FAMILY NAME: ....
OTHER NAME(S): ....
STUDENT NUMBER: ....
SIGNATURE: ....

THE UNIVERSITY OF NEW SOUTH WALES

SCHOOL OF MATHEMATICS AND STATISTICS

June 2010

MATH2089
Numerical Methods and Statistics

(1) TIME ALLOWED – 3 Hours

(2) TOTAL NUMBER OF QUESTIONS – 6

(3) ANSWER ALL QUESTIONS

(4) THE QUESTIONS ARE OF EQUAL VALUE

(5) THIS PAPER MAY **NOT** BE RETAINED BY THE CANDIDATE

(6) **ONLY** CALCULATORS WITH AN AFFIXED “UNSW APPROVED” STICKER MAY BE USED

(7) STATISTICAL FORMULAE ARE ATTACHED AT END OF PAPER
STATISTICAL TABLES ARE ATTACHED AT END OF PAPER

Part A – Numerical Methods consists of questions 1 – 3

Part B – Statistics consists of questions 4 – 6

Both parts must be answered

All answers must be written in ink. Except where they are expressly required pencils may only be used for drawing, sketching or graphical work.

Part A – Numerical Methods

1. Answer in a separate book marked Question 1

a) What are the values produced by the following MATLAB expressions:

i)
```
x = 1e+200;
ans1 = exp(x^3)
```

ii)
```
v = [-2:2:2];
ans2 = v./sqrt(v)
```

iii)
```
x = 0.1;
h = 1e-18;
ans3 = x + h <= x
```

b) The computational complexity of some common operations with n by n matrices are

Operation	Flops
Matrix multiplication	$2n^3$
LU factorization	$\frac{2n^3}{3} + O(n^2)$
Cholesky factorization	$\frac{n^3}{3} + O(n^2)$
Back/forward substitution	$\frac{n^2}{2} + O(n)$
Tridiagonal solve	$8n + O(1)$

i) Estimate the size of the largest linear system that can be solved in one minute on a 3 GHz quad core computer, where each core can do two floating point operations per clock cycle. Assume that the coefficient matrix has no special structure.

ii) If multiplying two n by n matrices takes 12 seconds, estimate how long it will take to solve an n by n symmetric positive definite linear system.

c) The coefficient matrix A and the right-hand-side vector $\mathbf{b}$ are known to 8 significant figures, and

$$\|A\| = 1.9 \times 10^1, \qquad \|A^{-1}\| = 2.2 \times 10^3.$$

i) What is the condition number $\kappa(A)$?

ii) How many significant figures are there in a computed solution to $A\mathbf{x} = \mathbf{b}$?

d) You want to find an intersection of the functions

$$f_1(x) = \frac{1}{1+x^2} \quad \text{and} \quad f_2(x) = \log(|x|).$$

i) Define a function f to reformulate this as the problem of solving the nonlinear equation $f(x) = 0$.

ii) Write a MATLAB anonymous function `myfun` to evaluate f. Your function must work for a vector of input values, as well as a scalar input argument.

iii) The Secant method produces iterates x_k with errors $e_k = |x^* - x_k|$, where $x^* = 1.40132162219381544 21$ has been found to high accuracy by another method. The errors satisfy

k	e(k)	e(k+1)/e(k)	e(k+1)/e(k)^2	e(k+1)/e(k)^3
1	4.01e-01	1.49e+00	3.72e+00	9.26e+00
2	5.99e-01	1.71e-01	2.85e-01	4.76e-01
3	1.02e-01	2.53e-01	2.48e+00	2.43e+01
4	2.58e-02	4.37e-02	1.69e+00	6.55e+01
5	1.13e-03	1.11e-02	9.80e+00	8.67e+03
6	1.25e-05	4.84e-04	3.87e+01	3.09e+06
7	6.06e-09	5.35e-06	8.83e+02	1.46e+11
8	3.24e-14			

Giving reasons, answer the following questions.

A) Is the method converging?

B) Is the order of convergence linear, super-linear or quadratic?

2. Answer in a separate book marked Question 2

a) You have measurements at times t_i of the concentration $c_i = c(t_i)$ of a chemical which is being consumed, but not produced, in a reaction. The data is given in Table 2.1 and plotted in Figure 2.1.

