HW1

1. Try to use the values of u at $x+h, x, x-h$ and $x-2h$ to approximate $u'(x)$. Find the finite difference scheme that achieves the best accuracy. What is the order of accuracy for your scheme?

2. Show that the matrix constructed using the Matlab code

   ```
   e = ones(m,1);
   A = spdiags( [e -2*e e], -1:1, m, m)/h^2;
   ```

   is nonsingular (invertible).

3. Suppose that we want to discretize the following equation

   $$-(\kappa(x)u')' = f, \quad u(0) = u(1) = 0.$$

   The first way is $-\kappa u'' - \kappa' u' = f$ and

   $$-\kappa_j \frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} - \kappa_j' \frac{u_{j+1} - u_{j-1}}{2h} = f(x_j).$$

   A second way is to discretize the conservative form directly:

   $$-\frac{(\kappa u')_{j+1/2} - (\kappa u')_{j-1/2}}{h} = -\frac{\kappa_{j+1/2}(u^{j+1} - u_j) - \kappa_{j-1/2}(u^j - u^{j-1})}{h^2} = f(x_j).$$

   Which one is better? Give your reason.

4. Consider the elliptic equation with Neumann boundary condition on $\Omega = [0,1] \times [0,1]$.

   $$-\Delta u = f(x), \quad x \in \Omega$$
   $$\frac{\partial u}{\partial n} = 0, \quad x \in \partial\Omega.$$

   - Use 5-point discretization to formulate an FDM for this problem. Find the conditions on the point values $\{f_{ij}\}$ so that your numerical scheme has a solution $U = (u_{ij})$. (Note that for the boundary conditions, you may want to introduce the so-called ghost points.)
   - Try to explain how we can use Discrete Cosine Transform to solve your FDM in the first part. (If you like, you can code up, but this is not required.)

5. Consider the 2D elliptic equation

$$-(a(x,y)u_x)_x - (a(x,y)u_y)_y = f(x,y), \Omega = [-1,1] \times [-1,1]$$
$$u = 0, \ \ on \ \partial\Omega$$

$a = 1 + 3\exp(-3(x+y)^2 - (x-y)^2)$ and $f = 1$. Apply a five-point scheme for this equation. Determine the order of accuracy of your scheme, using the calculation with a small h as the 'exact' solution.

Hint: To solve the linear system $AU = F$, one option is to construct A directly and do $A\backslash F$. Here, the coefficient is not a constant, constructing this matrix might be a little tricky (remember to keep the matrix sparse in Matlab) (One bad way is to set the point value of U to be 1 at a single point, then output the action of the scheme on this U, which will be the corresponding column of your matrix.) Another better option is to write a function that returns AU when the input is U and then apply an iterative method (such as conjugate gradient) to find the solution. By doing this, you don't have to construct A.

6. Consider the following ODE for $u : [0, \infty) \to \mathbb{R}^d$:

$$u'(t) = b(u), \quad u(0) = u_0.$$

Assume that $b(0) = 0$ and b is Lipschitz continuous, i.e. there exists $L > 0$ such that

$$|b(x) - b(y)| \leq L|x - y|.$$

(a) Given $T > 0$, show that the forward Euler scheme gives the first order accuracy for $u(T)$.

(b) If b satisfies

$$(b(x) - b(y)) \cdot (x - y) \leq -r|x - y|^2$$

for some $r > 0$, show that the $O(k)$ error for the forward Euler scheme is global. This means that there exists $C > 0$ such that

$$\sup_{n \geq 0} |u_n - u(nk)| \leq Ck.$$

Hint: There are two options. First possible way is to consider

$$\dot{v} = b(v([t/k]k), [t/k]k)$$

where $[x]$ means the largest integer not exceeding x. Let $E = (v(t) - u(t))^2$. Control the derivative of E using E and k. The Gronwall's ienquality then gives the result. A second way is like this: one first shows that $|u(t_n) - u_n + k(b(u(t_n)) - b(u_n))| \le |u(t_n) - t_n|$; then show that the local truncation error is like $Ce^{-\gamma t_n} k^2$. Then, establish the following relation:

