Math 4650 Homework #2 Solutions

(1.3 #3) The Maclaurin series for the arctangent function converges for $-1 < x \leq 1$ and is given by

$$\arctan x = \lim_{n\to\infty} P_n(x) = \lim_{n\to\infty} \sum_{i=1}^{n} (-1)^{i+1} \frac{x^{2i-1}}{2i-1}.$$

a. Use the fact that $\tan \pi/4 = 1$ to determine the number n of terms of the series that need to be summed to ensure that $|4P_n(1) - \pi| < 10^{-3}$.

Solution: Because $P_n(1)$ is an alternating series, we know the last two approximations $P_n(1)$ and $P_{n-1}(1)$ always bound the true value $\frac{\pi}{4}$. Therefore, we know

$$|P_n(1) - \tfrac{\pi}{4}| < |P_n(1) - P_{n-1}(1)| \leq \frac{1}{2n-1},$$

which implies

$$|4P_n(1) - \pi| < \frac{4}{2n-1}.$$

To be guaranteed that this is less than 10^{-3}, we need to choose n large enough that $\frac{4}{2n-1} \leq \frac{1}{1000}$, i.e., $n \geq 4001/2$. Hence $n = 2001$ will work.

b. The C++ programming language requires the value of π to be within 10^{-10}. How many terms of the series would we need to sum to obtain this degree of accuracy?

Solution: We would need $\frac{4}{2n-1} \leq 10^{-10}$, so that $n \geq 2 \times 10^{10} + 1$.

(1.3 #7) Find the rates of convergence of the following functions as $h \to 0$.

a. $\lim\limits_{h\to 0} \dfrac{\sin h}{h} = 1$

Solution: The Maclaurin series for $\sin h$ is $\sin h = h - \frac{h^3}{6} + \cdots$, so that we have

$$\frac{\sin h}{h} - 1 = 1 - \frac{h^2}{6} + \cdots - 1 = -\frac{h^2}{6} + \cdots = O(h^2).$$

b. $\lim\limits_{h\to 0} \dfrac{1-\cos h}{h} = 0$

Solution: The Maclaurin series for $\cos h$ is $\cos h = 1 - \frac{h^2}{2} + \cdots$, so that

$$\frac{1-\cos h}{h} - 0 = \frac{h}{2} + \cdots = O(h).$$

c. $\lim\limits_{h\to 0} \dfrac{\sin h - h\cos h}{h} = 0$

Solution: Combining the Maclaurin series for $\sin h$ and $\cos h$, we get

$$\sin h - h\cos h = \left(h - \frac{h^3}{6} + \frac{h^5}{120} - \cdots\right) - h\left(1 - \frac{h^2}{2} + \frac{h^4}{24} - \cdots\right) = \frac{h^3}{3} - \frac{h^5}{30} + \cdots,$$

and therefore

$$\frac{\sin h - h\cos h}{h} - 0 = \frac{h^2}{3} + \cdots = O(h^2).$$

d. $\lim_{h\to 0} \frac{1-e^h}{h} = -1$

Solution: The Maclaurin series for e^h is $e^h = 1 + h + \frac{h^2}{2} + \cdots$, so that

$$\frac{1-e^h}{h} + 1 = \frac{1-1-h-h^2/2-\cdots}{h} + 1 = -\frac{h}{2} - \cdots = O(h).$$

(1.3 #8) a. How many multiplications and additions are required to determine a sum of the form

$$\sum_{i=1}^{n}\sum_{j=1}^{i} a_i b_j?$$

Solution: To make it clearer, let's write $c_{ij} = a_i b_j$ and $D_i = \sum_{j=1}^{i} c_{ij}$, so that the final answer is $E = \sum_{i=1}^{n} D_i$.

To get *each* c_{ij} we need one multiplication, and the total number of c_{ij} terms is

$$\sum_{i=1}^{n}\sum_{j=1}^{i}(1) = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

So there are $n(n+1)/2$ multiplications.

To get each D_i we have to do $i-1$ additions. Then to get the total E, we have to do another $n-1$ additions, so that the total number of additions is

$$\sum_{i=1}^{n}(i-1) + (n-1) = n(n+1)/2 - n + (n-1) = (n^2+n-2)/2 = (n+2)(n-1)/2.$$

b. Modify the sum in part (a) to an equivalent form that reduces the number of computations.

