Instructions: Show your work. You may use a calculator but neither books nor notes.

1. Consider a binary floating-point number system containing numbers of the form

\[ \pm 0.d_1d_2d_3 \times 2^e, \quad d_1 \neq 0; \quad d_1, d_2, d_3 \in \{0, 1\}; \quad -4 \leq e \leq 6. \]

Suppose that the system uses a conventional rounding to the nearest policy to convert a real number to its binary floating-point number and to do floating-point arithmetic.

(a) What are the smallest and largest positive numbers (in decimal) in this floating-point system?

(b) What is the smallest floating-point number (in decimal) among those greater than 1?

(c) What is the unit roundoff of this floating-point system?

(d) What is the floating-point number representation of the number 9?

(e) Give an example that shows \(fl(fl(a + b) + c) \neq fl(a + fl(b + c))\), where \(a, b, c\) are floating-point numbers contained in this system.

2. Suppose that \(f(x)\) is continuously differentiable on an interval \([a, b]\) and \(a \leq x_0 < x_1 < x_2 \leq b\).

(a) Use either Lagrange's form or Newton's form to determine the polynomial \(p_2(x)\) that interpolates \(f(x)\) at \(x_0, x_1, x_2\).

(b) Prove that for all \(t \in [a, b]\),

\[ f(t) - p_2(t) = f[x_0, x_1, x_2, t](t - x_0)(t - x_1)(t - x_2). \]

(Hint: Consider a polynomial of degree at most 3 that interpolates \(f(x)\) at \(x_0, x_1, x_2, t\).)

(c) Compute \(p_2(x)\) for \(f(x) = \sin(\pi x)\) and \(x_0 = 1/4, x_1 = 1/2,\) and \(x_2 = 3/4\).

(d) Use the formula in (b) to derive a bound for \(\max_{0 \leq x \leq 1} |\sin(\pi x) - p_2(x)|\) with \(p_2(x)\) being obtained in (c).

(Hint: \(f[x_0, x_1, x_2, t] = \frac{f'''(\eta)}{3!}\) for some \(\eta\) satisfying \(\min\{x_0, x_1, x_2, t\} \leq \eta \leq \max\{x_0, x_1, x_2, t\}\).)

3. (a) Describe the secant method for finding a root of an equation \(f(x) = 0\).

(b) Let \(x_0, x_1, x_2, \ldots\) be the number sequence computed by applying the secant method to the equation \(x^2 - 1 = 0\). What is the (iterative) formula for computing \(x_k\)?

(c) Suppose that the sequence generated in (b) satisfies \(\lim_{k \to \infty} x_k = 1\). Prove

\[ \lim_{k \to \infty} \frac{|x_{k+1} - 1|}{|x_k - 1| \cdot |x_{k-1} - 1|} = \frac{1}{2} \]

and

\[ \lim_{k \to \infty} \frac{|x_{k+1} - 1|}{|x_k - 1|} = 0. \]
(d) Based on the results in (c), does the secant method applied to \( x^2 - 1 = 0 \) converge linearly, sublinearly, superlinearly, quadratically, or with other rates? Explain.

4. Consider numerical integration of the initial value problem

\[ y' = f(t, y), \quad y(0) = y_0. \]

Let \( \Delta t \) be a time step size and \( t_n = n\Delta t, \ n = 0, 1, \ldots \)

(a) Determine the parameters \( a_0, a_1, \) and \( \beta \) such that the local truncation error of the scheme

\[ y_{n+1} = a_0y_{n-1} + a_1y_n + \beta \Delta t f(t_{n+1}, y_{n+1}) \]

is of highest order.

(b) What is the local truncation error? What is the order of the scheme? Is the scheme consistent with the underlying differential equation? Justify your answer to the last question.

(c) Is the derived method an implicit method or an explicit method?

(d) Use the root condition to determine whether or not the scheme is stable.

(e) If the scheme is stable, determine if it is weakly stable or strongly stable.

5. (a) Describe an LU factorization of a matrix \( A = [a_{ij}] \in \mathbb{R}^{n,n}. \)

(b) Suppose that \( A = [a_{ij}] \in \mathbb{R}^{n,n} \) is diagonally dominant (by columns), i.e., \( |a_{ii}| > \sum_{j=1 \atop j \neq i}^n |a_{ji}|, \)

for \( i = 1, \ldots, n. \) Let

\[ A = \begin{bmatrix} a_{11} & \alpha^T \\ \beta & A_1 \end{bmatrix}, \]

where \( \alpha, \beta \in \mathbb{R}^{n-1} \) and \( A_1 \in \mathbb{R}^{n-1,n-1}. \) Prove that \( a_{11} \neq 0 \) and the matrix (of size \( n-1 \))

\[ B := A_1 - \frac{1}{a_{11}} \beta \alpha^T \]

is diagonally dominant.

(c) Suppose that \( A \in \mathbb{R}^{n,n} \) is diagonally dominant. Prove that \( A \) has an LU factorization.

(d) Suppose that \( A = [a_{ij}] \in \mathbb{R}^{n,n} \) is symmetric and diagonally dominant and has the property \( a_{ii} > 0 \) for \( i = 1, \ldots, n. \) Prove that \( A \) is positive definite.

6. (a) Define a Schur form (or Schur decomposition) of a real symmetric matrix \( A. \)

(b) Describe an algorithm for computing a Schur form of a real symmetric matrix \( A, \) including main strategies for storage and cost saving.

(c) Define a singular value decomposition (SVD) of a general real matrix.

(d) When \( A \) is real symmetric, can an SVD of \( A \) be obtained from a Schur form of \( A? \) Explain.
Numerical Analysis PhD Qualification Examination
August, 2008

Instructions: Show your work. You may use a calculator but neither books nor notes.

1. Consider the floating point number system where a non-zero number $x$ is stored in the form

$$x = \pm(0.a_1 \ldots a_t)2^e$$

with $a_i = 0$ or 1 ($i = 1, \ldots, t$), $a_1 \neq 0$, $t = 53$, and $-1021 \leq e \leq 1024$.

(a) Find the greatest and smallest positive numbers of this number system.
(b) Find the unit roundoff.
(c) Which of the following are numbers in this number system? Explain.

$$10, \quad 1 + 2^{-53}, \quad 1 - 10^{-53}, \quad 2^{1024}$$

2. (a) Consider Newton’s iteration

$$x_{n+1} = x_n - f(x_n)/f'(x_n)$$

for finding a root of $f(x) = 0$. Suppose that $x_n \to \alpha$ and $f'({\alpha}) \neq 0$. Show that the convergence order of the method is quadratic.

(b) Consider a variant of Newton’s method in which only one derivative is needed:

$$x_{n+1} = x_n - f(x_n)/f'(x_0).$$

Suppose that $x_n \to \alpha$ and $f'({\alpha}) \neq 0$. Find $C$ and $s$ such that

$$e_{n+1} = Ce_n^s,$$

where $e_n = x_n - \alpha$, for $n = 0, 1, \ldots$. What is the convergence order of this method? Explain.

(c) The secant method,

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}f(x_n),$$

requires no derivatives. Suppose that $x_n \to \alpha$ and $f'({\alpha}) \neq 0$. Show the convergence order of the method is $(1 + \sqrt{5})/2 \approx 1.62$.

3. Consider the polynomial interpolation problem: Given $n+1$ data points $(x_i, y_i)$, $i = 0, ..., n$, where $x_0, ..., x_n$ are distinct, find a polynomial $p(x)$ of degree no more than $n$ such that

$$p(x_i) = y_i, \quad i = 0, ..., n.$$

(a) Prove that such a polynomial $p(x)$ always exists and is unique.

(b) Suppose that $f(x)$ is a sufficiently smooth function and $y_i = f(x_i)$, $i = 0, ..., n$. Give the error formula for the interpolation.
(c) What is the Newton form of polynomial \( p(x) \)? Using this formula, find \( p(x) \) for the dataset

<table>
<thead>
<tr>
<th>( x )</th>
<th>( 5 )</th>
<th>-7</th>
<th>-6</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y )</td>
<td>1</td>
<td>-23</td>
<td>-54</td>
<td>-954</td>
</tr>
</tbody>
</table>

4. Consider Euler’s scheme

\[ y_{n+1} = y_n + hf(t_n, y_n) \]  \hspace{1cm} (1)

for integrating the initial value problem

\[ y' = f(t, y), \quad 0 < t \leq T < \infty \]
\[ y(0) = y_0. \]  \hspace{1cm} (2)

where \( f \) is assumed to be sufficiently smooth.

(a) What are the definition and meaning of the local truncation error? Find the local truncation error for the scheme.

(b) What is the order of the local truncation error and what is the order of the method? Explain.

(c) Find a bound on the global error. What can you say about the convergence of the method based on your result? [Hint: Convergence analysis typically starts with deriving the error equation.]

5. Consider an LU decomposition of a square matrix \( A \).

(a) Give the detailed iteration formulas for implementing an LU decomposition.

(b) If \( A \) is non-singular, is it always possible to carry out the above described LU decomposition without break-down? Explain.

(c) Is an LU decomposition unique if it exists? Explain.

(d) Describe two different types of matrices for which LU decomposition is guaranteed to exist. Give an example for each type.

6. (a) Show that every matrix \( A \in \mathbb{R}^{n \times n} \) has a QR factorization \( A = QR \), where \( Q \in \mathbb{R}^{n \times n} \) is an orthogonal matrix and \( R \in \mathbb{R}^{n \times n} \) is an upper triangular matrix with non-negative diagonal entries.

(b) Show that \( A \) is nonsingular if and only if the above factorization is unique.

(c) What is the relationship between the above QR factorization and the Gram-Schmidt process applied to the columns of \( A \)? Explain.
NUMERICAL ANALYSIS PHD QUALIFICATION EXAMINATION
MAY, 2008

Instructions: Show your work. You may use a calculator but neither books nor notes.

1. (a) Describe when a problem is ill-conditioned and when it is well-conditioned. (If you use
the phrase "conditioning number", then you need to give its definition or description.)
(b) Show that the recurrence relation

\[ x_n = 2x_{n-1} + x_{n-2} \]

with initial values \( x_0 \) and \( x_1 \) has a solution of the form

\[ x_n = A\lambda^n + B\mu^n. \]

(c) Consider the problem of computing \( x_n \) from \( x_0 \) and \( x_1 \) using the above recurrence. For
what values (or ranges) of \( x_0 \) and \( x_1 \) is the problem well-conditioned? Justify your
answer rigorously.

2. Consider the iterative scheme

\[ x_{n+1} = \frac{1}{2 + x_n}, \quad n = 0, 1, ... \]

for computing the continued fraction

\[ \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \ldots}}}, \]

where \( x_0 \geq 0 \) is an initial guess.

(a) Take an initial guess \( x_0 \geq 0 \) and carry out the calculations for two steps of the iteration.
(b) Show that the scheme is linearly convergent when it is convergent.
(c) State the contractive mapping theorem for a general mapping.
(d) Use the theorem to show that the above iterative scheme is convergent for any initial
guess \( x_0 \geq 0 \).

3. (a) Derive the error formula for the two-point Gaussian quadrature formula

\[ \int_{-1}^{1} f(x)dx \approx f(-s) + f(s), \]

where \( s = 1/\sqrt{3} \). What is the algebraic degree of precision of the formula? Explain it
briefly.
(b) Given \( N \) nodes \( x_i = -1 + 2i/N, \quad i = 0, ..., N, \) determine the composite rule for
\( \int_{-1}^{1} f(x)dx \)
using the two-point Gaussian quadrature formula.
(c) Derive the error formula for the composite rule. Determine the convergence order as
\( N \to \infty \).
4. Consider the scheme

\[ y_{n+1} = y_n + \frac{h}{2} \left[ f(t_n, y_n) + f(t_{n+1}, y_{n+1}) \right] \]  

for integrating the initial value problem \( y' = f(t, y), \ y(0) = y_0 \).

