微積分 MA1001-A 上課筆記（精簡版）

2018.10.30.

Ching-hsiao Arthur Cheng 鄭經斅

3.7 Newton's Method

The Newton method is a numerical method for finding zeros of differentiable functions. Let $f : (a, b) \to \mathbb{R}$ be a differentiable function, and $c \in (a, b)$ is a zero of f. To find an approximated value of c, the Newton method is the following iterative scheme:

1. Make an initial estimate $x_1 \in (a, b)$ that is close to c.

2. Determine a new approximation using the iterative relation:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

3. When $|x_n - x_{n+1}|$ is within the desired accuracy, let x_{n+1} serve as the final approximation.

Example 3.32. To find the square root of a positive number A is equivalent to finding zeros of the function $f(x) = x^2 - A$ in $(0, \infty)$. The Newton method provides the iterative scheme

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - A}{2x_n} = \frac{x_n}{2} + \frac{A}{2x_n}$$

to find approximated value of $\sqrt{A}$.

It can be shown that when $\left|\frac{f(x)f''(x)}{f'(x)^2}\right| < 1$ for all $x \in (a, b)$, then the Newton method produces a convergent sequence which approaches a zero in (a, b).

Chapter 4. Integration

• **The Σ notation**: The sum of n-terms $a_1, a_2, \cdots, a_n$ is written as $\sum_{i=1}^{n} a_i$. In other words,

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n \,.$$

Here i is called the index of summation, a_i is the i-th terms of the sum.

• **Basic properties of sums**: $\sum_{i=1}^{n}(ka_i + b_i) = k\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i$.

Theorem 4.1: Summation Formula

1. $\sum_{i=1}^{n} c = cn$ if c is a constant;
2. $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$;
3. $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$;
4. $\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$.

4.1 The Area under the Graph of a Non-negative Continuous Function

Let $f : [a, b] \to \mathbb{R}$ be a non-negative continuous function, and R be the region enclosed by the graph of the function f, the x-axis and straight lines $x = a$ and $x = b$. We consider computing $\mathcal{A}(\mathrm{R})$, the area of R. Generally speaking, since the graph of $y = f(x)$ is in general not a straight line, the computation of $\mathcal{A}(\mathrm{R})$ is not straight-forward. How do we compute the area $\mathcal{A}(\mathrm{R})$?

Partition $[a, b]$ into n sub-intervals with equal length, and let $\Delta x = \frac{b-a}{n}$, $x_i = a + i\Delta x$. By the Extreme Value Theorem, for each $1 \leqslant i \leqslant n$ f attains its maximum and minimum on $[x_{i-1}, x_i]$; thus for $1 \leqslant i \leqslant n$, there exist $M_i, m_i \in [x_{i-1}, x_i]$ such that

$$f(M_i) = \text{the maximum of } f \text{ on } [x_{i-1}, x_i]$$

and

$$f(m_i) = \text{the minimum of } f \text{ on } [x_{i-1}, x_i].$$

The sum $S(n) \equiv \sum_{i=1}^{n} f(M_i)\Delta x$ is called the upper sum of f for the partition $\{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$, and $s(n) \equiv \sum_{i=1}^{n} f(m_i)\Delta x$ is called the lower sum of f for the partition

$\{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$. By the definition of the upper sum and lower sum, we find that for each $n \in \mathbb{N}$,

$$\sum_{i=1}^{n} f(m_i)\Delta x \leqslant \mathcal{A}(\mathrm{R}) \leqslant \sum_{i=1}^{n} f(M_i)\Delta x\,.$$

If the limits of the both sides exist and are identical as Δx approaches 0 (which is the same as n approaches infinity), by the Squeeze Theorem we can conclude that $\mathcal{A}(\mathrm{R})$ is the same as the limit.

