Sums, Approximations, and Asymptotics II

1 Block Stacking

How far can a stack of identical blocks overhang the end of a table without toppling over? Can a block be suspended beyond the table’s edge?

The stack falls off the desk if its center of mass lies beyond the desk’s edge. Moreover, the center of mass of the top $k$ blocks must lie above the $(k+1)$-st block; otherwise, the top $k$ fall off. In order to find the best configuration of blocks, we’ll need a fact from physics about centers of mass.

**Fact 1.** If two objects have masses $m_1$ and $m_2$ and centers-of-mass at positions $x_1$ and $x_2$, then the center of mass of the two objects together is at position:

$$\frac{z_1m_1 + z_2m_2}{m_1 + m_2}$$

For this problem, only the horizontal dimension is relevant, and we’ll use the width of a block as our measure of distance. Define the **offset** of a particular configuration of blocks to be the horizontal distance from its center of mass to its rightmost edge. The offset measures how far the configuration can extend beyond the desk since at best the center of mass lies right at the desk’s edge.

---

If you suspect you’ve found an error in these notes or you just find a part hopelessly confusing, please send email to e_lehman@mit.edu, and I’ll try to fix the problem.
We can find the greatest possible offset of a stack of \( n \) blocks with an inductive argument. This is an instance where induction not only serves as a proof technique, but also turns out to be a great tool for reasoning about the problem.

**Theorem 1.** The greatest possible offset of a stack of \( n \geq 1 \) blocks is:

\[
X_n = \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \ldots + \frac{1}{2n}
\]

**Proof.** We use induction on \( n \), the number of blocks. Let \( P(n) \) be the proposition that the greatest possible offset of a stack of \( n \geq 1 \) blocks is \( 1/2 + 1/4 + \ldots + 1/(2n) \).

**Base case:** For a single block, the center of mass is distance \( X_1 = 1/2 \) from its rightmost edge. So \( P(1) \) is true.

**Inductive step:** For \( n \geq 2 \), assume that \( P(n - 1) \) is true in order to prove \( P(n) \). A stack of \( n \) blocks consists of the bottom block together with a stack of \( n - 1 \) blocks on top.

In order to achieve the greatest possible offset with \( n \) blocks, the top \( n - 1 \) blocks should have the greatest possible offset, which is \( X_{n-1} \); otherwise, we could do better by replacing the top \( n - 1 \) blocks with a different configuration that has greater offset. Furthermore, the center of mass of the top \( n - 1 \) blocks should lie directly above the right edge of the bottom block; otherwise, we could do better by sliding the top \( n - 1 \) blocks farther to the right.

---

1A different analysis was presented in lecture.
Thus, by the physics fact, the maximum possible offset of a stack of \( n \) blocks is:

\[
X_n = \frac{X_{n-1} \cdot (n-1) + (X_{n-1} + \frac{1}{2}) \cdot 1}{n} \\
= X_{n-1} + \frac{1}{2n} \\
= \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \ldots + \frac{1}{2n}
\]

We use the assumption \( P(n - 1) \) in the last step. This proves \( P(n) \).

The theorem follows by the principle of induction. \( \square \)

1.1 Harmonic Numbers

Sums similar to the one in Theorem ?? come up all the time in computer science. In particular,

\[
\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}
\]

is called a harmonic sum. Its value is called the \( n \)-th harmonic number and is denoted \( H_n \).

In these terms, the greatest possible offset of a stack of \( n \) blocks is \( \frac{1}{2}H_n \). We can tabulate the greatest overhang achievable with \( n = 1, 2, 3 \) and 4 blocks by computing harmonic numbers:

<table>
<thead>
<tr>
<th># of blocks</th>
<th>maximum overhang</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \frac{1}{2}H_1 = \frac{1}{2}(\frac{1}{1}) = \frac{1}{2} )</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{2}H_2 = \frac{1}{2}(\frac{1}{1} + \frac{1}{2}) = \frac{3}{4} )</td>
</tr>
<tr>
<td>3</td>
<td>( \frac{1}{2}H_3 = \frac{1}{2}(\frac{1}{1} + \frac{1}{2} + \frac{1}{3}) = \frac{11}{12} )</td>
</tr>
<tr>
<td>4</td>
<td>( \frac{1}{2}H_4 = \frac{1}{2}(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}) = \frac{25}{24} &gt; 1 )</td>
</tr>
</tbody>
</table>

The last line reveals that we can suspend the fourth block beyond the edge of the table! Here’s the configuration that does the trick:
We’ll have to study harmonic sums more closely to determine what can be done with large numbers of blocks. Let’s use integration to bound the $n$-th harmonic number. A picture is extremely helpful for getting the right functions and limits of integration.

