CPSC 303 HOMEWORK 4 SOLUTIONS, SPRING 2020

JOEL FRIEDMAN

(1) Problem 1 (in the Exercise section) of the handout "Normal and Subnormal Numbers in Double Precision."

Solution:

MATLAB returns 1.7977e+308 for $n = 50, 51, 52$, and Inf for $n = 53, 54, 55$. The largest number in double precision is $2^{1023}(2 - 2^{-52})$ (see equation (1) on page 2 of the handout), which is why MATLAB reports Inf for $n = 53, 54, 55$.

Research supported in part by an NSERC grant.

(2) Problem 2 (in the Exercise section) of the handout "Normal and Subnormal Numbers in Double Precision."

Solution:

(a) For small n you see the exact solution, i.e., $C_1 + C_2(3/4)^n$ with $C_2 = 1$ and C_1 zero or negligible. Since MATLAB reports $x\{250\}$ as 9.3378e-20, for large n you numerically see the solution C_1 equal to 9.3378e-20 and C_2 zero or negligible. Hence the values of C_1, C_2 (as far as we can numerically observe) that explain both the small n and large n values are 9.3378e-20 and 1.

(b) MATLAB reports columns 1 through 25 of `ratio` as 1, and the larger columns as 1.0000. Similarly, it reports the first 25 values of `ratio_versus_one` as 0, and the remaining values begin at 1.1102e-16, for the 26th value, and ending at 8.3189e-13, in a fashion that roughly increases but does not strictly increase from the 26th value to the 250th value (the fact that the values sometimes decrease can be viewed in many places along the sequence).

(c) MATLAB reports 1 only when the numerical difference between $(3/4)^n$ and $C_1 + (3/4)^n$ with $C_1 = \texttt{9.3378e-20}$ is 0. When this difference is nonzero numerically, i.e. 1.1102e-16 or larger, MATLAB reports 1.0000. So MATLAB reports 1 when a value is exactly 1 in double precision, and 1.0000 when a value is not exactly 1 in double precision (even if this difference is 1.1102e-16, and can be accounted by the limit of finite precision).

(d) Here MATLAB reports 1 for `ratio` and 0 for `ratio_versus_one` again for the first 25 values, but also for 283rd values and higher.

[The homework does not require you to explain this, but the reason that the higher values are 1 and 0 in the experiment with 400 is that $x\{250\}$ is not the (double precision) limit of the sequence, whereas as $x\{400\}$ is; you can see this not (you can see this by typing $x\{400\} - x\{250\}$ and $x\{400\} - x\{399\}$, etc.).

(3) Problem 3 (in the Exercise section) of the handout "Normal and Subnormal Numbers in Double Precision."

Solution:

(a) MATLAB reports $x\{200\}$ as 9.3378e-20.
(b) MATLAB reports $x\{200\}$ as 3.6854e-30, -3.6019e-39, -7.5418e-49 for $r = 3/8, 3/16, 3/32$ respectively. Hence the reported values of $x\{200\}$ descrease in absolute value by roughly 10^{-10} in each successive experiment.
(c) The values of 3^n times the powers of 2 involved fail to be exact (in double precision) when it takes more than 53 bits to represent 3^n (These numbers involved in the experiments are all normal, and so you can expect 53 bits of precision, i.e., "1." followed by 52 bits.) So you should imprecision near the value of n where $3^n \geq 2^{54}$, which is roughly $n = 54\log(2)/\log(3) = 34.07\ldots$ Hence with each successive experiment, which divides the result by an additional 2^n, you would expect of imprecision to be on the order of magnitude (very roughly) by $2^{34.07}$, i.e., 1.8036e+10.
Hence, in rough terms, you'd expect $x\{200\}$, which represents imprecision around $n = 34$, to drop by roughly 10^{10}. [You can't get more precise than this rough estimate, unless you are willing to dig in deeply to the roundoff/truncation errors that are specified in double precision.]

