Final #1 - Questions and Solutions

1. a. Let x_0 be a real number. Define a sequence $(x_n)_{n=1}^{\infty}$ by:

$$x_{n+1} = \sin x_n, \qquad n = 0, 1, 2, \dots \ .$$

Using the fact that $|\sin x| < |x|$ for every $x \neq 0$, show that the sequence x_n converges to the unique fixed point $x = 0$ of the function $g(x) = \sin x$.

We first note that, since $x_1 = \sin x_0$, we must have $|x_1| \leq 1$. If $x_1 = 0$, then all subsequent terms of the sequence are 0. Let us deal with the case $x_1 > 0$. (The case $x_1 < 0$ is analogous.) A straightforward induction shows that in this case the sequence is decreasing, and all its terms are positive. Hence it converges to a finite limit ξ. As g is continuous, ξ is a fixed point of g. Since $|\sin x| < |x|$ for every $x \neq 0$, the unique fixed point is 0, whence $\xi = 0$.

b. Is the convergence linear, slower or faster?

Let e_n be the distance from the fixed point after n steps, namely $e_n = \xi - x_n$. Since $\frac{\sin x}{x} \underset{x \to 0}{\longrightarrow} 1$, we have

$$\frac{e_{n+1}}{e_n} \underset{n \to \infty}{\longrightarrow} 1 \ .$$

Thus, the convergence is sub-linear.

c. Write MATLAB code designed to calculate (approximately) the fixed point in question. Given an initial value a for x_0, the x_n's are consecutively calculated up to the first index L for which $x_{L+1} = x_L$.

```
function [x, niter] = prob1c(a)
niter = 0;
x = [a];
tmp = a;
ftmp = sin(tmp);
while tmp ~= ftmp
  x = [x ftmp];
  tmp = ftmp;
  ftmp = sin(tmp);
  niter = niter + 1;
end
```

d. Suppose the above MATLAB code is run on a system working with the IEEE standard, double prcision and round-to-nearest mode, and the initial value of x_0 lies within the interval $\left[\frac{\pi}{4}, \frac{\pi}{2}\right]$. Find (the order of magnitude of) x_L.

In this case the sequence x_n decreases to zero. On the computer, it will decrease as long as x_k and $\sin x_k$ are distinguishable. Since, for $x_k \approx 0$ we $\sin x_k \approx x_k - \frac{x_k^3}{6}$, we have to find when $x_k - \frac{x_k^3}{6}$ is rounded to x_k itself and when to a number strictly smaller. Let $x_k = m \cdot 2^E$ with $1 \le m \le 2$. Then $\frac{x_k^3}{6}$ will be big enough to change x_k exactly if

$$\frac{x_k^3}{6} \ge 2^{-54} * 2^E .$$

To find the order of magnitude of the required threshold, we replace 2^E on the right hand side by x_k to obtain the condition

$$\frac{x_k^3}{6} \ge 2^{-54} x_k .$$

Equivalently, we see that $\sin x_k$ is rounded to a number smaller than x_k (approximately) if $x_k \ge \sqrt{6} \cdot 2^{-27}$.

The above calculations actually suffice to obtain the precise result. In fact, it hints that the threshold is somehwere between 2^{-26} and 2^{-25}. For x_k in this region, the exact condition is

$$\frac{x_k^3}{6} \ge \frac{1}{2} \cdot 2^{-52} \cdot 2^{-26} ,$$

which gives

$$x_k \ge \sqrt[3]{3} \cdot 2^{-26} .$$

Thus $x_L = \sqrt[3]{3} \cdot 2^{-26}$.

2. Let $(x_n)_{n=1}^{\infty}$ be a sequence of distinct points on the real line, and let $f(x) = e^x \cos x$. For each n, let P_n denote the interpolation polynomial of degree not exceeding n, coinciding with f at the points $x_0, x_1, \ldots, x_n$.

a. Prove that we do not necessarily have $P_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ for every $x \in \mathbf{R}$.

In fact, let $x_n = n\pi + \frac{\pi}{2}$. Since f vanishes at all the points x_n, each P_n is the zero polynomial, and therefore the required convergence convergence does not hold at points which are not zeros of f.

b. Show that if $(x_n)_{n=1}^{\infty}$ then $P_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ for every $x \in \mathbf{R}$; moreover, the convergence is uniform on any finite interval.

