**Department of Applied Mathematics**
**Preliminary Examination in Numerical Analysis**
**January 2024**

## Instructions

**You have three hours to complete this exam. Submit solutions to four (and no more) of the following six problems. Please start each problem on a new page. You MUST prove your conclusions or show a counter-example for all problems unless otherwise noted. Write your student ID number (not your name!) on your exam.**

**Problem 1: Rootfinding** (25 points)

Consider a rootfinding problem for $f \in C^\infty$ near a simple root $f(\alpha) = 0$. We apply a few iterative methods to solve it; each produces iterates $x_n$ with errors defined by $\epsilon_n = x_n - \alpha$.

1. Assume $x_n \to \alpha$. Write down a definition for *order* of convergence of an iterative method.
2. We wish to consider algorithms with iterates of the form:

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

   where $\gamma_n$ is some function of $x_n, x_{n-1}, \ldots, x_0$. Explain how the Newton-Raphson and Secant methods are both examples; for each, indicate their order of convergence.
3. Let $M(\alpha) = \frac{f''(\alpha)}{2f'(\alpha)}$, and assume $f''(\alpha) \neq 0$. Using an argument similar to the one used for Newton (involving Taylor expansions of $f(x)$ around $\alpha$), we can conclude that if the secant method is convergent,

   $$\epsilon_{n+1} = M(\alpha)\epsilon_n\epsilon_{n-1}$$

   Based on this formula, give intuition as to how we know whether the order is linear or superlinear (Hint: use an example). Derive an equation for the order of convergence $p$.

**Solution:**

1. The order of convergence $p > 0$ is defined by $p$ such that

$$\lim_{n\to\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = \lim_{n\to\infty} \frac{|\epsilon_{n+1}|}{|\epsilon_n|^p} = L$$

with $L \neq 0$. This tells us that, as $n$ goes to infinity, $|e_{n+1}| \simeq L|e_n|^p$.

2. Newton-Raphson sets $1/\gamma_n = f'(x_n)$, the slope of the line tangent to $f$ at $(x_n, f(x_n))$. If convergent to a simple root, its order is quadratic ($p = 2$). Secant method, as its name suggests, uses $1/\gamma_n = \frac{f(x_n)-f(x_{n-1})}{x_n - x_{n-1}}$, the slope of the secant based on the two most recent iterations. If convergent to a simple root, its order is superlinear, but smaller than quadratic ($p = \phi = 0.5(1 + \sqrt{5}) \simeq 1.618\ldots$).

3. Based on the formula, we observe that if $\ell_n = \log_{10} \epsilon_n$, we have that:

$$\ell_{n+1} = \ell_n + \ell_{n-1} + \log_{10} M(\alpha)$$

For large enough $n$, $\ell_{n+1}$ is negative and is smaller than $\ell_n, \ell_{n-1}$. Given an example where $M = 10, \epsilon_0 = 1e-2, \epsilon_1 = 1e-2$, we see that the exponents for the errors are $\{-2, -2, -3, -4, -6, -9, -14, -22, \ldots\}$, which is faster than linear (the number of digits gained is increasing and looks like Fibonacci).

To obtain the formula for $p$, we assume that $n$ is large enough that $\epsilon_{n+1} = L\epsilon_n$ (to first order), and so, $\ell_{n+1} = p\ell_n + \log_{10} L = p\ell_n + \lambda$. We can use either of these to find equations for $p$ and for $L$.

Using the first, we plug it in on both sides of $\epsilon_{n+1} = M\epsilon_n\epsilon_{n-1}$. We get the following:

$$\begin{aligned} L\epsilon_n^p &= LM\epsilon_{n-1}^{p+1} \\ L^{p+1}\epsilon_{n-1}^{p^2} &= LM\epsilon_{n-1}^{p+1} \\ \epsilon_{n-1}^{(p^2-p-1)} &= ML^{-p} \end{aligned}$$

Since this equation is approximately true for all $n$ sufficiently big, we must have $p^2 - p - 1 = 0$ and $ML^{-p} = 1$. The only positive root of $p^2 - p - 1$ is $p = \phi = 0.5(1 + \sqrt{2})$, and $L = M^{1/p}$.

**Alternatively:** we could have followed this derivation using the log errors. In that case,

$$\begin{aligned} p\ell_n + \lambda &= (p+1)\ell_{n-1} + \log_{10} M + \lambda \\ p^2\ell_{n-1} + (p+1)\lambda &= (p+1)\ell_{n-1} + \log_{10} M + \lambda \\ (p^2 - p - 1)\ell_{n-1} &= \log_{10} M - p\lambda \end{aligned}$$

the reasoning is the same: since $\ell_{n-1} \to -\infty$, we must have $p^2 - p - 1 = 0$ and $\log_{10} M - p\lambda = \log_{10} M - p\log_{10} L = 0$.

