COMPSCI 230: Discrete Mathematics for Computer Science	April 3, 2019

Lecture 21

Lecturer: Debmalya Panigrahi — *Scribe: Erin Taylor*

1 Overview

In this lecture, we will study some common sums and products that have closed form expressions. In some cases, a closed form for a sum or product may not exist, and so we will study a general method to approximate sums and products.

2 Sequences and Series

Sums and products arise frequently in computer science, and often we would like to evaluate them. Expressions that do not make use of summations or products are called closed forms. Typically, closed forms are easier to evaluate. We begin by introducing several common series that have closed form representations.

2.1 Arithmetic Progressions

We will start by considering a familiar sum.

Example 1: Recall that we proved the identity $\sum_{i=1}^{n} i = n(n+1)/2$ by induction. Here is another simple proof.

$$\begin{array}{rcccc} S = 1+ & 2+\cdots+ & (n-1)+n \\ S = n+ & (n-1)+\cdots+ & 2+1 \\ \hline 2S = (n+1)+ & (n+1)+\cdots+ & (n+1)+(n+1) \end{array}$$

$$\Rightarrow 2S = (n+1)n \Rightarrow S = \frac{n(n+1)}{2}.$$

This is an example of an arithmetic progression. More generally, an arithmetic progression (or arithmetic sequence) is a sequence of numbers such that the difference between consecutive terms is constant. The arithmetic progression in the first example is the first n natural numbers, $1, 2, 3, \ldots, n$; the constant difference between terms is 1.

More generally, the first term of an arithmetic progression could be any number a. We some constant x to obtain each consecutive term. Let

$$A = (a,\ a+x,\ a+2x, \ldots,\ b),$$

where b is the last term in the sequence. If there are n terms in the sequence, then $b = a + (n-1)x$. Then, we can evaluate the sum as follows:

$$\begin{array}{rcc} S_A = a+ & (a+x)+\cdots+ & (b-x)+b \\ S_A = b+ & (b-x)+\cdots+ & (a+x)+a \\ \hline \end{array}$$

$$\Rightarrow 2S = (a+b)n \Rightarrow S = \frac{(a+b)n}{2}.$$

Since $b = a + (n-1)x$, we can also write the above sum as

$$S = \frac{(2a + (n-1)x)n}{2} = (a + (n-1)x/2)n.$$

2.2 Geometric Progressions

A geometric progression, or geometric sequence, is a sequence of numbers in which each consecutive term is found by multiplying the previous by a non-zero constant.

$$G = (a, ax, ax^2, \ldots, ax^n) \text{ for some } x \neq 0.$$

We will see a closed-form expression for the sum of a geometric progression:

$$S_G = a + ax + \cdots + ax^n.$$

Lemma 1. *If $x \neq 1$,*

$$S_G = a\Big(\frac{1 - x^{n+1}}{1 - x}\Big).$$

Proof. We will prove this identity by induction on n.
Base Case: When $n = 0$, $S_G = a = a\frac{(1-x)}{(1-x)}$.
Inductive Hypothesis: Let n be an arbitrary positive integer. Assume

$$S_G = a + ax + \cdots + ax^{n-1} = a\Big(\frac{1-x^n}{1-x}\Big).$$

Inductive Case:

$$\begin{aligned}
&a + ax + \cdots + ax^{n-1} + ax^n \\
&= a\Big(\frac{1-x^n}{1-x}\Big) + ax^n && \text{applying the inductive hypothesis to the first } n \text{ terms} \\
&= a\Big(\frac{1-x^n}{1-x} + x^n\Big) \\
&= a\Big(\frac{1-x^n + (1-x)x^n}{1-x}\Big) \\
&= a\Big(\frac{1-x^n + x^n - x^{n+1}}{1-x}\Big) = a\Big(\frac{1-x^{n+1}}{1-x}\Big). \quad \square
\end{aligned}$$

What about an infinite geometric series? Such a series may not always converge. For example, if $x > 1$, S_G does not converge. If the sequence is infinite, then we can write the sum as:

$$S = \lim_{n\to\infty} S_G$$

If $x \geq 1$, the limit tends to infinity. When $x < 1$:

$$S = \lim_{n\to\infty} \frac{x^{n+1} - 1}{x - 1} = \frac{1}{1-x}.$$

2.3 Arithmetic-Geometric Progressions

An arithmetic-geometric progression (AGP) is a progression in which each term can be represented as the product of the terms of an arithmetic progression and a geometric progression. Consider the series

$$S = 1 + 2x + 3x^2 + \dots nx^{n-1}.$$

This looks a lot like the previous geometric series we considered. We will try to directly solve for a closed-form representation of S. Consider xS:

$$xS = x + 2x^2 + \cdots + (n-1)x^{n-1} + nx^n.$$

Notice that we could subtract off xS from the sum to obtain a geometric series.

$$S - xS = 1 + x + x^2 + \cdots + x^{n-1} - nx^n$$

Assuming $x \neq 1$, we now plug in the closed-form expression for the geometric sum from Lemma 1.

$$S = \left[\frac{1-x^n}{1-x} - nx^n\right]\left(\frac{1}{1-x}\right) = \frac{1-x^n-(1-x)nx^n}{(1-x)^2} = \frac{1-(n+1)x^n+nx^{n+1}}{(1-x)^2}.$$

If $x < 1$, the sum of the corresponding infinite AGP is $1/(1-x)^2$.

