This collection of 15 problems is intended to help you review and prepare for the preliminary exam on *Numerical Analysis*. The exam covers material in MAT 128A (Numerical Analysis), MAT 128B (Numerical Analysis in Solutions of Equations), and MAT 128C (Numerical Analysis in Differential Equations). Some problems are harder and longer than what you will find on the exam. Try to write the solutions as you would in the exam: Write out all details when solving each problem. Be organized and use the notation appropriately. Initially try to solve the problems without any assistance.

The prelim exam will be 3 hours long and have 6 problems. In what follows we present in consecutive order six problems from MAT 128A, six problems from MAT 128B, and 6 problems from MAT 128C.

PART I: Problems from Numerical Analysis

1. Let be a floating point system with base/radix β. Let $\widehat{x}$ and $\widehat{y}$ be floating point numbers such that
$$\frac{1}{\beta} \le \widehat{y} < 1 \le \widehat{x} < \beta.$$
Show that $\widehat{x} - \widehat{y}$ is a floating point number.

2. Consider the approximation

$$f'(x) \approx \frac{-3f(x) + 4f(x+h) - f(x+2h)}{2h}.$$

(a) Using Taylor's Theorem, derive the error term for the above approximation.

(b) What is the round-off error in the above finite difference? (You can ignore the round-off error in computing h, x, $x+h$, and $x+2h$.

(c) Using the above finite difference approximation for f', find a finite difference approximation for f''.

3. Let $\mathcal{P}_n$ denote the space of all polynomials of degree at most n. Further, let $x_0, x_1, \ldots, x_n \in [0,1]$ be distinct. Define

$$\ell_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}, \quad i = 0, 1, \ldots, n,$$

and $P_n : C([0,1]) \to \mathcal{P}_n$ by

$$[P_n f](x) = \sum_{i=0}^{n} f(x_i)\ell_i(x).$$

(a) Show that $p(x) = [P_n f](x)$ is the unique element of $\mathcal{P}_n$ that satisfies

$$p(x_i) = f(x_i), \quad i = 0, 1, \ldots, n.$$

(b) Show that P_n is a linear operator on $C([0,1])$.

(c) Define

$$\|P_n\|_\infty = \sup_{f \in C([0,1])} \frac{\|P_n f\|_\infty}{\|f\|_\infty}.$$

Show that

$$\|P_n\|_\infty \leq \Lambda_n,$$

where

$$\Lambda_n = \max_{x \in [0,1]} \sum_{i=0}^{n} |\ell_i(x)|.$$

(d) Show that for any $f \in C([0,1])$

$$\|f - P_n f\|_\infty \leq (1 + \Lambda_n) \min_{p_* \in \mathcal{P}_n} \|f - p_*\|_\infty.$$

4. Let $w \in C([-1,1])$ satisfy $w(x) > 0$ for all $x \in [-1,1]$. Suppose that $P_0, P_1, P_2, \ldots$ is a family of polynomials such that P_n is of degree exactly n for all $n \geq 0$ and

$$\int_{-1}^{1} P_i(x) P_j(x) w(x)\, \mathrm{d}x = 0, \quad i \neq j,$$

for all $i, j \in \{0, 1, \ldots\}$.

(a) Show that there exist sequences of numbers $\{\alpha_n\}_{n=1}^{\infty}$, $\{\beta_n\}_{n=1}^{\infty}$, $\{\gamma_n\}_{n=1}^{\infty}$ such that

$$P_{n+1}(x) = \alpha_n x P_n(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x), \quad n \geq 1.$$

(b) Let $N \geq 2$. Let $\{x_i\}_{i=1}^{N}$ denote the roots of P_N. Suppose $\{w_i\}_{i=1}^{N}$ are chosen such that

$$\sum_{i=1}^{N} w_i p(x_i) = \int_{-1}^{1} p(x) w(x)\, \mathrm{d}x,$$

where p is any polynomial of degree $N-1$ or lower.