**Problem 2: Quadrature** (25 points)

In a number of applications in probability and statistics, we need to compute integrals of the form

$$E[h(y)] = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} h(y) e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy$$

for $h(y) \in C^\infty(\mathbb{R})$. This yields the expectation of $h(y)$ for a normally distributed random variable with mean $\mu \in \mathbb{R}$ and standard deviation $\sigma > 0$.

1. The Gauss-Hermite quadrature is a gaussian quadrature for integrals

$$I[f] = \int_{-\infty}^{\infty} f(x) e^{-x^2} dx$$

   with weight function $w(x) = e^{-x^2}$. Given nodes and weights $\{x_i, w_i\}_{i=1}^n$ for the n-point quadrature, write down an approximation for $E[h(y)]$ (Hint: use a change of variable).

2. Hermite polynomials are an orthonormal basis for the inner product $< f, g >_w = \int_{-\infty}^{\infty} f(x)g(x)w(x)dx$. Given that $H_0(x) = 1$, derive formulas for $H_1(x), H_2(x)$ (e.g. using Gram-Schmidt). You may use the formula

$$\int_{-\infty}^{\infty} x^n e^{-x^2} dx = \frac{((-1)^n + 1)}{2} \Gamma\left(\frac{n+1}{2}\right)$$

   with $\Gamma(\frac{1}{2}) = \sqrt{\pi}, \Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$, $\Gamma(\frac{5}{2}) = \frac{3\sqrt{\pi}}{4}$ (generally, $\Gamma(x+1) = x\Gamma(x)$).

3. Find the nodes and weights for the 2 point Gauss-Hermite quadrature.

**Solution:**

1. We first choose the linear change of variable $x = (y - \mu)/(\sqrt{2}\sigma)$, so that the exponential matches the Gauss-Hermite weight function. This corresponds to $y = \sqrt{2}\sigma x + \mu$ and $dy = \sqrt{2}\sigma dx$. So,

$$\begin{aligned} E[h(y)] &= \int_{\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} h(y) e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\ &= \int_{-\infty}^{\infty} \frac{\sigma\sqrt{2}}{\sigma\sqrt{2\pi}} h(\sqrt{2}\sigma x + \mu) e^{-x^2} dx \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} h(\sqrt{2}\sigma x + \mu) e^{-x^2} dx \\ &= \frac{1}{\sqrt{\pi}} I[h(\sqrt{2}\sigma x + \mu)] \simeq \frac{1}{\sqrt{\pi}} \sum_{i=1}^{n} w_i h(\sqrt{2}\sigma x_i + \mu) \end{aligned}$$

2. We apply Gram-Schmidt to the canonical basis $\{1, x, x^2\}$ with the associated inner product. We readily notice that $< x, 1 >_w = 0$ (using the formula, or that the integrand is odd), and so, $H_1(x) = x$. Finally, we get:

$$\begin{aligned} H_2(x) &= x^2 - \frac{< x^2, x >_w}{< x, x >_w} x - \frac{< x^2, 1 >_w}{< 1, 1 >_w} 1 \\ &= x^2 - \frac{< x^2, 1 >_w}{< 1, 1 >_w} 1 \\ &= x^2 - \frac{\sqrt{\pi}/2}{\sqrt{\pi}} 1 = x^2 - \frac{1}{2} \end{aligned}$$

3. The nodes for the 2 point rule are the roots of $H_2(x)$, that is, $\pm\frac{1}{\sqrt{2}} \simeq \pm 0.7071$. The weights $w_1, w_2$ must be such that this rule integrates constants and linear polynomials exactly. This gives us the system of $2 \times 2$ equations:

$$\begin{aligned} \sqrt{\pi} &= I[1] = w_1 + w_2 \\ 0 &= I[x] = -\frac{1}{\sqrt{2}} w_1 + \frac{1}{\sqrt{2}} w_2 \end{aligned}$$

The second equation simplifies to $w_2 - w_1 = 0$, meaning both weights are identical. We can easily obtain from this that $w_1 = w_2 = \sqrt{\pi}/2$.

