LECTURE 36 THE DEFINITE INTEGRAL

In previous sections, we learned about using a finite amount of rectangles to approximate the area under the graph of a function. At the end of section 5.2, we took this finite number of rectangles to infinity, while shrinking the size of the subintervals. The goal is to obtain the most accurate approximations. With the limiting process, we are now in position to define the area under the graph of a function.

From Finite to Infinite

Let's recall what the finite sum looks like. Consider a function $f(x)$ on the interval $[a,b]$. We consider n subintervals on $[a,b]$, and suppose we use the right endpoint rule. Each subinterval has length $\Delta x = \frac{b-a}{n}$. Our partition then is

$$P = \{a, a+\Delta x, a+2\Delta x, \dots, a+(n-1)\Delta x, a+n\Delta x\}$$

where we note the last term

$$a + n\Delta x = a + n\frac{b-a}{n} = a + b - a = b.$$

Using the right endpoints, we simply ignore $x = a$. Thus, the sum of all these rectangles is

$$\begin{aligned} Area(n) &= \Delta x f(a+\Delta x) + \Delta x f(a+2\Delta x) + \cdots + \Delta x f((n-1)\Delta x) + \Delta x f(a+n\Delta x) \\ &= \Delta x \left(f(a+\Delta x) + f(a+2\Delta x) + \cdots + f((n-1)\Delta x) + f(a+n\Delta x)\right) \\ &= \Delta x \sum_{k=1}^{n} f(a+k\Delta x) \end{aligned}$$

If we know $f(x)$ and the interval $[a,b]$ explicitly, then we know the expression for $Area(n)$. The next step is to take $n \to \infty$, or equivalently shrink the size of the subintervals.

General Partitions and The Limit of a Riemann Sum

Here, most usually, we are doing equal-width subintervals. However, one can do a more general partition using subintervals of various sizes, as long as the sum of the lengths add up to $b - a$. This means, each subinterval does not necessarily have to have the same length, and we label them

$$\Delta x_k = x_k - x_{k-1}, \quad k = 1, 2, \dots, n,$$

that is, the k^{th} rectangle has width Δx_k. Therefore, the partition is also generalized (total number of points didn't change, $n+1$ points),

$$P = \{x_0, x_1, x_2, \dots, x_n\}$$

or in other words, the interval $[a,b]$ now can be written as a union of

$$[x_0, x_1], \quad [x_1, x_2], \dots, [x_{n-1}, x_n].$$

Here, choosing the left or right endpoint is up to you. But we are also allowed to choose any point in the interior of each subinterval, such as the midpoint. We call this arbitrary choice of evaluation point c_k, in the k^{th} interval. Once an evaluation point c_k is chosen, the area of the rectangle then is $f(c_k)\Delta x_k$. We simply add from $k = 1$ up to $k = n$, i.e. the **Riemann sum**

$$Area(n) = \sum_{k=1}^{n} f(c_k)\Delta x_k.$$

Since the subintervals may be of different lengths, the way to interpret $n \to \infty$ then is to consider the maximal length of the subintervals, written as $\|P\|$. If $\|P\| \to 0$, that is, the maximal subinterval length goes to 0, then all the smaller ones go to zero as well.

Definition. If

$$\lim_{\|P\| \to 0} \sum_{k=1}^{n} f(c_k)\Delta x_k = J$$

exists, then we define

$$J = \int_a^b f(x)\, dx.$$

Note that this definition is regardless of the choice of partition, whether equal-width or not, and hence regardless of what rule you use (that decides the location of c_k). This notation reads "the integral from a to b of f of x dee x/with respect to x.

The parallel here is between the symbol $\sum$ and $\int$ (which both relate to the word **sum**).

If the subintervals are of equal width, then the general form above reduces to

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \frac{b-a}{n} \sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right).$$

One last thing to note is the differential dx in association with the independent variable x. The value of the definite integral depends on the function form and the interval, but not on the variable we use to represent the independent variable. In other words,

$$\int_a^b f(t)\,dt = \int_a^b f(u)\,du = \int_a^b f(x)\,dx$$

which all are the same number, if there is one. This variable of integration is called a **dummy variable** (t, u, x in the above three cases).

Theorem. *If a function is continuous on $[a,b]$, or it has at most finitely many jump discontinuities there, then the definite integral $\int_a^b f(x)\,dx$ exists and f is said to be* ***integrable*** *over $[a,b]$.*

Properties of Definite integrals

(1)
$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

(2)
$$\int_a^a f(x)\,dx = 0.$$

(3) The definite integral is linear.

(4)
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

(5) Max-min inequality: if f achieves a max and a min on $[a,b]$, then
$$(\min f)(b-a) \le \int_a^b f(x)\,dx \le (\max f)(b-a).$$
Note that with the current form, this inequality doesn't make too much intuitive sense. However, since $[a,b]$ is not a trivial interval (such that $a=b$), we know $b-a>0$. Therefore, we are allowed to divide $b-a$ from both sides without affecting the two inequalities,
$$\min f \le \frac{1}{b-a}\int_a^b f(x)\,dx \le \max f.$$
Does the middle part look familiar? Yes, that is the average value of the function $f(x)$ on $[a,b]$. Certainly, it should be sandwiched between the extrema.

(6) Domination: if $f(x) \ge g(x)$ on $[a,b]$, then
$$\int_a^b f(x)\,dx \ge \int_a^b g(x)\,dx.$$
A special case is that if $f(x) \ge 0$, then
$$\int_a^b f(x)\,dx \ge 0,$$
that is, non-negative functions have non-negative area under it over $[a,b]$.