(a) Describe a convergent, iterative scheme for solving (1) for \( y_{n+1} \). Give the convergence proof of your scheme assuming that \( f(t, y) \) is sufficiently smooth and \( h \) is sufficiently small.

(b) Show that if (1) is solved "exactly" for \( y_{n+1} \), then the scheme is an \( O(h^2) \) method.

(c) Take (1) as a corrector and Euler's scheme as a predictor to construct the following predictor-corrector integrator:

\[
\begin{align*}
\tilde{y}_{n+1} & = y_n + hf(t_n, y_n) \\
y_{n+1} & = y_n + \frac{h}{2} \left[ f(t_n, y_n) + f(t_{n+1}, \tilde{y}_{n+1}) \right].
\end{align*}
\]

Show that the predictor-corrector scheme is an \( O(h^2) \) method.

(d) Explain for general predictor-corrector methods why \( O(h^p) \) predictors are often combined with \( O(h^{p+1}) \) correctors.

5. (a) Describe the singular value decomposition of a matrix.

(b) If \( \sigma_1, \sigma_2, \ldots, \sigma_r \) are the nonzero singular values of \( A \), show that \( \|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2} \).

(c) If \( \sigma_1 \) is the largest singular value of \( A \), show that \( \|A\|_2 = \sigma_1 \).

(d) Show that for a matrix \( A \in \mathbb{R}^{m \times n} \),

\[ \|A\|_2 \leq \|A\|_F \leq \sqrt{\text{rank}(A)} \|A\|_2. \]

(e) Given a matrix \( A \in \mathbb{R}^{n \times n} \), show

\[ \min_{\forall B: \text{rank}(B) = n-1} \|A - B\|_2 = \sigma_n, \]

where \( \sigma_n \) is the smallest singular value of \( A \). (Hint: Consider \( \|(A - B)x\|_2 \) where \( x \) satisfies \( Bx = 0 \) and \( \|x\|_2 = 1 \).)

6. This problem is about the QR algorithm for computing the eigenvalues of a real matrix \( A \in \mathbb{R}^{n \times n} \).

(a) Define the upper Hessenberg form of a square matrix.

(b) Describe an algorithm for reducing \( A \) to an upper Hessenberg matrix via orthogonal similarity transformations using Householder reflections.

(c) Describe two common shift strategies for the QR algorithm.

(d) Give a general procedure of the shifted QR algorithm based on the Hessenberg form. You don’t need to include deflation and stopping criterion.

(e) Explain the importance for using the shift strategy and the Hessenberg form in the QR algorithm.
January 2008 Numerical Analysis Qualifying Exam

Instructions: Complete all 6 questions. You may use a calculator but neither books nor notes.

1. Consider a $C^2$ function $f : \mathbb{R}^n \to \mathbb{R}^n$, $n \geq 1$.
   (i) Derive Newton’s method for approximating solutions $x$ of $f(x) = 0$.
   (ii) State a local convergence result for Newton’s method.
   (iii) Prove a local convergence result for Newton’s method.

2. The two-point Gaussian quadrature formula with error term has the form
   \[ \int_{-1}^{1} f(x)dx = A_1 f(-x_1) + A_1 f(x_1) + E(f), \quad E(f) = cf^{(d+1)}(\xi). \]
   Find $A_1, x_1, c$, and $d$.

3. The $s$-step backward differentiation formula (BDF) method is a multistep method of the form
   \[ \sum_{m=0}^{s} a_m y_{n+m} = h\beta f(t_{n+s}, y_{n+s}), \quad n = 0, 1, \ldots, \]
   where $a_s = 1$, for solving the initial value problem
   \[ y'(t) = f(t, y), \quad y(t_0) = y_0. \]
   (i) Find $a_0, \ldots, a_{s-1}$ and $\beta$ for $s$-order, $s$-step BDF methods.
   (ii) Show that the BDF methods are stable and hence convergent if and only if $1 \leq s \leq 6$.

4. Assume that $A \in \mathbb{R}^{m \times m}$ is symmetric and positive definite.
   (i) Let $A_1$ and $A_2$ be the matrices generated by two Cholesky steps:
   \[ A = G_1 G_1^T, \quad A_1 = G_1^T G_1, \]
   \[ A_1 = G_2 G_2^T, \quad A_2 = G_2^T G_2, \]
   where $G_i \in \mathbb{R}^{m \times m}, i = 1, 2$, are lower triangular with positive elements on the main diagonal.
   Show that $A_1$ and $A_2$ are symmetric and positive definite. (The first equations in (1) and (2) are Cholesky factorizations of $A$ and $A_1$, respectively, while the second equations in (1) and (2) define $A_1$ and $A_2$.)
   (ii) Let $B$ be the matrix produced by one $QR$ step applied to $A$, that is,
   \[ A = QR, \quad B = RQ, \]
   where $Q \in \mathbb{R}^{m \times m}$ is orthogonal and $R \in \mathbb{R}^{m \times m}$ is upper triangular with positive elements on the main diagonal.
   Show that $B = A_2$ where $A_2$ is defined in (i).
5. For $A \in \mathbb{R}^{m \times m}$ and $A = A^T$, the Rayleigh Quotient Iteration to approximate eigenvalues and eigenvectors of $A$ is defined by:

Select $\vec{u}_0 \in \mathbb{R}^m$ such that $\|\vec{u}_0\|_2 = 1$.

for $k = 1, 2, \ldots$

Compute $\lambda_{k-1} = \vec{u}_{k-1}^T A \vec{u}_{k-1}$.

Solve $(A - \lambda_{k-1}I)\vec{w} = \vec{u}_{k-1}$ for $\vec{w}$.

Compute $\vec{u}_k = \vec{w}/\|\vec{w}\|_2$.

Consider a $2 \times 2$ matrix $A = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$ with $d_1 > d_2$. Let $\vec{u}_{k-1} = [\alpha_{k-1}, \beta_{k-1}]^T \in \mathbb{R}^2$ with $\alpha_{k-1} \neq 0$ and $\beta_{k-1} \neq 0$.

(i) Express components of $\vec{u}_k = [\alpha_k, \beta_k]^T$ in terms of $\alpha_{k-1}$ and $\beta_{k-1}$.

(ii) Assume, in addition, that $\alpha_{k-1} \approx 1$ and $\beta_{k-1} \approx 0$. Show that $\|\vec{u}_k - \vec{e}_1\|_2 \approx \|\vec{u}_{k-1} - \vec{e}_1\|_2^3$, where $\vec{e}_1 = [1, 0]^T$. What is the significance of this result? **Hint:** $\sqrt{1 + x} \approx 1 + x/2$ for $x \approx 0$.

6.

(i) Assume that $A \in \mathbb{R}^{m \times n}$, $\vec{b} \in \mathbb{R}^m$ are given. Let $\vec{z} \in \mathbb{R}^n$ be such that $A^T A \vec{z} = A^T \vec{b}$. Show that

$$\|\vec{b} - A \vec{z}\|_2^2 = \|\vec{b} - A \vec{z}\|_2^2 + \|A(\vec{z} - \vec{x})\|_2^2 \quad \forall \vec{x} \in \mathbb{R}^n.$$ 

Use this result to prove that $\vec{z}$ minimizes $\|\vec{b} - A \vec{z}\|_2$ with respect to $\vec{x} \in \mathbb{R}^n$.

(ii) Assume that $\vec{b} \in \mathbb{R}^4$, $A = [\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4] \in \mathbb{R}^{4 \times 4}$ are given and that $\vec{a}_2 = \alpha \vec{a}_1 + \beta \vec{a}_3$ for some $\alpha, \beta \in \mathbb{R}$, and $\vec{a}_1, \vec{a}_3, \vec{a}_4$ are linearly independent. Reduce the problem of minimizing $\|\vec{b} - A \vec{z}\|_2$ with respect to $\vec{z} \in \mathbb{R}^4$ to the problem of solving the linear system $B^T B \vec{y} = B^T \vec{b}$, where $B = [\vec{a}_1, \vec{a}_3, \vec{a}_4]$. Express all solutions $\vec{x}$ of the minimization problem in terms of the unique solution $\vec{y}$ of the system $B^T B \vec{y} = B^T \vec{b}$. 
NUMERICAL ANALYSIS PHD QUALIFICATION EXAMINATION
AUGUST 2007

Do the following exercises correctly. Write your answers in your examination booklet. You may
use a graphing calculator but you may not use books, notes, or other assistance.

1. Consider the iteration scheme

\[ x_{k+1} = \frac{1}{2} \left( \sqrt{x_k} + \frac{1}{\sqrt{x_k}} \right), \quad x_0 = a > 0. \]

(a) Prove that \( x_1 > x_2 > \ldots > x_k > \ldots > 1 \) for any positive initial value \( a \neq 1 \). (HINT: First show that if \( a \neq 1 \), then \( x_k > 1 \) for \( k = 1, 2, \ldots \))

(b) Find the limit \( \lim_{k \to \infty} x_k \).

(c) Determine the convergence rate of the iteration.

2. Consider the real matrix

\[ A = \begin{bmatrix}
1 & 0 \\
3 & \ddots \\
& \ddots & \ddots \\
0 & \cdots & 2n-1
\end{bmatrix} - \mathbf{bb}^T, \]

where \( \mathbf{b} = [b_1, b_2, \ldots, b_n]^T \) is a real column vector satisfying \( b_1, \ldots, b_n > 0 \) and \( \sum_{k=1}^{n} b_k = 1 \).

(a) For \( k = 1, 2, \ldots, n \) let \( D_k \) be the disk in the complex plane with center \( 2k - 1 \) and radius \( b_k \), i.e.,

\[ D_k = \{ \lambda : |\lambda - (2k - 1)| \leq b_k \}. \]

Use the Gerschgorin circle theorem to show that each disk \( D_k \) contains one and only one eigenvalue of \( A \).

(b) Prove the stronger result: each of the \( n \) intervals

\[ [(2k-1) - b_k, (2k-1) + b_k], \quad k = 1, 2, \ldots, n \]

contains one and only one eigenvalue of \( A \).

3. (a) Show that \textit{polynomial interpolation is linear} in the following sense. Let \( x_0, x_1, x_2, \ldots, x_n \) be \( n+1 \) distinct interpolation points and let \( f \) and \( g \) be two arbitrary functions whose domain includes the interpolation points. Then, for any linear combination of \( f \) and \( g \),

\[ h = \alpha f + \beta g \] (where \( \alpha \) and \( \beta \) are constants), we have

\[ p_h = \alpha p_f + \beta p_g, \]

where \( p_h, p_f, \) and \( p_g \) are the polynomials interpolating \( h, f, \) and \( g \), respectively, at the interpolation points.

(b) Show that \textit{the n-th divided difference is linear} in the following sense. Let \( x_0, x_1, x_2, \ldots, x_n \) be \( n+1 \) distinct interpolation points and let \( f \) and \( g \) be two arbitrary functions whose domain includes the interpolation points. Then, for any linear combination of \( f \) and \( g \),

\[ h[x_0, x_1, x_2, \ldots, x_n] = \alpha f[x_0, x_1, x_2, \ldots, x_n] + \beta g[x_0, x_1, x_2, \ldots, x_n] \]
4. Consider a numerical integration rule of the form

\[ \int_0^2 f(x)dx \approx I(f) = w_1f(x_1) + w_2f(x_2). \]

(a) Determine the weights \(w_1, w_2\) and nodes \(x_1, x_2\) such that the rule \(I(f)\) has a possibly highest degree of precision.

(b) Determine the degree of precision of \(I(f)\) derived in 4a and justify.

(c) Suppose that \(f\) is sufficiently smooth on \([0,2]\). Find an expression in terms of derivatives of \(f(x)\) for the error \(\int_0^2 f(x)dx - I(f)\).

5. Let \(A \in \mathbb{R}^{n \times n}\) be symmetric, \(b, y \in \mathbb{R}^n\), and \(r = b - Ay\). Define the real symmetric matrix

\[ \delta A = \alpha [ry^T + yr^T - \alpha (r^Ty)yy^T], \]

where \(\alpha = \|y\|_2^{-2}\).