Show that

$$\sum_{i=1}^{N} w_i q(x_i) = \int_{-1}^{1} q(x) w(x)\, \mathrm{d}x,$$

where q is any polynomial of degree $2N-1$ or lower.

5. Simpson's rule for computing the integral

$$I = \int_a^b f(x)\,\mathrm{d}x$$

is to approximate I as

$$I \approx I_h = \frac{h}{6}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right)$$

where $h = b - a$.

(a) For smooth functions, what is the approximation error of Simpson's rule with respect to the interval width h?

(b) What is the form of one step of Richardson extrapolation of I_h? What approximation error to I is obtained in the result?

PART II: Problems from Numerical Analysis in Solutions of Equations

1. Let $S \in \mathbb{C}^{n\times n}$ be skew-Hermitian, i.e., $S^* = -S$.

   (a) Show that if λ is an eigenvalue of S then $\operatorname{Re}\lambda = 0$.

   (b) Show that $I - S$ is nonsingular.

   (c) Show that $Q = (I - S)^{-1}(I + S)$ is unitary.

2. Sometimes in numerical computations, the naïve heuristic for the accuracy of the computed solution based on condition numbers is too pessimistic. In this problem we will examine perhaps the most famous example of this phenomenon.

   (a) Let $x \in \mathbb{R}$. Let $\widehat{x} = x(1+\delta_1)$ where $|\delta_1| \leq u$ and $0 < u \ll 1$. Find δ_2 such that

   $$\frac{1}{\widehat{x}} = \frac{1}{x}(1+\delta_2).$$

   Furthermore, show that $|\delta_2| \leq 2u$ as long as $u < \frac{1}{2}$.

   (b) Let $D \in \mathbb{R}^{n\times n}$ be a diagonal matrix with all diagonal elements positive. Let $\kappa_2(D)$ denote the matrix condition number of D in the 2-norm. Compute $\kappa_2(D)$.

   (c) Now, suppose you store D on a computer. Let $\widehat{D} \in \mathbb{R}^{n\times n}$ be the rounded version of D with entries given by

   $$\widehat{D}_{ij} = D_{ij}(1+\eta_{ij}), \quad |\eta_{ij}| \leq u, \quad i,j = 1,2,\ldots,n.$$

   Let $x, \widehat{x}, b \in \mathbb{R}^n$ satisfy $Dx = b$ and $\widehat{D}\widehat{x} = b$. Based on your answer to (b), what does our condition number heuristic say about how large we should expect

   $$\frac{\|x-\widehat{x}\|_2}{\|x\|_2}$$

   to be?

   (d) Now, suppose you compute $\widehat{x}$ using the formula

   $$\widehat{x}_i = \frac{b_i}{\widehat{D}_{ii}}, \quad i = 1,2,\ldots,n.$$

   Show that if $u < \frac{1}{2}$, then

   $$\frac{\|x-\widehat{x}\|_2}{\|x\|_2} \leq 2u.$$

   (e) To make this tangible let $n = 100$ and suppose $D_{ii} = 10^{-(i-1)}$ for $i = 1,2,\ldots,100$. For this specific example work out the heuristic in (c) and the estimate in (d) when IEEE double precision floating point arithmetic is used.

   (f) How can you reconcile the discrepancy between (c) and (d)?

3. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix with eigenpairs (λ_i, q_i) for $i = 1, 2, \ldots, n$. Further assume that $q_1, q_2, \ldots, q_n$ is an orthonormal basis for $\mathbb{R}^n$ and $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$. Define

$$P = q_1 q_1^T.$$

   (a) Suppose $x \in \mathbb{R}^n$ satisfies $Px \neq 0$. Define

$$y^{(k)} = \frac{A^k x}{\|A^k x\|_2}, \quad k = 0, 1, 2, \ldots.$$

   Show that

$$\|y^{(k)} - Py^{(k)}\|_2 = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right).$$

   (b) Show that

$$|\lambda_1 - (y^{(k)})^T A y^{(k)}| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right).$$

   (c) Is it guaranteed that

$$\|y^{(k+1)} - y^{(k)}\|_2 \to 0$$

   as $k \to \infty$? Explain your answer.