Example 4.2. Let $f(x) = x^2$, and R be the region enclosed by the graph of $y = f(x)$, the x axis, and the straight lines $x = a$ and $x = b$, where we assume that $0 \leqslant a < b$. Then the lower sum is obtained by the "left end-point rule" approximation of $\mathcal{A}(\mathrm{R})$

$$\sum_{i=1}^{n} \Big(a + \frac{(i-1)(b-a)}{n}\Big)^2 \frac{b-a}{n}$$

and the upper sum is obtained by the "right end-point rule" approximation

$$\sum_{i=1}^{n} \Big(a + \frac{i(b-a)}{n}\Big)^2 \frac{b-a}{n}\,.$$

By Theorem 4.1,

$$\begin{aligned}\sum_{i=1}^{n} \Big(a + \frac{i(b-a)}{n}\Big)^2 \frac{b-a}{n} &= \sum_{i=1}^{n} \Big[a^2 + \frac{2a(b-a)i}{n} + \frac{a^2(b-a)^2 i^2}{n^2}\Big]\frac{b-a}{n} \\ &= a^2(b-a) + \frac{a(b-a)^2 n(n+1)}{n^2} + \frac{a^2(b-a)^3}{n^3}\frac{n(n+1)(2n+1)}{6} \\ &= a^2(b-a) + a(b-a)^2\Big(1+\frac{1}{n}\Big) + \frac{a^2(b-a)^3}{6}\Big(1+\frac{1}{n}\Big)\Big(2+\frac{1}{n}\Big)\,.\end{aligned}$$

Letting $n \to \infty$, we find that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \Big(a + \frac{i(b-a)}{n}\Big)^2 \frac{b-a}{n} = a^2(b-a) + a(b-a)^2 + \frac{a^2(b-a)^3}{3} = \frac{b^3-a^3}{3}\,.$$

Similarly,

$$\begin{aligned}\sum_{i=1}^{n} \Big(a + \frac{(i-1)(b-a)}{n}\Big)^2 \frac{b-a}{n} &= \frac{a^2(b-a)}{n} + \sum_{i=1}^{n} \Big(a + \frac{i(b-a)}{n}\Big)^2 \frac{b-a}{n} - \frac{b^2(b-a)}{n} \\ &= a^2(b-a) + \frac{a(b-a)^2 n(n+1)}{n^2} + \frac{a^2(b-a)^3}{n^3}\frac{n(n+1)(2n+1)}{6} + \frac{(a^2-b^2)(b-a)}{n}\,;\end{aligned}$$

thus

$$\lim_{n\to\infty}\sum_{i=1}^{n}\Big(a+\frac{(i-1)(b-a)}{n}\Big)^2\frac{b-a}{n}=\frac{b^3-a^3}{3}\,.$$

Therefore, $\mathcal{A}(\mathrm{R})=\dfrac{b^3-a^3}{3}$.

Remark 4.3. Let R_1 be the region enclosed by $f(x)=x^2$, the x-axis and $x=a$, the R_2 be the region enclosed by $f(x)=x^2$, the x-axis and $x=b$, then intuitively $\mathcal{A}(\mathrm{R})=\mathcal{A}(\mathrm{R}_2)-\mathcal{A}(\mathrm{R}_1)$ and this is true since $\mathcal{A}(\mathrm{R}_1)=\dfrac{a^3}{3}$ and $\mathcal{A}(\mathrm{R}_2)=\dfrac{b^3}{3}$.

If f is not continuous, then f might not attain its extrema on the interval $[x_{i-1},x_i]$. In this case, it might be impossible to form the upper sum or the lower sum for a given partition. On the other hand, the left end-point rule $\sum\limits_{i=1}^{n}f(x_{i-1})\Delta x$ and the right end-point rule $\sum\limits_{i=1}^{n}f(x_i)\Delta x$ of approximating the area are always possible. We can even consider the "mid-point rule" approximation given by

$$\sum_{i=1}^{n}f\Big(\frac{x_{i-1}+x_i}{2}\Big)\Delta x$$

and consider the limit of the expression above as n approaches infinity.

4.2 Riemann Sums and Definite Integrals

In general, in order to find an approximation of $\mathcal{A}(\mathrm{R})$, the interval $[a,b]$ does not have to be divided into sub-intervals with equal length. Assume that $[a,b]$ are divided into n sub-intervals and the end-points of those sub-intervals are ordered as $a=x_0<x_1<x_2<\cdots<x_n=b$, here the collection of end-points $\mathcal{P}=\{x_0,x_1,\cdots,x_n\}$ is called a ***partition*** of $[a,b]$. Then the "left end-point rule" approximation for the partition $\mathcal{P}$ is given by

$$\ell(\mathcal{P})=\sum_{i=1}^{n}f(x_{i-1})(x_i-x_{i-1})$$

and the "right end-point rule" approximation for the partition $\mathcal{P}$ is given by

$$r(\mathcal{P})=\sum_{i=1}^{n}f(x_i)(x_i-x_{i-1})\,.$$