Analysis Notes
Grinshpan

Riemann Sums

Let $f(x)$ be a bounded function on a bounded interval $[a, b]$. Let

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$$

be points partitioning $[a, b]$ into n subintervals $[x_{k-1}, x_k]$, $k = 1, \ldots, n$.

In each subinterval $[x_{k-1}, x_k]$ choose a point x_k^*, $x_{k-1} \le x_k^* \le x_k$. The sum

$$S = \sum_{k=1}^{n} (x_k - x_{k-1}) f(x_k^*)$$

is called the Riemann sum of $f(x)$ on $[a, b]$ corresponding to the partition $\{x_k, x_k^*\}$.

If $f(x) > 0$, S represents the sum of areas of rectangles with base $[x_{k-1}, x_k]$ and height $f(x_k^*)$. The union of these rectangles approximates the region between the graph of $f(x)$ and $[a, b]$.

If x_k are equally spaced, then the mesh is $h = \frac{1}{n}(b - a)$, $x_k = a + kh$, and

$$S = \tfrac{b-a}{n} \sum_{k=1}^{n} f(x_k^*).$$

In this case S is $(b - a)$ times the average of $f(x_1^*), \ldots, f(x_n^*)$.

Left, Right, Middle

The most common choices for x_k^* are

$$x_k^* = x_{k-1}, \quad x_k^* = x_k, \quad x_k^* = \tfrac{1}{2}(x_{k-1} + x_k).$$

In the equispaced case, the corresponding Riemann sums are

$$L = \tfrac{b-a}{n} \sum_{k=0}^{n-1} f(x_k) \qquad \text{(left endpoint sum)}$$

$$R = \tfrac{b-a}{n} \sum_{k=1}^{n} f(x_k) \qquad \text{(right endpoint sum)}$$

$$M = \tfrac{b-a}{n} \sum_{k=0}^{n-1} f(\tfrac{x_{k-1}+x_k}{2}) \qquad \text{(midpoint sum).}$$

Upper and lower

Let M_k be the supremum and m_k be the infimum of $f(x)$ on $[x_{k-1}, x_k]$, $k = 1, \ldots, n$. Consider the corresponding sums

$$S_{\max} = \sum_{k=1}^{n} (x_k - x_{k-1}) M_k \qquad \text{(upper sum)}$$
$$S_{\min} = \sum_{k=1}^{n} (x_k - x_{k-1}) m_k \qquad \text{(lower sum).}$$

Evidently, for any choice of points x_k^*, we have

$$S_{\min} \leq S \leq S_{\max}$$

and the area under the graph is also captured between $S_{\min}$ and $S_{\max}$. Note that $S_{\max} - S_{\min}$ is an upper bound for the area between the graph and rectangles.

Limit

Consider a sequence of partitions $\{x_k^{(j)}, x_k^{*(j)}\}$ whose maximum interval length $\max_k(x_k^{(j)} - x_{k-1}^{(j)})$ approaches zero. If the corresponding Riemann sums

$$S_j = \sum_{k=1}^{n_j} (x_k^{(j)} - x_{k-1}^{(j)}) f(x_k^{*(j)})$$

converge to the same limit for all choices of $x_k^{*(j)}$, we say that $f(x)$ is Riemann integrable on $[a, b]$ and that the limit is

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

If the limit exists, it then exists for any sequence of partitions whose maximum interval length approaches zero and

$$\sup_{\{x_k\}} S_{\min} = \int_a^b f(x)dx = \inf_{\{x_k\}} S_{\max}.$$

Crude error estimate

We will derive an error bound for the Riemann sum approximations and show that it cannot be improved for the left/right endpoint sums.

