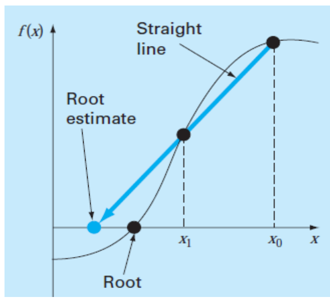
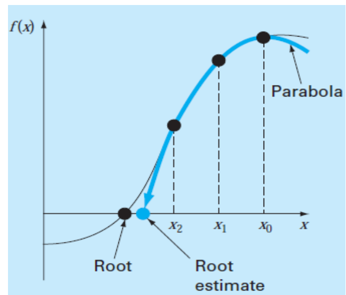
Lecture 07: Roots of Polynomials





Polynomials

$$f_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

- where *n* = the order of the polynomial and the *a's* = constant coefficients.
- *Although the* coefficients can be complex numbers, we will limit our discussion to cases where they are real.
- For such cases, the roots can be real and/or complex.

Polynomials

Å The roots of such polynomials follow these rules:

1. For an *n*th-order equation, there are *n* real or complex roots. It should be noted that these roots will not necessarily be distinct.
2. If *n* is odd, there is at least one real root.
3. If complex roots exist, they exist in conjugate pairs (that is, *λ* + *μi* and *λ* ī *μi*), where *i* = ãī 1.

Polynomials in Engineering and Science

Å Polynomials have many applications in engineering and science.

Å For example, they are used extensively in curve-fitting.

Å One of their most interesting and powerful applications is in characterizing dynamic systems and, in particular, linear systems.

Å Examples include mechanical devices, structures, and electrical circuits.

$$y_1(t) = H\{x_1(t)\} \qquad y_2(t) = H\{x_2(t)\}$$

$$\alpha y_1(t) + \beta y_2(t) = H\{\alpha x_1(t) + \beta x_2(t)\}$$

Polynomials in Engineering and Science

Å Consider a simple second-order system defined by the following linear *ordinary differential equation* (or ODE):

$$a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = F(t)$$

Å This equation can be alternatively expressed as a pair of first-order ODEs by defining a new variable *z*,

$$z = \frac{dy}{dt}$$

$$\frac{dz}{dt} = \frac{F(t) - a_1 z - a_0 y}{a_2}$$

$$\frac{dy}{dt} = z$$

Å In a similar fashion, an *n*th-order linear ODE can always be expressed as a system of *n* first-order ODEs.

Polynomials in Engineering and Science

$$a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = F(t)$$

Å The forcing function represents the effect of the external world on the system.

Å The homogeneous or general solution of the equation deals with the case when the forcing function is set to zero,

$$a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = 0$$

Å Thus, the general solution should tell us something very fundamental about the system being simulatedð that is, how the system responds in the absence of external stimuli.

Polynomials in Engineering and Science

$$a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = 0$$

- Now, the general solution to all unforced linear systems is of the form $y = e^{rt}$.

$$a_2 r^2 e^{rt} + a_1 r e^{rt} + a_0 e^{rt} = 0$$

$$\boxed{a_2 r^2 + a_1 r + a_0 = 0}$$

- Notice that the result is a polynomial called the characteristic equation.
- The roots of this polynomial are the values of *r* that satisfy the above equation.
- These *r*'s are referred to as the system's characteristic values, or eigenvalues.

Polynomials in Engineering and Science

- So, here is the connection between roots of polynomials and engineering and science.
- The eigenvalue tells us something fundamental about the system we are modeling, and finding the eigenvalues involves finding the roots of polynomials.
- Whereas finding the root of a second-order equation is easy with the quadratic formula, finding roots of higher-order systems (and hence, higher-order polynomials) is arduous analytically.
- Thus, the best general approach requires numerical methods of the type described in this lecture.

Polynomials in Engineering and Science

$$a_2 r^2 + a_1 r + a_0 = 0$$

$$\frac{r_1}{r_2} = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2 a_0}}{a_0}$$

- If the discriminant ($a_1^2 - 4a_2a_0$) is positive, the roots are real and the general solution can be represented as

$$y = c_1 e^{r_1 t} + c_2 e^{r_2 t}$$

- where the c's = constants that can be determined from the initial conditions.
- This is called the overdamped case.

