Approximation and asymptotics

1. **Sums and integrals.** Let $f\colon [1,\infty) \to \mathbb{R}$ be decreasing: i.e. such that $f(x) \geq f(y)$ when $x \leq y$. Then, as shown in Figure 1, for any positive integer n, we have

$$\sum_{j=1}^{n} f(j) \geq \int_{1}^{n+1} f(x)\, dx \geq \sum_{j=2}^{n+1} f(j). \tag{1.1}$$

FIGURE 1. An example of (1.1) with $n = 4$.

Rearranging yields

$$0 \leq \sum_{j=1}^{n} f(j) - \int_{1}^{n+1} f(x)\, dx \leq f(1) - f(n+1).$$

EXAMPLE 1.2. Let us apply this to the harmonic series.

$$0 \leq \sum_{j=1}^{n} \frac{1}{j} - \ln(n+1) \leq 1 - \frac{1}{n+1}.$$

A simplified version of this result is

$$\sum_{j=1}^{n} \frac{1}{j} \sim \ln(n+1) \sim \ln n,$$

where we use the notation

$$a_n \sim b_n \qquad \Longleftrightarrow \qquad \lim_{n \to \infty} \frac{a_n}{b_n} = 1.$$

In words, we say that a_n and b_n are *asymptotic* to one another.

This concept is helpful for many problems, but can lead to pitfalls. For example, let us try to use it to evaluate the alternating harmonic series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots .$$

The even partial sums are

$$s_{2k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots - \frac{1}{2k} = \sum_{j=1}^{k} \frac{1}{2j-1} - \frac{1}{2j},$$

and we have

$$\sum_{j=1}^{k} \frac{1}{2j} \sim \frac{1}{2} \ln k, \qquad \sum_{j=1}^{k} \frac{1}{2j-1} \sim \int_{1}^{k+1} \frac{dx}{2x-1} = \frac{1}{2} \ln(2k+1).$$

However, it is incorrect to conclude that

$$s_{2k} \sim \frac{1}{2} \ln(2k+1) - \frac{1}{2} \ln k = \frac{1}{2} \ln(2 + \tfrac{1}{k}) \sim \frac{1}{2} \ln 2, \tag{1.3}$$

as we can see from the fact that

$$s_{2k} > s_{2(k-1)} > \cdots > s_2 = \frac{1}{2}, \qquad \text{while} \qquad \frac{1}{2} \ln 2 = \frac{1}{2} \int_{1}^{2} dx/x < \frac{1}{2} \int_{1}^{2} dx = \frac{1}{2}.$$

The mistake is subtracting the asymptotics in the first step of (1.3). A simpler version of the same mistake is subtracting

$$n + 2 \sim n, \qquad n + 1 \sim n,$$

to obtain $1 \sim 0$.

Instead, replace (1.3) by

$$s_{2k} = \sum_{j=1}^{2k} \frac{1}{j} - 2 \sum_{j=1}^{k} \frac{1}{2j} = \sum_{j=k+1}^{2k} \frac{1}{j},$$

where we took the partial sum for the harmonic series with all positive signs, and then subtracted the even terms twice to correct their signs to negative. To evaluate the last sum, observe that (1.1) implies

$$\int_{k+1}^{2k+1} \frac{dx}{x} \leq \sum_{j=k+1}^{2k} \frac{1}{j} \leq \int_{k}^{2k} \frac{dx}{x}.$$

Evaluating the integrals and taking $k \to \infty$ yields

$$s_{2k} = \sum_{j=k+1}^{2k} \frac{1}{j} \to \ln 2. \tag{1.4}$$

It remains to compute the limit of the odd partial sums:

$$s_{2k+1} = s_{2k} + \frac{1}{2k+1} \to \ln 2 + 0. \tag{1.5}$$

Combining (1.4) and (1.5) yields

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \ln 2. \tag{1.6}$$

EXERCISE 1.7. Consider the rearrangement

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} \cdots \qquad (1.8)$$

of the alternating harmonic series: in this arrangement, terms number $3k-2$, $3k-1$, and $3k$ are $\frac{1}{4k-3}$, $\frac{1}{4k-1}$, and $-\frac{2}{4k}$. Let s_n be the sum of the first n terms of this series.

(1) Use mathematical induction to show that for any positive integer k, we have

$$s_{3k} = \sum_{j=1}^{2k} \frac{(-1)^{j+1}}{j} + \sum_{j=k}^{2k-1} \frac{1}{2j+1}.$$

(2) Find $\lim s_{3k}$, using the result of part (1), (1.6), and the method used to obtain (1.4).
(3) Prove that s_{3k+1} and s_{3k+2} have the same limit as s_{3k}, and thereby evaluate the sum (1.8)

A variant of (1.1) that leads to bounds on the factorial function for $n \geq 2$ is

$$\int_1^n \ln x \, dx \leq \sum_{j=1}^n \ln j \leq \ln n + \int_2^n \ln x \, dx, \qquad (1.9)$$

or, using $\frac{d}{dx}(x \ln x - x) = \ln x$,

$$n \ln n - n + 1 \leq \sum_{j=1}^n \ln j \leq (n+1) \ln n - n - 2 \ln 2 + 2. \qquad (1.10)$$

