M.Sc. Course in
Applied Mathematics
with Oceanology and Computer Programming
Vidyasagar University

Semester-II

Paper-MTM 202 **Paper Name: Numerical Analysis**

Solution of Non-Linear Equation

Module No. 1

Roots of a Polynomial Equation

Objective

(a) Roots of polynomial equation

(b) Bairstow method

(c) Convergence of the method

Keywords

polynomial equation, Bairstow method, Convergence of the method

Developed by

Professor Madhumangal Pal
Department of Applied Mathematics
Vidyasagar University, Midnapore-721102
email: mmpalvu@gmail.com

. . . .

Derivation of all roots of a polynomial equation is a very important task. In many applications of science and engineering all roots of a polynomial equations are needed to solve a particular problem. For example, to find the poles, singularities, etc. of a function, the zeros of the denominator (polynomial) are needed. The available analytic methods are useful when the degree of the polynomial is at most four. So, numerical methods are required to find the roots of the higher degree polynomial equations.

Fortunately, many direct and iterated numerical methods are developed to find all the roots of a polynomial equation. In this module, we discuss only Bairstow methods.

1.1 Roots of polynomial equations

Let $P_n(x)$ be a polynomial in x of degree n. If $a_0, a_1, \ldots, a_n$ are coefficients of $P_n(x)$, then equation $P_n(x) = 0$ can be written in explicit form as

$$P_n(x) \equiv a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = 0. \tag{1.1}$$

Here, we assumed that the coefficients $a_0, a_1, \ldots, a_n$ are real numbers.

A number ξ (may be real or complex) is a root of the polynomial equation $P_n(x) = 0$ if and only if $P_n(\xi) = 0$. That is, $P_n(x)$ is exactly divisible by $x - \xi$. If $P_n(x)$ is exactly divisible by $(x-\xi)^k$ $(k \geq 1)$, but it is not divisible by $(x-\xi)^{k+1}$, then ξ is called a root of multiplicity k. The roots of multiplicity $k = 1$ are called **simple roots** or **single roots**.

From fundamental theorem of algebra, we know that every polynomial equation has a root. More precisely, every polynomial equation $P_n(x) = 0$, $(n \geq 1)$ with any numerical coefficients has exactly n, real or complex roots.

The roots of any polynomial equation are either real or complex. If the coefficients of the equation are real and it has a complex root $\alpha + i\beta$ of multiplicity k, then $\alpha - i\beta$ must be another complex root of the equation with multiplicity k.

Let

$$a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = 0, \tag{1.2}$$

be a polynomial equation, where $a_0, a_1, \ldots, a_n$ are real coefficients. Also, let $A = \max\{|a_1|, |a_2|, \ldots, |a_n|\}$ and $B = \max\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}$. Then the magnitude of a root of the equation (1.2) lies between $\dfrac{1}{1 + B/|a_n|}$ and $1 + \dfrac{A}{|a_0|}$.

The other methods are also available to find the upper bound of the positive roots of the polynomial equation. Two such results are stated below:

Theorem 1.1 (Lagrange's). *If the coefficients of the polynomial*

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$$

satisfy the conditions $a_0 > 0, a_1, a_2, \ldots, a_{m-1} \geq 0, a_m < 0$, for some $m \leq n$, then the upper bound of the positive roots of the equation is $1 + \sqrt[m]{B/a_0}$, where B is the greatest of the absolute values of the negative coefficients of the polynomial.

Theorem 1.2 (Newton's). *If for $x = c$ the polynomial*

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$

and its derivatives $f'(x), f''(x), \ldots$ assume positive values then c is the upper bound of the positive roots of the equation $f(x) = 0$.

In the following section, Bairstow method is discussed to find all the roots of a polynomial equation of degree n.

1.2 Bairstow method

This method is also an iterative method. In this method, a quadratic factor is extracted from the polynomial $P_n(x)$ by iteration. As a by product the deflated polynomial (the polynomial obtained by dividing $P_n(x)$ by the quadratic factor) is also obtained. It is well known that the determination of roots (real or complex) of a quadratic equation is easy. Therefore, by extracting all quadratic factors one can determine all the roots of a polynomial equation. This is the basic principle of Bairstow method.