(a) Show that \(y\) is the exact solution to \((A + \delta A)y = b\).

(b) Show \(\|uv^T\|_2 = \|u\|_2\|v\|_2\) for any real vectors \(u, v\).

(c) Show \(\|\delta A\|_2 \leq 3\|r\|_2/\|y\|_2\). (Hint: Use the result in 5b.)

6. (a) Define the region of absolute stability of a linear, multi-step method for the numerical solution of the initial value problem \(y' = f(x, y), y(x_0) = y_0\).

(b) Sketch the absolute stability region of Euler’s method \(y_{n+1} = y_n + hf(x_n, y_n)\).

(c) Sketch the absolute stability region of Backward Euler’s method \(y_{n+1} = y_n + hf(x_{n+1}, y_{n+1})\).
Numerical Analysis PhD Qualification Examination
May 2007

Write your answers in your examination booklet. You may use a graphing calculator but you may not use books, notes, or other assistance.

1. Consider the interpolation problem: Find a polynomial of degree less than or equal to three interpolating the table

<table>
<thead>
<tr>
<th>x</th>
<th>0.0</th>
<th>0.5</th>
<th>1.0</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>f(x)</td>
<td>1.00000</td>
<td>1.64872</td>
<td>2.71828</td>
<td>7.38906</td>
</tr>
</tbody>
</table>

This table is generated using the function \( f(x) = e^x \).

(a) Prove that the solution to the interpolation problem exists and is unique.

(b) Find the interpolation polynomial in Newton’s form.

(c) Find the approximation at \( x = 0.8 \) using the interpolation polynomial.

(d) Estimate the error for the approximation found in Question 1c using the interpolation error term. Compare the error estimate with the exact error.

2. A fixed point of function \( f \) is a number \( x^* \) satisfying

\[
 f(x^*) = x^*.
\]

The functional iteration is defined as: \( x_{n+1} = f(x_n), \ n = 0, 1, 2, \ldots \) for a given initial approximation \( x_0 \). Suppose that \( f \) is arbitrarily differentiable on an interval \([a, b]\).

(a) State and justify sufficient condition(s) under which there exists a fixed point in \([a, b]\).

(b) State and justify additional condition(s) under which there exists a unique fixed point in \([a, b]\).

(c) Prove that, under the conditions stated in Question 2b, for any \( x_0 \in [a, b] \) the functional iteration produces a sequence of points \( \{x_n\} \) which converges to the fixed point of \( f \).

(d) Prove that the sequence \( \{x_n\} \) converges linearly (or faster) to the unique fixed point.

3. Let \( A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n \) and define \( \delta A \in \mathbb{R}^{n \times n} \) by \( \delta A = -\alpha y^T \), where \( r = Ay - b \) and \( \alpha = \|y\|^{-\beta} \).

(a) Show that \( y \) is the exact solution to \((A + \delta A)y = b\).

(b) Show that \( \|\delta A\|_F = \|r\|_2/\|y\|_2 \) \( (\|\cdot\|_F \) is the Frobenius norm) and

\[
\frac{\|\delta A\|_2}{\|A\|_2} \leq \frac{\|r\|_2}{\|A\|_2 \|y\|_2}.
\]
4. Prove the Bauer-Fike Theorem:

Let $A$ be an $n \times n$ matrix with a complete set of linearly independent eigenvectors and suppose the $V^{-1}AV = D$ where $V$ is nonsingular and $D$ is diagonal. Let $\delta A$ be a perturbation of $A$ and let $\mu$ be an eigenvalue of $A + \delta A$. Then $A$ has an eigenvalue $\lambda$ such that

$$|\mu - \lambda| \leq \kappa_p(V)\|\delta A\|_p, \quad 1 \leq p \leq \infty$$

where $\kappa_p(V)$ is the $p$-norm condition number of $V$.

5. (a) Consider a numerical integration rule of the form

$$\int_{-1}^{1} f(x)dx \approx Af\left(-\sqrt{\frac{3}{5}}\right) + Bf(0) + Cf\left(\sqrt{\frac{3}{5}}\right).$$

What is the linear system that must be solved in order to determine the coefficients $A, B, C$ so that the formula has a possibly highest algebraic degree of precision? Solve for $A, B, C$. What is the algebraic degree of precision?

(b) Derive the composite quadrature rule resulting from the application of the rule in Question 5a to the general integral $\int_{a}^{b} f(x)dx$ on a uniform partition $x_0 = a < x_1 < \cdots < x_n = b$.

6. Consider the numerical method

$$y_{n+1} = y_n + \frac{h}{2}[y_n' + y_{n+1}'] + \frac{h^2}{12}[y_n'' - y_{n+1}'']$$

for solving the initial value problem $y' = f(t, y), y(t_0) = y_0$ where $y_n' = f(t_n, y_n)$ and

$$y_n'' = \frac{\partial f(t_n, y_n)}{\partial t} + f(t_n, y_n)\frac{\partial f(t_n, y_n)}{\partial y}.$$
Numerical Analysis PhD Qualification Examination
January 2007

Do the following exercises correctly. Write your answers in your examination booklet. You may use a graphing calculator but you may not use books, notes or other assistance.

Part I: Definitions and basic concepts. Give short answers of between one and ten sentences to the following questions. There is no need to rederive formulas that you can remember. Little explanation is necessary. Do not write a computer program.

1. The following questions refer to numerical solutions of the initial value problem \( y' = -y^2, \ y(0) = 1 \). using a fixed step size \( h \). If appropriate, use the following notation. For \( x_j = x_0 + jh \), denote a particular method's approximation to \( y(x_j) \) by \( y_j \) and denote \( f(x_j, y_j) \) by \( f_j \).

(a) Define what is the local truncation error of a numerical method for "solving" the initial value problem.

(b) Define what is a convergent numerical method for solving the initial value problem. (Give a rigorously correct definition—don't just state a theorem.)

(c) Define what is a consistent numerical method. (Give a rigorously correct definition—don't just state a theorem.)

(d) i. Write down the Taylor method with \( O(h^3) \) local truncation error.
   ii. Write down any convergent Runge-Kutta method with local truncation error \( O(h^3) \) or smaller.
   iii. Write down any convergent linear multi-step method with local truncation error \( O(h^3) \) or smaller.

2. (a) Give an example of a matrix all of whose eigenvalues have condition number equal to one. Justify that your answer is correct.

(b) Give an example of a matrix with an eigenvalue which has condition number infinity. Justify that your answer is correct.

3. Use a divided difference table to find a polynomial \( p(x) \) of the lowest possible degree such that

\[ p(-1) = 4, \quad p'(-1) = -11, \quad p(1) = 2, \quad p'(1) = 1, \quad p''(1) = 2. \]

Show your work but do not bother to simplify the answer. (HINT: Check the answer.)

4. Determine constant weights \( w_0 \) and \( w_1 \) so that the quadrature rule

\[ \int_{-1}^{1} f(x) \, dx \approx I(f) = w_0 f \left( \frac{-1}{\sqrt{3}} \right) + w_1 f \left( \frac{1}{\sqrt{3}} \right) \]

has as degree of precision as large as possible.
II: Theory. Do the following exercises correctly. Explain and justify your answers fully.

1. (a) Briefly define what is the matrix norm subordinate to a vector norm \( \| \cdot \| \). (These are sometimes called operator norms.)
   (b) Let \( \| A \|_\infty \) denote the matrix norm subordinate to the vector norm \( \| x \|_\infty = \max_{1 \leq i \leq n} |x_i| \). Show that for all \( A \in \mathbb{R}^{m \times n} \),
   \[
   \| A \|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|,
   \]
   i.e., \( \| A \|_\infty \) is the maximum absolute row sum norm.

2. (a) State the Gerschgorin circle theorem. (HINT: Don't forget to state both parts of the theorem.)
   (b) Use the Gerschgorin circle theorem to show that if
   \[
   \| E \|_\infty < \min_{i \neq j} \left| \frac{d_{ii} - d_{jj}}{2} \right|
   \]
   then there is an ordering of the eigenvalues of \( D + E \), \( \mu_1, \mu_2, \mu_3, \ldots, \mu_n \) such that
   \[
   |d_{ii} - \mu_i| < \| E \|_\infty
   \]
   for all \( i = 1, 2, 3, \ldots, n \). (HINT: First show that the \( i \)-th Gerschgorin disk is contained in the disk with center \( d_{ii} \) and radius \( \sum_{j=1}^{n} |e_{ij}| \).)

3. Let \( A \in \mathbb{R}^{n \times n} \) be a nonsingular matrix and let \( b \in \mathbb{R}^n \) and \( r \in \mathbb{R}^n \) be nonzero vectors. Suppose that \( x \) is solution of the system of equations \( Ax = b \) and that \( \hat{x} \) is a solution to the perturbed system of equations \( A\hat{x} = b - r \).
   (a) Show that
   \[
   \left( \frac{1}{\| A^{-1} \|_2 \| A \|_2} \right) \frac{\| r \|_2}{\| b \|_2} \leq \frac{\| x - \hat{x} \|_2}{\| x \|_2}
   \]  
   (1)
   (b) Show that (1) is tight in the sense that for every \( \rho > 0 \) there exist vectors \( b \in \mathbb{R}^n \) and \( r \in \mathbb{R}^n \) for which \( \rho = \| r \|_2/\| b \|_2 \) and the inequality (1) holds with equality. (HINT: Set \( r \) and \( b \) to be scalar multiples of the first and last columns of \( U \) in the singular value decomposition \( A = U\Sigma V^T \).)

4. (a) Is floating point addition commutative? In other words, is it true that \( u \oplus v = v \oplus u \) for any pair of floating point numbers \( u \) and \( v \)? If so, prove it. If not, give a counter example.
   (b) Is floating point addition associative? In other words, is it true that \( u \oplus (v \oplus w) = (u \oplus v) \oplus w \) for any three floating point numbers \( u \), \( v \), \( w \)? If so, prove it. If not, give a counter example.
5. The fixed point iterations $x_{n+1} = x_n + \lambda(x_n)f(x_n)$ are a frequently reinvented set of methods for "solving" the nonlinear equation $f(x) = 0$. Each choice of $\lambda(x)$ gives a different numerical method. Assume that $f(x)$ and $\lambda(x)$ have at least two continuous derivatives and that the roots of $f(x)$ have multiplicity one.

(a) Under what condition on $\lambda(x)$ does the fixed point iteration converge linearly (or faster) to a root $x_*$ from initial conditions close enough $x_*$? Explain.

(b) Under what condition on $\lambda(x)$ does the fixed point iteration converge quadratically (or faster) to a root $x_*$ from initial conditions close enough $x_*$?

6. Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and let $b \in \mathbb{R}^n$ be arbitrary.

(a) What are the Krylov subspaces generated by $A$ and $b$.

(b) Give a concise definition of a Krylov subspace method for solving a system of linear equations $Ax = b$. What distinguishes one Krylov subspace method from another?

(c) What can you say about how many steps are needed to solve $Ax = b$ exactly (using mathematically exact arithmetic)? Prove it.

III: Practice. Do the following exercises correctly. Explain and justify your answers fully.

A typical floating point number system uses $t = 53$ base $\beta = 2$ digits with exponent bounds $-1021 \leq e \leq 1024$ and symmetric rounding. The following short MATLAB program is run using this typical floating point system.

1. Is it an infinite loop?

2. Does it generate a floating point overflow?

3. Approximately what is the final value of $k$ and what is its significance.

Explain your answers.

```matlab
M = 1;
k = 0;
while 1 + M > M
    M = 2*M;
k = k + 1;
end
k
```
Do the following exercises correctly. Write your answers in your examination booklet. You may use a graphing calculator but you may not use books, notes or other assistance.

**Part I: Definitions and basic concepts.** Give short answers of between one and ten sentences to the following questions. There is no need to rederive formulas that you can remember. Little explanation is necessary. Do not write a computer program.