One shows easily by induction that for all n we have

$$f^{(n)}(x) = a_n e^x \cos x + b_n e^x \sin x$$

for appropriately chosen constants $a_n,\ b_n$.

It is possible, in various ways, to give explicit expressions for a_n and b_n. For example, differentiating the formula for $f^{(n)}(x)$, we easily obtain the following recursion:

$$\begin{aligned} a_{n+1} &= a_n + b_n, \\ b_{n+1} &= -a_n + b_n \ . \end{aligned} \tag{1}$$

As the recursion is linear, and with constant coefficients, it is possible to solve explicitly for a_n and b_n. Another possibility is by starting with:

$$f^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} e^{x^{(k)}} \sin^{(n-k)} x = \sum_{k=0}^{n} \binom{n}{k} e^{x} \sin^{(n-k)} x \ . \tag{2}$$

(Still another option is by rewriting the function in the form

$$f(x) = \frac{1}{2}\left(e^{(1+i)x} + e^{(1-i)x}\right) \ ,$$

and differentiating as a complex function.)

In any case, the important thing is to observe that $|f^{(n)}(x)|$ grows at most exponentially fast as a function of n. For example, (1) gives easily by induction that $|a_n| + |b_n| < 2^n$. (The induction step is

$$|a_{n+1}| + |b_{n+1}| = |a_n + b_n| + |-a_n + b_n| \leq 2|a_n| + 2|b_n| \ .)$$

Alternatively, (2) implies

$$|f^{(n)}(x)| \leq \sum_{k=0}^{n} \binom{n}{k} e^{x} = 2^n e^x \ .$$

Let $[a, b]$ be any finite interval. Without loss of generality we may assume the interval to contain all points (x_n). Then for any $x \in [a, b]$:

$$\begin{aligned} |E_n(x)| &= \left| \frac{(x - x_0)(x - x_1)\dots(x - x_n)}{(n+1)!} \right| \cdot \left| f^{(n+1)}(c) \right| \\ &\leq \frac{(b-a)^{n+1}}{(n+1)!} \cdot 2^{n+1} e^{\max\{a,b\}} \underset{n\to\infty}{\longrightarrow} 0 \ . \end{aligned}$$

3. Let $(x_1, y_1), \dots (x_n, y_n)$ be n data points. The least-squares line corresponding to these points is known to be $y = -2x + 5$. A point (x_{n+1}, y_{n+1}) is added. Formulate and prove a (simple) necessary and sufficient condition on (x_{n+1}, y_{n+1}) for the least-squares line corresponding to all $n + 1$ data points to still be $y = -2x + 5$.

The condition is $y_{n+1} = -2x_{n+1} + 5$. Let us prove it.

Sufficiency: Suppose $y_{n+1} = -2x_{n+1} + 5$. Let

$$D_m(a,b) = \sum_{k=1}^{m} (y_k - ax_k - b)^2 \ .$$

Obviously, for all a and b we have $D_{n+1}(a,b) \geq D_n(a,b)$. For $(a,b) = (-2,5)$ we have:

$$D_{n+1}(-2,5) = D_n(-2,5) = \min_{a,b \in \mathbf{R}} D_n(a,b) \leq \min_{a,b \in \mathbf{R}} D_{n+1}(a,b) \ .$$

Therefore $D_{n+1}(-2,5) = \min_{a,b \in \mathbf{R}} D_n(a,b)$.

Necessity: Suppose $y = -2x + 5$ is also the new least-squares line. This means that:

$$\overline{y}_n = -2\overline{x}_n + 5, \qquad \overline{y}_{n+1} = -2\overline{x}_{n+1} + 5 \ .$$

Now

$$\overline{x}_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} x_i = \frac{n}{n+1}\overline{x}_n + \frac{1}{n+1} x_{n+1}$$

and similarly

$$\overline{y}_{n+1} = \frac{n}{n+1}\overline{y}_n + \frac{1}{n+1} y_{n+1} \ .$$

Consequently:

$$\left(\frac{n}{n+1}\overline{y}_n + \frac{1}{n+1} y_{n+1}\right) + 2\left(\frac{n}{n+1}\overline{x}_n + \frac{1}{n+1} x_{n+1}\right) = 5 \ .$$

It follows that

$$\frac{n}{n+1}\left(\overline{y}_n + 2\overline{x}_n\right) + \frac{1}{n+1}\left(y_{n+1} + 2x_{n+1}\right) = 5 \ ,$$

and hence

$$\frac{n}{n+1} \cdot 5 + \frac{1}{n+1}\left(y_{n+1} + 2x_{n+1}\right) = 5 \ .$$

This easily yields

$$y_{n+1} + 2x_{n+1} = 5 \ .$$

4. The integral $\int_0^1 \sqrt{1-x}\, dx$ is estimated by dividing the interval $[0,1]$ into n equal subintervals and employing the rectangle rule for each. For which n will the error be at most 0.001?

