Estimating roots of polynomials using perturbation theory

M. Pakdemirli a,*, H.A. Yurtsever b

a Department of Mechanical Engineering, Celal Bayar University, 45140 Muradiye, Manisa, Turkey
b Department of Mathematics, Georgetown University, Washington DC, USA

Abstract

Perturbation theory and the order of magnitude of terms are employed to develop two theorems. The theorems may be useful to estimate the order of magnitude of the roots of a polynomial a priori before solving the equation. The theorems are developed for two special types of polynomials of arbitrary order with their coefficients satisfying certain conditions. Numerical applications of the theorems are presented as examples.

Keywords: Roots of polynomials; Perturbation theory

1. Introduction

Finding roots of polynomials is a problem of interest both in mathematics and in application areas such as physical systems. For example, finding natural frequencies of a vibrating system may reduce to a polynomial equation which has to be solved for its roots. For cubic and higher order polynomials, numerical techniques are used to find the roots. Many different algorithms (Newton-Raphson, Mullers, Secant, Householders’ Iteration etc.) were already developed for this issue which was discussed in detail in numerical analysis books (see [1] and [2] for example). The algorithms are mainly iterative in nature and the convergence to a root requires a good initial estimate.

Two theorems are presented and proven here to guide the researcher for a good initial estimate. The theorems are developed based on the order of magnitude concept of the perturbation theory. For details of perturbation theory and applications, the beginner reader is referred to Nayfeh [3] and Hinch [4]. The link between perturbation theory and root finding algorithms were exploited in a recent paper by Pakdemirli and Boyaci [5]. Since the focus in this paper is on the initial estimate rather than the root itself, interested readers may refer to [5] for the relation between the perturbation theory and various root finding algorithms. After proving the theorems, numerical applications of the theorems are given for both cases.

* Corresponding author.
E-mail address: mpak@bayar.edu.tr (M. Pakdemirli).

0096-3003/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.amc.2006.11.071
2. Estimation of the roots

In perturbation theory, the magnitude of terms are ordered with respect to a small parameter usually expressed as $\varepsilon$, $\varepsilon$ being a much smaller quantity than 1 ($\varepsilon \ll 1$). Therefore a term of order $1/\varepsilon$, denoted by $O(1/\varepsilon)$ is much bigger than 1. Depending on the magnitude of coefficients of polynomials, two classes of polynomial equations will be defined and theorems for estimating the magnitude of roots will be given with their proofs.

2.1. Polynomial with all coefficients the same order of magnitude

In this subsection, one chooses a polynomial equation in which all the coefficients have the same order of magnitudes. The magnitudes of roots follow from the below theorem:

**Theorem 1.** For the polynomial equation

$$a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \cdots + a_1 x + a_0 = 0, \tag{1}$$

if all coefficients $a_i$ ($i = 0 \ldots n$) are of the same order of magnitude, then the root is of $O(1)$.

**Proof.** Divide the equation by $a_n$.

$$x^n + \bar{a}_{n-1} x^{n-1} + \bar{a}_{n-2} x^{n-2} + \cdots + \bar{a}_1 x + \bar{a}_0 = 0. \tag{2}$$

In this case, all $\bar{a}_i$ will be of order 1, i.e. $\bar{a}_i \sim O(1)$. There may be three distinct cases for the roots.

(i) $|x| \gg 1$, (ii) $|x| \ll 1$, (iii) $|x| \sim O(1)$. If one shows that the first two cases are impossible, then the third case would be true.

(i) $|x| \gg 1$ case

Divide Eq. (2) by $x^n$. The equation becomes

$$1 + \bar{a}_{n-1} \frac{1}{x} + \bar{a}_{n-2} \frac{1}{x^2} + \cdots + \bar{a}_1 \frac{1}{x^{n-1}} + \bar{a}_0 \frac{1}{x^n} = 0.$$

Since $|x| \gg 1$, then $1/x$ should be much smaller than 1, define $\frac{1}{x} = \varepsilon \ll 1$. Then the equation is

$$1 + \bar{a}_{n-1} \varepsilon + \bar{a}_{n-2} \varepsilon^2 + \cdots + \bar{a}_1 \varepsilon^{n-1} + \bar{a}_0 \varepsilon^n = 0.$$

Each term is of different magnitude, becoming gradually smaller to the right. The terms cannot be balanced with each other yielding zero on the right hand side. So this case should be discarded.