Dividing through by the left side in (1.10) yields

$$1 \leq \frac{\sum_{j=1}^n \ln j}{n \ln n - n + 1} \leq \frac{(n+1) \ln n - n - 2 \ln 2 + 2}{n \ln n - n + 1} = 1 + r_n,$$

where

$$r_n = \frac{\ln n - 2 \ln 2 + 1}{n \ln n - n + 1} \sim \frac{1}{n}.$$

Consequently,

$$n \ln n - n + 1$$

is not only a lower bound to but also a good approximation to

$$\sum_{j=1}^n \ln j = \ln n!$$

as n gets large.

EXERCISE 1.11. Let $p > -1$ be given. Find constants a, b, B, c, and C in terms of p, such that

$$an^{p+1} + bn^p + c \leq \sum_{j=1}^n j^p \leq an^{p+1} + Bn^p + C$$

for every positive integer n.

To convert (1.10) to bounds on $n!$, we exponentiate , which yields

$$e^{n \ln n - n + 1} \leq n! \leq e^{(n+1) \ln n - n - 2 \ln 2 + 2}. \qquad (1.12)$$

This can be written more compactly as

$$n^n e^{1-n} \le n! \le n^{n+1} e^{2-n}/4.$$

The problem is that these bounds are very far apart: if we divide them we get

$$\frac{n^{n+1} e^{2-n}/4}{n^n e^{1-n}} = \frac{en}{4}.$$

Since the right side goes to ∞ as $n \to \infty$, we see that we are only recovering the rough order of magnitude of $n!$. To get a decent result for $n!$, and not just for $\ln n!$ requires finer analysis. We use Taylor approximation.

2. **Taylor approximation and Newton's method.** If f is differentiable near p, then the mean value theorem yields

$$f(p+h) = f(p) + hf'(c),$$

where c is some number between p and $p+h$. Assuming f' is continuous near p, we can rewrite this in a more complicated but more explicit way using the fundamental theorem of calculus:

$$f(p+h) = f(p) + \int_p^{p+h} f'(s)\, ds = f(p) + h \int_0^1 f'(p+ht)\, dt.$$

Equating the two yields

$$f'(c) = \int_0^1 f'(p+ht)\, dt;$$

thus instead of evaluating f' and an unspecified c between p and $p+h$, we can average f' over the values it takes on between p and $p+h$.

Assuming f'' exists and is continuous, we can integrate by parts:

$$\int_p^{p+h} f'(s)\, ds = hf'(p) + \int_p^{p+h} (p+h-s) f''(s)\, ds = hf'(p) + h^2 \int_0^1 (1-t) f''(p+ht)\, dt.$$

This yields the *second order Taylor approximation*:

$$f(p+h) = f(p) + hf'(p) + R_2(p,h), \qquad R_2(p,h) = h^2 \int_0^1 (1-t) f''(p+ht)\, dt. \qquad (2.1)$$

EXERCISE 2.2. Let I be an open interval, p a point in I, n a positive integer. Let $f : I \to \mathbb{R}$ be $n+1$ times differentiable with $f^{(n+1)}$ continuous. Use mathematical induction and integration by parts to prove that

$$f(p+h) = f(p) + f'(p)h + \frac{f''(p)}{2!} h^2 + \cdots + \frac{f^{(n)}(p)}{n!} h^n + R_{n+1}(p,h),$$

where

$$R_{n+1}(p,h) = \frac{1}{n!} \int_p^{p+h} (p+h-s)^n f^{(n+1)}(s)\, ds = \frac{h^{n+1}}{n!} \int_0^1 (1-t)^n f^{(n+1)}(p+ht)\, dt.$$

The second order Taylor approximation can be used to prove convergence of Newton's method for solving equations, under favorable conditions. Let f be continuous and twice differentiable. Given an approximate solution x_n to $f(x) = 0$, we find an improved solution x_{n+1} as follows: let $\ell(x)$ be the tangent line at x_n to $f(x)$, and let x_{n+1} be the solution to the linear equation $\ell(x_{n+1}) = 0$:
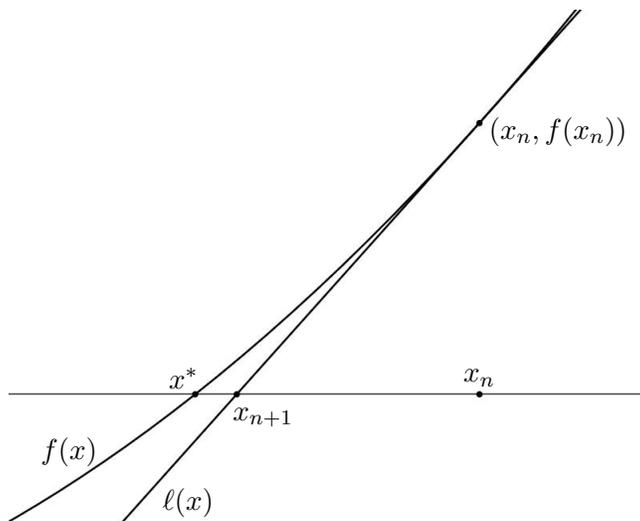
Explicitly, $\ell(x) = f'(x_n)x + f(x_n) - f'(x_n)x_n$, so