Let the polynomial $P_n(x)$ of degree n be

$$x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n. \tag{1.3}$$

Let $x^2 + px + q$ be a factor of the polynomial $P_n(x)$, $n > 2$. When this polynomial is divided by the factor $x^2 + px + q$, then the quotient is a polynomial of degree $(n-2)$ and remainder is a linear polynomial. Let the quotient and the remainder be denoted by $Q_{n-2}(x)$ and $Mx + N$, where M and N are two constants.

. . . .

Using this notation, $P_n(x)$ can be written as

$$P_n(x) = (x^2 + px + q)Q_{n-2}(x) + Mx + N. \tag{1.4}$$

The polynomial $Q_{n-2}(x)$ is called deflated polynomial and let it be

$$Q_{n-2}(x) = x^{n-2} + b_1x^{n-3} + \cdots + b_{n-3}x + b_{n-2}. \tag{1.5}$$

It is obvious that the values of M and N depends on p and q. If $x^2 + px + q$ is an exact factor of $P_n(x)$, then the remainder $Mx + N$, i.e. M and N must be zero. Thus the main aim of Bairstow method is to find the values of p and q such that

$$M(p, q) = 0 \qquad \text{and} \qquad N(p, q) = 0. \tag{1.6}$$

These are two non-linear equations in p and q and these equations can be solved by Newton-Raphson method for two variables (discussed in Module 3 of this chapter).

Let (p_T, q_T) be the exact values of p and q and Δp, Δq be the (errors) corrections to p and q. Therefore,

$$p_T = p + \Delta p \qquad \text{and} \qquad q_T = q + \Delta q.$$

Hence,

$$M(p_T, q_T) = M(p + \Delta p, q + \Delta q) = 0 \qquad \text{and} \qquad N(p_T, q_T) = N(p + \Delta p, q + \Delta q) = 0.$$

By Taylor's series expansion, we get

$$M(p + \Delta p, q + \Delta q) = M(p, q) + \Delta p\frac{\partial M}{\partial p} + \Delta q\frac{\partial M}{\partial q} + \cdots = 0$$
$$\text{and } N(p + \Delta p, q + \Delta q) = N(p, q) + \Delta p\frac{\partial N}{\partial p} + \Delta q\frac{\partial N}{\partial q} + \cdots = 0.$$

All the derivatives are evaluated at the approximate value (p, q) of (p_T, q_T). Neglecting square and higher powers of Δp and Δq, as they are small, the above equations become

$$\Delta pM_p + \Delta qM_q = -M \tag{1.7}$$

$$\Delta pN_p + \Delta qN_q = -N. \tag{1.8}$$

Therefore, the values of Δp and Δq are obtained by the formulae

$$\Delta p = -\frac{MN_q - NM_q}{M_pN_q - M_qN_p}, \quad \Delta q = -\frac{NM_p - MN_p}{M_pN_q - M_qN_p}. \tag{1.9}$$

It is expected that in this stage the values of Δp and Δq are either 0 or very small.

Now, the coefficients of the deflated polynomial $Q_{n-2}(x)$ and the expressions for M and N in terms of p and q are computed below.

From equation (1.4)

$$\begin{aligned} &x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n \\ &\quad = (x^2 + px + q)(x^{n-2} + b_1x^{n-3} + \cdots + b_{n-3}x + b_{n-2}) + Mx + N. \end{aligned} \tag{1.10}$$

Comparing both sides, we get

$$\begin{array}{ll} a_1 = b_1 + p & b_1 = a_1 - p \\ a_2 = b_2 + pb_1 + q & b_2 = a_2 - pb_1 - q \\ \vdots & \vdots \\ a_k = b_k + pb_{k-1} + qb_{k-2} & b_k = a_k - pb_{k-1} - qb_{k-2} \\ \vdots & \vdots \\ a_{n-1} = M + pb_{n-2} + qb_{n-3} & M = a_{n-1} - pb_{n-2} - qb_{n-3} \\ a_n = N + qb_{n-2} & N = a_n - qb_{n-2}. \end{array} \tag{1.11}$$

In general,

$$b_k = a_k - pb_{k-1} - qb_{k-2}, \qquad k = 1, 2, \ldots, n. \tag{1.12}$$

The values of b_0 and b_{-1} are taken as 1 and 0 respectively.

