Math 6500 Additional Material: Numerical Improper Integrals

Suppose we are interesting in integrating

$$\int_0^\infty f(x)\,dx$$

numerically. At first, you might think that this is impossible: how are we supposed to divide $(0, \infty)$ into subregions? There are two elementary approaches.

1. Make a u-substitution to change to a finite domain

If we make the substitution $x = \tan u$, $dx = \sec^2 u\,du$ we can transform

$$\int_0^\infty f(x)\,dx = \int_0^{\frac{\pi}{2}} f(\tan u)\sec^2 u\,du.$$

We can then apply the trapezoid rule as usual to the right-hand integral. This is not the kind of u-substitution that we're used to, because the function on the right doesn't look any easier to integrate than the function on the left– in fact, it probably looks harder! However, the only thing we need to do with the right-hand side is to compute values of the integrand. We usually can (although we may run into trouble evaluating at $u = \pi/2$ where the value has to be computed by doing the limit). The error will be controlled by the second derivative of the transformed function $f(\tan u)\sec^2 u$ using the usual formula[1].

For example, suppose we consider the *Gaussian Probability Integral* $\int_0^\infty e^{-x^2}\,dx$ which turns out[2] to be exactly equal to $\frac{\sqrt{\pi}}{2}$. If we make the u-substitution $x = \tan u$,

$$\int_0^\infty e^{-x^2}\,dx = \int_0^{\frac{\pi}{2}} e^{-\tan^2 u}\sec^2 u\,du.$$

We can compute the value at $\pi/2$ by proving that $\lim_{u\to\pi/2} e^{-\tan^2 u}\sec^2 u = 0$.

---

[1]We could improve this process by choosing a different u-substitution which reduces the second derivative of this function, but doing so takes some knowledge of the function f. For instance, the best possible thing would be to solve the differential equation

$$f(g(u))g'(u) = C \quad \text{or} \quad \frac{d}{du}f(g(u))g'(u) = 0.$$

for some function $g(u)$ with $g(a) = 0$ and $g(b) = \infty$ and then make the u-substitution $x = g(u)$. That would transform the integrand to a constant function, and make the error in the trapezoid rule vanish altogether. However, if you could do that, you could probably do the integral exactly anyway.

[2]You can prove this using complex analysis by rewriting this as a contour integral; you might have done this in MATH 4150.

2. IGNORE THE TAIL

Assuming the improper integral converges at all, we know that by definition

$$\int_0^\infty f(x)\,dx = \lim_{b\to\infty}\int_0^b f(x)\,dx.$$

Therefore, since we know that

$$\int_0^\infty f(x)\,dx = \int_0^b f(x)\,dx + \int_b^\infty f(x)\,dx,$$

we can conclude that if $T_f(b) = \int_b^\infty f(x)\,dx$[3], then

$$\lim_{b\to\infty} T_f(b) = 0.$$

Therefore, we can approximate the improper integral $\int_0^\infty f(x)\,dx$ by the integral $\int_0^b f(x)\,dx$ and consider $T_f(b)$ to be part of the approximation error. Our usual error bound for the trapezoid rule applies to $\int_0^b f(x)\,dx$, but we need a bound for $|T_f(b)|$ to control the overall error. Such a bound is called a "tail bound" for f. Finding tail bounds is not something for which there's a general procedure which always works; in each case, you have to think about the function.

However, a method that works often is to find a function $g(x) \geq f(x)$ for which you can compute the integral $\int_b^\infty g(x)\,dx$. Then you can integrate both sides of the inequality $g(x) \geq f(x)$ to conclude that

$$\int_b^\infty g(x)\,dx \geq \int_b^\infty f(x)\,dx = T_f(b).$$

For instance, for the Gaussian probability integral $\int_0^\infty e^{-x^2}\,dx$, if $x > 1/2$, we might take[4] $2x\,e^{-x^2} > e^{-x^2}$. Integrating both sides of the equation, (if $b \geq 1/2$) and observing that the left-hand integral is easy, we have

$$\int_b^\infty 2x\,e^{-x^2}\,dx > \int_b^\infty e^{-x^2}\,dx \implies e^{-b^2} > \int_b^\infty e^{-x^2}\,dx.$$

However, now that we know that $e^{-b^2} > T_f(b)$, we can already see that $T_f(b)$ is *really* small even for quite modest b. For instance, if $b = 5$, we see $e^{-25} \sim 1.3\times 10^{-11} > T_f(b)$. This is already a lot smaller than the approximation error in using the trapezoid rule unless we take a very large number of subdivisions[5].

