Techniques of Integration: II

Math 10A

October 12, 2017

By Parts

We've calculated a few integrals by using the formula

$$\int u\,dv = uv - \int v\,du.$$

Summary: we write the integrand in an antiderivative problem as the product $f(x)g(x)$ where we can figure out an antiderivative $G(x)$ of the factor $g(x)$. We put $u = f(x)$, $v = G(x)$, $dv = g(x)\,dx$. Then

$$\int f(x)g(x)\,dx = f(x)G(x) - \int f'(x)G(x)\,dx.$$

By expressing $\int f(x)g(x)\,dx$ in terms of $\int f'(x)G(x)\,dx$, we can replace f by its derivative and g by an antiderivative G of g.

We will look at a few more examples before moving on to numerical integration.

We already saw when discussing such integrals as $\int x^2 e^{-x}\,dx$ and $\int x^2 e^{-x}\,dx$ that repeated applications of integration by parts can pay off. Next is another problem with this feature.

Find $\int e^x \sin x \, dx$.

The integrand is a product, so we can try setting $u = e^x$, $dv = \sin x \, dx$, $v = -\cos x$. We get

$$\int e^x \sin x \, dx = -e^x \cos x + \int e^x \cos x \, dx.$$

The right-hand integral is a sibling of the initial integral, with sin replaced by cos. Have we made any progress?

In $\int e^x \cos x \, dx$, put $u = e^x$, $dv = \cos x \, dx$, $v = \sin x$. Then

$$\int e^x \cos x \, dx = e^x \sin x - \int \sin x \, e^x \, dx.$$

Then

$$\begin{aligned}\int e^x \sin x \, dx &= -e^x \cos x + \int e^x \cos x \, dx \\ &= -e^x \cos x + e^x \sin x - \int \sin x \, e^x \, dx\end{aligned}$$

This gives

$$2 \int e^x \sin x \, dx = -e^x \cos x + e^x \sin x,$$

$$\int e^x \sin x \, dx = \frac{-e^x \cos x + e^x \sin x}{2}.$$

An example gotten by googling: Find $\int x^5 \sqrt{x^3+1}\, dx$.

Yuck!

For lack of a better idea, we can try to put $dv = x^2\sqrt{x^3+1}\, dx$. This sounds promising because we can integrate $x^2\sqrt{x^3+1}$: an antiderivative is $\frac{2}{9}(x^3+1)^{\frac{3}{2}}$. (See doc camera.)

We need to take $u = x^3$; then

$$\int x^5\sqrt{x^3+1}\, dx = \frac{2}{9}x^3(x^3+1)^{\frac{3}{2}} - \int \frac{2}{9}(x^3+1)^{\frac{3}{2}} 3x^2\, dx.$$

An example gotten by googling: Find $\int x^5 \sqrt{x^3+1}\, dx$.

Yuck!

For lack of a better idea, we can try to put $dv = x^2\sqrt{x^3+1}\, dx$. This sounds promising because we can integrate $x^2\sqrt{x^3+1}$: an antiderivative is $\frac{2}{9}(x^3+1)^{\frac{3}{2}}$. (See doc camera.)

We need to take $u = x^3$; then

$$\int x^5 \sqrt{x^3+1}\, dx = \frac{2}{9}x^3(x^3+1)^{\frac{3}{2}} - \int \frac{2}{9}(x^3+1)^{\frac{3}{2}} 3x^2\, dx.$$

This is good because

$$\int (x^3+1)^{\frac{3}{2}} 3x^2\,dx = \frac{2}{5}(x^3+1)^{\frac{5}{2}} + C.$$

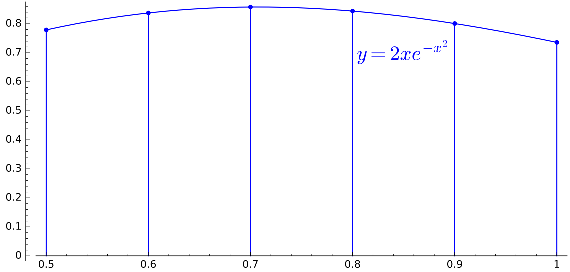
Putting everything together gives us

$$\begin{aligned}\int x^5\sqrt{x^3+1}\,dx &= \frac{2}{9}x^3(x^3+1)^{\frac{3}{2}} - \int \frac{2}{9}(x^3+1)^{\frac{3}{2}} 3x^2\,dx \\ &= \frac{2}{9}x^3(x^3+1)^{\frac{3}{2}} - \frac{2}{9}\cdot\frac{2}{5}(x^3+1)^{\frac{5}{2}} + C.\end{aligned}$$

This was Example 7 in Paul's tutorial.

Numerical integration

Our aim is to find a numerical approximation to a definite integral like $\int_{1/2}^{1} 2xe^{-x^2}\,dx$. The definite integral in question represents the area under the curve $y = 2xe^{-x^2}$:



We will introduce a number of approximation methods and use them to compute approximate values of the integral $\int_{1/2}^{1} 2xe^{-x^2}\,dx$. I picked this particular integral because we happen to know an antiderivative of $2xe^{-x^2}$, namely $-e^{-x^2}$. Hence it is possible to compute the area exactly as $e^{-1/4} - e^{-1}$; numerically, this quantity works out to be $\approx 0.410921341899963$.