With this notation, the expressions for M and N are

$$M = b_{n-1}, \qquad N = b_n + pb_{n-1}. \tag{1.13}$$

Note that M and N depend on b's. Differentiating the equation (1.12) with respect to p and q to find the partial derivatives of M and N.

$$\frac{\partial b_k}{\partial p} = -b_{k-1} - p\frac{\partial b_{k-1}}{\partial p} - q\frac{\partial b_{k-2}}{\partial p}, \qquad \frac{\partial b_0}{\partial p} = \frac{\partial b_{-1}}{\partial p} = 0 \tag{1.14}$$

$$\frac{\partial b_k}{\partial q} = -b_{k-2} - p\frac{\partial b_{k-1}}{\partial q} - q\frac{\partial b_{k-2}}{\partial q}, \qquad \frac{\partial b_0}{\partial q} = \frac{\partial b_{-1}}{\partial q} = 0 \tag{1.15}$$

For simplification, we denote

$$\frac{\partial b_k}{\partial p} = -c_{k-1}, \qquad k = 1, 2, \ldots, n \tag{1.16}$$

$$\text{and } \frac{\partial b_k}{\partial q} = -c_{k-2}. \tag{1.17}$$

....

With this notation, the equation (1.14) simplifies as

$$c_{k-1} = b_{k-1} - pc_{k-2} - qc_{k-3}. \tag{1.18}$$

Also, the equations (1.15) becomes

$$c_{k-2} = b_{k-2} - pc_{k-3} - qc_{k-4}. \tag{1.19}$$

Hence, the recurrence relation for c_k is

$$c_k = b_k - pc_{k-1} - qc_{k-2}, k = 1, 2, \ldots, n-1 \text{ and } c_0 = 1, c_{-1} = 0. \tag{1.20}$$

Therefore,

$$\begin{aligned}
M_p &= \frac{\partial b_{n-1}}{\partial p} = -c_{n-2} \\
N_p &= \frac{\partial b_n}{\partial p} + p\frac{\partial b_{n-1}}{\partial p} + b_{n-1} = b_{n-1} - c_{n-1} - pc_{n-2} \\
M_q &= \frac{\partial b_{n-1}}{\partial q} = -c_{n-3} \\
N_q &= \frac{\partial b_n}{\partial q} + p\frac{\partial b_{n-1}}{\partial q} = -(c_{n-2} + pc_{n-3}).
\end{aligned}$$

From the equation (1.9), the explicit expressions for Δp and Δq, are obtained as follows:

$$\begin{aligned}
\Delta p &= -\frac{b_n c_{n-3} - b_{n-1} c_{n-2}}{c_{n-2}^2 - c_{n-3}(c_{n-1} - b_{n-1})} \\
\Delta q &= -\frac{b_{n-1}(c_{n-1} - b_{n-1}) - b_n c_{n-2}}{c_{n-2}^2 - c_{n-3}(c_{n-1} - b_{n-1})}.
\end{aligned} \tag{1.21}$$

Therefore, the improved values of p and q are $p + \Delta p$ and $q + \Delta q$. Thus if p_0, q_0 be the initial guesses of p and q, then the first approximate values of p and q are

$$p_1 = p_0 + \Delta p \qquad \text{and} \qquad q_1 = q_0 + \Delta q. \tag{1.22}$$

Table 1.1 is helpful to calculate the values of b_k's and c_k's, where p_0 and q_0 are taken as initial values of p and q.

The second approximate values p_2, q_2 of p and q are determined from the equations: $p_2 = p_1 + \Delta p, \qquad q_2 = q_1 + \Delta q.$

In general,

$$p_{k+1} = p_k + \Delta p, \qquad q_{k+1} = q_k + \Delta q, \tag{1.23}$$

	1	a_1	a_2	$\cdots$	a_k	$\cdots$	a_{n-1}	a_n
$-p_0$		$-p_0$	$-p_0b_1$	$\cdots$	$-p_0b_{k-1}$	$\cdots$	$-p_0b_{n-2}$	$-p_0b_{n-1}$
$-q_0$			$-q_0$	$\cdots$	$-q_0b_{k-2}$	$\cdots$	$-q_0b_{n-3}$	$-q_0b_{n-2}$
	1	b_1	b_2	$\cdots$	b_k	$\cdots$	b_{n-1}	b_n
$-p_0$		$-p_0$	$-p_0c_1$	$\cdots$	$-p_0c_{k-1}$	$\cdots$	$-p_0c_{n-2}$	
$-q_0$			$-q_0$	$\cdots$	$-q_0c_{k-2}$	$\cdots$	$-q_0c_{n-3}$	
	1	c_1	c_2	$\cdots$	c_k	$\cdots$	c_{n-1}	

Table 1.1: Tabular form of b's and c's.

the values of Δp and Δq are calculated at $p = p_k$ and $q = q_k$.

