Strategies for estimating integrals and sums

In previous courses in analysis, you have all evaluated a variety of sums and integrals. In this course, we frequently encounter complicated sums and integrals. In most cases, I do not know how to evaluate these sums and integrals exactly. Instead, we try to estimate the integrals, by proving upper and lower bounds for them. In this course, if we estimate a certain integral up to a factor of 100, I usually count it as a big success. For example, on Problem Set 3, the following integral appeared.

\[ I_N = \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 + \cos x)^N \, dx. \]

It's fairly difficult to estimate the integral \( I_N \) up to a factor of 100. We will use it as an example from time to time as we go through some methods of estimating sums and integrals. The goal of this essay is to systematically discuss the most fundamental methods of proving upper and lower bounds for sums and integrals. This first draft probably does not live up to the goal. Your feedback is welcome.

The first thing I do when I look at a sum or integral is to check whether all of the terms have the same sign. The situation is much simpler if all the terms are positive than if the terms change sign.

1. Positive sums and integrals

We consider here either sums \( \sum_{n=1}^{N} c_n \) or integrals \( \int_{a}^{b} f(x) \, dx \). When looking at sums, we assume \( c_n \) is positive, and when looking at integrals we assume \( f(x) \) is positive at every point \( x \).

We begin with the simplest bounds and slowly work towards more complicated strategies. The simplest strategies work in the most general situations, and you should always check them before trying more complicated strategies.

U1. The First Upper Bound

The simplest upper bound controls a sum in terms of its largest summand.

\[ \sum_{n=1}^{N} c_n \leq N \max(c_n). \quad U1 \]

\[ \int_{a}^{b} f(x) \, dx \leq (b - a) \max(f(x)). \quad U1 \]

We have used this estimate many times in class. We will call it U1 for the first upper bound.

For example, using this estimate, it follows that the integral \( I_N \) is at most \( 2^N \). That estimate turns out to be off by more than a factor of 100 for large \( N \), but it’s not too bad. An example where the bound U1 works well is the sum \( \sum_{n=1}^{N} n^2 \). The largest term in the sum is \( N^2 \), so U1 gives an upper bound of \( N^3 \). The actual value of the sum is roughly \((1/3)N^3 \) for large \( N \). An example where the bound U1 works badly is the sum \( \sum_{n=1}^{N} n^{-2} \). The largest term in the sum is 1, so U1 gives an upper bound of \( N \). The actual value of the sum is less than 2, regardless of \( N \). The bound U1 works badly when the sum consists mostly of terms that are much smaller than the largest term. The bound U1 says nothing at all for infinite sums.
or integrals with infinite limits (i.e. \( \int_0^\infty \)). It usually does a good job for estimating the integrals of functions on the circle.

L1. The First Lower Bound

Now we turn to lower bounds. There is a lower bound exactly analogous to the first upper bound. Namely, \( \sum_{n=1}^N c_n \geq N \min(c_n) \). This estimate is often lousy, even by the standards of doing the simplest thing you can think of. The reason is that adding a single small term to the end of the sum \( \sum_{n=1}^N c_n \) makes our estimate tiny. It’s amusing to think what this estimate tells us about \( \sum_{n=1}^N n^{-2} \). The smallest term in the sum is \( N^{-2} \), and so our lower bound is \( N^{-1} \). As we add more positive numbers in our sum, our lower bound gets smaller and goes to zero! The basic lower bound estimate is therefore the following:

If \( c_n \geq H \) for \( M \) values of \( n \), then

\[
\sum_{n=1}^N c_n \geq MH. \quad L1
\]

There is a similar estimate for integrals. If \( f(x) \geq H \) on an interval \([c, d]\) contained in \([a, b]\), then

\[
\int_a^b f(x) dx \geq (d-c)H.
\]

The letter \( H \) here stands for height. The region where \( f(x) \geq H \) may not be a single interval. More generally, if there is a union of disjoint intervals in \([a, b]\) with total length \( W \) so that \( f(x) \geq H \) on each interval then

\[
\int_a^b f(x) dx \geq WH.
\]

We also call these lower bounds \( L1 \), for the first lower bound.

