Week 1 Lectures

Janko Gravner

MAT 21B
Fall 2023

5.1–5.3. Area and definite integral

Suppose you have a function $y = f(x)$ on the interval $[a, b]$. At first, assume that this is a positive function. We want to approximate the area below its graph.

5.1–5.3. Area and definite integral

A partition P of $[a, b]$ is given by the numbers $a = x_0 < x_1 < \ldots < x_n = b$, so it is a set $P = \{x_0, x_1, \ldots, x_n\}$. Note that these numbers divide $[a, b]$ into n intervals $[x_{k-1}, x_k]$, with length $\Delta x_k = x_k - x_{k-1}$, where $k = 1, \ldots, n$. We call the maximal length of these intervals the *norm* of the partition P:

$$||P|| = \max_k \Delta x_k$$

5.1–5.3. Area and definite integral

Now choose a point $c_k \in [x_{k-1}, x_k]$ for each $k = 1, \ldots, n$. This points are often called *evaluation points*.

Approximate the area under the graph on $[x_{k-1}, x_k]$ by the area of the rectangle with width $x_k - x_{k-1} = \Delta x_k$ and height $f(c_k)$.

The area of the rectangle is

$$f(c_k)(x_k - x_{k-1}) = f(c_k)\Delta x_k$$

5.1–5.3. Area and definite integral

Now add up to approximate the area of $y = f(x)$ on $[a, b]$ by

$$
\begin{aligned}
& f(c_1)\Delta x_1 + f(c_2)\Delta x_2 + \ldots + f(c_n)\Delta x_n \\
& = \sum_{k=1}^{n} f(c_k)\Delta x_k \\
& = \sum_{k=1}^{n} f(c_k)(x_k - x_{k-1})
\end{aligned}
$$

This is called the *Riemann sum*, or sometimes *approximating sum*, for f on $[a, b]$.
We say that we approximate the area by a Riemann sum.

Bernhard Riemann (1826–1866) was a German mathematician.

A detour on the $\sum$ ("Sigma") notation.
If we have numbers $a_1, \ldots, a_n$, we write

$$a_1 + a_2 + \ldots + a_n = \sum_{i=1}^{n} a_i = \sum_{k=1}^{n} a_k$$

Similarly

$$a_0 + a_1 + \ldots + a_n = \sum_{i=0}^{n} a_i$$

Compute:

$$\sum_{i=0}^{3} i, \quad \sum_{i=0}^{3} i^2, \quad \sum_{i=0}^{3} 5, \quad \sum_{i=1}^{3} \frac{1}{i}$$

Here is a famous formula:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

For example,
$1 + 2 + 3 + \cdots + 100 = 5050.$
As the story goes, C. F. Gauss proved this as a schoolchild by regrouping:

Carl Friedrich Gauss (1777–1855) was arguably the greatest mathematician of all time. He was Riemann's adviser.

$$(1 + 100) + (2 + 99) + (3 + 98) + \cdots + (50 + 51) = 50 \cdot 101$$

We prove this by "telescoping," writing

$$i = \frac{1}{2}\left[(i+1)^2 - i^2\right] - \frac{1}{2}$$

and then

$$\begin{aligned}
\sum_{i=1}^{n} i &= \frac{1}{2}\sum_{i=1}^{n}\left[(i+1)^2 - i^2\right] - \frac{1}{2}\sum_{i=1}^{n} 1 \\
&= \frac{1}{2}\left[(2^2 - 1^2) + (3^2 - 2^2) + (4^2 - 3^2) + \ldots + ((n+1)^2 - n^2)\right] \\
&\qquad - \frac{1}{2}\cdot n \\
&= \frac{1}{2}\left[-1^2 + (n+1)^2\right] - \frac{1}{2}\cdot n \\
&= \frac{1}{2}\left[n^2 + 2n - n\right] = \frac{n(n+1)}{2}
\end{aligned}$$

Some other famous formulas:

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{i=1}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2$$

Back to approximating an area by a Riemman sum.

$$\text{area below } y = f(x) \text{ on [a,b]} \approx \sum_{k=1}^{n} f(c_k)\Delta x_k$$
$$= \sum_{k=1}^{n} f(c_k)(x_k - x_{k-1})$$

Example. Take the function $y = x^2$ on $[0, 2]$. Approximate the area below the graph of this function with the Riemann sum with $n = 4$ intervals and evaluation points c_k chosen to be the midpoints of each interval.

So, all $\Delta x_k = 1/2$ and $c_1 = 1/4$, $c_2 = 3/4$, $c_3 = 5/4$ and $c_4 = 7/4$. The Riemann sum is

$$\frac{1}{2}\left[(1/4)^2 + (3/4)^2 + (5/4)^2 + (7/4)^2\right] = 21/8 = 2.625.$$

(The exact area is $8/3 \approx 2.667$.)

We may write the Riemann sum when f is not always positive, in which case it gives negatives contributions from parts of the graph below the x-axis. So the Riemann sum approximates the difference between the area above the x-axis and the area below the x-axis.