Let $f(x) = \sqrt{1-x}$. Then $f^{'}(x) = -\frac{1}{2\sqrt{1-x}}$.

Let E_1 be the error in the interval $\left[0, 1-\frac{1}{n}\right]$ and E_2 the error in the interval $\left[1-\frac{1}{n}, 1\right]$. We have:

$$\begin{aligned}|E_1| &\leq \frac{1}{2} \max_{x\in[0,1-1/n]} \left|f^{'}(x)\right| \cdot \left(1-\frac{1}{n}\right) \cdot \frac{1}{n} \\ &= \frac{1}{2} \cdot \frac{1}{2\sqrt{1-(1-1/n)}} \cdot \left(1-\frac{1}{n}\right) \cdot \frac{1}{n} = \frac{1-1/n}{4\sqrt{n}} \ .\end{aligned}$$

Since f is decreasing and non-negative in $[0,1]$, the error on $\left[1-\frac{1}{n}, 1\right]$ is at most $f(1-1/n)\cdot\frac{1}{n}$, so that:

$$|E_2| \leq \sqrt{1-(1-1/n)} \cdot \frac{1}{n} = \frac{1}{n^{3/2}} \ .$$

Hence:

$$|E| \leq |E_1| + |E_2| \leq \frac{1-1/n}{4\sqrt{n}} + \frac{1/n}{\sqrt{n}} = \frac{1+3/n}{4\sqrt{n}} \ .$$

Thus n has to be a little more than $250^2 = 62500$.

A better bound on the error may be obtained if we divide the interval $[0,1]$ not to the parts $\left[0, 1-\frac{1}{n}\right]$ and $\left[1-\frac{1}{n}, 1\right]$, but to $\left[0, 1-\frac{a_n}{n}\right]$ and $\left[1-\frac{a_n}{n}, 1\right]$, where a_n is to be determined. We obtain

$$|E_1| \leq \frac{1}{2} \max_{x\in[0,1-a_n/n]} \left|f^{'}(x)\right| \cdot \left(1-\frac{a_n}{n}\right) \cdot \frac{1}{n} \leq \frac{1}{4\sqrt{na_n}}$$

and

$$|E_2| \leq \sqrt{1-(1-a_n/n)} \cdot \frac{1}{n} = \left(\frac{a_n}{n}\right)^{3/2} \ .$$

Taking $a_n \approx \sqrt{n}/2$ we have:

$$|E| \leq \frac{1}{4\sqrt{n\sqrt{n}/2}} + \left(\frac{\sqrt{n}}{2n}\right)^{3/2} = \frac{1}{\sqrt{2}n^{3/4}} \ .$$

With this bound, we see that already $n \geq \frac{10000}{2^{2/3}}$ provides an estimate with the required accuracy.

A still better bound is obtained by observing that the function is concave, so that, using the preceding method, the error on $\left[1-\frac{a_n}{n}, 1\right]$ is actually at most half of our previous bound:

$$|E_2| \leq \frac{1}{2} \cdot \left(\frac{a_n}{n}\right)^{3/2} \ .$$

This time we take $a_n \approx \sqrt{n/2}$ to get $|E| \leq \left(\frac{1}{2n}\right)^{3/4}$, which shows that it suffices to take $n \geq 5000$.

An altogether different possibility is to notice that, since f is decreasing, the rectangle rule yields an upper bound on the value of the integral, whereas the analogous method of taking the right endpoint of each subinterval yields a lower bound. Hence:

$$|E| < \sum_{i=0}^{n-1} \left(f\left(\frac{i}{n}\right) - f\left(\frac{i+1}{n}\right) \right) \cdot \frac{1}{n} = (f(1) - f(0)) \cdot \frac{1}{n} = \frac{1}{n} .$$

Thus even $n \geq 1000$ suffices.

One can improve the last method even further by observing that the concavity of f implies that the error on each subinterval is at most half our previous bound. This yields $|E| \leq \frac{1}{2n}$, which shows that already $n \geq 500$ is good enough.