We can not integrate the function $1/x$ starting from zero. So, for the upper bound on $H_n$ we’ll take the first term explicitly ($1/1$) and then upper bound the remaining terms normally. This gives:

\[
\int_0^n \frac{1}{x+1} \, dx \leq H_n \leq 1 + \int_1^n \frac{1}{x} \, dx \\
\ln(x+1) \bigg|_0^n \leq H_n \leq 1 + \left( \ln x \bigg|_1^n \right) \\
\ln(n+1) \leq H_n \leq 1 + \ln n
\]

There are good bounds; the difference between the upper and lower values is never more than 1.

Suppose we had a million blocks. Then the overhang we could achieve—assuming no breeze or deformation of the blocks—would be $\frac{1}{2} H_{1000000}$. According to our bounds, this is:

at least $\frac{1}{2} \ln(1000001) = 6.907\ldots$

at most $\frac{1}{2} (1 + \ln(1000000)) = 7.407\ldots$

So the top block would extend about 7 lengths past the end of the table! In fact, since the lower bound or $\frac{1}{2} \ln(n+1)$ grows arbitrarily large, there is no limit on how far the stack can overhang!
Mathematicians have worked out some extremely precise approximations for the \( n \)-th harmonic number. For example:

\[
H_n = \ln(n) + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{\epsilon(n)}{120n^4}
\]

where \( \gamma = 0.577215664 \ldots \) is Euler’s constant and \( \epsilon(n) \) is between 0 and 1. Interestingly, no one knows whether Euler’s constant is rational or irrational.

## 2 Products

We’ve now looked at many techniques for coping with sums, but no methods for dealing with products. Fortunately, we don’t need to develop an entirely new set of tools. Instead, we can first convert any product into a sum by taking a logarithm:

\[
\ln \left( \prod f(n) \right) = \sum \ln f(n)
\]

Then we can apply our summing tools and exponentiate at the end to undo the logarithm.

Let’s apply this strategy to a product that you’ll encounter almost daily hereafter:

\[ n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n - 1) \cdot n \]

First, we take a logarithm:

\[
\ln n! = \ln 1 + \ln 2 + \ln 3 + \ldots + \ln n
\]

This sum is rather nasty, but we can still get bounds by integrating. First, we work out appropriate functions and limits of integration with a picture.

Now we integrate to get bounds on \( \ln n! \).

\[
\int_1^n \ln x \, dx \leq \ln n! \leq \int_0^n \ln(x + 1) \, dx
\]

\[
x \ln x - x \bigg|_1^n \leq \ln n! \leq (x + 1) \ln(x + 1) - (x + 1) \bigg|_1^n
\]

\[
n \ln n - n \leq \ln n! \leq (n + 1) \ln(n + 1) - n
\]
Finally, we exponentiate to get bounds on $n!$.

$$\frac{n^n}{e^n} \leq n! \leq \frac{(n + 1)^{(n+1)}}{e^n}$$

This gives some indication how big $n!$ is: about $(n/e)^n$. This estimate is often good enough. However, as with the harmonic numbers, more precise bounds are known.

**Fact 2.**

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{1/(12n+1)} \leq n! \leq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{1/(12n)}$$

These bounds are ridiculously close. For example, if $n = 100$, then we get:

$$100! \geq \left(\frac{100}{e}\right)^{100} \sqrt{200\pi} e^{1/1201} = 1.000832...$$

$$100! \leq \left(\frac{100}{e}\right)^{100} \sqrt{200\pi} e^{1/1200} = 1.000833...$$

The upper bound is less than 7 hundred-thousandths of 1% greater than the lower bound!

Taken together, these bounds imply **Stirling’s Formula**:

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

Stirling’s formula is worth committing to memory; we’ll often use it to rewrite expressions involving $n!$. Now, one might object that the expression on the left actually looks a lot better than the expression on the right. But that’s just an artifact of notation. If you actually wanted to compute $n!$, you’d need $n - 1$ multiplications. However, the expression on the right is a closed form; evaluating it requires only a handful of basic operations, regardless of the value of $n$. Furthermore, when $n!$ appears inside a larger expression, you usually can’t do much with it. It doesn’t readily cancel or combine with other terms. In contrast, the expression on the right looks scary, but melds nicely into larger formulas. So don’t be put off; the expression on the right is your true friend.

Stepping back a bit, Stirling’s formula is fairly amazing. Who would guess that the product of the first $n$ positive integers could be so precisely described by a formula involving both $e$ and $\pi$?

### 3 Asymptotic Notation

Approximation is a powerful tool. It lets you sweep aside irrelevant detail without losing track of the big picture.