1. This question concerns numerical methods for solving an initial value problem $y' = f(x, y)$, $y(0) = y_0$.
   (a) Define what is a Runge-Kutta method?
   (b) Define what is a linear, multi-step method.
   (c) Define what is a “convergent” method. (Give a rigorously correct definition. Don’t state a theorem.)
   (d) Define “local truncation error”. (Don’t waste time deriving the local truncation error of your example, just give a definition.)

2. (a) What is a Krylov subspace?
   (b) Give a concise definition of a Krylov subspace method for solving a system of linear equations $Ax = b$. What distinguishes one Krylov subspace method from another?
   (c) What is a preconditioner?

3. (a) Define the “Schur decomposition” of an $n$-by-$n$ matrix.
   (b) Briefly describe the QR algorithm with shifts of origin.

4. (a) Define “condition number”.
   (b) Give an example of a well-conditioned computational problem.
   (c) Give an example of an ill-conditioned computational problem.

**II: Theory.** Do the following exercises correctly. Explain and justify your answers fully.

1. (a) Consider the matrix

   $\begin{bmatrix}
   10 & 6 & 0 \\
   2 & 8 & \varepsilon \\
   0 & \varepsilon & 2
   \end{bmatrix}$, \quad \text{where} \ |\varepsilon| < 1/2.

   i. Sketch the Gerschgorin disks.
   
   ii. What is the minimum and maximum number of eigenvalues that Gerschgorin’s theorem states must lie in each of the Gerschgorin disks. Explain.

   (b) Suppose that $A \in \mathbb{C}^{n \times n}$ has an eigenvalue-eigenvector pair, $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ such that $\|x\|_\infty = |x_1| = |x_2|$. Show that $\lambda$ lies inside or on the boundary of at least two Gerschgorin disks.
2. (a) Let $A \in \mathbb{R}^{n \times n}$ be a nonsingular matrix and let $b \in \mathbb{R}^n$ be a nonzero vector. Suppose that $x$ is solution of the system of equations $Ax = b$ and that $\hat{x}$ is a solution to the perturbed system of equations $A\hat{x} = b - r$. Show that

$$\frac{\|x - \hat{x}\|_2}{\|x\|_2} \leq (\|A^{-1}\|_2 \|A\|_2) \frac{\|r\|_2}{\|b\|_2} \quad (1)$$

(b) Let $A \in \mathbb{R}^{n \times n}$, $b, r, x, \hat{x} \in \mathbb{R}^n$ be as in Exercise 2a and let $A = U\Sigma V^T$ be a singular value decomposition of $A$. Show that if $b$ is a scalar multiple of the last column of $U$, then

$$\frac{\|x - \hat{x}\|_2}{\|x\|_2} \leq \frac{\|r\|_2}{\|b\|_2}$$

independent of the condition number of $A$.

3. (a) Show that the $n$-th divided difference $f[x_0, x_1, \ldots, x_n]$ is independent of the order of $x_0, x_1, \ldots, x_n$.

(b) Show that $f[x_0, x_1, x_2, \ldots, x_n] = f^{(n)}(\xi)/n!$ for some number $\xi$ between the minimum and the maximum of the $x_i$'s.

4. Let $f(x)$ be a sufficiently smooth function on $\mathbb{R}$ with a multiple root at $\alpha \in \mathbb{R}$, i.e., suppose $f(x) = (x - \alpha)^m g(x)$ where $m > 0$ is an integer and $g(x)$ is a sufficiently smooth function such that $g(\alpha) \neq 0$.

(a) For each positive integer $m$, determine the rate at which Newton's method applied to $f(x)$ will converge starting from from $x_0 \neq \alpha$ but "close enough" to $\alpha$.

(b) Define $g(x)$ by

$$g(x) = \begin{cases} f(x)/f'(x) & \text{if } x \neq \alpha \\ 0 & \text{if } x = \alpha \end{cases}$$

Show that regardless of the value of the positive integer $m$, Newton's method applied to $g(x)$ converges at least quadratically starting from $x_0 \neq \alpha$ "close enough" to $\alpha$. 

---

2
5. (a) Determine constant weights \( w_0 \) and \( w_1 \) so that the quadrature rule
\[
\int_{-1}^{1} f(x) \, dx \approx I(f) = w_0 f\left(\frac{-1}{\sqrt{3}}\right) + w_1 f\left(\frac{1}{\sqrt{3}}\right)
\]
has as degree of precision as large as possible.
(b) Determine the actual degree of precision of the quadrature rule in Exercise 5a.
(c) Show that the quadrature formula has error in the form
\[
\int_{-1}^{1} f(x) \, dx - I(f) = cf^{(d+1)}(\xi)
\]
where \( d \) is the degree of precision, \( \xi \in [-1, 1] \), and \( c \) is a constant.
(HINT: Try Hermite interpolation. Don’t waste time evaluating messy integrals whose value is a constant.)

III: Practice. Do the following exercises correctly. Explain and justify your answers fully.

A typical floating point number system uses \( t = 53 \) base \( \beta = 2 \) digits with exponent bounds \(-1021 \leq e \leq 1024\) and symmetric rounding. The following short MATLAB program is run using this typical floating point system.

1. Is it an infinite loop?
2. Does it generate a floating point overflow?
3. Approximately what is the final value of \( k \) and what is its significance.

Explain your answers.

\[
\begin{align*}
M &= 1; \\
\text{k} &= 0; \\
\textbf{while} \ 1 \ + \ M \ > \ M \\
M &= 2 \* M; \\
\text{k} &= \text{k} \ + \ 1; \\
\textbf{end} \\
\text{k}
\end{align*}
\]
Numerical Analysis PhD Qualification Examination
January, 2006

Instructions: Show your work. You may use a calculator but neither books nor notes.

1. Consider function \( f(x) = \sin(x) \).
   (a) Show that \( f(x) \) is well conditioned when \( |x| < 1/10000 \).
   (b) Show that \( f(x) \) is ill-conditioned when \( |x - 2\pi| < 1/10000 \).
   (c) What can be done to improve the condition number of \( f(x) \) when \( |x - 2\pi| < 1/10000 \)? Justify your answer.

2. (a) Use the Gram-Schmidt process to find an orthogonal set of functions based on the given set of three functions \( \{1, x, x^2\} \) defined on \([0, 1]\).
   (b) Find the least-squares approximation to function \( f(x) = e^x \) defined on \([0, 1]\) using the orthogonal set of functions obtained in (a).

3. Consider the interpolation problem: Find a polynomial of degree less than or equal to three interpolating the table

   \[
   \begin{array}{c|ccc}
   x & 0 & 1 & 7 & 2 \\
   f(x) & 51 & 3 & 201 & 1 \\
   \end{array}
   \]

   (a) Prove that the solution to the interpolation problem exists and is unique.
   (b) Find the interpolation polynomial in Newton’s form.
   (c) Derive the error for the interpolation polynomial provided that function \( f = f(x) \) is sufficiently smooth.

4. (a) Find the coefficients and nodes of a Gaussian quadrature formula of the form

   \[
   \int_a^b f(x)dx \approx A_0 f(x_0) + A_1 f(x_1)
   \]

   (b) What is the degree of (algebraic) precision of the above formula with the found coefficients and nodes?
   (c) Prove that no Gaussian quadrature formula of the above form can be exact for all polynomials of degree \( \leq 4 \).

5. Let \( A \) be a real Hessenberg matrix of order \( n \).
   (a) Describe how to calculate the QR factorization of matrix \( A, A = QR \), using Householder reflections.
   (b) Show that the orthogonal matrix \( Q \) in the QR factorization is a Hessenberg matrix.
   (c) Show that the product \( RQ \) is also a Hessenberg matrix.
   (d) Describe the QR method with shift for computing all of the eigenvalues of matrix \( A \).
6. Let $A$ be a nonsingular square matrix of order $n$.

(a) Assume that $A$ can be expressed as

$$A = \begin{bmatrix} \hat{A} & d \\ c^T & \alpha \end{bmatrix} \quad c, d \in \mathbb{R}^{n-1}, \quad \alpha \in \mathbb{R}, \quad \hat{A} \in \mathbb{R}^{(n-1) \times (n-1)}$$

where $\alpha \neq 0$ and $\hat{A}$ is non-singular and has $LU$ factorization, $\hat{A} = \hat{L}\hat{U}$. Look for $A = LU$ in the form

$$A = \begin{bmatrix} \hat{L} & 0 \\ m^T & 1 \end{bmatrix} \begin{bmatrix} \hat{U} & q \\ 0 & \gamma \end{bmatrix} \quad m, q \in \mathbb{R}^{n-1} \quad \gamma \in \mathbb{R}$$

Show that $m$, $q$, and $\gamma$ can be found, and describe how to do so.

(b) Use the result in (a) to prove that a real, symmetric positive definite square matrix $A$ of order $n$ is guaranteed to have an $LU$ factorization, $A = LU$. 
Part I. Basic Concepts Give short answers to the following questions.

1. Briefly describe a practical, well-known, numerical method other than the rectangular rule, the trapezoid rule, and Simpson's rule for approximating the integral \( I(f) = \int_0^1 f(x) \, dx \).

2. Let \( x, y \in \mathbb{R}^2 \). Assume that the elements of \( x, y \) are floating-point numbers in a typical floating point system. The following algorithm computes the inner product \( x^T y \) using finite precision arithmetic,

\[
\begin{align*}
  s_1 &\leftarrow x_1 \otimes y_1, \\
  s_2 &\leftarrow s_1 \oplus x_2 \otimes y_2,
\end{align*}
\]

where \( \oplus \) and \( \otimes \) denote floating-point addition and floating-point multiplication, respectively. Suppose the algorithm is run using finite precision arithmetic with machine precision \( \mu \), and neither overflows nor underflows occur. (For the purpose of this exercise, machine precision and the unit round are the same.)

(a) Show that \( |s_2 - x^T y| \leq 2\mu(|x_1 y_1| + |x_2 y_2|) + O(\mu^2) \).

(b) Show that \( s_2 = x^T \bar{y} \) where \( \bar{y}_j = y_j (1 + \varepsilon_j) \) and \( \varepsilon_j \) is some number satisfying \( |\varepsilon_j| \leq 2\mu + O(\mu^2) \) for \( j = 1, 2 \).

3. (a) State sufficient condition(s) under which a function \( g \) has a unique fixed point \( x^* \) (i.e., \( x^* = g(x^*) \)) in an interval \([a, b]\), and for any initial \( x_0 \in [a, b] \), the number sequence \( \{x_n\} \) computed by the fixed-point iteration converges to \( x^* \).

(b) Let \( f(x) = e^x - x - 2 \) for \( x > 0 \). Function \( f(x) \) has a unique positive zero \( x^* \). The equation \( f(x) = 0 \) can be written as

\[
x = e^x - 2 =: g_1(x) \quad \text{or} \quad x = \ln(x + 2) =: g_2(x).
\]

So the number \( x^* \) is also a fixed point of both \( g_1 \) and \( g_2 \). If you have to apply the fixed-point iteration to one of \( g_1 \) and \( g_2 \) to compute \( x^* \), which function will you choose, \( g_1 \) or \( g_2 \), based on your answer in (a)? Explain.
4. Find the following factorizations for

\[ A = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}. \]

(a) Singular value decomposition,
(b) QR factorization,
(c) Schur decomposition.

Part II: Theory and Practice

Problem 1.

(a) Describe Newton's method for finding a root \( x^* \) of equation \( f(x) = 0 \).

(b) Suppose that \( f(x) \) is an arbitrary differentiable function and \( f'(x^*) \neq 0 \). Prove that the convergence rate of Newton's method is quadratic.

(c) Let \( x_0, x_1, \ldots, x_k, \ldots \) be a number sequence generated by applying Newton's method to the equation

\[ f(x) = x^2 - 1 = 0. \]

Suppose \( x_0 > 0 \). Prove \( \lim_{k \to \infty} x_k = 1 \).

Problem 2.