Since we know the area exactly in this example, why do we want to approximate it? We are introducing methods that can be used for other integrals that can't be calculated exactly:

1. We might want to calculate areas like $\int_{1/2}^{1} e^{-x^2}\,dx$, for which we don't know an antiderivative.
2. We might be scientists and have access only to data associated with an unknown function—we'd never know an explicit function.

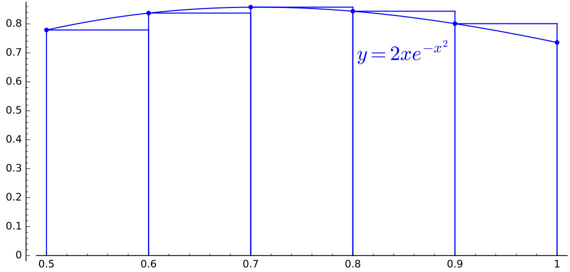
Riemann sums

We can use Riemann sums with left or right endpoints to approximate the definite integral as a sum of areas of rectangles. Note that, in the picture, I've taken the interval $[1/2, 1]$ and divided it up into five equal segments of length 0.1. With left endpoints we get the approximation

$$0.1 \times [(f(0.5) + f(0.6) + \cdots + f(0.9)] \approx 0.4118$$

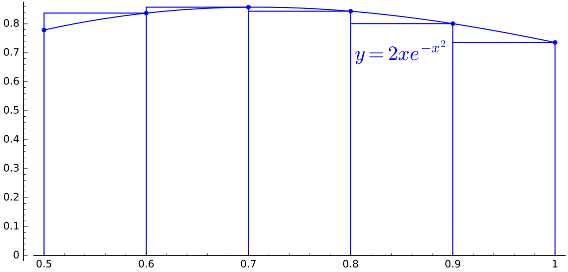
for the integral; I've put $f(x) = 2xe^{-x^2}$ for convenience.

The approximate value 0.4118 compares well with the actual value $0.41092\cdots$ because two of the approximating rectangles under-estimate the area while the three others give over-estimates.



We are slightly lucking out.

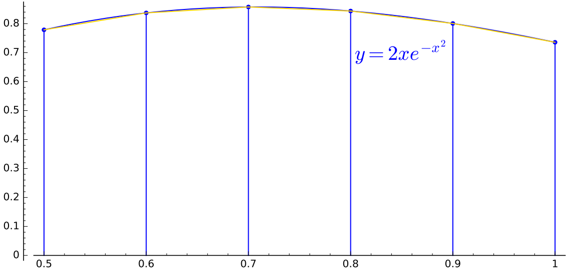
In this example, we could use right endpoints instead of left endpoints. The approximation to the area that we obtain is $0.4075\ldots$; again, the actual value is $0.41092\ldots$.



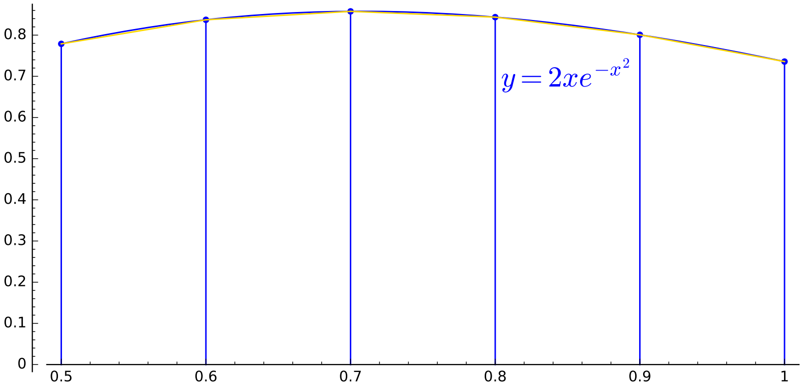
This time we have two over-estimates and three under-estimates.

Trapezoidal rule

The trapezoidal rule amounts to approximating thin strips of area under the curve by areas of trapezoids running between the left and right endpoints of each strip. The resulting *trapezoidal approximation* is just the average of the left- and right-endpoint approximations.

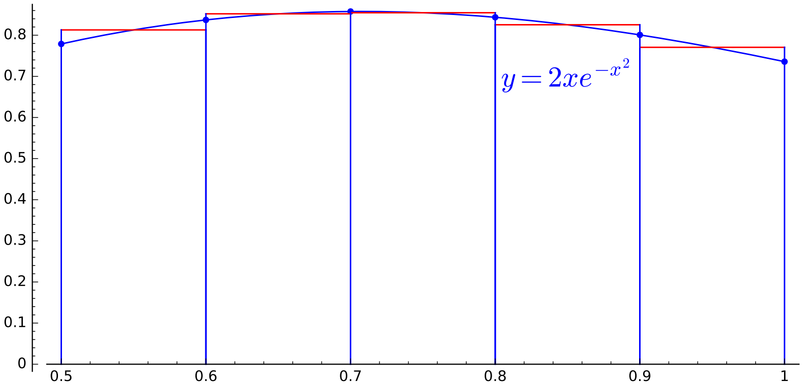


Here's the diagram again, slightly blown up. The tops of the trapezoids have been colored gold; otherwise, we'd never be able to pick them out. The trapezoidal approximation looks really good, doesn't it?



