MATH 387 ASSIGNMENT 3

DUE TUESDAY MARCH 29

Part I. Analytic exercises

1. (Hermite/osculatory interpolation) Let $x_0, x_1, \ldots, x_n$ be distinct points in $[a, b]$, and let $r_0, r_1, \ldots, r_n$ be positive integers. Show that if f and its derivatives of order $r_0 - 1, r_1 - 1, \ldots, r_n - 1$ are defined respectively at the points $x_0, x_1, \ldots, x_n$, then there exists a unique polynomial q of degree at most N, where $N = r_0 + \ldots + r_n - 1$, such that
$$q^{(k)}(x_j) = f^{(k)}(x_j), \qquad \text{for} \quad k = 0, 1, \ldots, r_j - 1, \quad j = 0, 1, \ldots, n.$$
Furthermore, prove that if $f \in C^{N+1}([a, b])$, then for any $x \in [a, b]$, there exists $\xi \in [a, b]$ such that
$$f(x) - q(x) = \frac{(x - x_0)^{r_0} \cdots (x - x_n)^{r_n}}{(N + 1)!} f^{(N+1)}(\xi).$$
Show that this process contains both Lagrange interpolation and Taylor polynomials as special cases.
2. Let f be a C^3 function on some interval $[-a, a]$, and consider the first order centred difference formula
$$D_h = \frac{f(h) - f(-h)}{2h},$$
for approximating $f'(0)$. In exact arithmetic, we have $D_h - f'(0) = O(h^3)$. Here it is obvious that the accuracy of D_h can be made arbitrarily small by choosing the parameter h small. Now, suppose that the function evaluations are performed inexactly, so that instead of f, we only have access to some function $f + \delta f$ with $|\delta f(x)| \leq \varepsilon$ for $x \in (-a, a)$, where $\varepsilon > 0$ is a small constant. The aforementioned centred difference should then be replaced by
$$\tilde{D}_h = \frac{f(h) + \delta f(h) - f(-h) - \delta f(-h)}{2h}.$$
   (a) Show that for each $0 < h < a$, there exists $\xi \in (-h, h)$ such that
$$\tilde{D}_h - f'(0) = \frac{h^2}{6} f'''(\xi) + \frac{\delta f(h) - \delta f(-h)}{2h}.$$
   (b) Conclude that
$$|\tilde{D}_h - f'(0)| \leq \frac{\varepsilon}{h} + \frac{M h^2}{6}, \qquad \text{where} \qquad M = \max_{x \in [-a, a]} |f'''(x)|.$$

Date: Winter 2016.

Supposing that ε and M are fixed, sketch the graph of the bound $B(h) = \frac{\varepsilon}{h} + \frac{Mh^2}{6}$ as a function of h. Find the value of h that minimizes $B(h)$. Explain why it is not a good idea to choose h too small.

3. Give a constructive proof of the Weierstrass approximation theorem for the unit square $Q = [0,1] \times [0,1]$, by using products of Bernstein polynomials. That is, show that if f is a continuous function on Q, then for any $\varepsilon > 0$, there exists m and n, such that

$$\max_{(x,y)\in Q} |f(x,y) - B_{mn}(x,y)| \leq \varepsilon,$$

where

$$B_{mn}(x,y) = \sum_{j=0}^{m} \sum_{k=0}^{n} f(x_j, y_k)\beta_{mj}(x)\beta_{nk}(y),$$

with $x_j = j/m$, $y_k = k/n$, and β_{nk} being the Bernstein polynomials we defined in class.

4. (Süli-Mayers) Show that the minimax polynomial approximation of $f \in C([-a,a])$ of any degree is an even function, if f is even. Deduce that if f is even, the minimax polynomial approximation of degree $2n$ is also the minimax polynomial approximation of degree $2n+1$. What does this imply about the sequence of critical points of the minimax polynomial of degree $2n$?
5. (Süli-Mayers) Find the minimax polynomial approximation of degree n for the function $f(x) = a_0 + a_1 x + \ldots + a_{n+1}x^{n+1}$ on the interval $[-1,1]$, where $a_{n+1} \neq 0$. (You need to prove that the polynomial you found is indeed the minimax polynomial.)
6. Show that the least squares approximation on $(-a,a)$ with respect to the weight function w preserves parity of a function, if w is an even function.

Part II. Computational exercises

Note: Along with your well explained illustrations, please submit the code you have written for the following experiments.

1. It is well known that Lagrange interpolation with equally spaced nodes is susceptible to the so-called *Runge phenomenon*. In this exercise, you are asked to do some reading to familiarize yourself with this phenomenon, and illustrate the phenomenon by a well-chosen numerical experiment. You are also asked to perform Lagrange interpolation with Chebyshev nodes on the same examples, in order to see how Chebyshev nodes compare to equally spaced nodes. The design of the experiment, and how you showcase it are entirely up to you, with the only requirement being that you need to show some originality. In particular, the standard example $f(x) = \frac{1}{1+x^2}$ and its close relatives should not be used. You are not required to provide a rigorous mathematical proof that your example indeed exhibits the Runge phenomenon.
2. In this exercise, we will do an experimental study of the growth factor for Gaussian elimination with partial pivoting. Given the factorization $PA = LU$, define the growth factor by

$$g(A) = \frac{\max_{i,j} |u_{ij}|}{\max_{i,j} |a_{ij}|},$$

where u_{ij} and a_{ij} are the elements of U and A, respectively.

(a) Write a function/subroutine that takes n as its parameter and generates an $n \times n$ matrix, whose entries are random numbers uniformly distributed in $[-1, 1]$.

(b) Implement Gaussian elimination with partial pivoting.

(c) Plot the growth factor against the matrix size, in logarithmic scale, where the matrices are generated as in (a). The sample of matrices should be large enough to be a good representative of random matrices with n ranging between, say 10 and 1000 (larger sizes are welcome, if it does not take too long). In particular, for each n, one should experiment on a generous number of matrices. From the experimental plot, estimate the power α in the assumed dependence $g \approx cn^{\alpha}$, where g is the growth factor, n is the matrix size, and c is a constant. Compare this with the worst case scenario $g \approx 2^n$.

(d) Now we study the probability distribution of the growth factor for a fixed n. Fix n, say $n = 10$ or $n = 16$, and generate an abundant number of random matrices, as in (a), to measure their growth factors. Then by subdividing the value-space of the growth factors into small subintervals of equal length, and by counting the number of matrices with growth factor lying in each of those subintervals, produce an approximation of the probability density function of the growth factor (the usual "histogram" technique). Plot it against the growth factor value, with the vertical axis in logarithmic scale. Make a conjecture on how the probability density decays as the growth factor becomes large. Note that the number of matrices and the length of the subintervals should be so that most of the subintervals individually contain a large number of matrices, and that there are enough subintervals to give a meaningful approximation of the probability density function (i.e., the width of bars in the histogram must be small, but most of the bars must still include a large number of cases). Repeat the experiment for several values of n, say $n = 20, 40, 80$, or $n = 32, 64, 128$.