Error Bounds on the Loggamma Function Amenable to Interval Arithmetic

Russell Leidich
pkejjy@gmail.com

September 12, 2016

Keywords: loggamma, logarithm, log, gamma, interval, range reduction, error, bound, sum, factorial, logΓ

Abstract

Unlike other common transcendental functions such as log and sine, James Stirling's convergent series for the loggamma ("logΓ") function suggests no obvious method by which to ascertain meaningful bounds on the error due to truncation after a particular number of terms. ("Convergent" refers to the fact that his original formula appeared to converge, but ultimately diverged [1].) As such, it remains an anathema to the interval arithmetic algorithms which underlie our confidence in its various numerical applications.

Certain error bounds do exist in the literature [1], but involve branches and procedurally generated rationals which defy straightforward implementation via interval arithmetic.

In order to ameliorate this situation, we derive error bounds on the loggamma function which are readily amenable to such methods.

Acknowledgements

Thanks to Raphael Schumacher, who published several useful transcendental summation series [2], for outlining an as yet unpublished approach to loggamma error analysis involving a Weniger transform of the Euler-

Maclaurin summation formula. It would be interesting to compare those results with the bounds described herein, if and when they are published.

Thanks also to Mike Irwin for his thoughtful recommendations on the structure of this paper.

Finally, be it known that WolframAlpha managed to reduce a particular infinite sum to a finite expression, without which this paper would not exist.

Applications

The loggamma function finds broad use in number theory and statistics, for example, in approximating the gamma function, profiling the energy states in Einstein solids [3], and detecting buried signals in my open source Dyspoissometer noise analysis software [4]. It also enables the efficient computation of the sums of the logs of the first natural numbers, otherwise known as the log of a factorial:

$$\sum_{a=1}^{A} \ln a \equiv \ln(A!) \equiv log\Gamma(A+1)$$

Method

Stirling's convergent formula (hereinafter “the formula”) is the right side of the following identity [1]:

$$log\Gamma(x) \equiv \left(x-\frac{1}{2}\right)\ln x - x + \frac{1}{2}\ln 2\pi + \sum_{a=1}^{\infty} \frac{1}{(2a)(x+\overline{a})} \sum_{b=1}^{a} \frac{b|S(a,b)|}{(b+1)(b+2)}$$

where: (x>0); $|S(a, b)|$ are the magnitudes of the Stirling numbers of the first kind [5], e.g. $|S(5, 2)|$=50; and

$$(x+\overline{a})$$

means the product given by

$$\prod_{b=1}^{a}(x+b)$$

All of the terms on the right side of the formula except the nested infinite sum have been well studied and are amenable to interval arithmetic. Nevertheless, for the sake of completeness, we will examine each transcendental expression in detail.

First of all, the log operands must be multiplied or divided by a suitable whole power of 2 in order to scale them into the open-closed domain (1, 2]:

$$\ln x \equiv \ln 2^n x - n \ln 2\,;\, n = \lfloor \log_2 \frac{1}{x} \rfloor + 1\,;\, 0 < x \leq 1$$

or

$$\ln x \equiv \ln \frac{x}{2^n} + n \ln 2\,;\, n = \lfloor \log_2 x \rfloor\,;\, x > 1$$

This results in an expression of involving (ln *y*), which is guaranteed to converge according to the Taylor series

$$\ln y \equiv \sum_{a=1}^{\infty} \frac{(-1)^{a+1}(y-1)^a}{a}$$

(This is true even if (y=2), even though the radius of convergence is one due to the singularity at (y=0).) This series involves alternating signs and terms with monotonically decreasing magnitude, which by the alternating series test [6] is sufficient to guarantee that the magnitude of the error due to truncation after a given number of terms is less than the magnitude of the last term computed. (In fact it's also less than the postterminal term, but if one were to compute that term, then it might as well be added to the series, so in practice we compute the confidence interval from the last term added.) Furthermore if the last term in the partial sum was positive, then that sum is an upper bound; otherwise, it's a lower bound.