(ii) $|x| \ll 1$ case

Since $x$ is small, define $x = \varepsilon$. Eq. (2) now reads

$$\varepsilon^n + \bar{a}_{n-1} \varepsilon^{n-1} + \bar{a}_{n-2} \varepsilon^{n-2} + \cdots + \bar{a}_1 \varepsilon + \bar{a}_0 = 0.$$

Each term is of different magnitude, the leading terms being smaller. The terms cannot be balanced with each other yielding zero on the right hand side. So this case should also be discarded.

Since the first two cases are impossible, the third case immediately follows, i.e. $|x| \sim O(1)$. \qed

2.2. Polynomial with one relatively large coefficient

In this subsection, one chooses a polynomial equation in which one coefficient is substantially larger than the others with all the remaining coefficients being of order one. The theorem for estimating the roots is posed below:

**Theorem 2.** For the polynomial equation

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_m x^m + \cdots + a_1 x + a_0 = 0, \tag{3}$$

if $a_m \sim O(\frac{1}{\varepsilon^k})$ ($k > 0$) with all other coefficients being of $O(1)$, then the possible roots may be of either $O(\varepsilon^{k/m})$ ($m \neq 0$ case) or $O(\frac{1}{\varepsilon^{k/(k-m)}})$ ($m \neq n$ case).
Proof. Three distinct cases have to be examined separately:

(i) $|x| \sim O(1)$, (ii) $|x| \sim O(\varepsilon^p)$, $p > 0$, (iii) $|x| \sim O(1/\varepsilon^p)$, $p > 0$.

(i) $|x| \sim O(1)$.

If $|x|$ is of order one, then each term in Eq. (3) can be written in their orders of magnitudes

$$O(1) + O(1) + \cdots + O(1/\varepsilon^k) + \cdots + O(1) + O(1) = 0.$$ 

The middle term is much larger than the others and a balance of terms is impossible. Therefore the roots can not be of order one.

(ii) $|x| \sim O(\varepsilon^p)$, $p > 0$.

If one substitutes for magnitude of the root as $O(\varepsilon^p)$ with $p$ a positive real number to Eq. (3), the order of terms can be written as

$$O(\varepsilon^p) + O(\varepsilon^{p-n-1}) + \cdots + O(\varepsilon^{m-k}) + \cdots + O(\varepsilon^p) + O(1) = 0.$$ 

The only possible balancing can be done with the dominating middle term and the last term which yields $pm - k = 0$ or $p = k/m$. Therefore some of the roots are expected to be of $O(\varepsilon^{km})$. For this case, if $m = 0$, that is the largest term in the polynomial is the last term, than this order of magnitude does not apply.

(iii) $|x| \sim O(1/\varepsilon^p)$, $p > 0$.

For this case, the orders of magnitudes of all terms read

$$O(1/\varepsilon^p) + O(1/\varepsilon^{p-n}) + \cdots + O(1/\varepsilon^{m+k}) + \cdots + O(1/\varepsilon^p) + O(1) = 0.$$ 

The dominating largest terms have to be balanced, or equating the magnitudes of the first and middle term yields $np = pm + k$ or $p = k/(n-m)$. Some roots then have to be of $O(1/\varepsilon^{k(n-m)})$. For this case, if $m = n$, that is the largest term in the polynomial is the leading term, than this order of magnitude does not apply. □

3. Numerical examples

In this section numerical examples for the theorems will be given. The small parameter $\varepsilon$ is assumed to be 0.1. Therefore a term of order $1/\varepsilon$ is expected to be near 10 and that of order $1/\varepsilon^2$ to be near 100. There are intermediate orderings also such as $\sqrt{\varepsilon} \approx 0.3162$ or $1/\sqrt{\varepsilon} \approx 3.162$. Note that the theorems are universal and independent of the selection of the parameter $\varepsilon$ as long as this parameter is much smaller than 1.