The iteration process to find the values of p and q will be terminated when both $|\Delta p|$ and $|\Delta q|$ are very small.

The next quadratic factor can be obtained by similar process from the deflated polynomial $Q_{n-2}(x)$.

The values of Δp and Δq are obtained by applying Newton-Raphson method for two variables case. Also, the rate of convergence of Newton-Raphson method is quadratic. Hence, the rate of convergence of this method is quadratic.

Example 1.1 *Extract all the quadratic factors from the equation $x^4+2x^3+3x^2+4x+1=0$ by using Bairstow method and hence solve this equation.*

Solution. Let the initial guess of p and q be $p_0 = 0.5$ and $q_0 = 0.5$.

First iteration

	1	2.000000	3.000000	4.000000	1.000000
−0.500000		−0.500000	−0.750000	−0.875000	−1.187500
−0.500000			−0.500000	−0.750000	−0.875000
	1	1.500000	1.750000	2.375000	−1.062500
−0.500000		−0.500000	−0.500000	−0.375000	
−0.500000			−0.500000	−0.500000	
	1	1.000000	0.750000	1.500000	
		$= c_1$	$= c_2$	$= c_3$	

$$\Delta p = -\frac{b_4c_1 - b_3c_2}{c_2^2 - c_1(c_3 - b_3)} = 1.978261, \qquad \Delta q = -\frac{b_3(c_3 - b_3) - b_4c_2}{c_2^2 - c_1(c_3 - b_3)} = 0.891304$$

. . . .

Therefore, $p_1 = p_0 + \Delta p = 2.478261, \quad q_1 = q_0 + \Delta q = 1.391304.$

Second iteration

	1.00000	4.00000	−7.00000	−22.00000	24.00000
−1.38095		−1.38095	−3.61678	11.11025	5.73794
2.57143			2.57143	6.73469	−20.68805
	1.00000	2.61905	−8.04535	−4.15506	9.04989
−1.38095		−1.38095	−1.70975	9.92031	
2.57143			2.57143	3.18367	
	1.00000	1.23810	−7.18367	8.94893	

$\Delta p = 0.52695, \quad \Delta q = -0.29857.$
$p_2 = p_1 + \Delta p = 1.90790, q_2 = q_1 + \Delta q = -2.86999.$

Third iteration

	1	2.000000	3.000000	4.000000	1.000000
−2.478261		−2.478261	1.185256	−6.924140	5.597732
−1.391304			−1.391304	0.665407	−3.887237
	1	−0.478261	2.793951	−2.258734	2.710495
−2.478261		−2.478261	7.327033	−21.634426	
−1.391304			−1.391304	4.113422	
	1	−2.956522	8.729680	−19.779737	

$\Delta p = -0.479568, \quad \Delta q = -0.652031.$
$p_3 = p_2 + \Delta p = 1.998693, q_3 = q_2 + \Delta q = 0.739273.$

Fourth iteration

	1	2.000000	3.000000	4.000000	1.000000
−1.998693		−1.998693	−0.002613	−4.513276	1.027812
−0.739273			−0.739273	−0.000967	−1.669363
	1	0.001307	2.258114	−0.514242	0.358449
−1.998693		−1.998693	3.992159	−11.014794	
−0.739273			−0.739273	1.476613	
	1	−1.997385	5.511000	−10.052423	

$\Delta p = -0.187110, \;\; \Delta q = -0.258799.$

$p_4 = p_3 + \Delta p = 1.811583, q_4 = q_3 + \Delta q = 0.480474.$

Fifth iteration

	1	2.000000	3.000000	4.000000	1.000000
−1.811583		−1.811583	−0.341334	−3.945975	0.066131
−0.480474			−0.480474	−0.090530	−1.046566
	1	0.188417	2.178192	−0.036504	0.019565
−1.811583		−1.811583	2.940498	−8.402511	
−0.480474			−0.480474	0.779889	
	1	−1.623165	4.638216	−7.659126	

$\Delta p = -0.015050, \;\; \Delta q = -0.020515.$

$p_5 = p_4 + \Delta p = 1.796533, q_5 = q_4 + \Delta q = 0.459960.$

Sixth iteration

	1	2.000000	3.000000	4.000000	1.000000
−1.796533		−1.796533	−0.365535	−3.906570	0.000282
−0.459960			−0.459960	−0.093587	−1.000184
	1	0.203467	2.174505	−0.000157	0.000098
−1.796533		−1.796533	2.861996	−8.221908	
−0.459960			−0.459960	0.732746	
	1	−1.593066	4.576541	−7.489319	

$\Delta p = -0.000062, \;\; \Delta q = -0.000081.$

$p_6 = p_5 + \Delta p = 1.796471, q_6 = q_5 + \Delta q = 0.459879.$

Note that, Δp and Δq are correct up to four decimal places. Thus $p = 1.7965, q = 0.4599$ correct up to four decimal places.