Of course, to have gotten here in the first place required the scaling by a power of 2, which as shown above may have produced a multiple of (ln 2), which is of course transcendental. It's tempting to dismiss this as a constant which one could look up in a table, but this paper concerns range reduction via interval arithmetic, so we need to perform explicit error analysis.

The easiest method is via the identity

$$\ln 2 \equiv \eta(1)$$

where η(x) is the Dirichlet eta function [7]. That is:

$$\ln 2 \equiv \sum_{a=1}^{\infty} \frac{(-1)^{(a+1)}}{a}$$

Note that this is identical to the Taylor series for the log at (y=2). Therefore, as with (ln y), the magnitude of the error is bounded by the last term computed, and the same alternating upper and lower bounds apply.

There is one other transcendental constant the range of which must be reduced according to interval arithmetic, namely, π. There are a wide variety of series that we might use, but we should choose the one with the most straightforward range reduction. One such approach is to scale the arctangent of one:

$$\pi \equiv 4 \arctan 1$$

$$\equiv 4 \sum_{a=0}^{\infty} \frac{(-1)^{a}}{2a+1}$$

One again we have an alternating series consisting of terms monotonically decreasing in magnitude. Therefore the magnitude of the error in this approximation is less than 4 times the magnitude of the term last computed, on account of the scaling factor. And once again, alternating upper and lower bounds apply.

Now, finally, we have the nested infinite series:

$$\sum_{a=1}^{\infty} \frac{1}{(2a)(x+\bar{a})} \sum_{b=1}^{a} \frac{b|S(a,b)|}{(b+1)(b+2)}$$

According to [1], this series has been proven to converge for (x>0), but what is the maximum possible error magnitude due to truncation?

We begin by looking at the coefficients C_b of the $|S(a, b)|$ terms, that is

$$C_b \equiv \frac{b}{(b+1)(b+2)}; b \in \mathbb{N}$$

which have no singularities in the domain in question (namely, the naturals). Now let's consider the ratio (C_{b+1}/C_b):

$$\frac{C_{b+1}}{C_b} \equiv \frac{(b+1)^2}{b(b+3)}$$

For which (b=1) is the only value which makes the ratio equal to one; greater values of b result in monotonically nonincreasing values of thereof, in particular

$$\left(C_1 \equiv C_2 \equiv \frac{1}{6}\right) \wedge \left(C_{b+2} < C_{b+1}\right)$$

Now consider the $|S(a, b)|$ terms themselves. First of all, by definition:

$$\sum_{b=1}^{a} |S(a,b)| \equiv a!; a,b \in \mathbb{N}$$

simply because, for each value of a, the sum above counts the number of possible permutations of a nodes in a graph. (The upper limit of the sum is (b=a) because all terms beyond that index are zero.) Furthermore $|S(a, b)|$ is constrained by the generating function:

$$|S(a+1,b+1)| \equiv a|S(a,b+1)| + |S(a,b)|$$

Now suppose that we define $P(a, b)$ as a fraction of a probability density function:

$$P(a,b) \equiv \frac{|S(a,b)|}{a!}$$

This is indeed a valid probability density function because its sum over *b* is one, for any given value of *a*.

Now, due to the generating function above, we know that *P*(*a*, *b*) is monotonically nonincreasing as *a* increases (as one moves down the table of |*S*(*a*, *b*)|, as presented in [5], with *b* fixed). That is:

$$P(a+1,b+1) \equiv \frac{|S(a+1,b+1)|}{(a+1)!}$$

$$\equiv \frac{a|S(a,b+1)|+|S(a,b)|}{(a+1)!}$$

$$\equiv \frac{1}{a+1}\left(\frac{a|S(a,b+1)|}{a!}+\frac{|S(a,b)|}{a!}\right)$$

$$\equiv \frac{1}{a+1}(aP(a,b+1)+P(a,b))$$

which is to say that probability flows monotonically to the right (toward greater *b*) as *a* increases. (This is purely a net flux argument based on a normalized linear combination of prior probabilities; probability does not *decrease* monotonically to the right. Indeed |S(a,b)| looks curiously like a Poisson distribution, which is a whole other investigation in itself: the Stirling numbers of the *second* kind relate to Poisson distributions via the Bell numbers [8]. But I digress.) Critically, this probability density is *not* replenished from the top or the left, as *P*(0, 0) is one and *P*(*a*, 0) and *P*(0, *b*) are both zero. Thus *P*(0, 0) only serves to seed *P*(1, 1) with one, after which recursion applies in both dimensions; whereas neither *P*(*a*, 0) nor *P*(0, *b*) supply any other *P*(*a*, *b*) with incoming flux.

