NUMERICAL COMPUTING

Introduction

Numerical methods are mathematical techniques used for solving mathematical problems that cannot be solved or are difficult to solve analytically.

An analytical solution is an exact answer in the form of a mathematical expression in terms of the variables associated with the problem that is being solved. A numerical solution is an approximate numerical value (a number) for the solution. Although numerical solutions are an approximation, they can be very accurate. In many numerical methods, the calculations are executed in an iterative manner until a desired accuracy is achieved.

Errors: Absolute and Relative

Suppose that α and β are two numbers, of which one is regarded as an approximation to the other. The **error** of β as an approximation to α is $\alpha - \beta$; that is, the error equals the exact value minus the approximate value.

The **absolute error** of β as an approximation to α is $|\alpha - \beta|$.

The **relative error** of β as an approximation to α is $|\alpha - \beta| / |\alpha|$.

In summary, we have

$$\text{error} = \text{exact value} - \text{approximate value}$$

$$\text{absolute error} = |\text{exact value} - \text{approximate value}|$$

$$\text{relative error} = \frac{|\text{exact value} - \text{approximate value}|}{|\text{exact value}|}$$

$$\text{Percentage Error} = \text{Relative Error} \times 100$$

$$= \frac{|\text{Exact value} - \text{Approximate Value}|}{|\text{Exact valu}e|} \times 100$$

Relative Error is more meaningful than the Absolute Error because it takes into consideration the size of the number being approximated

For practical reasons, the relative error is usually more meaningful than the absolute error. For example, if $\alpha_1 = 1.333$, $\beta_1 = 1.334$, and $\alpha_2 = 0.001$, $\beta_2 = 0.002$, then the absolute error of β_i as an approximation to α_i is the same in both cases—namely, 10^{-3}. However, the relative errors are $\frac{3}{4} \times 10^{-3}$ and 1, respectively. The relative error clearly indicates that β_1 is a good approximation to α_1 but that β_2 is a poor approximation to α_2.

Local Error

local error measures the accuracy of the method at a specific step, assuming that the method was exact at the previous step.

Global Error

At a given step, the global error of an approximate solution contains both the local error at that step and the accumulative effect of all the local errors at all previous steps.

Truncation Error

This terminology originates from the technique of replacing a complicated function with a truncated Taylor series. For example, the infinite Taylor series

$$e^{x^2} = 1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \frac{x^8}{4!} + \cdots + \frac{x^{2n}}{n!} + \cdots$$

might be replaced with just the first five terms $1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \frac{x^8}{4!}$. This might be done when approximating an integral numerically.

Term-by-term integration produces

$$\int_0^{1/2}\left(1+x^2+\frac{x^4}{2!}+\frac{x^6}{3!}+\frac{x^8}{4!}\right)dx=\left(x+\frac{x^3}{3}+\frac{x^5}{5(2!)}+\frac{x^7}{7(3!)}+\frac{x^9}{9(4!)}\right)_{x=0}^{x=1/2}$$

$$=\frac{1}{2}+\frac{1}{24}+\frac{1}{320}+\frac{1}{5376}+\frac{1}{110,592}$$

$$=\frac{2,109,491}{3.870.720}=0.544986720817$$

Scientific Notation

Scientific notation is commonly used to represent very large or very small numbers in a convenient way.

Scientific notation is written in the following format using a power with base 10.

$$m \times 10^{b}$$

where $1 \le m < 10$

large numbers will have a positive exponent.

small numbers will have a negative exponent.

To write a number in scientific notation:

- Move the decimal to a position immediately to the right of the first nonzero digit.
- Count the number of place values you had to move the decimal point. This is the value of the exponent.
    - If you moved the decimal point to the **left**, make the **exponent positive**.
    - If you moved the decimal point to the **right**, make the **exponent negative**.

e.g.

$$0.0000747 = 7.47 \times 10^{-5},$$

$$31.4159265 = 3.14159265 \times 10,$$

$$9{,}700{,}000{,}000 = 9.7 \times 10^{9}.$$

Rounding and Chopping

Consider the normalized form of a positive real number

$$y = 0.d_1 d_2 \ldots d_k d_{k+1} d_{k+2} \ldots \times 10^n$$

$$(\; 1 \le d_1 \le 9, \quad \text{and} \quad 0 \le d_i \le 9 \;)$$

The floating-point form of y, denoted $fl(y)$, is obtained by terminating the mantissa of y at k decimal digits. There are two common ways of performing this termination. One method, called **chopping**, is to simply chop off the digits $d_{k+1} d_{k+2} \ldots$. This produces the floating-point form

$$fl(y) = 0.d_1 d_2 \ldots d_k \times 10^n.$$

The other method, called **rounding**,

For rounding, when $d_{k+1} \geq 5$, we add 1 to d_k to obtain $fl(y)$; that is, we *round up*. When $d_{k+1} < 5$, we simply chop off all but the first k digits; so we *round down*.

Problem

Determine the five-digit **(a)** chopping

and **(b)** rounding values of the irrational number π.

Solution

The number π has an infinite decimal expansion of the form

$$\pi = 3.14159265\ldots.$$

Written in normalized decimal form, we have

$$\pi = 0.314159265\ldots \times 10^1$$

(a) The floating-point form of π using five-digit chopping is

$$fl(\pi) = 0.31415 \times 10^1 = 3.1415$$

(b) The sixth digit of the decimal expansion of π is a 9, so the floating-point form of π using five-digit rounding is

$$fl(\pi) = (0.31415 + 0.00001) \times 10^1 = 3.1416$$

Round-off Error

The error that results from replacing a number with its floating-point form is called **round-off error** regardless of whether the rounding or chopping method is used.

Round-off Errors and Computer Arithmetic

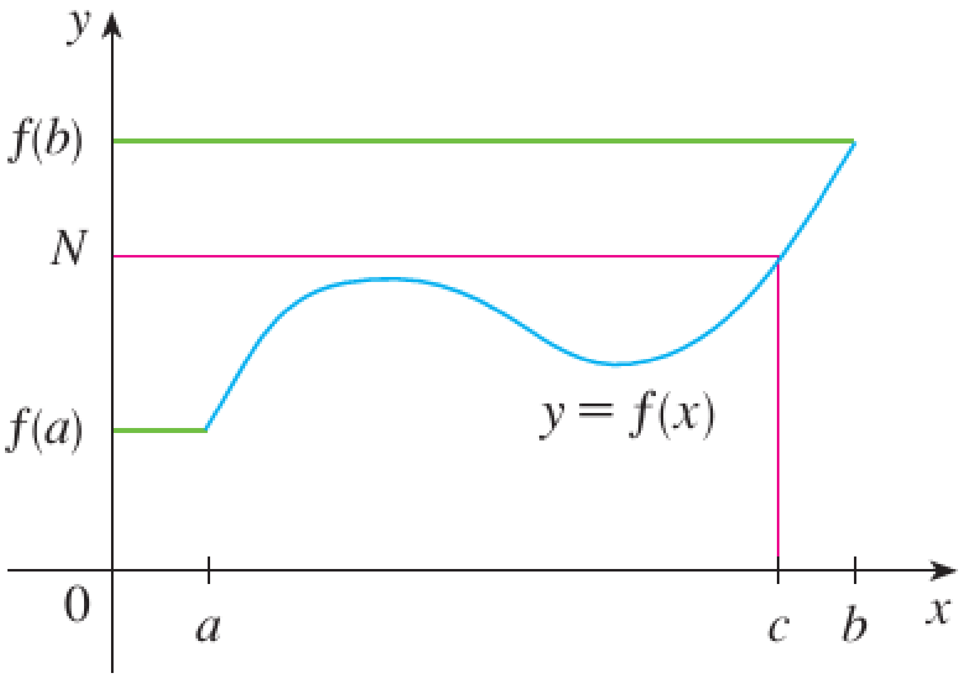
The error that is produced when a calculator or computer is used to perform real-number calculations is called **round-off error**. It occurs because the arithmetic performed in a machine involves numbers with only a finite number of digits, with the result that calculations are performed with only approximate representations of the actual numbers.

Algorithms

An **algorithm** is a procedure that describes, in an unambiguous manner, a finite sequence of steps to be performed in a specified order.

The Intermediate Value Theorem Suppose that f is continuous on the closed interval $[a, b]$ and let N be any number between $f(a)$ and $f(b)$, where $f(a) \neq f(b)$. Then there exists a number c in (a, b) such that $f(c) = N$.

The Intermediate Value Theorem states that a continuous function takes on every intermediate value between the function values $f(a)$ and $f(b)$. It is illustrated by Figure 8. Note that the value N can be taken on once [as in part (a)] or more than once [as in part (b)].



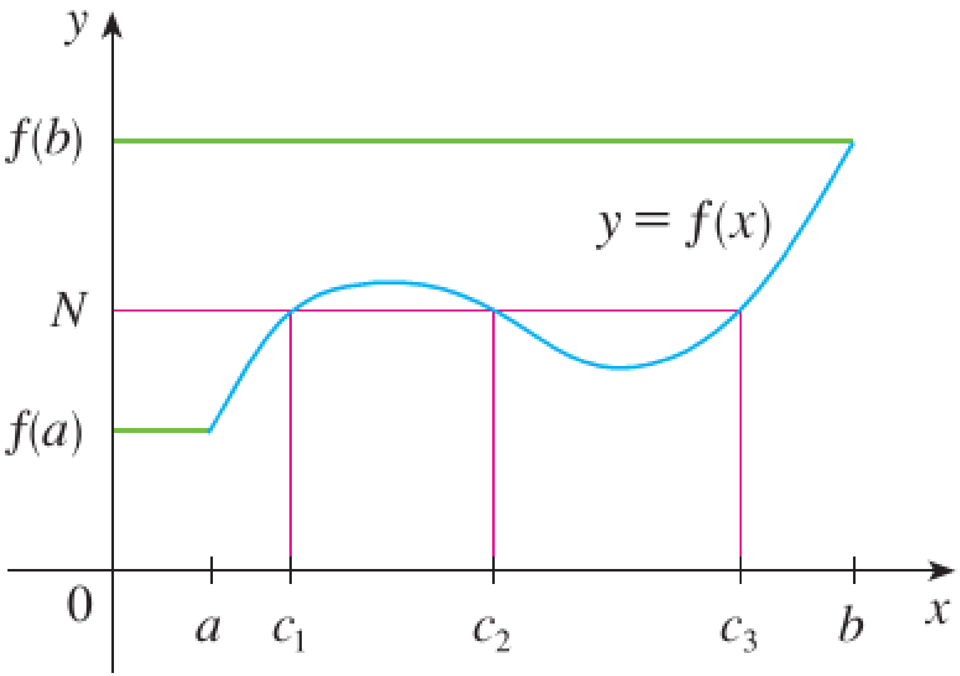


FIGURE 8

Root of an Equation & Zero of a Function

Any number r for which $f(r) = 0$ is called a ***root of the equation*** $f(x) = 0$. Also, we say that r is a ***zero of the function*** $f(x)$.

For example, the equation $2x^2 + 5x - 3 = 0$ has two real roots $r_1 = 0.5$ and $r_2 = -3$, whereas the corresponding function $f(x) = 2x^2 + 5x - 3 = (2x - 1)(x + 3)$ has two real zeros, $r_1 = 0.5$ and $r_2 = -3$.

THEOREM

Let f be a continuous function on $[a, b]$, satisfying $f(a)f(b) < 0$. Then f has a root between a and b, that is, there exists a number r satisfying $a < r < b$ and $f(r) = 0$.

Show that there is a root of the equation

$$4x^3 - 6x^2 + 3x - 2 = 0$$

between 1 and 2.

SOLUTION Let $f(x) = 4x^3 - 6x^2 + 3x - 2$. We are looking for a solution of the given equation, that is, a number c between 1 and 2 such that $f(c) = 0$. Therefore we take $a = 1$, $b = 2$, and $N = 0$ in Theorem 10. We have

$$f(1) = 4 - 6 + 3 - 2 = -1 < 0$$

and

$$f(2) = 32 - 24 + 6 - 2 = 12 > 0$$

Thus $f(1) < 0 < f(2)$; that is, $N = 0$ is a number between $f(1)$ and $f(2)$. Now f is continuous since it is a polynomial, so the Intermediate Value Theorem says there is a number c between 1 and 2 such that $f(c) = 0$. In other words, the equation $4x^3 - 6x^2 + 3x - 2 = 0$ has at least one root c in the interval $(1, 2)$.

***Tolerance in the solution*:** A tolerance is the maximum amount by which the true solution can deviate from an approximate numerical solution.

Solutions of Nonlinear Equations

THE BISECTION METHOD

The first technique, based on the Intermediate Value Theorem, is called the **Bisection**, or **Binary-search, method**.

In computer science, the process of dividing a set continually in half to search for the solution to a problem, as the bisection method does, is known as a *binary search* procedure.

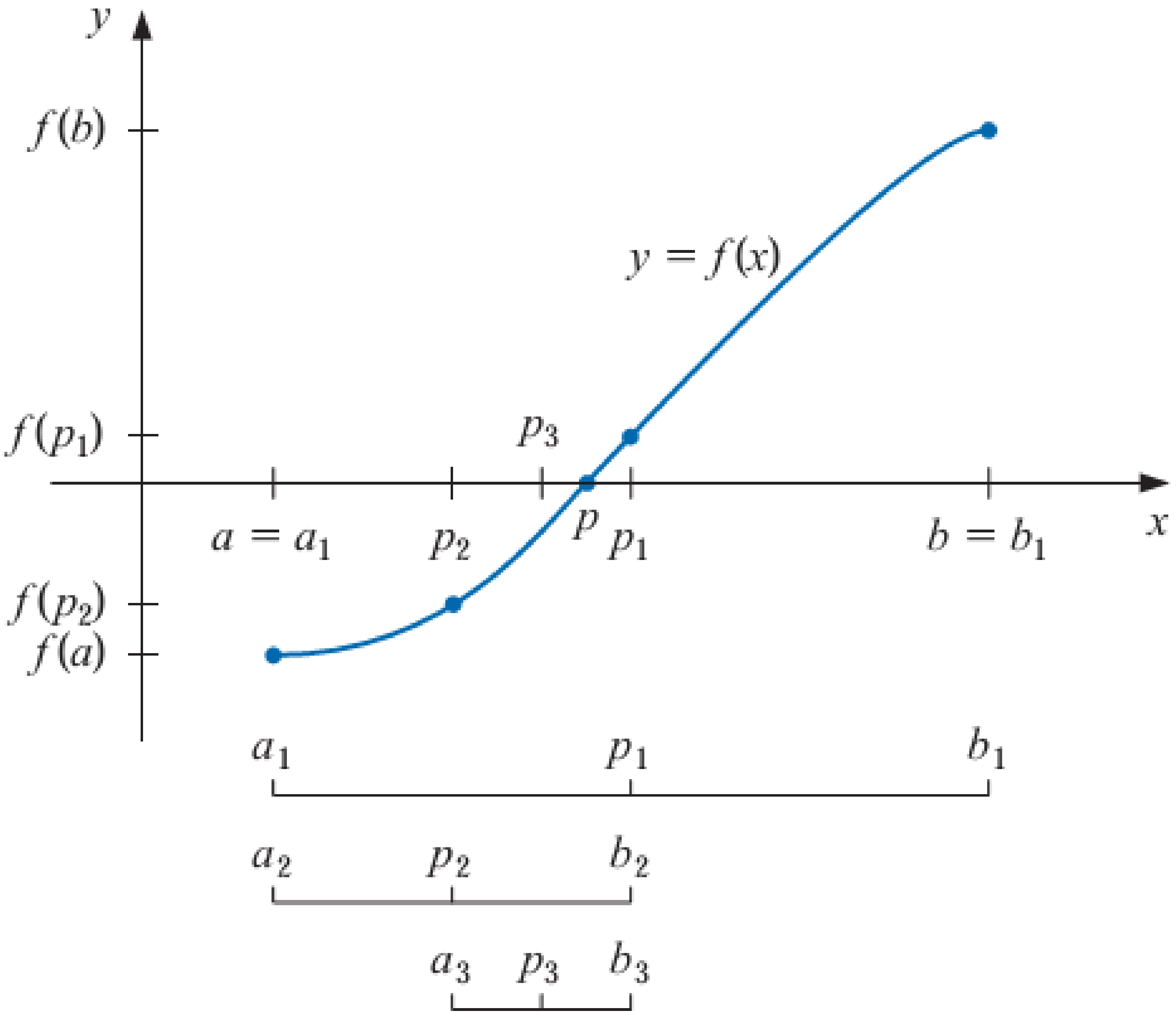
Suppose f is a continuous function defined on the interval $[a, b]$, with $f(a)$ and $f(b)$ of opposite sign. The Intermediate Value Theorem implies that a number p exists in (a, b) with $f(p) = 0$. Although the procedure will work when there is more than one root in the interval (a, b), we assume for simplicity that the root in this interval is unique. The method calls for a repeated halving (or bisecting) of subintervals of $[a, b]$ and, at each step, locating the half containing p.

To begin, set $a_1 = a$ and $b_1 = b$, and let p_1 be the midpoint of $[a, b]$; that is,

$$p_1 = a_1 + \frac{b_1 - a_1}{2} = \frac{a_1 + b_1}{2}.$$

- If $f(p_1) = 0$, then $p = p_1$, and we are done.
- If $f(p_1) \neq 0$, then $f(p_1)$ has the same sign as either $f(a_1)$ or $f(b_1)$.
  - If $f(p_1)$ and $f(a_1)$ have the same sign, $p \in (p_1, b_1)$. Set $a_2 = p_1$ and $b_2 = b_1$.
  - If $f(p_1)$ and $f(a_1)$ have opposite signs, $p \in (a_1, p_1)$. Set $a_2 = a_1$ and $b_2 = p_1$.

Then reapply the process to the interval $[a_2, b_2]$.


y
f(b)
y = f(x)
f(p1)
p3
a = a1
p2
p
p1
b = b1
x
f(p2)
f(a)
a1
p1
b1
a2
p2
b2
a3
p3
b3

Stopping Criteria

When using a computer to generate approximations, it is good practice to set an upper bound on the number of iterations. This eliminates the possibility of entering an infinite loop

Theorem

Suppose that $f \in C[a, b]$ and $f(a) \cdot f(b) < 0$. The Bisection method generates a sequence $\{p_n\}_{n=1}^{\infty}$ approximating a zero p of f with

$$|p_n - p| \leq \frac{b-a}{2^n}, \quad \text{when} \quad n \geq 1.$$

Problem

Determine the number of iterations necessary to solve $f(x) = x^3 + 4x^2 - 10 = 0$ with accuracy 10^{-3} using $a_1 = 1$ and $b_1 = 2$.

Solution We we will use logarithms to find an integer N that satisfies

$$|p_N - p| \le 2^{-N}(b - a) = 2^{-N} < 10^{-3}.$$

Logarithms to any base would suffice, but we will use base-10 logarithms because the tolerance is given as a power of 10. Since $2^{-N} < 10^{-3}$ implies that $\log_{10} 2^{-N} < \log_{10} 10^{-3} = -3$, we have

$$-N \log_{10} 2 < -3 \quad \text{and} \quad N > \frac{3}{\log_{10} 2} \approx 9.96.$$

Hence, ten iterations will ensure an approximation accurate to within 10^{-3}.

Formula for calculating number of iteration when accuracy within 10^{-p} is required:

$$N > \frac{p + \log(b - a)}{\log(2)}$$

Note: In above formula base of the log is 10.

Practice Problem

Determine the number of iterations required to find the zero of

$$f(x) = x^3 - x^2 - 1$$

in $[1, 2]$ with an absolute error of no more than 10^{-6}.

Ans: 20 iterations

Problem

Show that $f(x) = x^3 + 4x^2 - 10 = 0$ has a root in $[1, 2]$, and use the Bisection method to determine an approximation to the root that is accurate to at least within 10^{-4}.

Solution

Because $f(1) = -5$ and $f(2) = 14$ the Intermediate Value Theorem ensures that this continuous function has a root in $[1, 2]$.

Using Formula

$$N > \frac{p + \log(b - a)}{\log(2)}$$

We can say that 14 iteration will ensure approximation accurate to within 10^{-4}

n	a_n	b_n	p_n	$f(p_n)$
1	1.0	2.0	1.5	2.375
2	1.0	1.5	1.25	−1.79687
3	1.25	1.5	1.375	0.16211
4	1.25	1.375	1.3125	−0.84839
5	1.3125	1.375	1.34375	−0.35098
6	1.34375	1.375	1.359375	−0.09641
7	1.359375	1.375	1.3671875	0.03236
8	1.359375	1.3671875	1.36328125	−0.03215
9	1.36328125	1.3671875	1.365234375	0.000072
10	1.36328125	1.365234375	1.364257813	−0.01605
11	1.364257813	1.365234375	1.364746094	−0.00799
12	1.364746094	1.365234375	1.364990235	−0.00396
13	1.364990235	1.365234375	1.365112305	−0.00194
14	1.365112305	1.365234375	**1.365173340**	**-0.0009358**

Exact Value: p=1.365230013 (Calculated Using MATLAB)

Approximate Value: p_{14} =1.365173340

Absolute Error= 0.00005667

Desired Accuracy achieved

The Latin word *signum* means "token" or "sign". So the signum function quite naturally returns the sign of a number (unless the number is 0).

$$\operatorname{sgn}(x) = \begin{cases} -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0, \\ 1, & \text{if } x > 0. \end{cases}$$

MATLAB Code For Bisection Method

```
%Program for Bisection Method
%To compute approximate solution of f(x)=0
%Input: function handle f; a,b such that f(a)*f(b)<0,
% and tolerance tol
%Output: Approximate solution pn
function pn=bisec(f,a,b,tol)
if f(a)==0
    disp('a is a root of the given equation')
end
if f(b)==0
    disp('b is a root  of the given equation')
end
if sign(f(a))*sign(f(b))>=0
error('Bisection Method is not Applicable Here') %ceases execution
end
```

Continue in next slide

```
fa=f(a);
fb=f(b);
while (b-a)/2>tol
pn=(a+b)/2;
fpn=f(pn);
if fpn == 0 %pn is a solution, done
break
end
if sign(fpn)*sign(fa)<0 %a and pn make the new interval
b=pn;fb=fpn;
else %pn and b make the new interval
a=pn;fa=fpn;
end
end
pn=(a+b)/2; %new midpoint is best estimate
```

EXERCISE SET 2.1

Recommended Problems: 1 to 6

Q3.

Use the Bisection method to find solutions accurate to within 10^{-2} for $x^3 - 7x^2 + 14x - 6 = 0$ on each interval.

a. $[0, 1]$ **b.** $[1, 3.2]$ **c.** $[3.2, 4]$

Practice Problem

Consider the equation $x^4 = x^3 + 10$.

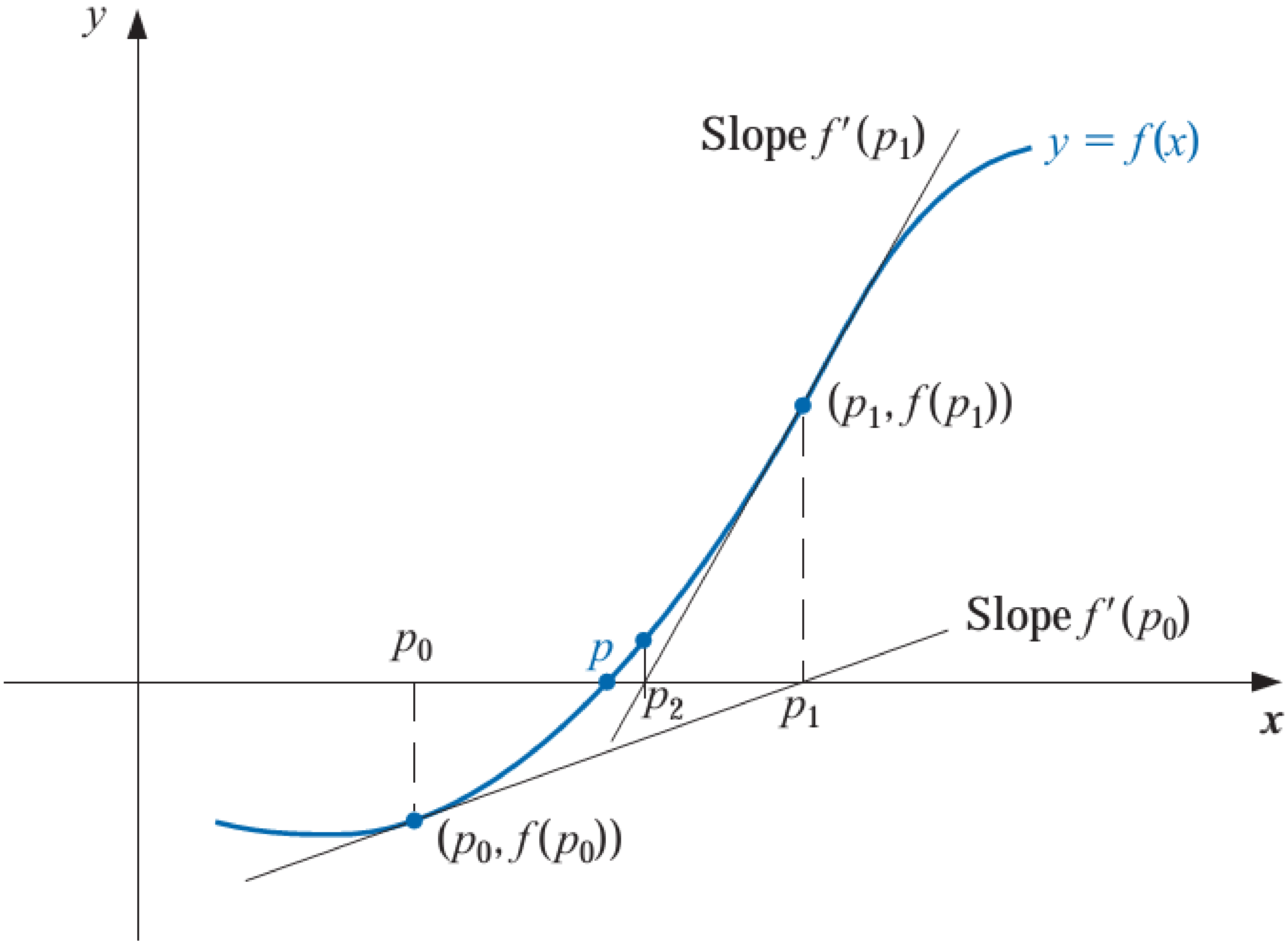
(a) Find an interval $[a, b]$ of length one inside which the equation has a solution.

(b) Starting with $[a, b]$, how many steps of the Bisection Method are required to calculate the solution within 10^{-10}? Answer with an integer.

Newton's Method

Newton's (or the *Newton-Raphson*) **method** is one of the most powerful and well-known numerical methods for solving a root-finding problem.

Starting with the initial approximation p_0, the approximation p_1 is the x-intercept of the tangent line to the graph of f at $(p_0, f(p_0))$. The approximation p_2 is the x-intercept of the tangent line to the graph of f at $(p_1, f(p_1))$ and so on.



y
Slope $f'(p_1)$
$y = f(x)$
$(p_1, f(p_1))$
Slope $f'(p_0)$
p_0
p
p_2
p_1
x
$(p_0, f(p_0))$

Formula for calculating approximate root of f(x)=0 using Newton-Raphson Method

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad \text{for } n \geq 1.$$

Stopping Criterias

$$|p_N - p_{N-1}| < \varepsilon,$$

$$\frac{|p_N - p_{N-1}|}{|p_N|} < \varepsilon, \quad p_N \neq 0,$$

$$|f(p_N)| < \varepsilon.$$

Problem

Use Newton's Method to approximate a root of the equation $\cos x - x = 0$. Take initial guess $p_0 = \frac{\pi}{4}$.

Solution

To apply Newton's method to this problem we need $f'(x) = -\sin x - 1$. Starting with $p_0 = \pi/4$, we generate the sequence defined, for $n \geq 1$, by

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f(p'_{n-1})} = p_{n-1} - \frac{\cos p_{n-1} - p_{n-1}}{-\sin p_{n-1} - 1}.$$

This gives the approximations in Table 2.4. An excellent approximation is obtained with $n = 3$. Because of the agreement of p_3 and p_4 we could reasonably expect this result to be accurate to the places listed.

	Newton's Method
n	p_n
0	0.7853981635
1	0.7395361337
2	0.7390851781
3	0.7390851332
4	0.7390851332

Problem

Find the Newton's Method formula for the equation $x^3 + x - 1 = 0$.

Division-by-Zero

One obvious pitfall of the Newton-Raphson method is the possibility of division by zero in formula which would occur if $f'(p_{k-1}) = 0$.

