STIRLING'S APPROXIMATION

NICHOLAS F. MARSHALL

1. Introduction

1.1. Stirling's approximation. Stirling's approximation states that

$$n! \sim \sqrt{2\pi} n^{n+1/2} e^{-n},$$

where $f(n) \sim g(n)$ means that $\lim_{n\to\infty} f(n)/g(n) = 1$. In this note, we give an elementary proof that this approximation is always correct up to a factor between 0.9 and 1.1.

Proposition 1.1. *We have*

$$(0.9)\sqrt{2\pi} n^{n+1/2} e^{-n} \leq n! \leq (1.1)\sqrt{2\pi} n^{n+1/2} e^{-n},$$

for all $n \geq 1$.

The proof involves some basic facts about concave functions, trapezoid rule, and midpoint rule, which are stated and proved in §2. In particular, see Lemmas 2.1 and 2.2.

Proof of Proposition 1.1. First we establish the upper bound. We have

$$\ln(n!) = \left(\sum_{k=1}^{n} \ln(k) - \frac{1}{2}\ln(n) \right) + \frac{1}{2}\ln(n) \leq \int_1^n \ln(x)dx + \frac{1}{2}\ln(n),$$

where the final inequality follows from the fact that the term in the parentheses is trapezoid rule for the integral $\int_1^n \ln(x)dx$ whose integrand $\ln(x)$ is concave, and that trapezoid rule is a lower bound for the integral of concave functions by Lemma 2.1. Exponentiating both sides and using the relation $(x\ln x - x)' = \ln x$ gives

$$n! \leq e n^{n+1/2} e^{-n} \leq (1.1)\sqrt{2\pi} n^{n+1/2} e^{-n},$$

where the final inequality follows because $e \approx 2.718$ and $1.1\sqrt{2\pi} \approx 2.757$. It remains to establish the lower bound. We have

$$\ln(n!) = \sum_{k=2}^{n} \ln(k) \geq \int_{3/2}^{n+1/2} \ln(x)dx,$$

where the inequality follows from the fact that the sum is midpoint rule for the integral $\int_{3/2}^{n+1/2} \ln(x)dx$ whose integrand $\ln(x)$ is concave, and that midpoint rule is an upper bound for the integral of concave functions by Lemma 2.2. Exponentiating both sides of the equation and using the inequality $\int_{3/2}^{n+1/2} \ln(x)dx \geq \int_{3/2}^{n} \ln(x)dx + \frac{1}{2}\ln(n)$ (which holds since $\ln(x)$ is increasing) gives

$$n! \geq (3/2)^{-3/2} e^{3/2} n^{n+1/2} e^{-n} \geq 0.9\sqrt{2\pi} n^{n+1/2} e^{-n},$$

where the final inequality follows because $(3/2)^{-3/2} e^{3/2} \approx 2.4395$ and $0.9\sqrt{2\pi} \approx 2.25597$. This completes the proof. $\square$

2. Background

2.1. Concave functions. A function $f : [a, b] \to \mathbb{R}$ is concave if

$$f((1-\alpha)x + \alpha y) \geq (1-\alpha)f(x) + \alpha f(y), \tag{1}$$

for all $a \leq x < y \leq b$ and $0 < \alpha < 1$. If we set $w := (1-\alpha)x + \alpha y$, then this inequality is equivalent to the statement that

$$f(w) \geq \frac{y-w}{y-x} f(x) + \frac{w-x}{y-x} f(y) := l(w), \tag{2}$$

for $a \leq x < w < y \leq b$. Note that $l(w)$ is the line passing through $(x, f(x))$ and $(y, f(y))$ so this inequality, informally speaking, says that line segments between points on the graph of f are below the graph of f. Alternatively, we can write (2) in the equivalent form

$$\frac{f(w) - f(x)}{w - x} \geq \frac{f(y) - f(w)}{y - w},$$

for all $a \leq x < w < y \leq b$. This inequality implies that a function is concave if its average change on $[x, w]$ is at least its average change on $[w, y]$ whenever $x < w < y$. If we assume that f is differentiable, then considering this inequality in the limit as $y \to w$ and the limit as $x \to w$ separately gives

$$\frac{f(w) - f(x)}{w - x} \geq f'(w) \geq \frac{f(y) - f(w)}{y - w} \tag{3}$$

for all $x < w < y$. It follows that differentiable concave functions have nonincreasing derivatives. Moreover, if f is differentiable and its derivative is nonincreasing, then (3) follows from the mean-value theorem. Finally, observe that the pair of inequalities in (3) are equivalent to the inequality:

$$f'(w)(t - w) + f(w) \geq f(t), \tag{4}$$

for all $a \leq t \leq b$, which informally speaking says that tangent lines to f are always above f. Indeed, the equivalence between (3) and (4) can be established by considering the cases $t > w$ and $t \leq w$ separately. The key ideas from the above discussion can be summarized as follows.