(a) Let \( A \) be an \( m \)-by-\( n \) matrix and \( B \) an \( n \)-by-\( p \) matrix..

(i) Show that \( \| A \|_2 \leq \| A \|_F \leq \sqrt{\min(m, n)} \| A \|_2 \).

(ii) Show that \( \| A \|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}| \).

(b) Let \( A \in \mathbb{R}^{n \times n} \) be a nonsingular matrix and \( b \in \mathbb{R}^n \) any vector. Suppose that \( x \) is solution of the system of equations \( Ax = b \) and that \( \hat{x} \) is an approximate solution with residual \( r = b - A\hat{x} \). Show that

\[ \frac{\| x - \hat{x} \|_2}{\| x \|_2} \leq \left( \| A^{-1} \|_2 \| A \|_2 \right) \frac{\| r \|_2}{\| b \|_2}. \]
Problem 3.

(a) State Gershgorin's theorem. (HINT: There are two parts in the theorem.)

(b) Prove Gershgorin's theorem.

(c) Consider the matrix

\[
A = \begin{bmatrix}
20 & 60 & 0 \\
10 & -20 & 1 \\
0 & 1 & -30
\end{bmatrix}
\]

Sketch the Gershgorin disks. State the minimum possible number of eigenvalues that each disk may contain. Use Gershgorin’s theorem to justify your answer.

Problem 4. Consider the numerical method

\[
y_{n+1} = \frac{1}{3}y_{n-1} + \frac{4}{3}y_n + \frac{2h}{3} f(t_n+1, y_{n+1}), \quad n = 0, 1, 2, \ldots
\]

for solving the initial value problem

\[
\frac{dy}{dt} = f(t, y), \quad y(0) = y_0,
\]

where \(h\) is a constant step size and \(y_{n-1}, y_n, y_{n+1}\) are approximations of the solution \(y(t)\) at \(t = t_{n-1}, t_n, t_{n+1}\), respectively.

(a) Is this a one-step method or multi-step method?

(b) Is this an explicit method or implicit method?

(c) What is the order of this method? Show your work.

(d) Is this method unstable, weakly stable, or strongly stable? Prove it.
Basic Concepts: Give short answers of between 1 and 10 sentences to the following questions.

1. (a) What is a simple, polynomial interpolative quadrature rule? Give a definition and an example. (HINT: Do not confuse this with composite quadrature rules.)
(b) It is known that high degree of precision Newton-Cotes integration formulas are unstable and cannot be used to obtain highly accurate approximations to integrals. Name and briefly explain two different approaches to get high accuracy approximations to integrals.

2. Name and briefly describe any numerical method other than the power method for calculating one or more eigenvalues of a symmetric (or nonsymmetric) matrix.

3. (a) Give an example of a “well-conditioned” system of equations \( Ax = b \). Justify by calculating a condition number.
(b) Give an example of an “ill-conditioned” system of equations \( Ax = b \). Justify by calculating a condition number.
(c) Give an example of a matrix with “well-conditioned” eigenvalues. 
   Briefly Explain.
(d) Give an example of a matrix with an “ill-conditioned” eigenvalue. 
   Briefly Explain.

4. (a) Give examples of a sublinearly convergent sequence, a linearly convergent sequence and a super linearly convergent sequence?
(b) Give an example of any super linearly convergent method to find a solution of a scalar nonlinear equation \( f(x) = 0 \) other than Newton's method.

5. A typical floating point number system uses \( t = 53 \) base \( \beta = 2 \) digits with exponent bounds \(-1021 \leq e \leq 1024\).
(a) Which real numbers are contained in the floating point system?
(b) A typical floating point system also contains some elements that are not numbers. What are they?
Theory:

1. (a) Show that if \( \| \cdot \| \) is the derived, subordinate matrix norm to a vector norm \( \| \cdot \| \), then \( \| f \| = 1 \).

(b) Let \( \| \cdot \| \) be a vector norm and let \( W \in \mathbb{R}^{n \times n} \) be a matrix. Define \( \nu(x) \) by \( \nu(x) = \| Wx \| \).
   i. Show that \( \nu(\cdot) \) is a vector norm if and only if \( W \) is nonsingular.
   ii. In the derived subordinate matrix norm, show that \( \nu(A) = \| WAW^{-1} \| \) for all \( A \in \mathbb{R}^{n \times n} \).

2. Assume \( b \in \mathbb{R}^m \), \( A \in \mathbb{R}^{m \times n} \), \( m < n \), the rank of \( A = m \). Then any \( x \in \mathbb{R}^n \) that minimizes \( \| b - Ax \|_2 \) is called a least squares solution of the under determined system \( Ax = b \). Show how to obtain the minimum 2-norm least squares solution of \( Ax = b \) using \( SVD \) of \( A^T \) given by \( A^T = USV^T \).

3. The fixed point iterations \( x_{n+1} = x_n + \lambda(x_n)f(x_n) \) are a frequently reinvented set of methods for "solving" the nonlinear equation \( f(x) = 0 \). Each choice of \( \lambda(x) \) gives a different numerical method. Assume that \( f(x) \) and \( \lambda(x) \) have at least two continuous derivatives and that the roots of \( f(x) \) have multiplicity one.
   (a) Under what condition on \( \lambda(x) \) does the fixed point iteration converge linearly (or faster) to a root \( x^* \) from initial conditions close enough \( x_0 \)? Explain.
   (b) Under what condition on \( \lambda(x) \) does the fixed point iteration converge quadratically (or faster) to a root \( x^* \) from initial conditions close enough \( x_0 \)?
   (c) Stephenson’s method uses \( \lambda(x) = -f(x)/(f(x + f(x)) - f(x)) \). What can you say about its rate of convergence?

4. Numerical Differential Equations. This problem is concerned with numerical methods for solving the initial value problem

\[
y' = f(x, y), \quad y(x_0) = y_0.
\]  \tag{1}

Consider the following implicit linear multi-step method

\[
y_{k+1} = y_{k-1} + \frac{h}{3}(f(x_{k+1}, y_{k+1}) + 4f(x_{k}, y_{k}) + f(x_{k-1}, y_{k-1})).
\]

(a) Determine the order of the local truncation error, i.e., find an integer \( p \) for which the local truncation error is \( O(h^{p+1}) \).

(b) Is this method stable? Weakly stable? Strongly stable?

(c) Is this method consistent?

(d) Is this method convergent?

(e) Could this method be recommended for use on general ordinary differential equations? Explain.

(f) To get started, the method needs \( y_0 \) and \( y_1 \). The initial value problem includes a given choice of \( y_0 \). Suggest a good way to obtain \( y_1 \).
Summer 2004 NA Qualifying Exam

Instructions: Complete all 6 questions.

Problem 1. In this problem, $\| \cdot \|$ denotes an arbitrary norm on $\mathbb{R}^n$. Throughout this exercise $A$ is a nonsingular $n \times n$ real matrix and $B$ is a singular $n \times n$ real matrix.

1. State the definition of the norm $\| A \|$ of an $n \times n$ matrix that is induced by the norm $\| \cdot \|$ on $\mathbb{R}^n$.

2. Let $x \in \mathbb{R}^n$ be the (exact) solution of the linear system of equations $Ax = b$, let $\hat{x}$ denote an approximate solution (e.g., $\hat{x}$ may have been computed using finite precision arithmetic). State the definitions of the relative error, the relative residual, and a definition of the condition number of $A$.

3. Derive an upper bound and a lower bound on the relative error in terms of the relative residual and a condition number of the matrix $A$.

4. Let $A$ be a nonsingular $n \times n$ matrix. Prove that for every singular $n \times n$ matrix $B$, $\|A - B\| \geq 1/\|A^{-1}\|$. (Alternatively, prove this for the 2-norm using the singular value decomposition.)

Problem 2. A fixed point of a function $f$ is a number $x^*$ such that

$$f(x^*) = x^*.$$  

The method of functional iteration is as follows. Starting with $x_0$, we define $x_{n+1} = f(x_n)$, where $n = 0, 1, 2, \cdots$. Suppose that $f$ is an arbitrary differentiable function defined on an interval $[a, b]$.

1. State sufficient condition(s) under which there exists a fixed point in $[a, b]$.

2. State additional condition(s) under which there exists a unique fixed point in $[a, b]$.

3. Prove that there exists a unique fixed point given the conditions you have stated. Prove that for any $x_0 \in [a, b]$ the method of functional iteration produces a sequence of points $\{x_n\}$ which converges to the fixed point of $f$.

4. Prove that the sequence $\{x_n\}$ converges linearly (or faster) to the unique fixed point.
Problem 3. Consider the problem
\[
\frac{dy}{dt} = -t^2 y \equiv f(t, y),
\]
y(0) = 1,
the implicit trapezoid method given by
\[
y_{n+1} = y_n + \frac{h}{2} [f(t_n, y_n) + f(t_{n+1}, y_{n+1})] \quad n = 0, 1, ...
\]
and the midpoint method
\[
y_{n+1} = y_{n-1} + 2hf(t_n, y_n) \quad n = 1, 2, ...
\]
where \( h = t_n - t_{n-1} = t_{n+1} - t_n \) and \( y_{n-1}, y_n, y_{n+1} \) are approximations of the exact solution \( y = y(t) \) of the continuous problem at \( t = t_{n-1}, t_n, t_{n+1} \), respectively.

1. Compute \( y_1 \) and \( y_2 \) for this problem using the implicit trapezoid method.
2. Determine the form of the local truncation error for the first step computed in part 1 using the implicit trapezoid method.
3. Derive the order of the implicit trapezoid method.
4. Using the value of \( y_1 \) obtained using the implicit trapezoid method in part 1, find \( y_2 \) using the midpoint method.
5. Derive the order of the midpoint method.

Problem 4. Let \( A \in C^{n \times n} \) and \( \rho = \max |\lambda_i| \) where \( \{\lambda_i\}_{i=1}^n \) are the eigenvalues of \( A \).

1. Assume that \( A \) is upper triangular. Show that for any \( \epsilon > 0 \) there exists a nonsingular \( X \in C^{n \times n} \) such that
\[
||X^{-1}AX||_2 \leq \rho + \epsilon.
\]

**HINT:** Set \( X = \text{diag}(1, a, ..., a^{n-1}) \) for some real number \( a \).

2. Prove this for the case of general matrix \( A \in C^{n \times n} \).
3. Conclude by showing that there is a constant \( M \) such that \( ||A^k||_2 \leq M(\rho + \epsilon)^k \) for \( k = 0, 1, ... \).

Problem 5. Consider an \( m \times n \) matrix \( A \) with \( m \geq n \) and rank(\( A \)) = \( n \). The matrix \( C = (A^T A)^{-1} \) arises in many statistical applications and is known as the variance-covariance matrix.

1. If the \( QR \) decomposition of \( A \) is given by \( A = QR \), then show that \( C = (R^T R)^{-1} \).
2. Assuming that \( R \) is given, provide an algorithm requiring about \( n^3/6 \) flops for computing the diagonal entries of matrix \( C, c_{11}, ..., c_{nn} \).
Problem 6. Consider the integral
\[ \int_{-1}^{1} w(x)f(x)\,dx. \]

1. Define Gaussian quadrature.

2. What are reasonable conditions on the weight function \( w(x) \). Explain.

3. Derive the two-point Gaussian quadrature formula with its associated error
\[ \int_{-1}^{1} f(x)\,dx = A_0 f(x_0) + A_1 f(x_1) + E(f), \quad E(f) = cf^{(d+1)}(\xi). \]

Find \( A_0, A_1, x_0, x_1 \) and \( c, d \).
Numerical Analysis PhD Qualification Examination
August 2002

Do the following exercises correctly. Show your work. You may use a calculator but neither books nor notes.

This

Part I: Definitions and Fundamental Concepts

Give short answers of between 1 and 10 sentences to the following questions. There is no need to derive formulae that you can remember. Little explanation is necessary. Do not write a computer program.

1. (a) What is a natural cubic spline?
   (b) Describe any spline that is not a natural cubic spline.