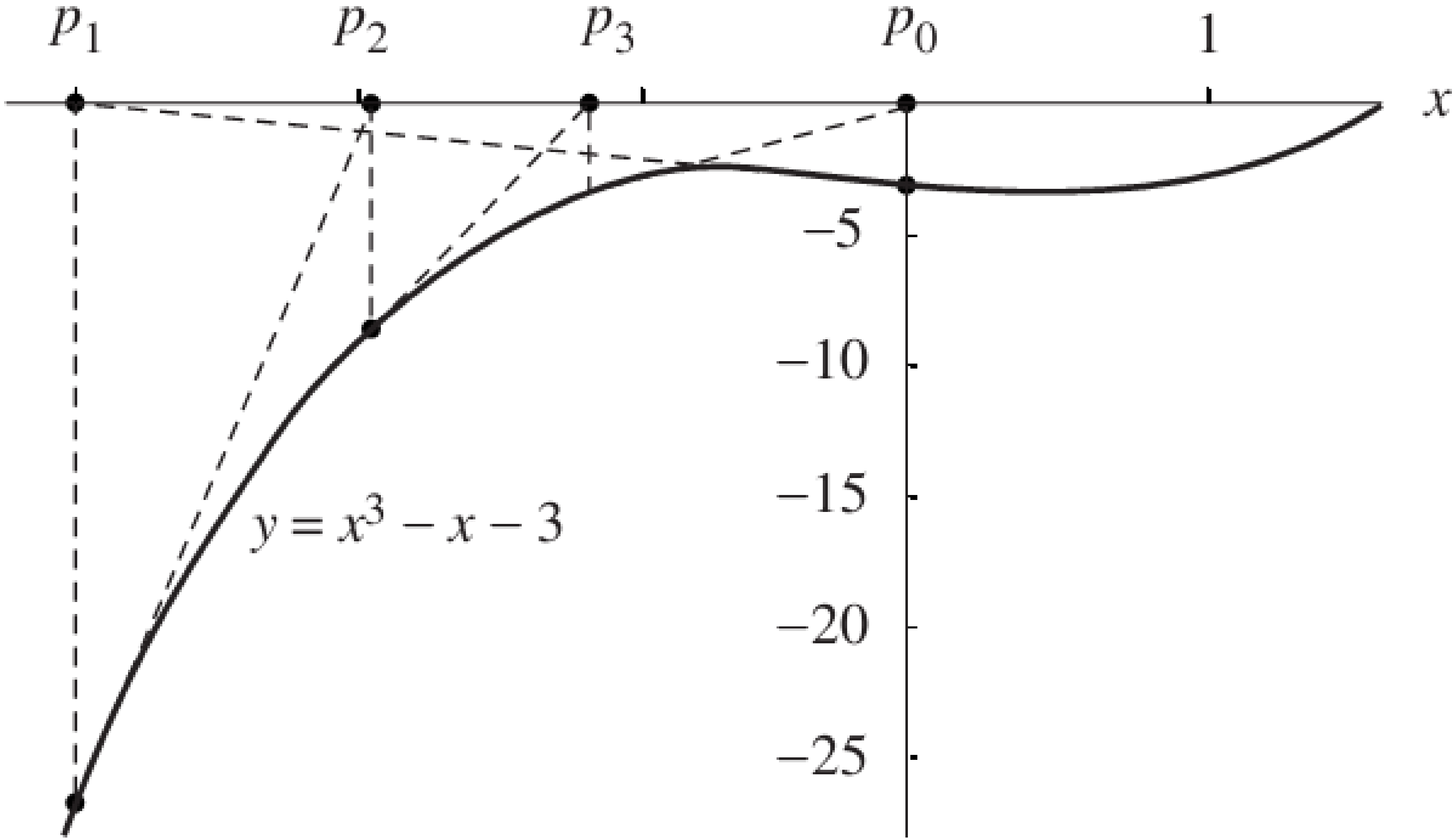
Note: The method may converge to a root different from the expected one or diverge if the starting value is not close enough to the root.

★ ***cycling*** occurs when the terms in the sequence $\{p_k\}$ tend to repeat or almost repeat. For example, if $f(x) = x^3 - x - 3$ and the initial approximation is $p_0 = 0$, then the sequence is

$$p_1 = -3.000000, \quad p_2 = -1.961538, \quad p_3 = -1.147176, \quad p_4 = -0.006579,$$
$$p_5 = -3.000389, \quad p_6 = -1.961818, \quad p_7 = -1.147430, \quad \ldots$$

and we are stuck in a cycle where $p_{k+4} \approx p_k$ for $k = 0, 1, \ldots$

Graph is shown in the next slide.



p_1
p_2
p_3
p_0
1
x
−5
−10
−15
−20
−25
$y = x^3 - x - 3$

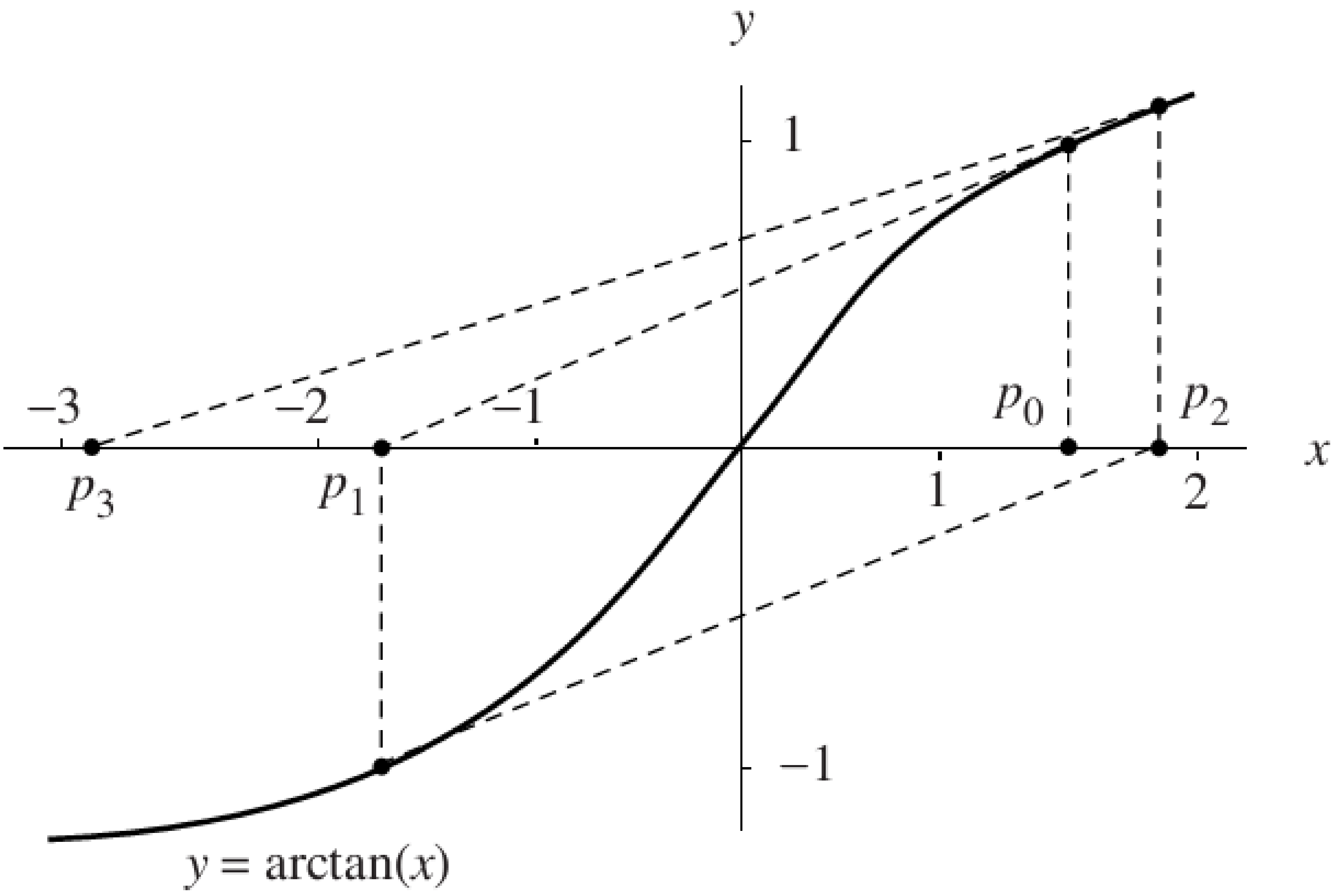
But if the starting value p_0 is sufficiently close to the root $p \approx 1.671699881$, then $\{p_k\}$ converges. If $p_0 = 2$, the sequence converges: $p_1 = 1.72727272$, $p_2 = 1.67369173$, $p_3 = 1.671702570$, and $p_4 = 1.671699881$.

★ there is a chance of divergent oscillation.

For example, let $f(x) = \arctan(x)$; then the Newton-Raphson iteration function is $g(x) = x - (1 + x^2)\arctan(x)$, and $g'(x) = -2x\arctan(x)$. If the starting value $p_0 = 1.45$ is chosen, then

$$p_1 = -1.550263297, \qquad p_2 = 1.845931751, \qquad p_3 = -2.889109054, \text{ etc.}$$

Graph is shown in the next slide.


y
1
−3
−2
−1
p_0
p_2
x
p_3
p_1
1
2
−1
y = arctan(x)

But if the starting value is sufficiently close to the root $p = 0$, a convergent sequence results. If $p_0 = 0.5$, then

$$p_1 = -0.079559511, \qquad p_2 = 0.000335302, \qquad p_3 = 0.000000000.$$

The situations above point to the fact that we must be honest in reporting an answer. Sometimes the sequence does not converge. It is not always the case that after N iterations a solution is found. The user of a root-finding algorithm needs to be warned of the situation when a root is not found. If there is other information concerning the context of the problem, then it is less likely that an erroneous root will be found. Sometimes $f(x)$ has a definite interval in which a root is meaningful. If knowledge of the behavior of the function or an "accurate" graph is available, then it is easier to choose p_0.

EXERCISE SET 2.3

1. Let $f(x) = x^2 - 6$ and $p_0 = 1$. Use Newton's method to find p_2.
2. Let $f(x) = -x^3 - \cos x$ and $p_0 = -1$. Use Newton's method to find p_2. Could $p_0 = 0$ be used?

5. Use Newton's method to find solutions accurate to within 10^{-4} for the following problems.
   a. $x^3 - 2x^2 - 5 = 0, \quad [1, 4]$
   b. $x^3 + 3x^2 - 1 = 0, \quad [-3, -2]$
   c. $x - \cos x = 0, \quad [0, \pi/2]$
   d. $x - 0.8 - 0.2 \sin x = 0, \quad [0, \pi/2]$

6. Use Newton's method to find solutions accurate to within 10^{-5} for the following problems.
   a. $e^x + 2^{-x} + 2\cos x - 6 = 0 \quad$ for $1 \leq x \leq 2$
   b. $\ln(x-1) + \cos(x-1) = 0 \quad$ for $1.3 \leq x \leq 2$
   c. $2x \cos 2x - (x-2)^2 = 0 \quad$ for $2 \leq x \leq 3$ and $3 \leq x \leq 4$

The Secant Method

The word secant is derived from the Latin word *secan*, which means to cut. The secant method uses a secant line, a line joining two points that cut the curve, to approximate a root.

Newton's method is an extremely powerful technique, but it has a major weakness: the need to know the value of the derivative of f at each approximation. Frequently, $f'(x)$ is far more difficult and needs more arithmetic operations to calculate than $f(x)$.

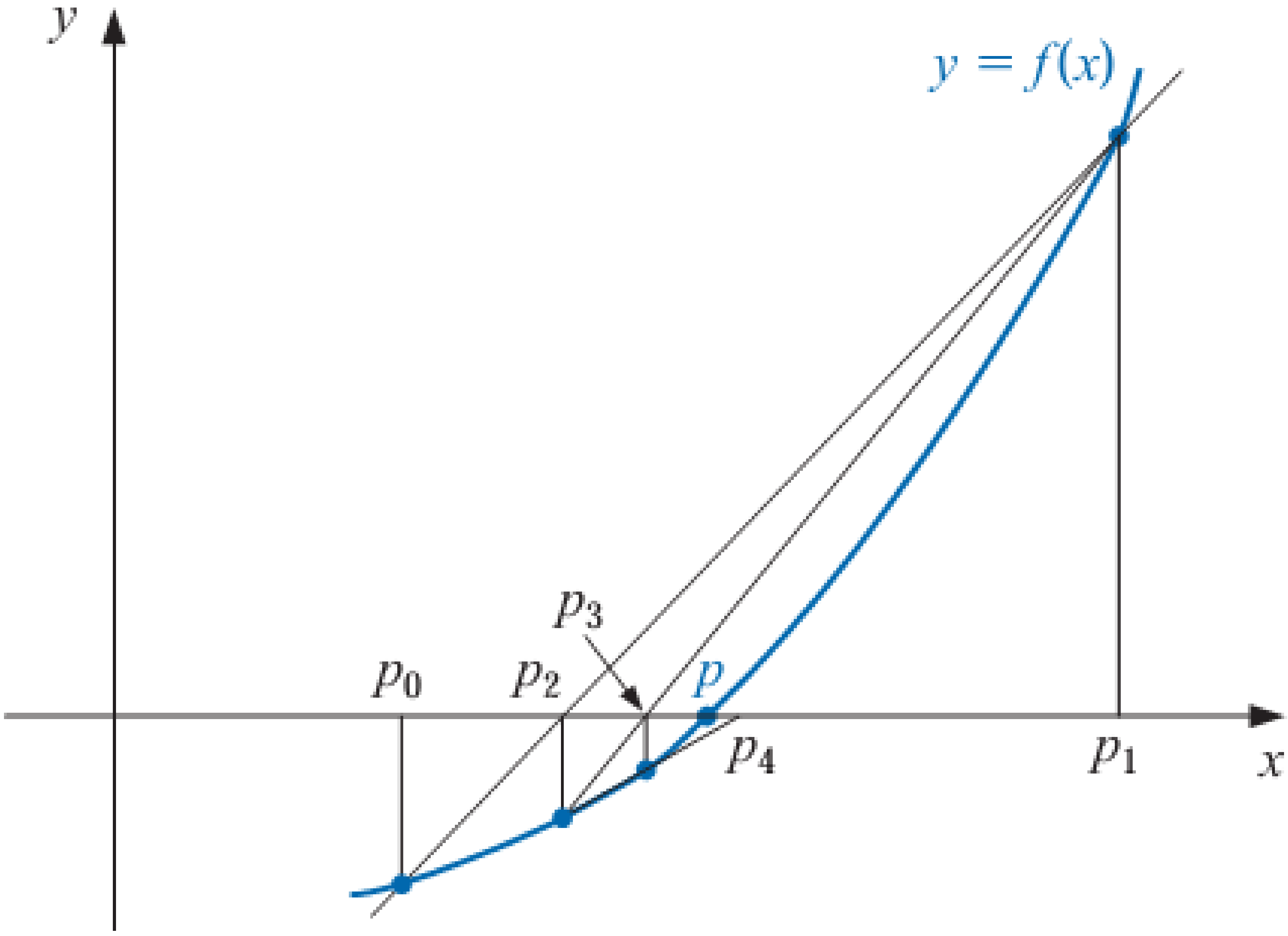


Figure 1

Starting with the two initial approximations p_0 and p_1, the approximation p_2 is the x-intercept of the line joining $(p_0, f(p_0))$ and $(p_1, f(p_1))$. The approximation p_3 is the x-intercept of the line joining $(p_1, f(p_1))$ and $(p_2, f(p_2))$, and so on. Note that only one function evaluation is needed per step for the Secant method after p_2 has been determined. In contrast, each step of Newton's method requires an evaluation of both the function and its derivative.

Formula for calculating approximate root of f(x)=0 using Secant Method

$$p_n = p_{n-1} - \frac{f(p_{n-1})(p_{n-1} - p_{n-2})}{f(p_{n-1}) - f(p_{n-2})}$$

$n \geq 2$

Problem

Use the Secant method to find a solution to $x = \cos x$. Use $p_0 = 0.5$ and $p_1 = \pi/4$.

Solution

Succeeding approximations are generated by the formula

$$p_n = p_{n-1} - \frac{(p_{n-1} - p_{n-2})(\cos p_{n-1} - p_{n-1})}{(\cos p_{n-1} - p_{n-1}) - (\cos p_{n-2} - p_{n-2})}, \quad \text{for } n \geq 2$$

n	Secant p_n
0	0.5
1	0.7853981635
2	0.7363841388
3	0.7390581392
4	0.7390851493
5	0.7390851332

Newton's method or the Secant method is often used to refine an answer obtained by another technique, such as the Bisection method, since these methods require good first approximations but generally give rapid convergence.

The Method of False Position

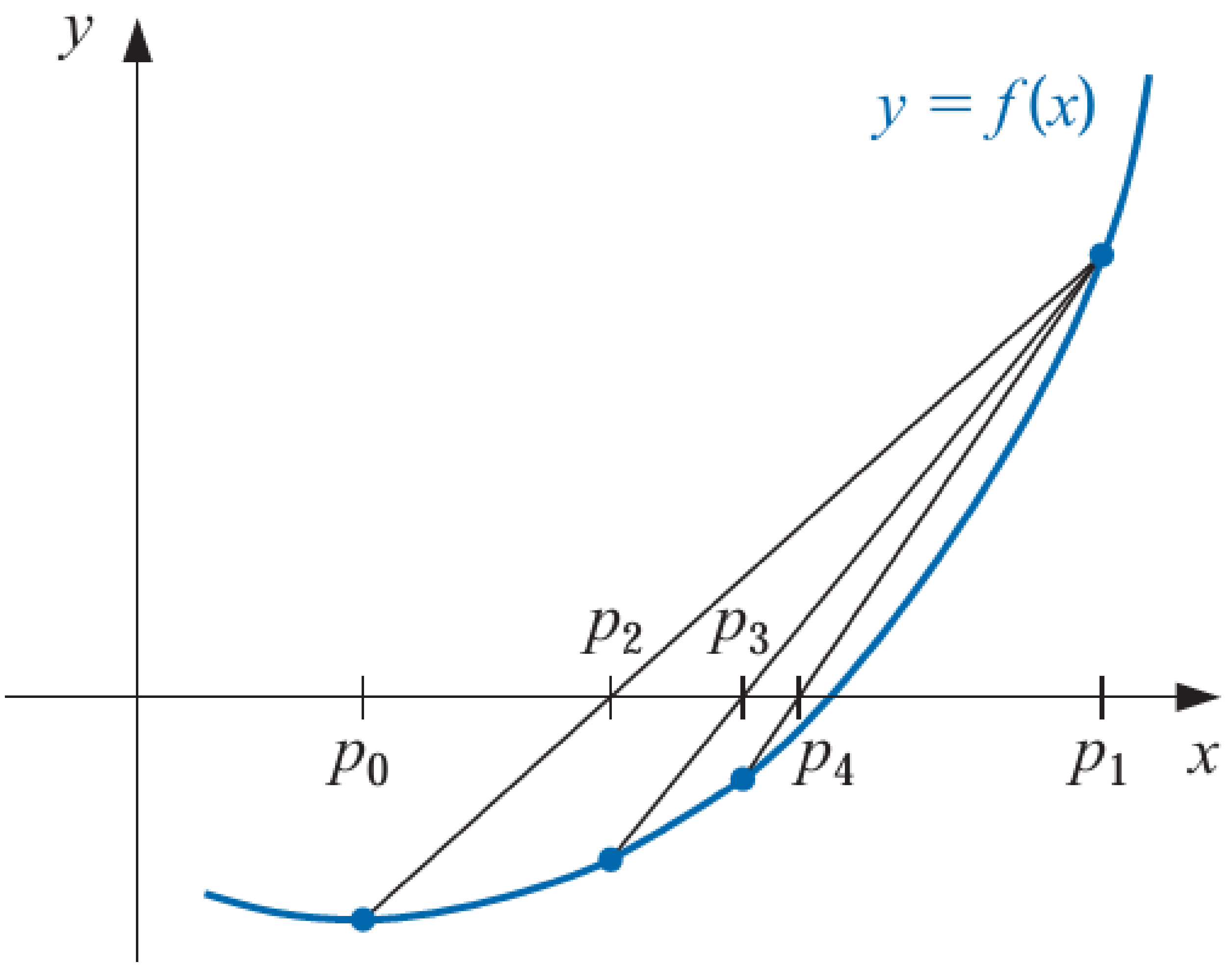
The **method of False Position** (also called *Regula Falsi*) generates approximations in the same manner as the Secant method, but it includes a test to ensure that the root is always bracketed between successive iterations. Although it is not a method we generally recommend, it illustrates how bracketing can be incorporated.

First choose initial approximations p_0 and p_1 with $f(p_0) \cdot f(p_1) < 0$. The approximation p_2 is chosen in the same manner as in the Secant method, as the x-intercept of the line joining $(p_0, f(p_0))$ and $(p_1, f(p_1))$. To decide which secant line to use to compute p_3, consider $f(p_2) \cdot f(p_1)$, or more correctly $\operatorname{sgn} f(p_2) \cdot \operatorname{sgn} f(p_1)$.

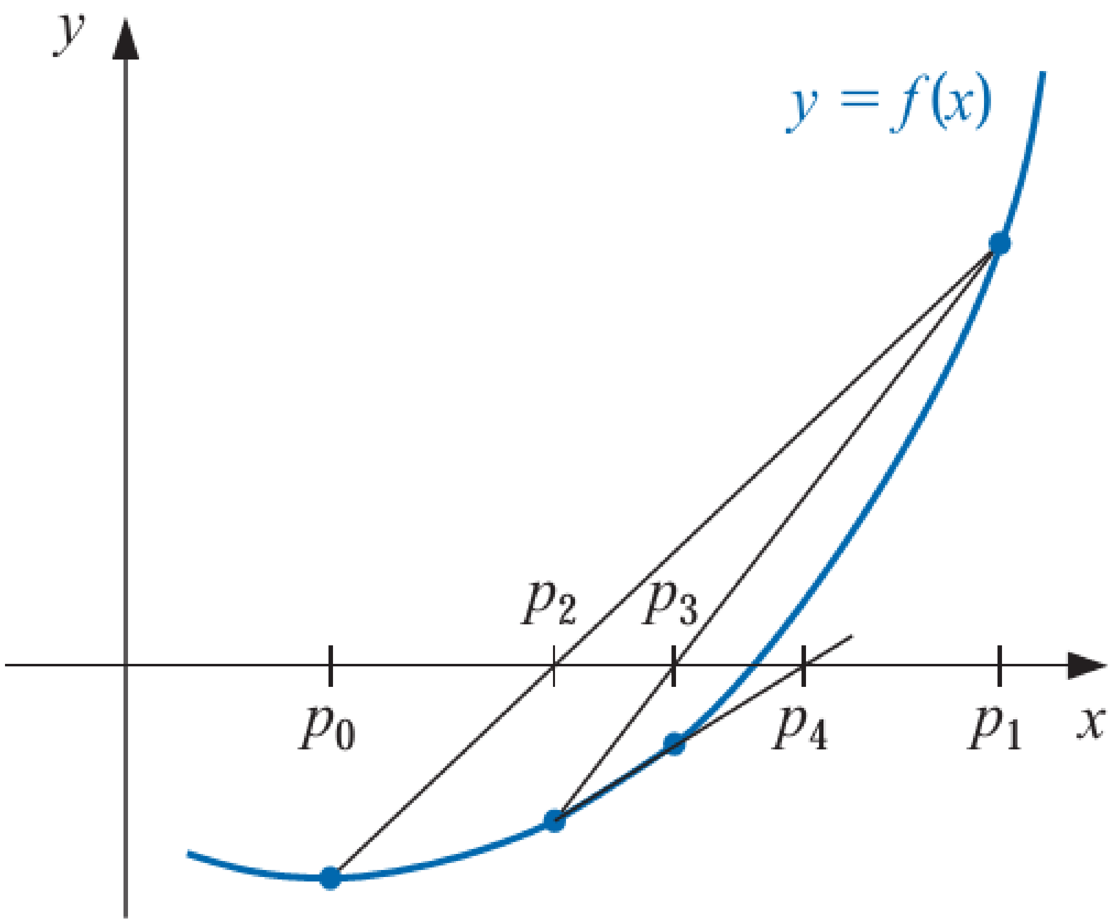
- If $\operatorname{sgn} f(p_2) \cdot \operatorname{sgn} f(p_1) < 0$, then p_1 and p_2 bracket a root. Choose p_3 as the x-intercept of the line joining $(p_1, f(p_1))$ and $(p_2, f(p_2))$.
- If not, choose p_3 as the x-intercept of the line joining $(p_0, f(p_0))$ and $(p_2, f(p_2))$, and then interchange the indices on p_0 and p_1.

In a similar manner, once p_3 is found, the sign of $f(p_3) \cdot f(p_2)$ determines whether we use p_2 and p_3 or p_3 and p_1 to compute p_4.

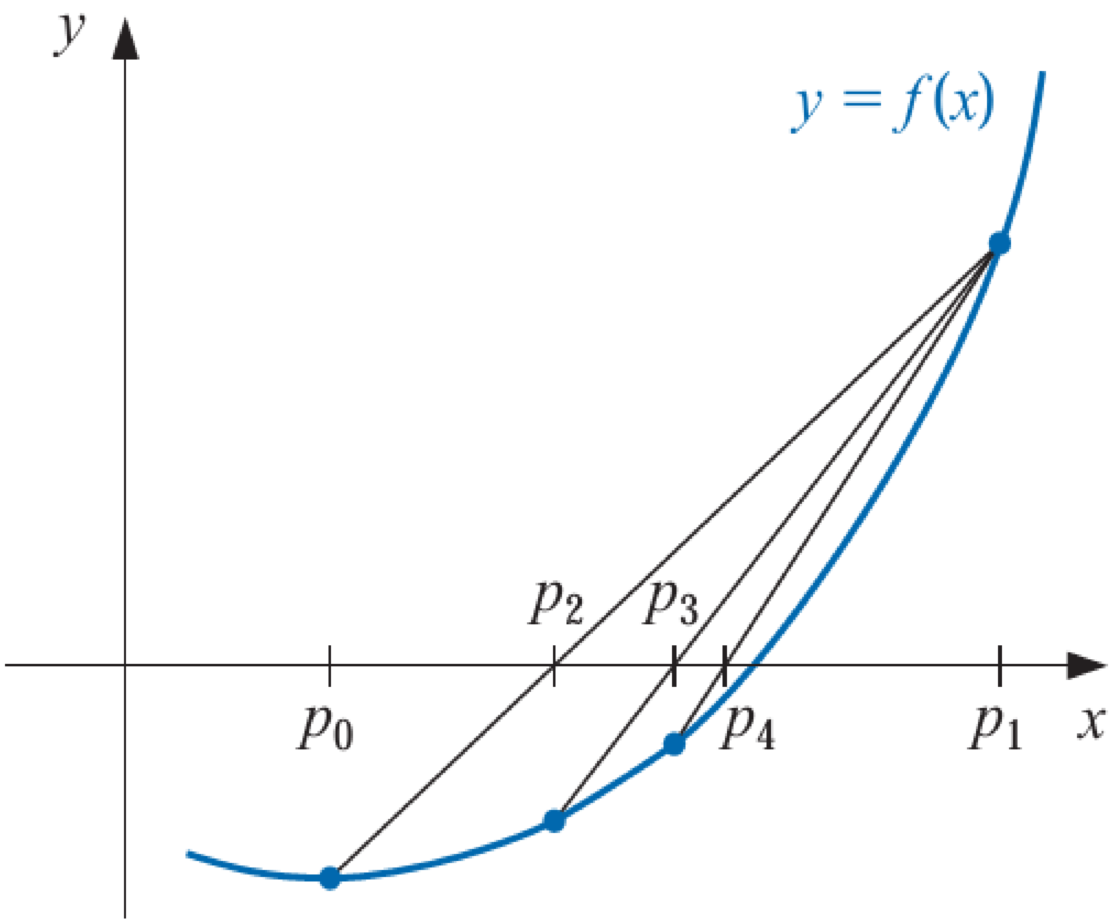
Method of False Position



Secant Method



Method of False Position



Problem

Use the false position method to find the root of $x \sin(x) - 1 = 0$ that is located in the interval $[0, 2]$ (the function $\sin(x)$ is evaluated in radians).

EXERCISE SET 2.3

3. Let $f(x) = x^2 - 6$. With $p_0 = 3$ and $p_1 = 2$, find p_3.

a. Use the Secant method.

b. Use the method of False Position.

c. Which of **a.** or **b.** is closer to $\sqrt{6}$?

4. Let $f(x) = -x^3 - \cos x$. With $p_0 = -1$ and $p_1 = 0$, find p_3.

a. Use the Secant method.

b. Use the method of False Position.

Fixed-Point Iteration

Definition

The number p is a **fixed point** for a given function g if $g(p) = p$.

e.g.

$p=1$ is a fixed point of the function $f(x)=x^2$ and

$q=0$ is a fixed point of the function $h(x)=sin(x)$

A fixed point for g occurs precisely when the graph of $y = g(x)$ intersects the graph of $y = x$

Problem

Determine any fixed points of the function $g(x) = x^2 - 2$.

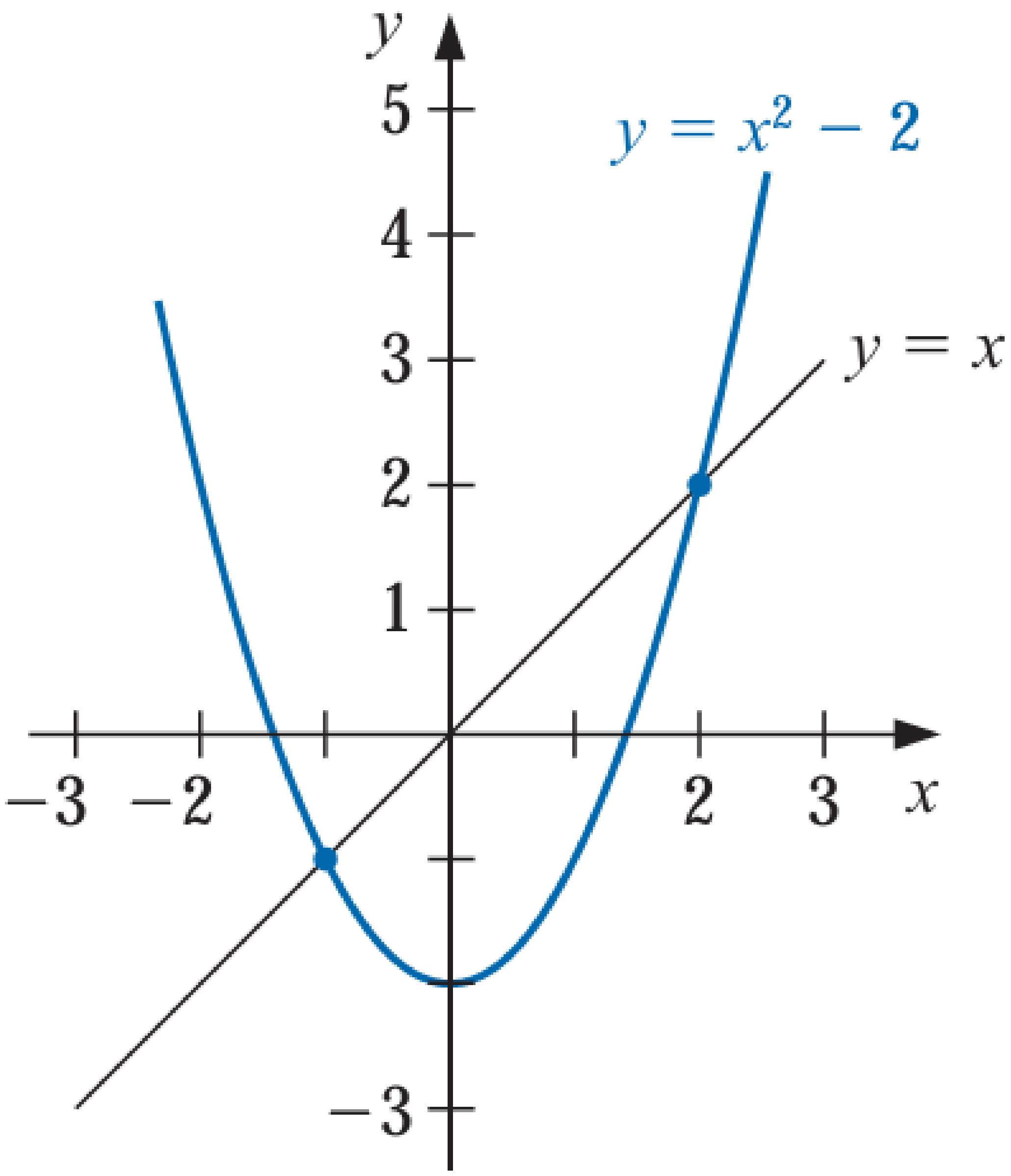
Solution

$$p = g(p) = p^2 - 2$$

implies that $0 = p^2 - p - 2 = (p + 1)(p - 2).$

$p = -1$ and the other **is** $p = 2.$

Graphical Discussion is in next slide.



y
5
4
3
2
1
−3
$y = x^2 - 2$
$y = x$
−3
−2
2
3
x

Connection b/w root finding problem & Fixed point problem

- Given a root-finding problem $f(p) = 0$, we can define functions g with a fixed point at p in a number of ways, for example, as

$$g(x) = x - f(x) \quad \text{or as} \quad g(x) = x + 3f(x).$$

- Conversely, if the function g has a fixed point at p, then the function defined by

$$f(x) = x - g(x)$$

has a zero at p.