This simple lower bound is remarkably good. It’s actually a challenge to think of an example where the lower bound is misleading. Before we do a few examples, you may want to try for a minute.

To show how well the lower bound \( L1 \) works, consider the sum \( \sum_{n=1}^N n^\alpha \) for some number \( \alpha \). In order to use the lower bound, we have to pick a number \( H \). (It takes a little practice to pick \( H \) in a strategic way. First we give some examples, and then we discuss strategy.)

If \( \alpha > 0 \), then the terms in the sum are increasing. We choose \( M \) to be close to the largest value \( N^\alpha \), but a little bit less. More exactly, take \( H = (1/2)N^\alpha \). The summand \( N^\alpha \) is at least \( H \) whenever \( n \) is at least \( (1/2)^{\alpha-1} N \). Therefore, we get a lower bound for our sum of \( (1/2)N^\alpha(1-(1/2)^{\alpha-1})N \), which is roughly \( N^{1+\alpha} \). This answer is fairly accurate. I believe that it is correct up to a factor of 2 for all values of \( N \) and all values of \( \alpha > 0 \).

If \( \alpha < 0 \), then the terms in the sum are decreasing. In this case, we take \( H \) to be the lowest term in the sum, \( N^\alpha \). Our lower bound is then \( NN^\alpha = N^{1+\alpha} \). This lower bound is also pretty good. The worst case is when \( \alpha = -1 \). In this case, the sum is actually close to \( \log N \), but our lower bound is 1. For every other value of \( \alpha \), the estimate is within some factor of the right answer uniformly for all \( N \). For example, if \( 0 \geq \alpha > -0.9 \) or \( \alpha < -1.1 \), then our lower bound is within a factor of 20 of the actual sum.

So far, we have used two strategies to guess \( H \). The first strategy is to pick \( H \) to be half the maximal value of the summands. The second strategy was to pick \( H \) to be the minimum of the summands. I’d like to say the second strategy a little bit differently: take \( H \) to be the most common value of the summands, or maybe half of that. One of these strategies usually works well. Although it is possible to design
sums where some other value of $H$ works much better than these two strategies, such sums rarely appear in practice.

Next, let’s apply this lower bound to the integral $I_N$. We have to consider the integrand $(1 + \cos x)^N$ on the interval $[-\pi, \pi]$. The maximum value is $2^N$ which occurs at $x = 0$. Trying the first strategy, we take $H$ to be $(1/2)^{2N}$. Now we have to estimate the size of the interval where $(1 + \cos x)^N \geq H$. In other words, we are looking at the inequality

$$(1 + \cos x)^N \geq (1/2)^{2N}.$$  

Taking $N^{th}$ roots of both sides gives us the inequality

$$\cos x \geq 2^{1/(2N)} - 1.$$  

For large $N$, the right hand side is a little bit less than 1. The inequality is satisfied only when $x$ is close to zero. We can simplify the equation, introducing an error of no more than a factor of 2, by saying that $\cos(x)$ is roughly $1 - x^2/2$. Putting this in the above inequality gives us

$$x^2/2 \leq 2(1 - (1/2)^{1/N}).$$  

Next we say that $1 - (1/2)^{1/N}$ is roughly $(1/N)$, introducing another error of a factor of 2 or so. Finally we get the estimate

$$|x| \leq 2N^{-1/2}.$$  

Since we had a little error at each step, we can safely say that $(1 + \cos x)^N \geq H$ as long as $|x| < (1/10)N^{-1/2}$. The lower bound $L1$ now gives $I_N \geq (1/20\pi)N^{-1/2}2^N$.

We can summarize our upper and lower bounds for $I_N$ as follows.

$$(1/20\pi)N^{-1/2}2^N \leq I_N \leq 2^N.$$  

These bounds are already quite good. We have determined $I_N$ up to a factor of $20\pi N^{1/2}$. Considering that the value of $I_N$ is so high, the size of this factor is quite small, even though for most $N$ it is bigger than 100.

