Probabilistic Roundoff Error Analysis

Johnathan Rhyne
Advisor: Ilse Ipsen

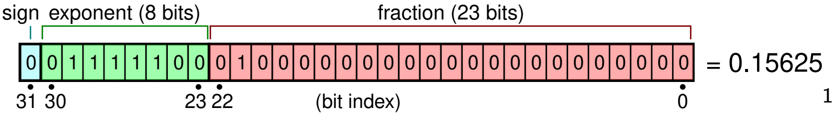
September 22, 2020

Outline

Other Work in This Area

1. Ilse C.F. Ipsen and Hua Zhou
   - Ipsen, I. C. F. & Zhou, H. Probabilistic Error Analysis for Inner Products SIAM J. Matrix Anal. Appl., 2020, to appear
2. Nick J. Higham and Mary
   - Nicholas J. Higham, Théo Mary. A New Approach to Probabilistic Rounding Error Analysis. SIAM Journal on Scientific Computing, Society for Industrial and Applied Mathematics, 2019, 41 (5), pp.A2815-A2835. ff10.1137/18M1226312ff. ffhal-02311269f

Floating Point Representation



[1]

This is represents:

$$(-1)^0 \left(1 + 2^{-1}\right) \times 2^{2^2+2^3+2^4+2^5+2^6-127} = 2^{-3} \times 1.25 = \frac{5}{2^5} = \frac{5}{32}$$

[1] Image from `https://commons.wikimedia.org/wiki/File:Float_example.svg`

Error in Floating Point Addition

We assume a and b are floating point numbers.

$$\mathrm{fl}(a+b) = (a+b)(1+\delta).$$

We also assume that $|\delta| \leq u$ where u is unit roundoff.

Unit roundoff for IEEE single and double precision floating point numbers.

Single Precision	Double Precision
$u = 2^{-24} \approx 5.96 \times 10^{-8}$	$u = 2^{-53} \approx 1.11 \times 10^{-16}$

Summation Algorithm

Algorithm 1 Sequential Summation

Inputs: n real numbers $x_1, \dots, x_n$

Outputs: The sum: $\sum_{k=1}^{n} x_k$

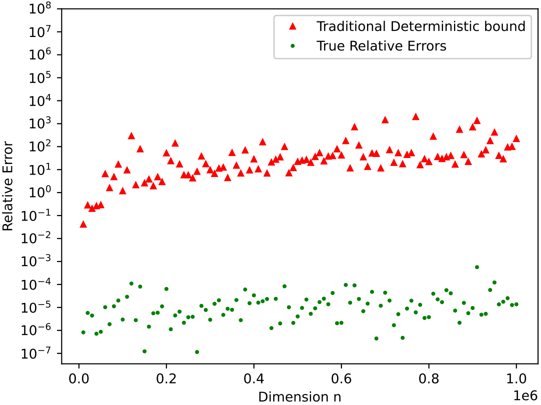
```
Sum ← 0
for k = 1 up to n do
    Sum ← Sum + x_k
end for
return Sum
```

Here is how we represent the partial sums

Exact computation	Floating point arithmetic	Index range
$z_1 = x_1$	$\hat{z}_1 = x_1$	
$z_2 = x_1 + x_2$	$\hat{z}_2 = (x_1 + x_2)(1 + \delta_2)$	
$z_k = z_{k-1} + x_k$	$\hat{z}_k = (\hat{z}_{k-1} + x_k)(1 + \delta_k)$	$2 \le k \le n$
$z_n = \sum_{k=1}^{n} x_k$	$\hat{z}_n = \mathrm{fl}\left(\sum_{k=1}^{n} x_k\right)$	

Traditional Deterministic Roundoff Error Bounds

$$\left|\frac{z_n - \hat{z}_n}{z_n}\right| \leq u(n-1)\frac{\sum_{k=1}^{n}|x_k|}{|z_n|} + \mathcal{O}\left(u^2\right)$$



