AIMS Exercise Set # 1

Peter J. Olver

1. Determine the form of the single precision floating point arithmetic used in the computers at AIMS. What is the largest number that can be accurately represented? What is the smallest positive number n_1? The second smallest positive number n_2? Which is larger: the gap between n_1 and 0 or the gap betweeen n_1 and n_2? Discuss.

2. Determine the value of each of the following quantities using 4 digit rounding and four digit chopping arithmetic. Find the abolute and relative errors of your approximation. (a) $\pi + e - \cos 22^\circ$, (b) $\dfrac{e^\pi - \pi^e}{\log \frac{10}{11}}$.

3. (a) To how many significant decimal digits do the numbers $\sqrt{10002}$ and $\sqrt{10001}$ agree? (b) Subtract the two numbers. How many significant decimal digits are lost in the computation? (c) How might you rearrange the computation to obtain a more accurate answer.

4. (a) Verify that $f(x) = 1 - \sin x$ and $g(x) = \dfrac{\cos^2 x}{1 + \sin x}$ are identical functions.
(b) Which function should be used for computations when x is near $\frac{1}{2}\pi$? Why?
(c) Which function should be used for computations when x is near $\frac{3}{2}\pi$? Why?

5. *Horner's Method*

(a) Suppose x is a real number and n a positive integer. How many multiplications are need to efficiently compute x^n? *Hint*: The answer is not $n - 1$.

(b) Verify the polynomial identity

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = a_0 + x(a_1 + x(a_2 + x(\cdots + x\, a_n) \cdots)).$$

Explain why the right hand side is to be preferred when computing the values of the polynomial $p(x)$.

6. Let

$$f(x) = e^x - \cos x - x.$$

(*a*) Using calculus, what should the graph of $f(x)$ look like for x near 0?

(*b*) Using both single and double precision arithmetic, graph $f(x)$ for $|\,x\,| \le 5 \times 10^{-8}$ and discuss what you observe.

(*c*) How might you obtain a more realistic graph?

7. Consider the linear system of equations

$$1.1\,x + 2.1\,y = a, \qquad 2\,x + 3.8\,y = b.$$

Solve the system for the following right hand sides: (*i*) $a = 3.2,\ \ b = 5.8$; (*ii*) $a = 3.21$, $b = 5.79$; (*iii*) $a = 3.1,\ \ b = 5.7$. Discuss the conditioning of this system of equations.

AIMS Exercise Set # 2

Peter J. Olver

1. Explain why the equation $e^{-x} = x$ has a solution on the interval $[0,1]$. Use bisection to find the root to 4 decimal places. Can you prove that there are no other roots?

2. Find $\sqrt[6]{3}$ to 5 decimal places by setting up an appropriate equation and solving using bisection.

3. Find all real roots of the polynomial $x^5 - 3x^2 + 1$ to 4 decimal places using bisection.

4. Let $g(u)$ have a fixed point $u^\star$ in the interval $[0,1]$, with $g'(u^\star) \neq 1$. Define

$$G(u) = \frac{u\,g'(u) - g(u)}{g'(u) - 1}.$$

(a) Prove that, for an initial guess $u^{(0)}$ near $u^\star$, the fixed point iteration scheme $u^{(n+1)} = G(u^{(n)})$ converges to the fixed point u. (b) What is the order of convergence of this method? (c) Test this method on the non-convergent cubic scheme in Example 2.16.

5. Let $g(u) = 1 + u - \frac{1}{8}u^3$. ($a$) Find all fixed points of $g(u)$. (b) Does fixed point iteration converge? If so, to which fixed point(s)? What is the rate of convergence? (c) Predict how may iterates will be needed to get the fixed point accurate to 4 decimal places starting with the initial guess $u^{(0)} = 1$. (d) Check your prediction by performing the iteration.

6. Solve Exercise 1–3 by Newton's Method.

7. (a) Let $u^\star$ be a simple root of $f(u) = 0$. Discuss the rate of convergence of the iterative method (sometimes known as *Olver's Method*, in honor of the author's father) based on $g(u) = u + \dfrac{f(u)^2 f''(u) - 2\,f(u)\,f'(u)^2}{2\,f'(u)^3}$ to $u^\star$. (b) Try this method on the equation in Exercise 3, and compare the speed of convergence with that of Newton's Method.