5.1–5.3. Area and definite integral

Definition

The *definite* (or *Riemann*) integral of the function $y = f(x)$ from a to b (or over the interval $[a, b]$) is denoted by

$$\int_a^b f(x)\, dx$$

and defined as the limit of Riemann sums:

$$\int_a^b f(x)\, dx = \lim_{\text{norm}\to 0} \sum_{k=1}^{n} f(c_k)\Delta x_k,$$

provided that the limit is independent of the choice of evaluation points c_k.

5.1–5.3. Area and definite integral

Theorem (Existence of Riemann integral)

If $y = f(x)$ is continuous on $[a, b]$, then it has

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

The same is true if $y = f(x)$ is continuous except for finitely many finite jumps.

The endpoints a and b are called *bounds* (or, confusingly, *limits*) of integration. The area below the graph of $y = f(x)$ is *defined through the integral*.

Example. (1) Compute $\int_a^b 3\,dx$. (2) Compute $\int_0^1 x\,dx$.
First get the answer by computing the area, and then by the limit of Riemann sums!

(1) We have $f(x) = 3$ for all x, so any Riemann sum:

$$\sum_{k=1}^{n} f(c_k)\Delta x_k = 3\sum_{k=1}^{n} \Delta x_k = 3(b-a)$$

and so

$$\int_a^b 3\,dx = \lim_{\text{norm}\to 0} \sum_{k=1}^{n} f(c_k)\Delta x_k = 3(b-a).$$

(2) Now $f(x) = x$. We choose:

- a partition of n intervals of equal length so that $\Delta x_k = 1/n$ for all k;
- the evaluation points c_k to be the *right* endpoints, so that $c_1 = 1/n$, $c_2 = 2/n, \ldots, c_n = 1$, that is $c_k = k/n$.

So,

$$\begin{aligned}
\int_0^1 x\,dx &= \lim_{\text{norm}\to 0} \sum_{k=1}^{n} f(c_k)\Delta x_k \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \frac{k}{n} \cdot \frac{1}{n} \\
&= \lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} k = \lim_{n\to\infty} \frac{1}{n^2} \cdot \frac{n(n+1)}{2} = \frac{1}{2}
\end{aligned}$$

For a general interval, we get in a similar way

$$\int_a^b x\,dx = \frac{b^2}{2} - \frac{a^2}{2}$$

and

$$\int_a^b x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3}.$$

We will soon have a much better methods to compute these that by computing limits of Riemann sums.

Note that the name of independent variable makes no difference:

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(\square)\,d\square.$$

We *define*

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

and so

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

Then we have *interval additivity*, valid for any a, b, c:

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

We also have *linearity*: for any constant k

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

and for any functions f and g such that the integral exists

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

5.1–5.3. Area and definite integral

Next is *monotonicity*. Assume $a < b$. If $f(x) \le g(x)$ for all $x \in [a, b]$, then

$$\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx.$$

In particular, if M is a constant and $f(x) \le M$ for for all $x \in [a, b]$, then

$$\int_a^b f(x)\,dx \le M(b-a)$$

If m is a constant and $f(x) \ge m$ for for all $x \in [a, b]$, then

$$\int_a^b f(x)\,dx \ge m(b-a)$$

The *average* av(f) of $y = f(x)$ on $[a, b]$ is a *constant* such that its integral equals $\int_a^b f(x)\,dx$. So,

$$(b - a) \cdot \text{av}(f) = \int_a^b f(x)\,dx,$$

and so

$$\text{av}(f) = \frac{1}{b - a} \int_a^b f(x)\,dx$$

Note that, when $\Delta x_k = (b - a)/n$ (all intervals are of equal length),

$$\text{av}(f) = \frac{1}{b - a} \int_a^b f(x)\,dx \approx \frac{1}{b - a} \sum_{i=1}^{n} f(c_i) \frac{b - a}{n} = \frac{1}{n} \sum_{i=1}^{n} f(c_i).$$

Example. Compute the average of $f(x) = x^2 + 3x$ on $[0, 2]$.

Example. Compute the average of $f(x) = x^2 + 3x$ on $[0, 2]$.
We apply the formula, for which we need to compute

$$\int_0^2 (x^2 + 3x)\, dx = \int_0^2 x^2\, dx + 3 \int_0^2 x\, dx$$
$$= \frac{2^3 - 0}{3} + 3 \cdot \frac{2^2 - 0}{2} = \frac{26}{3}$$

and so

$$\operatorname{av}(f) = \frac{1}{2} \cdot \frac{26}{3} = \frac{13}{3}.$$

Other than area, the definite integral has many other applications. We can use the definite integral to compute the change in distance if an object moves on a straight line with variable velocity $v(t)$ between two times $t = a$ and $t = b$.