i	0	1	2	3	4
t_i	0	0.5	1	1.5	2
c_i	1.0000	0.4283	0.5297	0.1344	0.0549

Table 2.1: Measured data values

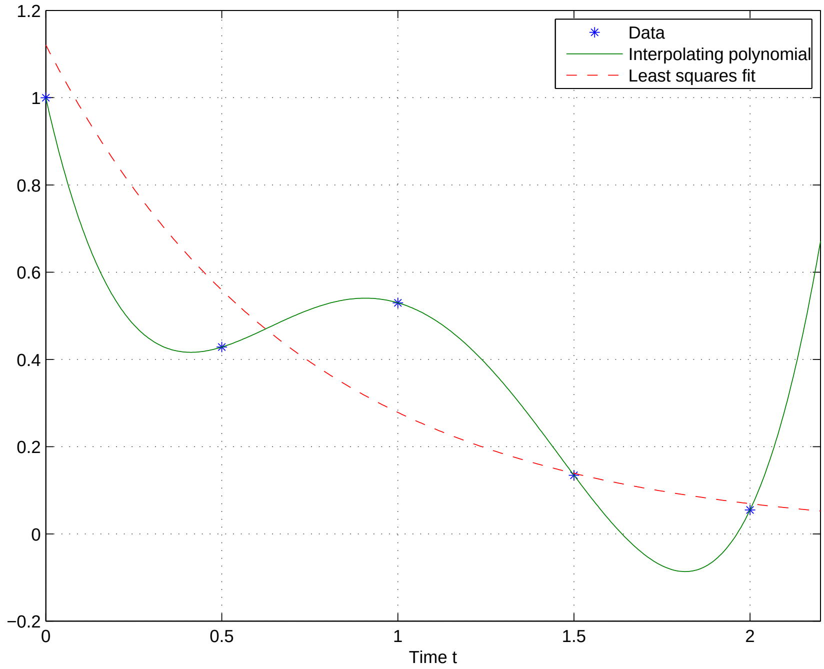

Figure 2.1: Data, polynomial interpolant and least squares approximation

i) You are asked to approximate $c(t)$ by the interpolating polynomial, which is illustrated in Figure 2.1.

A) What is the degree of the interpolating polynomial?

B) Give **two** reasons why this is not a sensible approximation.

ii) As the chemical is being consumed, an approximation of the form

$$c(t) = \alpha e^{-\beta t}$$

is thought to be more appropriate. Give MATLAB commands to calculate the least squares approximation of this form, assuming the data is in the column vectors `tdat` and `cdat`. You do **not** need to calculate values for α and β.

iii) Consider the problem of numerically approximating the integral

$$\int_a^b f(x)dx$$

using the values $f_i = f(x_i), i = 0, \ldots, N$ at the equally spaced points $x_i = a + ih, i = 0, \ldots, N$ where $h = (b-a)/N$. The Trapezoidal rule and Simpson's rule are, assuming N is even,

$$\begin{aligned} Q_N^{Trap}(f) &= h\left(\frac{f_0}{2} + f_1 + f_2 + \cdots + f_{N-1} + \frac{f_N}{2}\right) + O(h^2) \\ Q_N^{Simp}(f) &= \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + \cdots + 4f_{N-1} + f_N) + O(h^4) \end{aligned}$$

A) Use the data in Table 2.1 and Simpson's rule to estimate

$$\int_0^2 c(t)dt. \tag{2.1}$$

B) If additional data values at $t = 0.25, 0.75, 1.25, 1.75$ are measured, estimate how much the error in Simpson's method will decrease.

C) Can a Gauss-Legendre rule be used to estimate the integral (2.1) using just the data values in Table 2.1?

b) Consider the initial value problem (IVP)

$$\begin{aligned} &y''' + 2y'' - (\pi^2 + 1)y = \pi(\pi^2 + 1)e^{-t}\sin(\pi t) \\ &y(0) = 1, \quad y'(0) = -1, \quad y''(0) = 1 - \pi^2. \end{aligned}$$

i) Reformulate the IVP as a system of first-order differential equations

$$\mathbf{x}'(t) = \mathbf{f}(t, \mathbf{x})$$

with the appropriate initial condition.

ii) Write a MATLAB function M-file **or** anonymous function

`myode(t, x)`

to specify the function $\mathbf{f}$ in your first order system.

3. Answer in a separate book marked Question 3

Fick's second law, which predicts how the concentration of a chemical changes with time under the influence of diffusion, is

$$\frac{\partial c}{\partial t} = D\left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right), \tag{3.1}$$

where (x, y) is the position, t is the time, $c(x, y, t)$ is the concentration at position (x, y) and time t, and D is the constant diffusion coefficient. The