Proposition 2.1. *Suppose that f is a differentiable real-valued function on $[a, b]$. The following are equivalent:*

- *The function f is concave in the sense of* (1)
- *The derivative of f is nonincreasing*
- *Line segments connecting points on f are below f; more precisely* (2) *holds*
- *Lines tangent to f are above f; more precisely* (4) *holds*

2.2. Trapezoid rule and midpoint rule. Suppose that $f : [a, b] \to \mathbb{R}$, and set $h := (b - a)/n$ for a fixed integer $n > 0$. Trapezoid rule $T(h)$ for the integral of f on $[a, b]$ is defined by

$$T(h) = \left(\sum_{k=0}^{n} f(a + kh)h \right) - \frac{1}{2} \left(f(a) + f(b) \right) h, \tag{5}$$

and midpoint rule $M(h)$ for the integral of f on $[a, b]$ is defined by

$$M(h) = \sum_{k=1}^{n} f(a + kh - h/2)h.$$

Both of these rules have geometric interpretations that we describe in the following.

2.3. **Geometric interpretations.** Trapezoid rule can be rewritten as

$$T(h) = \sum_{k=1}^{n} A_k, \quad \text{for} \quad A_k := \frac{f(x_{k-1}) + f(x_k)}{2} h,$$

where $x_k = a + kh$. The quantity A_k can be interpreted as the area of a trapezoid with bases $x_{k-1} \times [0, f(x_{k-1})]$ and $x_k \times [0, f(x_k)]$, see Figure 1.

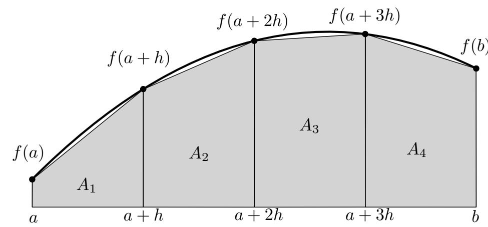

FIGURE 1. Trapezoid rule is the sum of signed areas of trapezoids between line segments connecting points on the graph of f and the x-axis.

In particular, the top side of each trapezoid is a line segment connecting points on the graph of f, so when f is concave the following result follows from the third equivalence in Proposition 2.1.

Lemma 2.1. *If f is concave on $[a, b]$, then trapezoid rule $T(h) \le \int_a^b f(x)dx$.*

The formula for midpoint rule has a similar useful geometric interpretation in terms of trapezoids, which can be observed by rewriting $M(h)$ as

$$M(h) = \sum_{k=1}^{n} B_k, \quad \text{for} \quad B_k := \frac{f(m_k) - \frac{h}{2}f'(m_k) + f(m_k) + \frac{h}{2}f'(m_k)}{2} h$$

where $m_k = a + kh - h/2$. The quantity B_k can be interpreted as the area of the trapezoid with bases $x_{k-1} \times [0, f(m_k) - \frac{h}{2}f'(m_k)]$ and $x_k \times [0, f(m_k) + \frac{h}{2}f'(m_k)]$, see Figure 2.

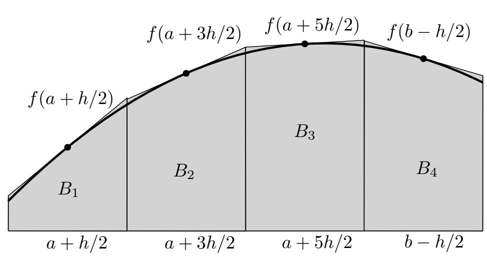

FIGURE 2. Midpoint rule is the sum of signed areas of trapezoids between line segments tangent to the graph of f and the x-axis.

The top side of each trapezoid is a segment on a line tangent to f, and thus, the following lemma is an immediate consequence of this geometric interpretation of midpoint rule and the fourth equivalence in Proposition 2.1.

Lemma 2.2. *If f is concave on $[a, b]$, then midpoint rule $M(h) \ge \int_a^b f(x)dx$.*

2.4. **Numerics.** To get an idea of the actual accuracy of Stirling's approximation we compute

$$r_n := \frac{n!}{\sqrt{2\pi} n^{n+1/2} e^{-n}}.$$

numerically. In particular, we compute r_n for $n = 1, \ldots, 10$ and plot $\log_{10} |r_n - 1|$, see Figure 3.

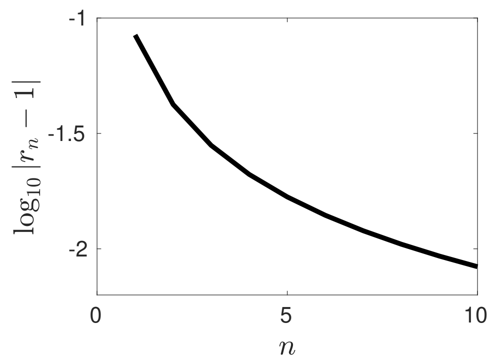

FIGURE 3. A plot of $\log_{10} |r_n - 1|$ for $n = 1, \ldots, 10$.

Observe that the quantity $|r_n - 1|$ starts slightly below 0.1 at $n = 1$ and then decreases to slightly below 0.01 at $n = 10$. In general, the error $|r_n - 1|$ decreases linearly with $1/n$, see for example https://dlmf.nist.gov/5.11.10 which gives an asymptotic series for $r_n - 1$ stated in terms of the Gamma function Γ (recall that $\Gamma(n + 1) = n!$ for positive integers n).