Bangladesh University of Eng. & Tech. | Slide 9 of 37 | Industrial &Production Engineering

Polynomials in Engineering and Science

- If the discriminant is zero, a single real root results, and the general solution can be formulated as

$$y = (c_1 + c_2 t)e^{\lambda t}$$

- This is called the critically damped case.
- If the discriminant is negative, the roots will be complex conjugate numbers,

$$\frac{r_1}{r_2} = \lambda \pm \mu i$$

- The general solution can be formulated as

$$y = c_1 e^{(\lambda+\mu i)t} + c_2 e^{(\lambda-\mu i)t}$$

Bangladesh University of Eng. & Tech. | Slide 10 of 37 | Industrial &Production Engineering

Polynomials in Engineering and Science

$$y = c_1 e^{(\lambda + \mu i)t} + c_2 e^{(\lambda - \mu i)t}$$

Å The physical behavior of this solution can be elucidated by using Eulerôs formula

$$e^{\mu i t} = \cos \mu t + i \sin \mu t$$

$$y = c_1 e^{\lambda t} \cos \mu t + c_2 e^{\lambda t} \sin \mu t$$

Å This is called the underdamped case.

Polynomials in Engineering and Science

$$y = c_1 e^{r_1 t} + c_2 e^{r_2 t}$$

$$y = (c_1 + c_2 t) e^{\lambda t}$$

$$y = c_1 e^{\lambda t} \cos \mu t + c_2 e^{\lambda t} \sin \mu t$$

Å These equations express the possible ways that linear systems respond dynamically.

Å The exponential terms mean that the solutions are capable of decaying (negative real part) or growing (positive real part) exponentially with time.

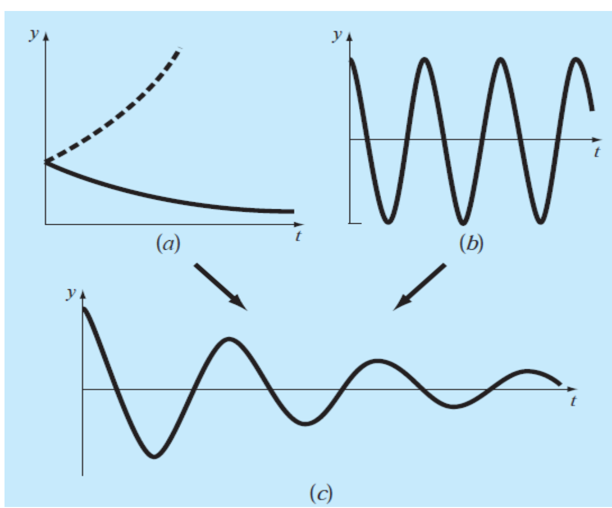
Å The sinusoidal terms (imaginary part) mean that the solutions can oscillate.

Polynomials in Engineering and Science

Å If the eigenvalue has both real and imaginary parts, the exponential and sinusoidal shapes are combined.

Å Because such knowledge is a key element in understanding, designing, and controlling the behavior of a physical system, characteristic polynomials are very important in engineering and many branches of science.

Polynomials in Engineering and Science



Computing with Polynomials

- **Polynomial Evaluation and Differentiation**

$$f_3(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

- It involves six multiplications and three additions.
- In general, for an *n*th-order polynomial, this approach requires *n*(*n* + 1)/2 multiplications and *n* additions.

$$f_3(x) = ((a_3x + a_2)x + a_1)x + a_0$$

- It involves three multiplications and three additions.
- For an *n*th-order polynomial, this approach requires *n* multiplications and *n* additions.

Computing with Polynomials

- Because the nested format minimizes the number of operations, it also tends to minimize round-off errors.
- Note that, depending on your preference, the order of nesting can be reversed:

$$f_3(x) = a_0 + x(a_1 + x(a_2 + xa_3))$$

Polynomial Deflation

- Suppose that you determine a single root of an *n*th-order polynomial.
- If you repeat your root location procedure, you might find the same root.
- Therefore, it would be nice to remove the found root before proceeding.
- This removal process is referred to as polynomial deflation.

Polynomial Deflation

- For example, a fifth-order polynomial could be written as

$$f_5(x) = -120 - 46x + 79x^2 - 3x^3 - 7x^4 + x^5$$

- This fifth-order polynomial might be expressed alternatively as

$$f_5(x) = (x+1)(x-4)(x-5)(x+3)(x-2)$$

- This is called the factored form of the polynomial.

