Aero 320: Numerical Methods

Homework 2

Name: ....

Due: September 24, 2013

NOTE: All problems are to be done by hand (with the help of a calculator) but you need to show all the steps. Turn in a hard copy of your HW stapled with this as cover sheet with your name written in the above field. Submit your HW in the lab next Tuesday. Late submissions or failure to submit in the required format will receive no credit.

Problem 1

More on round-off error

The Taylor series of degree n for $\exp(x)$ is $\displaystyle\sum_{j=1}^{n} \frac{x^j}{j!}$. Use this polynomial and *rounding-off to three digits*, to find an approximate value of $\exp(-5)$ using

(a) $\exp(-5) \approx \displaystyle\sum_{j=1}^{9} \frac{(-5)^j}{j!} = \sum_{j=1}^{9} \frac{(-1)^j\, 5^j}{j!}$,

(b) $\exp(-5) = \dfrac{1}{\exp(5)} \approx \dfrac{1}{\displaystyle\sum_{j=1}^{9} \frac{5^j}{j!}}$.

The *exact value* of $\exp(-5)$, correct to three digits, is 6.74×10^{-3}. Which formula (a) or (b) gives more accurate result? Why?

Solution

Unfortunately, there was a typo in this problem. The Taylor series summation should start from $j = 0$, instead of $j = 1$. You will receive full credit if you attempt this problem.

From (a),

$$\begin{aligned}\exp(-5) \approx \sum_{j=0}^{9} \frac{(-1)^j 5^j}{j!} &= 1 - 5 + \frac{25}{2} - \frac{125}{6} + \frac{625}{24} - \frac{3125}{120} + \frac{15625}{720} - \frac{78125}{5040} + \frac{390625}{40320} - \frac{1953125}{362880} \\ &\approx 1 - 5 + 12.5 - 20.8 + 26.0 - 26.0 + 21.7 - 15.5 + 9.69 - 5.38 \\ &= -1.79.\end{aligned}$$

From (b),

$$\exp(-5) \approx \frac{1}{\sum_{j=0}^{9} \frac{5^j}{j!}} = \frac{1}{143.35} \approx \text{(up to three significant digits) } 0.00697 = 6.97 \times 10^{-3}.$$

Clearly, part (b) produces more accurate answer. Since both approximations use 10 terms in the Taylor series, both (a) and (b) have same truncation error. However, (a) produces more round-off error since the approximating Taylor series have terms with alternate signs and 3 digit round-off error affects each term. On contrary, the Taylor series in part (b) has all positive signs and appear in the denominator. So the formula (b) is less affected by the accumulation of round-off errors.

Notice that if the Taylor summation started indeed from $j = 1$ (as appeared in the problem statement due to the typo), then you would get the opposite conclusion, i.e. part (a) would give better approximation than part (b).

Problem 2

(Problem # 10, p. 67) Convergence of secant versus bisection method

Explain why the secant method (see p. 39) usually converges to a given stopping tolerance faster than bisection.

Solution

For bisection method, $e_n = |x_n - x| \leq \frac{b-a}{2^n}$. As $n \to \infty$, $\frac{b-a}{2^n} \to 0$, and hence $e_n \to 0$, meaning the bisection method always converges, no matter what nonlinear equation $f(x) = 0$ we are solving for (the error estimate is independent of the nonlinear function $f(x)$). Further, $\frac{e_{n+1}}{e_n} \leqslant \frac{1}{2}$. Hence, $\alpha_{\text{bisection}} = 1$, and $\lambda_{\text{bisection}} = \frac{1}{2}$. In other words, bisection method has *linear convergence*.

It can be shown that the secant method has *super-linear convergence*, i.e. $1 < \alpha_{\text{secant}} < 2$, meaning the error e_n in secant method converges *faster than linear, but slower than quadratic*. More specifically, it can be proved that $\alpha_{\text{secant}} = \frac{\sqrt{5}+1}{2} \approx 1.618$. This number is called the *Golden ratio*. Further, if we assume that x is the root, and $f'(x), f''(x) \neq 0$, then

$$\lambda_{\text{secant}} \approx \left| \frac{f''(x)}{2f'(x)} \right|^{\frac{\sqrt{5}-1}{2}}.$$

Since the secant method error decays faster than linear, hence it usually reaches the error tolerance faster than bisection method. You do not need to derive the value of α_{secant} and λ_{secant} for this homework.