AIMS Exercise Set # 3

Peter J. Olver

1. Which of the following matrices are regular? If reguolar, write down its LU factorization. (a) $\begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$, (b) $\begin{pmatrix} 0 & -1 \\ 3 & -2 \end{pmatrix}$, (c) $\begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}$.

2. In each of the following problems, find the $A = LU$ factorization of the coefficient matrix, and then use Forward and Back Substitution to solve the corresponding linear systems $A\mathbf{x} = \mathbf{b}$ for each of the indicated right hand side:

$$(a)\ A = \begin{pmatrix} -1 & 3 \\ 3 & 2 \end{pmatrix},\ \mathbf{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}; \qquad (b)\ A = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 2 & 3 & -1 \\ -1 & 3 & 2 & 2 \\ 0 & -1 & 2 & 1 \end{pmatrix},\ \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

3. Find the LDL^T factorization of the matrix $\begin{pmatrix} 1 & -1 & -1 \\ -1 & 3 & 2 \\ -1 & 2 & 0 \end{pmatrix}$.

4. (a) Find the LU factorization of the $n \times n$ tridiagonal matrix A_n with all 2's along the diagonal and all -1's along the sub- and super-diagonals for $n = 3, 4$ and 5. (b) Use your factorizations to solve the system $A_n \mathbf{x} = \mathbf{b}$, where $\mathbf{b} = (1, 1, 1, \ldots, 1)^T$. (c) Can you write down the LU factorization of A_n for general n? Do the entries in the factors approach a limit as n gets larger and larger?

5. *True or false*: (a) The product of two tridiagonal matrices is tridiagonal. (b) The inverse of a tridiagonal matrix is tridiagonal.

6. (a) Find the exact solution to the linear system $x - 5y - z = 1$, $\frac{1}{6}x - \frac{5}{6}y + z = 0$, $2x - y = 3$. (b) Solve the system using Gaussian Elimination with 4 digit rounding. (c) Solve the system using Partial Pivoting and 4 digit rounding. Compare your answers.

7. Implement the computer experiment with Hilbert matrices outlined in the last paragraph of the section.

AIMS Exercise Set # 4

Peter J. Olver

1. Find the explicit formula for the solution to the following linear iterative system:
$$u^{(k+1)} = u^{(k)} - 2\,v^{(k)}, \qquad v^{(k+1)} = -\,2\,u^{(k)} + v^{(k)}, \qquad u^{(0)} = 1, \qquad v^{(0)} = 0.$$

2. Determine whether or not the following matrices are convergent:
$$(a)\quad \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}, \qquad\qquad (b)\quad \frac{1}{5}\begin{pmatrix} 5 & -3 & -2 \\ 1 & -2 & 1 \\ 1 & -5 & 4 \end{pmatrix}.$$

3. (a) Find the spectral radius of the matrix $T = \begin{pmatrix} 1 & 1 \\ -1 & -\frac{7}{6} \end{pmatrix}$. (b) Predict the long term behavior of the iterative system $\mathbf{u}^{(k+1)} = T\,\mathbf{u}^{(k)} + \mathbf{b}$, where $\mathbf{b} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, in as much detail as you can.

4. Consider the linear system $A\,\mathbf{x} = \mathbf{b}$, where $A = \begin{pmatrix} 4 & 1 & -2 \\ -1 & 4 & -1 \\ 1 & -1 & 4 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix}$. (a) First, solve the equation directly by Gaussian Elimination. (b) Using the initial approximation $\mathbf{x}^{(0)} = \mathbf{0}$, carry out three iterations of the Jacobi algorithm to compute $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}$ and $\mathbf{x}^{(3)}$. How close are you to the exact solution? (c) Write the Jacobi iteration in the form $\mathbf{x}^{(k+1)} = T\,\mathbf{x}^{(k)} + \mathbf{c}$. Find the 3×3 matrix T and the vector $\mathbf{c}$ explicitly. (d) Using the initial approximation $\mathbf{x}^{(0)} = \mathbf{0}$, carry out three iterations of the Gauss–Seidel algorithm. Which is a better approximation to the solution — Jacobi or Gauss–Seidel? (e) Write the Gauss–Seidel iteration in the form $\mathbf{x}^{(k+1)} = \widetilde{T}\,\mathbf{x}^{(k)} + \widetilde{\mathbf{c}}$. Find the 3×3 matrix $\widetilde{T}$ and the vector $\widetilde{\mathbf{c}}$ explicitly. (f) Determine the spectral radius of the Jacobi matrix T, and use this to prove that the Jacobi method iteration will converge to the solution of $A\,\mathbf{x} = \mathbf{b}$ for any choice of the initial approximation $\mathbf{x}^{(0)}$. (g) Determine the spectral radius of the Gauss–Seidel matrix $\widetilde{T}$. Which method converges faster? (h) For the faster method, how many iterations would you expect to need to obtain 5 decimal place accuracy? (i) Test your prediction by computing the solution to the desired accuracy.

