CS412 Spring Semester 2013

Practice Midterm #3

1. MULTIPLE CHOICE SECTION. Circle or underline the correct answer (or answers). You do not need to provide a justification for your answer(s).

   (1) Numerical integration rules which use just one point per interval are:
   (Circle or underline the ONE most correct answer)

   (a) Always first order accurate. Methods that use 2 points are second order accurate and so on.

   (b) At most first order accurate, but can never have an order of accuracy of 2 or more.

   (c) Generally first order accurate, but sometimes could also be second order due to fortuitious cancellation.

   Comments: The rectangle rule is first order, while the midpoint rule is second order. Both use just a single point

   (2) A certain numerical integration rule integrates all cubic polynomials exactly. Which of the following property describes its accuracy?
   (Circle or underline the ONE most correct answer)

   (a) The method is exactly third order accurate.

   (b) The method is at least fourth order accurate.

   (c) The global error is proportional to h^3.

   Comments: When any polynomial of up to $n-1$ degree is integrated exactly, the method is at least n-order accurate. This also means the global error scales like $O(h^n)$, the local like $O(h^{n+1})$

   (3) Which of the following methods are good choices for solving $\mathbf{Ax} = \mathbf{b}$, where A is symmetric and positive definite?
   (Circle or underline ALL correct answers)

   (a) **LU** factorization with full pivoting.

   (b) System of normal equations.

   (c) Gauss-Seidel method.

   (d) Jacobi method.

   Comments: Even if it were a viable choice, LU factorization would not require pivoting on a symmetric, positive definite matrix. (b) is for least squares problems and would unnecessarily square the condition number if applied to a square matrix. The Jacobi method requires diagonal dominance to guarantee convergence.

(4) If the $n \times n$ matrix $\mathbf{A}$ is poorly conditioned (i.e. it has a very large condition number), then ...
(Circle or underline the ONE most correct answer)

(a) Solving $\mathbf{Ax} = \mathbf{b}$ accurately would be difficult with LU decomposition or Gauss elimination, but iterative methods (Jacobi, Gauss-Seidel) would not have a problem.

(b) Solving $\mathbf{Ax} = \mathbf{b}$ accurately with iterative methods (Jacobi, Gauss-Seidel) would be difficult, but LU decomposition with pivoting would not have a problem.

(c) Solving $\mathbf{Ax} = \mathbf{b}$ accurately will be challenging regardless of the method we use.

(5) Imagine that we perform an evaluation of a certain composite integration rule, partitioning the integration interval $[a, b]$ into equal size subintervals, with length h. We observe that by doubling the number of data points, the error in the approximation of the integral is reduced by a factor of eight. Which of the following are true?
(Circle or underline ALL correct answers)

(a) The integration rule is third order accurate.

(b) The global integration error scales proportionately to h^4.

(c) The local integration error scales proportionately to h^4.

(6) Which of the following are good reasons for using an iterative method (e.g. Jacobi or Gauss-Seidel) instead of a direct method (e.g. Gauss Elimination or LU factorization) to solve the $n \times n$ system $\mathbf{Ax} = \mathbf{b}$?
(Circle or underline ALL correct answers)

(a) When an iterative method is convergent and the matrix $\mathbf{A}$ is relatively sparse, the computational cost of finding a good approximation of the simultion using an iterative approach could be significantly lower than using a direct method.

(b) Iterative methods work very well with poorly conditioned matrices, while direct methods face problems in this case.

(c) Iterative methods do not require pivoting when $\mathbf{A}$ is diagonally dominant or symmetric positive definite, while a direct method would require pivoting in this case.

(7) Consider the rectangular $m \times n$ matrix $\mathbf{A}$ (with $m > n$) and the vector $\mathbf{b} \in \mathbf{R}^m$. If $\mathbf{x}$ is the *least squares solution* to $\mathbf{Ax} \approx \mathbf{b}$, can we say that $\mathbf{x}$ is an actual solution to $\mathbf{Ax} = \mathbf{b}$?
(Circle or underline the ONE most correct answer)

(a) Yes, in fact $\mathbf{Ax} = \mathbf{b}$ has many solutions and the "least squares solution" is the one with the smallest 2-norm of the residual vector $\|\mathbf{r}\|_2$.

