Lecture 30: Numerical integration

We briefly look at some numerical techniques for computing integrals. There are variations of basic Riemann sums but speed up the computation.

Riemann sum with nonuniform spacing

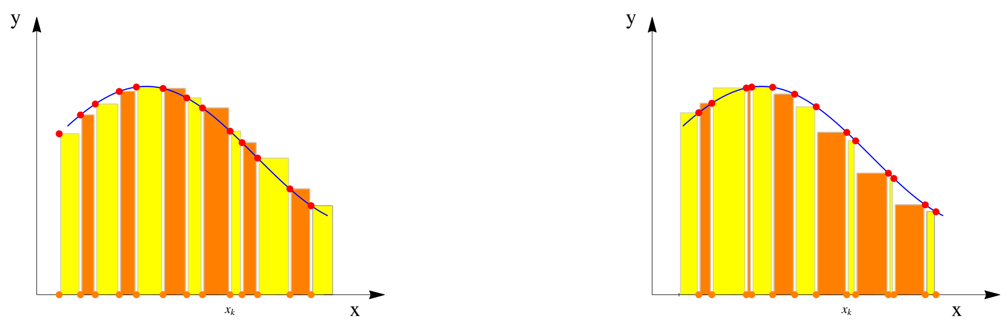
A more general Riemann sum is obtained by choosing n points $\{x_j\}$ in $[a,b]$ and then to look at

$$S_n = \sum f(y_j)(x_{j+1} - x_j) = \sum_{y_j} f(y_j)\Delta x_j ,$$

where y_j is in (x_j, x_{j+1}).

This generalization allows to use a small mesh size where the function fluctuates a lot. The function $f(x) = \sin(1/(x^2 + 0.1))$ for example fluctuates near the origin more and would need more division points there.

The sum $\sum f(x_j)\Delta x_j$ is called the **left Riemann sum**, the sum $\sum f(x_{j+1})\Delta x_j$ the **right Riemann sum**.



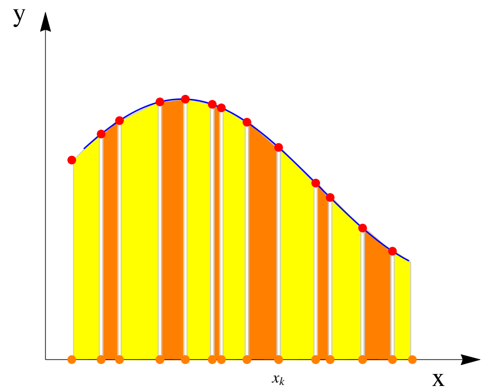
If $x_0 = a, x_n = b$ and $\max_j \Delta x_j \to 0$ for $n \to \infty$ then S_n converges to $\int_a^b f(x)\, dx$.

1 If $x_j - x_k = 1/n$ and $z_j = x_j$, then we have the Riemann sum as we defined it initially.

2 You numerically integrate $\sin(x)$ on $[0, \pi/2]$ with a Riemann sum. What is better, the left Riemann sum or the right Riemann sum? Look also at the interval $[\pi/2, \pi]$? **Solution:** you see that in the first case, the left Riemann sum is smaller than the actual integral. In the second case, the left Riemann sum is larger than the actual integral.

Trapezoid rule

The average between the left and right hand Riemann sum is called the **Trapezoid rule**. Geometrically, it sums up areas of trapezoids instead of rectangles.



The Trapezoid rule does not change things much in the case of equal spacing $x_k = a + (b-a)k/n$.

$$\frac{1}{2n}[f(x_0) + f(x_n)] + \frac{1}{n}\sum_{k=1}^{n-1} f(x_k) \; .$$

Simpson rule

The **Simpson rule** computes the sum

$$S_n = \frac{1}{6n}\sum_{k=1}^{n}[f(x_k) + 4f(y_k) + f(x_{k+1})] \; ,$$

where y_k are the midpoints between x_k and x_{k+1}.

The Simpson rule is good because it is exact for quadratic functions: for $f(x) = ax^2 + bx + c$, the formula

$$\frac{1}{v-u}\int_u^v f(x)\, dx = [f(u) + 4f((u+v)/2) + f(v)]/6$$

holds exactly. To prove it just run the following two lines in Mathematica: (== means "is equal")

```
f[x_] := a x^2 + b x + c;
Simplify[(f[u]+f[v]+4f[(u+v)/2])/6==Integrate[f[x],{x,u,v}]/(v-u)]
```

This actually will imply (as you will see in Math 1b) that the numerical integration for functions which are 4 times differentiable gives numerical results which are n^{-4} close to the actual integral. For 100 division points, this can give accuracy to 10^{-8} already.

There are other variants which are a bit better but need more function values. If x_k, y_k, z_k, x_{k+1} are equally spaced, then

The **Simpson 3/8 rule** computes

$$\frac{1}{8n}\sum_{k=1}^{n}[f(x_k) + 3f(y_k) + 3f(z_k) + f(x_{k+1})] \; .$$

This formula is again exact for quadratic functions: for $f(x) = ax^2 + bx + c$, the formula

$$\frac{1}{v-u}\int_u^v f(x)\, dx = [f(u) + 3f((2u+v)/3) + 3f((u+2v)/3) + f(v)]/6$$

holds. If you are interested, run the two Mathematica lines:

```
f[x_] := a x^2 + b x + c; L=Integrate[f[x],{x,u,v}]/(v-u);
Simplify[(f[u]+f[v]+3f[(2u+v)/3]+3f[(u+2v)/3])/8==L]
```

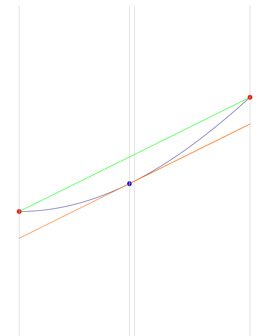
This 3/8 method can be slightly better than the first Simpson rule.

