

**NATIONAL UNIVERSITY OF SINGAPORE**

PC5215 – NUMERICAL RECIPES WITH APPLICATIONS

(Semester I: AY 2013-14)

Name of Examiner: Prof. Wang Jian-Sheng

Time Allowed: 2 Hours

---

**INSTRUCTIONS TO CANDIDATES**

1. Please write your student number only. **Do not write your name.**
2. This examination paper contains FIVE questions and comprises THREE printed pages.
3. Answer ALL the questions; questions carry equal marks.
4. Answers to the questions are to be written in the answer books.
5. Please start each question on a new page.
6. This is a CLOSED BOOK assessment.
7. Non-programmable calculators are allowed.

1. Answer briefly the following questions.

- The IEEE 754 quadruple-precision binary floating-point format (e.g., REAL\*16 in Fortran) has 128 bits (16 bytes). Among them the highest bit is for the sign, the next 15 bits are for a biased exponent with a bias of 16383. The rest of the 112 bits is for the fractional part. What is the machine epsilon of this format? What is the number of significant figures in decimal? What is the largest possible value in this format?
- What is a pivoting in LU decomposition of a square matrix? What is the purpose or role in pivoting? What is the computational complexity of an LU decomposition?
- Give the polynomial for a three-point interpolation passing through exactly  $(x_n, y_n) = (0,1), (3,-1), (7,2)$ .
- Define the discrete Fourier transform. Discuss qualitatively why fast Fourier transform is fast?

- Machine epsilon is  $2^{-112} = 1.9 \times 10^{-34}$ . The number of significant figures is 34 digits. And the largest possible value is approximately  $2^{16384} \approx 10^{4932}$ .*
- Pivoting is a way to move another matrix element to the diagonal either by column or row exchange/permutation. This is to improve the numerical stability of the method. The computational complexity of LU decomposition is  $O(N^3)$  for  $N \times N$  matrix.*
- $F(x) = 17/84 x^2 - 107/84 x + 1$ .*
- $y_k = \sum_{j=0}^{N-1} x_j e^{i2\pi kj/N}$ . It is fast because of the divide-and-conquer strategy like that for quick sort. The algorithmic complexity is  $O(N \ln N)$ .*

2. Consider a Markov chain with state space being the integers 1, 2, 3, and 4 and a transition matrix given by

$$W = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

- Define recurrence and answer if the state 1 is recurrent or not.

- b. Is there an absorbing state among 1, 2, 3, and 4?
  - c. Define irreducibility and answer if the Markov chain defined above by  $W$  is irreducible.
  - d. Determine the invariant or equilibrium distribution  $P$ , if it exists, such that  $P = PW$ .
- a. A state is recurrent if it is visited an infinite number of times as the steps in a Markov chain go to infinite. The state 1 is recurrent.
  - b. There is no absorbing state.
  - c. The  $W$  defined for the Markov chain is irreducible, because if  $n > 4$ , we can reach any state from any other state. I.e.,  $[W^n]_{ij} > 0$  for  $n > 4$ .
  - d.  $P = (1/4, 1/4, 1/4, 1/4)$ .
3. In this problem, we prove some of the important relations in the conjugate gradient method for iteratively finding the minimum  $x_{\min}$  for the quadratic form  $(1/2) x^T A x - b^T x$ . We assume that the matrix  $A$  is  $N$  by  $N$ . Let  $d_{(i)}$  be the search direction,  $e_{(i)} = x_{(i)} - x_{\min}$  be the error, and  $r_{(i)} = b - A x_{(i)}$  the residue, in  $i$ -th step,  $i = 0, 1, 2, \dots, N$ . We assume  $d_{(j)}$  is  $A$ -orthogonal,  $d_{(i)}^T A d_{(j)} = 0$  if  $i \neq j$ , and assume  $e_{(i)} = \sum_{j=i}^{N-1} \delta_j d_{(j)}$  without proof.
- a. Show that  $d_{(i)}^T r_{(j)} = 0$  if  $i < j$ .
  - b. Show that  $r_{(i)}^T r_{(j)} = 0$  if  $i \neq j$ .
  - c. Show that  $d_{(i)}^T r_{(i)} = r_{(i)}^T r_{(i)}$ .
- a. Multiply  $d_{(i)}^T A$  to  $e_{(j)}$  from the left, using the property of  $A$ -orthogonality, and  $r_{(j)} = -Ae_{(j)}$ , we obtain the result.
  - b. The key equation is to realize that  $r_{(i)}$  is in the space (Krylov space) spanned by  $d_{(0)}$  to  $d_{(i)}$ . Using this and the result of part a, we obtain the result.
  - c.  $d_{(i)}$  is built from the Gram-Schmidt process from  $r_{(j)}$ , so multiply  $r_{(i)}^T$  we get the result using part a again. Further details can be found in Shewchuk's note "an introduction to the conjugate gradient method without the agonizing pain".
4. Consider least-squares fit to experimental data to determine the value of  $f(0)$ . We assume that the data points are pair  $(x_i, y_i)$ ,  $i = 1, 2, \dots, N$ , where  $N$  is much larger than 3. We also assume that the fitting function is a quadratic polynomial,  $f(x) = a + bx + cx^2$ . Derive a formula for the estimate of  $f(0)$ , assuming equal weight (the same standard deviation  $\sigma$  for all  $y_i$ ).

Define the chi-square as  $\chi^2(a,b,c) = \sum_{i=1}^N (y_i - a - bx_i - cx_i^2)^2$ . Then take partial derivatives with respect to  $a$ ,  $b$ , and  $c$ . We get three linear equations. The equations are then solved formally (or concretely) using, e.g., Cramer's rule. Then  $f(0)=a$ .

5. Consider a one degree of freedom Hamiltonian system with the equations of motion

$$\dot{p} = -\frac{\partial H}{\partial q} = f(q),$$

$$\dot{q} = \frac{\partial H}{\partial p} = p,$$

where  $f(q)$  is the force. We propose the following numerical algorithm to solve it:

$$p_{n+1} = c_1 p_n + c_2 f(q_n),$$

$$q_{n+1} = c_3 q_n + c_4 p_n + c_5 p_{n+1}.$$

- Determine a relation among the coefficients  $c_1$  to  $c_5$  such that the algorithm becomes symplectic.
  - Determine or suggest a set of  $c_1$  to  $c_5$  such that it is consistent with the original differential equations.
  - Finally, propose an algorithm based on part a and b above that is both consistent with the Hamilton's equations and symplectic, and determine its order of accuracy.
- The condition for symplecticity is  $c_1 c_3 - c_2 c_4 f'(q) = 1$ .
  - Take  $c_1 = 1$ ,  $c_2 = h$ ,  $c_3 = 1$ ,  $c_4 + c_5 = h$ .
  - Since the symplectic condition needs to be satisfied for all  $q$ , we must take  $c_4 = 0$ , and take the rest as in part b. The order of the algorithm is first order, with local errors of  $O(h^2)$ .

--- the end ---

[WJS]