WHEN IS $f(1) + \cdots + f(n)$ ASYMPTOTICALLY OF THE SAME ORDER AS $nf(n)$?

JOHN M. HOLTE
NCS-MAA MEETING AT CARLETON COLLEGE, 17 APRIL 1999

1. Introduction

Two sequences of positive numbers $(f(n))_{n=1}^{\infty}$ and $(g(n))_{n=1}^{\infty}$ are asymptotically of the same order—written $f(n) \asymp g(n)$ or $f(n) = \Theta(g(n))$—if

$$f(n) = O(g(n)) \quad \text{and} \quad g(n) = O(f(n)) \quad \text{as} \quad n \to \infty,$$

i.e., there exist constants c_1 and c_2 and an index n_0 such that

$$0 < c_1 \leq f(n)/g(n) \leq c_2 < \infty \quad \text{for} \quad n \geq n_0.$$

It is easy to show that for positive sequences this is equivalent to the existence of constants C_1 and C_2 such that

$$0 < C_1 \leq f(n)/g(n) \leq C_2 < \infty \quad \text{for } \textit{every } n \in \mathbb{N}.$$

One frequently encounters sequences $(f(n))$ of positive numbers for which the sequence of partial sums $f(1) + \cdots + f(n)$ is asymptotically of the same order as $(nf(n))$, or, in other words, for which the sequence of arithmetic means $(\{f(1) + \cdots + f(n)\}/n)$ is asymptotically of the same order as the original sequence.

For example, in considering algorithms for sorting arbitrary lists of n items, information theory tells us that this task has complexity $\log_2 n!$, and then it follows that

$$\log_2 n! = \sum_{k=1}^{n} \log_2 k = \int_1^n \log_2 x\, dx + O(\ln n) = n \log_2 n - n/\ln 2 + O(\ln n),$$

so that

$$\sum_{k=1}^{n} \log_2 k \asymp n \log_2 n.$$

Polynomial sequences all exhibit this behavior. For example,

$$\sum_{k=1}^{n} k^3 = \frac{n^4}{4} + \text{ lower order terms},$$

so that

$$\sum_{k=1}^{n} k^3 \sim \frac{1}{4} \cdot n \cdot n^3,$$

meaning that the ratio of the two sides has limit 1 as $n \to \infty$, whence

$$\sum_{k=1}^{n} k^3 \asymp n \cdot n^3.$$

Divergent p-series also exhibit this behavior for $p < 1$ (but not for $p = 1$):

$$\sum_{k=1}^{n} k^{-p} \sim \int_1^n x^{-p}\,dx \sim \frac{n^{-p+1}}{-p+1} \asymp n \cdot n^{-p}.$$

Fluctuating sequences $(f(n))$ may also exhibit this behavior. For example, if $f(n) = 2 + \sin n$, then $1 \leq f(n) \leq 3$, so that

$$n \leq \sum_{k=1}^{n} f(k) \leq 3n,$$

whence

$$\frac{1}{3} \leq \frac{\sum_{k=1}^{n} f(k)}{nf(n)} \leq 3,$$

and so $f(1) + \cdots + f(n) \asymp nf(n)$. More generally, if $f(n) = n^{-p}(2 + \sin n)$ where $p < 1$, then we still have $f(1) + \cdots + f(n) \asymp nf(n)$.

On the other hand, convergent series of positive numbers never exhibit this behavior. For if $f(1) + \cdots + f(n) \leq Cnf(n)$ for some $C > 0$, then $f(n) \geq f(1)/Cn$, whence $\sum f(n)$ must diverge.

Also, sequences that increase geometrically do not exhibit this behavior: If $r > 1$, then

$$\sum_{k=1}^{n} r^k = \frac{r^{n+1} - 1}{r - 1},$$

so

$$\frac{\sum_{k=1}^{n} r^k}{nr^n} = \frac{r - r^{-n}}{n(r-1)},$$

and as $n \to \infty$, this approaches 0, so it is not bounded away from 0.