2. Classify each of the following sequences as “not convergent”, “sublinearly convergent”, “linearly convergent”, “super linearly convergent”, “quadratically convergent”, or “cubically convergent.” If more than one adjective applies, then list all that do apply.
   (a) 2.1, 2.01, 2.0001, 2.00000001, …, $2 + 10^{-2n}$, …
   (b) $x_0 = 0$, $x_1 = 1/2$, $x_2 = 2/3$, …, $x_n = n/(n+1)$, …
   (c) $x_0 = 16$ and for $n = 0, 1, 2, …$, $x_{n+1} = \sqrt{x_n}$.

3. Consider the real matrix $A = \begin{bmatrix} 0 & \beta \\ \gamma & \delta \end{bmatrix}$.
   (a) For which values of $\beta$, $\gamma$ and $\delta$ is $A$ an orthogonal matrix?
   (b) For which values of $\beta$, $\gamma$ and $\delta$ is $A$ a rotation? (Such matrices are sometimes called “plane rotations”, “Givens rotations” or “Jacobi rotations.”)
   (c) For which values of $\beta$, $\gamma$ and $\delta$ is $A$ a reflection? (Such matrices are sometimes called “elementary reflectors” or “Householder reflections.”)

4. A typical floating point number system uses $t = 53$ digit base $\beta = 2$ arithmetic with exponent bounds $-1021 \leq e \leq 1024$. Let $\oplus$, $\ominus$, $\otimes$ and $\oslash$ be the operations of floating point addition, subtraction, multiplication and division, respectively. Using the typical floating point system evaluate the following expressions. (Assume that powers are computed by repeated floating point multiplication.)

$$10^{50} \oplus 1 \oplus 10^{50}, \quad 10^{-1000}, \quad 0 \otimes 0, \quad -2 \oslash \infty,$$

$$10^{50} \otimes 10^{50} \oplus 1, \quad 10^{1000}, \quad \infty \oslash \infty, \quad \text{NaN} \ominus \text{NaN}.$$

5. Briefly describe any super linearly convergent numerical method for finding a root of a nonlinear equation $f(x) = 0$ other than Newton's method.
6. (a) Define what is the singular value decomposition of an $m$-by-$n$ matrix.
   
   (b) Let $A$ be an $m$-by-$n$ matrix and let $b$ be a vector of length $m$. Suppose that the singular value decomposition of an $m$-by-$n$ matrix $A$ is known. Explain how to use the known singular value decomposition to solve the linear least squares problem $\min_{x \in \mathbb{R}^n} ||b - Ax||_2$. (Don't forget to include the possibility of zero singular values.)

Part II: Theory and Practice

Write thorough answers to the following questions. Rigorously justify everything.

1. (a) Derive any convergent numerical method other than a Taylor method for numerical solution of a general initial value problem $y' = f(x, y)$, $y(x_0) = y_0$ that is of second or higher order.
   
   (b) Justify the claim that your method is of second or higher order.
   (HINT: Euler's method is a first order method.)

2. (a) State and prove the Gerschgorin circle theorem.
   
   (b) The roots of $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_0$ are the eigenvalues of the companion matrix

   $$C = \begin{bmatrix}
   0 & 1 & 0 & \cdots & 0 \\
   0 & 0 & 1 & \cdots & 0 \\
   \vdots & \vdots & \ddots & \vdots \\
   0 & 0 & \cdots & 0 & 1 \\
   -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1}
   \end{bmatrix}$$

   (If $n = 1$, then the companion matrix is $C = [-a_0].$) Throughout this exercise, assume that $n > 1$.

   i. Show that the roots of $p(\lambda)$ lie in the union of the disks $D_1 = \{z \in \mathbb{C} \mid |z + a_{n-1}| \leq |a_0| + |a_1| + \cdots + |a_{n-2}|\}$ and $D_0 = \{z \in \mathbb{C} \mid |z| \leq 1\}$.

   ii. If $|a_{n-1}| > 1 + |a_0| + |a_1| + \cdots + |a_{n-2}|$, how many eigenvalues lie in $D_1$? How many eigenvalues lie in $D_0$. (HINT: Draw a picture.)

   iii. Is it possible for $D_0$ to contain no root of $p(\lambda)$? If so, give an example. If not, prove it. Is it possible for $D_1$ to contain no root of $p(\lambda)$? If so, give an example. If not, prove it. (HINT: Don't forget to consider the case in which the hypothesis of Exercise 2(b)ii fails.)
3. (a) Let $r$ and $x$ be any vectors in $\mathbb{R}^n$ and consider the matrix $E = rx^T$. Show that $\|E\|_2 = \|E\|_F = \|r\|_2 \|x\|_2$.
(b) Suppose that $x \in \mathbb{R}^n$, $\|x\|_2 = 1$ is an approximate normalized eigenvector of $A \in \mathbb{R}^{n \times n}$ with approximate eigenvalue $\rho \in \mathbb{R}$. Show that $(A + E)x = \rho x$ for a matrix $E$ satisfying $\|E\|_2 = \|E\|_F = \|\rho x - Ax\|_2$.

4. Suppose that $f(x)$ is continuously differentiable on an interval $[a, b]$ and $p_1(x)$ is a polynomial of degree at most 1 that interpolates $f(x)$ at $x = a$ and $x = b$. Let $f[a, b, t]$ be the second divided difference of $f(x)$ evaluated at $a$, $b$ and $t \in [a, b]$.
   (a) Show that for all $t \in [a, b]$, $f(t) - p_1(t) = f[a, b, t](t - a)(t - b)$.
      (HINT: What degree at most 2 polynomial interpolates $f(x)$ at $x = a$, $x = b$ and $x = t$?)
   (b) Use Exercise 4a to show that
      \[
      \int_a^b f(t) \, dt = (b - a) \frac{f(a) + f(b)}{2} - \frac{(b - a)^3}{6} f[a, b, c]
      \]
      for some number $c$ in the interval $(a, b)$. Hint: It may be convenient to recall that
      \[
      \int_a^b (x - a) \, dx = \frac{(b - a)^2}{2},
      \]
      \[
      \int_a^b (x - a)(x - b) \, dx = \frac{-(b - a)^3}{6}.
      \]

5. Possibly the simplest way to compute the sum of $N + 1$ floating point numbers $t_0, t_1, t_2, \ldots t_N$ is the following repeated addition algorithm.

   $s \leftarrow t_0$
   for $n = 1, 2, 3, \ldots N$
   $s \leftarrow s \oplus t_n$
   end

   where $\oplus$ represents $t$ digit, base $\beta$ floating point arithmetic. Show that if $\frac{N+1}{2} \beta^{1-t} < .1$, then the computed value of $s$ satisfies
   \[
   s = \hat{t}_0 + \hat{t}_1 + \cdots \hat{t}_N
   \]
   where for $n = 0, 1, 2, \ldots N$, $\hat{t}_n = t_n(1 + \delta_n)$, $\delta_0 = \delta_1$ and for $n = 1, 2, 3, \ldots, N$, $|\delta_n| \leq 1.06(N - n + 1) \left(\frac{1}{2} \beta^{1-t}\right)$. 

3
Numerical Analysis PhD Qualification Examination
May 2002

Do the following exercises correctly. Show your work. You may use a calculator but
neither books nor notes.
This is a three hour, pass/fail examination.

Part I: Definitions and Fundamental Concepts
Give short answers of between 1 and 10 sentences to the following questions. There is no need to derive formulae that you can remember. Little explanation is necessary. Do not write a computer program.

1. Using any ad hoc method, find each of the following factorizations of
A = \[
\begin{bmatrix}
0 & 0 \\
-1 & 0
\end{bmatrix}
\].
   (a) Singular value decomposition,
   (b) QR factorization,
   (c) Schur decomposition.

2. Very briefly describe the following eigenvalue algorithms. Describe which eigenvalues the algorithm is designed to calculate.
   (a) The power method.
   (b) The QR algorithm.

3. (a) Define what is a floating point number in a typical floating point system? (Include in your description a typical base, a typical number of digits and typical exponent bounds.)
   (b) Using the typical floating point system you described in Question 3a, evaluate the following expressions. (Assume that powers are computed by repeated multiplication.)
   \[-2/0, \infty - \infty, 1 + 10^{-50}, 10^{-1000}, 10^{1000}\]

4. Briefly describe any practical, well-known, numerical method for approximating the integral \(\int_a^b f(x) \, dx\) that is not Trapezoid rule, Simpson’s rule or a Riemann sum.

5. Describe any super linearly convergent numerical method for finding a root of a nonlinear equation \(f(x) = 0\) other than Newton’s method.

6. (a) Give an example of a “well-conditioned” system of equations 
\[Ax = b.\] Justify by calculating a condition number.
   (b) Give an example of an “ill-conditioned” system of equations 
\[Ax = b.\] Justify by calculating a condition number.
   (c) Give an example of a matrix with “well-conditioned” eigenvalues. 
Briefly Explain.
   (d) Give an example of a matrix with an “ill-conditioned” eigenvalue. 
Briefly Explain.
Part II: Theory and Practice

Write thorough answers to the following questions. Rigorously justify everything. (Do not write a computer program.)

1. (a) Show that if \( x_0 \neq 2 \) is “close enough” to 2, then the sequence
\[
x_{n+1} = 2x_n^4 - 5x_n^3 + 6x_n^2 + 4x_n - 6
\]
converges cubically to 2.
(b) Find
\[
\lim_{n \to \infty} \frac{2 - x_{n+1}}{(2 - x_n)^3}
\]
assuming \( x_0 \) is close enough to 2.

2. Let \( \sigma_{\text{min}} \) be the smallest singular value of \( A \in \mathbb{R}^{n \times n} \).
   (a) Show that if \( B \in \mathbb{R}^{n \times n} \) is singular, then \( \| A - B \|_2 \geq \sigma_{\text{min}} \).
   (b) Show that there exists a singular matrix \( B \in \mathbb{R}^{n \times n} \) for which \( \| A - B \|_2 = \sigma_{\text{min}} \).

3. Suppose that \( p_n(x) \) is a polynomial of degree at most \( n \) that interpolates \( \sinh(x) = (e^x - e^{-x})/2 \) at \( n + 1 \) points in the interval \([-1, 1]\).
   Show that if \( x = 0 \) is one of the interpolation points, then for all \( x \in [-1, 1]\), the relative error satisfies the bound
\[
\left| \frac{\sinh(x) - p_n(x)}{\sinh(x)} \right| \leq \frac{2^{n-1}(e + 1/e)}{(n+1)!} < \frac{2^{n+1}}{(n+1)!}.
\]
(HINT: Recall \( |\sinh(x)| \geq |x| \).)

4. The following algorithm computes the inner product \( x^T y \) using finite precision arithmetic.

\[
s_0 \leftarrow 0 \\
\text{For } j = 1, 2, 3, \ldots, n \\
\quad s_j \leftarrow s_{j-1} \oplus x_j \odot y_j
\]
(Here \( \oplus \) and \( \odot \) denote floating point addition and floating point multiplication.)

Suppose the algorithm is run using finite precision arithmetic with unit round \( \mu, n \mu < 1/10 \), and neither overflows nor underflows occur.
   (a) Show that \( s_n = x^T y + \epsilon \) where \( |\epsilon| \leq 1.06n\mu \sum_{j=1}^{n} |x_j y_j| \).
   (b) Show that \( s_n = x^T \tilde{y} \) where \( \tilde{y}_j = y_j(1 + \epsilon) \) where \( |\epsilon| \leq 1.06n\mu \).

5. (a) State and prove the Gershgorin circle theorem.
   (b) Let \( A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \) and let \( E \in \mathbb{R}^{2 \times 2} \) be a perturbation matrix with \( |e_{ij}| \leq \epsilon, \ i, j = 1, 2 \). Use the Gershgorin theorem to show that the eigenvalues of \( A + E \) lie in a disk with center 1 and radius \( \epsilon + \sqrt{\epsilon}(1 + \epsilon) \). Justify fully. (HINT: Use a diagonal similarity transformation.)
6. (a) Derive any method other than a Taylor method for numerical solution of a general initial value problem $y' = f(x, y)$, $y(x_0) = y_0$ that has global error of second or higher order.