Now we connect this flux argument with C_b:

$$\sum_{b=1}^{a+1} C_b P(a+1,b) \leq \sum_{b=1}^{a} C_b P(a,b)$$

because we already established that the coefficients C_b are monotonically nonincreasing and probability is flowing toward greater *b* while *a* increases. Therefore, from the definition of *P*(*a*, *b*),

$$a!\sum_{b=1}^{a+1} C_b |S(a+1,b)| \leq (a+1)!\sum_{b=1}^{a} C_b |S(a,b)|$$

which implies that

$$F(a) \equiv \frac{\sum_{b=1}^{a+1} C_b |S(a+1,b)|}{\sum_{b=1}^{a} C_b |S(a,b)|} \leq (a+1)$$

But the situation is complicated by the (1/(2*a*)) in the formula. Let

$$G(a) \equiv \frac{1}{2a}\sum_{b=1}^{a} C_b |S(a,b)| \quad .$$

Then:

$$\frac{G(a+1)}{G(a)} \equiv \frac{2a}{2a+2} F(a)$$

$$\frac{G(a+1)}{G(a)} \leq \frac{a}{a+1}(a+1)$$

$$\frac{G(a+1)}{G(a)} \leq a$$

Therefore if we evaluate the *H*(*A*, *x*) given by

$$H(A,x) \equiv \frac{1}{(2A)(x+\overline{A})}\sum_{b=1}^{A} \frac{b|S(A,b)|}{(b+1)(b+2)}$$

then the above constraint on the growth rate of *G*(*a*), and the definition of $(x+\overline{a})$, jointly imply that

$$\sum_{a=A+1}^{\infty}\frac{1}{(2a)(x+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}<H(A,x)\sum_{a=A+1}^{\infty}\frac{(a-1)!/(A-1)!}{(x+a)!/(x+A)!}$$

$$\sum_{a=A+1}^{\infty}\frac{1}{(2a)(x+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}<H(A,x)\sum_{a=A+1}^{\infty}\frac{(a-1)!(x+A)!}{(A-1)!(x+a)!}$$

$$\sum_{a=A+1}^{\infty}\frac{1}{(2a)(x+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}<H(A,x)\frac{(x+A)!}{(A-1)!}\sum_{a=A+1}^{\infty}\frac{(a-1)!}{(x+a)!}$$

Note that "≤" has changed to "<" because equality only applies to C_1 and C_2. According to [9], it turns out that

$$\sum_{a=1}^{A}\frac{(a-1)!}{(x+a)!}\equiv\frac{(x+1)(x+A+1)!-A!(x+1)!(x+A+1)}{x(x+1)!(x+A+1)!}$$

where the definition of (x!) is analytically extended to noninteger values via the identity presented in [10], namely

$$x!\equiv\Gamma(x+1)$$

but otherwise behaves in its usual manner. The expression second above further simplies to:

$$\sum_{a=1}^{A}\frac{a!}{(x+a)!}\equiv\frac{1}{x}\left(\frac{1}{x!}-\frac{A!}{(x+A)!}\right)$$