Mean value method

The mean value theorem shows that for $x_k = k/n$, there are points $y_k \in [x_k, x_{k+1}]$ such that $f(y_j) = F'(y_j) = f(x_{j+1}) - f(x_j)$ and so

$$\frac{1}{n}\sum_{k=1}^{n} f(y_k) = F(x_n) - F(x_0) \ .$$

This is a version of the fundamental theorem of calculus which is exact in the sense that for every n, this is a correct formula. Lets call y_k the **Rolle points**.



The Rolle point is close to the interval midpoint.

For any partition x_k on $[a, b]$ with $x_0 = a, x_n = b$, there is a choice of Rolle points $y_k \in [x_k, x_{k+1}]$ such that the Riemann sum $\sum_k f(y_k)\Delta(x)_k$ is equal to $\int_a^b f(x)\, dx$.

For linear functions the Rolle points are the midpoints. In general, the deviation $g(t)$ from the midpoint is small if the interval is $[x_0 - t, x_0 + t]$. One can estimate $g(t)$ to be of the order $t^2 \frac{f'''(x_0)}{6f''(x_0)}$. We could modify the trapezoid rule and replace the line through the points by a Taylor polynomial. The Rolle point method is useful for functions which can have poles.

Monte Carlo Method

A powerful integration method is to chose n random points x_k in $[a, b]$ and look at the sum divided by n. Because it uses randomness, it is called **Monte Carlo method**.

The **Monte Carlo** integral is the limit S_n to infinity

$$S_n = \frac{1}{n}\sum_{k=1}^{n} f(x_k) \ ,$$

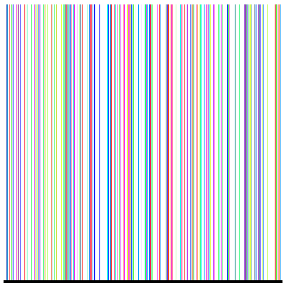
where x_k are n random values in $[a, b]$.

The law of large numbers in probability shows that the **Monte Carlo integral** is equivalent to the **Lebesgue integral** which is more powerful than the Riemann integral. Monte Carlo integration is interesting especially if the function is complicated.

3 Lets look at the **salt and pepper** function

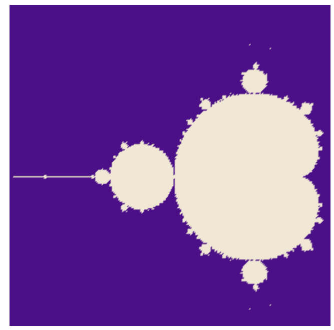
$$f(x) = \begin{cases} 1 & x \text{ rational} \\ 0 & x \text{ irrational} \end{cases}$$

The Riemann integral with equal spacing k/n is equal to 1 for every n. But this is only because we have evaluated the function at rational points, where it is 1.
The Monte Carlo integral gives zero because if we chose a random number in $[0, 1]$ we hit an irrational number with probability 1.




The Salt and Pepper function and the Boston Salt and Pepper bridge (Anne Heywood).

4 The following two lines evaluate the **area of the Mandelbrot fractal** using Monte Carlo integration. The function F is equal to 1, if the parameter value c of the quadratic map $z \to z^2 + c$ is in the Mandelbrot set and 0 else. It shoots $100'000$ random points and counts what fraction of the square of area 9 is covered by the set. Numerical experiments give values close to the actual value around 1.51.... One could use more points to get more accurate estimates.



```
F[c_]:=Block[{z=c,u=1},Do[z=N[z^2+c];If[Abs[z]>3,u=0;z=3],{99}];u];
M=10^5; Sum[F[-2.5+3 Random[]+I(-1.5+3 Random[])],{M}]*(9.0/M)
```

Homework

1 Use a computer to generate 20 random numbers x_k in [0,1]. Sum up the square x_k^2 of these numbers and divide by 20. Compare your result with $\int_0^1 x^2 \, dx$. **Remark.** If using a program, increase the value of n as large as you can. Here is a Mathematica code:

```
n=20; Sum[ Random[]^2,{n}]/n
```

Here is an implementation in Perl. Its still possible to cram the code into one line:

```
#!/usr/bin/perl
$n=20;$s=0;for($i=0;$i<$n;$i++){$f=rand();$s+=$f*$f;} print $s/$n;
```

2 a) Use the Simpson rule to compute $\int_0^\pi \sin(x) \, dx$ using $n = 2$ intervals $[0, \pi/2]$ and $[\pi/2, \pi]$. On each of these intervals $[a, b]$ compute the Simpson sum $[f(a) + 4f((a + b)/2) + f(b)]/6$ with $f(x) = \sin(x)$. Compare with the actual integral.
b) Now use the 3/8 Simpson rule to estimate $\int_0^\pi f(x) \, dx$ using $n = 1$ intervals $[0, \pi]$. Again compare with the actual integral.

Instead of adding more numerical methods exercices, we want to practice a bit more integration. The challenge in the following problems is to find out which integration method is bets suited. This is good preparation for the final, where we will not reveal which integration method is the best.

3 Integrate $\tan(x)/\cos(x)$ from 0 to $\pi/6$.

4 Find the antiderivative of $\sqrt{x} \log x$.

5 Find the antiderivative of $x/\sin(x)^2$.