**Problem 3: Numerical Linear Algebra** (25 points)

Let $A \in \mathbb{R}^{n\times n}$, non-singular (invertible) matrix. Let $u, v \in \mathbb{R}^n$; we define the rank 1 perturbation $\widehat{A} = A + uv^T$. The *Sherman-Morrison* formula gives us a formula for $\widehat{A}^{-1}$, if it exists:

$$\widehat{A}^{-1} = (A + uv^T)^{-1} = A^{-1} - \left(\frac{1}{s_A}\right) A^{-1}uv^T A^{-1} \quad \text{with} \quad s_A = 1 + v^T A^{-1} u$$

1. Let $(A + uv^T)x = b$, and define $y = v^T x$. Based on this, write down a a $n+1 \times n+1$ linear system of the form

$$M \begin{bmatrix} x \\ y \end{bmatrix} = c, \quad c \in \mathbb{R}^{n+1}$$

   satisfied by $x$ and $y$.

2. Use Gaussian elimination on the *last row* of $M$ (i.e. so that $M(n+1, 1:n)$ becomes zero). Using the resulting linear system, find a necessary and sufficient condition for $\widehat{A}$ to be invertible. You may assume that $A^{-1}$ is available.

3. Assume that $A$ is such that we can solve a system of the form $Ax = b$ in $O(n)$ floating point operations, or flops (e.g. $A$ is tridiagonal). What is then the computational cost to solve a system $\widehat{A}x = \widehat{b}$ using Sherman-Morrison? Indicate computational cost (flops) on each step.

**Solution:**

1. We can write both linear equations as the following system:

$$\begin{bmatrix} A & u \\ v^T & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

2. We apply Gauss Elimination to zero out all entries under the leading $n \times n$ block equal to $A$. To do so, we must multiply the first $n$ rows by something such that, adding them to the $n+1$ row, gives us 0. Since $A$ is invertible, we can multiply by $-v^T A^{-1}$ (this is equivalent to first multiplying by $A^{-1}$ to get an identity matrix on the leading block, then eliminating the last row). This gives us the equivalent system:

$$\begin{bmatrix} A & u \\ 0 & -1 - v^T A^{-1} u \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ -v^T A^{-1} b \end{bmatrix}$$

   Alternatively, if we had done the whole elimination, we would have:

$$\begin{bmatrix} I & A^{-1} u \\ 0 & -1 - v^T A^{-1} u \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A^{-1} b \\ -v^T A^{-1} b \end{bmatrix}$$

   Looking at either of these systems, we notice that they are block upper triangular (and the second one is also upper triangular), where the leading block is invertible. These systems of equations are uniquely solvable if and only if the last pivot is non-zero, that is, we require $1 + v^T A^{-1} u \neq 0$.

   Looking at the Sherman-Morrison formula, this makes sense. If $s_A = 0$, the formula would be asking us to divide by zero. Note that, while *while this was not required in this exercise*, the formula can be quickly derived from solving the equivalent system above for $x$.

3. We notice that the formula may be factored as: $\widehat{A} = (I - \frac{1}{s_A}(A^{-1}u)v^T)A^{-1}$. Given our input right-hand-side $\widehat{b}$, we perform the following steps:

   (a) Solve $Ax_1 = b$, which can be done in $O(n)$ flops.

   (b) Compute $c = v^T x_1$, which is $2n = O(n)$ flops (inner product)

   (c) Solve $Aw = u$, which means $w = A^{-1}u$, which can be done in $O(n)$ flops.

   (d) Finally, $x = x_1 - (c/s_A)w$, which is $2n = O(n)$ flops (product by scalar and sum).

   We note that this allows us to write down a solver for $\widehat{A}$ which is also $O(n)$.

**Problem 4: Interpolation/Approximation** (25 points)

Obtain the first degree polynomial achieving the minimax approximation for $f(x) = 1/(1+x)$ on $[0, 1]$. Formulate the theorem describing properties of the minimax error.

**Solution:**

Since $f(x)$ is monotone, the answer is obtained by directly matching the requirements of Chebyshev Equioscillation theorem. We consider

$$f(x) = 1/(x+1) - (ax+b),$$

compute its values at $x = 0$ and $x = 1$,

$$f(0) = 1 - b$$

and

$$f(1) = \frac{1}{2} - a - b.$$

Then, computing the derivative of $f$ and setting it to zero, we find a point in $[0, 1]$ where $f$ achieves its minimum,

$$x = \sqrt{2} - 1.$$

The Chebyshev Equioscillation theorem requires that $f(0) = f(1) = -f\left(\sqrt{2} - 1\right)$. Solving these equations, we obtain

$$ax + b = -\frac{x}{2} + \frac{1}{\sqrt{2}} + \frac{1}{4}.$$

That is, $a = -1/2$, $b = \frac{1}{\sqrt{2}} + \frac{1}{4}$, and $f(0) = f(1) = -f(x) = \frac{3}{4} - \frac{1}{\sqrt{2}}$.

