Homework 1: Problem 4

David Denberg

March 12, 2019

Part A

$$
\begin{aligned}
f(x) &= 1 - e^{-x} \\
f'(x) &= e^{-x} \\
(\text{cond } f)(x) &= \left|\frac{xf'(x)}{f(x)}\right| = \left|\frac{xe^{-x}}{1-e^{-x}}\right| = \frac{xe^{-x}}{1-e^{-x}} \quad \text{on } x \in (0,1)
\end{aligned}
$$

To show $(\text{cond } f)(x)$ is less then 1 on (0,1), we first assume:

$$
\begin{aligned}
\frac{xe^{-x}}{1-e^{-x}} &< 1 \\
xe^{-x} &< 1 - e^{-x} \\
x &< e^{x} - 1 = x + \frac{x^2}{2!} + \frac{x^3}{3!} + ... \\
0 &< \frac{x^2}{2!} + \frac{x^3}{3!} + ...
\end{aligned}
$$

This last expression is true for $x \in (0,1)$, so $(\text{cond } f)(x) < 1$ on (0,1).

Part B

$$
\begin{aligned}
\text{fl}(e^{-x}) &= e^{-x}(1+\epsilon_{exp}) \\
\text{fl}(1 - \text{fl}(e^{-x})) &= \text{fl}\left((1-e^{-x})\left(1 + \frac{\epsilon_1}{1-e^{-x}} - \frac{\epsilon_{exp}e^{-x}}{1-e^{-x}}\right)\right) \\
f_A(x) &= (1-e^{-x})\left(1 + \frac{\epsilon_1}{1-e^{-x}} - \frac{\epsilon_{exp}e^{-x}}{1-e^{-x}} + \epsilon_{rnd}\right)
\end{aligned}
$$

To get an upper bound on the error we let $\epsilon_1 = \epsilon_{rnd} = -\epsilon_{exp} = \text{eps}$. Then the maximum bounded $f_A(x)$ is:

$$
f_A(x) = (1-e^{-x})\left(1 + \frac{2 \cdot \text{eps}}{1-e^{-x}}\right) = 1 - e^{-x} + 2 \cdot \text{eps}
$$

The condition of the algorithm is calculated as:

$$
(\text{cond } A)(x) = \frac{1}{\text{eps}}\left|\frac{x_A - x}{x}\right| = \frac{1}{\text{eps}}\left|\frac{-\ln(1-f_A(x)) - x}{x}\right|
$$

First we plug in $f_A(x)$ into the expression:

$$\begin{aligned}
\frac{1}{\text{eps}}\left|\frac{-\ln(1-1+e^{-x}-2\cdot\text{eps})-x}{x}\right| &= \frac{1}{\text{eps}}\left|\frac{-\ln(e^{-x}(1-2\cdot\text{eps}\cdot e^{x}))-x}{x}\right| \\
&= \frac{1}{\text{eps}}\left|\frac{x-\ln(1-2\cdot\text{eps}\cdot e^{x})-x}{x}\right| \\
&= \frac{1}{\text{eps}}\left|\frac{-\ln(1-2\cdot\text{eps}\cdot e^{x})}{x}\right|
\end{aligned}$$

For eps $\ll 1$ we can expand the logarithm:

$$\begin{aligned}
\frac{1}{\text{eps}}\left|\frac{2\cdot\text{eps}\cdot e^{x} + \cancelto{\approx 0}{2\cdot\text{eps}^2\cdot e^{2x}} + \ldots}{x}\right| &= \frac{1}{\text{eps}}\left|\frac{2\cdot\text{eps}\cdot e^{x}}{x}\right| \\
&= \frac{2e^{x}}{x}
\end{aligned}$$

on (0,1). Thus (cond $A)(x)$ is bounded by $\frac{2e^x}{x}$ which is always greater than 1 on (0, 1).

Part C

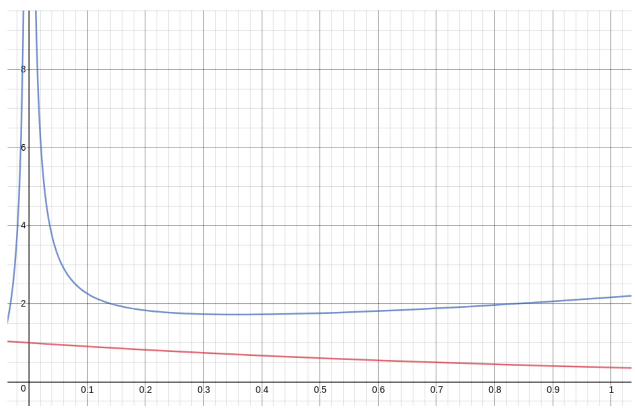

Figure 1: (cond $f)(x)$ is in red and (cond $A)(x)$ is in blue

The cause of the poor conditioning when x becomes small in (cond $A)(x)$ is the floating point error due to subtraction.

Part D

To find the value of x at which b bits of significance are lost we use the upper bounded $f_A(x)$:

$$f_A(x) = (1-e^{-x})\left(1+\frac{2\cdot\text{eps}}{1-e^{-x}}\right)$$

And

$$\epsilon_{max} = \frac{2\cdot\text{eps}}{1-e^{-x}}$$

Is the maximum error. Then for b bits of significance lost:

$$\log_2\left(\frac{2\cdot \text{eps}}{1-e^{-x}}\right) = b$$

$$x = -\log\left(1 - \frac{\cdot\text{eps}}{2^{b-1}}\right)$$

For 1, 2, 3, and 4 bits lost we need:

$$x_1 = -\log\left(1 - \text{eps}\right)$$

$$x_2 = -\log\left(1 - \frac{\text{eps}}{2}\right)$$

$$x_3 = -\log\left(1 - \frac{\text{eps}}{4}\right)$$

$$x_4 = -\log\left(1 - \frac{\text{eps}}{8}\right)$$

Part E

The maximum relative error is:

$$\frac{2\cdot \text{eps}}{x(1-e^{-x})}$$

Part F

An alternative function to evaluate would be:

$$f(x) = \ln\left(\frac{e}{e^{e^{-x}}}\right)$$

Because there isn't subtraction present then there shouldn't be a source of floating point error when x goes to 0.