4. Let $A \in \mathbb{R}^{m\times n}$ where $m > n$, and assume that $\operatorname{rank}(A) = n$. Let $b \in \mathbb{R}^m$. Consider the least-squares problem

$$\widehat{x} = \arg\min_{x\in\mathbb{R}^n} \|Ax - b\|_2. \tag{1}$$

   (a) Let A be as above, and let $O \in \mathbb{R}^{m\times m}$ be an orthogonal matrix. Show that $\widehat{x}$ satisfies (1) if and only if $\widehat{x}$ satifies

$$\widehat{x} = \arg\min_{x\in\mathbb{R}^n} \|OAx - Ob\|_2.$$

   (b) Let $S \in \mathbb{R}^{m\times n}$ be diagonal (i.e., $S_{ij} \neq 0$ only if $i = j$), and suppose $S_{ii} \neq 0$ for all $i = 1, 2, \ldots, n$. Find the solution to the least-squares problem

$$\widehat{y} = \arg\min_{y\in\mathbb{R}^n} \|Sy - b\|_2.$$

   (c) Let A be as above and suppose $P \in \mathbb{R}^{n\times n}$ is an invertible matrix. Suppose $\widehat{z}$ satisfies

$$\widehat{z} = \arg\min_{z\in\mathbb{R}^n} \|APz - b\|_2.$$

   Find the solution to (1) in terms of $\widehat{z}$.

   (d) Suppose we have a factorization $A = U\Sigma V^T$ where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal, and $\Sigma \in \mathbb{R}^{m\times n}$ with entries satisfying $\Sigma_{ij} = 0$ if $i \neq j$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Using (a)–(c), show how this factorization can be used to solve the least-squares problem.

5. The first stage in most symmetric eigenvalue computations is to first tridiagonalize the matrix.

   (a) Consider the symmetrix 3×3 matrix given by

   $$A = \begin{bmatrix} a_1 & b & c \\ b & a_2 & d \\ c & d & a_3 \end{bmatrix}.$$

   Find an explicit formula for an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ such that

   $$QAQ^T = \begin{bmatrix} \widetilde{a}_1 & \widetilde{b} & 0 \\ \widetilde{b} & \widetilde{a}_2 & \widetilde{d} \\ 0 & \widetilde{d} & \widetilde{a}_3 \end{bmatrix},$$

   where the tilde variables denote arbitrary real numbers.

   (b) We can use induction to show that this approach can be extended to any arbitrary symmetric $n \times n$ matrix. For an arbitrary $n \times n$ matrix, we assume that we can tridiagonalize the leading $n-1 \times n-1$ submatrix. Thus, we may assume that $A \in \mathbb{R}^{n \times n}$ is of the form

   $$A = \begin{bmatrix} a_1 & b_2 & & & & \\ b_2 & a_2 & b_3 & & & \\ & b_3 & a_3 & \ddots & & \\ & & \ddots & \ddots & b_{n-1} & d \\ & & & b_{n-1} & a_{n-1} & c \\ & & & d & c & a_n \end{bmatrix},$$

   where all omitted entries are zero. Find an explicit formula for an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ such that QAQ^T is tridiagonal.

PART III: Problems from Numerical Analysis in Differential Equations

1. Consider the scalar initial value problem (IVP)

$$\begin{aligned} y' &= f(t, y(t)), \quad t \in [0, 1] \\ y(0) &= y_0. \end{aligned}$$

Let N be a positive integer, let $h = 1/N$, and $t_j = jh$ for $j = 0, 1, 2, \ldots, N$.