Other strategies of choosing the height $H$ do not work very well. The most common value of $(1 + \cos x)^N$ is pretty near to 1, which gives an estimate $I_N \geq 1/2$. Other values of $H$ also do not do as well as the estimate above using $H = (1/2)^{2N}$.

As often happens, the lower bound $L1$ is more accurate than the upper bound $U1$. With more work, you can show that $(1/2\pi)(1/10)N^{-1/2}2^N \leq I_N \leq 100N^{-1/2}2^N$, which estimates $I_N$ up to a factor of 100 in either direction. I want to emphasize, though, that the bounds U1 and L1 are very good in this case. The integral $I_N$ appeared in Problem Set 3, and these bounds are easily good enough to solve the problem involving $I_N$.

**U2. The Second Upper Bound**

The upper bound U1 did a respectable job in estimating $I_N$, but it does not work at all for infinite sums or integrals. The most common technique for infinite sums is to compare them with simple convergent infinite sums. If $c_n \leq d_n$, then of course $\sum_{n=1}^{N} c_n \leq \sum_{n=1}^{N} d_n$. One approach is to search for an appropriate sum $d_n$ which is much easier to calculate than the original sum. In theory, many sums could be used for comparison, but in practice only the simplest few are usually needed.
These are $d_n = a^{-n}$ for some number $a > 1$ or $d_n = n^{-\alpha}$ for some power $\alpha > 0$. The sum $\sum_{n=1}^{N} d_n$ is easy to estimate in these cases: in the first case by geometric series and in the second by approximating it with the integral $\int_1^{N} x^{-\alpha} dx$. We call this upper bound $U_2$ for the second upper bound.

**L2. The Second Lower Bound**

This strategy can also be applied to get lower bounds, but it is needed less often because $L_1$ is already surprisingly good. Nevertheless, if $c_n \geq d_n$, then $\sum_{n=1}^{N} c_n \geq \sum_{n=1}^{N} d_n$. The most useful example is when $d_n = 1/n$, in which case the lower bound $L_1$ had some room for improvement. We call this lower bound $L_2$. It’s not needed very often, but there was at least one example in our course. In Lecture 7, we estimated $\int_{-\pi}^{\pi} |D_N(x)| dx$. We broke the integral into roughly $N$ pieces and showed that the $j^{th}$ term was bigger than $\frac{1}{100}$. In this way we checked that the integral is greater than $(1/100) \log N$, using $L_2$. This lower bound is correct up a factor of 1000 or so.

**B1. Breaking the Integral into Two Parts**

There is just one more common technique for getting upper bounds. This technique is to break the integral or sum into pieces and use different strategies on the different pieces. For example, consider the following integral involving a constant $h > 0$.

$$\int_0^{\infty} |\sin (hx)| x^{-3/2} dx.$$ 

Since the domain of integration is infinite, the bound $U_1$ does not work at all. We pass to $U_2$. We notice that the integrand is bounded by the simple function $x^{-3/2}$. Therefore, our integral is bounded by the integral

$$\int_0^{\infty} x^{-3/2} dx.$$ 

Unfortunately, this integral is infinite. We notice, though, that $|\sin (hx)| \leq hx$, and therefore the integrand is bounded by the simple function $hx^{-1/2}$. Therefore, our integral is bounded by the integral

$$\int_0^{\infty} hx^{-1/2} dx.$$ 

Unfortunately, this integral is also infinite. There is a sense, however, in which we have made progress. Our first strategy produced an integral with a very big contribution near zero but a decent contribution from the large values of $x$. Our second strategy produced an integral with a decent size for small values of $x$ but a very big contribution from large values of $x$. We should combine these two strategies.

In other words, we have two bounds on $|\sin (hx)|$. The first bound is 1. The second bound is $hx$. The second bound is better when $x < (1/h)$, and the first bound is better when $x > (1/h)$. Combining our two bounds, our original integral is at most the sum
\[ \int_0^{1/h} h x^{-1/2} \, dx + \int_{1/h}^{\infty} x^{-3/2} \, dx. \]

This sum works out to \(2 h^{1/2} + (2/3) h^{1/2} = (8/3) h^{1/2}\). Using L1 to get a lower bound, we can check that this estimate is correct up to a factor of 100. Integrals of this kind appear in Fourier analysis - for example in exercise 3 of Problem Set 4.