But

$$\sum_{a=A+1}^{\infty}\frac{a!}{(x+a)!}\equiv\sum_{a=1}^{\infty}\frac{a!}{(x+a)!}-\sum_{a=1}^{A}\frac{a!}{(x+a)!}$$

so

$$\sum_{a=A+1}^{\infty}\frac{a!}{(x+a)!}\equiv\frac{1}{x}\frac{1}{x!}-\frac{1}{x}\left(\frac{1}{x!}-\frac{A!}{(x+A)!}\right)$$

$$\equiv\frac{A!}{x(x+A)!}$$

which implies that

$$\frac{(x+A)!}{(A-1)!}\sum_{a=A+1}^{\infty}\frac{(a-1)!}{(x+a)!}\equiv\frac{A}{x}$$

and thus

$$\sum_{a=A+1}^{\infty}\frac{1}{(2a)(x+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}<H(A,x)\frac{A}{x}$$

which makes it appear as though the error magnitude actually *increases* with each computed term. However, expanding $H(A, x)$, gives

$$\sum_{a=A+1}^{\infty}\frac{1}{(2a)(x+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}<\frac{1}{(2x)(x+\overline{A})}\sum_{b=1}^{A}\frac{b|S(A,b)|}{(b+1)(b+2)}$$

which is in turn bound by

$$\frac{1}{(2x)(x+\overline{A})}\sum_{b=1}^{A}\frac{b|S(A,b)|}{(b+1)(b+2)}<\frac{(A-1)!}{(2x)A!}$$

because the sum is bounded by $((A\text{-}1)!)$ according to the limit on the growth rate of $G(a)$, while the denominator must be at least $((2x)(A!))$ due to the expansion of $(x+\overline{A})$. This reduces to

$$\frac{1}{(2x)(x+\overline{A})}\sum_{b=1}^{A}\frac{b|S(A,b)|}{(b+1)(b+2)}<\frac{1}{2Ax}$$

which clearly approaches zero in the limit of infinite A, which is what we want from an error bound. However, for the sake of accuracy and efficiency, the error bound should be computed as

$$H(A,x)\frac{A}{x}$$

using straightforward interval arithmetic because $H(A, x)$ is just the last term computed.

Furthermore, considering that all of the terms in the original nested infinite sum are positive, their partial sum through (*a*=*A*) constitutes a lower bound to which this error bound should be added in order to obtain an interval for the entire nested series.

For example, we can compute H(A, x) for the nested series of (*x*=7.6) to (*A*=3) terms:

$$H(A,x) \equiv \frac{1}{(2A)(x+\overline{A})} \sum_{b=1}^{A} \frac{b|S(A,b)|}{(b+1)(b+2)}$$

which, according to [11] and upon substituting for *A* and *x* gives

$$H(A,x) = \frac{1}{(2*3)(7.6+\overline{3})} \sum_{b=1}^{3} \frac{b|S(3,b)|}{(b+1)(b+2)}$$

$$H(3,7.6) = \frac{1475}{7876224}$$

which implies that

$$H(A,x)\frac{A}{x} = H(3,7.6)\frac{3}{7.6}$$

$$= \frac{1475}{7876224} * \frac{3}{7.6}$$

$$= \frac{7375}{99765504} \in [0.0000739233, 0.0000739234]$$

which is an upper bound on the error induced by truncating the nested series after (*a*=3). The first 3 terms add up as follows:

$$\sum_{a=1}^{3} \frac{1}{(2a)(x+\overline{a})} \sum_{b=1}^{a} \frac{b|S(a,b)|}{(b+1)(b+2)} = \frac{5}{516} + \frac{25}{24768} + \frac{1475}{7876224}$$

$$= \frac{85745}{7876224} \in [0.0108865618, 0.0108865619]$$

which provides an interval for the nested infinite series for (x=7.6):

$$\sum_{a=1}^{\infty}\frac{1}{(2a)(7.6+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}\in[0.0108865618,0.0108865619+0.0000739234]$$

$$\sum_{a=1}^{\infty}\frac{1}{(2a)(7.6+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}\in[0.0108865618,0.0109604853]$$

But is this correct? Well, the value of the nested infinite series is exactly

$$\sum_{a=1}^{\infty}\frac{1}{(2a)(7.6+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}=log\Gamma(7.6)-\log(7.6-\frac{1}{2})\ln 7.6+7.6-\frac{1}{2}\ln 2\pi$$

which according to [12] is bounded by

$$\sum_{a=1}^{\infty}\frac{1}{(2a)(7.6+\overline{a})}\sum_{b=1}^{a}\frac{b|S(a,b)|}{(b+1)(b+2)}\in[0.0109586153,0.0109586154]$$

which, as expected, is entirely included by the interval second above.