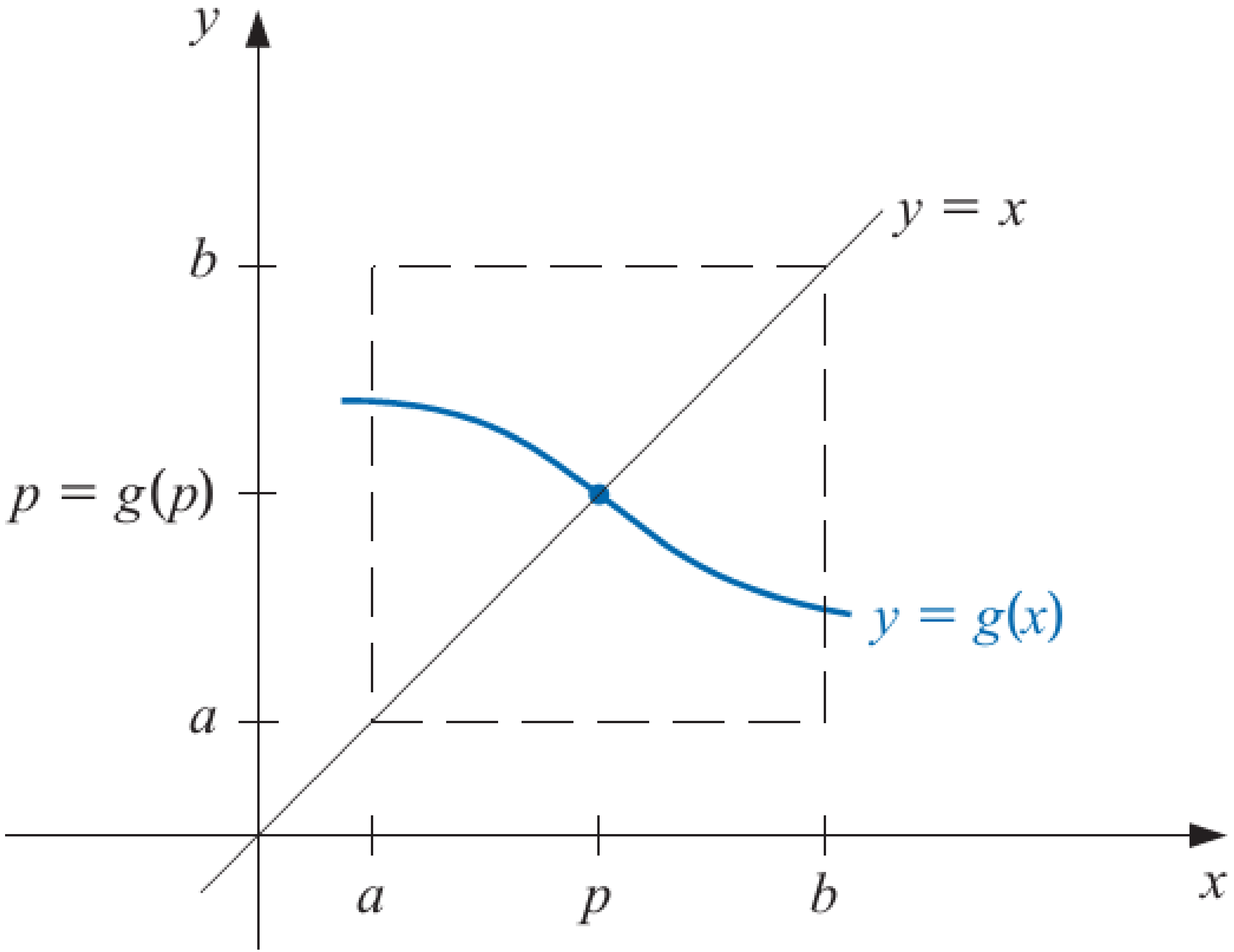
Theorem

(i) If $g \in C[a, b]$ and $g(x) \in [a, b]$ for all $x \in [a, b]$, then g has at least one fixed point in $[a, b]$.

(ii) If, in addition, $g'(x)$ exists on (a, b) and a positive constant $k < 1$ exists with

$$|g'(x)| \leq k, \quad \text{for all } x \in (a, b),$$

then there is exactly one fixed point in $[a, b]$.


y
y = x
b
p = g(p)
y = g(x)
a
a
p
b
x

Problem

Show that $g(x) = (x^2 - 1)/3$ has a unique fixed point on the interval $[-1, 1]$.

Solution The maximum and minimum values of $g(x)$ for x in $[-1, 1]$ must occur either when x is an endpoint of the interval or when the derivative is 0. Since $g'(x) = 2x/3$, the function g is continuous and $g'(x)$ exists on $[-1, 1]$. The maximum and minimum values of $g(x)$ occur at $x = -1$, $x = 0$, or $x = 1$. But $g(-1) = 0$, $g(1) = 0$, and $g(0) = -1/3$, so an absolute maximum for $g(x)$ on $[-1, 1]$ occurs at $x = -1$ and $x = 1$, and an absolute minimum at $x = 0$.

Moreover

$$|g'(x)| = \left|\frac{2x}{3}\right| \le \frac{2}{3}, \quad \text{for all } x \in (-1, 1).$$

So g satisfies all the hypotheses of Theorem and has a unique fixed point in $[-1, 1]$.

the unique fixed point p in the interval $[-1, 1]$ can be determined algebraically.

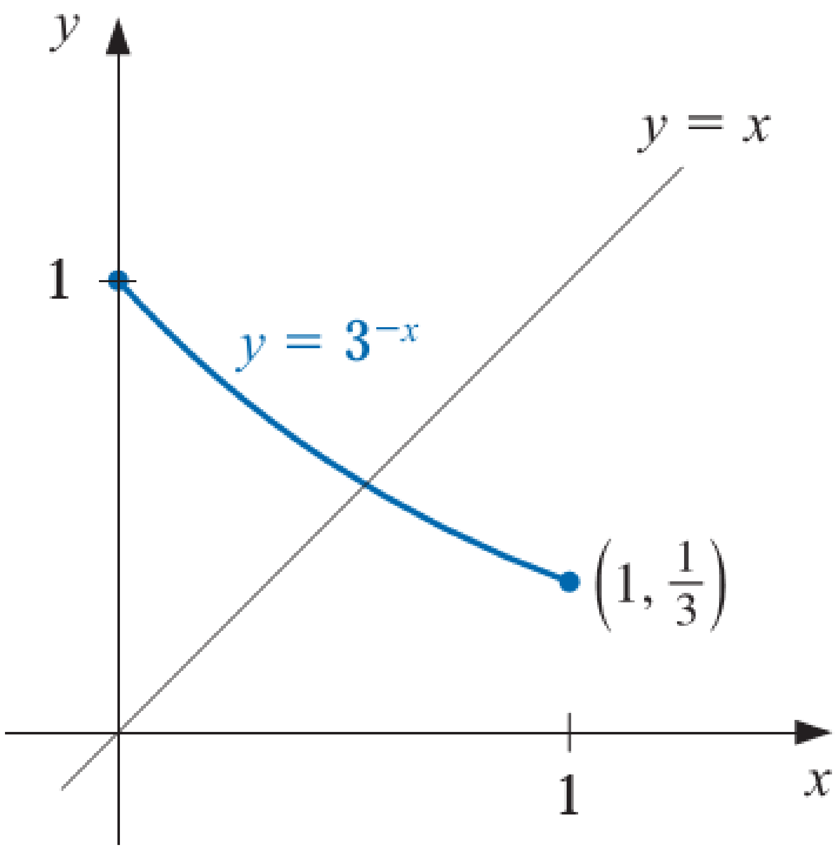
If

$$p = g(p) = \frac{p^2 - 1}{3}, \quad \text{then} \quad p^2 - 3p - 1 = 0,$$

$$\Longrightarrow \quad p = \frac{1}{2}(3 - \sqrt{13}).$$

Note that g also has a unique fixed point $p = \frac{1}{2}(3 + \sqrt{13})$ for the interval $[3, 4]$. However, $g(4) = 5$ and $g'(4) = \frac{8}{3} > 1$, so g does not satisfy the hypotheses of Theorem on $[3, 4]$. This demonstrates that the hypotheses of Theorem are sufficient to guarantee a unique fixed point but are not necessary.

Consider the graph of the function $g(x) = 3^{-x}$.



It is clear from the graph that g has a unique fixed point in [0,1]

But We cannot explicitly determine this fixed point because we have no way to solve for p in the equation $p = g(p) = 3^{-p}$.

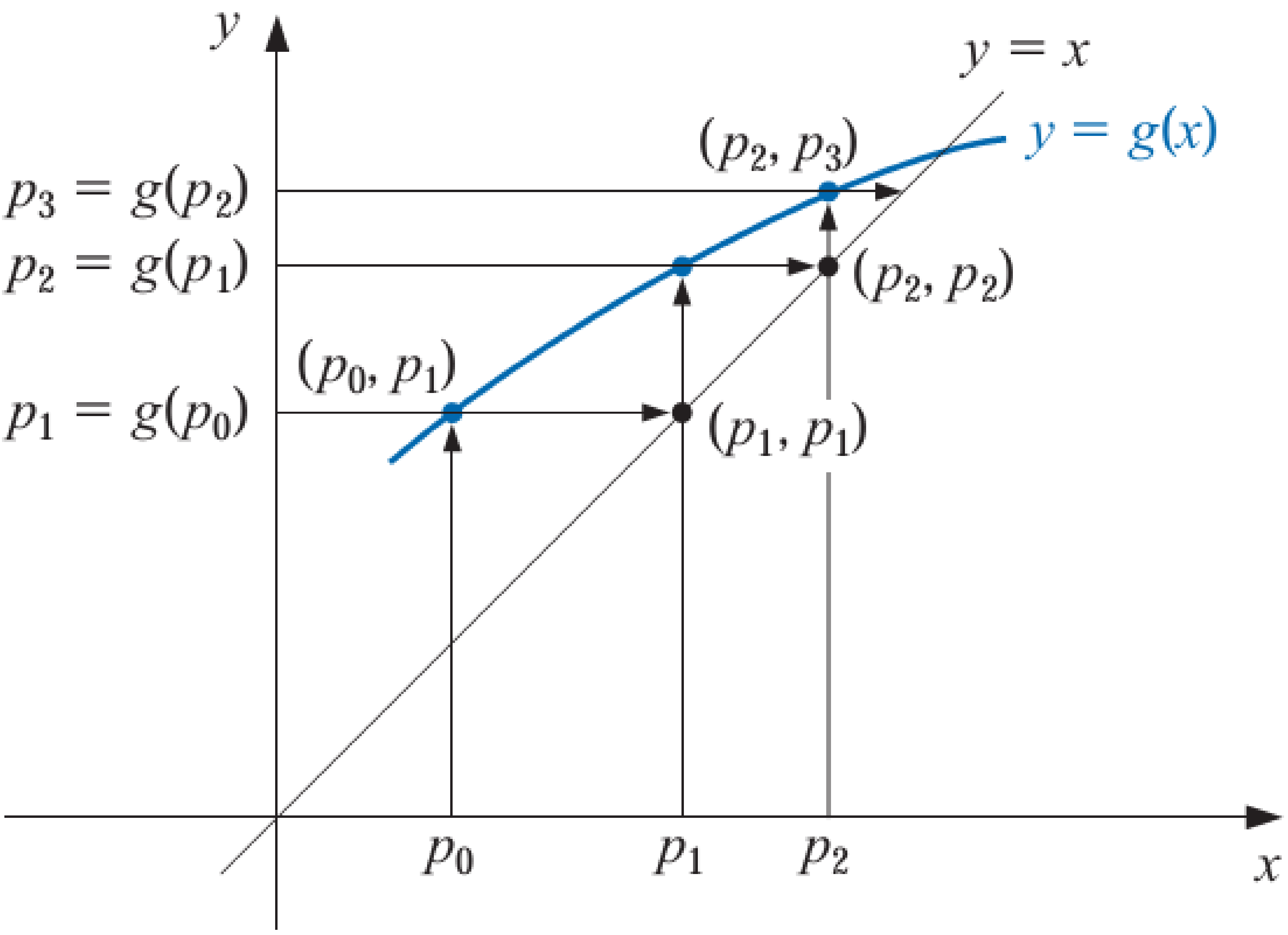
We can, however, determine approximations to this fixed point to any specified degree of accuracy. We will now consider how this can be done.

Fixed-Point Iteration

To approximate the fixed point of a function g, we choose an initial approximation p_0 and generate the sequence $\{p_n\}_{n=0}^{\infty}$ by letting $p_n = g(p_{n-1})$, for each $n \geq 1$. If the sequence converges to p and g is continuous, then

$$p = \lim_{n\to\infty} p_n = \lim_{n\to\infty} g(p_{n-1}) = g\left(\lim_{n\to\infty} p_{n-1}\right) = g(p),$$

and a solution to $x = g(x)$ is obtained. This technique is called **fixed-point**, or **functional iteration**.



y
$y = x$
$y = g(x)$
(p_2, p_3)
$p_3 = g(p_2)$
$p_2 = g(p_1)$
(p_2, p_2)
(p_0, p_1)
$p_1 = g(p_0)$
(p_1, p_1)
p_0
p_1
p_2
x

The equation $x^3 + 4x^2 - 10 = 0$ has a unique root in $[1, 2]$. There are many ways to change the equation to the fixed-point form $x = g(x)$ using simple algebraic manipulation.

(a) $x = g_1(x) = x - x^3 - 4x^2 + 10$

(b) $x = g_2(x) = \left(\dfrac{10}{x} - 4x\right)^{1/2}$

(c) $x = g_3(x) = \dfrac{1}{2}(10 - x^3)^{1/2}$

(d) $x = g_4(x) = \left(\dfrac{10}{4 + x}\right)^{1/2}$

(e) $x = g_5(x) = x - \dfrac{x^3 + 4x^2 - 10}{3x^2 + 8x}$

With $p_0 = 1.5$, Table 2.2 lists the results of the fixed-point iteration for all five choices of g.

n	(a)	(b)	(c)	(d)	(e)
0	1.5	1.5	1.5	1.5	1.5
1	−0.875	0.8165	1.286953768	1.348399725	1.373333333
2	6.732	2.9969	1.402540804	1.367376372	1.365262015
3	−469.7	$(-8.65)^{1/2}$	1.345458374	1.364957015	1.365230014
4	1.03×10^8		1.375170253	1.365264748	1.365230013
5			1.360094193	1.365225594	
6			1.367846968	1.365230576	
7			1.363887004	1.365229942	
8			1.365916734	1.365230022	
9			1.364878217	1.365230012	
10			1.365410062	1.365230014	
15			1.365223680	1.365230013	
20			1.365230236		
25			1.365230006		
30			1.365230013		

Table 2.2

Although the various functions we have given are fixed-point problems for the same root-finding problem, they differ vastly as techniques for approximating the solution to the root-finding problem. Their purpose is to illustrate what needs to be answered:

- Question: How can we find a fixed-point problem that produces a sequence that reliably and rapidly converges to a solution to a given root-finding problem?

Theorem (Fixed-Point Theorem)

Let $g \in C[a,b]$ be such that $g(x) \in [a,b]$, for all x in $[a,b]$. Suppose, in addition, that g' exists on (a,b) and that a constant $0 < k < 1$ exists with

$$|g'(x)| \leq k, \quad \text{for all } x \in (a,b).$$

Then for any number p_0 in $[a,b]$, the sequence defined by

$$p_n = g(p_{n-1}), \quad n \geq 1,$$

converges to the unique fixed point p in $[a,b]$.

- Answer: Manipulate the root-finding problem into a fixed point problem that satisfies the conditions of Fixed-Point Theorem and has a derivative that is as small as possible near the fixed point.

EXERCISE SET 2.2

5. Use a fixed-point iteration method to determine a solution accurate to within 10^{-2} for $x^4 - 3x^2 - 3 = 0$ on $[1, 2]$. Use $p_0 = 1$.

6. Use a fixed-point iteration method to determine a solution accurate to within 10^{-2} for $x^3 - x - 1 = 0$ on $[1, 2]$. Use $p_0 = 1$.

CHAPTER

6

Direct Methods for Solving Linear Systems

Linear Systems of Equations

$$
\begin{aligned}
E_1: &\quad a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\
E_2: &\quad a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\
&\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
E_n: &\quad a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n.
\end{aligned} \tag{6.1}
$$

We use three operations to simplify the linear system given in (6.1):

1. Equation E_i can be multiplied by any nonzero constant λ with the resulting equation used in place of E_i. This operation is denoted $(\lambda E_i) \rightarrow (E_i)$.

2. Equation E_j can be multiplied by any constant λ and added to equation E_i with the resulting equation used in place of E_i. This operation is denoted $(E_i + \lambda E_j) \rightarrow (E_i)$.

3. Equations E_i and E_j can be transposed in order. This operation is denoted $(E_i) \leftrightarrow (E_j)$.

The pivot element for a specific column is the entry that is used to place zeros in the other entries in that column.

Problem

Represent the linear system

$$
\begin{aligned}
E_1 &: \quad x_1 - x_2 + 2x_3 - x_4 = -8, \\
E_2 &: \quad 2x_1 - 2x_2 + 3x_3 - 3x_4 = -20, \\
E_3 &: \quad x_1 + x_2 + x_3 = -2, \\
E_4 &: \quad x_1 - x_2 + 4x_3 + 3x_4 = 4,
\end{aligned}
$$

as an augmented matrix and use Gaussian Elimination to find its solution.

Solution The augmented matrix is

$$\tilde{A} = \tilde{A}^{(1)} = \left[\begin{array}{rrrr:r} 1 & -1 & 2 & -1 & -8 \\ 2 & -2 & 3 & -3 & -20 \\ 1 & 1 & 1 & 0 & -2 \\ 1 & -1 & 4 & 3 & 4 \end{array}\right].$$

Performing the operations

$$(E_2 - 2E_1) \to (E_2), \ (E_3 - E_1) \to (E_3), \quad \text{and} \quad (E_4 - E_1) \to (E_4),$$

gives

$$\tilde{A}^{(2)} = \left[\begin{array}{rrrr:r} 1 & -1 & 2 & -1 & -8 \\ 0 & 0 & -1 & -1 & -4 \\ 0 & 2 & -1 & 1 & 6 \\ 0 & 0 & 2 & 4 & 12 \end{array}\right].$$

The diagonal entry $a_{22}^{(2)}$, called the **pivot element**, is 0, so the procedure cannot continue in its present form. But operations $(E_i) \leftrightarrow (E_j)$ are permitted, so a search is made of the elements $a_{32}^{(2)}$ and $a_{42}^{(2)}$ for the first nonzero element. Since $a_{32}^{(2)} \neq 0$, the operation $(E_2) \leftrightarrow (E_3)$ is performed to obtain a new matrix,

$$\tilde{A}^{(2)'} = \left[\begin{array}{cccc:c} 1 & -1 & 2 & -1 & -8 \\ 0 & 2 & -1 & 1 & 6 \\ 0 & 0 & -1 & -1 & -4 \\ 0 & 0 & 2 & 4 & 12 \end{array}\right].$$

Since x_2 is already eliminated from E_3 and E_4, $\tilde{A}^{(3)}$ will be $\tilde{A}^{(2)'}$, and the computations continue with the operation $(E_4 + 2E_3) \rightarrow (E_4)$, giving

$$\tilde{A}^{(4)} = \left[\begin{array}{rrrr:r} 1 & -1 & 2 & -1 & -8 \\ 0 & 2 & -1 & 1 & 6 \\ 0 & 0 & -1 & -1 & -4 \\ 0 & 0 & 0 & 2 & 4 \end{array}\right].$$

Finally, the matrix is converted back into a linear system that has a solution equivalent to the solution of the original system and the backward substitution is applied:

$$x_4 = \frac{4}{2} = 2,$$

$$x_3 = \frac{[-4 - (-1)x_4]}{-1} = 2,$$

$$x_2 = \frac{[6 - x_4 - (-1)x_3]}{2} = 3,$$

$$x_1 = \frac{[-8 - (-1)x_4 - 2x_3 - (-1)x_2]}{1} = -7.$$

6.2 Pivoting Strategies

Problem

Apply Gaussian elimination to the system

$$
\begin{aligned}
E_1 &: \quad 0.003000x_1 + 59.14x_2 = 59.17 \\
E_2 &: \quad 5.291x_1 - 6.130x_2 = 46.78,
\end{aligned}
$$

using four-digit arithmetic with rounding, and compare the results to the exact solution $x_1 = 10.00$ and $x_2 = 1.000$.

Solution

The first pivot element, $a_{11}^{(1)} = 0.003000$, is small, and its associated multiplier,

$$m_{21} = \frac{5.291}{0.003000} = 1763.6\overline{6},$$

rounds to the large number 1764.

Perform $(E_2 - m_{21}E_1) \to (E_2)$:

Rounding with four-digit arithmetic:

Coefficient of x_2:

$$\begin{aligned} & -6.130 - 1764 \times 59.14 = -6.130 - 104322.96 \\ \approx & \quad -6.130 - 104300 = -104306.13 \\ \approx & \quad -104300. \end{aligned}$$

Right hand side:

$$
\begin{aligned}
& 46.78 - 1764 \times 59.17 = 46.78 - 104375.88 \\
\approx \; & 46.78 - 104400 = -104353.22 \\
\approx \; & -104400.
\end{aligned}
$$

New linear system which we have got by doing calculation with four-digit arithmetic with rounding is:

$$\begin{aligned} 0.003000x_1 + 59.14x_2 &\approx 59.17 \\ -104300x_2 &\approx -104400, \end{aligned}$$

Approximated solution:

$$\begin{aligned} x_2 &= \frac{104400}{104300} \approx 1.001, \\ x_1 &= \frac{59.17 - 59.14 \times 1.001}{0.003000} = \frac{59.17 - 59.19914}{0.003000} \\ &\approx \frac{59.17 - 59.20}{0.003000} = -10.00. \end{aligned}$$

This ruins the approximation to the actual value $x_1 = 10.00$.

Partial Pivoting

Previous problem shows how difficulties can arise when the pivot element $a_{kk}^{(k)}$ is small relative to the entries $a_{ij}^{(k)}$, for $k \le i \le n$ and $k \le j \le n$.

The simplest strategy is to select an element in the same column that is below the diagonal and has the largest absolute value; specifically, we determine the smallest $p \ge k$ such that

$$|a_{pk}^{(k)}| = \max_{k \le i \le n} |a_{ik}^{(k)}|$$

and perform $(E_k) \leftrightarrow (E_p)$. In this case no interchange of columns is used.

The technique just described is called **partial pivoting** (or *maximal column pivoting*)

Problem

Apply Gaussian elimination to the system

$$
\begin{aligned}
E_1 : &\quad 0.003000x_1 + 59.14x_2 = 59.17 \\
E_2 : &\quad 5.291x_1 - 6.130x_2 = 46.78,
\end{aligned}
$$

using partial pivoting and four-digit arithmetic with rounding, and compare the results to the exact solution $x_1 = 10.00$ and $x_2 = 1.000$.

Solution

The partial-pivoting procedure first requires finding

$$\max\left\{|a_{11}^{(1)}|, |a_{21}^{(1)}|\right\} = \max\{|0.003000|, |5.291|\} = |5.291| = |a_{21}^{(1)}|.$$

This requires that the operation $(E_2) \leftrightarrow (E_1)$ be performed to produce the equivalent system

$$\begin{aligned} E_1: \quad & 5.291x_1 - 6.130x_2 = 46.78, \\ E_2: \quad & 0.003000x_1 + 59.14x_2 = 59.17. \end{aligned}$$

The multiplier for this system is

$$m_{21} = \frac{a_{21}^{(1)}}{a_{11}^{(1)}} = 0.0005670,$$

and the operation $(E_2 - m_{21}E_1) \rightarrow (E_2)$ reduces the system to

$$\begin{aligned} 5.291x_1 - 6.130x_2 &\approx 46.78, \\ 59.14x_2 &\approx 59.14. \end{aligned}$$

The four-digit answers resulting from the backward substitution are the correct values $x_1 = 10.00$ and $x_2 = 1.000$.

Note

Each multiplier m_{ji} in the partial pivoting algorithm has magnitude less than or equal to 1. Although this strategy is sufficient for many linear systems, situations do arise when it is inadequate.

Illustration

Consider the linear system

$$
\begin{aligned}
E_1: &\quad 30.00x_1 + 591400x_2 = 591700, \\
E_2: &\quad 5.291x_1 - \quad 6.130x_2 = 46.78,
\end{aligned}
$$

The above system is the same as that in previous problem except that all entries in the first equation is multiplied by 10^4.

Let's solve the system by Gauss Elimination using partial pivoting and four-digit arithmetic with rounding.

The maximal value in the first column is 30.00, and the multiplier

$$m_{21} = \frac{5.291}{30.00} = 0.1764$$

leads to the system

$$\begin{aligned} 30.00x_1 + 591400x_2 &\approx 591700, \\ -104300x_2 &\approx -104400, \end{aligned}$$

which has the same inaccurate solutions

$$x_2 \approx 1.001 \text{ and } x_1 \approx -10.00.$$

Scaled Partial Pivoting

Scaled partial pivoting (or *scaled-column pivoting*) is needed for the system in the Illustration. It places the element in the pivot position that is largest relative to the entries in its row. The first step in this procedure is to define a scale factor s_i for each row as

$$s_i = \max_{1 \le j \le n} |a_{ij}|$$

If we have $s_i = 0$ for some i, then the system has no unique solution since all entries in the ith row are 0. Assuming that this is not the case, the appropriate row interchange to place zeros in the first column is determined by choosing the least integer p with

$$\frac{|a_{p1}|}{s_p} = \max_{1 \le k \le n} \frac{|a_{k1}|}{s_k}$$

and performing $(E_1) \leftrightarrow (E_p)$. The effect of scaling is to ensure that the largest element in each row has a *relative* magnitude of 1 before the comparison for row interchange is performed.

In a similar manner, before eliminating the variable x_i using the operations

$$E_k - m_{ki}E_i, \quad \text{for } k = i+1, \ldots, n,$$

we select the smallest integer $p \geq i$ with

$$\frac{|a_{pi}|}{s_p} = \max_{i \leq k \leq n} \frac{|a_{ki}|}{s_k}$$

Note that the scale factor s_1, s_2,…s_n is row dependent, so they must be interchanged when the row interchanges are performed.

Solution to the problem in illustration

Applying scaled partial pivoting to the previous Illustration gives

$$s_1 = \max\{|30.00|, |591400|\} = 591400$$

and

$$s_2 = \max\{|5.291|, |-6.130|\} = 6.130.$$

Consequently

$$\frac{|a_{11}|}{s_1} = \frac{30.00}{591400} = 0.5073 \times 10^{-4}, \qquad \frac{|a_{21}|}{s_2} = \frac{5.291}{6.130} = 0.8631,$$

and the interchange $(E_1) \leftrightarrow (E_2)$ is made.

Applying Gaussian elimination to the new system

$$5.291x_1 - 6.130x_2 = 46.78$$

$$30.00x_1 + 591400x_2 = 591700$$

produces the correct results: $x_1 = 10.00$ and $x_2 = 1.000$.

Practice Problems

Problem 1

Use Gaussian elimination and three-digit chopping arithmetic to solve the following linear system and compare the approximations to the actual solution.

$$
\begin{aligned}
3.03x_1 - 12.1x_2 + 14x_3 &= -119, \\
-3.03x_1 + 12.1x_2 - 7x_3 &= 120, \\
6.11x_1 - 14.2x_2 + 21x_3 &= -139.
\end{aligned}
$$

Actual solution $[0, 10, \frac{1}{7}]$.

Problem 2

Use Gaussian elimination and three-digit chopping arithmetic to solve the following linear system and compare the approximations to the actual solution.

$$0.03x_1 + 58.9x_2 = 59.2,$$

$$5.31x_1 - 6.10x_2 = 47.0.$$

Actual solution [10, 1].

Also solve the system using Gaussian elimination with partial pivoting and three-digit rounding arithmetic.

Problem 3

Use Gaussian elimination with scaled partial pivoting and three-digit rounding arithmetic to solve the following linear system and compare the approximations to the actual solution.

$$\begin{aligned}
3.3330x_1 + 15920x_2 + 10.333x_3 &= 7953, \\
2.2220x_1 + 16.710x_2 + 9.6120x_3 &= 0.965, \\
-1.5611x_1 + 5.1792x_2 - 1.6855x_3 &= 2.714.
\end{aligned}$$

Actual solution $[1, 0.5, -1]$.

Diagonally Dominant Matrices

The $n \times n$ matrix A is said to be **diagonally dominant** when

$$|a_{ii}| \geq \sum_{\substack{j=1,\\ j\neq i}}^{n} |a_{ij}| \quad \text{holds for each } i = 1, 2, \cdots, n. \tag{6.10}$$

In other words, a diagonally dominant matrix is one where the absolute value of each diagonal entry is greater than or equal to the sum of the absolute values of all the other entries in that row.