Solution: We can factor out the a_i from the j-sum, which gives

$$E = \sum_{i=1}^{n} a_i \sum_{j=1}^{i} b_j.$$

Write this in the form $E = \sum_{i=1}^{n} a_i F_i$, where $F_i = \sum_{j=1}^{i} b_j$. Then to obtain all the F_i terms, we need to do $i-1$ additions for each i, which gives a total of $\sum_{i=1}^{n}(i-1) = n(n-1)/2$ additions for all F_i terms. Then we do n multiplications. Then we do another sum of $n-1$ terms, giving us a total of

$$n(n-1)/2 + n - 1 = (n+2)(n-1)/2 \text{ additions and } n \text{ multiplications.}$$

This method reduces the number of multiplications but does not save any additions. A more clever way is to obtain each F_i not by summing up $F_i = \sum_{j=1}^{i} b_j$ each time, but rather by using the recursion $F_{i+1} = F_i + b_{i+1}$ with $F_1 = b_1$. Then we will use $n-1$ additions to get all the F_i, another n multiplications to get the products $a_i F_i$, and finally a sum of $n-1$ terms to get the final answer E. This gives a total of

$$2(n-1) \text{ additions and } n \text{ multiplications.}$$

(2.1 #6) Use the Bisection method to find solutions, accurate to within 10^{-5}, for the following problems. *Use mathematical software and the bisection program from class.*

Solution: We use the programs from class, with the tolerance Tol $= 0.00001$. By default MATLAB only prints four digits.

a. $3x - e^x = 0$ for $1 \le x \le 2$

Solution: The program from class, used in MATLAB, gives 1.5121. The similar program in Mathematica gives 1.51214.

b. $2x + 3\cos x - e^x = 0$ for $0 \le x \le 1$

Solution: The program in MATLAB gives 1.0000. Mathematica gives 0.999992. Notice that this is wrong! The function is certainly not even close to zero at $x = 1$; we have $f(1) \approx 0.903$. The problem is that our program does not test whether the initial conditions actually satisfy $f(a)f(b) < 0$, and the bounds in this problem do not satisfy that. (We have $f(0) = 2$.)

c. $x^2 - 4x + 4 - \ln x = 0$ for $1 \le x \le 2$ and $2 \le x \le 4$

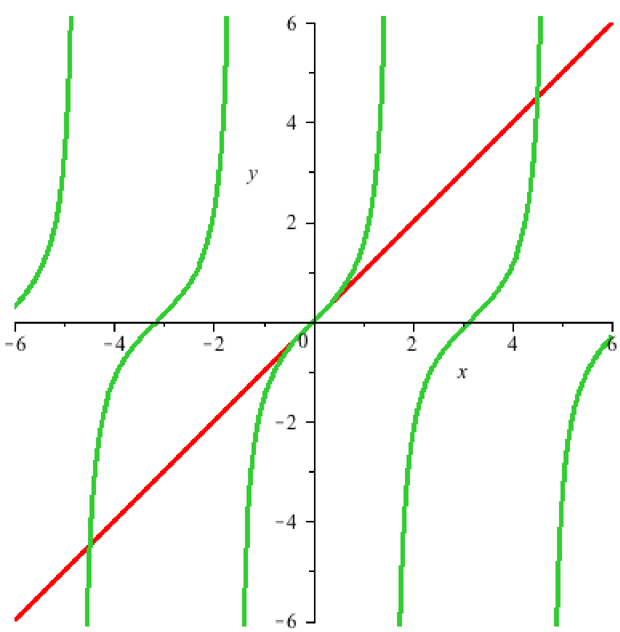
Solution: MATLAB gives 1.4124 and 3.0571. Mathematica gives 1.41239 and 3.05711.

d. $x + 1 - 2\sin \pi x = 0$ for $0 \le x \le 0.5$ and $0.5 \le x \le 1$

Solution: MATLAB gives 0.2060 and 0.6820. Mathematica gives 0.206032 and 0.681969.

(2.1 #8) a. Sketch the graphs of $y = x$ and $y = \tan x$ (by hand).

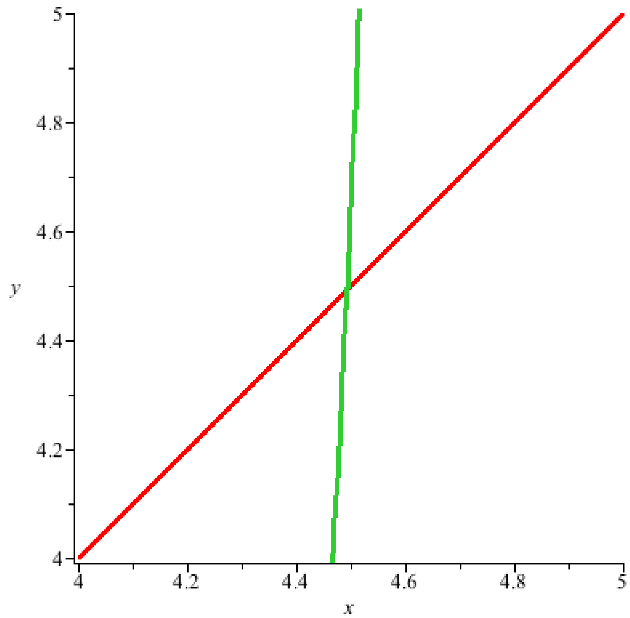
Solution: Here's a graph generated by Maple on the domain $-6 \le x \le 6$.



b. Use the Bisection method to find an approximation to within 10^{-5} to the first positive value of x with $x = \tan x$. *Use mathematical software and the bisection program from class.*

Solution: We want to find a positive root of the function $f(x) = \tan x - x$. From the big graph, it looks like this happens around 4.5 or so. Of course we have to watch out for the asymptote of $\tan x$ (which happens at $3\pi/2 \approx 4.7$), since we certainly do not want the function being evaluated there (which would lead to overflow).