3.1. Polynomial with all coefficients the same order of magnitude

Example cubic equations are given in Table 1. Since all coefficients are of order 1 with some values in the far limits of order one, the roots are of order 1 as expected. Note that the magnitudes of complex roots has to be considered and it is seen that for complex roots the magnitudes are still of order 1. For various fourth order

<table>
<thead>
<tr>
<th>$a_2$</th>
<th>$a_1$</th>
<th>$a_0$</th>
<th>Roots</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-2</td>
<td>1.2056, -1.1028 ± 0.6655i</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>-2.2470, 0.8019, -0.5550</td>
</tr>
<tr>
<td>0.5</td>
<td>-1</td>
<td>-2</td>
<td>1.3463, -0.9231 ± 0.7959i</td>
</tr>
<tr>
<td>0.5</td>
<td>-1</td>
<td>2</td>
<td>-1.7378, 0.6189 ± 0.8763i</td>
</tr>
<tr>
<td>0.5</td>
<td>-1.1</td>
<td>0.7</td>
<td>-1.5236, 0.5118 ± 0.4444i</td>
</tr>
<tr>
<td>0.5</td>
<td>-1.1</td>
<td>0.3</td>
<td>-1.4220, 0.5000, 0.4220</td>
</tr>
<tr>
<td>0.8</td>
<td>1.4</td>
<td>0.6</td>
<td>-0.4813, -0.1593 ± 1.1051i</td>
</tr>
<tr>
<td>-0.8</td>
<td>-1.4</td>
<td>-0.6</td>
<td>1.7775, -0.4888 ± 0.3141i</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1.8393, -0.4196 ± 0.6063i</td>
</tr>
<tr>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>2.3901, -0.4451 ± 0.6554i</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1.0000, ±i</td>
</tr>
</tbody>
</table>
polynomials, examples for Theorem 1 are given in Table 2. Similar conclusions can be drawn from the table. It should be noted that, our estimates are rough initial estimates for a search of roots.

3.2. Polynomial with one relatively large coefficient

Several examples will be given for the applications of Theorem 2.

**Example 1.** Consider the equation \(x^3 + 0.5x^2 - x + 1200 = 0\). Here \(a_0 = 1200 \sim O(1/e^3)\). Hence \(k = 3, m = 0\) and \(n = 3\). Since \(m = 0\), one does not expect a root of \(O(e^{km})\) and the only magnitude of root expected is of \(O(1/e)\) or inserting \(k, n\) and \(m\), the expected magnitude is \(O(1/e)\) or \(|x| \approx 10\). The numerically calculated roots are \(-10.8278, 5.1639 \pm 9.1739i\) which are all approximately 10 in magnitude.

**Example 2.** Consider the equation \(x^3 + 12x^2 + x - 1.5 = 0\). Here \(a_2 = 12 \sim O(1/e)\). Hence \(k = 1, m = 2\) and \(n = 3\). Two different magnitudes of roots are expected to be of order \(O(e^{km})\) or \(O(1/e)\). Inserting \(k, n\) and \(m\), the expected magnitudes are \(O(\sqrt{e})\) or \(O(1/e)\). Estimated roots are \(|x| \approx 0.3162\) and \(|x| \approx 10\). The numerically calculated roots are \(0.3108, -0.4054\) and \(-11.9054\). The first two belong to \(O(\sqrt{e})\) class and the last to \(O(1/e)\) class.

**Example 3.** Consider the equation \(x^4 + 1100x^3 - x^2 + 1.2x - 0.7 = 0\). Here \(a_3 = 1100 \sim O(1/e^3)\). Hence \(k = 3, m = 3\) and \(n = 4\). Two different magnitudes of roots are expected to be of order \(O(e^{km})\) or \(O(1/e)\). Inserting \(k, n\) and \(m\), the expected magnitudes are \(O(e)\) or \(O(1/e)\). Estimated roots are \(|x| \approx 0.1\) or \(|x| \approx 1000\). The numerically calculated roots are \(0.08208, \pm 0.1i\) and \(-1100.00091\). The first three belong to \(O(e)\) and the last belong to \(O(1/e^3)\) class.

**Example 4.** As a final example, consider the equation \(x^5 + 1.1x^4 + 0.8x^3 - 0.6x^2 - 116x - 2 = 0\). Here \(a_1 = 116 \sim O(1/e^2)\). Hence \(k = 2, m = 1\) and \(n = 5\). Two different magnitudes of roots are expected to be of order \(O(e^{km})\) or \(O(1/e)\). Inserting \(k, n\) and \(m\), the expected magnitudes are \(O(e^2)\) or \(O(1/\sqrt{e})\). Estimated roots are \(|x| \approx 0.01\) or \(|x| \approx 3.162\). The numerically calculated roots are \(-3.504, 3.004, -0.2911 \pm 3.3065i\) and \(-0.0172\). The first four belong to \(O(1/\sqrt{e})\) class and the last belong to \(O(e^2)\) class.

**References**