Our Deterministic Bound

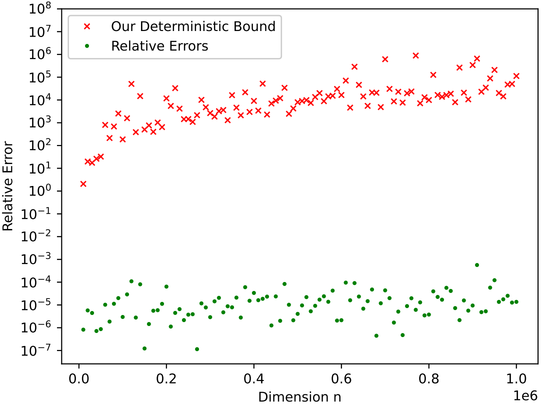
Construction	Valid Range
$c_1 = \|x_1\|((1+u)^{n-1} - 1)$	
$c_k = \|x_k\|((1+u)^{n-k+1} - 1)$	$2 \leq k \leq n$

Note that for $1 \leq k \leq n$, c_k are multiples of u, that is, $c_k = |x_k|(u + \cdots)$.

$$\left|\frac{z_n - \hat{z}_n}{z_n}\right| \leq \sqrt{n}\frac{\sum_{k=1}^{n} c_k}{|z_n|}$$

Numerical Experiment for Our Deterministic Bound

$$\left|\frac{\sum_{k=1}^{n} x_k - \mathrm{fl}\left(\sum_{k=1}^{n} x_k\right)}{\sum_{k=1}^{n} x_k}\right| \leq \sqrt{n}\,\frac{\sum_{k=1}^{n} c_k}{\left|\sum_{k=1}^{n} x_k\right|}$$



Probabilistic Model for Roundoff Errors

- We model roundoff errors as bounded zero mean random variables
  - $|\delta_k| \leq u$ for $2 \leq k \leq n$
  - $\mathbb{E}(\delta_k) = 0$ for $2 \leq k \leq n$

Construction	**Valid Range**
$Z_1 = x_1 \prod_{l=2}^{n} (1+\delta_l) - x_1$ $Z_k = x_k \prod_{l=k}^{n} (1+\delta_l) - x_k$	$2 \leq k \leq n$
$Z = \sum_{k=1}^{n} Z_k$	

Linearity of expectation implies

$$\mathbb{E}(Z) = 0$$
$$\mathbb{E}(Z_k) = 0. \qquad 1 \leq k \leq n$$

Azuma's Inequality[2]

If $A = A_1 + \cdots + A_n$ is a sum of independent real-valued random variables, $0 \leq a_k$ for $1 \leq k \leq n$, $0 < \delta < 1$, and

$$|A_k - \mathbb{E}[A_k]| \leq a_k \qquad 1 \leq k \leq n.$$

Then with probability at least $1 - \delta$

$$|A - \mathbb{E}[A]| \leq \sqrt{2 \ln \frac{2}{\delta}} \sqrt{\sum_{k=1}^{n} a_k^2}.$$

[2] Theorem 5.3 in Concentration Inequalities and Martingale Inequalities: A Survey by Chung, F. & Lu, L. 2006

First Probabilistic Bound

Construction	**Valid Range**
$c_1 = \|x_1\|((1+u)^{n-1} - 1)$	
$c_k = \|x_k\|((1+u)^{n-k+1} - 1)$	$2 \leq k \leq n$

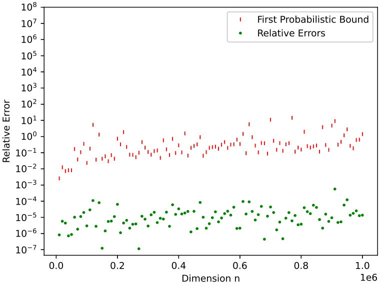
For any $0 < \delta < 1$, with probability at least $1 - \delta$

$$\left| \frac{z_n - \hat{z}_n}{z_n} \right| \leq \sqrt{2 \ln \frac{2}{\delta}} \frac{\sqrt{\sum_{k=1}^{n} c_k^2}}{|z_n|}.$$

Numerical Experiment for our First Probabilistic Bound

With probability at least $1 - \delta$,

$$\left|\frac{z_n - \hat{z}_n}{z_n}\right| \leq \sqrt{2 \ln \frac{2}{\delta}} \frac{\sqrt{\sum_{k=1}^{n} c_k^2}}{|z_n|}.$$



Here, we use $\delta = 10^{-16}$ as our failure probability.

Second Probabilistic Bound: Martingale[3]

A collection of random variables, $M_1, M_2, \ldots, M_n$ is called Martingale if the following are satisfied

1. $\mathbb{E}[|M_n|]$ is finite.
2. $\mathbb{E}[M_k | M_1, \ldots, M_{k-1}] = M_{k-1}$

This is also referred to as being a Martingale with respect to itself.