(4) (a) Consider for a sequence $\dots, y_{-1}, y_0, y_1, y_2, \dots$, consider the recurrence relation:

$$y_{n+1} = y_n \quad \forall n \in \mathbb{Z}$$

(i.e., for all integers n). For for any $a \in \mathbb{R}$, what is the unique solution to this recurrence that satisfies the condition $y_0 = a$? Briefly justify your answer.

(b) Consider another recurrence relation:

$$y_{n+2} = 2y_{n+1} - y_n \quad \forall n \in \mathbb{Z}.$$

For any $a, b \in \mathbb{R}$, what is the unique solution to this recurrence that satisfies the condition $y_0 = a$ and $y_1 = a + b$? Briefly justify your answer.

(c) Consider another recurrence relation:

(1)
$$y_{n+3} = 3y_{n+2} - 3y_{n+1} + y_n \quad \forall n \in \mathbb{Z}.$$

Show that the following sequences satisfy this recurrence relation: $y_n = 1$, $\tilde{y}_n = n$, $\hat{y}_n = n^2$.

(d) Show that for any $C_1, C_2, C_3 \in \mathbb{R}$, the sequence

$$y_n = C_1 + C_2 n + C_3 n^2$$

satisfies (1) [you may use the previous part].

(e) If we define for each sequence

$$\mathbf{y} = \{y_n\}_{n\in\mathbb{Z}} = \{\dots, y_{-1}, y_0, y_1, y_2, \dots\}$$

a new sequence $D\mathbf{y}$ (known as the "(forward) difference of $\mathbf{y}$") defined by

$$(D\mathbf{y})_n = y_{n+1} - y_n,$$

we can "apply D twice" to get $D^2\mathbf{y}$, meaning $D(D\mathbf{y})$, given by

$$D(D\mathbf{y}) = (D\mathbf{y})_{n+1} - (D\mathbf{y})_n = \big(y_{n+2} - y_{n+1}\big) - \big(y_{n+1} - y_n\big).$$

Simplify this formula, and then compute a similarly simplified formula for $D^3\mathbf{y}$, meaning $D(D^2\mathbf{y})$. How do your formulas for relate to the previous parts (a)–(c) of this problem?

(f) For any $a, b \in \mathbb{R}$, what is the exact solution of (1) given the conditions $y_0 = a$, $y_1 = a + b$, $y_2 = a + 2b$?

(g) Run the following MATLAB code to test what happens numerically in the previous part for $a = \sqrt{2}$ and $b = \sqrt{7}$:

```
clear

n = 1
sequence_length = 30 * n

x{1}= sqrt(2)
x{2}= sqrt(2) + sqrt(7)
x{3}= sqrt(2) + 2 * sqrt(7)

for i=4:sequence_length, x{i}=3*x{i-1} - 3*x{i-2}+x{i-3}; end

for i=1:30; should_be_zero{i} = x{i*n}-sqrt(2)-(i*n-1)*sqrt(7); end
```

`should_be_zero`

and then run the entire code again (starting from the `clear` at the top) with $n = 10$, $n = 100$, $n = 1000$, and $n = 10000$. Explain why the cell array `should_be_zero` should be zero if MATLAB were computing in "exact" (or "infinite precision") arithmetic.
Roughly speaking, what do you think you are seeing numerically?

(h) Run the above code again for $n = 10000$ (i.e., 10,000), and at the bottom add the lines

```
for i=1:20; rough_effect_of_error{i} = should_be_zero{i} / (i*n-1)^2; end
rough_effect_of_error
```

Then run all of the above code with $n = 100000$ (i.e. 100,000), and $n = 1000000$ (i.e., 1,000,000). Is the description of the results as $C_1 + C_2 n + C_3 n^2$ with $C_1 = \sqrt{2}$, $C_2 = \sqrt{7}$, and C_3 a very small constant seem completely correct for $n = 10^4$? Is there any trend to how the apparent C_3 is behaving in the various entries of `rough_effect_of_error`? What about for $n = 10^5$? What about for $n = 10^6$? [Your answers may or may not be the same for these three values of n.]