Problem 3

(Problem # 15, p. 67) Newton's method for finding n^{th} root of a number

Applying Newton's method to the equation $x^2 = N$, gives the following algorithm to compute the square root of N (Problem # 14, p. 67):

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{N}{x_n}\right).$$

(a) Find an algorithm for getting the *cubic* and *quartic* roots of N that have a similar form to the one above for the square root. Can you generalize your answer for the n^{th} root?

(b) Starting with $x_0 = 2$, perform 3 Newton iterations to find $3^{1/3}$. Compare your result with the *actual value* of $3^{1/3}$ (use your calculator), and determine the *number of significant digits* in your answer.

Solution

To avoid confusion, let us use the symbol k as the iteration index for Newton's method, i.e. write

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

(a) For *cubic root*, the nonlinear equation to solve is $f(x) = x^3 - N = 0$, $f'(x) = 3x^2$. Hence, the Newton's iteration becomes

$$\begin{aligned} x_{k+1} &= x_k - \frac{x_k^3 - N}{3x_k^2} \\ &= x_k - \frac{x_k}{3} + \frac{N}{3x_k^2} \\ &= \frac{1}{3}\left(2x_k + \frac{N}{x_k^2}\right). \end{aligned}$$

For *quartic root*, the nonlinear equation to solve is $f(x) = x^4 - N = 0$, $f'(x) = 4x^3$. Hence, the Newton's iteration becomes

$$\begin{aligned} x_{k+1} &= x_k - \frac{x_k^4 - N}{4x_k^3} \\ &= x_k - \frac{x_k}{4} + \frac{N}{4x_k^3} \\ &= \frac{1}{4}\left(3x_k + \frac{N}{x_k^3}\right). \end{aligned}$$

In general, to find the n^{th} root, we need to solve the equation $f(x) = x^n - N = 0$, $f'(x) = nx^{n-1}$. Hence, the Newton's iteration becomes

$$\begin{aligned} x_{k+1} &= x_k - \frac{x_k^n - N}{nx_k^{n-1}} \\ &= x_k - \frac{x_k}{n} + \frac{N}{nx_k^{n-1}} \\ &= \frac{1}{n}\left((n-1)\,x_k + \frac{N}{x_k^{n-1}}\right). \end{aligned}$$

(b) To solve $x = 3^{1/3}$ is same as solving $x^3 - 3 = 0$. Thus, $N = 3$ in this case. From part (a), the Newton iteration for this case becomes

$$x_{k+1} = \frac{1}{3}\left(2x_k + \frac{3}{x_k^2}\right), \quad \text{with initial guess } x_0 = 2.$$

Hence, $x_1 = \frac{1}{3}\left(4 + \frac{3}{4}\right) = \frac{4.75}{3} \approx 1.5833$, up to 5 significant digits. Similarly, $x_2 = \frac{1}{3}\left(2 \times 1.5833 + \frac{3}{(1.5833)^2}\right) \approx 1.4544$, up to 5 significant digits. Next, $x_3 = \frac{1}{3}\left(2 \times 1.4544 + \frac{3}{(1.4544)^2}\right) \approx 1.4424$, up to 5 significant digits.

The exact value of $3^{\frac{1}{3}}$ (verify this from your calculator), is 1.44225 (this has six significant digits). If we keep up to 5 significant digits, then its value is 1.4422.

Problem 4

(Problem # 36, p. 69) Fixed point method

Most equations of the form $f(x) = 0$ can be rearranged in the form $x = g(x)$, with which to begin the fixed-point method. For $f(x) = \exp(x) - 2x^2$, one possible way to rewrite $f(x) = 0$ is:

$$x = \pm\sqrt{\frac{\exp(x)}{2}}.$$

(a) Show that this fixed point iteration converges to a root near 1.5 if the positive value is used and to the root near -0.5 if the negative value is used. (Show 3 iterations).

(b) There is a third root near 2.6. Show that we do not converge to this root even though values near to the root are used to begin the iterations. Where does it converge if $x_0 = 2.5$? If $x_0 = 2.7$? (Show 3 iterations).

(c) Find another rearrangement of the form $x = g(x)$, that converges correctly to the third root.

Solution

(a) First, we show that, the equation $x = +\sqrt{\frac{\exp(x)}{2}}$ has a root between 1 and 2. This is same as showing that the equation $f_+(x) = x - \sqrt{\frac{\exp(x)}{2}} = 0$ has a root between 1 and 2. Notice that

$$f_+(1)\, f_+(2) = \left(1 - \sqrt{\frac{e}{2}}\right)\left(2 - \frac{e}{\sqrt{2}}\right).$$

But we know that $2 < e < 3 \Rightarrow 1 < \frac{e}{2} \Rightarrow 1 < \sqrt{\frac{e}{2}}$. Hence $f_+(1) < 0$. On the other hand, $0.693 \approx \ln(2) > \frac{2}{3} \approx 0.667$. As a result, $\frac{3}{2}\ln(2) > 1 \Rightarrow \ln(2) > 1 - \frac{1}{2}\ln(2) = \ln\left(\frac{e}{\sqrt{2}}\right) \Rightarrow 2 > \frac{e}{\sqrt{2}}$. Thus, $f_+(2) > 0$. Hence, $f_+(x) = 0$ has a root between 1 and 2, possibly near 1.5.