(b) Justify the claim that your method is of second or higher order.

(HINT: Euler's method is a first order method.)
NUMERICAL ANALYSIS PHD QUALIFICATION EXAMINATION
JANUARY 2002

Do the following exercises correctly. Show your work. You may use a calculator but neither books nor notes. This is a three hour, pass/fail examination.

Part I: Definitions and Fundamental Concepts

Give short answers of between 1 and 10 sentences to the following questions. There is no need to rederive formulae that you can remember. Little explanation is necessary. Do not write a computer program.

1. It is known that high degree of precision Newton-Cotes integration formulae are unstable and cannot be used to obtain highly accurate approximations to integrals. Name and briefly explain two different approaches to get high accuracy approximations to integrals.

2. For each of the following matrices state whether it is an elementary reflector, a plane rotation, or an elementary eliminator.

   a. \( \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \)
   b. \( \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \)
   c. \( \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \)
   d. \( \begin{bmatrix} 1 & 0 \\ \frac{1}{\sqrt{2}} & 1 \end{bmatrix} \)
   e. \( \begin{bmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & \frac{3}{5} \end{bmatrix} \)
   f. \( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 1 \end{bmatrix} \)

(Elementary reflectors are also called a "Householder reflections". Plane rotations are also called "Jacobi rotations" or "Givens rotation". Elementary eliminators are also called "Gauss transformations:].

3. Let \( A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \).
   a. Find an \( LU \) factorization of \( A \) or briefly explain why it doesn't exist.
   b. Find a \( QR \) factorization of \( A \) or briefly explain why it doesn't exist.
   c. Find a singular value decomposition of \( A \) or briefly explain why it doesn't exist.

4. (a) Give an example of a "well-conditioned" system of equations \( Ax = b \). Justify by calculating a condition number.
   (b) Give an example of an "ill-conditioned" system of equations \( Ax = b \). Justify by calculating a condition number.
   (c) Give an example of a matrix with "well-conditioned" eigenvalues. Briefly Explain.
   (d) Give an example of a matrix with an "ill-conditioned" eigenvalue. Briefly Explain.
5. Consider the system of linear equations

\begin{align*}
2x + y &= 4 \\
x + 3y &= 3
\end{align*}

(a) Calculate two steps of the Jacobi method for solving a system of linear equations. Start with \(x_0 = y_0 = 0\).

(b) Calculate two steps of the Gauss-Seidel method for solving a system of linear equations. Start with \(x_0 = y_0 = 0\).

6. (a) Suppose that the sequence \(x_n\) converges superlinearly to a limit \(\alpha\). Evaluate \(\lim_{n \to \infty} \frac{|\alpha - x_{n+1}|}{|x_{n+1} - x_n|}\).

(b) Describe a practical application of this limit in a computer program to solve nonlinear equations.

Part II: Theory and Practice

Write thorough answers to the following questions. Rigorously justify everything. (Do not write a computer program.)

1. (a) Derive any second order method other than a Taylor method for numerical solution of a general initial value problem \(y' = f(x, y)\), \(y(x_0) = y_0\).

(b) Prove that your method is actually of second order.

2. Define \(z_+\) by \(z_+ = (z + |z|)/2\).

(a) Show that for \(w \in \mathbb{R}\), \((z - w)^3_+\) is a cubic spline with a single knot at \(z = w\).

(b) Show that if \(S(z)\) is a cubic spline with a knot at \(z = w\), a second knot at \(z = v\), \(w < v\), then \(S(z)\) has the form \(S(z) = p_3(z) + \alpha_1(z - w)^3_+ + \alpha_2(z - v)^3_+\) where \(p_3(z)\) is a cubic polynomial and \(\alpha_1, \alpha_2 \in \mathbb{R}\) are constants.

3. (a) State and prove the Gerschgorin circle theorem.

(b) Let \(D, S, T \in \mathbb{R}^{50 \times 50}\) be defined by

\[
D = \text{diag}(1, 2, 3, \ldots, 50)
\]

\[
T = \begin{bmatrix}
50 & -1 & -1 & \cdots & -1 \\
1 & 49 & -1 & \cdots & -1 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
1 & 1 & 1 & 2 & -1 \\
1 & 1 & 1 & 1 & 1
\end{bmatrix}
\]

Estimate each of the eigenvalues of the matrix \(B = D + 10^{-3}T\) to within an absolute error of \(10^{-4}\). Justify fully. (HINT: Use a diagonal similarity transformation.)
4. In this exercise, \( A \in \mathbb{R}^{n \times n} \), \( x, b, r \in \mathbb{R}^n \) satisfy \( Ax = b - r \). The Euclidean norm on \( \mathbb{R}^n \) is \( \|z\| = \sqrt{z^Tz} \). The notation \( \|z\|_2 \) is used unambiguously for the corresponding subordinate matrix norm.

(a) Show that matrix \( E = r x^T / x^T x \) is a matrix of smallest norm with the property that \((A + E)x = b\).

(b) Show that if \( A \) is nonsingular, then

$$\frac{\|A^{-1}b - z\|_2}{\|A^{-1}b\|_2} \leq \frac{\|A^{-1}\|_2 \|A\|_2 \|r\|_2}{\|b\|_2}.$$ 

5. (a) Use an interpolatory polynomial to derive the two point closed Newton-Cotes quadrature formula \( T_1(f) \) that uses \( f(a) \) and \( f(b) \) to estimate \( \int_a^b f(s) \, ds \).

(b) What is the degree of precision \( d \) of \( T_1(f) \). Justify.

(c) Derive an expression for the error \( \int_a^b f(s) \, ds - T_1(f) \) in terms of \( f^{(d+1)} \).

(d) Derive an expressions for the corresponding \( n \)-panel composite quadrature rule \( T_n(f) \) and the error \( \int_a^b f(s) \, ds - T_n(f) \).

(e) Show

$$\lim_{n \to \infty} \left( \frac{n}{b - a} \right)^2 \left( \int_a^b f(s) \, ds - T_n(f) \right) = -\frac{1}{12} \left( f'(b) - f'(a) \right).$$

(f) Use Exercise 5e to derive a corrected trapezoid quadrature rule with higher degree of precision.
Do the following exercises correctly. You may use a calculator but neither books nor notes. This is a three hour, pass/fail examination.

**Part I: Definitions and Fundamental Concepts**

Briefly answer the following questions. Give short answers of between 1 and 10 sentences. There is no need to rederive formulae that you can remember. Little explanation is necessary. Do not write a computer program. Keep answers short.

1. (a) What is a cubic spline?
   (b) What is Hermite interpolation? (This is sometimes called osculatory interpolation.)

2. Describe *exactly one* of the following common numerical quadrature procedures: Gaussian quadrature, Romberg integration, adaptive quadrature.

3. (a) Symmetric matrices are sometimes said to have “well-conditioned” or “perfectly conditioned” eigenvalues. Explain.
   (b) In the context of systems of linear equations $Ax = b$, orthogonal matrices are sometimes said to be “well-conditioned” or “perfectly conditioned”. Explain.

4. Describe any practical, superlinearly convergent, numerical method for finding a root of a nonlinear scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ other than Newton’s method. At what rate would one ordinarily expect the method to converge? (You may assume that $f$ is as smooth as necessary, has only a single, simple root, and that practical methods to evaluate $f$ and derivatives up to any order are available.)

5. Name and briefly describe any practical, numerically stable, numerical method for solving the nontrivial, general system of linear equations $Ax = b$ that is *not* a variation of Gaussian Elimination. A direct method, an iterative, or a semi-iterative numerical method or some other kind numerical method are all acceptable answers. (HINT: Some variations of Gaussian Elimination are also sometimes called LU, PLU or PLUQ factorizations.)
Part II: Theory and Practice

Give detailed, rigorous answers to the following questions.

1. A typical floating point number system uses 53 digit base 2 arithmetic, symmetric rounding, and exponent bounds $-1022, 1023$ and (mostly) conforms to IEEE standard. The binary operations $\oplus$ and $\otimes$ refer to floating point addition and multiplication of floating point numbers respectively.

(a) i. What is the smallest floating point number greater than one?
ii. What is the largest floating point number less than one?
iii. What is the largest, positive, finite floating point number?
iv. What is the smallest, positive, nonzero, normalized floating point number?

(b) Is floating point addition commutative? In other words, is it true that $u \oplus v = v \oplus u$ for any pair of floating point numbers $u$ and $v$? If so, prove it. If not, give a counter example.

(c) Is floating point addition associative? In other words, is it true that $u \oplus (v \oplus w) = (u \oplus v) \oplus w$ for any three floating point numbers $u, v, w$? If so, prove it. If not, give a counter example.

2. Let $A$ be an $n$-by-$n$ matrix.

(a) Describe the singular value decomposition (SVD) of $A$.

(b) Show that $\min_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_n$ where $\sigma_n$ is the smallest singular value of $A$.

(c) Use the singular value decomposition to construct a matrix $E \in \mathbb{R}^{n \times n}$ such that $A + E$ is singular and $\|E\|_2 = \sigma_n$.

(d) Show that if $E \in \mathbb{R}^{n \times n}$ is any matrix with the property that $A + E$ is singular, then $\|E\|_2 \geq \sigma_n$. (HINT: $A + E$ singular if and only if there is a nonzero vector $x \in \mathbb{R}^n$ for which $(A + E)x = 0$.)

(This shows that $\sigma_n$ is the 2-norm distance from $A$ to a 2-norm nearest singular matrix.)
3. For each \( \theta \in [0, 1] \), consider the \( \theta \)-method

\[
    y_{n+1} = y_n + h \left( \theta f(t_{n+1}, y_{n+1}) + (1 - \theta) f(t_n, y_n) \right)
\]

for numerical integration of the initial value problem \( y' = f(t, y) \), \( y(t_0) = y_0 \). For simplicity, assume that \( f(t, y) \) has infinitely many derivatives in both \( t \) and \( y \).

(a) For which value(s) of \( \theta \) is the \( \theta \)-method explicit? For which value(s) of \( \theta \) is the \( \theta \)-method implicit?

(b) What are the common names of the numerical methods obtained from the choices \( \theta = 0, \theta = \frac{1}{2} \) and \( \theta = 1 \)?

(c) Find an expression for the local truncation error in terms of \( \theta \), the stepsize \( h \), and derivatives of the exact solution \( y(x) \) to the initial value problem. What are the local truncation errors for the particular choices \( \theta = 0, \theta = \frac{1}{2}, \) and \( \theta = 1 \)?

(d) For which value(s) of \( \theta \) is the \( \theta \)-method convergent? Justify rigorously.

(e) For which values of \( \theta = 0, \theta = \frac{1}{2}, \) and \( \theta = 1 \) is the \( \theta \)-method A-stable?

4. Throughout this exercise, assume that \( f(x) = (x - \alpha)^m h(x) \) where \( m \) is a positive integer and \( h(x) \) is an infinitely differentiable function with \( h'(\alpha) \neq 0 \). In other words, \( f \) has a root of positive integer multiplicity \( m \) as many derivatives as you may need.

(a) Prove that if \( m > 1 \) and if one starts from any initial “guess” that is “close enough” to \( \alpha \), then Newton’s method converges linearly but not faster.

(b) Prove that if one starts from any initial “guess” that is “close enough” to \( \alpha \), then the modified Newton’s method

\[
    x_{n+1} = x_n - \frac{mf(x_n)}{f'(x_n)}
\]

converges superlinearly.

(HINT: Use a well known theorem about stationary iterative methods of the form \( x_{n+1} = g(x_n) \). Optionally prove the theorem.)
NUMERICAL ANALYSIS PHD QUALIFICATION EXAMINATION
JUNE 2001

Do the following exercises correctly. You may use a calculator but neither
books nor notes.
This is a three hour, pass/fail examination.