Solution:

(a) This recurrence implies that $y_1 = y_0 = a$ and $a = y_0 = y_{-1}$, and similary $y_i = a$ for $i = 2, 3, \ldots$ and $i = -2, -3, \ldots$ (i.e., for all i).

(b) In class we remarked that the general solution is $y_n = C_1 + C_2 n$ (see the last two pages of notes for Feb 5); and hence for $y_0 = a$ we get $C_1 = a$, and $y_1 = a + b$ gives $a + b = C_1 + C_2 = a + C_2$ and hence $C_2 = b$. Hence the general solution is $y_n = a + nb$.
[Alternatively you could check that $y_2 = 2y_1 - y_0 = a + 2b$, $y_3 = 2y_2 - y_1 = a + 3b$, and guess that the general solution is $y_n = a + bn$: to check it you just have to see that $y_0 = a$ and $y_1 = a + b$ (as required), and see that this function of y_n satisfies the recurrence $y_{n+2} = 2y_{n+1} - y_n$.]

(c) For $y_n = 1$: for all $n \in \mathbb{Z}$,

$$y_{n+3} - 3y_{n+2} + 3y_{n+1} - y_n = 1 - 3 \cdot 1 + 3 \cdot 1 - 1 = 0.$$

For $\tilde{y}_n = n$: for all $n \in \mathbb{Z}$,

$$\tilde{y}_{n+3} - 3\tilde{y}_{n+2} + 3\tilde{y}_{n+1} - \tilde{y}_n = (n+3) - 3(n+2) + 3(n+1) - n = 0.$$

For $\hat{y}_n = n^2$: for all $n \in \mathbb{Z}$,

$$\hat{y}_{n+3} - 3\hat{y}_{n+2} + 3\hat{y}_{n+1} - \hat{y}_n = (n+3)^2 - 3(n+2)^2 + 3(n+1)^2 - n^2$$

$$= (n^2 + 6n + 9) - 3(n^2 + 4n + 4) + 3(n^2 + 2n + 1) - n^2 = 0.$$

(d) If

$$y_n = C_1 + C_2 n + C_3 n^2$$

then

$$y_{n+3} - 3y_{n+2} + 3y_{n+1} - y_n$$

$$= C_1\Big(1{-}3{\cdot}1{+}3{\cdot}1{-}1\Big){+}C_2\Big((n{+}3){-}3(n{+}2){+}3(n{+}1){-}n\Big){+}C_3\Big((n{+}3)^2{-}3(n{+}2)^2{+}3(n{+}1)^2{-}n^2\Big)$$

$$= C_1 \cdot 0 + C_2 \cdot 0 + C_3 \cdot 0 = 0$$

by the calculations in part (b).

[Alternatively, you can say that the equation

$$y_{n+3} - 3y_{n+2} + 3y_{n+1} - y_n = 0$$

is a *linear equation*, in that for any two sequences $\{u_n\}_{n\in\mathbb{Z}}$ and $\{v_n\}_{n\in\mathbb{Z}}$ satisfy then equation, then for any reals α, β, the sequence

$$x_n = \alpha u_n + \beta v_n$$

satisfies the equation, since

$$x_{n+3} - 3x_{n+2} + 3x_{n+1} - x_n$$

$$= \alpha\big(u_{n+3} - 3u_{n+2} + 3u_{n+1} - u_n\big) + \beta\big(v_{n+3} - 3v_{n+2} + 3v_{n+1} - v_n\big) = \alpha\cdot 0 + \beta\cdot 0 = 0.$$

It follows that any *linear combination* (sometimes called *superposition*) of solutions to this equation yields another solution. So since $1, n, n^2$ are solutions, so is $C_1 + C_2 n + C_3 n^2$.]

