Homework 2. Due Friday Sept. 13 by 10am in my mailbox

Always provide complete, well written solutions.

The solution to one of the following problems (of your choice) has to be typed in Latex, while the rest of the solutions can be handwritten.

Always submit printouts of the output of your codes. Moreover your matlab codes have to be submitted via ELMS.

1. Ref. [1], Chapter 2, Problem 6 (page 26).

   Use the fact that the floating point sum `x + y` has relative error ϵ such that $|\epsilon| < \epsilon_M$ to show that the absolute error in the sum `S`$= \sum_{i=0}^{n-1} x_i$ computed below is no worse than $(n-1)\epsilon_M \sum_{i=0}^{n-1} |x_i|$:

```
function S = compute_sum(x)
%% S = compute_sum(x)
% compute the sum of the entries in the vector x
%
% INPUT
% x vector of n scalars
%
% OUTPUT
% S scalar
%    sum of the n entries of x
n = length(x);
S = 0;
for i = 1 : n
    S = S + x(i);
end
end
```

2. Ref. [1], Chapter 2, Problem 9 (page 26).

   The binomial coefficients, $a_{n,k}$, are defined by

$$a_{n,k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

   To compute the $a_{n,k}$, for a given n, start with $a_{n,0} = 1$ and then use the recurrence relation $a_{n,k+1} = \frac{n-k}{k+1} a_{n,k}$.

   (a) For a range of n values, compute the $a_{n,k}$ this way, noting the largest $a_{n,k}$ and the accuracy with which $a_{n,n} = 1$ is computed. Why is roundoff not a problem here?

(b) Use the algorithm of part (a) to compute

$$E(n) = \frac{1}{2^n} \sum_{k=0}^{n} k a_{n,k} = \frac{n}{2} \quad . \tag{1}$$

Write a program without any safeguards against overflow or zero divide (this time only!) Show that the computed answer has high accuracy as long as the intermediate results are within the range of floating point numbers. As with (a), explain how the computer gets an accurate, small, answer when the intermediate numbers have such a wide range of values. Why is cancellation not a problem? Note the advantage of a wider range of values: we can compute $E(n)$ for much larger n in double precision. Print $E(n)$ as computed by (1) and $M_n = \max_k a_{n,k}$. For large n, one should be `inf` and the other `NaN`. Why?

3. Consider the integrals

$$E_n = \int_0^1 x^n e^{x-1} dx.$$

(a) Compute E_0.

(b) For $n \in \mathbb{N}^*$, express E_n in terms of E_{n-1}.

(c) Show that for all $n \in \mathbb{N}$ we have

$$\frac{1}{e(n+1)} \leq E_n \leq \frac{1}{n+1}.$$

(d) Assuming that the only error made is in the evaluation of the first term, explain what can go wrong if you implement the formula obtained in 3b.

(e) Write a code, with a parameter `nmax` defining the number of iterations, implementing the formulas obtained in 3a and 3b to compute E_n for n up to `nmax`. Plot the approximated values of E_n as a function of n as well as the bounds obtained in 3c, choosing an appropriate value of `nmax` to evidence something going wrong. Hand in a print out of the figure.

4. **Manipulating basic linear algebra concepts**

(a) Is $V' := \{u \in \mathbb{R}^n, \sum_{k=1}^n u_k = 0\}$ a subspace of $\mathbb{R}^n$? Prove it.

(b) Is $V' := \{u \in \mathbb{C}^n, \sum_{k=1}^n u_k = 1\}$ a subspace of $\mathbb{C}^n$? Prove it.

(c) What is the standard inner product on $\mathbb{R}^n$?

(d) Prove that the space of polynomials with real coefficients is a vector space.

(e) Assume $d \in \mathbb{N}^*$. Prove that the space of polynomial of degree no larger than d with real coefficients is a subspace of the vector space of polynomials with real coefficients.

(f) If A is a real matrix, what's the difference between its transpose and its adjoint?

References

[1] Bindel and Goodman, Principles of scientific computing