Polynomial Deflation

$$f_5(x) = (x + 1)(x - 4)(x - 5)(x + 3)(x - 2)$$

- Now, suppose that we divide this fifth-order polynomial by any of its factors, for example, $x + 3$.
- For this case the result would be a fourth-order polynomial with a remainder of zero.

$$f_4(x) = (x + 1)(x - 4)(x - 5)(x - 2) = -40 - 2x + 27x^2 - 10x^3 + x^4$$

- Dividing polynomials using the approach called synthetic division.
- Several computer algorithms (based on both synthetic division, as well as other methods) are available for performing the operation.

Polynomial Deflation

- It is also possible to divide by polynomials of higher order.
- The most common task involves dividing by a second-order polynomial or parabola.
- A more general case is dividing an *n*thorder polynomial *a* by an *m*th-order polynomial *d*.

Polynomial Deflation

- Because each calculated root is known only approximately, it should be noted that deflation is sensitive to round-off error.
- In some cases, they can grow to the point that the results can become meaningless.
- Some general strategies can be applied to minimize this problem.

Polynomial Deflation

- For example, round-off error is affected by the order in which the terms are evaluated.
- Forward deflation refers to the case where new polynomial coefficients are in order of descending powers of *x* (that is, from the highest-order to the zero-order term).
- For this case, it is preferable to divide by the roots of smallest absolute value first.
- Conversely, for backward deflation (that is, from the zero-order to the highest-order term), it is preferable to divide by the roots of largest absolute value first.

Polynomial Deflation

- Another way to reduce round-off errors is to consider each successive root estimate obtained during deflation as a good first guess.
- These can then be used as a starting guess, and the root determined again with the original nondeflated polynomial.
- This is referred to as root polishing.
- Finally, a problem arises when two deflated roots are inaccurate enough that they both converge on the same undeflated root.
- In that case, you might be erroneously led to believe that the polynomial has a multiple root.
- One way to detect this problem is to compare each polished root with those that were located previously.

Conventional Methods

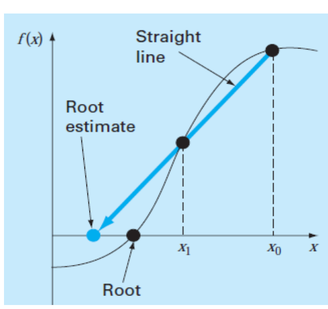
- If only real roots exist, any of the previously described methods could have utility.
- However, the problem of finding good initial guesses complicates both the bracketing and the open methods, whereas the open methods could be susceptible to divergence.
- When complex roots are possible, the bracketing methods cannot be used because of the obvious problem that the criterion for defining a bracket (that is, sign change) does not translate to complex guesses.

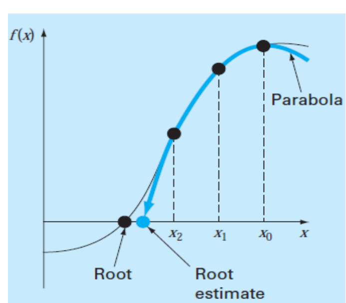
Conventional Methods

- Of the open methods, the conventional Newton-Raphson method would provide a viable approach.
- In particular, concise code including deflation can be developed.
- If a language that accommodates complex variables is used, such an algorithm will locate both real and complex roots.
- However, as might be expected, it would be susceptible to convergence problems.
- For this reason, special methods have been developed to find the real and complex roots of polynomials.
- Two such methods are: Müller and Bairstow methods.

Müller's Method

- Recall that the secant method obtains a root estimate by projecting a straight line to the x axis through two function values.
- Müllerôs method takes a similar approach, but projects a parabola through three points.





Müller's Method

Å The method consists of deriving the coefficients of the parabola that goes through the three points.

Å These coefficients can then be substituted into the quadratic formula to obtain the point where the parabola intercepts the *x* axisð that is, the root estimate.

Å The approach is facilitated by writing the parabolic equation in a convenient form,

$$f_2(x) = a(x - x_2)^2 + b(x - x_2) + c$$

Å We want this parabola to intersect the three points [x_0, f(x_0)], [x_1, $f(x_1)$], and [x_2, $f(x_2)$].

Müller's Method

$$f(x_0) = a(x_0 - x_2)^2 + b(x_0 - x_2) + c$$
$$f(x_1) = a(x_1 - x_2)^2 + b(x_1 - x_2) + c$$
$$f(x_2) = a(x_2 - x_2)^2 + b(x_2 - x_2) + c$$

Å Because two of the terms in the last equation are zero, it can be immediately solved for $c = f(x_2)$.