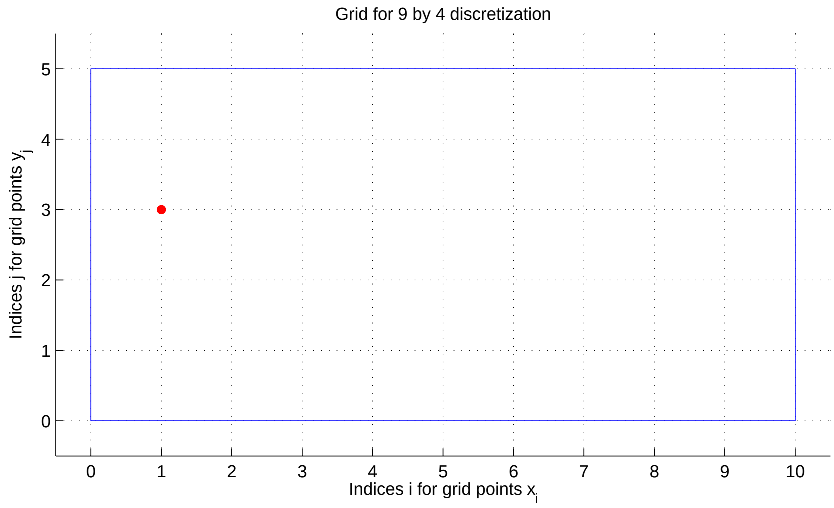

Figure 3.1: Discretization of space domain

space domain

$$\Omega = \{(x, y) : 0 \leq x \leq 2, \quad 0 \leq y \leq 1\}$$

is discretised using $h = 1/n$ and

$$x_i = ih, \quad i = 0, \ldots, 2n \qquad y_j = jh, \quad j = 0, \ldots, n.$$

This is illustrated in Figure 3.1 for $n = 5$. The time steps $t_\ell = \ell\Delta t$, for $\ell = 0, 1, 2, \ldots$ are used to calculate the approximate concentrations

$$c_{i,j}^{\ell} \approx c(x_i, y_j, t_\ell).$$

Standard finite difference approximations for a function f of **one** variable are

$$\begin{aligned}
f'(x) &= \frac{f(x+h) - f(x)}{h} + O(h), \\
f'(x) &= \frac{f(x+h) - f(x-h)}{2h} + O(h^2), \\
f''(x) &= \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + O(h^2).
\end{aligned}$$

a) What additional information is needed to completely specify the problem?

b) What is the steady-state version of Fick's second law (3.1)?

c) At the point (x_i, y_j) and time $t_{\ell+1}$ give a backward difference approximation of $O(\Delta t)$ to $\frac{\partial c}{\partial t}$.

d) At the point (x_i, y_j) and time $t_{\ell+1}$ give central difference approximations of $O(h^2)$ to $\frac{\partial^2 c}{\partial x^2}$ and $\frac{\partial^2 c}{\partial y^2}$.

e) Using the finite difference approximations from the previous two parts, show that Fick's second law (3.1) is approximated by

$$\alpha c_{i,j}^{\ell+1} + \beta c_{i+1,j}^{\ell+1} + \beta c_{i-1,j}^{\ell+1} + \beta c_{i,j+1}^{\ell+1} + \beta c_{i,j-1}^{\ell+1} = c_{i,j}^{\ell}$$

and find α, β in terms of $s = (D\Delta t)/h^2 > 0$.

f) Suppose that $c(0, y, t) = 6$ for $0 < y < 1$ and $t > 0$, and $c(x, y, 0) = 3$ for all $(x, y) \in \Omega$. Find the equation at the grid point (x_1, y_3) marked in Figure 3.1 and at time t_1. Clearly indicate what the unknowns are.

g) Is this an explicit method or an implicit method? Briefly discuss the relative advantages and disadvantages of an implicit method compared to an explicit method.

h) You are given that using a row-ordering of the variables $c_{i,j}^{\ell}$ produces the coefficient matrix A whose non-zero entries are illustrated in Figure 3.2.

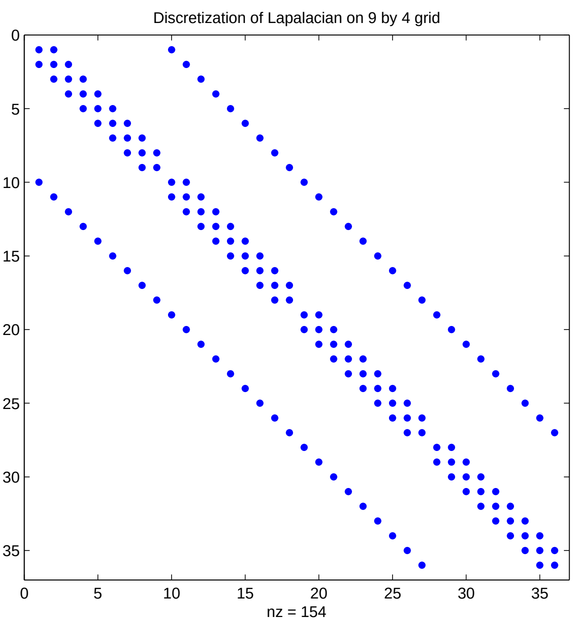

Figure 3.2: Spy plot of coefficient matrix

i) Calculate the sparsity of A.

ii) Why is explicitly calculating A^{-1} not a good idea?

iii) What other structure does A have that will make solving a linear system $A\mathbf{x} = \mathbf{b}$ more efficient?