Therefore, a quadratic factor is $x^2 + 1.7965x + 0.4599$ and the deflated polynomial is $Q_2(x) = P_4(x)/(x^2 + 1.7965x + 0.4599) = x^2 + 0.2035x + 2.1745.$

Thus, $P_4(x) = (x^2 + 1.7965x + 0.4599)(x^2 + 0.2035x + 2.1745).$

Hence, the roots of the given equation are

$-0.309212, \; -1.487258, \; (-0.1018, 1.4711), \; (-0.1018, -1.4711).$

Self Assessment (MCQ/Short answer questions)

1. The rate of convergence of Bairstow method is
   (a) 1 (b) 2 (c) 3 (d) none of these

2. Bairstow method is used to find
   (a) all roots of a polynomial equation
   (b) only real roots of a polynomial equation
   (c) only complex roots of a polynomial equation
   (d) only one real root of a polynomial equation

3. Let $a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$, where $a_0, a_1, \ldots, a_n$ are real coefficients, be a polynomial equation. Also, let $r = \frac{1}{1+B/|a_n|}$ and $R = 1 + \frac{A}{|a_0|}$, where $A = \max\{|a_1|, |a_2|, \ldots, |a_n|\}$ and $B = \max\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}$. Then
   (a) all roots lie in the interval $(0, r)$
   (b) positive roots lie in the interval $(0, R)$
   (c) positive roots lie in the interval (r, R) and negative roots lie in the interval $(-R, -r)$
   (d) all roots lie in the interval (r, R)

4. Every polynomial equation of degree at least one has a root.
   (a) true (b) false

5. If for $x = c$ the polynomial $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ and its derivatives $f'(x), f''(x), \ldots$ are positive. Then the upper bound of the positive roots of the equation $f(x) = 0$ is $\cdots\cdots$.

6. Let $a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$ be a polynomial equation, where $a_0 > 0, a_1, a_2, \ldots, a_{m-1} \geq 0, a_m < 0$, for some $m \leq n$, and B is the greatest of the absolute values of the negative coefficients. Then the upper bound of the positive roots of the equation is $\cdots\cdots$.

Answer to the questions

1. (b)
2. (a)
3. (c)
4. (a)
5. c
6. $1 + \sqrt[m]{B/a_0}$

Self Assessment (Long answer questions)

1. Find all quadratic factors of the following polynomial equations using Bairstows method.
   (i) $x^4 8x^3 + 39x^2 62x + 50 = 0$
   (ii) $x^3 2x^2 + x 2 = 0$.

2. Solve the following polynomial equations using Bairstows method.
   (i) $x^4 6x^3 + 18x^2 24x + 16 = 0$
   (ii) $x^3 2x + 1 = 0$.

Learn More

1. Danilina, N.I., Dubrovskaya, S.N., and Kvasha, O.P., and Smirnov, G.L. Computational Mathematics. Moscow: Mir Publishers, 1998.

2. Hildebrand, F.B. Introduction of Numerical Analysis. New York: London: McGraw-Hill, 1956.

3. Jain, M.K., Iyengar, S.R.K., and Jain, R.K. Numerical Methods for Scientific and Engineering Computation. New Delhi: New Age International (P) Limited, 1984.

4. Krishnamurthy, E.V., and Sen, S.K. Numerical Algorithms. New Delhi: Affiliated East-West Press Pvt. Ltd., 1986.

5. Mathews, J.H. Numerical Methods for Mathematics, Science, and Engineering, 2nd ed., NJ: Prentice-Hall, Inc., 1992.

6. Pal, M. Numerical Analysis for Scientists and Engineers: Theory and C Programs. New Delhi: Narosa, Oxford:Alpha Sciences, 2007.

7. Volkov, E.A. Numerical Methods. Moscow: Mir Publishers, 1986.