A diagonally dominant matrix is said to be **strictly diagonally dominant** when the inequality in (6.10) is strict for each n, that is, when

$$|a_{ii}| > \sum_{\substack{j=1,\\ j\neq i}}^{n} |a_{ij}| \quad \text{holds for each } i = 1, 2, \cdots, n.$$

Example 1

The matrix $A = \begin{bmatrix} 3 & -2 & 1 \\ 1 & -3 & 2 \\ -1 & 2 & 4 \end{bmatrix}$ is diagonally dominant

because

$|a_{11}| \geq |a_{12}| + |a_{13}|$ since $|3| \geq |-2| + |1|$ (True)

$|a_{22}| \geq |a_{21}| + |a_{23}|$ since $|-3| \geq |1| + |2|$ (True)

$|a_{33}| \geq |a_{31}| + |a_{32}|$ since $|4| \geq |-1| + |2|$. (True)

But A is not strictly diagonally dominant because $|a_{11}| \ngtr |a_{12}| + |a_{13}|$

Example 2

Consider the matrices

$$A = \begin{bmatrix} 7 & 2 & 0 \\ 3 & 5 & -1 \\ 0 & 5 & -6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 6 & 4 & -3 \\ 4 & -2 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

The matrix A is strictly diagonally dominant because

$$|7| > |2| + |0|, \quad |5| > |3| + |-1|, \quad \text{and} \quad |-6| > |0| + |5|.$$

The matrix B is not strictly diagonally dominant because, for example, in the first row the absolute value of the diagonal element is $|6| < |4| + |-3| = 7$.

Positive Definite Matrices

A matrix A is **positive definite** if it is symmetric and if $\mathbf{x}^t A\mathbf{x} > 0$ for every n-dimensional vector $\mathbf{x} \neq \mathbf{0}$.

To be precise, above Definition specify that the 1×1 matrix generated by the operation $\mathbf{x}^t A\mathbf{x}$ has a positive value for its only entry since the operation is performed as follows:

$$\mathbf{x}^t A\mathbf{x} = [x_1, x_2, \cdots, x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

The name positive definite refers to the fact that the number $\mathbf{x}^t A \mathbf{x}$ must be positive whenever $\mathbf{x} \neq \mathbf{0}$.

Note: Not all authors require symmetry of a positive definite matrix.

Problem

Show that the matrix $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ is positive definite.

Solution

Suppose $\mathbf{x}$ is any three-dimensional column vector.

Then

$$\begin{aligned}
\mathbf{x}^t A \mathbf{x} &= [x_1, x_2, x_3] \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= [x_1, x_2, x_3] \begin{bmatrix} 2x_1 & - & x_2 & & \\ -x_1 & + & 2x_2 & - & x_3 \\ -x_2 & + & 2x_3 & & \end{bmatrix} \\
&= 2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2.
\end{aligned}$$

Rearranging the terms gives

$$\begin{aligned}\mathbf{x}^t A\mathbf{x} &= x_1^2 + (x_1^2 - 2x_1x_2 + x_2^2) + (x_2^2 - 2x_2x_3 + x_3^2) + x_3^2 \\ &= x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2,\end{aligned}$$

which implies that

$$x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 > 0$$

unless $x_1 = x_2 = x_3 = 0$.

Thus A is a positive definite matrix.

Theorem

A symmetric matrix A is positive definite if and only if all the eigenvalues of A are positive.

Definition

A **leading principal submatrix** of a matrix A is a matrix of the form

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix},$$

for some $1 \leq k \leq n$.

Theorem

A symmetric matrix A is positive definite if and only if each of its leading principal submatrices has a positive determinant.

Problem

Show that the matrix $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ is positive definite.

Solution

$$\det A_1 = \det[2] = 2 > 0,$$

$$\det A_2 = \det\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 4 - 1 = 3 > 0,$$

and

$$\det A_3 = \det\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 2\det\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} - (-1)\det\begin{bmatrix} -1 & -1 \\ 0 & 2 \end{bmatrix}$$

$$= 2(4-1) + (-2+0) = 4 > 0.$$ Thus A is positive definite.

Method for Solving a linear system by LU-Decomposition

Suppose that A has been factored into the triangular form $A = LU$, where L is lower triangular and U is upper triangular. Then we can solve a linear system $A\mathbf{x} = \mathbf{b}$ more easily by using a two-step process.

- First we let $\mathbf{y} = U\mathbf{x}$ and solve the lower triangular system $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$.
- Once $\mathbf{y}$ is known, then solve the upper triangular system $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$.

Why LU Factorization?

The LU method is particularly useful when it is necessary to solve a whole series of systems

$$A\mathbf{x} = \mathbf{b}_1,\ A\mathbf{x} = \mathbf{b}_2,\ \ldots,\ A\mathbf{x} = \mathbf{b}_n$$

each of which has the same square coefficient matrix A.

Doolittle, Crout and Cholesky Factorization

Definition (LU-Factorization):

The nonsingular matrix **A** has an LU-factorization if it can be expressed as the product of a lower-triangular matrix **L** and an upper triangular matrix **U.**

When this is possible we say that **A** has an **LU**-decomposition. It turns out that this factorization (when it exists) is not unique.

If **L** has 1's on its diagonal, then it is called a **Doolittle factorization.**

If **U** has 1's on its diagonal, then it is called a **Crout factorization.**

If $\boldsymbol{U} = \boldsymbol{L}^{t}$, then it is called **Cholesky factorization.**

Theorem

If all n leading principal submatrix of an $n \times n$ matrix A are nonsingular, then A has an LU-decomposition.

Theorem

Every diagonally dominant matrix is nonsingular and has an LU-factorization.

Theorem (Cholesky Factorization)

The matrix A is positive definite if and only if A can be factored in the form LL^t, where L is lower triangular with nonzero diagonal entries.

Problem

Determine the Cholesky LL^t factorization of the positive definite matrix

$$A = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 4.25 & 2.75 \\ 1 & 2.75 & 3.5 \end{bmatrix}.$$

Solution

The LL^t factorization does not necessarily has 1s on the diagonal of the lower triangular matrix L so we need to have

$$A = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

$$= \begin{bmatrix} l_{11}^2 & l_{11}l_{21} & l_{11}l_{31} \\ l_{11}l_{21} & l_{21}^2 + l_{22}^2 & l_{21}l_{31} + l_{22}l_{32} \\ l_{11}l_{31} & l_{21}l_{31} + l_{22}l_{32} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix}$$

Thus

$$a_{11}: \quad 4 = l_{11}^2 \implies l_{11} = 2, \qquad a_{21}: \quad -1 = l_{11}l_{21} \implies l_{21} = -0.5$$

$$a_{31}: \quad 1 = l_{11}l_{31} \implies l_{31} = 0.5, \qquad a_{22}: \quad 4.25 = l_{21}^2 + l_{22}^2 \implies l_{22} = 2$$

$$a_{32}: \quad 2.75 = l_{21}l_{31} + l_{22}l_{32} \implies l_{32} = 1.5, \quad a_{33}: \quad 3.5 = l_{31}^2 + l_{32}^2 + l_{33}^2 \implies l_{33} = 1,$$

and we have

$$A = LL^t = \begin{bmatrix} 2 & 0 & 0 \\ -0.5 & 2 & 0 \\ 0.5 & 1.5 & 1 \end{bmatrix} \begin{bmatrix} 2 & -0.5 & 0.5 \\ 0 & 2 & 1.5 \\ 0 & 0 & 1 \end{bmatrix}.$$

Problem Use Crout's method to solve the system

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 5 \\ -1 & 1 & -5 & 3 \\ 3 & 1 & 7 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 10 \\ 31 \\ -2 \\ 18. \end{bmatrix}$$

Solution

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 5 \\ -1 & 1 & -5 & 3 \\ 3 & 1 & 7 & -2 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

By multiplying L with U and comparing the elements of the product matrix with those of A, we obtain:

(i) Multiplication of the first row of L with the columns of U gives

$$
\begin{aligned}
&l_{11} = 1, \\
&l_{11}u_{12} = 1 \Longrightarrow u_{12} = 1, \\
&l_{11}u_{13} = 1 \Longrightarrow u_{13} = 1, \\
&l_{11}u_{14} = 1 \Longrightarrow u_{14} = 1.
\end{aligned}
$$

(ii) Multiplication of the second row of L with the columns of U gives

$$
\begin{aligned}
&l_{21} = 2, \\
&l_{21}u_{12} + l_{22} = 3 \Longrightarrow l_{22} = 3 - l_{21}u_{12} = 1, \\
&l_{21}u_{13} + l_{22}u_{23} = 1 \Longrightarrow u_{23} = (1 - l_{21}u_{13})/l_{22} = -1, \\
&l_{21}u_{14} + l_{22}u_{24} = 5 \Longrightarrow u_{24} = (5 - l_{21}u_{14})/l_{22} = 3.
\end{aligned}
$$

(iii) Multiplication of the third row of L with the columns of U gives

$$
\begin{aligned}
&l_{31} = -1, \\
&l_{31}u_{12} + l_{32} = 1 \Longrightarrow l_{32} = 1 - l_{31}u_{12} = 2, \\
&l_{31}u_{13} + l_{32}u_{23} + l_{33} = -5 \Longrightarrow l_{33} = -5 - l_{31}u_{13} - l_{32}u_{23} = -2, \\
&l_{31}u_{14} + l_{32}u_{24} + l_{33}u_{34} = 3 \Longrightarrow u_{34} = (3 - l_{31}u_{14} - l_{32}u_{24})/l_{33} = 1.
\end{aligned}
$$

(iv) Multiplication of the fourth row of L with the columns of U gives

$$
\begin{aligned}
& l_{41} = 3, \\
& l_{41}u_{12} + l_{42} = 1 \implies l_{42} = 1 - l_{41}u_{12} = -2, \\
& l_{41}u_{13} + l_{42}u_{23} + l_{43} = 7 \implies l_{43} = 7 - l_{41}u_{13} - l_{42}u_{23} = 2, \\
& l_{41}u_{14} + l_{42}u_{24} + l_{43}u_{34} + l_{44} = -2 \implies \\
& l_{44} = -2 - l_{41}u_{14} - l_{42}u_{24} - l_{43}u_{34} = -1.
\end{aligned}
$$

Hence,

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 2 & -2 & 0 \\ 3 & -2 & 2 & -1 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

By applying the forward substitution to the lower triangular system $L\mathbf{y} = \mathbf{b}$, we get

$$\begin{aligned} y_1 &= 10 \\ y_2 &= 31 - 2(10) = 11 \\ y_3 &= [-2 + 10 - 2(11)]/(-2) = 7 \\ y_4 &= -[18 - 3(10) + 2(11) - 2(7)] = 4 \end{aligned}$$

Finally, by applying the back substitution to the upper triangular system $U\mathbf{x} = \mathbf{y}$, we get

$$
\begin{aligned}
x_1 &= 10 - 4 - 3 - 2 = 1 \\
x_2 &= -[11 - 4 - 3(3)] = 2 \\
x_3 &= 7 - 4 = 3 \\
x_4 &= 4.
\end{aligned}
$$

Practice Problems

Problem 1

Determine which of the following matrices are (i) symmetric, (ii) singular, (iii) strictly diagonally dominant, (iv) positive definite.

a. $\begin{bmatrix} -2 & 1 \\ 1 & -3 \end{bmatrix}$

b. $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 2 \\ 1 & 2 & 4 \end{bmatrix}$

c. $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 2 \end{bmatrix}$

d. $\begin{bmatrix} 2 & 3 & 1 & 2 \\ -2 & 4 & -1 & 5 \\ 3 & 7 & 1.5 & 1 \\ 6 & -9 & 3 & 7 \end{bmatrix}$

Problem 2

Let

$$A = \begin{bmatrix} \alpha & 1 & 0 \\ \beta & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

Find all values of α and β for which

a. A is singular.

b. A is strictly diagonally dominant.

c. A is symmetric.

d. A is positive definite.

Problem 3

Find the Cholesky decomposition—$A = LL^T$—of the matrix

$$A = \begin{pmatrix} 2 & -2 & -3 \\ -2 & 5 & 4 \\ -3 & 4 & 5 \end{pmatrix}$$

and hence, solve the system of equations $Ax = b$, where $b = \begin{pmatrix} 7 \\ -12 \\ -12 \end{pmatrix}$.

Problem 3 Answers:

$$L = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ -\sqrt{2} & \sqrt{3} & 0 \\ -3/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{pmatrix}$$

Required Sol. Of Linear System $\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$

Iterative Techniques for Solving a Linear System (Indirect Methods)

Vector Norms

Let $\mathbb{R}^n$ denote the set of all n-dimensional column vectors with real-number components. To define a distance in $\mathbb{R}^n$ we use the notion of a norm, which is the generalization of the absolute value on $\mathbb{R}$, the set of real numbers.

Notation

Vectors in $\mathbb{R}^n$ are column vectors, and it is convenient to use the transpose notation presented when a vector is represented in terms of its components. For example, the vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

will be written $\mathbf{x} = (x_1, x_2, \dots, x_n)^t$.

Definition

The l_2 and l_∞ norms for the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ are defined by

$$\|\mathbf{x}\|_2 = \left\{\sum_{i=1}^{n} x_i^2\right\}^{1/2} \quad \text{and} \quad \|\mathbf{x}\|_\infty = \max_{1\le i\le n} |x_i|.$$

Note that each of these norms reduces to the absolute value in the case $n = 1$.

Problem

Determine the l_2 norm and the l_∞ norm of the vector $\mathbf{x} = (-1, 1, -2)^t$.

Solution

The vector $\mathbf{x} = (-1, 1, -2)^t$ in $\mathbb{R}^3$ has norms

$$\|\mathbf{x}\|_2 = \sqrt{(-1)^2 + (1)^2 + (-2)^2} = \sqrt{6}$$

and

$$\|\mathbf{x}\|_\infty = \max\{|-1|, |1|, |-2|\} = 2.$$

The l_2 norm is called the **Euclidean norm** of the vector $\mathbf{x}$ because it represents the usual notion of distance from the origin in case $\mathbf{x}$ is in $\mathbb{R}^1 \equiv \mathbb{R}$, $\mathbb{R}^2$, or $\mathbb{R}^3$.

Distance between Vectors in $\mathbb{R}^n$

The norm of a vector gives a measure for the distance between an arbitrary vector and the zero vector, just as the absolute value of a real number describes its distance from 0. Similarly, the **distance between two vectors** is defined as the norm of the difference of the vectors just as distance between two real numbers is the absolute value of their difference.

Definition

If $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^t$ are vectors in $\mathbb{R}^n$, the l_2 and l_∞ distances between $\mathbf{x}$ and $\mathbf{y}$ are defined by

$$\|\mathbf{x} - \mathbf{y}\|_2 = \left\{ \sum_{i=1}^{n} (x_i - y_i)^2 \right\}^{1/2} \quad \text{and} \quad \|\mathbf{x} - \mathbf{y}\|_\infty = \max_{1 \le i \le n} |x_i - y_i|.$$

Problem

The linear system

$$3.3330x_1 + 15920x_2 - 10.333x_3 = 15913,$$

$$2.2220x_1 + 16.710x_2 + 9.6120x_3 = 28.544,$$

$$1.5611x_1 + 5.1791x_2 + 1.6852x_3 = 8.4254$$

has the exact solution $\mathbf{x} = (x_1, x_2, x_3)^t = (1, 1, 1)^t$, and Gaussian elimination performed using five-digit rounding arithmetic and partial pivoting , produces the approximate solution

$$\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)^t = (1.2001, 0.99991, 0.92538)^t.$$

Determine the l_2 and l_∞ distances between the exact and approximate solutions.

Solution

Measurements of $\mathbf{x} - \tilde{\mathbf{x}}$ are given by

$$\|\mathbf{x} - \tilde{\mathbf{x}}\|_\infty = \max\{|1 - 1.2001|, |1 - 0.99991|, |1 - 0.92538|\}$$
$$= \max\{0.2001, 0.00009, 0.07462\} = 0.2001$$

and

$$\|\mathbf{x} - \tilde{\mathbf{x}}\|_2 = \left[(1 - 1.2001)^2 + (1 - 0.99991)^2 + (1 - 0.92538)^2\right]^{1/2}$$
$$= [(0.2001)^2 + (0.00009)^2 + (0.07462)^2]^{1/2} = 0.21356.$$

Although the components $\tilde{x}_2$ and $\tilde{x}_3$ are good approximations to x_2 and x_3, the component $\tilde{x}_1$ is a poor approximation to x_1, and $|x_1 - \tilde{x}_1|$ dominates both norms.

An iterative technique to solve the $n \times n$ linear system $A\mathbf{x} = \mathbf{b}$ starts with an initial approximation $\mathbf{x}^{(0)}$ to the solution $\mathbf{x}$ and generates a sequence of vectors $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ that converges to $\mathbf{x}$.

Jacobi's Method

The **Jacobi iterative method** is obtained by solving the ith equation in $A\mathbf{x} = \mathbf{b}$ for x_i to obtain (provided $a_{ii} \neq 0$)

$$x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \left(-\frac{a_{ij}x_j}{a_{ii}} \right) + \frac{b_i}{a_{ii}}, \qquad \text{for } i = 1, 2, \ldots, n.$$

For each $k \geq 1$, generate the components $x_i^{(k)}$ of $\mathbf{x}^{(k)}$ from the components of $\mathbf{x}^{(k-1)}$ by

$$x_i^{(k)} = \frac{1}{a_{ii}} \left[\sum_{\substack{j=1 \\ j \neq i}}^{n} \left(-a_{ij} x_j^{(k-1)} \right) + b_i \right], \qquad \text{for } i = 1, 2, \ldots, n.$$

Problem

The linear system $A\mathbf{x} = \mathbf{b}$ given by

$$
\begin{aligned}
E_1: &\quad 10x_1 - x_2 + 2x_3 &&= 6,\\
E_2: &\quad -x_1 + 11x_2 - x_3 + 3x_4 &&= 25,\\
E_3: &\quad 2x_1 - x_2 + 10x_3 - x_4 &&= -11,\\
E_4: &\quad 3x_2 - x_3 + 8x_4 &&= 15
\end{aligned}
$$

has the unique solution $\mathbf{x} = (1, 2, -1, 1)^t$. Use Jacobi's iterative technique to find approximations $\mathbf{x}^{(k)}$ to $\mathbf{x}$ starting with $\mathbf{x}^{(0)} = (0, 0, 0, 0)^t$ until

$$
\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty}{\|\mathbf{x}^{(k)}\|_\infty} < 10^{-3}.
$$

Solution

We first solve equation E_i for x_i, for each $i = 1, 2, 3, 4$, to obtain

$$\begin{aligned}
x_1 &= \phantom{-\frac{1}{5}x_1 +} \frac{1}{10}x_2 - \frac{1}{5}x_3 \phantom{+ \frac{1}{10}x_4} + \frac{3}{5}, \\
x_2 &= \frac{1}{11}x_1 \phantom{+ \frac{1}{10}x_2} + \frac{1}{11}x_3 - \frac{3}{11}x_4 + \frac{25}{11}, \\
x_3 &= -\frac{1}{5}x_1 + \frac{1}{10}x_2 \phantom{+ \frac{1}{11}x_3} + \frac{1}{10}x_4 - \frac{11}{10}, \\
x_4 &= \phantom{-\frac{1}{5}x_1} - \frac{3}{8}x_2 + \frac{1}{8}x_3 \phantom{+ \frac{1}{10}x_4} + \frac{15}{8}.
\end{aligned}$$

So we have the following scheme for calculating approximate solution vectors

$$
\begin{aligned}
x_1^{(k)} &= \frac{1}{10}x_2^{(k-1)} - \frac{1}{5}x_3^{(k-1)} + \frac{3}{5},\\
x_2^{(k)} &= \frac{1}{11}x_1^{(k-1)} + \frac{1}{11}x_3^{(k-1)} - \frac{3}{11}x_4^{(k-1)} + \frac{25}{11},\\
x_3^{(k)} &= -\frac{1}{5}x_1^{(k-1)} + \frac{1}{10}x_2^{(k-1)} + \frac{1}{10}x_4^{(k-1)} - \frac{11}{10},\\
x_4^{(k)} &= -\frac{3}{8}x_2^{(k-1)} + \frac{1}{8}x_3^{(k-1)} + \frac{15}{8}.
\end{aligned}
$$

From the initial approximation $\mathbf{x}^{(0)} = (0, 0, 0, 0)^t$ we have $\mathbf{x}^{(1)}$ given by

$$\begin{aligned}
x_1^{(1)} &= \frac{1}{10}x_2^{(0)} - \frac{1}{5}x_3^{(0)} + \frac{3}{5} = 0.6000, \\
x_2^{(1)} &= \frac{1}{11}x_1^{(0)} + \frac{1}{11}x_3^{(0)} - \frac{3}{11}x_4^{(0)} + \frac{25}{11} = 2.2727, \\
x_3^{(1)} &= -\frac{1}{5}x_1^{(0)} + \frac{1}{10}x_2^{(0)} + \frac{1}{10}x_4^{(0)} - \frac{11}{10} = -1.1000, \\
x_4^{(1)} &= -\frac{3}{8}x_2^{(0)} + \frac{1}{8}x_3^{(0)} + \frac{15}{8} = 1.8750.
\end{aligned}$$

Additional iterates, $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, x_4^{(k)})^t$, are generated in a similar manner and are presented in Table 7.1.

Table 7.1

k	0	1	2	3	4	5	6	7	8	9	10
$x_1^{(k)}$	0.0000	0.6000	1.0473	0.9326	1.0152	0.9890	1.0032	0.9981	1.0006	0.9997	1.0001
$x_2^{(k)}$	0.0000	2.2727	1.7159	2.053	1.9537	2.0114	1.9922	2.0023	1.9987	2.0004	1.9998
$x_3^{(k)}$	0.0000	−1.1000	−0.8052	−1.0493	−0.9681	−1.0103	−0.9945	−1.0020	−0.9990	−1.0004	−0.9998
$x_4^{(k)}$	0.0000	1.8750	0.8852	1.1309	0.9739	1.0214	0.9944	1.0036	0.9989	1.0006	0.9998

We stopped after ten iterations because

$$\frac{\|\mathbf{x}^{(10)} - \mathbf{x}^{(9)}\|_\infty}{\|\mathbf{x}^{(10)}\|_\infty} = \frac{8.0 \times 10^{-4}}{1.9998} < 10^{-3}.$$

In fact, $\|\mathbf{x}^{(10)} - \mathbf{x}\|_\infty = 0.0002$.

The Gauss-Seidel Method

One can see a possible improvement in the Jacobi's Method by examining its formula. The components of $\mathbf{x}^{(k-1)}$ are used to compute all the components $x_i^{(k)}$ of $\mathbf{x}^{(k)}$. But, for $i > 1$, the components $x_1^{(k)}, \ldots, x_{i-1}^{(k)}$ of $\mathbf{x}^{(k)}$ have already been computed and are expected to be better approximations to the actual solutions $x_1, \ldots, x_{i-1}$ than are $x_1^{(k-1)}, \ldots, x_{i-1}^{(k-1)}$. It seems reasonable, then, to compute $x_i^{(k)}$ using these most recently calculated values. That is, to use

$$x_i^{(k)} = \frac{1}{a_{ii}} \left[-\sum_{j=1}^{i-1} (a_{ij} x_j^{(k)}) - \sum_{j=i+1}^{n} (a_{ij} x_j^{(k-1)}) + b_i \right],$$

for each $i = 1, 2, \ldots, n$

This modification is called the **Gauss-Seidel iterative technique**

Problem

Use the Gauss-Seidel iterative technique to find approximate solutions to

$$
\begin{aligned}
10x_1 - \quad x_2 + \quad 2x_3 \qquad\qquad &= 6, \\
-x_1 + 11x_2 - \quad x_3 + 3x_4 &= 25, \\
2x_1 - \quad x_2 + 10x_3 - \quad x_4 &= -11, \\
3x_2 - \quad x_3 + 8x_4 &= 15
\end{aligned}
$$

starting with $\mathbf{x} = (0, 0, 0, 0)^t$ and iterating until

$$\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty}{\|\mathbf{x}^{(k)}\|_\infty} < 10^{-3}.$$

Solution

For the Gauss-Seidel method we write the system, for each $k = 1, 2, \ldots$ as

$$
\begin{aligned}
x_1^{(k)} &= \qquad\qquad\quad \frac{1}{10}x_2^{(k-1)} - \frac{1}{5}x_3^{(k-1)} \qquad\qquad + \frac{3}{5}, \\
x_2^{(k)} &= \frac{1}{11}x_1^{(k)} \qquad\qquad + \frac{1}{11}x_3^{(k-1)} - \frac{3}{11}x_4^{(k-1)} + \frac{25}{11}, \\
x_3^{(k)} &= -\frac{1}{5}x_1^{(k)} + \frac{1}{10}x_2^{(k)} \qquad\qquad + \frac{1}{10}x_4^{(k-1)} - \frac{11}{10}, \\
x_4^{(k)} &= \qquad\quad - \frac{3}{8}x_2^{(k)} + \frac{1}{8}x_3^{(k)} \qquad\qquad + \frac{15}{8}.
\end{aligned}
$$

When $\mathbf{x}^{(0)} = (0, 0, 0, 0)^t$, we have $\mathbf{x}^{(1)} = (0.6000, 2.3272, -0.9873, 0.8789)^t$.

Subsequent iterations give the values in below Table

k	0	1	2	3	4	5
$x_1^{(k)}$	0.0000	0.6000	1.030	1.0065	1.0009	1.0001
$x_2^{(k)}$	0.0000	2.3272	2.037	2.0036	2.0003	2.0000
$x_3^{(k)}$	0.0000	−0.9873	−1.014	−1.0025	−1.0003	−1.0000
$x_4^{(k)}$	0.0000	0.8789	0.9844	0.9983	0.9999	1.0000

Because

$$\frac{\|\mathbf{x}^{(5)} - \mathbf{x}^{(4)}\|_\infty}{\|\mathbf{x}^{(5)}\|_\infty} = \frac{0.0008}{2.000} = 4 \times 10^{-4},$$

$\mathbf{x}^{(5)}$ is accepted as a reasonable approximation to the solution.

Note that Jacobi's method in previous problem required twice as many iterations for the same accuracy.

Theorem

If A is strictly diagonally dominant, then for any choice of $\mathbf{x}^{(0)}$, both the Jacobi and Gauss-Seidel methods give sequences $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ that converge to the unique solution of $A\mathbf{x} = \mathbf{b}$.

Practice Problems

Problem 1

Find the first two iterations of the Jacobi method for the following linear systems, using $\mathbf{x}^{(0)} = \mathbf{0}$:

a.
$$\begin{aligned} 3x_1 - x_2 + x_3 &= 1, \\ 3x_1 + 6x_2 + 2x_3 &= 0, \\ 3x_1 + 3x_2 + 7x_3 &= 4. \end{aligned}$$

b.
$$\begin{aligned} 10x_1 + 5x_2 \qquad\qquad &= 6, \\ 5x_1 + 10x_2 - 4x_3 \qquad &= 25, \\ -4x_2 + 8x_3 - x_4 &= -11, \\ -x_3 + 5x_4 &= -11. \end{aligned}$$

Problem 2

Use the Jacobi method to solve the linear systems in Problem 1, with $TOL = 10^{-3}$ in the l_∞ norm.

Problem 3

Use the **Gauss-Seidel** method to solve the linear systems in Problem 1, with $TOL = 10^{-3}$ in the l_∞ norm.

Interpolation

Forward Differences

If $y_0, y_1, y_2, \ldots, y_n$ denote a set of values of y, then $y_1 - y_0, y_2 - y_1, \ldots, y_n - y_{n-1}$ are called the *differences* of y. Denoting these differences by $\Delta y_0, \Delta y_1, \ldots, \Delta y_{n-1}$ respectively, we have

$$\Delta y_0 = y_1 - y_0, \qquad \Delta y_1 = y_2 - y_1, \ldots, \qquad \Delta y_{n-1} = y_n - y_{n-1},$$

where Δ is called the *forward difference operator* and $\Delta y_0, \Delta y_1, \ldots,$ are called *first forward differences*. The differences of the first forward differences are called *second forward differences* and are denoted by $\Delta^2 y_0, \Delta^2 y_1, \ldots$ Similarly, one can define *third forward differences*, *fourth forward differences*, etc.

Note: The symbol Δ is read as 'delta'

Thus,

$$\Delta^2 y_0 = \Delta \Delta y_0 = \Delta (y_1 - y_0) = \Delta y_1 - \Delta y_0$$

$$= y_2 - y_1 - (y_1 - y_0)$$

$$\Rightarrow \boxed{\Delta^2 y_0 = y_2 - 2y_1 + y_0,}$$

$$\Delta^3 y_0 = \Delta^2 \Delta y_0 = \Delta^2 (y_1 - y_0) = \Delta^2 y_1 - \Delta^2 y_0$$

$$= y_3 - 2y_2 + y_1 - (y_2 - 2y_1 + y_0)$$

$$\Rightarrow \quad \Delta^3 y_0 = y_3 - 3y_2 + 3y_1 - y_0$$

$$\begin{aligned}\Delta^4 y_0 = \Delta^3 y_1 - \Delta^3 y_0 &= y_4 - 3y_3 + 3y_2 - y_1 - (y_3 - 3y_2 + 3y_1 - y_0) \\ &= y_4 - 4y_3 + 6y_2 - 4y_1 + y_0.\end{aligned}$$

Next slide show an easy method for calculating these differences using a table.

Forward Difference Table

x	y_0	Δ	Δ^2	Δ^3	Δ^4	Δ^5	Δ^6
x_0	y_0						
		Δy_0					
x_1	y_1		$\Delta^2 y_0$				
		Δy_1		$\Delta^3 y_0$			
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$		
		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_0$	
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$		$\Delta^6 y_0$
		Δy_3		$\Delta^3 y_2$		$\Delta^5 y_1$	
x_4	y_4		$\Delta^2 y_3$		$\Delta^4 y_2$		
		Δy_4		$\Delta^3 y_3$			
x_5	y_5		$\Delta^2 y_4$				
		Δy_5					
x_6	y_6						

Problem Values of x (in degrees) and $\sin x$ are given in the following

x (in degrees)	$y = \sin x$
15	0.2588190
20	0.3420201
25	0.4226183
30	0.5
35	0.5735764
40	0.6427876

Construct the forward difference table.

Solution:

x	y	Δ	Δ^2	Δ^3	Δ^4	Δ^5
15	0.2588190					
		0.0832011				
20	0.3420201		−0.0026029			
		0.0805982		−0.0006136		
25	0.4226183		−0.0032165		0.0000248	
		0.0773817		−0.0005888		0.0000041
30	0.5		−0.0038053		0.0000289	
		0.0735764		−0.0005599		
35	0.5735764		−0.0043652			
		0.0692112				
40	0.6427876					

Backward Differences

The differences $y_1 - y_0$, $y_2 - y_1$, …, $y_n - y_{n-1}$ are called first *backward differences* if they are denoted by ∇y_1, ∇y_2, …, ∇y_n respectively, so that

$$\begin{array}{llll} \nabla y_1 = y_1 - y_0, & \nabla y_2 = y_2 - y_1, \\ \vdots \quad \vdots \quad \vdots & \quad \vdots \\ \nabla y_n = y_n - y_{n-1}, & \end{array}$$

where ∇ is called the *backward difference operator.* In a similar way, one can define backward differences of higher orders.

