Resit Numerical Mathematics 1, July 9th 2018, University of Groningen

Use of a simple calculator is allowed. All answers need to be justified.
There are four exercises, detailed on two pages in total. Each exercise has a certain amount of points. The grade will be computed as $grade = 1 + (points\ obtained)/(total\ points) * 9$. Please answer each exercise on a different sheet to facilitate the grading.

Exercise 1 (7 points)

Consider the ODE for $r(t)$:

$$r'(t) = -c\ r(t)\ , r(0) = r_0 \tag{1}$$

with $c \in \mathbb{R}^+$.

(a) 3.5 Applying the θ-method for the ODE (1) the following algebraic equation is obtained for $r_{k+1} \approx r(t_{k+1})$:

$$r_{k+1} = r_k - hc\theta r_{k+1} - hc(1-\theta)r_k$$

with $h = t_{k+1} - t_k$, $0 \leq \theta \leq 1$. Obtain a condition for h in terms of c, θ such that $|r_{k+1}| \leq |r_k|$. Explain how, in numerical computations, that condition affects the choice of h.

Repeat the exercise of the preparation of lab 6, but now all number are reals so that the answer is even easier. First, we write **(0.5 pt)**

$$r_{k+1} = \frac{1 - hc(1-\theta)}{1 + ch\theta} r_k = \frac{1 + hc\theta - hc}{1 + ch\theta} r_k = \left(1 - \frac{hc}{1 + ch\theta}\right) r_k$$

then we need **(0.5 pt)**

$$\left|1 - \frac{hc}{1 + ch\theta}\right| \leq 1$$

and hence **(0.5 pt)**

$$\frac{hc}{1 + ch\theta} \leq 2$$

and then for finding a general condition **(0.5 pt)**

$$hc \leq 2 + 2ch\theta$$

$$hc(1 - 2\theta) \leq 2$$

For $\theta \geq 1/2$, this is always true, no matter the value of $h > 0$ **(0.5 pt)**. When $\theta < 1/2$, then the condition is obtained **(0.5 pt)**

$$h \leq \frac{2}{c(1-2\theta)}.$$

In numerical computations, if an explicit method is used (i.e. $\theta = 0$), h has to be chosen small enough so that numerical does not diverge **(0.5 pt)**.

(b) 3.5 For $c = 2$, $h = 0.1$, compute the approximations at $t = h$ using $\theta = 0$, $\theta = 1/2$ and $\theta = 1$. Verify that the exact solution of the ODE is $r(t) = r_0 e^{-ct}$. Which value of θ leads to more accurate results (with respect to the exact solution)? Are the results you obtained as expected from the order of accuracy of each of the methods?

Recall

$$r_{k+1} = \left(1 - \frac{hc}{1+ch\theta}\right) r_k$$

Then for $\theta = 0$: $r_1 = (1 - hc)\, r_0 = 0.8 r_0$. **(0.5 pt)**
For $\theta = 1/2$: $r_1 = \left(1 - \frac{0.2}{1+0.1}\right) r_0 = 0.8181\ldots r_0$. **(0.5 pt)**
For $\theta = 1$: $r_1 = \left(1 - \frac{0.2}{1+0.2}\right) r_0 = 0.83333\ldots r_0$. **(0.5 pt)**
For verifying analytical solution, just replace in the ODE **(0.25 pt)** and check the initial condition **(0.25 pt)**. Exact solution is 0.8187 **(0.5 pt)**. Clearly $\theta = 1/2$ is more accurate **(0.5 pt)**. This is because $\theta = 1/2$ is second order accurate, while the others are only first **(0.5 pt)**.

Exercise 2 (7 points)

Consider the non-linear ODE for $u(t)$:

$$u' = -u^3 - u\ , u(0) = u_0 \tag{2}$$

(a) 1.0 Using one time-step of the forward Euler method ($\theta = 0$), find an expression for $u_1 \approx u(t_1)$, $t_1 > 0$. If a non-linear equation for u_1 results after discretization, approximate u_1 using one Newton iteration with u_0 as initial guess.

We apply the FE method to the problem leading to **(0.5 pt)**

$$u_1 = u_0 - t_1 u_0^3 - t_1 u_0$$

since the FE is explicit, we do not need to use a root-finding method for this problem **(0.5 pt)**.

(b) 2.0 Using one time-step of the backward Euler method ($\theta = 1$), find an expression for $u_1 \approx u(t_1)$, $t_1 > 0$. If a non-linear equation for u_1 results after discretization, approximate u_1 using one Newton iteration with u_0 as initial guess.