3 Approximation by integrals

3.1 Approximating Sums

We cannot always find a closed-form expression for a sum. For instance, $S = \sum_{i=1}^{n} \sqrt{i}$ has no closed form. We may still be interested in evaluating such sums, in which case we can find an approximation. We will see a general technique for finding upper and lower bounds that provide good approximations for many sums. We will introduce the technique using an example.

Example 2:

$$S = \sum_{i=1}^{n} \sqrt{i}$$

Consider the following plot in Figure 1.

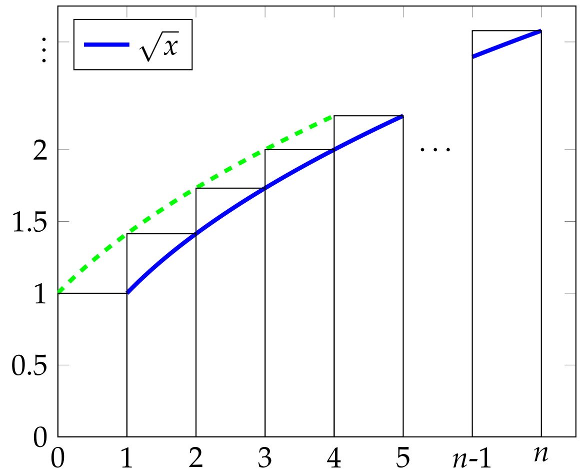

Figure 1: Area of rectangles corresponding to $\sum_{i=1}^{n} \sqrt{i}$. $f(x) = \sqrt{x}$ in blue, the same curve shifted to the left by 1 in green.

The area of each rectangle is the i^{th} term of the sum. In Figure 1, the first rectangle has area 1, the second $\sqrt{2}$, and so on. Approximating the area of these n rectangles is equivalent to approximating the sum. We draw the function $f(x) = \sqrt{x}$ in blue. We know we can calculate the area under the curve $f(x)$ using an integral:

$$\int_{x=1}^{n} \sqrt{x}\,\mathrm{d}x = \frac{2}{3}x^{3/2}\Big|_1^n = \frac{2}{3}\left(n^{3/2} - 1\right).$$

$f(x)$ is a lower bound for the area of the last $n-1$ rectangles. Notice that we integrated from 1 to n, so we need to add in the area of the first rectangle. After this, we obtain the following lower bound:

$$\sum_{i=1}^{n} \sqrt{i} \geq \frac{2}{3}n^{3/2} - \frac{2}{3} + 1 = \frac{2}{3}n^{3/2} + \frac{1}{3}.$$

Now, we are also interested in obtaining an upper bound for the sum. We could use the same integral to upper bound the area of the rectangles by shifting the curve to the left (see the dashed green curve in Figure 1). When we shift the curve, we now have an upper bound for the first $n-1$ rectangles, so we will add in the area for the n^{th} rectangle separately. After this, we obtain the following upper bound:

$$\sum_{i=1}^{n} \sqrt{i} \leq \frac{2}{3}n^{3/2} - \frac{2}{3} + \sqrt{n}.$$

Thus, the sum is $\Theta(n^{3/2})$, whereas the difference in the upper and lower bounds is $\Theta(n^{1/2})$. Thus, this approximation is quite tight in that its error is additive and of lower order than the sum itself.

For increasing functions, we can compute the integral, and adding in the first rectangle will give a lower bound.

Suppose $S = \sum_{i=1}^{n} f(i)$.
If f is non-decreasing then,

$$f(1) + \int_1^n f(x)\,\mathrm{d}x \le S \le \int_1^n f(x)\,\mathrm{d}x + f(n).$$

Note that $f(1)$ corresponds to the first term of the sum (the area of the first rectangle), and $f(n)$ corresponds to the last term of the sum (the area of the last rectangle).
If f is non-increasing then,

$$\int_1^n f(x)\,\mathrm{d}x + f(n) \le S \le f(1) + \int_1^n f(x)\,\mathrm{d}x\,.$$

Example 3: Next, consider the sum

$$H_n = \sum_{i=1}^{n} \frac{1}{i}.$$

The sum is denoted H_n, and it is known as the n^{th} Harmonic number. These numbers have been extensively studied in number theory. Asymptotically, the Harmonic numbers roughly approximate the natural logarithm. Consider Figure 2.