Note: The symbol ∇ is read as 'nabla' or 'Del'

Thus, we obtain

$$\nabla^2 y_2 = \nabla \nabla y_2 = \nabla(y_2 - y_1)$$

$$= \nabla y_2 - \nabla y_1$$

$$\Rightarrow \nabla^2 y_2 = y_2 - y_1 - (y_1 - y_0) = y_2 - 2y_1 + y_0,$$

$$\nabla^3 y_3 = \nabla^2 y_3 - \nabla^2 y_2$$

$$= y_3 - 3y_2 + 3y_1 - y_0, \text{ etc.}$$

Backward Difference Table

x	y	∇	∇^2	∇^3	∇^4	∇^5	∇^6
x_0	y_0						
x_1	y_1	∇y_1					
x_2	y_2	∇y_2	$\nabla^2 y_2$				
x_3	y_3	∇y_3	$\nabla^2 y_3$	$\nabla^3 y_3$			
x_4	y_4	∇y_4	$\nabla^2 y_4$	$\nabla^3 y_4$	$\nabla^4 y_4$		
x_5	y_5	∇y_5	$\nabla^2 y_5$	$\nabla^3 y_5$	$\nabla^4 y_5$	$\nabla^5 y_5$	
x_6	y_6	∇y_6	$\nabla^2 y_6$	$\nabla^3 y_6$	$\nabla^4 y_6$	$\nabla^5 y_6$	$\nabla^6 y_6$

Problem

Consider the following data:

x	**y**
1	24
3	120
5	336
7	720

Construct the Backward Difference Table.

Solution

x	y	∇	∇^2	∇^3
1	24			
3	120	120-24=96		
5	336	216	120	
7	720	384	168	48
	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$
	y_3	∇y_3	$\nabla^2 y_3$	$\nabla^3 y_3$

Definition of Interpolation

Interpolate (Dictionary Meaning): **To add something in the middle of a text, piece of music etc.**

Consider the following problem

x	y
1	24
3	120
5	336
6	Unknown
7	720

What is the value of y corresponding to x=6 ?

- The process of estimating the value of y for some x inside the given range of x-values is called **interpolation.**

- The process of estimating the value of y for some x outside the given range of x-values is called **extrapolation.**

INTERPOLATING FUNCTIONS

A function is said to interpolate a set of data points if it passes through those points. Suppose that a set of (x, y) data points has been collected, such as $(0, 1)$, $(2, 2)$, and $(3, 4)$. There is a parabola that passes through the three points, shown in Figure 3.1. This parabola is called the degree 2 interpolating polynomial passing through the three points.

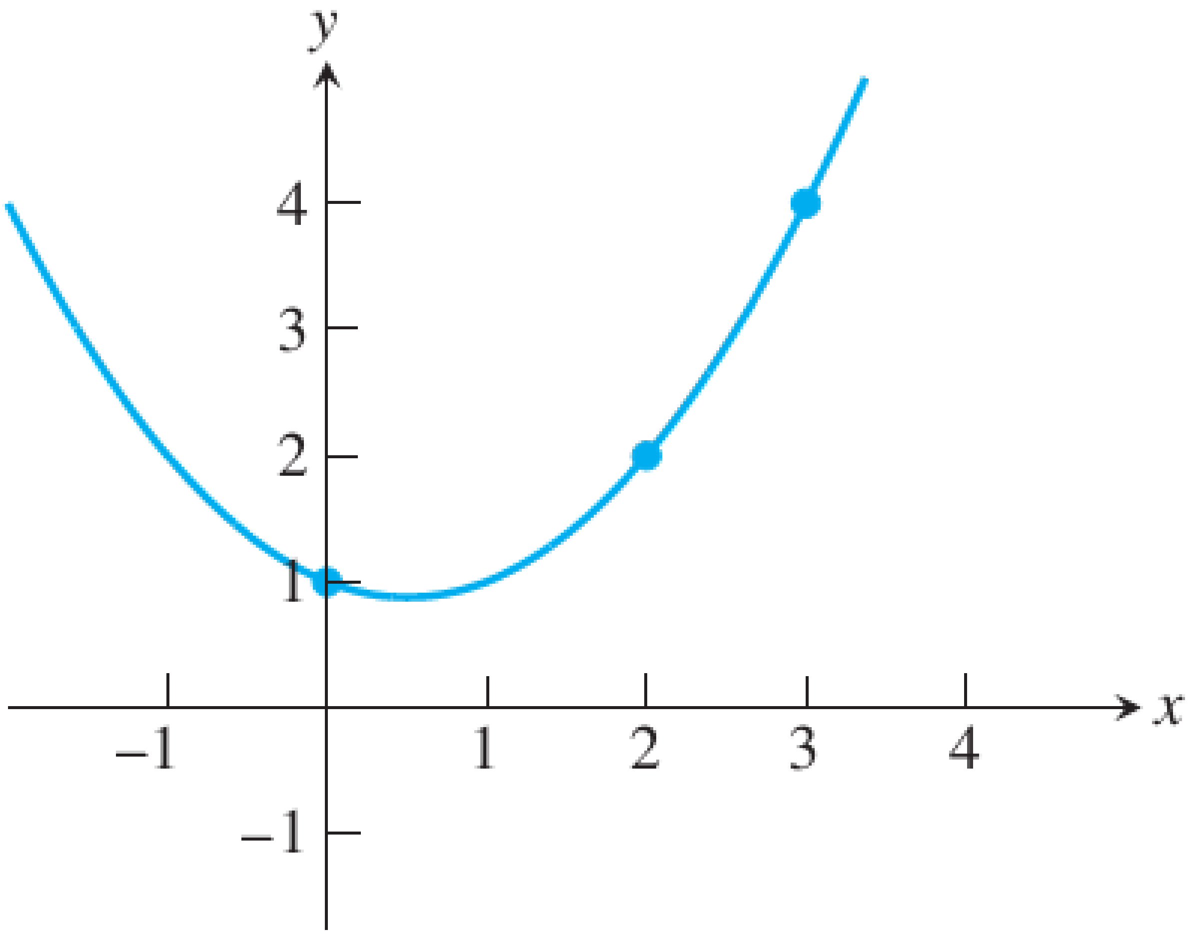


Figure 3.1 **Interpolation by parabola.** The points (0,1), (2,2), and (3,4) are interpolated by the function $P(x) = \frac{1}{2}x^2 - \frac{1}{2}x + 1$.

Polynomial Interpolation

In this section, we solve the following problem: We are given a table of $n + 1$ data points (x_i, y_i):

x	x_0	x_1	x_2	$\cdots$	x_n
y	y_0	y_1	y_2	$\cdots$	y_n

and we seek a polynomial p of lowest possible degree for which

$$p(x_i) = y_i \qquad (0 \leq i \leq n)$$

Such a polynomial is said to **interpolate** the data.

Theorem on Polynomial Interpolation

If $x_0, x_1, \ldots, x_n$ are distinct real numbers, then for arbitrary values $y_0, y_1, \ldots, y_n$, there is a unique polynomial p_n of degree at most n such that

$$p_n(x_i) = y_i \qquad (0 \leq i \leq n)$$

Problem 1

What is the maximum possible degree of the least degree interpolating polynomial that can be constructed from the following points

$$(1,2),\ (3,4),(5,6),(7,8),(9,10)$$

Ans: Since we are given 5 points so from Theorem of Polynomial Interpolation the maximum possible degree of the least degree interpolating polynomial is 4.

Problem 2

What are the possible degrees of the least degree interpolating polynomial that can be constructed from the following points

$$(1,2),\ (3,4),(5,6),(7,8),(9,10)$$

Ans: Since we are given 5 points so from Theorem of interpolating polynomial the possible degrees of the least degree interpolating polynomial are $0, 1, 2, 3,\ 4.$

Interpolation with Equally Spaced Data

Consider the following data

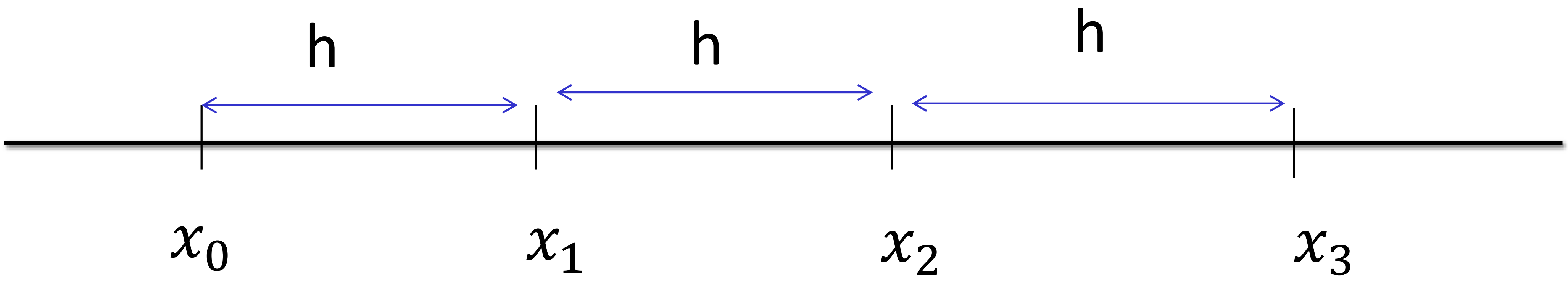
x	y
x_0	y_0
x_1	y_1
x_2	y_2
x_3	y_3

Let $x_0 < x_1 < x_2 < x_3$ then we say that the above data is equally spaced if $x_1 - x_0 = x_2 - x_1 = x_3 - x_2 = h$ (where h is some constant).

Simply speaking a set of points :

$$(x_0,\ y_0),\ (x_1,\ y_1),\ (x_2,\ y_2),\ \ldots,\ (x_n,\ y_n),$$

is equally spaced if the difference b/w successive x-values is a constant.



$x_1 = x_0 + h$, $x_2 = x_0 + 2h$, $x_3 = x_0 + 3h$

Generally,

$x_i = x_0 + ih, \quad i = 0, 1, 2, \dots n$

Examples of Equally spaced Data

x	y
1	43
2	32
3	-56
4	76

x	y
-2	43
0	32
2	-56
4	76

Names of formulas for Interpolation with Equally space data

- Newton's Forward Difference Interpolation Formula
- Newton's Backward Difference Interpolation Formula
- Gauss Forward Formula
- Gauss Backward Formula
- Stirling's Formula
- Bessel's Formula
- Laplace-Everett Formula

Newton's Forward Difference Interpolation Formula

Consider the set of (n+1) values :

$$(x_0,\ y_0),\ (x_1,\ y_1),\ (x_2,\ y_2),\ \ldots,\ (x_n,\ y_n),$$

and let

$$x_i = x_0 + ih, \qquad i = 0,\ 1,\ 2, \ldots, n.$$

i.e. it is equally spaced data. Then if x_i's are distinct then we can find the interpolating polynomial by Newton's Forward difference interpolation formula and is given by:

$$y_n(x) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \cdots$$

$$+ \frac{p(p-1)(p-2)\ldots(p-n+1)}{n!}\Delta^n y_0,$$

where $p = \dfrac{x - x_0}{h}$

For n=2 we have

$$y_n(x) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0$$

For n=3 we have

$$\boxed{y_n(x) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0}$$

Problem

Find the cubic polynomial which takes the following values:

$y(1) = 24$, $y(3) = 120$, $y(5) = 336$, and $y(7) = 720$.

Also compute $y(8)$.

Solution: We form the difference table:

x	y	Δ	Δ^2	Δ^3	
1	24				y_0
		96			Δy_0
3	120		120		$\Delta^2 y_0$
		216		48	$\Delta^3 y_0$
5	336		168		
		384			
7	720				

Here $h = 2$. With $x_0 = 1$, we have $x = 1 + 2p$ or $p = (x - 1)/2$.

Substituting values in the formula of Newton's Forward Difference Interpolation we get

$$y(x) = 24 + \frac{x-1}{2}(96) + \frac{\left(\frac{x-1}{2}\right)\left(\frac{x-1}{2}-1\right)}{2}(120) + \frac{\left(\frac{x-1}{2}\right)\left(\frac{x-1}{2}-1\right)\left(\frac{x-1}{2}-2\right)}{6}(48)$$

$$= x^3 + 6x^2 + 11x + 6.$$

$$y(8) = 990.$$

Problem 1 Values of x (in degrees) and $\sin x$ are given in the following

x (in degrees)	$\sin x$
15	0.2588190
20	0.3420201
25	0.4226183
30	0.5
35	0.5735764
40	0.6427876

Use Newton's Forward difference interpolation formula to determine the value of $\sin(38^0)$.

Problem 2

$Given \ that \ \sqrt{12500} = 111.8034, \sqrt{12510} = 111.8481, \sqrt{12520} = 111.8928,$
$\sqrt{12530} = 111.9375$, Use Newton's Forward Difference Interpolation formula to estimate the value of $\sqrt{12516}$:

Newton's Backward Difference Interpolation Formula

Consider the set of (n+1) values :

$$(x_0,\ y_0),\ (x_1,\ y_1),\ (x_2,\ y_2),\ \ldots,\ (x_n,\ y_n),$$

and let

$$x_i = x_0 + ih, \qquad i = 0,\ 1,\ 2, \ldots, n.$$

i.e. it is equally spaced data. Then if x_i's are distinct then we can find the interpolating polynomial by Newton's Backward Difference Interpolation Formula and is given by:

$$y_n(x) = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \cdots + \frac{p(p+1)\ldots(p+n-1)}{n!}\nabla^n y_n,$$

where $p = (x - x_n)/h$.

Problem

Consider the following data:

x	**y**
1	24
3	120
5	336
7	720

Use Newton's Backward Difference Interpolation Formula to find a polynomial that interpolates the given points and then calculate $y(8)$.

Solution

x	**y**	∇	∇^2	∇^3
1	24			
3	120	120-24=96		
5	336	216	120	
7	720	384	168	48
	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$
	y_3	∇y_3	$\nabla^2 y_3$	$\nabla^3 y_3$

Here $n=3$, $h=2$ & $x_n = x_3 = 7$ So $p = \frac{x - x_n}{h} \Rightarrow \boxed{p = \frac{x-7}{2}}$

For $n=3$

Newton's Backward Difference Interpolation Formula is

$$\boxed{y_3(x) = y_3 + P\nabla y_3 + \frac{P(P+1)}{2!}\nabla^2 y_3 + \frac{P(P+1)(P+2)}{3!}\nabla^3 y_3} \quad \text{——(i)}$$

Now

$$P+1 = \frac{x-7}{2} + 1 = \frac{x-7+2}{2} = \frac{x-5}{2}$$

$$\Rightarrow P(P+1) = \left(\frac{x-7}{2}\right)\left(\frac{x-5}{2}\right) = \frac{1}{4}(x-7)(x-5) = \frac{1}{4}\left(x^2 - 5x - 7x + 35\right)$$

$$\Rightarrow P(P+1) = \frac{1}{4}(x^2 - 12x + 35)$$

Also $$P(P+1)(P+2) = \frac{1}{4}(x^2-12x+35)\left(\frac{x-7}{2}+2\right)$$

$$= \frac{1}{4}(x^2-12x+35)\left(\frac{x-3}{2}\right)$$

$$= \frac{1}{8}(x^2-12x+35)(x-3)$$

$$= \frac{1}{8}(x^3-3x^2-12x^2+36x+35x-105)$$

$$\Rightarrow \boxed{P(P+1)(P+2) = \frac{1}{8}(x^3-15x^2+71x-105)}$$

Substituting values in eq (i) we get

$$y_3(x) = 720 + \left(\frac{x-7}{2}\right)(384) + \frac{1}{8}(x^2-12x+35)(168)$$
$$+ \frac{1}{\cancel{48}}(x^3-15x^2+71x-105)(\cancel{48})$$

$$\Rightarrow y_3(x) = 720 + 192(x-7) + 21(x^2-12x+35) + x^3 - 15x^2 + 71x - 105$$

$$= 720 + 192x - 1344 + 21x^2 - 252x + 735 + x^3 - 15x^2 + 71x - 105$$

$$= x^3 + 21x^2 - 15x^2 + 192x - 252x + 71x + 720 - 1344 + 735 - 105$$

$$\Rightarrow \boxed{y_3(x) = x^3 + 6x^2 + 11x + 6}$$

Required Interpolating Polynomial

Now

$$y_3(8) = 8^3 + 6(8^2) + 11(8) + 6$$

$$\Rightarrow \boxed{y(8) = 990}$$

Practice Problem

Values of x (in degrees) and $\sin x$ are given in the following

x (in degrees)	$y = \sin x$
15	0.2588190
20	0.3420201
25	0.4226183
30	0.5
35	0.5735764
40	0.6427876

Use Newton's Backward Difference Interpolation Formula to find a polynomial that interpolates the given points.

When to Use Newton's Forward & Backward Difference Interpolation Formulas

Specifically, Newton's Forward Difference Interpolation Formula is useful for interpolation *near the beginning* of a set of tabular values and Newton's Backward Difference Interpolation Formula is useful for interpolation *near the end* of a set of tabular values. However if we have no option then we can use Newton's Forward Difference Interpolation Formula for interpolation *near the end* of a set of tabular values and likewise one can use Newton's Backward Difference Interpolation formula for interpolation *near the beginning* of a set of tabular values.

GAUSS FORWARD INTERPOLATION FORMULA

Primarily, we use Gauss Forward Formula for interpolation near the center of a set of tabular values. e.g. Consider the following

x	y	
1	24	
3	120	
5	336	Center
7	720	
9	950	

and suppose we want to find $y(5.5)$ i.e. y value corresponding to $x = 5.5$

The Gauss forward formula is

$$y = y_0 + G_1 \Delta y_0 + G_2 \Delta^2 y_{-1} + G_3 \Delta^3 y_{-1} + G_4 \Delta^4 y_{-2} + \cdots$$

where

$$G_1 = p, \qquad G_2 = \frac{p(p-1)}{2!}, \qquad G_3 = \frac{(p+1)\ p(p-1)}{3!},$$

$$G_4 = \frac{(p+1)\ p(p-1)\ (p-2)}{4!} \quad \ldots \quad \text{where} \quad p = \frac{x - x_0}{h}$$

- The differences used in the Gauss forward formula lie on a line shown in a table called central difference table in the next slide.

- We have made the table in which the central point is chosen x_0.

x	y	Δ	Δ^2	Δ^3	Δ^4	Δ^5	Δ^6
x_{-3}	y_{-3}						
		Δy_{-3}					
x_{-2}	y_{-2}		$\Delta^2 y_{-3}$				
		Δy_{-2}		$\Delta^3 y_{-3}$			
x_{-1}	y_{-1}		$\Delta^2 y_{-2}$		$\Delta^4 y_{-3}$		
		Δy_{-1}		$\Delta^3 y_{-2}$		$\Delta^5 y_{-3}$	
x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$		$\Delta^6 y_{-3}$
		Δy_0		$\Delta^3 y_{-1}$		$\Delta^5 y_{-2}$	
x_1	y_1		$\Delta^2 y_0$		$\Delta^4 y_{-1}$		
		Δy_1		$\Delta^3 y_0$			
x_2	y_2		$\Delta^2 y_1$				
		Δy_2					
x_3	y_3						

Problem

From the following table, calculate $y(1.17)$ using Gauss forward formula:

x	y
1.00	2.7183
1.05	2.8577
1.10	3.0042
1.15	3.1582
1.20	3.3201
1.25	3.4903
1.30	3.6693

Solution

Here

$h = 0.05$ (Common Difference b/w successive x-values)

$x_0 = 1.15$

$x = 1.17$ (That x-value whose corresponding y value is unknown)

$$p = \frac{1.17 - 1.15}{0.05} = \frac{0.02}{0.05} = \frac{2}{5}$$

x	y	Δ	Δ^2	Δ^3	Δ^4	Δ^5	Δ^6
1.00	2.7183						
		0.1394					
1.05	2.8577		0.0071				
		0.1465		0.0004			
1.10	3.0042		0.0075		0		
		0.1540		0.0004		0	
1.15	3.1582		0.0079		0		0.0001
		0.1619		0.0004		0.0001	
1.20	3.3201		0.0083		0.0001		
		0.1702		0.0005			
1.25	3.4903		0.0088				
		0.1790					
1.30	3.6693						

$$\mathbf{y} = 3.1582 + \frac{2}{5}(0.1619) + \frac{(2/5)(2/5-1)}{2}(0.0079) + \frac{(2/5+1)\,(2/5)\,(2/5-1)}{6}(0.0004)$$

$$+\ 0\ +\ \frac{(2/5+2\ )\ (2/5+1)\ (2/5)\ (2/5-1)\ (2/5-2\ )\ \mathbf{0.0001}}{\mathbf{120}}\ +$$

$$\frac{(2/5+2\ )\ (2/5+1)\ (2/5)\ (2/5-1)\ (2/5-2\ )\ (2/5-3\ )\ \mathbf{0.0001}}{\mathbf{720}}$$

$$y = 3.2216535$$

Practice Problem

Use Gauss forward formula to find y for x = 30 given that

x	21	25	29	33	37
y	18.4708	17.8144	17.1070	16.3432	15.5154

Final Answer: $y = 16.9216$

Problem

Consider the following data:

x	**y**
25	0.2707
30	0.3027
35	0.3386
40	0.3794

Use Gauss Forward Interpolation Formula to calculate $y(32)$.

Solution Hints:

Since $x = 32$ is close to 30 so we will take $x_0 = 30$.

x	y
$x_{-1} = 25$	0.2707
$x_0 = 30$	0.3027
$x_1 = 35$	0.3386
$x_2 = 40$	0.3794

x	**y**	Δ	Δ^2	Δ^3
25	0.2707			
		0.032		
$x_0 = 30$	$y_0 = 0.3027$		0.0039	
		0.0359		0.001
35	0.3386		0.0049	
		0.0408		
35	0.3794			

GAUSS BACKWARD INTERPOLATION FORMULA

Primarily, we use Gauss Forward Formula for interpolation near the center of a set of tabular values. e.g. Consider the following

x	y	
1	24	
3	120	
5	336	Center
7	720	
9	950	

and suppose we want to find $y(4.5)$ i.e. y value corresponding to $x = 4.5$.

This formula uses the differences which lie on the line shown in Table

x	y	Δ	Δ^2	Δ^3	Δ^4	Δ^5	Δ^6
$\vdots$	$\vdots$						
x_{-1}	y_{-1}						
		Δy_{-1}		$\Delta^3 y_{-2}$		$\Delta^5 y_{-3}$	
x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$		$\Delta^6 y_{-3}$
		Δy_0		$\Delta^3 y_{-1}$		$\Delta^5 y_{-2}$	
x_1	y_1						
$\vdots$	$\vdots$						

The Gauss Backward formula is

$$y = y_0 + G_1'\Delta y_{-1} + G_2'\Delta^2 y_{-1} + G_3'\Delta^3 y_{-2} + G_4'\Delta^4 y_{-2} + \cdots$$

where

$$G_1' = p, \quad G_2' = \frac{p(p+1)}{2!}, \quad G_3' = \frac{(p+1)\,p(p-1)}{3!}$$

$$G_4' = \frac{(p+2)\,(p+1)\,p(p-1)}{4!} \qquad \text{where} \quad p = \frac{x - x_0}{h}$$

$$\vdots$$

Problem

Apply Gauss's backward interpolation formula and find the population of a town in 1946, with the help of following data

Year	1931	1941	1951	1961	1971
Population (in thousands)	15	20	27	39	52

Solution

We have $h = 10$ taking origin at 1951

$$p = \frac{1946 - 1951}{10} = -0.5$$

The difference table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
–2	15				
		5			
–1	20		2		
		7		3	
0	27		5		–7
		12		–4	
2	39		1		
		13			
2	52				

$$y = 27 + (-0.5) \times 7 + \frac{(0.5)(-0.5) \times 5}{2} + \frac{(0.5)(-0.5)(-1.5)}{6} \times 3 + \frac{(1.5)(0.5)(-0.5)(-1.5)(-7)}{24}$$

$$y = 22.898438.$$

$\therefore$ The population of the town in the year 1946 is 22.898 thousand, i.e., 22898.

Gauss's forward formula is used to interpolate the values of the function for the value of $\boldsymbol{p}$ such that $0 < \boldsymbol{p} < 1$, and Gauss's backward formula is used to interpolate line value of the function for a negative value of $\boldsymbol{p}$ which lies between –1 and 0 (i.e., $-1 < \boldsymbol{p} < 0$).

STIRLING'S FORMULA

$$y = y_0 + p\left[\frac{\Delta y_0 + \Delta y_{-1}}{2}\right] + \frac{p^2}{2}\Delta^2 y_{-1} +$$

$$\frac{p(p^2-1^2)}{3!}\frac{(\Delta^3 y_{-1} + \Delta^3 y_{-2})}{2} + \frac{p^2(p^2-1^2)}{4!}\Delta^4 y_{-2} + \frac{p(p^2-1^2)(p^2-2^2)}{5!}\frac{(\Delta^5 y_{-2} + \Delta^5 y_{-3})}{2} + \ldots$$

The above formula is called Stirling's Formula and it is useful if

$$-\frac{1}{4} \le p \le \frac{1}{4}$$

x	y	Δ	Δ^2	Δ^3	Δ^4	Δ^5	Δ^6
x_{-3}	y_{-3}						
		Δy_{-3}					
x_{-2}	y_{-2}		$\Delta^2 y_{-3}$				
		Δy_{-2}		$\Delta^3 y_{-3}$			
x_{-1}	y_{-1}		$\Delta^2 y_{-2}$		$\Delta^4 y_{-3}$		
		Δy_{-1}		$\Delta^3 y_{-2}$		$\Delta^5 y_{-3}$	
x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$		$\Delta^6 y_{-3}$
		Δy_0		$\Delta^3 y_{-1}$		$\Delta^5 y_{-2}$	
x_1	y_1		$\Delta^2 y_0$		$\Delta^4 y_{-1}$		
		Δy_1		$\Delta^3 y_0$			
x_2	y_2		$\Delta^2 y_1$				
		Δy_2					
x_3	y_3						

Problem

Use Stirling's formula to compute $y(12.2)$ *from the following table*

x	10	11	12	13	14
y	23967	28060	31788	35209	38368

Solution

$$p = \frac{x - x_0}{h} = \frac{12.2 - 12}{1} = 0.2, \quad \text{where } x_0 = 12 \text{ is the origin.}$$

x	y	Δ	Δ^2	Δ^3	Δ^4
10	23967				
		4093			
11	28060		–365		
		3728		58	
12	31788		–307		–13
		3421		45	
13	35209		–262		
		3159			
14	38368				

$$y = 31788 + (0.2)\left[\frac{3421+3728}{2}\right] + \frac{0.2^2}{2}(-307)$$

$$+ \frac{(0.2)(0.2^2-1^2)}{3!}\left[\frac{45+58}{2}\right]$$

$$+ \frac{0.2^2(0.2^2-1^2)}{4!}(-13)$$

$$= 31788 + (0.2)(3574.5) + 0.02(-307)$$

$$+ (-0.032)(51.5) + (-0.0016)(-13)$$

$$\boxed{y = 32495.1328}$$

Consider the following table constructed for some data in which x_0 is chosen close to that x whose corresponding y is unknown. How many values from the table we will be using in calculation ?

x	y	Δ	Δ^2	Δ^3	Δ^4	Δ^5
x_{-3}	y_{-3}					
		Δy_{-3}				
x_{-2}	y_{-2}		$\Delta^2 y_{-3}$			
		Δy_{-2}		$\Delta^3 y_{-3}$		
x_{-1}	y_{-1}		$\Delta^2 y_{-2}$		$\Delta^4 y_{-3}$	
		Δy_{-1}		$\Delta^3 y_{-2}$		$\Delta^5 y_{-3}$
x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$	
		Δy_0		$\Delta^3 y_{-1}$		
x_1	y_1		$\Delta^2 y_0$			
		Δy_1				
x_2	y_2					

BESSEL'S FORMULA

This is a very useful formula for interpolation, and it uses the differences as shown in the following table, where the brackets mean that the average of the values has to be taken.

$$
\begin{array}{cccccccc}
\vdots & \vdots & & & & & & \\
x_{-1} & y_{-1} & & & & & & \\
\begin{matrix} x_0 \\ x_1 \end{matrix} & \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} & \Delta y_0 & \begin{pmatrix} \Delta^2 y_{-1} \\ \Delta^2 y_0 \end{pmatrix} & \Delta^3 y_{-1} & \begin{pmatrix} \Delta^4 y_{-2} \\ \Delta^4 y_{-1} \end{pmatrix} & \Delta^5 y_{-2} & \begin{pmatrix} \Delta^6 y_{-3} \\ \Delta^6 y_{-2} \end{pmatrix} \\
\vdots & \vdots & & & & & &
\end{array}
$$

$$y = \frac{y_0 + y_1}{2} + B_1 \Delta y_0 + B_2 \frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + B_3 \Delta^3 y_{-1}$$

$$+ B_4 \frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} + \cdots$$

where

$$B_1 = p - \frac{1}{2}, \qquad B_2 = \frac{p(p-1)}{2!}, \qquad B_3 = \frac{(p-1/2)\, p(p-1)}{3!},$$

$$B_4 = \frac{(p+1)\; p(p-1)\,(p-2)}{4!} \qquad \ldots$$

The above formula is called Bessel's Formula and it is useful if $\frac{1}{4} \leq p \leq \frac{3}{4}$.