Approximations are particularly useful in the analysis of computer systems and algorithms. For example, suppose you wanted to know how long multiplying two $n \times n$
matrices takes. You could tally up all the multiplications and additions and loop variable increments and comparisons and perhaps factor in hardware-specific considerations such as page faults and cache misses and branch mispredicts and floating-point unit availability and all this would give you one sort of answer.

On the other hand, each of the $n^2$ entries in the product matrix takes about $n$ steps to compute. So the running time is proportional to $n^3$. This answer is certainly less precise. However, high-precision answers are rarely needed in practice. And this approximate answer is independent of tiny implementation and hardware details; it remains valid even after you upgrade your computer.

Computer scientists make heavy use of a specialized asymptotic notation to describe the growth of functions approximately. The notation involves six weird little symbols. We’ve already encountered one:

$$f(n) \sim g(n) \quad \text{means} \quad \lim_{n \to \infty} \frac{f(n)}{g(n)} = 1$$

Less formally, $f \sim g$ means that the functions $f$ and $g$ grow at essentially the same rate. We’ve already derived two important examples:

$$H_n \sim \ln n \quad \quad n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

Here are the other five symbols in the asymptotic notation system:

\[ O \quad \Omega \quad \Theta \quad o \quad \omega \]

oh omega theta little-oh little-omega

We’ll focus on the most important one, $O$. Here’s the definition: given functions $f, g : \mathbb{R} \to \mathbb{R}$, we say that

$$f(x) = O(g(x))$$

if there exist constants $x_0$ and $c > 0$ such that $|f(x)| \leq c \cdot g(x)$ for all $x \geq x_0$. Now this definition is pretty hairy. But what it’s trying to say, with all its cryptic little constants, is that $f$ grows no faster than $g$. A bit more precisely, it says that $f$ is at most a constant times greater than $g$, except maybe for small values of $x$. For example:

$$5x + 100 = O(x)$$

This holds because the left side is only about 5 times larger than the right. Of course, for small values of $x$ (like $x = 1$) the left side is many times larger than the right, but the definition of $O$ is cleverly designed to sweep such inconvenient details under the rug.

Let’s work through a sequence of examples carefully to better understand the definition.

Claim 2. $5x + 100 = O(x)$
Proof. We must show that there exist constants $x_0$ and $c > 0$ such that $|5x + 100| \leq c \cdot x$ for all $x \geq x_0$. Let $c = 10$ and $x_0 = 20$ and note that:

\[|5x + 100| \leq 5x + 5x = 10x \quad \text{for all } x \geq 20\]

\[\square\]

**Claim 3.** $x = O(x^2)$

Proof. We must show that there exist constants $x_0$ and $c > 0$ such that $|x| \leq c \cdot x^2$ for all $x \geq x_0$. Let $c = 1$ and $x_0 = 1$ and note that

\[|x| \leq 1 \cdot x^2 \quad \text{for all } x \geq 1\]

\[\square\]

What about the reverse? Is $x^2 = O(x)$? On an informal basis, this means $x^2$ grows no faster than $x$, which is false. Let’s prove this formally.

**Claim 4.** $x^2 \neq O(x)$

Proof. We argue by contradiction; suppose that there exist constants $x_0$ and $c$ such that:

\[|x^2| \leq c \cdot x \quad \text{for all } x \geq x_0\]

Dividing both sides of the inequality by $x$ gives:

\[x \leq c \quad \text{for all } x \geq x_0\]

But this is false when $x = \max(x_0, c + 1)$.

\[\square\]

We can show that $x^2 \neq O(100x)$ by essentially the same argument; intuitively, $x^2$ grows quadratically, while $100x$ grows only linearly. Generally, changes in multiplicative constants do not affect the validity of an assertion involving $O$. However, constants in exponentials are critical:

**Claim 5.**

\[4^x \neq O(2^x)\]

Proof. We argue by contradiction; suppose that there exist constants $x_0$ and $c > 0$ such that:

\[|4^x| \leq c \cdot 2^x \quad \text{for all } x \geq x_0\]

Dividing both sides by $2^x$ gives:

\[2^x \leq c \quad \text{for all } x \geq x_0\]

But this is false when $x = \max(x_0, 1 + \log c)$.

\[\square\]
While asymptotic notation is useful for sweeping aside irrelevant detail, it can be abused. For example, there are ingenious algorithms for multiplying matrices that are asymptotically faster than the naive algorithm. The “best” requires only $O(n^{2.376})$ steps. However interesting theoretically, these algorithms are useless in practice because the constants hidden by the $O$ notation are gigantic!

*For more information about asymptotic notation see the “Asymptotic Cheat Sheet”.**