We apply the BE method to the problem leading to **(0.5 pt)**

$$u_1 = u_0 - t_1 u_1^3 - t_1 u_1$$

so we obtain a root-finding problem for u_1: **(0.5 pt)**

$$0 = -u_1 + u_0 - t_1 u_1^3 - t_1 u_1$$

Using u_0 as initial guess, one Newton iteration reads:

$$u_1 = u_0 - \frac{1}{-1 - 3t_1 u_0^2 - t_1}(-u_0 + u_0 - t_1 u_0^3 - t_1 u_0)$$

(0.5 pt) for writing correctly the formula, **(0.5 pt)** for replacing the correct expressions in the formula.

(c) 0.5 Show that you can write the ODE (2) as

$$u(t) = u(0) - \int_0^t (u^3(s) + u(s))\ ds \tag{3}$$

Just integrate the ODE between some 0 and t. **(0.5 pt)**

(d) 1.5 Explain by using a drawing how the backward Euler, forward Euler, and Crank-Nicolson ($\theta = 1/2$) methods are built by approximating (without subintervals) the integral in (3).

Draw the area defined by the rectangles, **(0.5 pt)** for each of the methods.

(e) [2.0] Derive a numerical method for approximating $u(t_1)$ in terms of $u(0)$ from a Simpson's rule for the integral in (3), including the assumption that $u(t_1/2) \approx (u(t_1) + u(0))/2$. Remember that the Simpson's rule is the exact integral of the quadratic polynomial interpolation of the integrand. Is the resulting method explicit or implicit?

For applying (or deriving) Simpson to the integral **(1.0 pt)**. For replacing the assumption **(0.5 pt)**. For saying that it is implicit **(0.5 pt)**.

Exercise 3 (6 points)

Consider the matrix A given by:

$$A = \begin{bmatrix} a & -c \\ -c & a \end{bmatrix} \ , \ a, c > 0.$$

(a) [1.5] Compute the eigenvalues of A. Determine the range of values for a and c, such that A is positive definite.

The eigenvalues of A are given by:

$$\det \begin{bmatrix} a-\lambda & -c \\ -c & a-\lambda \end{bmatrix} = (a-\lambda)^2 - c^2 = a^2 - 2a\lambda + \lambda^2 - c^2 = 0 \ \textbf{(0.5 pt)}$$

leads to

$$\lambda_\pm = \frac{2a \pm \sqrt{4a^2 - 4(a^2 - c^2)}}{2} = a \pm c \ \textbf{(0.5 pt)}$$

$\lambda_+ > 0$ if $c > -a$ and $\lambda_- > 0$ if $a > c$ **(0.25 pt)**. Therefore, the matrix is positive definite if $a > c$. **(0.25 pt)**

Other ways to check positive definiteness are of course valid. But the eigenvalues need the eigenvalues later so computing them here is the shortest way.

(b) [1.0] For a and c in the previously computed range, calculate the 2-norm condition number $K_2(A)$ of A. Then, determine δ such that

$$\lim_{\frac{a}{c} \to \delta} K_2(A) = \infty$$

Since the matrix is symmetric positive definite

$$K_2(A_\varepsilon) = \frac{\lambda_{max}}{\lambda_{min}} = \frac{a+c}{a-c} \ \textbf{(0.5 pt)}$$

clearly, if $c \to a^+$ then $K_2 \to \infty$. Therefore, $\delta = 1^+$ **(0.5 pt)**. *There are ways to compute the condition number, the most general one being $K_2(A_\varepsilon) = \|A_\varepsilon\|_2 \|A_\varepsilon^{-1}\|_2$, what is fine if they get to the same result.*

(c) [1.0] Comment on two different issues that arise when solving linear systems (directly and/or iteratively) when the condition number of the matrix is large.

The issues that we have seen in the course are: (a) increased sensitivity to perturbations (e.g. due to rounding-off errors) **(0.5 pt)** and (b) slow convergence when using iterative methods for solving linear systems **(0.5 pt)**.

(d) [1.5] Compute the LU-factorization of A.