2. Increasing Sequences

Theorem 1. *Assume that f is a positive, nondecreasing sequence. Then $f(1) + \cdots + f(n) \asymp nf(n)$ if and only if the sequence $(f(2n)/f(n))$ is bounded.*

Proof. For a positive, nondecreasing f we always have

$$\frac{f(1) + \cdots + f(n)}{nf(n)} \leq 1, \tag{1}$$

so we need only show that there exists a constant $C > 0$ such that

$$\frac{f(1) + \cdots + f(n)}{nf(n)} \geq C \tag{2}$$

for all $n \geq 1$ if and only if there exists a constant B such that $f(2n)/f(n) \leq B$ for all $n \geq 1$. First assume $f(2n)/f(n) \leq B$ for $n \geq 1$. Write $n = 2m$ or $n = 2m+1$; then

$$\begin{aligned}
\frac{f(1)+\cdots+f(n)}{nf(n)} &\geq \frac{\overbrace{f(m+1)+\cdots+f(n)}^{m \text{ or } m+1 \text{ terms}}}{(2m \text{ or } 2m+1)f(n)} \\
&\geq \frac{(m \text{ or } m+1)f(m+1)}{[2m \text{ or } 2(m+1)]f(2(m+1))} \\
&\geq \frac{1}{2}\cdot\frac{1}{B} =: C > 0.
\end{aligned}$$

For the converse, assume (2). Choose integers p and q such that $p > q \geq 1$ and $1 - q/p < C^2$. Then, using (2) and (1) and monotonicity of f, we get

$$\begin{aligned}
\frac{C}{1} \leq \frac{\frac{f(1)+\cdots+f(pn)}{pnf(pn)}}{\frac{f(1)+\cdots+f(qn)}{qnf(qn)}} &= \frac{qf(qn)}{pf(pn)}\left\{1 + \frac{f(qn+1)+\cdots+f(pn)}{f(1)+\cdots+f(qn)}\right\} \\
&\leq \frac{qf(qn)}{pf(pn)}\left\{1 + \frac{(pn-qn)f(pn)}{Cqnf(qn)}\right\} = \frac{qf(qn)}{pf(pn)} + \frac{p-q}{pC},
\end{aligned}$$

where $(p-q)/pC < C$. Therefore,

$$\frac{qf(qn)}{pf(pn)} \geq C - \frac{p-q}{pC} =: D > 0,$$

so

$$\frac{f(pn)}{f(qn)} \leq \frac{q}{pD} =: E < \infty \quad \text{and} \quad f(pn) \leq Ef(qn).$$

If $f(p^{k-1}n) \leq E^{k-1}f(qn)$ for all n, then

$$f(p^k n) = f(p \cdot p^{k-1}n) \leq Ef(q \cdot p^{k-1}n) = Ef(p^{k-1}qn) \leq E \cdot E^{k-1}f(qn) = E^k f(qn)$$

for all n. By induction, $f(p^k n) \leq E^k f(qn)$ for all k and n. Choose integer $k \geq \log_p 4q$. If $n \geq q$, then $n = mq + r$ with $m \geq 1$ and $0 \leq r < q$, and

$$\frac{f(2n)}{f(n)} = \frac{f(2(mq+r))}{f(mq+r)} \leq \frac{f(2(m+1)q)}{f(mq)} \leq \frac{f(4mq)}{f(mq)} \leq \frac{f(p^k m)}{f(qm)} \leq E^k.$$

Thus, $f(2n)/f(n)$ is bounded by B defined as the maximum of E^k and the values of $f(2n)/f(n)$ for $n = 1, \ldots, q-1$. □

3. Positive Sequences in General

Theorem 2. *Let $f : \mathbb{N} \to \mathbb{R}^+$. The following are equivalent:*

1. $f(1)+\cdots+f(n) \asymp nf(n)$*;*
2. *there exist sequences $(g(n))$ and $(h(n))$ bounded away from 0 and ∞ (or, $g(n) \asymp 1$ and $h(n) \asymp 1$) such that*

$$f(n) = \frac{g(n)}{n}\exp\sum_{k=1}^{n}\frac{h(k)}{k};$$

3. *there exists* $\alpha > -1$ *such that* $n^{-\alpha} f(n)$ *is "almost decreasing," (i.e., there exist* $M > 0$ *and* $n_0 \in \mathbb{N}$ *such that* $n^{-\alpha} f(n) \le M k^{-\alpha} f(k)$ *whenever* $n_0 \le k \le n$*), and there exists* $\beta > -1$ *such that* $n^{-\beta} f(n)$ *is "almost increasing," (i.e., there exist* $m > 0$ *and* $n_0' \in \mathbb{N}$ *such that* $m k^{-\beta} f(k) \le n^{-\beta} f(n)$ *whenever* $n_0' \le k \le n$*).*

The proof is given via the following five lemmas showing that

$$\text{Statement 1} \Rightarrow \text{Statement 2} \Rightarrow \text{Statement 3} \Rightarrow \text{Statement 1}.$$