Second, we show that, the equation $x = -\sqrt{\frac{\exp(x)}{2}}$ has a root between -1 and 0. This is same as showing that the equation $f_-(x) = x + \sqrt{\frac{\exp(x)}{2}} = 0$ has a root between -1 and 0. Notice that

$$f_-(0) = \frac{1}{\sqrt{2}} > 0, \quad f_-(-1) = \frac{1}{\sqrt{2e}} - 1.$$

Notice that $\frac{1}{2} < e \Rightarrow 1 < 2e \Rightarrow \frac{1}{\sqrt{2e}} < 1 \Rightarrow f_-(-1) < 0$. As a result, $f_-(0)\, f_-(-1) < 0$, meaning $f_-(x) = 0$ has a root between -1 and 0, possibly near -0.5.

Now, we start the fixed point iterations with $x_0 = 1$. Also, let's work with 5 significant digits. First, we consider the fixed point iteration with *positive* sign, i.e.

$$x_{k+1} = +\sqrt{\frac{\exp(x_k)}{2}}.$$

Then $x_1 = \sqrt{\frac{\exp(1)}{2}} = 1.1658$. Next, $x_2 = \sqrt{\frac{\exp(1.1658)}{2}} = 1.2666$. Further, $x_3 = \sqrt{\frac{\exp(1.2666)}{2}} = 1.3321$.

If we use the fixed point iteration with *negative* sign, i.e.

$$x_{k+1} = -\sqrt{\frac{\exp(x_k)}{2}},$$

then $x_1 = -\sqrt{\frac{\exp(1)}{2}} = -1.1658$. Next, $x_2 = -\sqrt{\frac{\exp(-1.1658)}{2}} = -0.3948$. Further, $x_3 = -\sqrt{\frac{\exp(-0.3948)}{2}} = -0.5804$.

(b) We need to show that both the fixed point iterations (with positive and negative sign) will fail to capture the root near 2.6. For this purpose, it suffices to show that neither $f_+(x) = 0$, nor $f_-(x) = 0$, has any root around 2.6.

First, notice that $e \approx 2.71$, and $\ln(2) \approx 0.693$. Hence, $\ln(2) + 2 < e \Rightarrow \ln(2) + 2\ln(e) < e \Rightarrow 2e^2 < e^e \Rightarrow e < \sqrt{\frac{e^e}{2}} \Rightarrow e - \sqrt{\frac{e^e}{2}} < 0 \Rightarrow f_+(e) < 0$. Also, $18 < e^3 \Rightarrow 9 < \frac{e^3}{2} \Rightarrow 3 - \frac{e^{3/2}}{\sqrt{2}} < 0 \Rightarrow f_+(3) < 0$. However, we have shown in part (a) that $f_+(2) > 0$. Thus, there is a root for the equation $f_+(x) = 0$, in the interval $(2, e)$; but no root in $(e, 3)$. You can also confirm this by plotting the function $f_+(x)$. You may verify that $|f'_+(x)| < 1$, meaning if you choose x_0 in this interval, convergence is guaranteed. However, it is not clear from this analysis, which particular root will you converge to.

On the other hand, $f_-(2) = 2 + \frac{e}{\sqrt{2}} > 0$, and $f_-(3) = 3 + \frac{e^{3/2}}{\sqrt{2}} > 0$. Consequently, $f_-(2)\, f_-(3) > 0$. Hence, the equation $f_-(x) = 0$ has no root between 2 and 3.

Now, we start the fixed point iterations with $x_0 = 2.5$. As before, we work with 5 significant digits. First, we consider the fixed point iteration with *positive* sign, i.e.

$$x_{k+1} = +\sqrt{\frac{\exp(x_k)}{2}}.$$

Then $x_1 = 2.4680$, $x_2 = 2.4289$, $x_3 = 2.3819$. Seems like we are going away from the root near 2.6. Indeed, if you continue, then this iteration yields $x_{120} = 1.4880$, meaning we slowly converge toward the root at 1.5.