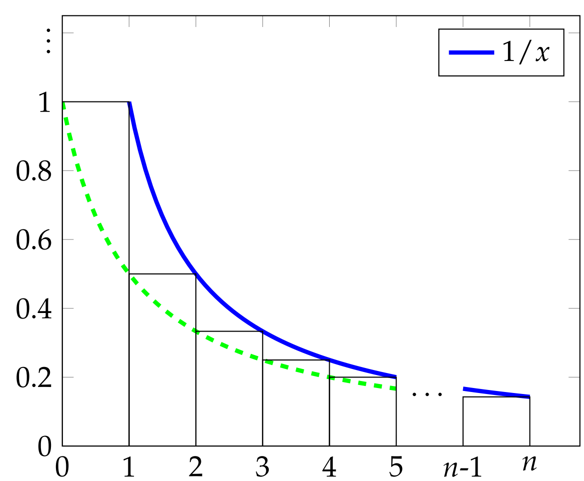

Figure 2: Area of rectangles corresponding to $\sum_{i=1}^{n} \frac{1}{i}$. $f(x) = \frac{1}{x}$ in blue, the same curve shifted to the left by 1 in green.

This function is decreasing, so for our upper bound we will add in the first rectangle, and the lower bound will add in the last. First computing the integral:

$$\int_{x=1}^{n} \frac{\mathrm{d}x}{x} = \ln(x)\Big|_1^n = \ln(n).$$

Then,

$$H_n \le \ln(n) + 1$$
$$H_n \ge \ln(n) + \frac{1}{n}$$

This shows that the n^{th} harmonic number is very close to $\ln(n)$.

3.2 Approximating Products

We will show that we can also use integrals to approximate products.

Example 4: Consider $n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n$. Naively, we can show that $n! = n^{O(n)}$ in the following way:

$$n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n \le \underbrace{n \cdot n \cdot n \cdot \cdots \cdot n}_{\text{n terms}} = n^n$$

and

$$\frac{n^{n/2}}{2^{n/2}} = \underbrace{\frac{n}{2} \cdot \frac{n}{2} \cdot \frac{n}{2} \cdot \cdots \cdot \frac{n}{2}}_{\text{n/2 terms}} \le \underbrace{\frac{n}{2} \cdot \left(\frac{n}{2} + 1\right) \cdot \left(\frac{n}{2} + 2\right) \cdot \cdots \cdot n}_{\text{n/2 terms}} \le 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n-1) \cdot n = n!.$$

Thus, we have the following upper and lower bounds:

$$\frac{n^{n/2}}{2^{n/2}} \le n! \le n^n.$$

However, we might want a tighter bound, which we can achieve using our method for approximating sums. We can take a logarithm to convert our product into a sum. Recall the identity $\log(a \cdot b) = \log(a) + \log(b)$. Applying this identity repeatedly we can write:

$$\ln(n!) = \sum_{i=1}^{n} \ln(i).$$

Now, we can apply our earlier method. Consider the plot in Figure 3.

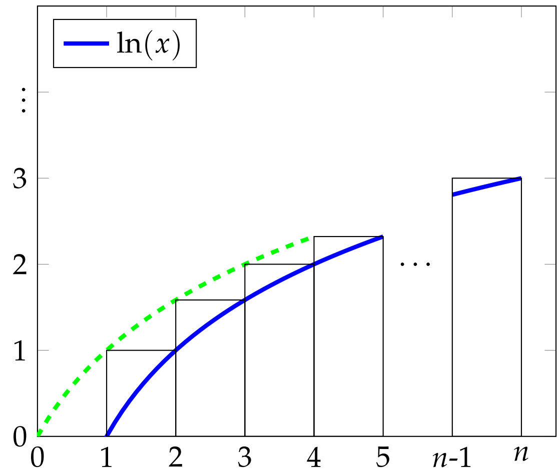

Figure 3: Area of rectangles corresponding to $\sum_{i=1}^{n} \ln i$. $f(x) = \ln x$ in blue, the same curve shifted to the left by 1 in green..

Recall integration by parts, $\int u \, dv = uv - \int v \, du$. Solving the integral:

$$\int_{x=1}^{n} \ln(x) \, dx = x \ln(x) \Big|_{1}^{n} - \int_{x=1}^{n} \frac{1}{x} x \, dx = n \ln(n) - n + 1.$$

We obtain the following upper and lower bounds by adding in the correct rectangle.

$$\ln(n!) \leq n \ln(n) - n + 1 + \ln(n)$$
$$\ln(n!) \geq n \ln(n) - n + 1 + 0$$

Now, to get an approximation of $n!$ instead of $\ln(n!)$, we simply exponentiate both sides.

$$n! \leq e^{n \ln(n) - n + 1 + \ln(n)} = \frac{n^{n+1}}{e^{n-1}}$$
$$n! \geq e^{n \ln(n) - n + 1} = \frac{n^{n}}{e^{n-1}}$$

Thus, we have the following approximation:

$$\frac{n^n}{e^{n-1}} \leq n! \leq \frac{n^{n+1}}{e^{n-1}}.$$

4 Summary

In this lecture, we saw arithmetic and geometric progressions. We also saw a general technique to approximate summations using integrals, and showed that this could be extended to approximate products.