Let $f : \mathbb{N} \to \mathbb{R}^+$, and define $b : \mathbb{N} \cup \{0\} \to \mathbb{R}^+$ as follows:

$$b(n) = \begin{cases} 1 \text{ if } n = 0; \\ (n+1) f(n+1) / \{f(1) + \cdots + f(n)\} \text{ if } n \in \mathbb{N}. \end{cases} \tag{3}$$

Lemma 1. *For* $n \in \mathbb{N}$,

$$\sum_{k=1}^{n} f(k) = f(1) \exp \sum_{k=1}^{n-1} \ln \left\{ 1 + \frac{b(k)}{k+1} \right\}$$

and

$$f(n) = f(1) \frac{b(n-1)}{n} \exp \sum_{k=1}^{n-2} \ln \left\{ 1 + \frac{b(k)}{k+1} \right\}. \tag{4}$$

Proof.

$$1 + \frac{b(k)}{k+1} = 1 + \frac{f(k+1)}{f(1) + \cdots + f(k)} = \frac{f(1) + \cdots + f(k+1)}{f(1) + \cdots + f(k)}.$$

$$\ln \left\{ 1 + \frac{b(k)}{k+1} \right\} = \ln\{f(1) + \cdots + f(k+1)\} - \ln\{f(1) + \cdots + f(k)\}.$$

$$\sum_{k=1}^{n} \ln \left\{ 1 + \frac{b(k)}{k+1} \right\} = \text{ telescoping sum} = \ln\{f(1) + \cdots + f(n+1)\} - \ln f(1).$$

$$\exp \sum_{k=1}^{n} \ln \left\{ 1 + \frac{b(k)}{k+1} \right\} = \frac{f(1) + \cdots + f(n+1)}{f(1)} \quad \text{for } n \ge 0.$$

The first conclusion now follows. By the definition of $b(\cdot)$ and the first conclusion, we have, for $n \ge 1$,

$$f(n+1) = \frac{b(n)}{n+1} \{f(1) + \cdots + f(n)\} = \frac{b(n)}{n+1} f(1) \exp \sum_{k=1}^{n-1} \ln \left\{ 1 + \frac{b(k)}{k+1} \right\},$$

and the first and last terms are also equal when $n = 0$. $\square$

Lemma 2. *If* $\sum_{k=1}^{n} f(k) \asymp n f(n)$*, then* $b(n) \asymp 1$.

Proof. Assume

$$0 < C_1 \le \frac{f(1)+\cdots f(n)}{nf(n)} \le C_2 < \infty$$

for $n \in \mathbb{N}$. Then

$$C_1 \le \frac{\sum_{k=1}^{n+1} f(k)}{(n+1)f(n+1)} = \frac{\sum_{k=1}^{n} f(k)}{(n+1)f(n+1)} + \frac{1}{n+1} \le C_2,$$

so

$$C_1 - \frac{1}{n+1} \le \frac{1}{b(n)} \le C_2 - \frac{1}{n+1} < C_2.$$

For sufficiently large n—$n \ge n_0$, say— $C_1 - 1/(n+1) \ge C_1 - 1/(n_0+1) > 0$, so that for $n \ge n_0$,

$$\infty > \frac{1}{C_1 - 1/(n_0+1)} \ge b(n) > \frac{1}{C_2} > 0.$$

Thus $b(n) \asymp 1$. □

Lemma 3. *If $b(n) \asymp 1$, then there exist sequences $(g(n))$ and $(h(n))$ such that $g(n) \asymp 1$, $h(n) \asymp 1$, and*

$$f(n) = \frac{g(n)}{n} \exp \sum_{k=1}^{n} \frac{h(k)}{k} \quad \textit{for } n \in \mathbb{N}. \tag{5}$$

Proof. For $n \ge 3$ let

$$h(n) = n \ln \left\{ 1 + \frac{b(n-2)}{n-1} \right\}$$

and let $h(1)$ and $h(2)$ be arbitrary positive numbers. For $n \in \mathbb{N}$ let

$$g(n) = f(1)b(n-1)\exp\{-h(1) - h(2)/2\}.$$

Then a simple calculation shows that Equation (4) implies Equation (5). Now suppose that $0 < b_l \le b(n) \le b_u < \infty$ for $n \ge 0$. Basic calculus shows that $\ln(1+x)$ is increasing and $x - x^2/2 \le \ln(1+x) \le x$ for $x \ge 0$. Therefore, for $n \ge 3$,

$$h(n) = n \ln \left\{ 1 + \frac{b(n-2)}{n-1} \right\} \le n \ln \left\{ 1 + \frac{b_u}{n-1} \right\} \le n \frac{b_u}{n-1} < 2b_u,$$

and

$$h(n) \ge n \ln \left\{ 1 + \frac{b_l}{n-1} \right\} \ge n \frac{b_l}{n-1} \left(1 - \frac{1}{2} \frac{b_l}{n-1} \right) > b_l \left(1 - \frac{b_l}{2} \right) > 0,$$

since $0 < b_l \le b(0) = 1$, so that for $n \ge 1$ we have

$$0 < \min\{h(1), h(2), b_l(1 - b_l/2)\} \le h(n) \le \max\{h(1), h(2), 2b_u\} < \infty.$$