Problem

From the following table, calculate $y(1.17)$ using Bessel's formula:

x	y
1.00	2.7183
1.05	2.8577
1.10	3.0042
1.15	3.1582
1.20	3.3201
1.25	3.4903
1.30	3.6693

Solution

Here

$h = 0.05$ (Common Difference b/w successive x-values)

$x = 1.17$ (That x-value whose corresponding y value is unknown)

$x_0 = 1.15$

$$p = \frac{1.17 - 1.15}{0.05} = \frac{0.02}{0.05} = \frac{2}{5} = 0.4$$

x	y	Δ	Δ^2	Δ^3	Δ^4	Δ^5	Δ^6
1.00	2.7183						
		0.1394					
1.05	2.8577		0.0071				
		0.1465		0.0004			
1.10	3.0042		0.0075		0		
		0.1540		0.0004		0	
1.15	3.1582		0.0079		0		0.0001
		0.1619		0.0004		0.0001	
1.20	3.3201		0.0083		0.0001		
		0.1702		0.0005			
1.25	3.4903		0.0088				
		0.1790					
1.30	3.6693						

In this case the Bessel's Formula is

$$y = \frac{3.1582 + 3.3201}{2} + B_1(0.1619) + B_2\left(\frac{0.0079 + 0.0083}{2}\right)$$

$$+ B_3(0.0004) + B_4\left(\frac{0 + 0.0001}{2}\right) + B_5(0.0001) \longrightarrow \text{(i)}$$

Now

$$B_1 = P - \frac{1}{2} = \frac{2}{5} - \frac{1}{2} = 0.4 - 0.5 = -0.1 \Rightarrow \boxed{B_1 = -0.1}$$

$$B_2 = \frac{P(P-1)}{2!} = \frac{0.4(0.4-1)}{2!} = -0.12 \Rightarrow \boxed{B_2 = -0.12}$$

$$B_3 = \frac{\left(P-\frac{1}{2}\right)P(P-1)}{3!} = \frac{(-0\cdot 1)(0\cdot 4)(0\cdot 4-1)}{3!} = \frac{0\cdot 024}{6} = 0\cdot 004$$

$$\Rightarrow \boxed{B_3 = 0\cdot 004}$$

Next

$$B_4 = \frac{(P+1)P(P-1)(P-2)}{4!} = \frac{(0\cdot 4+1)(0\cdot 4)(0\cdot 4-1)(0\cdot 4-2)}{4!} = 0\cdot 0224$$

$$\Rightarrow \boxed{B_4 = 0\cdot 0224}$$

Also

$$B_5 = \frac{\left(P-\frac{1}{2}\right)P(P-1)(P-2)(P-3)}{5!}$$ Substituting values we get

$$\boxed{B_5 = 0.000832}$$

Substituting values in eq(i) we get

$$y = 3.23915 + (-0.1 \times 0.1619) + (-0.12 \times 0.0081)$$
$$+ (0.004 \times 0.0004) + (0.0224 \times 0.00005)$$
$$+ (0.000832 \times 0.0001)$$

$$\boxed{y = 3.2219908}$$

LAPLACE–EVERETT FORMULA

It's one of the interpolation formula that uses only even order difference as shown in the following table:

x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$		$\Delta^6 y_{-3}$	
		–		–		–		–
x_1	y_1		$\Delta^2 y_0$		$\Delta^4 y_{-1}$		$\Delta^6 y_{-2}$	

x	y	Δ	Δ^2	Δ^3	Δ^4	Δ^5	Δ^6
x_{-3}	y_{-3}						
		Δy_{-3}					
x_{-2}	y_{-2}		$\Delta^2 y_{-3}$				
		Δy_{-2}		$\Delta^3 y_{-3}$			
x_{-1}	y_{-1}		$\Delta^2 y_{-2}$		$\Delta^4 y_{-3}$		
		Δy_{-1}		$\Delta^3 y_{-2}$		$\Delta^5 y_{-3}$	
x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$		$\Delta^6 y_{-3}$
		Δy_0		$\Delta^3 y_{-1}$		$\Delta^5 y_{-2}$	
x_1	y_1		$\Delta^2 y_0$		$\Delta^4 y_{-1}$		
		Δy_1		$\Delta^3 y_0$			
x_2	y_2		$\Delta^2 y_1$				
		Δy_2					
x_3	y_3						

Everett's formula is given by

$$y = qy_0 + \frac{q(q^2-1^2)}{3!}\Delta^2 y_{-1} + \frac{q(q^2-1^2)(q^2-2^2)}{5!}\Delta^4 y_{-2} + \cdots$$

$$+ py_1 + \frac{p(p^2-1^2)}{3!}\Delta^2 y_0 + \frac{p(p^2-1^2)(p^2-2^2)}{5!}\Delta^4 y_{-1} + \cdots$$

where $q = 1 - p$.

Problem

From the following table calculate $\boldsymbol{y(25)}$ using Everett's Formula.

x	**y**
20	2854
24	3162
28	3544
32	3992

Solution

Here

$h = 4$ (Common Difference b/w successive x-values)

$x = 25$ (That x-value whose corresponding y value is unknown)

$x_0 = 24$

$$p = \frac{25 - 24}{4} = \frac{1}{4} = 0.25 \quad \& \quad q = 1 - \frac{1}{4} = \frac{3}{4} = 0.75$$

The difference table is

x	y	Δ	Δ^2	Δ^3
20	2854			
		308		
24	3162		74	
		382		–8
28	3544		66	
		448		
32	3992			

Using Everett's Formulas, we have

$$y = \frac{3}{4}\cdot(3162) + \frac{\frac{3}{4}\left(\frac{9}{16}-1\right)}{6}(74) + \frac{1}{4}\cdot(3544) + \frac{\frac{1}{4}\left(\frac{1}{16}-1\right)}{6}(66)$$

$$\Rightarrow \quad y = 3254.875.$$

A note on the use of formulas for interpolation with equally space data

- Newton's Forward Difference Interpolation Formula (near the beginning)
- Newton's Backward Difference Interpolation Formula (near the end)

Methods for Interpolation near the center or intermediate of the data

- ❑ Gauss Forward Formula $(0 < p < 1)$
- ❑ Gauss Backward Formula $(-1 < p < 0)$
- ❑ Stirling's Formula $(-0.25 \leq p \leq 0.25)$
- ❑ Bessel's Formula $(0.25 \leq p \leq 0.75)$
- ❑ Laplace-Everett Formula

The value of p

- Newton's Forward Difference Interpolation Formula

$$p = \frac{x - x_0}{h}$$

- Newton's Backward Difference Interpolation Formula

$$p = \frac{x - x_n}{h}$$

- ❑ Gauss Forward Interpolation Formula
- ❑ Gauss Backward Interpolation Formula
- ❑ Stirling's Formula
- ❑ Bessel's Formula
- ❑ Laplace-Everett Formula

$$p = \frac{x - x_0}{h}$$

Interpolation with Unequally Spaced Data

A set of points

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

that is not equally spaced is called unequally spaced data. Here is an example:

x	y
1	50
3	31
4	354
9	3992

In the previous section we have studied various interpolation formulas that are useful for interpolation only if the data is equally spaced. Here in this section we are focusing on methods that will be useful even if the data is unequally spaced.

Lagrange's Interpolation Formula

Given the $(n+1)$ points (x_0, y_0), (x_1, y_1), …, (x_n, y_n) not necessarily be equally spaced, we wish to find a polynomial of degree n, say $L_n(x)$, such that

$$L_n(x_i) = y(x_i) = y_i, \quad i = 0, 1, \ldots, n$$

Before deriving the general formula, we first consider a simpler case, viz., the equation of a straight line (a linear polynomial) passing through two points (x_0, y_0) and (x_1, y_1). Such a polynomial, say $L_1(x)$, is easily seen to be

$$L_1(x) = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1 = l_0(x)\, y_0 + l_1(x)\, y_1$$

$$\Rightarrow \quad L_1(x) = \sum_{i=0}^{1} l_i(x)\, y_i, \qquad \text{(1)}$$

where

$$l_0(x) = \frac{x - x_1}{x_0 - x_1} \quad \text{and} \quad l_1(x) = \frac{x - x_0}{x_1 - x_0}.$$

Notice that

$$l_0(x_0) = 1, \qquad l_0(x_1) = 0, \qquad l_1(x_0) = 0, \qquad l_1(x_1) = 1.$$

These relations can be expressed in a more convenient form as

$$l_i(x_j) = \begin{cases} 1, \text{ if } i = j \\ 0, \text{ if } i \neq j. \end{cases}$$

The $l_i(x)$ also have the property

$$\sum_{i=0}^{1} l_i(x) = l_0(x) + l_1(x) = \frac{x - x_1}{x_0 - x_1} + \frac{x - x_0}{x_1 - x_0} = 1.$$

Equation (**1**) is called the *Lagrange polynomial of degree one passing through* two points (x_0, y_0) and (x_1, y_1). In a similar way, the *Lagrange polynomial of degree two* passing through three points (x_0, y_0), (x_1, y_1) and (x_2, y_2) is written as

$$L_2(x) = \sum_{i=0}^{2} l_i(x)\, y_i$$

$$= \frac{(x-x_1)\,(x-x_2)}{(x_0-x_1)\,(x_0-x_2)}\, y_0 + \frac{(x-x_0)\,(x-x_2)}{(x_1-x_0)\,(x_1-x_2)}\, y_1 + \frac{(x-x_0)\,(x-x_1)}{(x_2-x_0)\,(x_2-x_1)}\, y_2\,,$$

where the $l_i(x)$ satisfy

$$l_i(x_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j. \end{cases} \qquad \& \qquad \sum_{i=0}^{2} l_i(x) = =1.$$

Generally for (n+1) points

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

The Lagrange interpolating polynomial is given by

$$L_n(x) = \sum_{i=0}^{n} l_i(x)\, y_i,$$

where $l_i(x)$ is given by

$$l_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^{n} \frac{(x - x_k)}{(x_i - x_k)}.$$

Consider four Points

$(x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3)$

In this case the Lagrange Interpolating polynomial will be given by

$$L_3(x) = \sum_{i=0}^{3} l_i(x) y_i = l_0(x) y_0 + l_1(x) y_1 + l_2(x) y_2 + l_3(x) y_3$$

where

$$l_0(x) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}$$

$$l_1(x) = \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}$$

and similarly $l_2(x)$ & $l_3(x)$.

Problem

Use Lagrange's interpolation formula to find the value of y corresponding to $x = 10$ from the following table.

x	y
5	12
6	13
9	14
11	16

Solution

We know that Lagrange Interpolation Formula is

$$L_3(x) = \sum_{i=0}^{3} l_i(x) y_i = l_0(x) y_0 + l_1(x) y_1 + l_2(x) y_2 + l_3(x) y_3$$

Here we need $L_3(10)$ which is given by

$$\boxed{L_3(10) = l_0(10) y_0 + l_1(10) y_1 + l_2(10) y_2 + l_3(10) y_3 \quad \text{---} \quad (*)}$$

Now

$$l_0(x) = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)}$$

$$\Rightarrow l_0(10) = \frac{(10-6)(10-9)(10-11)}{(5-6)(5-9)(5-11)} = \frac{4(1)(-1)}{(-1)(-4)(-6)} = \frac{1}{6} \Rightarrow \boxed{l_0(10) = \frac{1}{6}}$$

Next

$$l_1(x) = \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}$$

$$\Rightarrow l_1(10) = \frac{(10-5)(10-9)(10-11)}{(6-5)(6-9)(6-11)} = \frac{(5)(1)(-1)}{(1)(-3)(-5)} = -\frac{1}{3} \Rightarrow \boxed{l_1(10) = \frac{-1}{3}}$$

Next $$l_2(x) = \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}$$

$$\Rightarrow l_2(10) = \frac{(10-5)(10-6)(10-11)}{(9-5)(9-6)(9-11)} = \frac{(5)(\cancel{4})(-1)}{(\cancel{+4})(3)(-2)} = \frac{5}{6} \Rightarrow \boxed{l_2(10) = \frac{5}{6}}$$

Also

$$l_3(x) = \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}$$

$$\Rightarrow l_3(10) = \frac{(10-5)(10-6)(10-9)}{(11-5)(11-6)(11-9)} = \frac{(5)(4)(1)}{(6)(5)(2)} = \frac{20}{60} = \frac{1}{3}$$

$$\Rightarrow \boxed{l_3(10) = \frac{1}{3}}$$

Substituting Values in eq * we get

$$L_3(10) = \frac{1}{6}(12) + \left(-\frac{1}{3}\right)13 + \frac{5}{6}(14) + \frac{1}{3}(16) = \frac{42}{3}$$

So the value of y corresponding to $x = 10$ is $\frac{42}{3}$.

Practice Problems

Problem 1

Use Lagrange's interpolation formula to find a polynomial which passes through the points $(\mathbf{0}, -\mathbf{12}), (\mathbf{1}, \mathbf{0}), (\mathbf{3}, \mathbf{6}), (\mathbf{4}, \mathbf{12})$.

Final Ans: $\boldsymbol{L_3(x) = x^3 - 7x^2 + 18x - 12}$

Problem 2

Let $f(x) = 2x^2 e^x + 1$. Construct a Lagrange polynomial of degree two or less using $x_0 = 0, x_1 = 0.5$, and $x_2 = 1$. Approximate $f(0.8)$.

The Lagrange interpolation formula, has the disadvantage that if another order pair were added to data, then the interpolation coefficients $l_i(x)$ will have to be recomputed. We therefore seek an interpolation polynomial which has the property that a polynomial of higher degree may be derived from it by simply adding new terms. Newton's divided-differnce formula is one such formula and it employs what are called *divided differences*.

Divided Differences

The *zeroth divided difference* of the function f with respect to x_i, denoted $f[x_i]$, is simply the value of f at x_i:

$$f[x_i] = f(x_i)$$

The remaining divided differences are defined recursively; the *first divided difference* f with respect to x_i and x_{i+1} is denoted $f[x_i, x_{i+1}]$ and defined as

$$f[x_i, x_{i+1}] = \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i}$$

The *second divided difference*, $f[x_i, x_{i+1}, x_{i+2}]$, is defined as

$$f[x_i, x_{i+1}, x_{i+2}] = \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{x_{i+2} - x_i}$$

The third *divided difference*, $f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$, is defined as

$$f[x_i, x_{i+1}, x_{i+2}, x_{i+3}] = \frac{f[x_{i+1}, x_{i+2}, x_{i+3}] - f[x_i, x_{i+1}, x_{i+2}]}{x_{i+3} - x_i}$$

and so on.

$$
\begin{array}{llccc}
x & f(x) & \begin{array}{c}\text{First}\\ \text{divided differences}\end{array} & \begin{array}{c}\text{Second}\\ \text{divided differences}\end{array} & \begin{array}{c}\text{Third}\\ \text{divided differences}\end{array} \\
\hline
x_0 & f[x_0] & & & \\
 & & f[x_0, x_1] = \dfrac{f[x_1] - f[x_0]}{x_1 - x_0} & & \\
x_1 & f[x_1] & & f[x_0, x_1, x_2] = \dfrac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} & \\
 & & f[x_1, x_2] = \dfrac{f[x_2] - f[x_1]}{x_2 - x_1} & & f[x_0, x_1, x_2, x_3] \\
x_2 & f[x_2] & & f[x_1, x_2, x_3] = \dfrac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} & \\
 & & f[x_2, x_3] = \dfrac{f[x_3] - f[x_2]}{x_3 - x_2} & & \\
x_3 & f[x_3] & & &
\end{array}
$$

Problem

Construct the divided difference table for the data.

x	-1	0	2	3
$f(x)$	-8	3	1	12

Solution

x	$f(x)$	First d.d	Second d.d	Third d.d
-1	-8			
		$\frac{3+8}{0+1} = 11$		
0	3		$\frac{-1-11}{2+1} = -4$	
		$\frac{1-3}{2-0} = -1$		$\frac{4+4}{3+1} = 2$
2	1		$\frac{11+1}{3-0} = 4$	
		$\frac{12-1}{3-2} = 11$		
3	12			

Practice Problem

Construct the divided difference table for the data.

x	-4	-1	0	2	5
$f(x)$	1245	33	5	9	1335

Newton's Divided-Difference Formula

Consider a set of unequally spaced data $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$. The polynomial which interpolates these points is given by the following formula called Newton's Divided-Difference Formula.

$$\boxed{\begin{aligned} P_n(x) = f(x_0) &+ (x - x_0) f[x_0, x_1] + (x - x_0)(x - x_1) f[x_0, x_1, x_2] + \dots \\ &+ (x - x_0)(x - x_1)\dots(x - x_{n-1}) f[x_0, x_1, \dots, x_n] \end{aligned}}$$

Problem

Use Newton's divided difference formula to find a polynomial which interpolates the following points :

x	-4	-1	0	2	5
$f(x)$	1245	33	5	9	1335

Solution

In this case, the Newton's divided difference formula is

$$
\begin{aligned}
P_4(x) = f(x_0) &+ (x - x_0)\, f[x_0, x_1] + (x - x_0)(x - x_1)\, f[x_0, x_1, x_2] \\
&+ (x - x_0)(x - x_1)(x - x_2) f[x_0, x_1, x_2, x_3] \\
&+ (x - x_0)(x - x_1)(x - x_2)(x - x_3)\, f[x_0, x_1, x_2, x_3, x_4]
\end{aligned}
$$

(1)

Divided difference table

x	$f(x)$	*First d.d*	*Second d.d*	*Third d.d*	*Fourth d.d*
– 4	1245				
		– 404			
– 1	33		94		
		– 28		– 14	
0	5		10		3
		2		13	
2	9		88		
		442			
5	1335				

Substituting values in eq. (1) we get

$$P_4(x) = 1245 + (x + 4)(-404) + (x + 4)(x + 1)(94) + (x + 4)(x + 1)\, x\, (-14)$$

$$+ (x + 4)(x + 1)\, x\, (x - 2)(3)$$

$$= 1245 - 404x - 1616 + (x^2 + 5x + 4)(94) + (x^3 + 5x^2 + 4x)(-14)$$

$$+ (x^4 + 3x^3 - 6x^2 - 8x)(3)$$

$$\boxed{P_4(x) = 3x^4 - 5x^3 + 6x^2 - 14x + 5.}$$

Practice Problem

Use Newton's divided difference formula to find a polynomial which interpolates the following points :

x	-2	-1	0	1	3	4
$f(x)$	9	16	17	18	44	81

Hence, interpolate at $x = 0.5$ *and* $x = 3.1.$

Final Answers: $x^3 + 17$, 17.125 , $47.791.$

INVERSE INTERPOLATION

Given a set of values of x and y, the process of finding the value of x for a certain value of y is called *inverse interpolation*.

For this problem, we consider the given data as

$$(y_0, x_0), (y_1, x_1), \dots, (y_n, x_n)$$

and construct the interpolation polynomial.

Problem

Find the value of x when y = 0.3 by applying Lagrange's formula inversely

x	0.4	0.6	0.8
y	0.3683	0.3332	0.2897

Solution

From Lagrange's inverse interpolation formula we get

$$x = \frac{(y - y_1)(y - y_2)}{(y_0 - y_1)(y_0 - y_2)} x_0 + \frac{(y - y_0)(y - y_2)}{(y_1 - y_0)(y_1 - y_2)} x_1 + \frac{(y - y_0)(y - y_1)}{(y_2 - y_0)(y_2 - y_1)} x_2$$

Substituting $x_0 = 0.4$, $x_1 = 0.6$, $x_2 = 0.8$,

$y_0 = 0.3683$, $y_1 = 0.3332$, $y_2 = 0.2899$, we get

$$x = \frac{(0.3-0.3332)(0.3-0.2897)}{(0.3683-0.3332)(0.3683-0.2897)} \times (0.4) + \frac{(0.3-0.3683)(0.3-0.2897)}{(0.3332-0.3683)(0.3332-0.2897)} \times (0.6)$$

$$+ \frac{(0.3-0.3683)(0.3-0.3332)}{(0.2897-0.3683)(0.2897-0.3332)} \times (0.8)$$

$$\boxed{x = 0.757358.}$$

Another Way……

Let

x	y
0.3683	0.4
0.3332	0.6
0.2897	0.8

then we need y value corresponding to $x = 0.3$.

As the data is unequally spaced, so let's use Lagrange interpolation formula, so we have

$$y = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}y_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}y_1 + \frac{(x-x_0)(x-x_2)}{(x_2-x_0)(x_2-x_1)}y_2$$

Substituting $x = 0.3,\ \ x_0 = 0.3683,\ \ x_1 = 0.3332,\ \ x_2 = 0.2897,$

$y_0 = 0.4, y_1 = 0.6, y_2 = 0.8,$ **we get**

$$\mathbf{y} = \frac{(0.3-0.3332)(0.3-0.2897)}{(0.3683-0.3332)(0.3683-0.2897)} \times (0.4) + \frac{(0.3-0.3683)(0.3-0.2897)}{(0.3332-0.3683)(0.3332-0.2897)} \times (0.6)$$

$$+ \frac{(0.3-0.3683)(0.3-0.3332)}{(0.2897-0.3683)(0.2897-0.3332)} \times (0.8)$$

$\Rightarrow$ $\mathbf{y} = 0.757358$. Thus the value of x corresponding to $y = 0.3$ is 0.757358.

Practice Problem

The following table gives the value of the elliptical integral

$$F(\phi) = \int_0^{\phi} \frac{dt}{1 - \frac{1}{2}\sin^2 t}$$

for certain values of ϕ. *Find the values of* ϕ *if* $F(\phi) = 0.3887$

ϕ	21	23	25
$F(\phi)$	0.3706	0.4068	0.4433

Final Ans:

$\boxed{\phi = 22\ .}$

Numerical Differentiation

The process of computing the value of the derivative $\frac{dy}{dx}$ for some particular value of x from the given data when the actual form of the function is not known is called *Numerical differentiation.*

As an example suppose we are ask to find

$$\frac{dy}{dx} \text{ at } x = 1.3$$

using the following data

$\boldsymbol{x}$	1	3	5	8
$\boldsymbol{y}$	0.34	0.48	0.96	0.99

The general method for deriving the numerical differentiation formulae is to differentiate the interpolating polynomial. Hence, corresponding to each of the interpolation formula we can derive a formula for the derivative.

We illustrate the derivation with Newton's forward difference formula only, the method of derivation being the same with regard to the other formulae.

Consider Newton's forward difference formula:

$$y = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \cdots$$

where

$$p = \frac{x - x_0}{h}$$

As we can see from above that y depends on p and p depends on x, so y depends on x indirectly. So using chain rule we have

$$\frac{dy}{dx} = \frac{dy}{dp}\frac{dp}{dx}$$

Now

$$\frac{dy}{dp} = \Delta y_0 + \frac{2p-1}{2}\Delta^2 y_0 + \frac{3p^2-6p+2}{6}\Delta^3 y_0 + \ldots$$

&

$$\frac{dp}{dx} = \frac{1}{h}$$

Thus

$$\frac{dy}{dx}=\frac{1}{h}\left(\Delta y_0+\frac{2p-1}{2}\Delta^2 y_0+\frac{3p^2-6p+2}{6}\Delta^3 y_0+\ldots\right)$$

where

$$p=\frac{x-x_0}{h}$$

The above formula is one of the numerical differentiation formula from which we can calculate $\frac{dy}{dx}$ at some given x.

As we can see from the previous slide that $\frac{dy}{dx}$ depends on p and p depends on x, so $\frac{dy}{dx}$ depends on x indirectly. So using chain rule we have

$$\frac{d}{dx}\left(\frac{dy}{dx}\right)=\frac{d}{dp}\left(\frac{dy}{dx}\right)\frac{dp}{dx}$$

$$\Rightarrow \frac{d^2y}{dx^2}=\frac{1}{h}\frac{d}{dp}\left(\Delta y_0+\frac{2p-1}{2}\Delta^2 y_0+\frac{3p^2-6p+2}{6}\Delta^3 y_0+\ldots\right)\frac{1}{h}$$

$$\Rightarrow \frac{d^2y}{dx^2}=\frac{1}{h^2}\frac{d}{dp}\left(\Delta y_0+\frac{2p-1}{2}\Delta^2 y_0+\frac{3p^2-6p+2}{6}\Delta^3 y_0+\ldots\right)$$

$$\Rightarrow \frac{d^2y}{dx^2}=\frac{1}{h^2}\left(\Delta^2 y_0+\frac{6p-6}{6}\Delta^3 y_0+\ldots\right)$$

Problem

Find dy/dx *at* $x = 1$ *from the following table of values*

x	1	2	3	4
y	1	8	27	64

Solution

We have the following forward difference table.

Forward difference table

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
1	1			
		7		
2	8		12	
		19		6
3	27		18	
		37		
4	64			

Here

$$h = 1$$

(Common Difference b/w successive x-values)

$$x = 1$$

(That x-value at which derivative is required)

$$x_0 = 1$$

$$p = \frac{1 - 1}{1} = 0$$

Substituting values in the formula

$$\frac{dy}{dx} = \frac{1}{h}\left(\Delta y_0 + \frac{2p-1}{2}\Delta^2 y_0 + \frac{3p^2-6p+2}{6}\Delta^3 y_0 + \ldots \right)$$

We get

$$\frac{dy}{dx} = 7 - \frac{1}{2}(12) + \frac{1}{3}(6) = 3.$$

Practice Problem

Find dy/dx *at* $x = 1.3$ *from the following table of values*

x	1	3	5	7
y	1	9	25	49

Problem

Find $f'(-1)$ using the following table

x	-4	-1	0	2	5
$f(x)$	1245	33	5	9	1335

Solution

In the previous slides using Newton's Divided Difference Formula we have already calculated the interpolating polynomial.

It was $P_4(x) = 3x^4 - 5x^3 + 6x^2 - 14x + 5.$

$$\Rightarrow P_4'(x) = 12x^3 - 15x^2 + 12x - 14$$

Now

$$P_4'(-1) = -12 - 15 - 12 - 14$$
$$\Rightarrow P_4'(-1) = -53.$$

So $f'(-1)$ is approximately equal to -53.

Practice Problem

Find dy/dx *at* $x = 1.3$ *from the following table of values*

x	1	2	5	7
y	1	9	25	49

Hint: **As the data is unequally spaced so first find the interpolating polynomial $p_3(\mathrm{x})$ and then differentiate it and substitute $x = 1.3$.**

Numerical Integration

Numerical Integration

The general problem of numerical integration may be stated as follows. Given a set of data points (x_0, y_0), (x_1, y_1), …, (x_n, y_n) of a function $y = f(x)$, where $f(x)$ is not known explicitly, it is required to compute the value of the definite integral

$$I = \int_a^b y \, dx.$$

Here we have derived a general formula for numerical integration using Newton's Forward Difference formula.

Let the interval $[a, b]$ be divided into n equal subintervals such that $a = x_0 < x_1 < x_2 < \cdots x_n = b$. Clearly, $x_n = x_0 + nh$. Hence the integral becomes

$$I = \int_{x_0}^{x_n} y\, dx.$$

Approximating y by Newton's forward difference formula, we obtain

$$I = \int_{x_0}^{x_n} \left[y_0 + p\Delta y_0 + \frac{p(p-1)}{2}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{6}\Delta^3 y_0 + \cdots \right] dx.$$

Since $x = x_0 + ph$, $dx = h\, dp$ and hence the above integral becomes

$$I = h \int_{0}^{n} \left[y_0 + p\Delta y_0 + \frac{p(p-1)}{2}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{6}\Delta^3 y_0 + \cdots \right] dp,$$

which gives on simplification

$$\int_{x_0}^{x_n} y\,dx = nh\left[y_0 + \frac{n}{2}\Delta y_0 + \frac{n(2n-3)}{12}\Delta^2 y_0 + \frac{n(n-2)^2}{24}\Delta^3 y_0 + \cdots\right].$$

From this general formula, we can obtain different integration formulas by taking various cases of n.

