Math 620 Comprehensive Examination

August 17, 2020

Instructions: You must show all your work to receive full credit. Partial answers will only receive partial credit. Please choose 3 of the 4 problems to solve. Please indicate which 3 problems you would like graded.

1. Consider the model of a binary computer number system, whose non-zero (and non-special) numbers are modeled by
$$\tilde{x} = (-1)^s \, (b_0.b_1 b_2 \dots b_{p-1})_2 \, 2^E$$
with sign bit $s \in \{0,1\}$, binary digits $b_i \in \{0,1\}$, $i = 0, 1, \dots, p-1$, and exponent $E_{\min} \le E \le E_{\max}$, where p, $E_{\min}$, $E_{\max}$ are the given parameters of the number system.

   (a) The normalized numbers $\tilde{x}$ are defined by $E_{\min} \le E \le E_{\max}$ and $b_0 \ne 0$ being the first non-zero digit of the number. Determine the formulas for the smallest and largest normalized numbers.

   (b) The de-normalized numbers $\tilde{x}$ are defined by $E = E_{\min}$ and the requirement that not all digits b_i, $i = 0, 1, \dots, p-1$, are zero. Determine the formulas for the smallest and largest de-normalized numbers.

   (c) Consider the very small number system with the given parameters $p = 3$, $E_{\min} = -1$, $E_{\max} = 2$. Compute and list **all** positive numbers of this number system (both normalized and de-normalized), and draw these positive numbers on a number line for $\tilde{x}$.

2. Consider the problem of approximation of a given function $f(x)$ by a polynomial $p(x)$.

   (a) State the problem precisely (with some key assumption). How is the goal of this problem different from conventional polynomial interpolation?

   (b) State the Weierstrass theorem that guarantees existence of a solution for this problem.

   (c) Summarize several methods that exist for the approach to this problem. Discuss the methods; are there differences in cost or other advantages and disadvantages?

3. (a) Consider a quadrature on $[0, 1]$ of the form:
$$Q(f) = w_0 f(0) + w_1 f(\xi), \tag{1}$$
meaning, $Q(f)$ is supposed to approximate $\int_0^1 f(x)dx$. Identify the values of w_0, w_1, and ξ so that the quadrature Q has the highest possible degree of precision. Determine the actual degree of precision of the quadrature you found.
Note: this is not a standard quadrature formula; there will be no partial credit if you assume that it is standard.

   (b) Translate the quadrature from the standard interval $[0, 1]$ to an arbitrary interval $[a, b]$, and derive (with a sufficiently plausible explanation, if not a complete proof) an approximation estimate for the quadrature Q in (1) of the form
$$\left| \int_a^b f(x)dx - Q(f) \right| \le C(b-a)^m \max_{x \in [a,b]} |f^{[n]}(x)|. \tag{2}$$
Hint: You may consider an interpolation polynomial of appropriate degree to produce the estimate, and you must identify the numbers m and n above.

   (c) Build a composite rule on an interval $[a, b]$ based on the quadrature Q in (1). Assume a uniform partition $a = x_0 < x_1 < x_2 < \cdots < x_k = b$, with $x_{j+1} - x_j = h$. Determine how many function evaluations are needed if k intervals are used, and state (based on your formula (2)) an approximation result of the form
$$\left| \int_a^b f(x)dx - Q(f) \right| \le Ch^r \max_{x \in [a,b]} |f^s(x)|. \tag{3}$$

4. (a) State the Contraction Mapping Theorem (CMT) and briefly describe the fixed point method.

   (b) Consider the equation

$$x = \lambda x(1 - x), \quad \lambda > 0, \tag{4}$$

   which has solutions $\xi_0 = 0$ and $\xi_1 = 1 - 1/\lambda$. Show that for $\lambda \in (0, 1)$ the CMT applies on the interval $[0, 1]$. To which solution does the fixed point method converge?

   (c) Show that for $\lambda \in (1, 3)$ the CMT applies on a small interval around ξ_1. (You may invoke a theorem related to fixed points).

   (d) Find $\lambda \in (1, 3)$ so that the fixed point method converges to ξ_1 quadratically.