Also for $n \ge 1$ we have

$$0 < f(1)b_l e^{-h(1)-h(2)/2} \le g(n) \le f(1)b_u e^{-h(1)-h(2)/2} < \infty. \quad \square$$

Lemma 4. *If $f(n)$ is given by (5) where $g(n) \asymp 1$ and $h(n) \asymp 1$, then there exist constants α and β, both greater than -1, such that $n^{-\alpha} f(n)$ is almost decreasing and $n^{-\beta} f(n)$ is almost increasing. If $0 < h_l \le h(n) \le h_u < \infty$ for all n, then we may take $\alpha = h_u - 1$ and $\beta = h_l - 1$.*

Proof. We shall show that there exist constants α and β and positive constants m and M such that

$$n^{-\alpha} f(n) \le M k^{-\alpha} f(k)$$

and

$$m k^{-\beta} f(k) \le n^{-\beta} f(n)$$

for $1 \le k \le n < \infty$. Assume that $0 < g_l \le g(n) \le g_u < \infty$ and $0 < h_l \le h(n) \le h_u < \infty$ for $n \ge 1$. Assume $1 \le k \le n$. By hypothesis,

$$\frac{f(n)}{f(k)} = \frac{g(n)/n}{g(k)/k} \cdot \frac{\exp \sum_{j=1}^{n} h(j)/j}{\exp \sum_{j=1}^{k} h(j)/j} = \frac{k g(n)}{n g(k)} \exp \sum_{j=k+1}^{n} \frac{h(j)}{j},$$

where

$$0 < \frac{g_l}{g_u} \le \frac{g(n)}{g(k)} \le \frac{g_u}{g_l} < \infty$$

and

$$h_l \int_{k+1}^{n+1} \frac{1}{x}\, dx \le h_l \sum_{j=k+1}^{n} \frac{1}{j} \le \sum_{j=k+1}^{n} \frac{h(j)}{j} \le h_u \sum_{j=k+1}^{n} \frac{1}{j} \le h_u \int_{k}^{n} \frac{1}{x}\, dx.$$

Then

$$h_l \ln \frac{n+1}{k+1} \le \sum_{j=k+1}^{n} \frac{h(j)}{j} \le h_u \ln \frac{n}{k},$$

whence

$$\left(\frac{n}{2k}\right)^{h_l} < \left(\frac{n+1}{k+1}\right)^{h_l} \le \exp \sum_{j=k+1}^{n} \frac{h(j)}{j} \le \left(\frac{n}{k}\right)^{h_u}.$$

Take $m = 2^{-h_l} g_l / g_u$ and $M = g_u / g_l$; then

$$m \left(\frac{n}{k}\right)^{h_l - 1} \le \frac{f(n)}{f(k)} \le M \left(\frac{n}{k}\right)^{h_u - 1}.$$

Let $\alpha = h_u - 1$ and $\beta = h_l - 1$; then

$$m k^{-\beta} f(k) \le n^{-\beta} f(n) \quad \text{and} \quad n^{-\alpha} f(n) \le M k^{-\alpha} f(k). \quad \square$$

Lemma 5. *If there exist constants α and β greater than -1 such that $n^{-\alpha} f(n)$ is almost decreasing and $n^{-\beta} f(n)$ is almost increasing, then $f(1) + \cdots + f(n) \asymp n f(n)$.*