2. Cancellation

We now turn to the more complicated situation of estimating an integral \(\int_a^b f(x) \, dx\) when the function \(f(x)\) changes sign. In this case, the integral may be positive or negative, and we often want to estimate \(|\int_a^b f(x) \, dx|\). This norm is certainly bounded by \(\int_a^b |f(x)| \, dx\), which we can estimate using the strategies in the last section. This method, however, disregards any cancellations in the integral. If the positive and negative parts of the function \(f\) cancel almost exactly, then \(|\int f|\) will be much smaller than \(\int |f|\). The main topic of this section is to find upper bounds for \(|\int f|\) that exploit cancellation in order to improve the bound \(\int |f|\).

Integrals with lots of cancellation appear frequently in Fourier analysis, and proving bounds that take cancellation into account is close to the heart of Fourier analysis. The first example appears in the formula for the Fourier coefficients,

\[ 2\pi \hat{f}(n) = \int_{-\pi}^{\pi} e^{-iny} f(y) \, dy. \]

If \(n\) is large, the integrand oscillates rapidly. Its real part goes back rapidly between positive and negative values, and the positive and negative parts cancel to a considerable degree, depending on the function \(f\).

IBP. Integration by Parts

The primary technique for dealing with oscillatory integrals is integration by parts. The point is that the antiderivative of \(e^{-iny}\) is \(-\frac{i}{n} e^{-iny}\), which is much smaller than the original function. For example, we can estimate the norm of the following integral.

\[ \int_0^1 \cos(nx) x^4 \, dx. \]

Integrating by parts yields a boundary term of norm at most \(1/|n|\) and a new integral

\[ - \int_0^1 (1/n) \sin(nx)(4x^3) \, dx. \]

The norm of this integral is bounded by \(4/|n|\), giving us a bound of \(5/|n|\) for the norm of the original integral. This estimate is within a factor of 100 of the actual norm for most values of \(n\).

B2. Breaking the Integral into Two Parts, continued

When we do integration by parts, we sometimes pay a price, because we have to differentiate the other term in the integrand, which may make this term much bigger. For example, let us try to estimate the norm of the following integral for large \(n\).
\[
\int_0^1 x^{-1/2} e^{inx} dx.
\]

If we don’t use integration by parts, we can estimate the norm of the integral by \( \int_0^1 x^{-1/2} dx \) which is equal to 2. We might suspect that the actual integral is much smaller because of cancellation. We try to exploit the cancellation by integrating by parts, but we find that the boundary terms that appear are infinite! Formally we get \( x^{-1/2} (1/in) e^{inx} \bigg|_0^1 - \int_0^1 x^{-3/2} (1/in) e^{inx} dx \). The boundary term is infinite and the integral diverges. The problem is that integration by parts works very badly near zero where \( x^{-3/2} \) is very large. But it works well away from zero. For example, we can estimate \( |\int_1^{1/2} x^{-1/2} e^{inx} dx| \) using integration by parts and show that it is bounded by \( 4/|n| \). For large \( n \), that estimate is much better than what we would get by ignoring the cancellation.

In other words, integration by parts is a good strategy away from zero but a bad strategy near zero. When we have two different strategies that work in different areas, we should consider breaking our integral into parts. A priori, I may not be sure where to break it, so I defer that decision by dividing the integral at a point called \( a \).

\[
\int_0^1 x^{-1/2} e^{inx} dx = \int_0^a x^{-1/2} e^{inx} dx + \int_a^1 x^{-1/2} e^{inx} dx.
\]

Then we use integration by parts on the second term on the RHS. It is equal to the following expression.

\[
x^{-1/2} (1/in) e^{inx} \bigg|_a^1 + (1/2) \int_a^1 x^{-3/2} (1/in) e^{inx} dx.
\]

The boundary term has norm roughly \( a^{-1/2} / |n| \), and the integral has norm at most \( (1/2) (1/|n|) \int_a^1 x^{-3/2} dx \), which is also roughly \( a^{-1/2} / |n| \). Finally, the first term on the RHS of equation 1, the integral \( \int_a^1 x^{-1/2} e^{inx} dx \), has norm at most \( 2a^{1/2} \). In total, the norm of our first integral is bounded by roughly \( a^{1/2} + a^{-1/2} / |n| \). This bound holds for every value of \( a \), and we get to choose the value of \( a \) that gives us the best estimate. We choose \( a = 1/|n| \), which gives an upper bound of roughly \( 2n^{-1/2} \).