(b) No, the system $\mathbf{Ax} = \mathbf{b}$ will generally not have a solution. What we call the "least squares solution" is the vector $\mathbf{x}$ with the smallest 2-norm of the error vector $\|\mathbf{x} - \mathbf{x}_{\text{exact}}\|_2$.

(c) No, the system $\mathbf{Ax} = \mathbf{b}$ will generally not have a solution. What we call the "least squares solution" is the vector $\mathbf{x}$ with the smallest 2-norm of the residual vector $\|\mathbf{b} - \mathbf{Ax}\|_2$.

2. SHORT ANSWER SECTION. Answer each of the following questions in no more than 1-2 sentences.

(a) Describe one scenario where you would prefer using the **LU** factorization, instead of an iterative method such as Jacobi or Gauss-Seidel, for solving a linear system $\mathbf{Ax} = \mathbf{b}$.

When we have many systems $\mathbf{Ax_k} = \mathbf{b_k}$ to solve, with the same coefficient matrix. Or when the matrix is neither symmetric, positive definite or diagonally dominant. It can also be a viable method when the size of the system is very small.

(b) What is one benefit of the **QR** factorization when compared to the normal equations, as methods for solving least squares problems?

The QR factorization does not square the condition number of the system, while the normal equations do. Thus, the QR approach is better conditioned.

(c) Describe one reason why solving a system $\mathbf{Ax} = \mathbf{b}$ could be extremely challenging when $\mathbf{A}$ has a very high condition number.

Because tiny inaccuracies in the right hand side, or the solution methodology can translate to gigantic errors in the computed solution.

(d) List one of the conditions that would guarantee convergence of the Jacobi method for solving a linear system $\mathbf{Ax} = \mathbf{b}$.

When the matrix $\mathbf{A}$ is diagonally dominant.

(e) Why is Simpson's rule potentially much more attractive than the trapezoidal rule, when approximating definite integrals?

For a small, almost negligible increase in algorithmic complexity, Simpson's rule offers 4th order accuracy, as opposed to the 2nd order accurate Trapezoidal Rule.

(f) Describe one plausible stopping criterion for determining when to stop an iterative solver for $\mathbf{Ax} = \mathbf{b}$, such as Jacobi or Gauss-Seidel.

$\|\mathbf{b} - \mathbf{A}\mathbf{x}_k\|$=*small. Or* $\|\mathbf{x}_{k+1} - \mathbf{x}_k\|$=*small.*

3. Consider the $n > 3$ data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. We want to find a cubic polynomial $p(x) = c_3x^3 + c_2x^2 + c_1x + c_0$ such that the graph of $p(x)$ approximates the given data points as much as possible. Write a Least Squares system $\mathbf{Ax} \approx \mathbf{b}$ which can be used to determine this cubic approximating polynomial. What does this system reduce to, in the case $n = 4$?

   Solution:

   We want to approximate:

$$c_3x_1^3 + c_2x_1^2 + c_1x_1 + c_0 \approx y_1$$
$$c_3x_2^3 + c_2x_2^2 + c_1x_2 + c_0 \approx y_2$$
$$\vdots$$
$$c_3x_n^3 + c_2x_n^2 + c_1x_n + c_0 \approx y_n$$

   Or in matrix form:

$$\begin{pmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ x_2^3 & x_2^2 & x_2 & 1 \\ x_3^3 & x_3^2 & x_3 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_n^3 & x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} c_3 \\ c_2 \\ c_1 \\ c_0 \end{pmatrix} \approx \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}$$

   Which is an $n \times 4$ least squares system $\mathbf{Ax} \approx \mathbf{b}$. When $n = 4$ the system becomes *square* (4×4) and an exact solution is obtainable. This is now simply the Vandermonde system for polynomial interpolation.