(e)

$$D^2\mathbf{y} = \big(y_{n+2} - y_{n+1}\big) - \big(y_{n+1} - y_n\big) = y_{n+2} - 2y_{n+1} - y_n,$$

and

$$D^3\mathbf{y} = D(D^2\mathbf{y}) = D\big(y_{n+2} - 2y_{n+1} - y_n\big)$$

$$= \big(y_{n+3} - 2y_{n+2} - y_{n+1}\big) - \big(y_{n+2} - 2y_{n+1} - y_n\big) = y_{n+3} - 3y_{n+2} + 3y_{n+1} - y_n$$

Parts (a)–(c) study the equations $D\mathbf{y} = 0$, $D^2\mathbf{y} = 0$, and $D^3\mathbf{y} = 0$.

(f) For $y_n = C_1 + C_2 n + C_3 n^2$, the equation $y_0 = a$ implies $C_1 = a$, and so $y_n = a + C_2 n + C_3 n^2$. The equations $y_1 = a + b$ and $y_2 = a + 2b$ then imply

$$a + C_2 + C_3 = a + b, \quad a + 2C_2 + 4C_3 = a + 2b,$$

so

$$C_2 + C_3 = b, \quad 2C_2 + 4C_3 = 2b,$$

and so $C_3 = 0$ and $C_2 = b$. Hence the exact solution is $y_n = a + bn$. [Alternatively, you can notice that from part (b), $y_n = a + bn$ fits the "initial" conditions $y_0 = a$, $y_1 = a+b$, and $y_2 = a+2b$, and that $a+bn$ (being of the form $C_1 + C_2 n + C_3 n^2$) satisfies (1).]

(g) In exact arithmetic, $x\{j\} = a + (j-1)b$, with $a = \sqrt{2}$ and $b = \sqrt{7}$ (don't forget the -1 in $j-1$, which is there since the cell array `x` begins in $x\{1\}$). Hence `should_be_zero` is 0 in exact arithmetic. Numerically you are seeing the effects of finite precision, and likely you are picking up a $C_3 n^2$ term where C_3 is small but nonzero. [This is the analog of what we've seen in three-term recurrences in earlier homework; however, as the next part shows, this is only roughly true.]

(h) The `rough_effect_of_error` is trying to find the value of C_3 if one models $x\{j\}$ as $a + (j-1)b + (j-1)^2 C_3$.

The $n = 10,000$ experiment gives C_3 ranging as small as 1.4681e-13 and -1.6657e-13 and as large as 9.5509e-12, but the C_3 value is not constant and has no particular pattern; since the e-13 values are for $j = n-1, 2n-1, \ldots, 6n-1$ and the e-12 values are for larger j, you could also say that C_3 may be *somewhat* increasing in j (although there is no clear pattern).

The $n = 100,000$ experiment shows a clearer pattern of increase in C_3, in that C_3 mostly increases for every additional n steps (starting at 6.6761e-12 and ending around 4.8926e-11) although C_3 does not increase at every step (e.g., at $j = n-1$ to $2n-1$, and it decreases between $6n-1$ to $9n-1$).

The $n = 1,000,000$ experiment shows a strict increase in C_3 until the very last step, from $j = 19n-1$ to $20n-1$, although before $19n-1$ some of the increases are quite small. [The first step is from 2.3017e-11 (at $n-1$) to 4.8926e-11 (at $2n-1$), but only reaches 8.1274e-10 at $19n-1$ and drops to 8.0879e-10 at $20n-1$.]

From n ranging from 10^4 to 10^5 and 10^6, the "observed C_3" is getting generally larger, but there is no simple pattern (e.g., strict increase) in the C_3. The model of this sequence by $a + (j-1)b + (j-1)^2 C_3$ is much more subtle than the models of $C_1 r_1^n + C_2 r_2^n$ for three-term recurrences in earlier homework.

Department of Computer Science, University of British Columbia, Vancouver, BC V6T 1Z4, CANADA.

E-mail address: jf@cs.ubc.ca

URL: http://www.cs.ubc.ca/~jf