We define the LU factorization such that $A = LU$

$$A = \begin{bmatrix} a & -c \\ -c & a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \ell & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{21} \\ 0 & u_{22} \end{bmatrix} \ \textbf{(0.5 pt)}$$

hence by solving a linear system for $\ell, u_{11}, u_{21}, u_{22}$ we obtain

$$L = \begin{bmatrix} 1 & 0 \\ -c/a & 1 \end{bmatrix} \ \textbf{(0.5 pt)}, \ U = \begin{bmatrix} a & -c \\ 0 & a - c^2/a \end{bmatrix} \ \textbf{(0.5 pt)}$$

(e) [1.0] Consider the following system of ODEs:

$$X'(t) = -KX(t) \ , \ X(0) = X_0 \ , \ K = \begin{bmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & A \end{bmatrix}.$$

If discretizing in time with the forward Euler method, determine the range of values for the time-step in terms of a and c such that the discrete solution $X_k \to 0$ when $k \to \infty$, with $X_k \approx X(t_k)$. Assume a, c such that A is positive definite.

K is positive definite, and the largest eigenvalue corresponds to the largest eigenvalue of A (since K is block-diagonal), so $\lambda_{max} = a + c$ **(0.5 pt)**. We know (for instance from the labs, lectures, that the stability condition for the θ-method for systems of ODEs (for instance the heat equation), and hence for the forward Euler is

$$h < \frac{2}{\lambda_{max}} = \frac{2}{a+c} \textbf{ (0.5 pt)}.$$

(0.25 pt) *For saying that the spectral radius of $I - h * A$ or $I - h * K$ has to be smaller than 1.*

Exercise 4 (5 points)

(a) 1.0 We want to solve the linear system $Ax = b$ for x by using stationary Richardson iterations:

$$x^k = x^{k-1} + \alpha \left(b - Ax^{k-1} \right) \tag{4}$$

using as initial guess the vector $[0, 0]^\intercal$. The matrix A is the one given in Exercise 3 (with a, c such that A is symmetric positive definite). Choose a value of α in terms of a and c so that convergence of the Richardson iterations towards $A^{-1}b$ is ensured. Justify your choice in view of the theory.

Any positive value smaller than $2/\lambda_+ = 2/(a + c)$ will ensure convergence. For instance, the "optimal" one $2/(a + c + a - c) = 1/a$ **(0.5 pt)**. Such values ensure that spectral radius of the iteration matrix $I - \alpha A$ for the error is smaller than 1 **(0.5 pt)**.

(b) 1.0 Find α_{max} in terms of a and c such that $\alpha < \alpha_{max}$ ensures convergence of the Richardson method.

Again, the answer is any positive value smaller than $2/\lambda_+ = 2/(a + c)$ will ensure convergence **(0.5 pt)**, such that the spectral radius of the iteration matrix for the error is smaller than 1 **(0.5 pt)**.

(c) 1.0 Explain if it is possible to directly apply a Gauss-Seidel preconditioner to the iterative method given by (4), with A as in Exercise 3. If you think that it is not possible, how can the linear system be rewritten such that we can use a Gauss-Seidel preconditioner? Write the preconditioner matrix according to your answer.

It is indeed possible to write a Gauss-Seidel preconditioner **(0.5 pt)**, which has the form

$$P = \begin{bmatrix} a & 0 \\ -c & a \end{bmatrix} \textbf{ (0.5 pt)}$$

(d) 1.0 Explain if it is possible to directly apply a Gauss-Seidel preconditioner to the iterative method given by (4), with A as:

$$A = \begin{bmatrix} a & -c \\ c & 0 \end{bmatrix}$$

(assume A invertible). If you think that it is not possible, how can the linear system be rewritten such that we can use a Gauss-Seidel preconditioner? Write the preconditioner matrix according to your answer.

It is not possible to use a Gauss-Seidel preconditioner since it is not invertible

$$P = \begin{bmatrix} a & 0 \\ -c & 0 \end{bmatrix} \textbf{ (0.25 pt)}$$

However, switching the first and second columns of the linear system (this means only changing the order of the unknowns in the vector x) leads to a new system matrix

$$A = \begin{bmatrix} -c & a \\ 0 & c \end{bmatrix} \textbf{ (0.5 pt)}$$

and the corresponding GS-preconditioner:

$$P = \begin{bmatrix} -c & 0 \\ 0 & c \end{bmatrix} \textbf{ (0.25 pt)}$$

(e) 1.0 Consider the scalar algebraic equation $ay = d$, for $a < 0, d$ given, and y being the unknown. Using stationary Richardson iterations, compute the range of values of the relaxation parameter α such that convergence is ensured. What is the value of that leads to the fastest convergence?

First rewrite the system such that $-ay = -d$, to make the system "matrix" positive definite. As in the previous questions, the range of values is $\alpha < 2/(-a)$ **(0.5 pt)**. Convergence in one iteration is achieved with $\alpha = 1/(-a)$ **(0.5 pt)**.