$$|u(t_{n+1}) - u_{n+1}| \le |u(t_n) - u_n| + Ce^{-\gamma t_n} k^2.$$

7. Below, we will look at a particular application of the discrete maximal principle.

   Consider the equation for π and $x \in \mathbb{T}$ (1D torus)

$$\frac{1}{2}\partial_x(a(x)\partial_x \pi(x)) - \partial_x(s(x)\pi) = 0,$$

   where $a(x) \ge a_0 > 0$. Both a and s are periodic smooth functions. Consider the following finite difference scheme

$$\begin{aligned}(L_h^* \pi)_j := \frac{1}{2h^2}&(a_{j+1/2}\pi_{j+1} - (a_{j+1/2} + a_{j-1/2})\pi_j + a_{j-1/2}\pi_{j-1}) \\ &- (\frac{s_j^+ \pi_j - s_{j-1}^+ \pi_{j-1}}{h} - \frac{s_{j+1}^- \pi_{j+1} - s_j^- \pi_j}{h}) = 0.\end{aligned}$$

   We want to show that for sufficiently small h

$$\max_{0 \le j \le N-1} \pi_j^h \le C \min_{0 \le j \le N-1} \pi_j^h. \tag{1}$$

   We prove this in the following steps:

   (a) The continuous problem has a unique solution up to a constant multiplicative factor. In particular, we choose the one with $\int \pi(x)\,dx = 1$. Show that this solution is positive everywhere.

   (b) The local truncation error for L_h^* satisfies

$$L_h^* \pi(\cdot) = \tau_j$$

   and we have $|\tau_j / h|$ is bounded uniformly in h.

   (c) The original scheme for π_j is equivalent to a scheme for $z_j = \pi_j / \pi(x_j)$:

$$(T_h z)_j = -z_j \tau_j.$$

   Show that T_h is a consistent difference operator for the operator

$$\tilde{L} := \frac{1}{2}\partial_x(a\pi\partial_x z) - (\frac{1}{2}a\partial_x \pi - \mu\pi)\partial_x z.$$

(d) Suppose z_j attains the maximal absolute value at j^*. Then, we consider

$$\zeta_j = \frac{z_j}{z_{j^*}} - 1.$$

Then,

$$T_h \zeta_j = -\frac{z_j}{z_{j^*}} \tau_j.$$

Construct a function ϕ such that

$$\tilde{L}\phi(x) = 1,\ x \neq x_{j^*},\ \phi(x_{j^*}) = 0.$$

Then, $\xi_j := 2\|\tau\|_\infty \phi(x_j) \pm \zeta_j$ satisfies

$$(T_h \xi)_j \geq 0,$$

when h is small enough.

(e) Show that both T_h and $\tilde{L}$ have some maximal principle, and conclude that

$$\max_j |\zeta_j| \leq 2\|\tau\|_\infty \|\phi\|_\infty.$$

Using the fact that $\|\phi\|_\infty$ is independent of x_{j^*} to conclude the claim.

8. Find the stability region of the following ODE solver for $u' = f(t, u)$:

$$u^{n+1} = u^n + k f(t^n + \frac{k}{2}, \frac{1}{2}(u^n + u^{n+1})).$$

Is this solver suitable for

$$f(t, u) = -100(u - e^{-t}) - e^{-t}?$$

Why?

9. This problem is to test your understanding of method of lines (MOL). Consider the transport equation $u_t + a u_x = 0$ with periodic boundary condition. We discretize the u_x at j using the scheme in Problem 1 to get a system of ODEs. We apply RK-r for time discretization. Compare the numerical results with RK2 and RK3 (since RK-r is not unique, you can use any one). Explain your observation.