4. Show that the coefficient matrix in the system of normal equations (used for solving least squares problems) is always symmetric and positive definite.

   Solution:

   The system of normal equations is $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$. We need to show that $\mathbf{A}^T\mathbf{A}$ is symmetric and positive definite. The symmetry is easy to show, since

$$\left(\mathbf{A}^T\mathbf{A}\right)^T = \mathbf{A}^T\left(\mathbf{A}^T\right)^T = \mathbf{A}^T\mathbf{A}.$$

   For positive definiteness, we need to show, that $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax} \geq 0$ for all $\mathbf{x} \in \mathbf{R}$. We have in fact:

$$\mathbf{x}^T\mathbf{A}^T\mathbf{Ax} = (\mathbf{Ax})^T(\mathbf{Ax}) = \|\mathbf{Ax}\|_2^2 \geq 0$$

5. Consider the $n \times n$ linear system $\mathbf{Ax} = \mathbf{b}$.

   (a) Show that if $\mathbf{A}$ is diagonal, the Jacobi method converges after just one iteration.

   (b) Show that if $\mathbf{A}$ is lower triangular, the Gauss-Seidel method converges after just one iteration.

   Solution:

   In both cases, we start by considering the decomposition $\mathbf{A} = \mathbf{D} - \mathbf{L} - \mathbf{U}$. Also let us denote the exact solution by $\mathbf{x}^*$, i.e. $\mathbf{Ax}^* = \mathbf{b}$.

   (a) the first iteration of Jacobi's method will give:

$$\mathbf{Dx}^{(1)} = (\mathbf{L} + \mathbf{U})\mathbf{x}^{(0)} + \mathbf{b}$$

   Since $\mathbf{A}$ is diagonal, we have $\mathbf{A} = \mathbf{D}$ and $\mathbf{U} = \mathbf{L} = \mathbf{0}$. Thus, the previous equation becomes:

$$\mathbf{Ax}^{(1)} = \mathbf{0} \cdot \mathbf{x}^{(0)} + \mathbf{b}$$

$$\mathbf{Ax}^{(1)} = \mathbf{Ax}^*$$

$$\mathbf{x}^{(1)} = \mathbf{x}^*$$

   (b) the first iteration of Gauss-Seidel will give:

$$(\mathbf{D} - \mathbf{L})\mathbf{x}^{(1)} = \mathbf{Ux}^{(0)} + \mathbf{b}$$

   Since $\mathbf{A}$ is lower triangular, we have $\mathbf{A} = \mathbf{D} - \mathbf{L}$ and $\mathbf{U} = \mathbf{0}$. Thus, the previous equation becomes:

$$\mathbf{Ax}^{(1)} = \mathbf{0} \cdot \mathbf{x}^{(0)} + \mathbf{b}$$

$$\mathbf{Ax}^{(1)} = \mathbf{Ax}^*$$

$$\mathbf{x}^{(1)} = \mathbf{x}^*$$

6. Consider the Ordinary Partial Differential Equation

$$y'(t) = -\sin^2(t)y^3(t)$$

(a) Write the formula for the Forward Euler method applied to this differential equation

(b) Write the formula for the Backward Euler method applied to this differential equation.

(c) Write the iterative formula for Newton's method, used to solve the equation derived in question (b) above.

(a)

$$y_{k+1} = y_k - \Delta t \sin^2(t_k) y_k^3$$

(b)

$$y_{k+1} = y_k - \Delta t \sin^2(t_{k+1}) y_{k+1}^3$$

(c) Let us replace the unknown value y_{k+1} in (b) with the variable x. Then, the equation becomes:

$$f(x) = x - y_k + \Delta t \sin^2(t_{k+1}) x^3 = 0$$

Newton's method yields:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{x_k - y_k + \Delta t \sin^2(t_{k+1}) x_k^3}{1 + 3\Delta t \sin^2(t_{k+1}) x_k^2}$$

7. The numerical integration rule known as *Milne's Rule* is defined as:

$$\int_a^b f(x)dx \approx \frac{b-a}{3}\left[2f\left(\frac{3a+b}{4}\right) - f\left(\frac{a+b}{2}\right) + 2f\left(\frac{a+3b}{4}\right)\right]$$

Determine the order of accuracy of this method.