[3] Theorem 12.1 in Probability and Computing: Randomized Algorithms and Probabilistic Analysis by Mitzenmacher, M. & Upfal, E.

Azuma-Hoeffding Inequality[4]

If $B_1, \ldots, B_n$ is a Martingale with respect to itself, $0 \leq b_k$ for $1 \leq k \leq n$. If

$$|B_k - B_{k-1}| \leq b_{k-1} \qquad \text{for } 2 \leq k \leq n,$$

then for any $0 < \delta < 1$, with probability at least $1 - \delta$

$$|B_n - B_1| \leq \sqrt{2 \ln \frac{2}{\delta}} \sqrt{\sum_{k=1}^{n-1} b_k^2}.$$

[4] Theorem 12.4 in Probability and Computing: Randomized Algorithms and Probabilistic Analysis by Mitzenmacher, M. & Upfal, E.

Second Probabilistic Bound

Construction	**Valid Range**
$m_k = \lvert x_1\rvert(1+u)^{k-1} + \sum_{j=2}^{k+1} \lvert x_j\rvert(1+u)^{k-j+1}$	$1 \leq k \leq n-1$

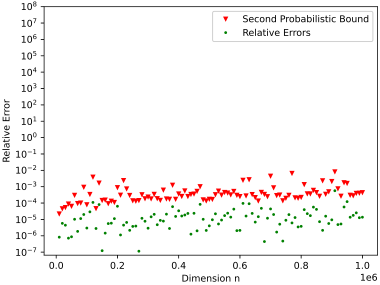
With probability at least $1-\delta$,

$$\left|\frac{z_n - \hat{z}_n}{z_n}\right| \leq u\sqrt{2\ln\frac{2}{\delta}}\,\frac{\sqrt{\sum_{k=1}^{n-1} m_k^2}}{|z_n|}$$

Second Probabilistic Bound Numerical Experiment

With probability at least $1-\delta$,

$$\left|\frac{z_n - \hat{z}_n}{z_n}\right| \leq u\sqrt{2\ln\frac{2}{\delta}}\,\frac{\sqrt{\sum_{k=1}^{n-1} m_k^2}}{|z_n|}$$



Here, we use $\delta = 10^{-16}$ as our failure probability.

Comparison of the Probabilistic Bounds

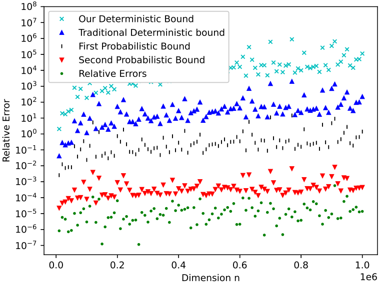
We derived two probabilistic bounds that hold with probability at least $1-\delta$. That are several orders of magnitude than the deterministic counterparts. We also found that our first bound is much more pessimistic than the second, more expensive one.

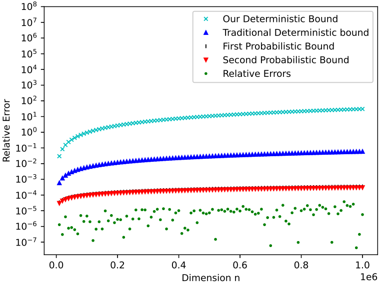
$$\left|\frac{z_n - \hat{z}_n}{z_n}\right| \leq \sqrt{2\ln\frac{2}{\delta}}\,\frac{\sqrt{\sum_{k=1}^{n} c_k^2}}{|z_n|}.$$

$$\left|\frac{z_n - \hat{z}_n}{z_n}\right| \leq u\sqrt{2\ln\frac{2}{\delta}}\,\frac{\sqrt{\sum_{k=1}^{n-1} m_k^2}}{|z_n|}$$

x_ks With Different Signs vs. x_ks With the Same Sign

With $\delta = 10^{-16}$ as our failure probability,





As shown above, when each of the x_k values are the same sign, all bounds are tighter, and our more pessimistic probabilistic bound becomes as accurate as the second.

Failure of the Second Bound

- Is this a fundamental problem with the bounds?
- Or do we need separate bounds depending on the structure of the data?