Assume for simplicity that $f(x)$ is continuously differentiable on $[a, b]$. The error of k-th rectangle is

$$e_k = \int_{x_{k-1}}^{x_k} f(x)dx - (x_k - x_{k-1}) f(x_k^*) = \int_{x_{k-1}}^{x_k} (f(x) - f(x_k^*))dx.$$

Hence

$$\begin{aligned}|e_k| &\le \int_{x_{k-1}}^{x_k} |f(x) - f(x_k^*)|\, dx \\ &\le \max_{[a,b]} |f'(x)| \int_{x_{k-1}}^{x_k} |x - x_k^*| dx \\ &= \frac{1}{2} \max_{[a,b]} |f'(x)| \left((x_k^* - x_{k-1})^2 + (x_k - x_k^*)^2\right) \\ &\le \frac{1}{2} \max_{[a,b]} |f'(x)| \, (x_k - x_{k-1})^2.\end{aligned}$$

Adding these error bounds over k, we see that the total error is bounded by

$$\frac{1}{2} \max_{[a,b]} |f'(x)| \sum_{k=1}^{n} (x_k - x_{k-1})^2 \le \frac{1}{2} \max_{[a,b]} |f'(x)| \max_k (x_k - x_{k-1})(b - a).$$

Thus

$$\left| \int_a^b f(x) dx - S \right| \le \frac{b-a}{2} \max_{[a,b]} |f'(x)| \, \delta_{\max},$$

where $\delta_{\max} = \max_k (x_k - x_{k-1})$. It is easy to check that this estimate, although crude, cannot be improved for endpoint Riemann sums.

EXAMPLE. Consider $f(x) = x$ on $[0, 1]$. We have

$$\int_0^1 x dx = \frac{1}{2}$$

$$L = \frac{1}{n} \sum_{k=0}^{n-1} \frac{k}{n} = \frac{1}{2} \left(1 - \frac{1}{n}\right)$$

$$R = \frac{1}{n} \sum_{k=1}^{n} \frac{k}{n} = \frac{1}{2} \left(1 + \frac{1}{n}\right),$$

so the absolute error in both cases is exactly

$$\frac{1}{2n} = \frac{b-a}{2} \max_{[a,b]} |f'(x)| \, \delta_{\max} \; .$$

Midpoint sum error estimate

The midpoint sum provides a better approximation to the integral.

Assume that $f(x)$ is twice continuously differentiable on $[a, b]$. Integrating Taylor's approximation

$$f(x) - f(x_k^*) = f'(x_k^*)(x - x_k^*) + \tfrac{1}{2} f''(c_x)(x - x_k^*)^2$$

over $[x_{k-1}, x_k]$ we find that the error of k-th rectangle is

$$e_k = f'(x_k^*) \int_{x_{k-1}}^{x_k} (x - x_k^*) dx + \tfrac{1}{2} \int_{x_{k-1}}^{x_k} f''(c_x)(x - x_k^*)^2 dx.$$

If $x_k^* = \frac{1}{2}(x_{k-1} + x_k)$ then the first integral on the right is 0 and

$$|e_k| \le \tfrac{1}{2} \max_{[a,b]} |f''(x)| \int_{x_{k-1}}^{x_k} (x - x_k^*)^2 dx = \frac{1}{24} \max_{[a,b]} |f''(x)| (x_k - x_{k-1})^3.$$

Summing over k, we obtain the error bound

$$\frac{1}{24}\max_{[a,b]}|f''(x)|\sum_{k=1}^{n}(x_k-x_{k-1})^3 \le \frac{1}{24}\max_{[a,b]}|f''(x)|\max_{k}(x_k-x_{k-1})^2(b-a).$$

Thus

$$\left|\int_a^b f(x)dx - M\right| \le \frac{b-a}{24}\ \max_{[a,b]}|f''(x)|\ \delta_{\max}^2.$$

In general, this bound cannot be improved.

EXAMPLE. For $f(x)=x^2$ on $[0,1]$ we have

$$\int_0^1 x^2dx = \frac{1}{3}$$

$$M = \frac{1}{n}\sum_{k=1}^{n}\frac{(k-\frac{1}{2})^2}{n^2} = \frac{1}{3}-\frac{1}{12n^2}.$$

So the absolute error is exactly

$$\frac{1}{12n^2} = \frac{b-a}{24}\ \max_{[a,b]}|f''(x)|\ \delta_{\max}^2\ .$$