5. The matrix $A = \begin{pmatrix} 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 4 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 4 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 4 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 \end{pmatrix}$ arises in the finite difference (and finite element) discretization of the Poisson equation on a nine point square grid. (*a*) Is A diagonally dominant? (*b*) Find the spectral radius of the Jacobi and Gauss–Seidel iteration matrices. (*c*) Use formula (7.69) to fix the optimal value of the SOR parameter. Verify that the spectral radius of the resulting iteration matrix agrees with the second formula in (7.69). (*d*) For each iterative scheme, predict how many iterations are needed to solve the linear system $A\mathbf{x} = \mathbf{e}_1$ to 3 decimal places, and then verify your predictions by direct computation.

6. The generalization of Exercise 5 to the Poisson equation on an $n \times n$ grid results in an $n^2 \times n^2$ matrix in block tridiagonal form $A = \begin{pmatrix} K & -\mathrm{I} & & & \\ -\mathrm{I} & K & -\mathrm{I} & & \\ & -\mathrm{I} & K & -\mathrm{I} & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$, in which K is the tridiagonal $n \times n$ matrix with 4's on the main diagonal and -1's on the sub- and super-diagonal, while I denotes an $n \times n$ identity matrix. Use the known value of the Jacobi spectral radius $\rho_J = \cos \dfrac{\pi}{n+1}$, [**47**], to design an SOR method to solve the linear system $A\mathbf{u} = \mathbf{f}$. Run the Jacobi, Gauss–Seidel, and SOR methods for the cases $n = 5$ and $\mathbf{f} = \mathbf{e}_{13}$ and $n = 25$ and $\mathbf{f} = \mathbf{e}_{313}$ corresponding to a unit force at the center of the grid.

AIMS Exercise Set # 5

Peter J. Olver

1. Use the power method to find the dominant eigenvalue and associated eigenvector of the following matrices: (a) $\begin{pmatrix} -2 & 0 & 1 \\ -3 & -2 & 0 \\ -2 & 5 & 4 \end{pmatrix}$, (b) $\begin{pmatrix} 4 & 1 & 0 & 1 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 1 & 0 & 1 & 4 \end{pmatrix}$.

2. Use Newton's Method to find all points of intersection of the following pairs of plane curves: $x^3 + y^3 = 3, \ \ x^2 - y^2 = 2$.

3. The system $x^2 + x\,z = 2$, $x\,y - z^2 = -1$, $y^2 + z^2 = 1$, has a solution $x^\star = 1$, $y^\star = 0$, $z^\star = 1$. Consider a fixed point iteration scheme with

$$\mathbf{g}(x,y,z) = \big(\, x + \alpha(x^2 + x\,z - 2), y + \alpha(x\,y - z^2 + 1), z + \alpha(y^2 + z^2 - 1)\,\big)^T,$$

where α is a constant. (a) For which values of α does the iterative scheme converge to the solution when the initial guess is nearby? (b) What is the best value of α as far as the rate of convergence goes? (c) For the value of α from part (a) (or another value of your own choosing) about how many iterations are required to approximate the solution to 5 decimal places when the initial guess is $x^{(0)} = \frac{5}{6}$, $y^{(0)} = -\frac{1}{3}$, $z^{(0)} = \frac{9}{8}$? Test your estimate by running the iteration. (d) Write down the Newton iteration scheme for this system. (e) Answer part (c) for the Newton scheme.

AIMS Exercise Set # 6

Peter J. Olver

1. Prove that the *Midpoint Method* (10.58) is a second order method.

2. Consider the initial value problem

$$\frac{du}{dt} = u(1-u), \qquad u(0) = .1,$$

for the *logistic differential equation*.

