ICSI 401 Midterm

James Oswald

October 9, 2020

1. (10 points) Which of the following is a valid expression for the function $f(x) = e^x x - x$ from Taylor's remainder theorem?

   A. $x^2 + \frac{x^3}{2} + (4e^0 + 0e^0)\xi^4/4!$, where ξ is some real number between 0 and x.

   B. $x^2 + \frac{x^3}{2} + (4e^0 + 0e^0)x^4/4!$.

   C. $x^2 + \frac{x^3}{2} + (4e^\xi + \xi e^\xi)x^4/4!$ where ξ is some real number between 0 and x.

   D. None of the above. Taylor's remainder theorem doesn't apply because $f(x)$ isn't differentiable everywhere.

   Hint: I've listed some derivatives of f: $f'(x) = -1 + xe^x + e^x; f''(x) = 2e^x; f'''(x) = 3e^x + xe^x; f^{(4)}(x) = 4e^x + xe^x$

   It appears we're computing the third order Taylor polynomial $P_3(x)$ for our function $f(x)$, thus our remainder $R_3(x)$ can be determined using the remainder formula $R_n(x) = \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1}$ which in our case for $R_3(x)$ will be $R_3(x) = \frac{f^{(4)}(\xi)x^4}{4!}$ where ξ is some real number between 0 and x. Expanding this out we get $\frac{(4e^\xi+\xi e^\xi)x^4}{4!}$, which we immediately see is answer choice C.

2. (10 points) The following code is meant to find a root of the input function F inside the interval $[a, b]$.

```
%
% Implements multisection search to find a root
% of F in the interval [a, b]. Runs for k iterations.
%
function x = multisection_search(F, a, b, k):
    if (F(a) <= 0 && F(b) > 0)
        direction = 1
    elseif (F(a) > 0 && F(b) <= 0)
        direction = -1
    end
    for i = 1:k
        z = a/10 + 9*b/10
        if (F(z) == 0)
            x = z
            return(x)
        end
        if (F(z) * direction < 0)
            a = z
        elseif (F(z) * direction > 0)
            b = z
        end
    end
    x = z
end
```

Assume infinite precision, and assume that F is a continuous function.

(a) Is z guaranteed to converge to a root of F in the interval $[a, b]$ if $F(a)$ and $F(b)$ have opposite signs and F is a continuous function? If so, what theorem guarantees this?

Yes, z is guaranteed to converge to a root of F in the interval $[a, b]$ if $F(a)$ and $F(b)$ have opposite signs and F is a continuous function. This is guaranteed by the Intermediate Value Theorem which states that if f is a continuous function on the interval $[a, b]$ and u is a number between $f(a)$ and $f(b)$ and $\exists c \in (a, b) : f(c) = u$. In our case since $F(a)$ and $F(b)$ have opposite signs, one will be > 0 and one < 0, and we can pick $u = 0$. Thus the IVT guarantees there exists a z in the interval $[a, b]$ such that $F(z) = 0$, or in other words, z will always have a root to converge to under these conditions.

(b) What is the maximum possible length of the new interval, in terms of $b - a$, after a single iteration of the i loop?

There are two possibilities for the length of the new interval after a single iteration of the loop, either we set a to z or b to z.

Case 1. Set $a = z$ new interval is $[a/10 + 9 * b/10, b]$, length of the new interval is $b - (a/10 + 9b/10) = \frac{1}{10}(b - a)$.

Case 2. Set $b = z$ new interval is $[a, a/10 + 9 * b/10]$, length of the new interval is $(a/10 + 9 * b/10) - a = \frac{9}{10}(b - a)$

We see that the maximum length of of an interval in terms of $b - a$ after the first iteration will be the case when we set $b = z$ and obtain an interval of length $\frac{9}{10}(b - a)$.

3. (10 points)

(a) Derive the Newton update equation (i.e., x_{k+1} in terms of x_k) for the function $F(x) = \sin(x)$.

The Newton iteration equation for a a function $F(x)$ is $x_{k+1} = x_k - \frac{F(x_k)}{F'(x_k)}$. Applying this to our function $F(x) = \sin(x)$ we get $x_{k+1} = x_k - \frac{\sin(x_k)}{\cos(x_k)}$ or even more concisely, $x_{k+1} = x_k - \tan(x_k)$.