Since there were a few approximations along the way, our initial integral is bounded by \( 10n^{-1/2} \). This estimate is right up to a factor of 100.

**O. Estimates Using Orthogonality**

Finally we mention that orthogonality can be used to estimate the size of certain integrals or sums taking into account some cancellation. Consider for example the following integral.

\[
\int_0^{2\pi} \left| \sum_{n=-N}^N e^{inx} \right|^2 dx.
\]

At first glance, this integral belongs in the first category, because the integrand is always positive, and so there is no cancellation in the integral. At each point, however, the integrand involves the sum \( \sum_{n=-N}^N e^{inx} \), and a good estimate of our integral needs to exploit some cancellation in the sum. If we estimate the sum
without taking into account any cancellation, we get \[ |\sum_{n=-N}^{N} e^{inx}| \leq 2N + 1. \]

Using this bound to estimate the integral, we see that it is at most \((2\pi)(2N + 1)^2\).

We can get a better estimate of our integral by exploiting the fact that the functions \(e^{inx}\) are orthogonal. Recall that in Lecture 9 we defined a Hermitian inner product for functions \(\langle f, g \rangle = \frac{1}{\pi} \int_{0}^{2\pi} f(x)\overline{g(x)} dx\). In terms of that inner product, our integral is \((2\pi)\langle \sum_{n=-N}^{N} e^{inx}, \sum_{n=-N}^{N} e^{inx} \rangle\). Using the orthogonality of the functions \(e^{inx}\), we see that this integral is equal to \((2\pi)(2N + 1)\). This estimate implies a degree of cancellation in the sum \( |\sum_{n=-N}^{N} e^{inx}| \). At least for most values of \(x\), it says that this sum is not bigger than \(10N^{1/2}\). If there were no cancellation, the sum would be \(2N + 1\), so our estimate implies quite a bit of cancellation for most values of \(x\). Nevertheless, there are some values of \(x\) where there is very little cancellation. For example, when \(x = 0\), the sum \( \sum_{n=-N}^{N} e^{inx} = 2N + 1\).

3. Summary and Review

The methods in this essay don’t give very accurate answers. Even for very simple integrals, our main bounds U1, L1, and U2 will be off by a factor of 5. They have two strengths. The first is that they can be applied in very many situations. The second is that they can be applied fairly quickly and the computations are fairly simple.

When you encounter a sum or integral that you want to estimate, you might want to approach it by following the sequence of strategies in this essay. Of course you don’t have to, and there may be some tough integrals which won’t crack this way. I would be interested to find examples of that kind. If you wanted to follow this essay, you might go through the following steps.

First, check whether the function is always positive or whether it changes sign.

I. If the function is always positive.

Try the upper bound U1 based on the maximum of the function.

Try the lower bound L1. In L1, you have to pick a value of the height \(H\). First try one half the maximum value of the integrand. Then try the most common value of the integrand.

Compare the upper bound U1 to the lower bound L1. If there is a big gap, you could try more refined techniques. More likely than not, the upper bound U1 is too big. Probably the next thing to try is U2 - especially if the domain of integration is very large. If that doesn’t solve the problem, you could try breaking the integral into pieces and using B1. Finally, if you are suspicious of the lower bound, you could try L2.

II. If the function changes sign

First try to upper bound \( |\int f| \) without taking cancellation into account by bounding \( f \) \(|f|\) using the techniques above. In many cases this will be very quick.

Next, try to improve the bound by using integration by parts. If that doesn’t work, but seems to work in some parts of the domain, you should try breaking the integral into pieces as in B2. If there is a sum in the integral, you could try to use orthogonality.