Proof. The hypothesis implies that there exist positive constants m and M and indices n_0' and n_0'' such that $n^{-\alpha} f(n) \le M k^{-\alpha} f(k)$ for $n_0' \le k \le n$ and $m k^{-\beta} f(k) \le n^{-\beta} f(n)$ for $n_0'' \le k \le n$. Then for $n \ge k \ge n_0 := \max\{n_0', n_0''\}$,

$$\frac{k^\alpha}{M n^\alpha} \le \frac{f(k)}{f(n)} \le \frac{k^\beta}{m n^\beta}$$

and $nf(n) \geq n^{1+\beta} m k^{-\beta} f(k) \geq n^{1+\beta} m^2 n_0^{-\beta} f(n_0) = cn^{1+\beta}$ (say), where $1+\beta > 0$. Then

$$\frac{\sum_{k=n_0}^{n} k^\alpha}{Mn^{1+\alpha}} \leq \frac{\sum_{k=n_0}^{n} f(k)}{nf(n)} \leq \frac{\sum_{k=n_0}^{n} k^\beta}{mn^{1+\beta}}$$

and

$$0 < \frac{f(1)+\cdots+f(n_0-1)}{nf(n)} \leq \frac{f(1)+\cdots+f(n_0-1)}{cn^{1+\beta}} \leq \frac{f(1)+\cdots+f(n_0-1)}{c}.$$

Now

$$\sum_{k=n_0}^{n} k^\alpha \sim \int_{n_0}^{n} x^\alpha \, dx \sim \frac{n^{1+\alpha}}{1+\alpha} - \frac{(n_0)^{1+\alpha}}{1+\alpha} \sim \frac{n^{1+\alpha}}{1+\alpha},$$

and similarly for $\sum_{k=n_0}^{n} k^\beta$, since both α and β are greater than -1. Therefore, there are constants C_1 and C_2 such that

$$0 < C_1 \leq \frac{f(1)+\cdots+f(n_0-1)+f(n_0)+\cdots+f(n)}{nf(n)} \leq C_2 < \infty$$

for all sufficiently large n. $\square$

Theorem 3. *Let $f : \mathbb{N} \to \mathbb{R}^+$. Then $f(1)+\cdots+f(n) \asymp nf(n)$ if and only if there exist constants $\alpha > -1$, $\beta > -1$, $m > 0$, $M > 0$, and $n_0 \in \mathbb{N}$ such that*

$$m\lambda^\beta \leq \frac{f(\lfloor \lambda n \rfloor)}{f(n)} \leq M\lambda^\alpha$$

for every real $\lambda \geq 1$ and integer $n \geq n_0$.

Proof. Use parts 1 and 3 of Theorem 2 ....
((More to come ...))

4. O-Regular Variation

Now consider functions of a real variable. A function $f : [X, \infty) \to \mathbb{R}^+$ for some real X is called a *regularly varying function* if it is measurable and

$$\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)} = \lambda^\rho$$

for some constant ρ and all $\lambda > 0$; it is called an *extended regularly varying function* if it is measurable and there exist constants c and d such that

$$\lambda^d \leq \liminf_{x\to\infty} \frac{f(\lambda x)}{f(x)} \leq \limsup_{x\to\infty} \frac{f(\lambda x)}{f(x)} \leq \lambda^c$$

for every $\lambda \geq 1$, and it is said to be a *O-regularly varying function* if it is measurable and satisfies

$$0 < \liminf_{x\to\infty} \frac{f(\lambda x)}{f(x)} \leq \limsup_{x\to\infty} \frac{f(\lambda x)}{f(x)} < \infty$$

for every $\lambda \geq 1$. Two such functions f and g are asymptotically of the same order—written $f \asymp g$—if $f(x) = O(g(x))$ and $g(x) = O(f(x)$ as $x \to \infty$, i.e.,

$$\limsup_{x\to\infty} \frac{f(x)}{g(x)} < \infty \quad \text{and} \quad \limsup_{x\to\infty} \frac{g(x)}{f(x)} < \infty.$$

These are definitions used in the theory of regular variation. This theory was in essence initiated by Jovan Karamata starting in 1930, and is now a well-developed part of real analysis; see [1] and [2]. The results given here provide a discrete counterpoint to results usually worked out for functions of a real variable.

As an illustration of the connection, we may use Theorem 3 to show that the function $g(x) := f(\lfloor x \rfloor)$ is a O-regularly varying function. For another example, note the parallel between the representation formula in part 2 of Theorem 2 and the representation formula for O-regularly varying functions f:

$$f(x) = \exp\left\{\eta(x) + \int_1^x \frac{\xi(t)}{t}\,dt\right\},$$

where η and ξ are bounded and measurable functions.

((More to come ...))

Acknowledgment

I thank my colleague Max Hailperin for bringing this problem to my attention and for helpful conversations along the way.

References

[1] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation*, Cambridge University Press, Cambridge, 1989.

[2] Eugene Seneta, *Regularly Varying Functions*, Springer-Verlag, 1976.

Gustavus Adolphus College, St. Peter MN 56082
E-mail address: holte@gustavus.edu