(b) Suppose that you choose an initial point $x_0 \in (\pi, 3\pi/2)$ sufficiently close to π. Can convergence of Newton's method be guaranteed? Why, or why not?

Yes, convergence of Newton's method is guaranteed due to the Newton's Method Convergence Theorem which states that if a function $f(x)$ has a continuous second derivative AND the derivative of $f(x)$ is non-zero at a root x_* AND x_0 is sufficiently close to to a root x_* then Newton's method is guaranteed to converge to the root x_*. We can show that we meet all of these criteria and that convergence will therefore be guaranteed. In our case $F(x) = \sin(x)$ does indeed have a continuous second derivative, $F''(x) = -\sin(x)$, its derivative is non-zero at the root $F'(\pi) = -1 \neq 0$, and we have been told we can choose an initial point $x_0 \in (\pi, 3\pi/2)$ sufficiently close to π which is our root. Therefore, Newton's method is guaranteed to converge to π under these conditions.

4. (10 points) Consider finding fixed points of the function $F(x) = 2xe^x + x$.

   (a) For which values of x is $F(x)$ a contraction?

   We know that a function F will be a contraction when $|F'(x)| < 1$. Thus, we begin by finding the absolute value of the derivative of $F(x)$. $|F'(x)| = |2xe^x + 2e^x + 1|$. We now want to find for what points $|2xe^x + 2e^x + 1| < 1$. We observe $F'(x)$ is always positive, so $|F'(x)| = F'(x)$ and thus we simplify the problem to finding the points for which $2xe^x + 2e^x + 1 < 1$. We can find the local minimum of $F'(x)$ by solving for where its derivative $F''(x) = 0$ and using a second derivative to verify it is a minimum. $F''(x) = 2xe^x + 4e^x$, now solve $2xe^x + 4e^x = 0$ and get $x = -2$. We check $F'(-2)$ and observe $2(-2)e^{-2} + 2e^{-2} + 1 = -2e^-2 + 1 \approx 0.729 < 1$. We now know there is an interval for which $|F'(x)| < 1$ and that our point $F'(-2)$ is the minimum inside of it. To find the bounds of the interval, I just solve for $2xe^x + 2e^x + 1 = 1 \Leftrightarrow 2xe^x + 2e^x = 0 \Leftrightarrow 2xe^x = -2e^x \Leftrightarrow x = -1$. We use this bound and our previous finding of $F'(-2)$ we know that for any $x < -1$ that $|F'(x)| < 1$. Therefore we know that $F(x)$ is a contraction for all $x < -1$

   (b) Will fixed point iteration starting with some point $x_0 \in [-1/2, 1]$ converge to a fixed point of F in that interval? If so, why? If not, why doesn't Banach's fixed point theorem apply?

   No, Banach's fixed point theorem does not apply here because we have not met its first hypothesis. The interval provided for x_0, $[-1/2, 1]$ is not a contraction on $F(x)$. As we proved in part (a) of this problem $F(x)$ will only be a contraction when $x < -1$ and the interval $[-1/2, 1]$ clearly does not meet this, therefore Banach's fixed point theorem would not apply.

5. (10 points) Consider the function $F(x) = \cot(x) = \frac{\cos(x)}{\sin(x)}$.

   (a) Derive an expression for the relative condition number $\kappa(x)$ of $F(x)$.

   We know that the relative condition number is determined via the formula $\kappa(x) = \left|\frac{x \cdot F'(x)}{F(x)}\right|$. Thus we compute $F'(x) = \frac{1}{\sin(x)^2}$ and simplify $\kappa(x)$ as follows:

$$\kappa(x) = \left|\frac{x \cdot F'(x)}{F(x)}\right| \qquad \text{The formula}$$

$$\kappa(x) = \left|\frac{x \cdot \frac{1}{\sin(x)^2}}{\frac{\cos(x)}{\sin(x)}}\right| \qquad \text{Plugging in}$$

$$\kappa(x) = \left|x \cdot \frac{1}{\sin(x)^2} \frac{\sin(x)}{\cos(x)}\right| \qquad \text{Simplify}$$

$$\kappa(x) = \left|\frac{x}{\sin(x)\cos(x)}\right| \qquad \text{Canceling out}$$

   Thus we have derived an expression for the relative condition number $\kappa(x)$ of $F(x)$.