[3]The notation $T_f(b)$ isn't particularly standard, but I chose it because this portion of the function $f(x)$ with $x \in [b,\infty)$ is called the "tail" of f.

[4]A sharp-eyed reader might observe that $2x\,e^{-x^2}$ is a *lot* bigger than e^{-x^2} for large x and presume that therefore this bound could be improved. See problem 3 for one approach to doing this.

[5]in which case, we'd just increase b accordingly …

3. MINIHOMEWORK

1. Use both strategies to compute the *Gaussian probability integral*:

$$\int_0^\infty f(x)\,dx = \int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}.$$

(use Mathematica to do the trapezoid rule integrations, of course) with 10, 50, and 100 subintervals. It will be necessary to use `Piecewise` to define the transformed integrand (because otherwise Mathematica won't be able to evaluate it at $\pi/2$).

a. For the "make a u-substitution" method, find a bound $M \geq \left|\frac{d^2}{du^2} f(\tan u)\sec^2 u\right|$ for u in $(0, \pi/2)$ (if one exists) and $f(x) = e^{-x^2}$ and use the formula

$$\left|\mathrm{Trap}(f,n) - \int_a^b f(x)\,dx\right| < \frac{1}{12}M(b-a)h^2$$

(with $(a,b) = (0,\pi/2)$) to bound the integration error for 10, 50 and 100 intervals.

b. For the "make a u-substitution" method, compute the actual error of the trapezoid rule

$$\left|\mathrm{Trap}(f,n) - \int_a^b f(x)\,dx\right|$$

(again, for $f(\tan u)\sec^2 u = e^{-(\tan u)^2}\sec^2 u$ and $(a,b) = (0,\pi/2)$) for 10, 50, and 100 intervals.

c. For the "ignore the tail" method, find a bound $M \geq \left|\frac{d^2}{dx^2} f(x)\right|$ for x in $(0,b)$ and choose b so that the combined error bound

$$\left|\mathrm{Trap}(f,n) - \int_0^b f(x)\,dx\right| + \left|\int_b^\infty f(x)\,dx\right| \leq \frac{1}{12}M(b-0)h^2 + e^{-b^2}$$

is minimized. Find the best combined error bound for 10, 50, and 100 intervals.

d. For the "ignore the tail" method, compute the actual error (with your choices of b) in using 10, 50 and 100 intervals.

e. Write a concluding paragraph: which method had a better error bound? Which method was actually better in practice? Why?

2. Complete steps a.-e. of Problem 1, for the new integral

$$\int_0^\infty f(x)\,dx = \int_0^\infty \frac{\sin x}{x^{3/2}}\,dx$$

Note that this time you'll have to find your own tail bound

$$T_f(b) \geq \left|\int_b^\infty \frac{\sin x}{x^{3/2}}\,dx\right|.$$

You can start by finding a function $g(x) \geq \left|\frac{\sin x}{x^{3/2}}\right|$ and observing that

$$\int_b^\infty g(x)\,dx \geq \int_b^\infty \left|\frac{\sin x}{x^{3/2}}\right|\,dx \geq \left|\int_b^\infty \frac{\sin x}{x^{3/2}}\,dx\right|$$

as we did above for e^{-x^2}.

3. Derive Feller's tail estimate[6] for the Gaussian probability integrand by integrating both sides of the inequality
$$e^{-x^2}\left(1+\frac{1}{2x^2}\right) \geq e^{-x^2}.$$
If you can do the integration on the left by hand, that's great! But at this point, doing the integral with a table or Mathematica is ok as long as you say that's how you did it.

4. (Extra credit) Choose new b values for the integration of the Gaussian probability integral by the "ignore the tail" method using this improved tail bound and redo the error bound for 10, 50, and 100 intervals accordingly. Does it make much of a difference?

[6] *Introduction to Probability*, Vol 1, Lemma 2 on p. 166