Numerically, we get $0.409658\ldots$.

Before discussing theoretical estimates for the accuracy of the approximation methods we've encountered so far, I'll draw a picture that illustrates the *midpoint rule*. For this rule, you use middle points instead of left or right endpoints.



The approximation here is 0.4115533986 . . . .

Numerical summary, so far

Actual value 0.410921341899963.

Rule	Left	Right	Midpoint	Trap
Estimate	0.4118	0.4075	0.41155	0.40966
Abs. Error	0.0008888	0.0034	0.000632	0.00126

In this particular case, the midpoint rule came out the winner. I would have bet on the trapezoids. I think that the midpoint rule, in general, is kind of a turkey.

Here's the reason for my bet: the left, right and midpoint rules approximate the function by a *constant* over each small interval. The trapezoidal rule approximates the function by a line ("$y = mx + b$") over each little interval. As we saw from the picture, the tops of the trapezoids hug the function very well.

Simpson's rule

Simpson's rule approximates the function by parabolas—and thus does even better.

For a parabola, you need three points (just as a line is determined by two points), so the picture looks different.

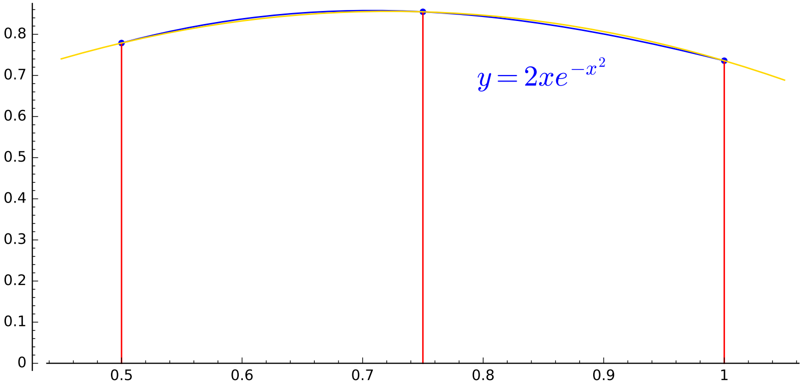
Before showing you the picture, I want to recommend a visit to the Wikipedia page for Simpson's rule (linked above). That page gives you the explicit formula for a quadratic $Ax^2 + Bx + C$ that passes through three specified points on the plane (with all different x-coordinates).

Simpson's rule

Simpson's rule approximates the function by parabolas—and thus does even better.

For a parabola, you need three points (just as a line is determined by two points), so the picture looks different.

Before showing you the picture, I want to recommend a visit to the Wikipedia page for Simpson's rule (linked above). That page gives you the explicit formula for a quadratic $Ax^2 + Bx + C$ that passes through three specified points on the plane (with all different x-coordinates).



For simplicity, I've divided the interval from 0.5 to 1 into only two segments, each of length 0.25. The gold parabola passes through the three relevant points of $y = f(x)$. Note the tight fit!

The area under the parabola is the product of two numbers: the length of the full interval for which we're approximating the area (0.5 in this case) and a weighted average of the three values of the function. The weights are 1, 4, 1, so the weighted average in this case is

$$\frac{1}{6}(f(0.5) + 4f(0.75) + f(1.0)) \approx 0.8222.$$

The area estimate is then

$$0.5 \times 0.822\ldots \approx 0.411104717816652.$$

The difference between this number and the true value is $0.000183\ldots$. If I'm not mistaken, the Simpson estimate is the best so far. And we've used only three points!

The general formula

Simpson's rule works when the number of slivers is even. (For us, it was 2.) Suppose the number of slivers is n, and let $\Delta x = \dfrac{b-a}{n}$, as usual. Then the area is estimated by:

$$\frac{\Delta x}{3}[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 4f(x_{n-1)} + f(x_n)].$$

The numbers x_0,. . . , x_n are the equally spaced endpoints of the small segments. Each number $x_1, \ldots, x_{n-1}$ is the end of one segment and the beginning of the next. The weights "2" in the formula come about as $1 + 1$: $f(x_2)$, for example, occurs for the first of the $n/2$ parabolas, but also for the second.

When $n = 2$, the general formula specializes to

$$\frac{\Delta x}{3}[f(a) + 4f(\frac{a+b}{2}) + f(b)].$$

The quantity $\frac{\Delta x}{3}$ may be rewritten $\frac{b-a}{6}$, which explains the "6" as our previous denominator.