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{2.3}$$

This clearly fails if $f'(x_n) = 0$ so let us assume that $f'(x) \neq 0$ for x near the desired solution x^*. To prove that x_{n+1} is a better approximation than x_n, we use (2.1) with $p = x_n$, $p + h = x^*$, to write

$$0 = f(x_n) + (x^* - x_n)f'(x_n) + R_2,$$

and thus

$$x_{n+1} - x^* = \frac{R_2}{f'(x_n)} = (x_n - x^*)^2 A_n, \qquad A_n = \frac{\int_0^1 (1-t)f''(x_n + (x^* - x_n)t))\, dt}{f'(x_n)}.$$

EXAMPLE 2.4. A particularly nice case is $f(x) = x^2 - q$, used for computing the square root of q. Then

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{q}{x_n}\right),$$

and

$$A_n = \frac{\int_0^1 (1-t)\, dt}{x_n} = \frac{1}{2x_n}.$$

Writing $x = \sqrt{q}$, and taking $x_n > \sqrt{q}$ yields

$$0 < \frac{x_{n+1} - \sqrt{q}}{\sqrt{q}} = \frac{(x_n - \sqrt{q})^2}{2x_n\sqrt{q}} < \frac{1}{2}\left(\frac{x_n - \sqrt{q}}{\sqrt{q}}\right)^2,$$

so

$$0 < \frac{x_n - \sqrt{q}}{2\sqrt{q}} < \left(\frac{x_n - \sqrt{q}}{2\sqrt{q}}\right)^2 < \cdots < \left(\frac{x_0 - \sqrt{q}}{2\sqrt{q}}\right)^{2^n}$$

Returning to a more general case, if $|f''| \leq M$ and $|f'| \geq m$ then

$$|A_n| \leq \frac{M}{m}\int_0^1 (1-t)\, dt = \frac{M}{2m}.$$

Setting $C = \frac{M}{2m}$ yields

$$C|x_n - x^*| \leq (C(x_{n-1} - x^*))^2 \leq \cdots \leq (C(x_0 - x^*))^{2^n}.$$

Thus assuming we are able to start with a good enough initial approximation that $C(x_0 - x^*)$ is small, the convergence is extremely rapid.

EXAMPLE 2.5. Now let us generalize Example 2.4 by putting $f(x) = x^\alpha - q$, where $\alpha \neq 0$. Then

$$x_{n+1} = \frac{1}{\alpha}\Big((\alpha - 1)x_n + qx_n^{1-\alpha}\Big).$$

A nice case is when $\alpha = -k$ for some positive integer k. Then x_{n+1} is obtained from x_n using only addition and multiplication. This is interesting even for $\alpha = -1$, since it allows us to do division to arbitrary accuracy without actually dividing. For example, with $q = 3.14$ and $x_1 = .3$, we have

$$x_2 = \frac{1}{-1}\Big((-1 - 1)(.3) + (3.14)(.3)^2\Big) = .3174,$$

while $1/3.14 = 0.31847\ldots$

EXERCISE 2.6. Simplify (2.3) when $f(x) = x^2 - x - 1$. Starting with $x_0 = 1$, find a formula for x_n in terms of the terms of the sequence

$$\frac{1}{1}, \frac{2}{1}, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \ldots$$

from Example 5.18 of the chapter on Sequences.[1] What is $\lim x_n$?

EXERCISE 2.7. Let $f\colon [a, b] \to \mathbb{R}$ be continuous with $f(a) < 0 < f(b)$. Suppose f is twice differentiable on (a, b), and $f' > 0$ and $f'' \geq 0$ on (a, b). Let x^* be the unique point in (a, b) such that $f(x^*) = 0$. Let x_0 be a point in (x^*, b) and let

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n = 0, 1, 2, \ldots.$$

Prove that $\lim x_n = x^*$. Note that under these conditions we get convergence of Newton's method without needing $C(x_0 - x^*)$ to be small.

3. **Stirling's formula for the factorial.** To apply the above to the factorial function, we use the integral expression

$$n! = \int_0^\infty x^n e^{-x}\, dx = \lim_{R\to\infty} \int_0^R x^n e^{-x}\, dx. \tag{3.1}$$

We justify (3.1) inductively, starting with $n = 0$:

$$\int_0^R e^{-x}\, dx = 1 - e^{-R} \to 1,$$

and, integrating by parts,

$$\int_0^R x^{n+1} e^{-x}\, dx = -x^{n+1} e^{-x}\Big|_0^R + (n + 1)\int_0^R x^n e^{-x}\, dx \to (n + 1)\int_0^\infty x^n e^{-x}\, dx.$$

---

[1]It may be easiest to write out terms explicitly until you are able to anticipate the formula, then prove it by induction.