Conclusion

By means of the foregoing techniques, Stirling's convergent formula for the loggamma function can be made amenable to interval arithmetic. The process entails range reduction on 4 different series, namely ln x, ln 2, π, and the nested infinite series.

It would be tempting to combine all of them into a single expression bounded by a unified error term. However, in practice, each series is likely to be computed in a separate process and in any event exhibits distinct asymptotic error behavior. So a reasonable practice is to evaluate them as separate intervals, then linearly combine them to finally obtain an interval for the loggamma.

Open Questions

Empirically, it appears that the loggamma function is more rapidly range reduced via a series of the form

$$log\Gamma(x+1)\equiv\left(x+\frac{1}{2}\right)\left(\ln\left(x+\frac{1}{2}\right)-1\right)+\frac{1}{2}\ln 2\pi-\frac{1}{24\left(x+\frac{1}{2}\right)}+\frac{7}{2880\left(x+\frac{1}{2}\right)\left(x+\frac{3}{2}\right)\left(x+\frac{5}{2}\right)}$$
$$+\frac{1}{144\left(x+\frac{1}{2}\right)\left(x+\frac{3}{2}\right)\left(x+\frac{5}{2}\right)\left(x+\frac{7}{2}\right)}\pm\ldots$$

where ($x>0$) and which resulted from using

$$log\Gamma(A+1)\equiv\ln(A!)\equiv\sum_{a=1}^{A}\ln a\approx\int_{\frac{1}{2}}^{A+\frac{1}{2}}\ln t\,dt$$

for improved accuracy in determining the log of a factorial, as opposed to the typical whole number integral bounds. (The integral above generated the log terms; the other terms emerged from iterative analysis of the residual error.) Curiously, the partial sum above, while reminiscent of the formula, exhibits strange coefficients with no clear consistency of sign, and a missing term in $O(1/x^2)$. Unfortunately, due to limited time and numerical precision, I have been unable to derive any further terms. However, the rate of convergence per arithmetic operation (essentially, the computational efficiency) of the series above appears to be superior to that of the formula. (I'm *almost* certain that the coefficients are correct as stated, but I'm unable to discern a pattern, let alone prove that it persists to infinity.) Perhaps someone can use these hints to discover a more efficient alternative, although I suspect that most of that improvement would occur due to the missing $O(1/x^2)$ term. The answer probably lies in expanding all the transcendentals and subtracting them from the formula in order to obtain a new series.

Bibliography

1. https://en.wikipedia.org/wiki/Stirling%27s_approximation

2. Raphael Schumacher, Rapidly Convergent Summation Formulas involving Stirling Series, “Convergent versions of Stirling’s formula”, arXiv:1602.00336v1 [math.NT], 31 Jan 2016.

3. http://hyperphysics.phy-astr.gsu.edu/hbase/Therm/einsol.html

4. https://dyspoissonism.blogspot.com

5. https://en.wikipedia.org/wiki/Stirling_numbers_of_the_first_kind

6. https://en.wikipedia.org/wiki/Alternating_series_test

7. https://en.wikipedia.org/wiki/Dirichlet_eta_function

8. https://en.wikipedia.org/wiki/Bell_number

9. https://www.wolframalpha.com/input/?i=sum+from+a%3D1+to+A+of+%5B%5Ba-1%5D!%2F%5Ba%2Bx%5D!%5D

10. http://mathworld.wolfram.com/GammaFunction.html

11. https://www.wolframalpha.com/input/?i=%5B1%2F%5B%5B2*3%5D%5B76%2F10%2B1%5D%5B76%2F10%2B2%5D%5B76%2F10%2B3%5D%5D%5D*Sum%5Bb*%7CStirlingS1%5B3,b%5D%7C%2F%5B%5Bb%2B1%5D*%5Bb%2B2%5D%5D,%7Bb,1,3%7D%5D

12. https://www.wolframalpha.com/input/?i=loggamma%5B7.6%5D-%5B7.6-1%2F2%5D*ln%5B7.6%5D%2B7.6-%5B1%2F2%5D*ln%5B2*pi%5D