During a short time interval $[t_{i-1}, t_i]$, we may assume that the velocity is nearly constant, equal to $v(c_i)$ for some $c_i \in [t_{i-1}, t_i]$. So, during this short interval, the change in distance is $\approx v(c_i)(t_i - t_{i-1}) = v(c_i)\Delta t_i$ and the total change in distance is approximately

$$\sum_{i=1}^{n} v(c_i)\Delta t_i$$

and thus equals

$$\int_a^b v(t)\, dt.$$

Similarly, if $a(t)$, $t \geq 0$ is the acceleration of an object started at rest $v(0) = 0$, then its velocity at time b is

$$v(b) = \int_0^b a(t)\,dt.$$

8.7. Estimation of the definite integral

We will learn how to compute the definite integral exactly in many cases. However, very often the exact computation is impossible. So we must rely on approximations. We will discuss three of them.

1. **Rectangular method.** This simply approximates the integral by a Riemann sum:

$$\int_a^b f(x)\,dx \approx \sum_{k=1}^{n} f(c_k)\Delta x_k.$$

8.7. Estimation of the definite integral

1. **Trapezoidal method.** Approximate f by a linear function connecting the endpoints of its graph on each $[x_{k-1}, x_k]$. So, on each $[x_{k-1}, x_k]$, the area is approximated by

$$\frac{1}{2}\left(f(x_{k-1}) + f(x_k)\right) \cdot \Delta x_k$$

and so we get

$$\int_a^b f(x)\, dx \approx \sum_{k=1}^{n} \frac{1}{2}\left(f(x_{k-1}) + f(x_k)\right) \cdot \Delta x_k.$$

8.7. Estimation of the definite integral

$$\int_a^b f(x)\,dx \approx \sum_{k=1}^{n} \frac{1}{2}\left(f(x_{k-1}) + f(x_k)\right) \cdot \Delta x_k.$$

If $\Delta x_k = \Delta x = (b-a)/n$ for all k, the approximation is:

$$\begin{aligned}\Delta x \Bigg[&\frac{1}{2} f(x_0) + \frac{1}{2} f(x_1)\\ &+ \frac{1}{2} f(x_1) + \frac{1}{2} f(x_2)\\ &+ \frac{1}{2} f(x_2) + \frac{1}{2} f(x_3)\\ &\dots\\ &+ \frac{1}{2} f(x_{n-1}) + \frac{1}{2} f(x_n)\Bigg],\ \text{so}\end{aligned}$$

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{2}\left[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)\right]$$

8.7. Estimation of the definite integral

3. Simpson's method. We assume that $\Delta x_k = \Delta x = (b-a)/n$ for all k, and that n is *even*. We approximate f on two successive intervals by quadratic parabola, and get

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{3}\,[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)]$$

For all three methods there are error estimates (see the book), which will not be required.

Thomas Simpson (1710–1761) was a British mathematician.

Example. Estimate

$$\int_0^1 \frac{x}{1+x^2}\,dx$$

using 4 intervals of equal length and each of the three methods:

(1) rectangular method with evaluation points the right endpoints;
(2) trapezoidal method;
(3) Simpson's method.

In the cases (1) and (2), also determine whether the approximation is smaller or larger than the true value.

We have

$$f(x) = \frac{x}{1+x^2}$$

We compute

$$f'(x) = \frac{1-x^2}{(1+x^2)^2}, \quad f''(x) = \frac{-2x(3-x^2)}{(1+x^2)^3}$$

and we conclude that the function is increasing and concave down on $[0, 1]$.

We also need these values:

x	$f(x)$
0	0
1/4	4/17
1/2	2/5
3/4	12/25
1	1/2

(1) Estimate

$$\int_0^1 \frac{x}{1+x^2}\, dx$$

using 4 intervals of equal length and rectangular method with evaluation points the right endpoints.

The answer is

$$\frac{1}{4}f(1/4) + \frac{1}{4}f(1/2) + \frac{1}{4}f(3/4) + \frac{1}{4}f(1) \approx 0.4038$$

and is *larger* than the true value as f is increasing.

(2) Estimate

$$\int_0^1 \frac{x}{1+x^2}\,dx$$

using 4 intervals of equal length and trapezoidal method. The answer is

$$\frac{1}{8}\left[f(0)+2f(1/4)+2f(1/2)+2f(3/4)+f(1)\right]\approx 0.3413$$

and is *smaller* than the true value as f is concave down.

(3) Estimate

$$\int_0^1 \frac{x}{1+x^2}\,dx$$

using 4 intervals of equal length and Simpson's method.
The answer is

$$\frac{1}{12}\left[f(0)+4f(1/4)+2f(1/2)+4f(3/4)+f(1)\right]\approx 0.3468$$

The correct value is $\frac{1}{2}\ln 2 \approx 0.3466$.

A follow-up question. Suppose you estimate

$$\int_0^1 \frac{x}{1+x^2}\,dx$$

using 4 intervals of equal length and rectangular method, but now evaluation points are *midpoints*. Is the approximation larger or smaller?

On each subinterval, the area of the rectangle is the same as the area of the trapezoid formed by the tangent at the evaluation point. By concavity, the tangent is above the graph, so the area of the trapezoid is larger.

Answer: the approximation is *larger*.