If we use the fixed point iteration with *negative* sign, i.e.

$$x_{k+1} = -\sqrt{\frac{\exp(x_k)}{2}},$$

then $x_1 = -2.4680$, $x_2 = -0.2058$, $x_3 = -0.6379$. We converge near the root at -0.5.

Now, we start the fixed point iterations with $x_0 = 2.7$, for $x_{k+1} = +\sqrt{\frac{\exp(x_k)}{2}}$. We get $x_1 = 2.7276$, $x_2 = 2.7655$, $x_3 = 2.8185$. If you continue, then you will find that you actually diverge to $+\infty$.

For $x_0 = 2.7$, but $x_{k+1} = -\sqrt{\frac{\exp(x_k)}{2}}$, we obtain $x_1 = -2.7276$, $x_2 = -0.1808$, $x_3 = -0.6460$. We converge toward the root at -0.5.

(c) We can rewrite our original equation as

$$\exp(x) = 2x^2 \Rightarrow x = \underbrace{\ln\left(2x^2\right)}_{g(x)}.$$

In order to converge, $|g'(x)| < 1 \Rightarrow \left|\frac{2}{x}\right| < 1$. In other words, this fixed point iteration will converge if $0 > x_0 > -2$, *or* if $x_0 > 2$. Hence, the domain of convergence is: $(-2, 0) \cup (2, \infty)$. Remember that not all x_0 from this domain may converge to the root near 2.6 (i.e. they may converge to some other root). We just need to show that there exists *at least* one $x_0 \in (-2, 0) \cup (2, \infty)$, such that the fixed point iteration $x_{k+1} = \ln\left(2x_k^2\right)$ converges near 2.6.

Let us try $x_0 = 3$. We get $x_1 = \ln\left(2 \times 3^2\right) = 2.8904$, $x_2 = \ln\left(2 \times 2.8904^2\right) = 2.8159$, $x_3 = \ln\left(2 \times 2.8159^2\right) = 2.7637$, etc. Indeed, we are converging towards the root near 2.6.

Problem 5

Newton's and Halley's method to solve Kepler's equation

Kepler's equation, devised by Johannes Kepler in 1609, rules the propagation of satellites (or planets) in their orbits. The equation is transcendental:

$$M = E - \epsilon \sin(E)$$

where M is an angle called *mean anomaly*, ϵ is a positive number called *orbit eccentricity*, and E is an angle called *eccentric anomaly* that indicates the location of the satellite in the orbit. Solving Kepler's equation means to find E, when M and ϵ are given.

Solve the Kepler's equation using **Newton's** and **Halley's method** for the following cases. Use $E = M + 1$ as the starting point, and the convergence is achieved if the absolute error $\leq 10^{-6}$. Also report the number of iterations needed to converge for each case.

(a) $\epsilon = 0.2, 0.4, 0.8$, for $M = 0, 1, 10, 100$ degrees.

(b) $\epsilon = 0.9999$ (orbit almost parabolic), for $M = 0.001$ degrees (passage close to perigee).

Solution

Use the code sent as the solution of Lab Assignment 7. Convert M in degrees to radian, then take the initial guess as E (in radian) $= M$ (converted from degree to radian) $+ 1$ radian. Then start iterating the Kepler's equation. The E in Kepler's equation will be computed in radians.

The Grading Rubric for Homework 2

Problem 1: **Total = 15 points**. Breakdown: part (a) = 5 points; part (b) = 5 points; which formula = 1 point; why = 4 points.

Problem 2: **Total = 5 points**.

Problem 3: **Total = 20 points**. Breakdown: part (a) = $3 \times 5 = 15$ points; part (b) = 5 points (3 Newton iterations = 3 points, actual value = 1 point, significant digits = 1 point).

Problem 4: **Total = 30 points**. Breakdown: part (a) = 10 points (proof for root near 1.5 = 2 points, proof for root near -0.5 = 2 points, 3 iterations with positive formula = 3 points, 3 iterations with negative formula = 3 points); part (b) = 10 points (proof that positive formula fails = 2 points, proof that negative formula fails = 2 points, iterations starting from 2.5 = $(3 \times 0.5) \times 2 = 3$ points, iterations starting from 2.7 = $(3 \times 0.5) \times 2 = 3$ points); part (c) = 10 points (finding rearrangement = 3 points, show that whether it converges at all ($|g'(x)| < 1$) = 3 points, choose appropriate x_0 and running few iterations to demonstrate convergence near 2.6 = 4 points).

Problem 5: **Total = 30 points**. Breakdown: correct understanding and problem set up = 4 points; numerical results = $(13 \times 2) \times 1 = 26$ points.