Integration of the ODE on each interval yields

$$y(t_{j+1}) = y(t_j) + \int_{t_j}^{t_{j+1}} f(t, y(t))\,\mathrm{d}t.$$

The s-step Adams–Bashforth method is a linear multistep method that discretizes the above integral by replacing f with its polynomial interpolant at the points $t_j, t_{j-1}, \ldots, t_{j-s+1}$.

(a) Derive the 2-step Adams–Bashsforth method.

(b) Find the rate at which the local truncation error goes to 0 for the 2-step Adams–Bashforth method as $h \to 0$ when f is smooth.

(c) Show that the 2-step Adams–Bashforth method is zero stable.

(d) At each time step, the 2-step Adams–Bashforth method requires information about previous two time steps. However, we only have a single initial condition. Explain how you would overcome this obstacle in practice.

2. Consider the scalar initial value problem (IVP)

$$\begin{aligned} y' &= f(y), \quad t \in [0,1] \\ y(0) &= y_0. \end{aligned}$$

Let N be a positive integer, let $h = 1/N$, and $t_j = jh$ for $j = 0, 1, 2, \ldots, N$. Consider the linear multistep method:

$$y_{n+2} - 3y_{n+1} + 2y_n = -hf(y_n), \quad n = 0, 1, \ldots, N-2.$$

(a) Find the rate at which the local truncation error goes to zero for this method as $h \to 0$ when f is smooth.

(b) Determine whether this linear multistep method is zero stable.

(c) Do you expect the method to converge as $h \to 0$? Explain your answer.

3. Consider the scalar initial value problem (IVP)

$$y' = f(t, y(t))$$
$$y(0) = y_0.$$

Let $h > 0$ and $t_n = nh$ for $j = 0, 1, 2, \ldots$.

Consider the linear multistep method

$$y_{n+1} + by_{n-1} + ay_{n-2} = hf(t_n, y_n), \quad n = 0, 1, 2, \ldots$$

where a and b are constants.

(a) For a certain (unique) choice of a and b, this method is consistent. Find these values of a and b and verify that the order of accuracy is 1.

(b) Although the method is consistent for the choice of a and b from (a), the method is not zero stable. Show this and describe quantitatively what the unstable solutions will look like for small h.

4. Consider the scalar initial value problem (IVP)

$$
\begin{aligned}
y' &= f(t, y(t)), \quad t \in [0,1] \\
y(0) &= y_0.
\end{aligned}
$$

Let N be a positive integer, let $h = 1/N$, and $t_j = jh$ for $j = 0, 1, 2, \ldots, N$. Given $\alpha > 0$, consider the linear two-step method

$$y_{n+2} - \alpha y_n = \frac{h}{3}\left[f(t_{n+2}, y_{n+2}) + 4f(t_{n+1}, y_{n+1}) + f(t_n, y_n)\right].$$

(a) Determine the set of all α such that the method is zero stable.

(b) Find α such that the order of accuracy is as high as possible.

(c) Is the method convergent for the value of α in (b)?

5. Consider the scalar initial value problem (IVP)

$$\begin{aligned} y' &= f(t, y(t)), \\ y(0) &= y_0. \end{aligned}$$

Let $h > 0$ and $t_n = nh$ for $j = 0, 1, 2, \ldots$

Consider the one-step method

$$y_{n+1} = y_n + \alpha h f(t_n, y_n) + \beta h f(t_n + \gamma h, y_n + \gamma h f(t_n, y_n)),$$

where α, β, and γ are real parameters.

(a) Show that the method is consistent if and only if $\alpha + \beta = 1$.

(b) show that the order of the method cannot exceed 2.

(c) In the above IVP, suppose that $f(t, y) = -\lambda y$ and $y_0 = 1$, where $\lambda > 0$. Show that the sequence $y_0, y_1, y_2, \ldots$ is bounded if and only if $h \leq \frac{2}{\lambda}$.