Part I: Definitions and Fundamental Concepts

Answer five out of the following six short questions. Indicate on
the front of the exam booklet which five questions you answered and
which question you did not answer.

Give short answers of between 1 and 10 sentences. There is no need
to rederive formulae that you can remember. Little explanation is
necessary. Do not write a computer program. Keep answers short.

1. Approximate \( \sqrt{3} \) using two steps of the secant method starting
   with initial "guesses" \( x_{-1} = 1, x_0 = 2 \).

2. Find a polynomial \( p(x) \) of the lowest possible degree such that
   \[ p(-2) = -12, \quad p(-1) = -2, \quad p(0) = 0, \quad p(1) = 0, \quad p(2) = 4 \]
   Show your work but do not bother to simplify your answer.

3. (a) Give any particular example of an elementary reflector. (It is
   also called a "Householder reflection").
   (b) Give any particular example of a plane rotation. (It is also
   called a "Jacobi rotation", "Givens rotation", and sometimes
   other names as well.)

4. (a) Describe any convergent method for the initial value problem
   \( y' = f(x, y), y(x_0) = y_0 \) other than Euler's method. (HINT:
   Any second order method is not Euler's method.)
   (b) Classify the method described in Question 4a as a Taylor
   method, a Runge-Kutta method, a linear multistep method,
   an implicit method, an explicit method, and/or a predictor/corrector method. (HINT: Possibly more than one of
   these terms applies.)
5. Name and briefly describe any practical, numerically stable, numerical method for solving the nontrivial, general system of linear equations $Ax = b$ other than Gaussian Elimination with pivoting. A direct method, an iterative, or a semi-iterative numerical method or some other kind numerical method are all acceptable answers. (HINT: Gaussian Elimination with pivoting is sometimes called a PLU or PLUQ factorization.)

6. Name and briefly describe any practical, numerically stable, numerical method for calculating an eigenvalue and a corresponding eigenvector of a general, real symmetric matrix $A = A^T$ other than the power method. (A method that calculates more than one eigenvalue-eigenvector pair is acceptable. A method which applies to both symmetric and non-symmetric matrices is also acceptable.)

Part II: Theory and Practice

Answer five out of the following six short questions. Indicate on the front of the exam booklet which five questions you answered and which question you did not answer.

Give short answers of between 1 and 10 sentences. Write thorough answers to the following questions. Rigorously justify everything. (Do not write a computer program.)

1. (a) Show that the equation $x = 1 + \arctan(x)$ has at least one solution $x = \alpha \in [1, 1 + \pi/2]$. (HINT: If $x \in [1, 1 + \pi/2]$, then $0 < \arctan(x) < \pi/2$.)

(b) Counted according to multiplicity, how many solutions lie in $[1, 1 + \pi/2]$. Justify.

(c) Prove the iteration $x_{n+1} = 1 + \arctan(x_n)$, will converge from every $x_0 \in [1, 1 + \pi/2]$.

(d) Find the order of convergence of the fixed point iteration in Exercise 1c. If convergence is linear, find the rate of linear convergence in terms of $\alpha$.

(HINT: $d(\arctan(x))/dx = 1/(1 + x^2)$.)
2. The following short MATLAB program is run using a typical floating point arithmetic. Is it an infinite loop? Does it generate a floating point underflow? Explain. Approximately what is the final value of \( E \) and explain its significance.

\[
E = 1 \\
\text{while } 1 + E > 1 \\
\quad E = E/2; \\
\text{end} \\
E = 2*E
\]

3. (a) Show that the \( (f, g) = \int_{-1}^{1} f g + f' g' \, dx \) is an inner product on the space of continuously differentiable functions. (COMMENT: This is the \( H^1 \) inner product.)

(b) Find two polynomials of degree less than or equal to one which are orthogonal with respect to this inner product.

(c) Find the polynomial of degree one or less that best approximates the function \( f(x) = x^3 \) with respect to this inner product.

4. The corrected trapezoid rule

\[
\int_{a}^{b} f(x) \, dx \approx (b - a) \left( \frac{f(b) + f(a)}{2} \right) + (b - a)^2 \left( \frac{f'(b) - f'(a)}{12} \right)
\]

can be derived by integrating the cubic Hermite interpolation polynomial that interpolates \( f(x) \) and \( f'(x) \) at \( x = a \) and \( x = b \). (Do not waste time rederiving the corrected trapezoid rule. Answer only the questions that are asked.)

(a) Use the error formula for Hermite interpolation to obtain an error formula for the corrected trapezoid rule. (COMMENT: There is partial credit for an "educated guess" with some justification.) (HINT: Avoid some messy calculus by substituting \( t = x - a \) before integrating a degree four polynomial.)

(b) Derive the \( n \)-panel composite corrected trapezoid rule. How many values of \( f'(x) \) does the composite rule use?

(c) Derive the error term for the \( n \)-panel composite corrected trapezoid rule.
5. Let $A$ be an $n$-by-$n$ matrix.
   (a) Describe the singular value decomposition (SVD) of $A$.
   (b) Show that $\min_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_n$ where $\sigma_n$ is the smallest singular value of $A$.
   (c) Use the singular value decomposition to construct a matrix $E \in \mathbb{R}^{n \times n}$ such that $A + E$ is singular and $\|E\|_2 = \sigma_n$.
   (d) Show that if $E \in \mathbb{R}^{n \times n}$ is any matrix with the property that $A + E$ is singular, then $\|E\|_2 \geq \sigma_n$. (HINT: $A + E$ singular if and only if there is a nonzero vector $x \in \mathbb{R}^n$ for which $(A + E)x = 0$.)
   (This shows that a 2-norm closed singular matrix to $A$ lies at $\sigma_n$.)

6. Consider the following linear multi-step method

$$y_{k+1} = -y_k + 2y_{k-1} + h \left( \frac{5}{2} f(x_k, y_k) + \frac{1}{2} f(x_{k-1}, y_{k-1}) \right).$$

   (a) Determine the order of this method.
   (b) Is this method weakly stable? Strongly stable?
   (c) Is this method consistent?
   (d) Is this method convergent?

HINT:

$$\sum_{j=0}^{p} a_j = 1$$

$$- \sum_{j=1}^{p} ja_j + \sum_{j=-1}^{p} b_j = 1$$

$$\sum_{j=1}^{p} (-j)^i a_j + i \sum_{j=-1}^{p} (-j)^{i-1} b_j = 1$$
NUMERICAL ANALYSIS PHD QUALIFICATION EXAMINATION
JANUARY 2001

Do the following exercises correctly. Show your work. You may use a calculator but neither books nor notes.

This is a three hour, pass/fail examination.

Part I: Definitions and Fundamental Concepts

Give short answers of between 1 and 10 sentences to the following questions. There is no need to rederive formulae that you can remember. Little explanation is necessary. Do not write a computer program. Keep answers short.

1. Find a polynomial \( p(x) \) of the lowest possible degree such that

\[
p(-1) = 4, \quad p'(-1) = -11, \quad p(1) = 2, \quad p'(1) = 1, \quad p''(1) = 2.
\]

Show your work but do not bother to simplify the answer. (HINT: Check the answer.)

2. (a) Give any particular example of an elementary reflector. (It is also called a "Householder reflection".)

(b) Give any particular example of a plane rotation. (It is also called a "Jacobi rotation", "Givens rotation", and sometimes other names as well.)

3. (a) Describe any convergent numerical method for the initial value problem \( y' = f(x, y), \ y(x_0) = y_0 \) other than Euler's method. (HINT: Any second order method is not Euler's method.)

(b) Classify the method described in Question 3a as a Taylor method, a Runge-Kutta method, a linear multistep method, an implicit method, an explicit method, and/or a predictor/corrector method. (HINT: Possibly more than one of these terms applies.)

4. Name and briefly describe any practical, numerically stable, numerical method for calculating an eigenvalue and a corresponding eigenvector of a general, real symmetric matrix \( A = A^T \) other than the power method. (A method that calculates more than one eigenvalue-eigenvector pair is acceptable. A method which applies to both symmetric and non-symmetric matrices is also acceptable.)
Part II: Theory and Practice

Write thorough answers to the following questions. Rigorously justify everything. (Do not write a computer program.)

1. (a) Find the weights \( w_0 \) and \( w_1 \) for which the simple quadrature rule

\[
\int_{-1}^{1} f(x) \, dx \approx 2 \left( w_0 f \left( \frac{-1}{\sqrt{3}} \right) + w_1 f \left( \frac{1}{\sqrt{3}} \right) \right)
\]

is exact for polynomials of degree as high as possible, i.e., has as high a degree of precision as possible.

(b) What is the degree of precision of the resulting quadrature rule in Question 1a? (HINT: The degree of precision is not one.)

(c) Find a bound on the error

\[
E(f) = \int_{-1}^{1} f(x) \, dx - 2 \left( w_0 f \left( \frac{-1}{\sqrt{3}} \right) + w_1 f \left( \frac{1}{\sqrt{3}} \right) \right)
\]

where \( w_0 \) and \( w_1 \) are the weights you found in Question 1a. Prove that your bound is correct. (A correct but weak bound may get only partial credit.)

2. Suppose that \( f(x) \) and \( g(x) \) are twice continuously differentiable and that for some number \( a \in \mathbb{R} \), \( f(a) = 0 \) and \( f'(a) \neq 0 \). Consider the iteration

\[
x_{n+1} = x_n + f(x_n)/g(x_n).
\]

(a) What condition on \( f(x) \) and \( g(x) \) guarantees that for any initial iterate \( x_0 \) that is "close enough" to \( a \), the iteration will converge linearly to \( a \)? Derive the rate constant.

(b) What conditions on \( f(x) \) and \( g(x) \) guarantee that for any initial iterate \( x_0 \) that is "close enough" to \( a \), the iteration will converge quadratically to \( a \)? Prove it.

(c) What conditions on \( f(x) \) and \( g(x) \) guarantee that for any initial iterate \( x_0 \) that is "close enough" to \( a \), the iteration will converge cubically to \( a \)? Prove it.
3. In this problem, $\| \cdot \|$ denotes an arbitrary norm on $\mathbb{R}^n$. Throughout this exercise $A$ is a nonsingular $n \times n$ real matrix and $B$ is a singular $n \times n$ real matrix.

(a) State the definition of the norm $\|A\|$ of an $n \times n$ matrix that is induced by the norm $\| \cdot \|$ on $\mathbb{R}^n$.

(b) Let $x \in \mathbb{R}^n$ be the (exact) solution of the linear system of equations $Ax = b$, let $\hat{x}$ denote an approximate solution (e.g., $\hat{x}$ may have been computed using finite precision arithmetic). State the definitions of the relative error, the relative residual, and two different kinds of condition numbers of $A$.

(c) Derive an upper bound and a lower bound on the relative error in terms of the relative residual and a condition number of the matrix $A$.

(d) Let $A$ be a nonsingular $n \times n$ matrix. Prove that for every singular $n \times n$ matrix $B$, $\|A - B\| \geq 1/\|A^{-1}\|$. (Alternatively, prove this for the 2-norm using the singular value decomposition.)

4. Consider the following implicit linear multi-step method

$$y_{k+1} = y_{k-1} + \frac{h}{3} \left( f(x_{k+1}, y_{k+1}) + 4 f(x_k, y_k) + f(x_{k-1}, y_{k-1}) \right).$$

(a) Determine the order of this method.

(b) Is this method weakly stable? Strongly stable?

(c) Is this method consistent?

(d) Is this method convergent?

(e) Could this method be recommended for use on general ordinary differential equations? Explain.

HINT:

$$\sum_{j=0}^{p} a_j = 1$$

$$- \sum_{j=1}^{p} ja_j + \sum_{j=-1}^{p} b_j = 1$$

$$\sum_{j=1}^{p} (-j)^i a_j + i \sum_{j=-1}^{p} (-j)^{i-1} b_j = 1$$