   (b) Determine all points $x_* \in \mathbb{R}$ near which $F(x)$ is ill-conditioned, in the sense that $\lim_{x \to x_*} \kappa(x) = \infty$

   Since we have $\kappa(x) = \left|\frac{x}{\sin(x)\cos(x)}\right|$ and $\lim_{x \to x_*} \kappa(x) = \infty$ we know this implies that as $\kappa(x) \to \infty$ we have $\sin(x)\cos(x) \to 0$. This will occur when $\sin(x) = 0$ and $\cos(x) = 0$ respectively at $k \cdot \pi$ and $k \cdot \pi + \frac{\pi}{2}$ where $k \in \mathbb{Z}$. We must also take into account that at $x = 0$ we will have $\frac{0}{0}$ and exclude this from our set of ill-conditioned points. Therefore in conclusion have the set of points when $x = k \cdot \pi$ with $k \neq 0$ and $x = k \cdot \pi + \frac{\pi}{2}$.

6. (10 points) If a function $F(x) = \Theta(x^2)$ as $x \to \infty$, is it always true that $3F(x) = \Theta(x^2)$? If so, explain why. If not, give a counterexample.
Hint: $F(x) = \Theta(G(x))$ as $x \to \infty$ means that there exist positive constants C_1, C_2 such that $0 < C_1 \le \left|\frac{F(x)}{G(x)}\right| < C_2$ for all large enough x.

Yes this is true. To prove it we use the definition provided in the hint.

$$F(x) = \Theta(x^2) \text{ as } x \to \infty$$
$$\Rightarrow \exists C_1, C_2 \in \mathbb{R}^+ \text{ s.t. } 0 < C_1 \le \left|\frac{F(x)}{x^2}\right| < C_2$$

From here we see that we can multiply all members of the inequality by 3 and it will still hold true.

$$\Rightarrow \exists C_1, C_2 \in \mathbb{R}^+ \text{ s.t. } (3)0 < 3C_1 \le 3\left|\frac{F(x)}{x^2}\right| < 3C_2$$

by multiplying C_1 and C_2 by 3 we have created two new positive constants which I will call C_3 and C_4 respectively with $C_3 = 3C_1$ and $C_4 = 3C_2$. We can also use properties of the absolute value to move the 3 inside and end up with:

$$\Rightarrow \exists C_3, C_4 \in \mathbb{R}^+ \text{ s.t. } 0 < C_3 \le \left|\frac{3F(x)}{x^2}\right| < C_4$$

Which is the definition of:

$$\Rightarrow 3F(x) = \Theta(x^2) \text{ as } x \to \infty$$

Thus we are able to derive that:
If $F(x) = \Theta(x^2)$ as $x \to \infty$ then $3F(x) = \Theta(x^2)$ as well.

7. (10 points) Consider a floating point number system with 8 mantissa bits and 8 exponent bits that works as follows: if a number N is$+(1.x_1x_2x_3...x_8)_2\times 2^{(y_7y_6y_5...y_0)_2}$, then its floating point representation is given $\boxed{1}\boxed{\mathrm{x}_1x_2...x_8}\boxed{\mathrm{y}_7y_6...y_0}$ where the first bit represents the positive sign. Note that the initial 1 in the mantissa is not represented explicitly, so we are using a hidden bit representation.

   (a) What is the floating point representation of the number 3.75? Write your answer in binary, clearly giving the mantissa and exponent bit strings.

   We begin by converting the whole portion of the number (3) to binary $3_{10} \to 11_2$. We then convert the decimal portion (.75) to binary as well $.75_{10} \to .11_2$ and join these to obtain a binary representation of $3.75_{10} \to (11.11)_2$. We then convert this to binary scientific notation $(1.111)_2 \times 2^{1_2}$, and with this we have now obtained our mantissa as 111_2 and our exponent as 1_2. Finally we place the bits of these apropriately in the representation as defined by $\boxed{1}\boxed{\mathrm{x}_1x_2...x_8}\boxed{\mathrm{y}_7y_6...y_0}$. Doing this we get that the floating point representation of the number 3.75 in this system is $\boxed{1}\boxed{11100000}\boxed{00000001}$

   (b) What is the smallest number larger than 3.75 that is representable in the floating point number system described above? Write your answer in the binary floating point format, clearly giving the mantissa and exponent bit strings.

   The next smallest number larger then 3.75 that is representable in the floating point number system described is simply the representation of 3.75 with 1 added to the mantissa, which would give us $\boxed{1}\boxed{11100001}\boxed{00000001}$