**Problem 5: Numerical ODE** (25 points)

Consider the two step method (Adams-Bashforth)

$$y_{n+2} = y_{n+1} + h\left[\frac{3}{2}f(t_{n+1}, y_{n+1}) - \frac{1}{2}f(t_n, y_n)\right].$$

- Show that it is convergent and find its order. State the relevant theorems.
- Find the interval on the real axis that is a part of the region of absolute stability.

**Solution:**

A multistep method is convergent if it is consistent (i.e. its order $p \geq 1$) and stable. Computing the order of the two step Adams-Bashforth method, we have

$$
\begin{aligned}
y(t_{n+2}) - y(t_{n+1}) - h\left[\frac{3}{2}f(t_{n+1}, y_{n+1}) - \frac{1}{2}f(t_n, y_n)\right] &= y(t_n) + 2hy'(t_n) + 2h^2y''(t_n) + \mathcal{O}\left(h^3\right) \\
&- \left(y(t_n) + hy'(t_n) + \frac{1}{2}h^2y''(t_n) + \mathcal{O}\left(h^3\right)\right) \\
&- h\left[\frac{3}{2}y'(t_{n+1}) - \frac{1}{2}y'(t_n)\right] \\
&= hy'(t_n) + \frac{3}{2}h^2y''(t_n) + \mathcal{O}\left(h^3\right) \\
&- h\left[\frac{3}{2}\left(y'(t_n) + hy''(t_n) + \mathcal{O}\left(h^2\right)\right) - \frac{1}{2}y'(t_n)\right] \\
&= \mathcal{O}\left(h^3\right)
\end{aligned}
$$

implying that it is a second order method. Applying the method to the test problem

$$
\begin{aligned}
y' &= \lambda y \\
y(0) &= y_0,
\end{aligned}
$$

we obtain the characteristic equation

$$
r^2 - \left(1 + \frac{3}{2}h\lambda\right) r + \frac{1}{2}h\lambda = 0
$$

or

$$
r^2 - \left(1 + \frac{3}{2}z\right) r + \frac{1}{2}z = 0.
$$

The root condition is satisfied ($\lambda = 0$), i.e. (i) all zeros are inside the unit disk and (ii) all zeros of unit modulus (i.e., on the boundary of the unit disk) are simple. Therefore the method is stable. Finally, to sketch the region of absolute stability, we look for $z$ such that the roots of the characteristic polynomial are of absolute value less or equal to 1. We have

$$
r_{\pm} = \frac{1}{4}\left(2 + 3z \pm \sqrt{4 + 4z + 9z^2}\right).
$$

On the real axis $|r_{\pm}| \leq 1$ for $-1 \leq z \leq 0$.

**Problem 6: Numerical PDE** (25 points)

Consider the heat equation

$$\frac{\partial \phi}{\partial t} = \partial_x^2 \phi,$$

with initial condition

$$\phi|_{t=0} = \phi_0,$$

and periodic boundary conditions on the interval $[0, 1]$. Fully describe the Crank-Nicolson scheme for this problem. Show that the scheme is unconditionally stable.

**Solution:**

Discretizing the heat equation in space, we obtain a system of ODEs,

$$\frac{d\phi_j}{dt} = \frac{\phi_{j-1} - 2\phi_j + \phi_{j+1}}{h^2},$$

where $\phi_j = \phi\,(hj)$, $h = 1/N$, and $j = 1, \ldots, N$. The periodicity of solution is imposed by setting $\phi_0 = \phi_{N+1}$. Writing this system

$$\frac{d\phi}{dt} = \frac{1}{h^2} A\phi,$$

where $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_N)^T$ and $A$ is a circulant tridiagonal matrix, we apply the trapezoidal rule in time,

$$\boldsymbol{\phi}^{k+1} - \boldsymbol{\phi}^k = \frac{1}{2}\frac{h_t}{h^2} A \left(\boldsymbol{\phi}^{k+1} + \boldsymbol{\phi}^k\right)$$

where $\boldsymbol{\phi}^0 = \boldsymbol{\phi}_0$. We have

$$\boldsymbol{\phi}^{k+1} = \left(I - \frac{1}{2}\frac{h_t}{h^2} A\right)^{-1} \left(I + \frac{1}{2}\frac{h_t}{h^2} A\right) \boldsymbol{\phi}^k.$$

A circulant matrix is diagonalized by the Discrete Fourier transform and, in this case, has non-positive eigenvalues. Since $A$ is negative semi-definite, it is easy to verify that the eigenvalues of

$$\left(I - \frac{1}{2}\frac{h_t}{h^2} A\right)^{-1} \left(I + \frac{1}{2}\frac{h_t}{h^2} A\right)$$

are all less or equal to 1, confirming unconditional stability.