We start testing this rule on monomials of the form $f(x) = x^d$. We have:

- $f(x) = 1$:

$$I_{\text{rule}} = \frac{b-a}{3}[2-1-2] = b-a \equiv \int_a^b 1 \cdot dx$$

- $f(x) = x$:

$$I_{\text{rule}} = \frac{b-a}{3}\left[2\left(\frac{3a+b}{4}\right) - \left(\frac{a+b}{2}\right) + 2\left(\frac{a+3b}{4}\right)\right] =$$

$$= \frac{b^2}{2} - \frac{a^2}{2} \equiv \int_a^b x dx$$

- $f(x) = x^2$:

$$I_{\text{rule}} = \frac{b-a}{3}\left[2\left(\frac{3a+b}{4}\right)^2 - \left(\frac{a+b}{2}\right)^2 + 2\left(\frac{a+3b}{4}\right)^2\right] =$$

$$= \frac{b^3}{3} - \frac{a^3}{3} \equiv \int_a^b x dx$$

- $f(x) = x^3$:

$$I_{\text{rule}} = \frac{b-a}{3}\left[2\left(\frac{3a+b}{4}\right)^3 - \left(\frac{a+b}{2}\right)^3 + 2\left(\frac{a+3b}{4}\right)^3\right] =$$

$$= \frac{b^4}{4} - \frac{a^4}{4} \equiv \int_a^b x dx$$

- $f(x) = x^4$:

$$I_{\text{rule}} = \frac{b-a}{3}\left[2\left(\frac{3a+b}{4}\right)^4 - \left(\frac{a+b}{2}\right)^4 + 2\left(\frac{a+3b}{4}\right)^4\right] =$$

$$= \frac{b-a}{192}\left[37a^4 + 44a^3b + 30a^2b^2 + 44ab^3 + 37b^4\right] \neq \int_a^b x^4 dx$$

Since this rule integrates up to cubic poynomials exactly, it is fourth-order accurate.

8. Consider a numerical integration rule defined as:

$$\int_a^b f(x)dx \approx w_1 f(a) + w_2 f\left(\frac{2a+b}{3}\right) + w_3 f(b)$$

where w_1, w_2, w_3 are undetermined constants. Find the values of w_1, w_2, w_3 such that this rule becomes 3rd order accurate.

Solution:

For 3rd order accuracy, this rule needs to integrate exactly the monomials $1, x$ and x^2. Thus, we have:

- $f(x) = 1$:

$$I_{\text{rule}} = w_1 + w_2 + w_3$$

Thus we need:

$$w_1 + w_2 + w_3 = b - a \tag{1}$$

- $f(x) = x$:

$$I_{\text{rule}} = aw_1 + \left(\frac{2a+b}{3}\right) w_2 + bw_3$$

Thus we need:

$$aw_1 + \left(\frac{2a+b}{3}\right) w_2 + bw_3 = \frac{b^2}{2} - \frac{a^2}{2} \tag{2}$$

- $f(x) = x^2$:

$$I_{\text{rule}} = a^2 w_1 + \left(\frac{2a+b}{3}\right)^2 w_2 + b^2 w_3$$

Thus we need:

$$a^2 w_1 + \left(\frac{2a+b}{3}\right)^2 w_2 + b^2 w_3 = \frac{b^3}{3} - \frac{a^3}{3} \tag{3}$$

Equations (1,2,3) can be solved for the values of w_1, w_2, w_3 (the solution would be considered correct if stopped here). The exact solutions are $w_1 = 0, w_2 = 3(b-a)/4, w_3 = (b-a)/4$.