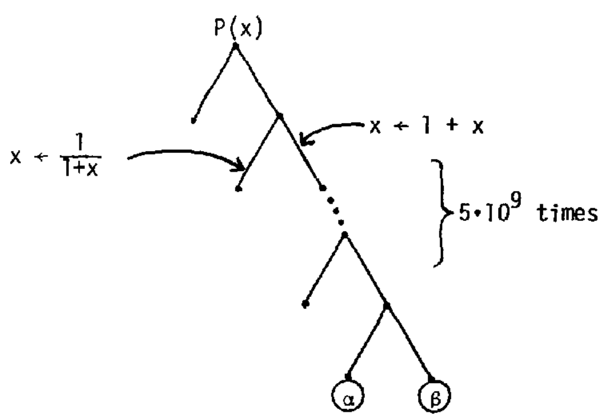

Figure 1

Clearly, in our example, there is only a_1 to be computed, which we know is equal to $5\cdot10^9$. If we assume that the average time for each substitution is 0.04 seconds, then Vincent's algorithm will terminate in 12 years!!

3. Akritas' Method

Our approach is quite different. Instead of calculating the a_i's by "brute force", we first "interpret" them and then compute their magnitude in an easy manner. It is therefore clear, that for each transformation $x = a + \frac{1}{y}$, only one general translation $(x \leftarrow a + x)$ will be performed in place of a unit ones. Moreover, we know from the previous section that, for each positive real root, the number of general translations and that of inversions are bounded by m. Taking, thus, inequality (11) into consideration, we "feel" that our method has a polynomial computing time.

The "interpretation" of the a_i's is derived by first stating our objective. In case of more than one positive real root, what we are trying to achieve with the transformations of the form $x = a + \frac{1}{y}$ (of Vincent's theorem), is to either have only one inside the interval (0,1) and all others in $(1,\infty)$, or to have one real root in $(1,\infty)$ and all others in (0,1). In the first case the substitution $x \leftarrow \frac{1}{1+x}$ will result in a polynomial with only one root in $(0,\infty)$, whereas, in the second case the same is attained with the substitution $x \leftarrow 1 + x$. The reader can easily see that our objective is easily met when the roots are "well separated" and close to the origin.

Lemma (Akritas). Let P(x) be an integral, univariate polynomial of degree $n \geq 2$, which has (at least) two positive real roots, α and β, where $0 < \alpha < \beta < 1$; set $\varepsilon = \beta - \alpha$. Then the substitution $x \leftarrow \frac{1}{1+x}$ will map α and β in $(0,\infty)$, where they will be a distance d apart, with $d > \varepsilon$.

This lemma indicates to us that the following has to be done for the realization of our objective: Whenever there is more than one real root simultaneously inside the interval (0,1), they will first get inverted in order to spread apart, and then they will be "moved" back to the origin, with the expectation that they will enter the interval (0,1) one at a time. (Recall that the general substitution $x \leftarrow a + x$, $a \geq 1$, when applied on a univariate polynomial P, will shift its positive roots to the left by a units.)

We view, therefore, each of the integral quantities $a_1, a_2, \ldots$ as the positive lower root bound (p.l.r.b.) of a polynomial, and we can easily compute their values. (Actually, depending on the method used and the polynomial, there will be more than one p.l.r.b. computation in order to determine a particular a_i; however, for theoretical purposes we will assume a_i = p.l.r.b.(P_i).)

We have implemented Cauchy's rule in order to compute the p.l.r.b. of a polynomial, whereas for the execution of the translations $x \leftarrow a + x$, $a \geq 1$, we used the Ruffini-Horner method. It is proven [1] that the former is

(12) $$O(n^2 L(|P|_\infty)),$$

whereas, the latter is

(13) $$O(n^3 L(a)^2 + n^2 L(a) L(|P|_\infty)) .$$

Our method then is implemented as follows (for the positive real roots): If the given polynomial has more than 1 sign variation, we enter a loop which examines a list of triplets

$((P_1,N_1,D_1), (P_2,N_2,D_2), \ldots (P_k,N_k,D_k))$, where, for all i, P_i is a polynomial with more than one sign variation, and N_i, D_i are the numerators and denominators, respectively, of the accompanying transformation. The main loop then, removes the first triplet and computes the p.l.r.b. of P_1. If the p.l.r.b. ≥ 1 only one new triplet is produced, corresponding to the general translation $x \leftarrow$ p.l.r.b. + x. (A test is made to see whether the p.l.r.b. itself is a root.) If the p.l.r.b. < 1 then two new triplets are produced corresponding to $x \leftarrow 1 + x$, and $x \leftarrow \frac{1}{1+x}$. The new triplets with more than one sign variation are returned to the head of the triplet list; the triplets with one sign variation contribute an isolating interval to the output list, by applying the accompanying transformation to $(0,\infty)$; finally, triplets with no sign variation are disregarded. The loop is repeated until the triplet list is empty.

