AMSC/CMSC 460 Dr. Wolfe ASSIGNMENT #1 Due June 13, 2008

1. Ex.1.35, p.47, *Numerical Computing with MATLAB* . Modify the programs by inserting a counter that will count the lines of output. Do not hand in any output. Just answer the questions.

2. Ex.1.38, p.48, *Numerical Computing with MATLAB*

3. The **error function** is an important function in many branches of applied mathematics. It is defined by an integral.

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\, dt.$$

The integral cannot be expressed in terms of more elementary functions. It is a library function in MATLAB . The Taylor series for the error function is

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{n!(2n+1)}.$$

The series converges for all x. Write a MATLAB function m-file to evaluate erf(x) using this series. Use as many terms in the series as are necessary so that the first neglected term does not alter the accumulated sum when it is added to it in floating point arithmetic. Since this is an alternating series, the error caused by truncating the infinite sum will then be less than the roundoff error. Investigate the effect of roundoff error by comparing the computed sum with the value given by the MATLAB function erf. Try $x = 0.5, 1.0, 2.0, 5.0, 7.0, 10.0$. Explain your results. *Hint:* The main loop of your program might look like this:

while $s \sim= so$
$so = s;$
$n = n + 1;$
$t = -xsq * t * (2.0 * n - 1.0)/(n * (2 * n + 1));$
$s = s + t;$
end

What values should be assigned to t, n, s and xsq before entering the loop ? Do not forget the factor $2/\sqrt{\pi}$. Before running the script, give the MATLAB command "format long".

4. Suppose a computer carries three decimal digits and rounds. If x and y are machine numbers, define the machine version of addition $x \oplus y$ to be the result of adding x and y and rounding to three digits. For example

$$49.3 \oplus 57.4 = 107.$$

Define machine multiplication $x \otimes y$ similarly. For example,

$$1.23 \otimes 4.86 = 5.98.$$

Constuct examples to show that, in general, the following statements are <u>not</u> true:
(a) $(x \oplus y) \oplus z = x \oplus (y \oplus z)$,
(b) $(x \otimes y) \otimes z = x \otimes (y \otimes z)$,
(c) $x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$.

5. For a set of measurements $x_1, x_2, \ldots, x_N$, the sample mean $\bar{x}$ is defined to be

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i.$$

The sample standard deviation s is defined to be

$$(N-1)s^2 = \sum_{i=1}^{N} (x_i - \bar{x})^2.$$

Another expression,

$$(N-1)s^2 = \sum_{i=1}^{N} x_i^2 - \frac{1}{N}\Big(\sum_{i=1}^{N} x_i\Big)^2,$$

is ofter recommended for hand computation of s. Show that these two expressions for s are mathematically equivalent. Explain why one of them may provide better numerical results than the other, and construct and example to illustrate your point.

6. If x_T represents the true value of a variable x and x_A represents an approximate value for x, and if f is a smooth function we have

$$f(x_T) - f(x_A) \approx f'(x_T)(x_T - x_A) \approx f'(x_A)(x_T - x_A). \qquad (*)$$

Let $f(x) = (x-1)(x-2)\ldots(x-n)$. Note that $f(1) = 0$. Estimate $f(1 + 10^{-4})$ by using (*) with $x_T = 1$, for $n = 2, 3, \ldots, 12$. Comment on the implications of this for finding the zeros of $f(x)$, say, for the case $n = 8$.
Hint: Do not calculate $f'(1)$ by first multiplying out $f(x)$. Instead, use the product rule for derivatives to evaluate $f'(x)$, and then obtain $f'(1)$.

7. Use three-digit arithmetic with rounding to compute the following sums (sum in the given order).

$$\text{(a)} \sum_{k=1}^{6} \frac{1}{3^k} \qquad\qquad \text{(b)} \sum_{k=1}^{6} \frac{1}{3^{7-k}}$$

Also, compare the answers with the exact sum. Which is better ?

8. Let a regular polygon of N sides be inscribed in a unit circle. If L_N denotes the length of one side, the circumference of the polygon, $N \times L_N$, approximates the circumference of the circle, 2π; hence $\pi \approx NL_N/2$ for large N. Using Pythagoras' thoerem it is easy to relate L_{2N} to L_N:

$$L_{2N}^2 = 2\Big(1 - \sqrt{1 - L_N^2/4}\Big).$$

Starting with $L_4 = \sqrt{2}$ for a square, approximate π by means of this recurrence. Explain why a straightforward implementation of the recurrence in floating point arithmetic does not yield an accurate value for π. (Keep in mind that $L_N \to 0$ as $N \to \infty$.) Show that the recurrence can be rearranged as

$$L_{2N}^2/4 = \frac{L_N^2/4}{2\left(1 + \sqrt{1 - L_N^2/4}\right)}$$

and demonstrate that this form works better.

9. What is the largest value of n so that $n!$ can be exactly represented in a floating point number system where $(\beta, t, m, M) = (2, 24, -100, 100)$? Show your work.