Lecture 8 : Fixed Point Iteration Method, Newton's Method

In the previous two lectures we have seen some applications of the mean value theorem. We now see another application.

In this lecture we discuss the problem of finding approximate solutions of the equation

$$f(x) = 0. \tag{1}$$

In some cases it is possible to find the exact roots of the equation (1), for example, when $f(x)$ is a quadratic or cubic polynomial. Otherwise, in general, one is interested in finding approximate solutions using some (numerical) methods. Here, we will discuss a method called fixed point iteration method and a particular case of this method called Newton's method.

Fixed Point Iteration Method : In this method, we first rewrite the equation (1) in the form

$$x = g(x) \tag{2}$$

in such a way that any solution of the equation (2), which is a fixed point of g, is a solution of equation (1). Then consider the following algorithm.

Algorithm 1: Start from any point x_0 and consider the recursive process

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, ... \tag{3}$$

If f is continuous and (x_n) converges to some l_0 then it is clear that l_0 is a fixed point of g and hence it is a solution of the equation (1). Moreover, x_n (for a large n) can be considered as an approximate solution of the equation (1).

First let us illustrate whatever we said above with an example.

Example 1: We know that there is a solution for the equation $x^3 - 7x + 2 = 0$ in $[0, 1]$. We rewrite the equation in the form $x = \frac{1}{7}(x^3 + 2)$ and define the process $x_{n+1} = \frac{1}{7}(x_n^3 + 2)$. We have already seen in a tutorial class that if $0 \leq x_0 \leq 1$ then (x_n) satisfies the Cauchy criterion and hence it converges to a root of the above equation. We also note that if we start with (for example) $x_0 = 10$ then the recursive process does not converge.

It is clear from the above example that the convergence of the process (3) depends on g and the starting point x_0. Moreover, in general, showing the convergence of the sequence (x_n) obtained from the iterative process is not easy. So we ask the following question.

Question : Under what assumptions on g and x_0, does Algorithm 1 converge ? When does the sequence (x_n) obtained from the iterative process (3) converge ?

The following result is a consequence of the mean value theorem.

Theorem 8.1: *Let $g : [a, b] \to [a, b]$ be a differentiable function such that*

$$\mid g'(x) \mid \leq \alpha < 1 \textit{ for all } x \in [a, b]. \tag{4}$$

Then g has exactly one fixed point l_0 in $[a, b]$ and the sequence (x_n) defined by the process (3), with a starting point $x_0 \in [a, b]$, converges to l_0.

Proof (*): By the intermediate value property g has a fixed point, say l_0. The convergence of (x_n) to l_0 follows from the following inequalities:

$$\mid x_n - l_0 \mid = \mid g(x_{n-1}) - g(l_0) \mid \leq \alpha \mid x_{n-1} - l_0 \mid \leq \alpha^2 \mid x_{n-2} - l_0 \mid .... \leq \alpha^n \mid x_0 - l_0 \mid \to 0.$$

If l_1 is a fixed point then $\mid l_0 - l_1 \mid = \mid g(l_0) - g(l_1) \mid \leq \alpha \mid l_0 - l_1 \mid < \mid l_0 - l_1 \mid$ This implies that $l_0 = l_1$. □

Example 2 : (i) Let us take the problem given in Example 1 where $g(x) = \frac{1}{7}(x^3 + 2)$. Then $g : [0, 1] \to [0, 1]$ and $\mid g'(x) \mid < \frac{3}{7}$ for all $x \in [0, 1]$. Hence by the previous theorem the sequence (x_n) defined by the process $x_{n+1} = \frac{1}{7}(x_n^3 + 2)$ converges to a root of $x^3 - 7x + 2 = 0$.

(ii) Consider $f : [0, 2] \to \mathbb{R}$ defined by $f(x) = (1 + x)^{1/5}$. Observe that f maps [0, 2] onto itself. Moreover $\mid f'(x) \mid \leq \frac{1}{5} < 1$ for $x \in [0, 2]$. By the previous theorem the sequence (x_n) defined by $x_{n+1} = (1 + x_n)^{1/5}$ converges to a root of $x^5 - x - 1 = 0$ in the interval $[0, 2]$.

In practice, it is often difficult to check the condition $f([a, b]) \subseteq [a, b]$ given in the previous theorem. We now present a variant of Theorem 1.

Theorem 8.2: *Let l_0 be a fixed point of $g(x)$. Suppose $g(x)$ is differentiable on $[l_0 - \varepsilon, l_0 + \varepsilon]$ for some $\varepsilon > 0$ and g satisfies the condition $\mid g'(x) \mid \leq \alpha < 1$ for all $x \in [l_0 - \varepsilon, l_0 + \varepsilon]$. Then the sequence (x_n) defined by (3), with a starting point $x_0 \in [l_0 - \varepsilon, l_0 + \varepsilon]$, converges to l_0.*

Proof : By the mean value theorem $g([l_0 - \varepsilon, l_0 + \varepsilon]) \subseteq [l_0 - \varepsilon, l_0 + \varepsilon]$ (Prove !). Therefore, the proof follows from the previous theorem. □

The previous theorem essentially says that if the starting point is sufficiently close to the fixed point then the chance of convergence of the iterative process is high.

Remark : If g is invertible then l_0 is a fixed point of g if and only if l_0 is a fixed point of g^{-1}. In view of this fact, sometimes we can apply the fixed point iteration method for g^{-1} instead of g. For understanding, consider $g(x) = 4x - 12$ then $\mid g'(x) \mid = 4$ for all x. So the fixed point iteration method may not work. However, $g^{-1}(x) = \frac{1}{4}x + 3$ and in this case $\mid (g^{-1})'(x) \mid = \frac{1}{4}$ for all x.

Newton's Method or Newton-Raphson Method :

The following iterative method used for solving the equation $f(x) = 0$ is called Newton's method.

Algorithm 2 : $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad n = 0, 1, 2, ...$

It is understood that here we assume all the necessary conditions so that x_n is well defined. If we take $g(x) = x - \frac{f(x)}{f'(x)}$ then Algorithm 2 is a particular case of Algorithm 1. So we will not get in to the convergence analysis of Algorithm 2. Instead, we will illustrate Algorithm 2 with an example.

Example 3: Suppose $f(x) = x^2 - 2$ and we look for the positive root of $f(x) = 0$. Since $f'(x) = 2x$, the iterative process of Newton's method is $x_{n+1} = \frac{1}{2}(x_n + \frac{2}{x_n}), n = 0, 1, 2, ...$ We have already discussed this sequence in a tutorial class. (Apparently, this process for calculating square roots was used in Mesopotamia before 1500 BC.)

Geometric interpretation of the iterative process of Newton's method : Suppose we have found $(x_n, f(x_n))$. To find x_{n+1}, we approximate the graph $y = f(x)$ near the point $(x_n, f(x_n))$ by the tangent : $y - f(x_n) = f'(x_n)(x - x_n)$. Note that x_{n+1} is the point of intersection of the x-axis and the tangent at x_n.