Å Thus, the coefficient *c* is merely equal to the function value evaluated at the third guess, x_2.

$$f(x_0) - f(x_2) = a(x_0 - x_2)^2 + b(x_0 - x_2)$$
$$f(x_1) - f(x_2) = a(x_1 - x_2)^2 + b(x_1 - x_2)$$

Müller's Method

$$f(x_0) - f(x_2) = a(x_0 - x_2)^2 + b(x_0 - x_2)$$
$$f(x_1) - f(x_2) = a(x_1 - x_2)^2 + b(x_1 - x_2)$$

Å One way to solve this involves defining a number of differences,

$$h_0 = x_1 - x_0 \qquad h_1 = x_2 - x_1$$
$$\delta_0 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} \qquad \delta_1 = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

$$(h_0 + h_1)b - (h_0 + h_1)^2 a = h_0\delta_0 + h_1\delta_1$$
$$h_1 \quad b - \quad h_1^2 \quad a = \quad h_1\delta_1$$

$$a = \frac{\delta_1 - \delta_0}{h_1 + h_0}$$
$$b = ah_1 + \delta_1$$
$$c = f(x_2)$$

Müller's Method

$$f_2(x) = a(x - x_2)^2 + b(x - x_2) + c$$

Å To find the root, we apply the quadratic formula to the above equation.

Å However, because of potential round-off error, rather than using the conventional form, we use the alternative formulation to yield

$$x_3 - x_2 = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}$$

$$x_3 = x_2 + \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}$$

Å Note that the use of the quadratic formula means that both real and complex roots can be located.

Å This is a major benefit of the method.

Müller's Method

- The error can be calculated as

$$\varepsilon_a = \left| \frac{x_3 - x_2}{x_3} \right| 100\%$$

Müller's Method

$$x_3 - x_2 = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}$$

$$x_3 = x_2 + \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}$$

- Now, a problem is that the equation yields two roots, corresponding to the ± term in the denominator.
- In Müller's method, the sign is chosen to agree with the sign of *b*.
- This choice will result in the largest denominator, and hence, will give the root estimate that is closest to x_2.

Müller's Method

- Once x_3 is determined, the process is repeated.
- This brings up the issue of which point is discarded.
- Two general strategies are typically used:
  1. If only real roots are being located, we choose the two original points that are nearest the new root estimate, x_3.
  2. If both real and complex roots are being evaluated, a sequential approach is employed. That is, just like the secant method, x_1, x_2, and x_3 take the place of x_0, x_1, and x_2.

Example: Müller's Method

Problem Statement. Use Müller's method with guesses of x_0, x_1, and $x_2 = 4.5$, 5.5, and 5, respectively, to determine a root of the equation

$$f(x) = x^3 - 13x - 12$$

Note that the roots of this equation are −3, −1, and 4.

Solution. First, we evaluate the function at the guesses

$$f(4.5) = 20.625 \qquad f(5.5) = 82.875 \qquad f(5) = 48$$

which can be used to calculate

$$h_0 = 5.5 - 4.5 = 1 \qquad h_1 = 5 - 5.5 = -0.5$$

$$\delta_0 = \frac{82.875 - 20.625}{5.5 - 4.5} = 62.25 \qquad \delta_1 = \frac{48 - 82.875}{5 - 5.5} = 69.75$$

Example: Müller's Method

$$a = \frac{69.75 - 62.25}{-0.5 + 1} = 15 \qquad b = 15(-0.5) + 69.75 = 62.25 \qquad c = 48$$

The square root of the discriminant can be evaluated as

$$\sqrt{62.25^2 - 4(15)48} = 31.54461$$

Then, because $|62.25 + 31.54451| > |62.25 - 31.54451|$, a positive sign is employed in the denominator of Eq. (7.27*b*), and the new root estimate can be determined as

$$x_3 = 5 + \frac{-2(48)}{62.25 + 31.54451} = 3.976487$$

Example: Müller's Method

$$\varepsilon_a = \left|\frac{-1.023513}{3.976487}\right| 100\% = 25.74\%$$

Because the error is large, new guesses are assigned; x_0 is replaced by x_1, x_1 is replaced by x_2, and x_2 is replaced by x_3. Therefore, for the new iteration,

$$x_0 = 5.5 \qquad x_1 = 5 \qquad x_2 = 3.976487$$

i	x_r	ε_a (%)
0	5	
1	3.976487	25.74
2	4.00105	0.6139
3	4	0.0262
4	4	0.0000119

Assignment-07

- Problems 7.1, 7.3, 7.4, 7.11.