Trapezoidal Rule

Setting $n = 1$ in the general formula all differences higher than the first will become zero and we obtain

$$\int_{x_0}^{x_1} y\,dx = h\left(y_0 + \frac{1}{2}\Delta y_0\right) = h\left[y_0 + \frac{1}{2}(y_1 - y_0)\right] = \frac{h}{2}(y_0 + y_1).$$

For the next interval $[x_1, x_2]$, we deduce similarly

$$\int_{x_1}^{x_2} y\,dx = \frac{h}{2}(y_1 + y_2)$$

and so on. For the last interval $[x_{n-1}, x_n]$, we have

$$\int_{x_{n-1}}^{x_n} y\,dx = \frac{h}{2}(y_{n-1} + y_n).$$

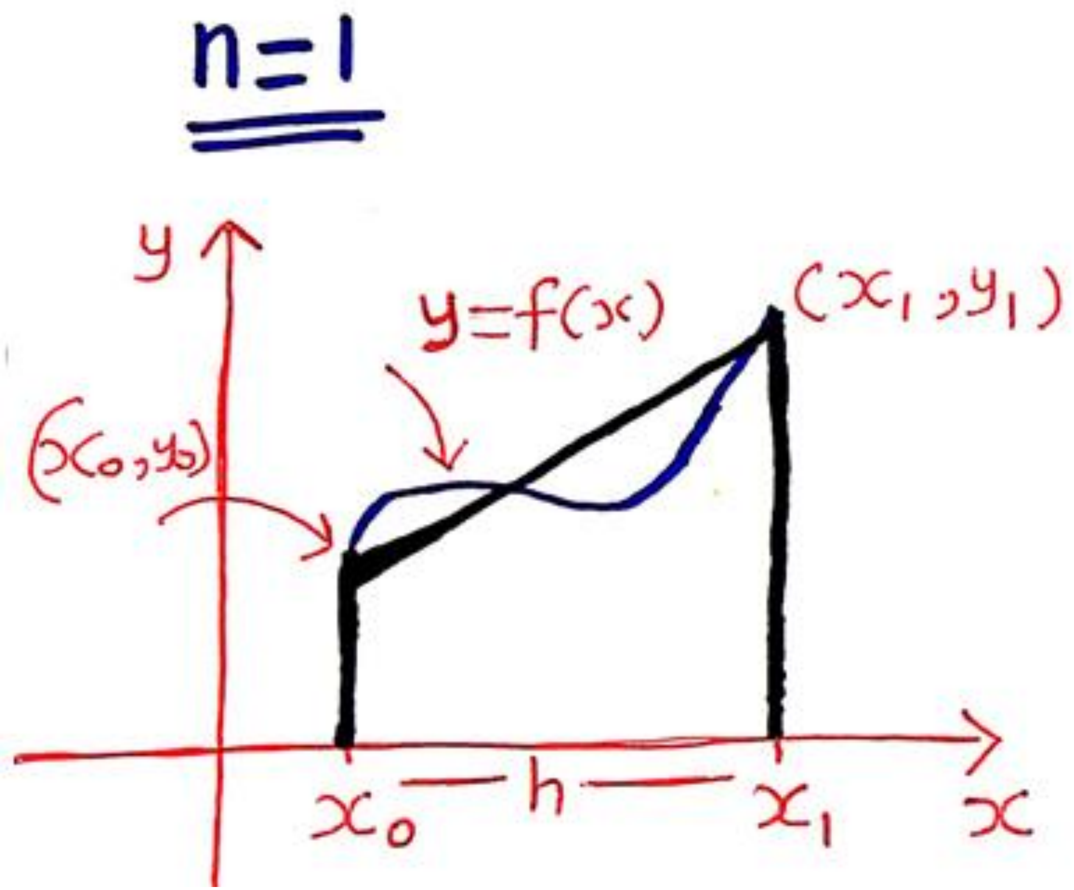
Combining all these expressions, we obtain the rule

$$\boxed{\int_{x_0}^{x_n} y\,dx = \frac{h}{2}\left[y_0 + 2\left(y_1 + y_2 + \cdots + y_{n-1}\right) + y_n\right],}$$

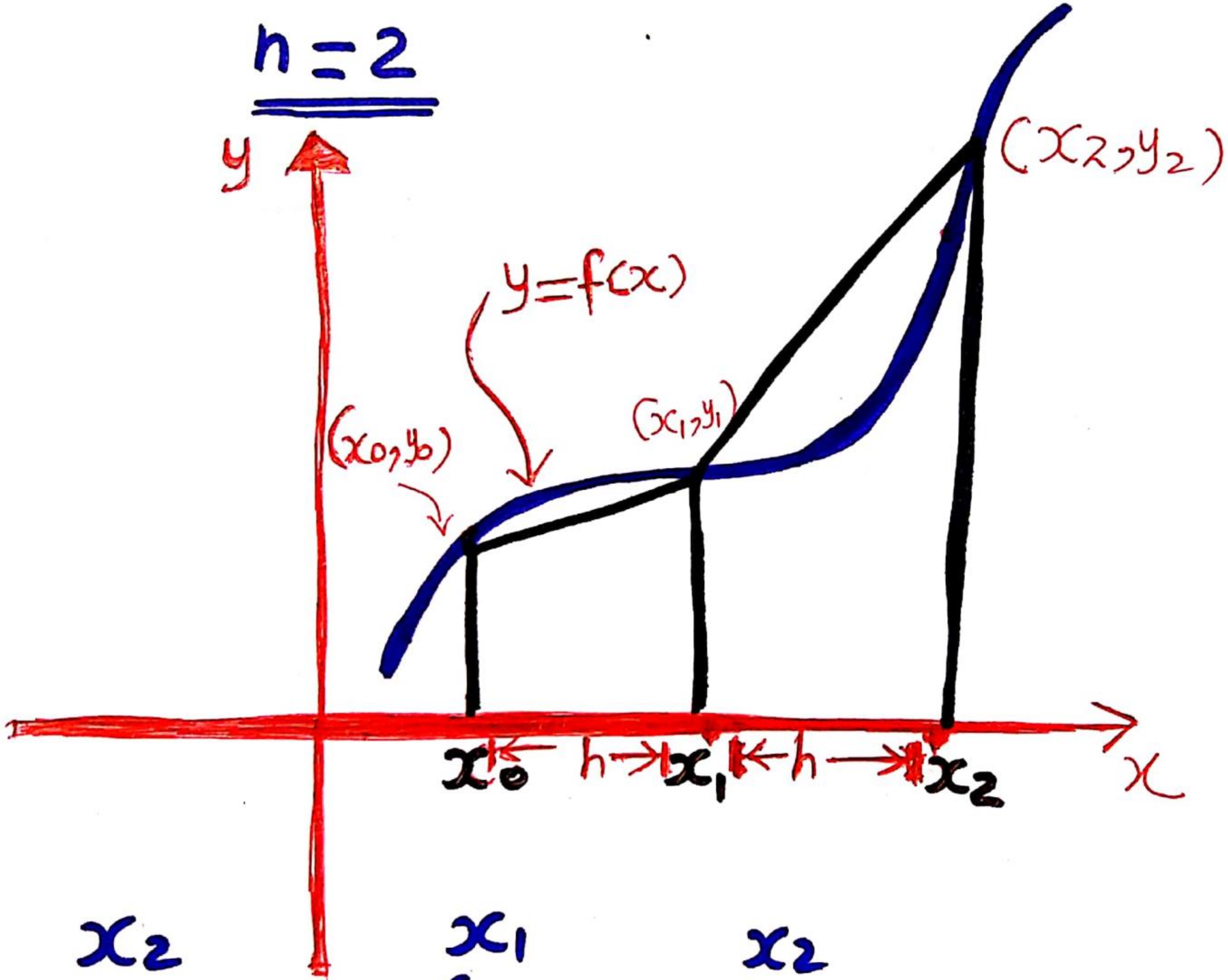
which is known as the *trapezoidal rule*.

The geometrical significance of this rule is that the curve $y = f(x)$ is replaced by n straight lines joining the points (x_0, y_0) and (x_1, y_1); (x_1, y_1) and (x_2, y_2), …, (x_{n-1}, y_{n-1}) and (x_n, y_n). The area bounded by the curve $y = f(x)$, the ordinates $x = x_0$ and $x = x_n$, and the x-axis is then approximately equivalent to the sum of the areas of the n trapeziums obtained.

Geometrical Interpretation of Trapezoidal Rule



$$\int_{x_0}^{x_1} y\,dx \simeq \frac{h}{2}(y_0+y_1)$$



$$\int_{x_0}^{x_2} y\,dx = \int_{x_0}^{x_1} y\,dx + \int_{x_1}^{x_2} y\,dx \simeq \frac{h}{2}(y_0+y_1)+\frac{h}{2}(y_1+y_2)$$

Simpson's Rule

This rule is obtained by putting $n = 2$ in the general formula. i.e. by replacing the curve by $n/2$ arcs of second-degree polynomials or parabolas.

We have then

$$\int_{x_0}^{x_2} y\,dx = 2h\left(y_0 + \Delta y_0 + \frac{1}{6}\Delta^2 y_0\right) = \frac{h}{3}(y_0 + 4y_1 + y_2).$$

<u>Note</u>: $\Delta^2(y_0) = \Delta(\Delta y_0) = \Delta(y_1 - y_0) = \Delta y_1 - \Delta y_0 = y_2 - y_1 - y_1 + y_0 = y_2 - 2y_1 + y_0$

Similarly,

$$\int_{x_2}^{x_4} y\,dx = \frac{h}{3}(y_2 + 4y_3 + y_4)$$

$$\vdots$$

and finally

$$\int_{x_{n-2}}^{x_n} y\,dx = \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n).$$

Summing up, we obtain

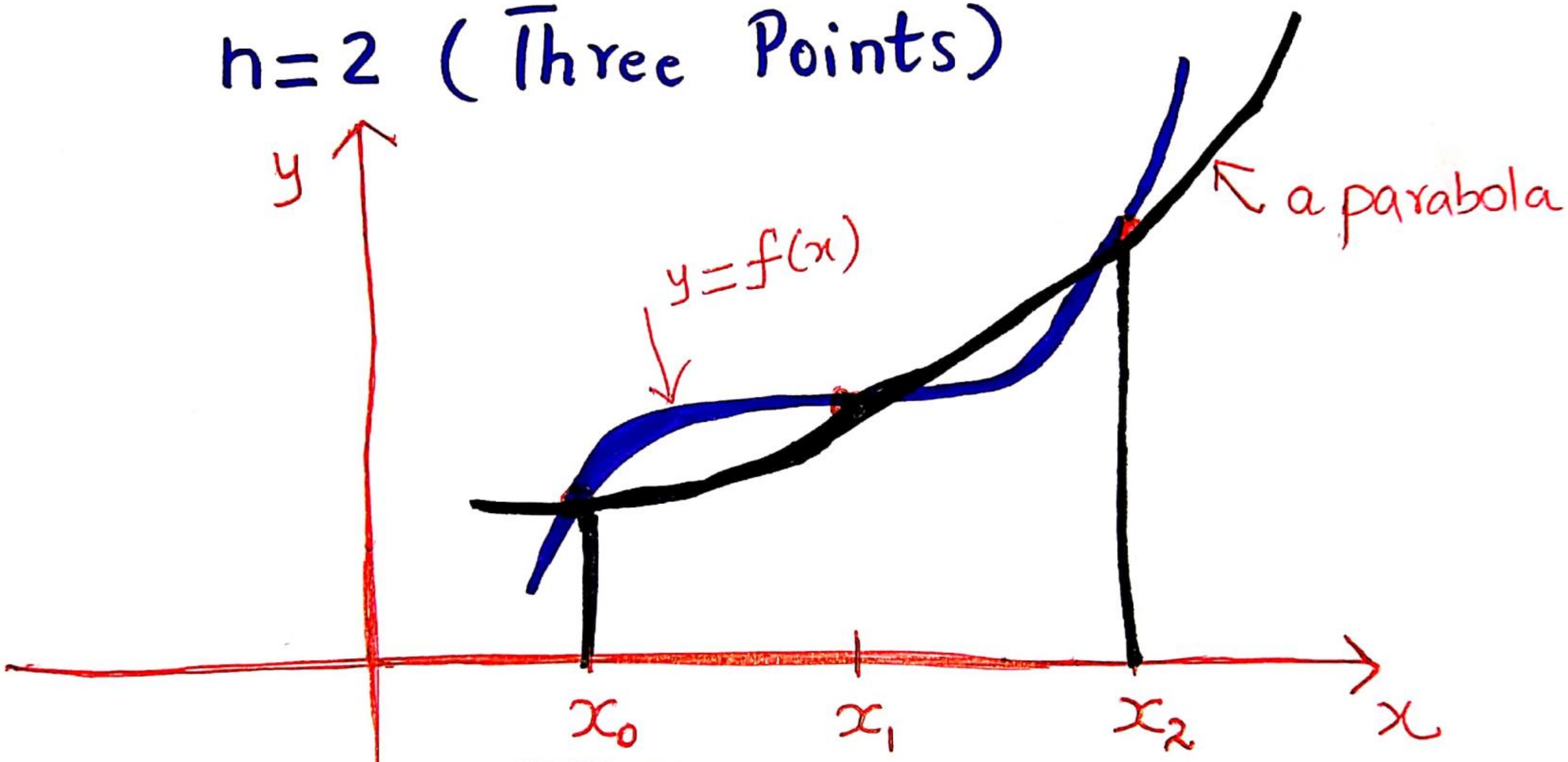
$$\boxed{\int_{x_0}^{x_n} y\,dx = \frac{h}{3}\left[y_0 + 4\left(y_1 + y_3 + y_5 + \cdots + y_{n-1}\right) + 2\left(y_2 + y_4 + y_6 + \cdots + y_{n-2}\right) + y_n\right],}$$

which is known as *Simpson's 1/3-rule*, or simply Simpson's rule.

It should be noted that this rule requires the division of the whole range into an even number of subintervals of width h.

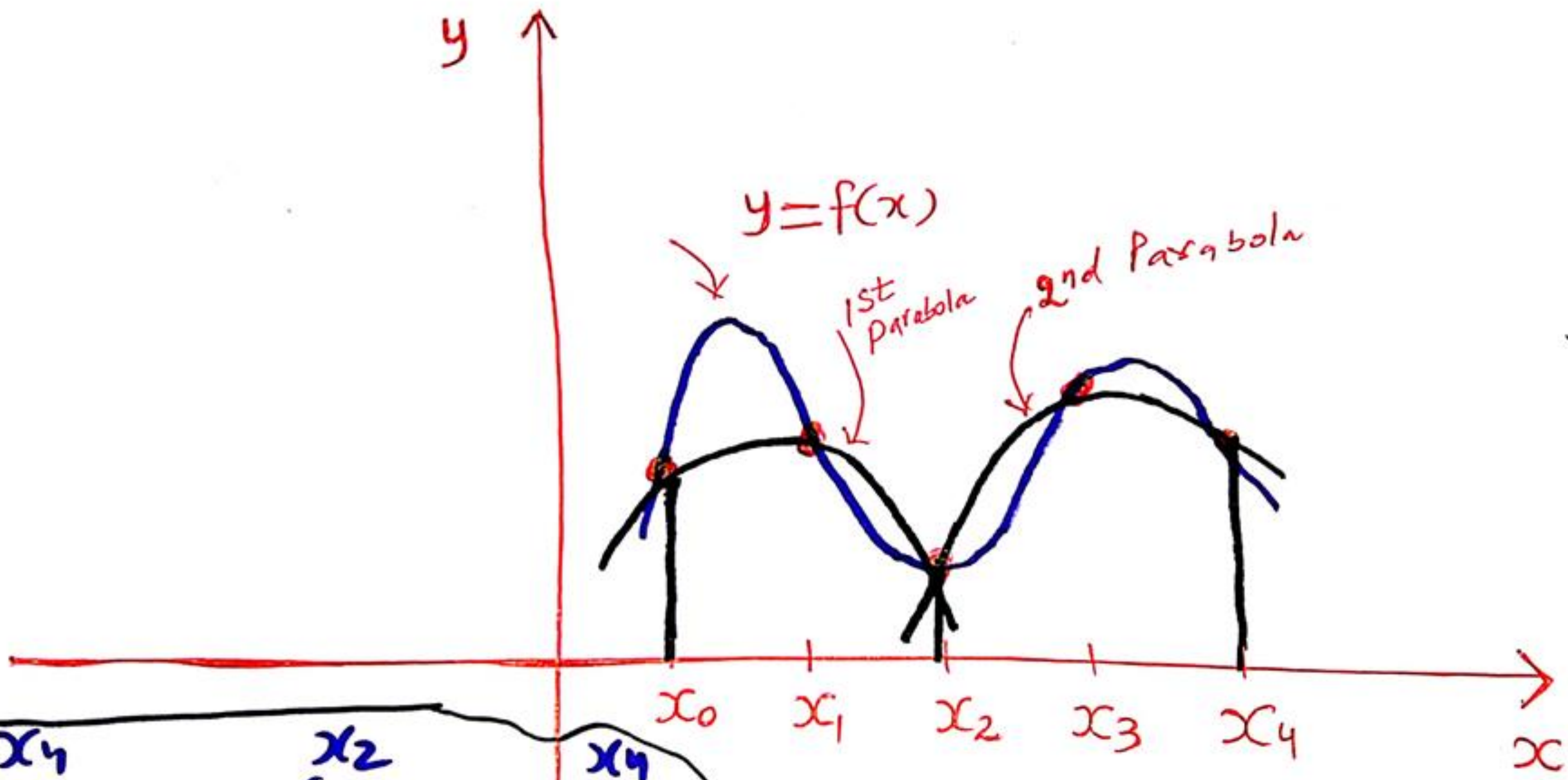
Geometrical Interpretation of Simpson's Rule

$n = 2$ (Three Points)



$$\boxed{\int_{x_0}^{x_2} y\,dx = \frac{h}{3}\left(y_0 + 4y_1 + y_2\right)}$$

$n=4$ (five Points)



$$\int_{x_0}^{x_4} y\,dx = \int_{x_0}^{x_2} y\,dx + \int_{x_2}^{x_4} y\,dx \approx \frac{h}{3}(y_0+4y_1+y_2)+\frac{h}{3}(y_2+4y_3+y_4)$$

Problem

Find, from the following table, the area bounded by the curve and the x-axis from $x = 7.47$ to $x = 7.52$

x	$f(x)$
7.47	1.93
7.48	1.95
7.49	1.98
7.50	2.01
7.51	2.03
7.52	2.06

Solution

$x_0 = 7.47$, $x_1 = 7.48$, $x_2 = 7.49$, $x_3 = 7.50$,

$x_4 = 7.51$, $x_5 = 7.52$ and

$y_0 = 1.93$, $y_1 = 1.95$, $y_2 = 1.98$, $y_3 = 2.01$,

$y_4 = 2.03$, $y_5 = 2.06$.

$h = 0.01$

From Trapezoidal Rule, We have

$$\text{Area} = \int_{7.47}^{7.52} f(x)\,dx \approx \frac{h}{2}\left[y_0 + 2(y_1 + y_2 + y_3 + y_4) + y_5 \right]$$

Substituting Values we get

$$\boxed{\text{Area} \simeq 0.0996}$$

Problem

Use Simpson's Rule to approximate the following integral from the data given below in the table.

$$\int_0^{20} f(x)dx$$

x	0	2	4	6	8	10	12	14	16	18	20
y	0	16	29	40	46	51	32	18	8	3	0

Solution

Here $x_0 = 0$, $x_1 = 2$, $x_2 = 4$, $x_3 = 6$, $x_4 = 8$, $x_5 = 10$,
$x_6 = 12$, $x_7 = 14$, $x_8 = 16$, $x_9 = 18$ and $x_{10} = 20$

Also $y_0 = 0$, $y_1 = 16$, $y_2 = 29$, $y_3 = 40$, $y_4 = 46$,
$y_5 = 51$, $y_6 = 32$, $y_7 = 18$, $y_8 = 8$, $y_9 = 3$ & $y_{10} = 0$.

$h = 2$

From Simpson's Rule We have

$$\int_0^{20} f(x)\,dx = \frac{h}{3}\left[y_0 + 4(y_1 + y_3 + y_5 + y_7 + y_9) + 2(y_2 + y_4 + y_6 + y_8) + y_{10}\right]$$

Substituting Values, we get

$$\boxed{\int_0^{20} f(x)\,dx = 494.667}$$

Practice Problems

Problem 1

Using the trapezium rule, find $\int_0^6 f(x)dx$, from the following set of values of x and $f(x)$.

x	0	1	2	3	4	5	6
$f(x)$	1.56	3.64	4.62	5.12	7.05	9.22	10.44

Problem 2

Use Trapezoidal Rule to approximate the integral $\int_0^6 e^{x^2}\,dx.$

Problem 3

Use Simpson's Rule to evaluate the integral in problem 1.

Problem 4

Use Simpson's Rule to evaluate the integral in problem 2.

Romberg Method for the trapezoidal Rule

Romberg method is a technique use for refining the approximation to the value of a definite integral. Let I denote the exact value of the definite integral i.e.

$$I = \int_a^b f(x)dx$$

First of all we will approximate I using trapezoidal Rule for different steps sizes such as $h, h/2, h/4, h/8$ etc. We will denote these approximations by $I(h), I(h/2), I(h/4)$ etc.

Romberg Method for Trapezoidal Rule

h	$I(h)$		
		$I^{(1)}(h) = \dfrac{4I(h/2) - I(h)}{3}$	
$h/2$	$I(h/2)$		$I^{(2)}(h) = \dfrac{16I^{(1)}(h/2) - I^{(1)}(h)}{15}$
		$I^{(1)}(h/2) = \dfrac{4I(h/4) - I(h/2)}{3}$	
$h/4$	$I(h/4)$		

Problem

(a)

Find the approximate value of $I = \int_0^1 \frac{dx}{1+x}$, *using the trapezium rule with* 2, 4 *and* 8 *equal subintervals.*

(b)

Apply the Romberg's method to improve the approximations to the values of the integrals.

Solution

With $N = 2, 4$ and 8, we have the following step lengths and nodal points.

$N = 2$: $h = \dfrac{b-a}{N} = \dfrac{1}{2}$. The nodes are 0, 0.5, 1.0.

$N = 4$: $h = \dfrac{b-a}{N} = \dfrac{1}{4}$. The nodes are 0, 0.25, 0.5, 0.75, 1.0.

$N = 8$: $h = \dfrac{b-a}{N} = \dfrac{1}{8}$. The nodes are 0, 0.125, 0.25, 0.375, 0.5, 0.675, 0.75, 0.875, 1.0.

We have the following tables of values.

$N = 2$:	x	0	0.5	1.0
	$f(x)$	1.0	0.666667	0.5

and similarly for **N=4** and **N=8**.

Using Trapezoidal Rule for all the three cases, we obtain

$I(1/2) = 0.708334$ $\quad$ $I(1/4) = 0.697024$ $\quad$ $I(1/8) = 0.694122$

Table for Romberg Method

1/2	0.708334		
		0.693254	
1/4	0.697024		0.693148
		0.693155	
1/8	0.694122		

A much refine approximation to I is 0.693148.

Numerical Solution of Ordinary Differential Equations

Taylor's Series Method

We can use Taylor's series method to solve a first order Initial Value Problem given by:

$$\frac{dy}{dx} = f(x, y), \qquad y(x_0) = y_0. \qquad \text{(I)}$$

If $y(x)$ is the exact solution of (I), then the Taylor's series for $y(x)$ about $x = x_0$ is given by

$$y(x) = y_0 + (x - x_0)y_0' + \frac{(x - x_0)^2}{2!} y_0'' + \frac{(x - x_0)^3}{3!} y_0''' + \cdots$$

(II)

By differentiating $\dfrac{dy}{dx} = y'(x) = f(x, y)$ several times we can get the values of $y'', y''', y^{(4)} \ldots$ The idea to get approximate solution from (II) is to calculate few derivatives at $x = x_0$,truncate the series after some terms and finally substitute values to get approximate $y(x)$.

Problem

From the Taylor series for $y(x)$, find $y(0.1)$ if $y(x)$ satisfies

$$y' = x - y^2 \quad \text{and} \quad y(0) = 1.$$

Solution

Let $y(x)$ is the exact solution of the given IVP.
The Taylor's Series of $y(x)$ about $x_0 = 0$ is

$$y(x) = 1 + x y_0' + \frac{x^2}{2!} y_0'' + \frac{x^3}{3!} y_0''' + \cdots \quad \text{——— (i)}$$

$$(\because \; x_0 = 0 \text{ and } y_0 = 1 \text{ are given})$$

As $y' = x - y^2 \Rightarrow y_0' = x_0 - y_0^2 = 0 - 1^2 = -1 \Rightarrow \boxed{y_0' = -1}$

So $y'' = 1 - 2y\frac{dy}{dx} = 1 - 2yy' \Rightarrow y_0'' = 1 - 2y_0 y_0' = 1 - 2 \cdot 1 \cdot -1$

$$\Rightarrow y_0'' = 1 + 2 = 3$$

$$\Rightarrow \boxed{y_0'' = 3}$$

$$y''' = 0 - 2[y y'' + y' \cdot y'] = -2[y y'' + (y')^2]$$

$$\Rightarrow y_0''' = -2[y_0 y_0'' + (y_0')^2] = -2[1 \cdot 3 + (-1)^2] = -2[3+1]$$

$$\Rightarrow \boxed{y_0''' = -8}$$

$$y^{(4)} = -2[\, y y''' + y'' y' + 2 y' \cdot y'' \,] = -2[y y''' + 3 y' y'']$$

$$\Rightarrow y_0^{(4)} = -2[y_0 y_0''' + 3 y_0' y_0''] = -2[1(-8) + 3(-1)(3)]$$

$$\Rightarrow y_0^{(4)} = -2[-8-9] = -2(-17) = 34 \Rightarrow \boxed{y_0^{(4)} = 34}$$

Substituting Values in (i) we get

$$y(x) = 1 + x(-1) + \frac{x^2}{2!}(3) + \frac{x^3}{3!}(-8) + \frac{x^4}{4!}(34) + \cdots$$

So

$$y(x) \simeq 1 + x + \frac{3}{2}x^2 - \frac{4}{3}x^3 + \frac{17}{12}x^4$$

Next

$$y(0.1) \simeq 1 + 0.1 + \frac{3}{2}(0.1)^2 - \frac{4}{3}(0.1)^3 + \frac{17}{12}(0.1)^4$$

$$\Rightarrow \boxed{y(0.1) \simeq 0.9138} \quad \text{Ans}$$

Practice Problem

Given that $\dfrac{dy}{dx} = 1 + xy$ with the initial condition $y(0) = 1.$

Compute $y(0.1)$ using Taylor's series method.

Another form of Taylor's series

Consider the Taylor's series of the function $y(x)$ about $x = x_0$

$$y(x) = y_0 + (x - x_0)y_0' + \frac{(x - x_0)^2}{2!}y_0'' + \frac{(x - x_0)^3}{3!}y_0''' + \cdots$$

Let $x = x_1 = x_0 + h$ then

$$y_1 = y(x_1) = y_0 + hy_0' + \frac{h^2}{2!}y_0'' + \frac{h^3}{3!}y_0''' + \cdots$$

Next to find y_2 we will first consider the Taylor's Series of $y(x)$

about $x = x_1$

$$y(x) = y_1 + (x - x_1) y_1' + \frac{(x - x_1)^2}{2!} y_1'' + \frac{(x - x_1)^3}{3!} y_1''' + \cdots$$

Let $x = x_2 = x_1 + h$ then

$$y_2 = y(x_2) = y_1 + h y_1' + \frac{h^2}{2!} y_1'' + \frac{h^3}{3!} y_1''' + \cdots$$

In general,

$$y_{n+1} = y_n + hy'_n + \frac{h^2}{2!} y''_n + \frac{h^3}{3!} y'''_n + \cdots \quad , n = 0, 1, 2, \ldots$$

Problem

Apply the Taylor's series method to find the value of y (1.1) and y (1.2) given that $\frac{dy}{dx} = xy^{\frac{1}{3}}, y(1) = y(1) = 1$ *taking the first three terms of the Taylor's series expansion.*

Solution

Given $\quad \dfrac{dy}{dx} = xy^{\frac{1}{3}},\ y_0 = 1,\ x_0 = 1,\ h = 0.1$

$$\longrightarrow \quad y'_0 = x_0\, y_0^{\frac{1}{3}} = 1.1^{\frac{1}{3}} = 1,$$

$$\frac{d^2y}{dx^2} = \frac{1}{3}\, xy^{\frac{-2}{3}}\ \frac{dy}{dx} + y^{\frac{1}{3}}$$

$$= \frac{1}{3}\, xy^{\frac{-2}{3}} \left(xy^{\frac{1}{3}} \right) + y^{\frac{1}{3}} = \frac{1}{3}\, x^2 y^{-\frac{1}{3}} + y^{\frac{1}{3}}$$

$$\Rightarrow \quad y_0'' = \frac{1}{3}1.1 + 1 = \frac{4}{3}.$$

Taking the first three terms of the Taylor's formula we get

$$y_{n+1} = y_n + hy_n' + \frac{h^2}{2!}y_n'', \quad n = 0,1,2,\ldots \qquad \text{(i)}$$

For *n=0* eq (i) is $\quad y_1 = y_0 = hy_0' + \frac{h^2}{2}y_0''$

Substituting values we get

$$y_1 = y(1.1) = 1 + (0.1) \times 1 + \frac{(0.1^2)}{2} \times \frac{4}{3} = 1.1066$$

$$\rightarrow \quad y_1(1.1) = 1.1066,$$

$$x_1 = x_0 + h = 1 + 0.1 = 1.1$$

$$y_1' = \left(x_1 \times y_1^{\frac{1}{3}} \right) = (1.1)(1.1066)^{\frac{1}{3}} = 1.138,$$

$$y_1'' = \frac{1}{3}x_1^2\, y_1^{\frac{1}{3}} + y_1^{\frac{1}{3}}$$

$$= \frac{1}{3}(1.1)^2(1.1066)^{\frac{-1}{3}} + (1.1066)^{\frac{1}{3}}$$

$$\longrightarrow \quad y_1'' = 1.4249.$$

Substituting in $\quad y_2 = y_1 + h y_1' + \dfrac{h^2}{2} y_1'',$

$$y_2 = y(1.2) = 1.1066 + 0.1 \times 1.138 + \frac{(0.1)^2}{2} \times 1.4249$$

$$\Rightarrow \quad y_2 = y(1.2) = 1.2275245 .$$

Practice Problem

Use Taylor series method of order four to solve

$$y' = x^2 + y^2, y(0) = 1$$

for $x \in [0, 0.4]$ with $h = 0.2$

Euler Method: We can use Euler Method to approximate solution of first-order IVP :

$$\frac{dy}{dx} = f(x, y), \qquad y(x_0) = y_0.$$

in some interval $[x_0, x_n]$ in the form of a set of tabulated values. If we divide the interval $[x_0, x_n]$ into N subintervals each of length $h = \frac{b-a}{N}$ and define mesh points $x_n = x_0 + nh, \quad n = 0, 1, 2, \ldots N,$ then the Euler formula is given by:

$$y_{n+1} = y_n + hf(x_n, y_n), \quad n = 0, 1, 2, \ldots, N-1.$$

Problem

Solve the initial value problem $yy' = x$, $y(0) = 1$, *using the Euler method in* $0 \le x \le 0.8$, *with* $h = 0.2$ *and* $h = 0.1$.

Solution

We have $\quad y' = f(x, y) = (x/y).$

Euler method gives $\quad y_{n+1} = y_n + hf(x_n, y_n) = y_{n+1} = y_n + \dfrac{hx_n}{y_n}, n = 0,1,2,\ldots$

Initial condition gives $\quad x_0 = 0, y_0 = 1.$

When $h = 0.2$, we get $$y_{n+1} = y_n + \frac{0.2x_n}{y_n}, n = 0,1,2,\ldots$$

We have the following results.

$$y(x_1) = y(0.2) \approx y_1 = y_0 + \frac{0.2\ x_0}{y_0} = 1.0.$$

$$y(x_2) = y(0.4) \approx y_2 = y_1 + \frac{0.2\ x_1}{y_1} = 1.0 + \frac{0.2(0.2)}{1.0} = 1.04.$$

$$y(x_3) = y(0.6) \approx y_3 = y_2 + \frac{0.2\, x_2}{y_2} = 1.04 + \frac{0.2(0.4)}{1.04} = 1.11692$$

$$y(x_4) = y(0.8) \approx y_4 = y_3 + \frac{0.2\, x_3}{y_3} = 1.11692 + \frac{0.2(0.6)}{1.11692} = 1.22436.$$

When $h = 0.1$, we get $y_{n+1} = y_n + \dfrac{0.1x_n}{y_n}, n = 0, 1, 2, \ldots$

We have the following results.

$$y(x_1) = y(0.1) \approx y_1 = y_0 + \frac{0.1x_0}{y_0} = 1.0.$$

$$y(x_2) = y(0.2) \approx y_2 = y_1 + \frac{0.1x_1}{y_1} = 1.0 + \frac{0.1(0.1)}{1.0} = 1.01.$$

$$y(x_3) = y(0.3) \approx y_3 = y_2 + \frac{0.1x_2}{y_2} = 1.01 + \frac{0.1(0.2)}{1.01} = 1.02980.$$

$$y(x_4) = y(0.4) \approx y_4 = y_3 + \frac{0.1x_3}{y_3} = 1.0298 + \frac{0.1(0.3)}{1.0298} = 1.05893.$$

$$y(x_5) = y(0.5) \approx y_5 = y_4 + \frac{0.1x_4}{y_4} = 1.05893 + \frac{0.1(0.4)}{1.05893} = 1.09670.$$

$$y(x_6) = y(0.6) \approx y_6 = y_5 + \frac{0.1x_5}{y_5} = 1.0967 + \frac{0.1(0.5)}{1.0967} = 1.14229.$$

$$y(x_7) = y(0.7) \approx y_7 = y_6 + \frac{0.1x_6}{y_6} = 1.14229 + \frac{0.1(0.6)}{1.14229} = 1.19482.$$

$$y(x_8) = y(0.8) \approx y_8 = y_7 + \frac{0.1x_7}{y_7} = 1.19482 + \frac{0.1(0.7)}{1.19482} = 1.25341.$$

Practice Problem

Find $y(4.4)$, by Euler's method taking $h = 0.2$ from the differential equation $\frac{dy}{dx} = \frac{2 - y^2}{5x}$, $y = 1$ when $x = 4$.