(a) Find an explicit formula for the solution. Describe in words the behavior of the solution for $t > 0$.

(b) Use the Euler Method with step sizes $h = .2$ and $.1$ to numerically approximate the solution on the interval $[0, 10]$. Does your numerical solution behave as predicted from part (a)? What is the maximal error on this interval? Can you predict the error when $h = .05$? Test your prediction by running the method and computing the error. Estimate the step size needed to compute the solution accurately to 10 decimal places (assuming no round off error)? How many steps are required? (Just predict — no need to test it.)

(c) Answer part (b) for the Improved Euler Method.

(d) Answer part (b) for the fourth order Runge–Kutta Method.

(e) Discuss the behavior of the solution, both analytical and numerical, for the alternative initial condition $u(0) = -.1$.

3. The nonlinear second order ordinary differential equation

$$\frac{d^2\theta}{dt^2} + \sin\theta = 0$$

describes the motion of a pendulum under gravity without friction, with $\theta(t)$ representing the angle from the vertical: $\theta = 0$ represents the stable equilibrium where the pendulum is hanging straight down, while $\theta = \pi$ corresponds to the unstable equilibrium where the pendulum is standing straight up.

(a) Write out an equivalent first order system of ordinary differential equations in $u(t) = \theta(t)$, $v(t) = \dfrac{d\theta}{dt}$.

(b) Prove that the total energy of the pendulum

$$E(u,v) = \tfrac{1}{2} v^2 + (1 - \cos u) = \frac{1}{2}\left(\frac{d\theta}{dt}\right)^2 + (1 - \cos\theta)$$

is constant on solutions. *Hint*: Show that $dE/dt = 0$. Explain why each solution moves along a single level curve $E(u,v) = c$ of the energy.

(c) Use either your physical intuition and/or part (b) to describe the motion of the pendulum for the following initial conditions:

(i) $u(0) = 0$, $v(0) = 1$; (ii) $u(0) = 0$, $v(0) = 1.95$; (iii) $u(0) = 0$, $v(0) = 2$.

(d) Use the Euler Method to integrate the initial value problems for $0 \le t \le 50$ with step sizes $h = .1$ and $.01$. How accurately do your numerical solutions preserve the energy? How accurately do your numerical solutions follow the behavior you predicted in part (b)?

(e) Answer part (d) using the fourth order Runge–Kutta Method.

AIMS Exercise Set # 7

Peter J. Olver

1. In this exercise, you are asked to find "one-sided" finite difference formulas for derivatives. These are useful for approximating derivatives of functions at or near the boundary of their domain. (a) Construct a second order, one-sided finite difference formula that approximates the derivative $f'(x)$ using the values of $f(x)$ at the points $x, x+h$ and $x+2h$. (b) Find a finite difference formula for $f''(x)$ that involves the same values of f. What is the order of your formula? (c) Test your formulas by computing approximations to the first and second derivatives of $f(x) = e^{x^2}$ at $x = 1$ using step sizes $h = .1, .01$ and $.001$. What is the error in your numerical approximations? Are the errors compatible with the theoretical orders of the finite difference formulae? Discuss why or why not. (d) Answer part (c) at the point $x = 0$.

2. (a) Design an explicit numerical method for solving the initial-boundary value problem

$$u_t = \gamma\, u_{xx} + s(x), \qquad u(t,0) = u(t,1) = 0, \qquad u(0,x) = f(x), \qquad 0 \le x \le 1,$$

for the heat equation with a *source term* $s(x)$. (b) Test your scheme on the particular problem for

$$\gamma = \frac{1}{6}, \qquad s(x) = x(1-x)(10-22x), \qquad f(x) = \begin{cases} 2\,\left|\,x - \frac{1}{6}\,\right| - \frac{1}{3}, & 0 \le x \le \frac{1}{3}, \\ 0, & \frac{1}{3} \le x \le \frac{2}{3}, \\ \frac{1}{2} - 3\,\left|\,x - \frac{5}{6}\,\right|, & \frac{2}{3} \le x \le 1, \end{cases}$$

using space step sizes $h = .1$ and $.05$, and a suitably chosen time step k. (c) What is the long term behavior of your solution? Can you find a formula for its eventual profile? (d) Design an implicit scheme for the same problem. Does the behavior of your numerical solution change? What are the advantages of the implicit scheme?