So let's zoom in and plot just for $4 \le x \le 5$.



Now it's clear that there will be a fixed point between 4.4 and 4.6, so we run bisection with these as our initial a and b. The MATLAB program spits out 4.4934, while the Mathematica program gives 4.49341.

(2.1 #16) Let $f(x) = (x-1)^{10}$, $p = 1$, and $p_n = 1 + 1/n$. Show that $|f(p_n)| < 10^{-3}$ whenever $n > 1$ but that $|p - p_n| < 10^{-3}$ requires that $n > 1000$.

Solution: This is easy: we have $f(p_n) = \frac{1}{n^{10}}$, so that $f(p_n) \le f(p_2)$ for every n. And since $f(p_2) = \frac{1}{2^{10}} = \frac{1}{1024} < 10^{-3}$, we're done.

On the other hand, $|p - p_n| = \frac{1}{n}$, and to make this less than $\frac{1}{1000}$ we need $n > 1000$.

The point of this problem is that the y-values may be very close to zero without the x-values actually being close to the root, which is why it's important in root-finding to not just check the y-values as a stopping criterion.

(2.2 #1) Use algebraic manipulation to show that each of the following functions has a fixed point at p precisely when $f(p) = 0$, where $f(x) = x^4 + 2x^2 - x - 3$.

a. $g_1(x) = (3 + x - 2x^2)^{1/4}$

Solution: Just set $g_1(p) = p$ and simplify.

$$\begin{aligned}(3 + p - 2p^2)^{1/4} &= p \\ 3 + p - 2p^2 &= p^4 \\ 0 &= p^4 + 2p^2 - p - 3 = f(p).\end{aligned}$$

If p is positive, we can reverse this computation to see that $f(p) = 0$ implies that $g_1(p) = p$. However there is a root $p = -0.876\ldots$ for which $g_1(p) \neq p$, so the problem is not quite correct as written.

b. $g_2(x) = \left(\dfrac{x + 3 - x^4}{2}\right)^{1/2}$

Solution: Again set $g_2(p) = p$, and we get

$$\begin{aligned}\left(\frac{p + 3 - p^4}{2}\right)^{1/2} &= p \\ \frac{p + 3 - p^4}{2} &= p^2 \\ p + 3 - p^4 &= 2p^2 \\ 0 &= p^4 + 2p^2 - p - 3 = f(p).\end{aligned}$$

Again the steps can be reversed if $p > 0$, but there is a root $p < 0$ of f that is not a fixed point of g_2.

c. $g_3(x) = \left(\dfrac{x + 3}{x^2 + 2}\right)^{1/2}$

Solution: Again set $g_3(p) = p$.

$$\begin{aligned}\left(\frac{p + 3}{p^2 + 2}\right)^{1/2} &= p \\ \frac{p + 3}{p^2 + 2} &= p^2 \\ p + 3 &= p^4 + 2p^2 \\ 0 &= p^4 + 2p^2 - p - 3 = f(p).\end{aligned}$$

Once again, we can reverse all the steps if $p > 0$, but not if $p < 0$.

d. $g_4(x) = \dfrac{3x^4 + 2x^2 + 3}{4x^3 + 4x - 1}$

Solution: Set $g_4(p) = p$. Then

$$\begin{aligned}\frac{3p^4 + 2p^2 + 3}{4p^3 + 4p - 1} &= p \\ 3p^4 + 2p^2 + 3 &= 4p^4 + 4p^2 - p \\ 0 &= p^4 + 2p^2 - p - 3 = f(p).\end{aligned}$$

In this case all the steps actually can be reversed, and so $g_4(p) = p$ is equivalent to $f(p) = 0$.

(2.2 #2) a. Perform four iterations, if possible, on each of the functions g defined in Exercise 1. Let $p_0 = 1$ and $p_{n+1} = g(p_n)$, for $n = 0, 1, 2, 3$.

Solution:

	g_1	g_2	g_3	g_4
p_1	1.189207	1.224745	1.154701	1.142857
p_2	1.080058	0.993666	1.116427	1.124482
p_3	1.149671	1.228569	1.126052	1.124123
p_4	1.107821	0.987506	1.123639	1.124123

b. Which function do you think gives the best approximation to the solution?

Solution: The solution is 1.124123..., and g_4 gives the best approximation (accurate to 7 digits already at p_3).

The reason is that we can compute $g_i'(p)$ for each i at the fixed point p, and we get

$$\begin{aligned} g_1'(p) &\approx -0.6154 \\ g_2'(p) &\approx -1.1041 \\ g_3'(p) &\approx -0.2509 \\ g_4'(p) &= 0. \end{aligned}$$

So iteration using g_2 will definitely not converge to the fixed point, iteration by g_1 will converge but rather slowly (slower than bisection), iteration by g_3 will converge fairly quickly, and iteration by g_4 will converge far faster than any of the others. (Notice that g_4 is just Newton's method.)