Modified Euler Method

The idea is to use the formula from Euler's method to obtain a first approximation to the solution $y(x_{n+1})$. We denote this approximation by y^*_{n+1}, so that

$$y^*_{n+1} = y_n + hf(x_n, y_n). \qquad \text{(i)}$$

We now improve (or "correct") this approximation by using the formula of modified Euler method which is

$$y_{n+1} = y_n + \tfrac{1}{2}h[f(x_n, y_n) + f(x_{n+1}, y^*_{n+1})]. \qquad \text{(ii)}$$

Problem

Apply the modified Euler method with $h = 0.1$ to determine an approximation to the solution to the initial-value problem

$$y' = y - x, \qquad y(0) = \frac{1}{2}$$

at $x = 1$.

Solution

Taking $h = 0.1$ and $f(x, y) = y - x$ in the modified Euler method yields

$$\begin{aligned} y^*_{n+1} &= y_n + 0.1(y_n - x_n), \\ y_{n+1} &= y_n + 0.05(y_n - x_n + y^*_{n+1} - x_{n+1}). \end{aligned}$$

For $n = 0$, Eq (i) becomes

$$y_1^* = y_0 + 0.1(y_0 - x_0)$$

Substituting values, we get

$$y_1^* = \frac{1}{2} + 0.1\left(\frac{1}{2} - 0\right) = \frac{1}{2} + \frac{0.1}{2} = 0.55 \Rightarrow \boxed{y_1^* = 0.55}$$

For $n = 0$, Eq (ii) becomes

$$y_1 = y_0 + 0.05(y_0 - x_0 + y_1^* - x_1)$$

Substituting values we get

$$y_1 = \frac{1}{2} + 0.05\left(\frac{1}{2} - 0 + 0.55 - 0.1\right) \Rightarrow \boxed{y_1 = 0.5475}$$

or $\boxed{y(0.1) = 0.5475}$

The following Table shows the results obtained by applying modified Euler Method. The required solution at x=1 is

$\boldsymbol{y(1) = 0.642959}$.

n	x_n	y_n
1	0.1	0.5475
2	0.2	0.589487
3	0.3	0.625384
4	0.4	0.654549
5	0.5	0.676277
6	0.6	0.689786
7	0.7	0.694213
8	0.8	0.688605
9	0.9	0.671909
10	1.0	0.642959

Practice Problem

Find $y(4.4)$, by Euler's modified method taking $h = 0.2$ from the differential equation $\frac{dy}{dx} = \frac{2 - y^2}{5x}$, $y = 1$ when $x = 4$.

Improved Euler's Method

$$y_{i+1} = y_i + hf\left(x_i + \frac{h}{2}, y_i + \frac{h}{2} f(x_i, y_i)\right)$$

PICARD'S METHOD OF SUCCESSIVE APPROXIMATIONS

$$y^{(n)} = y_0 + \int_{x_0}^{x} f(x, y^{(n-1)})\, dx \qquad \text{with } y^{(0)} = y_0$$

Runge-Kutta Method

The method is very simple. It is named after two German mathematicians Carl Runge (1856–1927) and Wilhelm Kutta (1867–1944). It was developed to avoid the computation of higher order derivations which the Taylor's method may involve. In the place of these derivatives extra values of the given function $f(x, y)$ are used. Runge–Kutta methods are designed to give greater accuracy.

The **fourth-order Runge-Kutta method** for approximating the solution to the initial-value problem

$$y' = f(x, y), \quad y(x_0) = y_0$$

at the points $x_{n+1} = x_0 + nh \; (n = 0, 1, \ldots)$ is

$$y_{n+1} = y_n + \frac{h}{6}\left(k_1 + 2k_2 + 2k_3 + k_4\right)$$

where

$$k_1 = f(x_n, y_n),$$

$$k_2 = f\left(x_n + \frac{h}{2}, y_n + \frac{hk_1}{2}\right),$$

$$k_3 = f\left(x_n + \frac{h}{2}, y_n + \frac{hk_2}{2}\right),$$

$$k_4 = f\left(x_n + h, y_n + hk_3\right).$$

Problem

Use the RK4 method with $h = 0.1$ to obtain an approximation to $y(1.5)$ for the solution of $y' = 2xy$, $y(1) = 1$.

Solution

For the sake of illustration let us compute the case when $n = 0$.

$$k_1 = f(x_0, y_0) = 2x_0y_0 = 2$$

$$k_2 = f\left(x_0 + \tfrac{1}{2}(0.1), y_0 + \tfrac{1}{2}(0.1)2\right)$$
$$= 2\left(x_0 + \tfrac{1}{2}(0.1)\right)\left(y_0 + \tfrac{1}{2}(0.2)\right) = 2.31$$

$$k_3 = f\left(x_0 + \tfrac{1}{2}(0.1), y_0 + \tfrac{1}{2}(0.1)2.31\right)$$

$$= 2\left(x_0 + \tfrac{1}{2}(0.1)\right)\left(y_0 + \tfrac{1}{2}(0.231)\right) = 2.34255$$

$$k_4 = f(x_0 + (0.1), y_0 + (0.1)2.34255)$$

$$= 2(x_0 + 0.1)(y_0 + 0.234255) = 2.715361$$

and therefore

$$y_1 = y_0 + \frac{0.1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$$= 1 + \frac{0.1}{6}(2 + 2(2.31) + 2(2.34255) + 2.715361) = 1.23367435.$$

The remaining calculations are summarized in a table in next slide.

RK4 Method with $h = 0.1$

x_n	y_n
1.00	1.0000
1.10	1.2337
1.20	1.5527
1.30	1.9937
1.40	2.6116
1.50	3.4902

Practice Problem

Use the RK4 method with $h = 0.1$ to obtain approximation of the indicated value.

(a) $y' = 2x - 3y + 1, y(1) = 5; \quad y(1.5)$

(b) $y' = xy + \sqrt{y}, y(0) = 1; \quad y(0.5)$

Solving a second-order IVP using Taylor's Series

Problem

Given the differential equation

$$y'' - xy' - y = 0$$

with the conditions $y(0) = 1$ and $y'(0) = 0$, use Taylor's series method to determine the value of $y\,(0.1)$.

Solution

Given that $y(0) = 1 \Rightarrow \boxed{x_0 = 0}$ and $\boxed{y_0 = 1}$

Also $y'(0) = 0 \Rightarrow \boxed{y'_0 = 0}$

<u>DE</u>

$$y'' - xy' - y = 0 \Rightarrow \boxed{y'' = xy' + y} \quad \text{—— (i)}$$

$$y''_0 = x_0 y'_0 + y_0 = 0 + 1 = 1$$

$$\Rightarrow \boxed{y''_0 = 1}$$

Next to find $y'''_0, y^{(IV)}_0, \ldots$, we will first differentiate (i) several times.

Consider (i) $y'' = xy' + y$

$$y''' = y' + xy'' + y' = xy'' + 2y'$$

$$\Rightarrow y''' = xy'' + 2y' \Rightarrow y_0''' = x_0 y_0'' + 2y_0' = 0 + 0$$

$$\Rightarrow \boxed{y_0''' = 0}$$

$$y^{(iv)} = y'' + xy''' + 2y'' = xy''' + 3y''$$

$$\Rightarrow y_0^{(iv)} = x_0 y_0''' + 3y_0'' = 3 \Rightarrow \boxed{y_0^{(iv)} = 3}$$

$$y^{(v)} = y''' + y''' + xy^{(iv)} + 2y''' = xy^{(iv)} + 4y'''$$

$$y^{(v)} = x_0 y_0^{(iv)} + 4y_0''' = 0 + 0 = 0 \Rightarrow \boxed{y_0^{(v)} = 0}$$

Also $y^{(vi)} = y^{(iv)} + xy^{(v)} + 4y^{(iv)} = xy^{(v)} + 5y^{(iv)}$

$\Rightarrow y_0^{(vi)} = xy_0^{(v)} + 5y_0^{(iv)} = 0 + 5(3) = 15$

$\Rightarrow \boxed{y_0^{(vi)} = 15}$

Let $y(x)$ is the solution of the given IVP. The Taylor's Series of $y(x)$ about $x_0 = 0$ is

$$y(x) = y_0 + (x-0)y_0' + \frac{(x-0)^2}{2!}y_0'' + \cdots$$

$$\Rightarrow y(x) = y_0 + x y_0' + \frac{x^2}{2!}y_0'' + \cdots$$

Substituting Values, we get

$$y(x) = 1 + x(0) + \frac{x^2}{2!}(1) + \frac{x^3}{3!}(0) + \frac{x^4}{4!}(3) + \frac{x^5}{5!}(0) + \frac{x^6}{6!}(15) + \cdots$$

$$\Rightarrow y(x) \simeq 1 + \frac{x^2}{2!} + \frac{3x^4}{4!} + \frac{15}{6!}x^6$$

and $y(0\cdot1) \simeq 1 + \dfrac{0\cdot1^2}{2!} + \dfrac{3(0\cdot1)^4}{4!} + \dfrac{15}{6!}(0\cdot1)^6$

Systems of Differential Equations

A **system of ordinary differential equations** is two or more equations involving the derivatives of two or more unknown functions of a single independent variable. For example, if x and y denote dependent variables and t denotes the independent variable, then a system of two first-order differential equations is given by

$$\frac{dx}{dt} = f(t, x, y) \qquad \text{(i)}$$

$$\frac{dy}{dt} = g(t, x, y).$$

A **solution** of a system such as **(i)** is a pair of differentiable functions $x = \phi_1(t)$, $y = \phi_2(t)$, defined on a common interval I, that satisfy each equation of the system on this interval.

Numerical Solution of a System

The solution of a system of the form

$$\boxed{\begin{aligned} x' &= f(t, x, y) \\ y' &= g(t, x, y) \\ x(t_0) = x_0, &\quad y(t_0) = y_0, \end{aligned}}$$

can be approximated by a version of Euler's **i.e.**

$$
\boxed{
\begin{aligned}
x_{n+1} &= x_n + hf(t_n, x_n, y_n)\\
y_{n+1} &= y_n + hg(t_n, x_n, y_n).
\end{aligned}
}
$$

We can also solve ☆ using Taylor series method of various order.

For instance the scheme for Taylor series of order 3 method will be

$$x_{n+1} = x_n + hx'_n + \frac{h^2}{2!}x''_n + \frac{h^3}{3!}x'''_n$$

$$y_{n+1} = y_n + hy'_n + \frac{h^2}{2!}y''_n + \frac{h^3}{3!}y'''_n$$

Problem

Use Taylor series method of order 2 to approximate $x(0.2)$ and $y(0.2)$ using $h = 0.1$.

$$x' = 6x + y + 6t$$

$$y' = 4x + 3y - 10t + 4$$

$$x(0) = 0.5, \quad y(0) = 0.2$$

Solution

Scheme of Taylor series method of order 2 is :

$$x_{n+1} = x_n + h x_n' + \frac{h^2}{2!} x_n'' \quad \text{———} \quad (i)$$

$$y_{n+1} = y_n + h y_n' + \frac{h^2}{2!} y_n''$$

Given that $h = 0.1$, $x_0 = 0.5$, $t_0 = 0$ and $y_0 = 0.2$.

For $\underline{n = 0}$ (i) becomes

$$x_1 = x_0 + h x_0' + \frac{h^2}{2!} x_0'' \quad \text{———} \quad (ii)$$

$$y_1 = y_0 + h y_0' + \frac{h^2}{2!} y_0''$$

As given that

$$x' = 6x + y + 6t \Rightarrow \boxed{x'' = 6x' + y' + 6}$$

$$y' = 4x + 3y - 10t + 4 \Rightarrow \boxed{y'' = 4x' + 3y' - 10}$$

$$x'_0 = 6x_0 + y_0 + 6t_0 = 6(0.5) + 0.2 + 6(0)$$

$$\Rightarrow \boxed{x'_0 = 3.2}$$

$$y'_0 = 4x_0 + 3y_0 - 10t + 4$$

$$= 4(0.5) + 3(0.2) - 10(0) + 4$$

$$\Rightarrow \boxed{y'_0 = 6.6}$$

$$x''_0 = 6x'_0 + y'_0 + 6 = 6(3.2) + 6.6 + 6$$

$$\boxed{x''_0 = 31.8}$$

$$y''_0 = 4x'_0 + 3y'_0 - 10 = 4(3.2) + 3(6.6) - 10$$

$$\Rightarrow \boxed{y''_0 = 22.6}$$

Substituting values in (ii) we get

$$x_1 = 0.5 + (0.1)(3.2) + \frac{(0.1)^2}{2!}(31.8) = 0.979$$

$$y_1 = 0.2 + (0.1)(6.6) + \frac{(0.1)^2}{2!}(22.6) =$$

So

$$\boxed{x_1 = x(0.1) = 0.979 \quad \text{and} \quad y_1 = y(0.1) = 0.973}$$

For $\underline{n=1}$ (i) becomes

$$x_2 = x_1 + h x_1' + \frac{h^2}{2!} x_1'' \quad \text{——— (iii)}$$

$$y_2 = y_1 + h y_1' + \frac{h^2}{2!} y_1''$$

Now

$$x_1' = 6x_1 + y_1 + 6t_1$$

Substituting values we get

$$x_1' = 6(0.979) + 0.973 + 6(0.1)$$

$$\Rightarrow \boxed{x_1' = 7.447}$$

Next

$$y_1' = 4x_1 + 3y_1 - 10t_1 + 4$$

$$= 4(0.979) + 3(0.973) - 10(0.1) + 4$$

$$\Rightarrow \boxed{y_1' = 9.835}$$

Also

$$x_1'' = 6x_1' + y_1' + 6 = 6(7.447) + 9.835 + 6$$

$$\Rightarrow \boxed{x_1'' = 60.517}$$

Next

$$y_1'' = 4x_1' + 3y_1' - 10 = 4(7.447) + 3(9.835) - 10$$

$$\Rightarrow \boxed{y_1'' = 49.293}$$

Substituting values in (iii) we get

$$x_2 = 0.979 + (0.1)(7.447) + \frac{(0.1^2)}{2!}(60.517)$$

$$y_2 = 0.973 + (0.1)(9.835) + \frac{(0.1)^2}{2!}(49.293)$$

So

$$\boxed{\begin{aligned} x_2 &= x(0.2) = 2.026285 \\ y_2 &= y(0.2) = 2.202965 \end{aligned}}$$

Required results.

Practice Problem

Use Taylor series method of order 3 to approximate $x(0.2)$ and $y(0.2)$ using $h = 0.1$.

$$
\begin{aligned}
x' &= -y + t \\
y' &= x - t \\
x(0) &= -3, \quad y(0) = 5
\end{aligned}
$$

MULTISTEP METHODS

All the methods discussed so far in the previous sections of this chapter have one thing in common, that is, y_{i+1} was computed knowing only y_i and values of f and its derivatives. So, these are called **single-step methods**. On the other hand, if the knowledge of $y_i, y_{i-1}, \ldots, y_{i-k+1}$ is required for the computation of y_{i+1}, the method is referred to as a **multistep method**.

A *predictor* formula is used to predict the value of y_{i+1} and then a *corrector* formula is used to improve the value of y_{i+1} .

Adams-Bashforth Methods (Predictor Method)

To solve a first order Initial Value Problem we can use Adams-Bashforth four-step method (AB4) is given by

$$y_{i+1} = y_i + \frac{h}{24}\left[55 f_i - 59 f_{i-1} + 37 f_{i-2} - 9 f_{i-3}\right]$$

provided that the values of y_0, y_1, y_2, y_3 are known.

Remark The required starting values for the application of the Adams-Bashforth methods are obtained by using any single step method like Euler's method, Taylor series method or Runge-Kutta method.

Problem

Find the approximate value of $y(0.4)$ *using the Adams-Bashforth method of fourth order for the initial value problem*

$$y' = x + y^2,\ y(0) = 1$$

with $h = 0.1$. *Calculate the starting values using the Euler's method with the same step length.*

Solution

The Adams-Bashforth method of fourth order is given by

$$y_{i+1} = y_i + \frac{h}{24}\left[55 f_i - 59 f_{i-1} + 37 f_{i-2} - 9 f_{i-3}\right]$$

We need the starting values, y_0, y_1, y_2, y_3.

The initial condition gives $y_0 = 1$.

Euler's method is given by

$$y_{i+1} = y_i + hf(x_i, y_i).$$

We obtain the following starting values.

$y(0.1) \approx y_1 = 1.1.$ $\quad$ $y(0.2) \approx y_2 = 1.231.$

$y(0.3) \approx y_3 = 1.402536.$

$x_3 = 0.3,\ y_3 = 1.402536, \quad y_3' = 2.267107.$

Now, we apply the given Adams-Bashforth method. We have

For i=3, we obtain

$$y(0.4) = y_4 = y_3 + \frac{h}{24}\left[55 f_3 - 59 f_2 + 37 f_1 - 9 f_0\right]$$

$$= 1.402536 + \frac{0.1}{24}\ [55(2.267107) - 59(1.715361) + 37(1.31) - 9(1)]$$

$$= 1.664847.$$

Adams-Moulton Methods (Corrector Methods)

Fourth order Adams-Moulton Method is given by

$$y_{i+1} = y_i + \frac{h}{24}\left[9 f_{i+1} + 19 f_i - 5 f_{i-1} + f_{i-2}\right]$$

Problem

Using the Adams-Bashforth predictor-corrector equations, evaluate $y(1.4)$, *if* y *satisfies*

$$\frac{dy}{dx} + \frac{y}{x} = \frac{1}{x^2}$$

and $y(1) = 1$, $y(1.1) = 0.996$, $y(1.2) = 0.986$, $y(1.3) = 0.972$.

Solution

The Adams-Bashforth predictor-corrector method is given by

Predictor P: Adams-Bashforth method of fourth order.

$$y_{i+1}^{(p)} = y_i + \frac{h}{24}[55f_i - 59f_{i-1} + 37f_{i-2} - 9f_{i-3}].$$

Corrector C: Adams-Moulton method of fourth order.

$$y_{i+1}^{(c)} = y_i + \frac{h}{24}[9f(x_{i+1}, y_{i+1}^{(p)}) + 19f_i - 5f_{i-1} + f_{i-2}]$$

With $h = 0.1$, we are given the values

$$y(1) = 1,\ y(1.1) = 0.996,\ y(1.2) = 0.986,\ y(1.3) = 0.972.$$

We have $f(x, y) = \dfrac{1}{x^2} - \dfrac{y}{x}$.

Predictor application

For $i = 3$, we obtain

$$y_4^{(0)} = y_4^{(p)} = y_3 + \frac{h}{24}[55f_3 - 59f_2 + 37f_1 - 9f_0]$$

We have

$$f_0 = f(x_0, y_0) = f(1, 1) = 1 - 1 = 0,$$

$$f_1 = f(x_1, y_1) = f(1.1, 0.996) = -0.079008,$$

$$f_2 = f(x_2, y_2) = f(1.2, 0.986) = -0.127222,$$

$$f_3 = f(x_3, y_3) = f(1.3, 0.972) = -0.155976.$$

$$y_4^{(0)} = 0.972 + \frac{0.1}{24}\left[55(-0.155976) - 59(-0.127222) + 37(-0.079008) - 9(0)\right]$$

$$= 0.955351.$$

Corrector application

Now, $f(x_4, y_4^{(0)}) = f(1.4, 0.955351) = -0.172189.$

First iteration

$$y_4^{(1)} = y_4^{(c)} = y_3 + \frac{h}{24}[9f(x_4, y_4^{(0)}) + 19f_3 - 5f_2 + f_1]$$

$$= 0.972 + \frac{0.1}{24}[9(-0.172189) + 19(-0.155976) - 5(-0.127222)$$

$$+ (-0.079008)] = 0.955516.$$

Second iteration

$$f(x_4, y_4^{(1)}) = f(1.4, 0.955516) = -0.172307.$$

$$y_4^{(2)} = y_3 + \frac{h}{24}[9f(x_4, y_4^{(1)}) + 19f_3 - 5f_2 + f_1]$$

$$= 0.972 + \frac{0.1}{24}[9(-0.172307) + 19(-0.155976) - 5(-0.127222)$$

$$+ (-0.079008)] = 0.955512.$$

Now, $\quad | \; y_4^{(2)} - y_4^{(1)} \; | = | \; 0.955512 - 0.955516 \; | = 0.000004.$

Therefore, $\quad y(1.4) = 0.955512.$

Milne's predictor-corrector method

Milne's predictor-corrector method is given by

Predictor P:

$$y_{i+1}^{(p)} = y_{i-3} + \frac{4h}{3}[2f_i - f_{i-1} + 2f_{i-2}] \, .$$

Corrector C:

$$y_{i+1}^{(c)} = y_{i-1} + \frac{h}{3}[f(x_{i+1}, y_{i+1}^{(p)}) + 4f_i + f_{i-1}]$$

Problem

Using Milne's predictor-corrector method, find $y(0.4)$ for the initial value problem

$y' = x^2 + y^2,\ y(0) = 1,\ \ with\ h = 0.1.$

Calculate all the required initial values by Euler's method.

Solution

Milne's predictor-corrector method is given by

Predictor P: $$y_{i+1}^{(p)} = y_{i-3} + \frac{4h}{3}\left[2f_i - f_{i-1} + 2f_{i-2}\right].$$

Corrector C: $$y_{i+1}^{(c)} = y_{i-1} + \frac{h}{3}\left[f(x_{i+1}, y_{i+1}^{(p)}) + 4f_i + f_{i-1}\right].$$

We are given that

$$f(x, y) = x^2 + y^2, x_0 = 0, y_0 = 1.$$

Euler's method gives

$$y_{i+1} = y_i + hf(x_i, y_i) = y_i + 0.1(x_i^2 + y_i^2).$$

With $x_0 = 0$, $y_0 = 1$, we get

$$y_1 = y_0 + 0.1\ (x_0^2 + y_0^2) = 1.0 + 0.1\ (0 + 1.0) = 1.1.$$

$$y_2 = y_1 + 0.1\ (x_1^2 + y_1^2) = 1.1 + 0.1\ (0.01 + 1.21) = 1.222.$$

$$y_3 = y_2 + 0.1\ (x_2^2 + y_2^2) = 1.222 + 0.1[0.04 + (1.222)^2] = 1.375328.$$

Predictor application

For $i = 3$, we obtain

$$y_4^{(0)} = y_4^{(p)} = y_0 + \frac{4(0.1)}{3}[2f_3 - f_2 + 2f_1]$$

We have

$$f_1 = f(x_1, y_1) = f(0.1,\ 1.1) = 1.22,$$

$$f_2 = f(x_2, y_2) = f(0.1,\ 1.222) = 1.533284,$$

$$f_3 = f(x_3, y_3) = f(0.3,\ 1.375328) = 1.981527.$$

$$y_4^{(0)} = 1.0 + \frac{0.4}{3}\ [2(1.981527) - 1.533284 + 2(1.22)] = 1.649303.$$

Corrector application

First iteration For $i = 3$, we get

$$y_4^{(1)} = y_2 + \frac{0.1}{3}[f(x_4, y_4^{(0)}) + 4f_3 + f_2]$$

Now, $f(x_4, y_4^{(0)}) = f(0.4, 1.649303) = 2.880200.$

$$y_4^{(1)} = 1.222 + \frac{0.1}{3}\ [2.880200 + 4(1.981527) + 1.533284] = 1.633320.$$

Second iteration

$$y_4^{(2)} = y_2 + \frac{0.1}{3}[f(x_4, y_4^{(1)}) + 4f_3 + f_2]$$

Now, $f(x_4, y_4^{(1)}) = f(0.4,\ 1.633320) = 2.827734.$

$$y_4^{(2)} = 1.222 + \frac{0.1}{3}\ [2.827734 + 4(1.981527) + 1.533284] = 1.631571.$$

We have $|\ y_4^{(2)} - y_4^{(1)}\ | = |\ 1.631571 - 1.633320\ | = 0.001749.$

Third iteration

$$y_4^{(3)} = y_2 + \frac{0.1}{3}[f(x_4, y_4^{(2)}) + 4f_3 + f_2]$$

Now, $f(x_4, y_4^{(2)}) = f(0.4,\ 1.631571) = 2.822024.$

$$y_4^{(3)} = 1.222 + \frac{0.1}{3}\ [2.822024 + 4(1.981527) + 1.533284] = 1.631381.$$

We have $|\ y_4^{(3)} - y_4^{(2)}\ | = |\ 1.631381 - 1.631571\ | = 0.00019.$

The required result can be taken as $y(0.4) \approx 1.63138.$

Boundary Value Problems

Finite Difference Approximations

The Taylor series expansion, centered at a point a, of a function $y(x)$ is

$$y(x) = y(a) + y'(a)\frac{x-a}{1!} + y''(a)\frac{(x-a)^2}{2!} + y'''(a)\frac{(x-a)^3}{3!} + \cdots$$

If we set $h = x - a$, then the preceding line is the same as

$$y(x) = y(a) + y'(a)\frac{h}{1!} + y''(a)\frac{h^2}{2!} + y'''(a)\frac{h^3}{3!} + \cdots$$

OR

$$y(a+h) = y(a) + y'(a)\frac{h}{1!} + y''(a)\frac{h^2}{2!} + y'''(a)\frac{h^3}{3!} + \cdots$$

OR

$$y(x+h) = y(x) + y'(x)h + y''(x)\frac{h^2}{2} + y'''(x)\frac{h^3}{6} + \cdots \qquad (1)$$

$$y(x-h) = y(x) - y'(x)h + y''(x)\frac{h^2}{2} - y'''(x)\frac{h^3}{6} + \cdots. \qquad (2)$$

Subtracting (1) and (2) also gives

$$y'(x) \approx \frac{1}{2h}[y(x+h) - y(x-h)]. \qquad (3)$$

adding (1) and (2) we obtain an approximation for the second derivative $y''(x)$:

$$y''(x) \approx \frac{1}{h^2}[y(x+h) - 2y(x) + y(x-h)]. \qquad (4)$$

The right sides of (3), (4), are called **difference quotients**.

The expressions $y(x+h) - y(x-h)$ and $y(x+h) - 2y(x) + y(x-h)$

are called **finite differences**.

Finite Difference Method

Consider now a linear second-order boundary-value problem

$$y'' + P(x)y' + Q(x)y = f(x), \qquad y(a) = \alpha, \qquad y(b) = \beta.$$

Suppose $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$ represents a regular partition of the interval $[a, b]$; that is, $x_i = a + ih$, where $i = 0, 1, 2, \ldots, n$ and $h = (b - a)/n$. The points

$$x_1 = a + h, \qquad x_2 = a + 2h, \qquad \ldots, \qquad x_{n-1} = a + (n - 1)h,$$

are called **interior mesh points** of the interval $[a, b]$.

If we let

$$y_i = y(x_i), \qquad P_i = P(x_i), \qquad Q_i = Q(x_i), \qquad \text{and} \qquad f_i = f(x_i)$$

and if y'' and y' in ☆ are replaced by the **difference quotients** we get

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + P_i \frac{y_{i+1} - y_{i-1}}{2h} + Q_i y_i = f_i$$

or, after simplifying,

$$\left(1 + \frac{h}{2}P_i\right)y_{i+1} + (-2 + h^2Q_i)\,y_i + \left(1 - \frac{h}{2}P_i\right)y_{i-1} = h^2f_i\,. \tag{5}$$

The last equation, known as a **finite difference equation**, is an approximation to the differential equation. It enables us to approximate the solution $y(x)$ of at the interior mesh points $x_1, x_2, \ldots, x_{n-1}$ of the interval $[a, b]$. By letting i take on the values $1, 2, \ldots, n-1$ in (5), we obtain $n - 1$ equations in the $n - 1$ unknowns $y_1, y_2, \ldots, y_{n-1}$. Bear in mind that we know y_0 and y_n, since these are the prescribed boundary conditions $y_0 = y(x_0) = y(a) = \alpha$ and $y_n = y(x_n) = y(b) = \beta$.

Problem

Use the difference equation with $n = 4$ to approximate the solution of the boundary-value problem

$$y'' - 4y = 0, \qquad y(0) = 0, \qquad y(1) = 5.$$

Solution

we identify $P(x) = 0$, $Q(x) = -4$, $f(x) = 0$, and $h = (1 - 0)/4 = \frac{1}{4}$.

Hence the difference equation is

$$y_{i+1} - 2.25y_i + y_{i-1} = 0. \tag{6}$$

Now the interior points are $x_1 = 0 + \frac{1}{4}$, $x_2 = 0 + \frac{2}{4}$, $x_3 = 0 + \frac{3}{4}$, and so for $i = 1, 2,$ and 3, (6) yields the following system for the corresponding y_1, y_2, and y_3:

$$\begin{aligned} y_2 - 2.25y_1 + y_0 &= 0 \\ y_3 - 2.25y_2 + y_1 &= 0 \\ y_4 - 2.25y_3 + y_2 &= 0. \end{aligned}$$

With the boundary conditions $y_0 = 0$ and $y_4 = 5$, the foregoing system becomes

$$
\begin{aligned}
-2.25y_1 + \quad y_2 \qquad\qquad &= 0 \\
y_1 - 2.25y_2 + \quad y_3 &= 0 \\
y_2 - 2.25y_3 &= -5.
\end{aligned}
$$

Solving the system gives $y_1 = 0.7256$, $y_2 = 1.6327$, and $y_3 = 2.9479$.

NUMERICAL STABILITY

Because of round-off error in the computer, errors are committed as the explicit calculation is carried out. Whether these errors amplify or decay characterizes the stability property of the scheme. A method is said to be **stable** if a small error at one step does not have increasingly great effect on subsequent calculations.

A stability analysis of single-step or self-starting methods such as Euler's and all Runge-Kutta methods has shown that these methods are stable for sufficiently small h.

A numerical method is said to be *convergent* if the numerical solution approaches the exact solution as the step size h goes to 0.

A numerical method is said to be *consistent* if all the approximations (finite difference, finite element, finite volume etc) of the derivatives tend to the exact value as the step size (Δt, Δx etc) tends to zero.

The End