Taking into account (11), (12), (13) and the fact that $L(a_i)$ and $L(|P|_\infty$ are deterministically related for all i, we see that for the worst case we have (given an integral, univariate, square-free polynomial P of degree n > 0)

Theorem 3. Akritas' method for polynomial real root isolation is $O(n^5 L(|P|_\infty)^3)$.

This is the best theoretical computing time achieved thus far.

4. Empirical Results and Conclusions

In the following we present three tables showing the observed computing times (in seconds) for three different classes of polynomials, for the methods of Sturm, Vincent and Akritas. The computer used was the IBM 370, Model 137 of North Carolina State University. The reader should bear in mind that these times are possibly 50% or more inflated due to the fact that we used SAC-1 (the computer algebraic system) written entirely in FORTRAN, having thus a fictitious word length of only 15 bits!!

Table 1

Chebyshev Polynomials

Degree	Sturm	Vincent	Akritas
5	.63	.16	.19
10	3.80	1.79	1.89
15	9.76	7.64	6.12
20	22.84	34.57	25.84
25	45.48	85.20	48.94

It is a known fact that Sturm's method is extremely fast in the case of Chebyshev's polynomials, but we observe that it is only slightly better than ours. Notice also that Vincent's method becomes exponential.

Table 2

Random Products

Degree	Sturm	Vincent	Akritas
5	2.86	.38	.50
10	36.97	2.15	2.70
15	173.28	17.50	15.10
20	*	31.47	26.94

*case not run

A random product is a polynomial of degree n, resulting from the multiplication of n linear random polynomials, each having coefficients of specified length. The coefficients of the resulting product are truncated, so that their maximum does not exceed the specified length of 10 decimal digits. Our present method is the best in this case.

Table 3

Random Polynomials

Degree	Sturm	Vincent	Akritas
5	2.05	.16	.26
10	33.28	.39	.45
15	156.40	.76	.92
20	*	1.97	2.30

*case not run

The coefficients of the random polynomials were all 10 decimal digits long. Notice that in these examples Vincent's method is slightly better than ours.

The Example mentioned in Section 2, for which Vincent's procedure would need 12 years, was solved by both Sturm's and ours in 1.4 seconds.

In general, we observe that our method is a very close second when the others are performing optimally, whereas, it is by far the best in the other cases.

References

[1] Akritas, A. G., "Vincent's Theorem in Algebraic Manipulation", Ph.D. Thesis (in preparation), North Carolina State University, 1978.

[2] Akritas, A. G., "A Correction on a Theorem by Uspensky", submitted to the Journal of the Greek Mathematical Society, 1978.

[3] Akritas, A. G. and Danielopoulos, S. D., "On the Forgotten Theorem of Mr. Vincent", Historia Mathematica, to appear.

[4] Collins, G. E. and Akritas, A. G., "Polynomial Real Root Isolation Using Descartes' Rule of Signs", Proceedings of the 1976 ACM Symposium on Symbolic and Algebraic Computation, Yorktown Heights, New York, pp. 272-275.

[5] Collins, G. E. and Loos, R., "Polynomial Real Root Isolation by Differentiation", Proceedings of the 1976 ACM Symposium on Symbolic and Algebraic Computation, Yorktown Heights, New York, pp. 15-20.

[6] Heindel, L. E., "Integer Arithmetic Algorithms for Polynomial Real Zero Determination", J.A.C.M., Vol. 18, No. 4, Oct. 1971, pp. 533-548.

[7] Kappos, D. A., Theory of Complex Functions, Athens, 1963 (in Greek).

[8] Uspensky, J. V., Theory of Equations, McGraw-Hill, New York, 1948.

[9] Vincent, "Sur la résolulion des équations numériques", Journal de Mathématiques Pures et Appliquées, Vol. 1, 1836, pp. 341-372.