Lösning: Tenta Numerical Analysis för D, L. FMN011, 090527

This exam starts at 8:00 and ends at 12:00. To get a passing grade for the course you need 35 points in this exam and an accumulated total (this exam plus the two computational exams) of 50 points.

Justify all your answers and write down all important steps. Unsupported answers will be disregarded.

During the exam you are allowed a pocket calculator, but no textbook, lecture notes or any other electronic or written material.

1. **(5p)** Name one numerical method that is suitable for solving the following problems:

   (a) $x - e^x = 0$

   (b) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$

   (c) Find a fourth degree polynomial passing through 5 given points.

   (d) Construct a smooth curve passing through 50 given points.

   (e) Solve the initial value problem $y' = y^2/t + \sin y$, $y(1) = 1$ in the interval $[1, 5]$.

   Solution:

   (a) Fixed-point iteration, Newton-Raphson, bisection.

   (b) Back substitution.

   (c) Polynomial interpolation (Lagrange interpolation, Newton divided differences).

   (d) Splines, Bézier curves.

   (e) Euler's method, Heun's method, a Runge-Kutta method.

2. **(5p)** Consider the three methods

   (a) $x_{n+1} = f(x_n)$

   (b) $p_{k+1} = p_k - \frac{f(p_k)}{f'(p_k)}$

   (c) $p_k = p_{k-1} - \frac{f(p_{k-1})(p_{k-1}-p_{k-2})}{f(p_{k-1})-f(p_{k-2})}$

State (all) those that have each of the following properties:

Solution:

(a) Solves $f(x) = 0$. *(b,c)*

(b) Usually has quadratic convergence. *(b)*

(c) Will converge only if $|f(x)| < 1$. *(none)*

(d) Is an iteration method. *(a,b,c)*

(e) Can be used to find a root of $16x^2 - 32x + 15 = 0$, with $f(x) = \sqrt{32x - 15}/4$. *(a)*

3. **(5p)** Suppose you have a square linear system of equations $Ax = b$. Name the three elementary row operations that yield an equivalent system.

Solution:

(a) Multiply a row by a scalar

(b) Interchange rows

(c) Subtract a multiple of one row from another

4. **(5p)** The following statement is wrong. Correct the statement.

The iterative method

$$x_{k+1} = Bx_k + c$$

is convergent (for any initial guess x_0) only if $\|B\|_p < 1$ for some p.

Solution:

The iterative method

$$x_{k+1} = Bx_k + c$$

is convergent (for any initial guess x_0) if $\|B\|_p < 1$ for some p.

Or:

The iterative method

$$x_{k+1} = Bx_k + c$$

is convergent (for any initial guess x_0) only if $\rho(B) < 1$.

5. **(5p)** Write the polynomial passing through the points (0, 1), (-2.5, 7.2) and (3.1, -2), if the Lagrange polynomials have been computed and are $x(x + 2.5)/17.36$, $x(x - 3.1)/14$ and $(x + 2.5)(3.1 - x)/7.75$.

Solution:

$$p(x) = \frac{(x + 2.5)(3.1 - x)}{7.75} + 7.2\,\frac{x(x - 3.1)}{14} - 2\,\frac{x(x + 2.5)}{17.36}$$

6. **(5p)** Consider the Bézier curve with control points (2,3), (1,-1), (4,1) and (5,0). Remembering that the cubic Bernstein polynomials are $(1-t)^3, 3(1-t)^2t, 3(1-t)t^2, t^3$, write down the parametric equations of the curve, and give the first control point of another cubic Bézier curve that starts at (5,0) and ends at (6,-2), and joins the previous curve as smoothly as possible.

   Solution:

$$
\begin{aligned}
X(t) &= 2(1-t)^3 + 3(1-t)^2t + 12(1-t)t^2 + 5t^3 \\
Y(t) &= 3(1-t)^3 - 3(1-t)^2t + 3(1-t)t^2
\end{aligned}
$$

   Any point collinear with (4, 1) and (5, 0) is OK, that is, of the form (x, 5-x).

7. **(5p)** The following table is found on an oatmeal porridge package:

Portions	Oats	Water
1	1	2.5
2	2	4.5
4	4	9

   Determine the straight line that passes through (0, 0) and best fits the given data.

   Solution:
   The equation of the line must be of the form $y = mx$, where y is the amount of oats and x is the number of portions (or the amount of water). The overdetermined is system is $\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} m = \begin{pmatrix} 2.5 \\ 4.5 \\ 9 \end{pmatrix}$, *and its least squares solution is approximately $m = 2.26$, so the straight line is $y = 2.26x$.*

8. **(5p)** Write an algorithm that solves the least squares problem $Ax = b$, where A is an 25×3 matrix, using QR factorization.

   Solution:

```
[Q,R] = qr(A);
y = Q(:,1:3)'*b;
x = R(1:3,1:3)\y;
```

9. **(5p)** Consider the formula

$$f''(t) = \frac{f(t+2h) - 2f(t+h) + f(t)}{h^2} - hf'''(\zeta)$$

Suppose that the rounding errors are such that

$$\hat{f}(t) \approx f(t) + \epsilon$$

Calculate the total error function $E(h)$, such that $|f''(t) - \hat{f}''(t)| \leq E(h)$.

Solution:

$$\begin{aligned} |error_{truncation}| &= h\,|f'''(\zeta)| \\ |error_{rounding}| &\leq (\epsilon + 2\epsilon + \epsilon)/h^2 \end{aligned}$$

$$|f''(t) - \hat{f}''(t)| \leq h\,|f'''\zeta)| + \frac{4\epsilon}{h^2}$$

10. **(5p)** Describe how the approximation

$$\int_a^b f(x)dx \approx S(a,c) + S(c,b) + \frac{1}{15}[S(a,c) + S(c,b) - S(a,b)]$$

can be used to construct an adaptive Simpson quadrature algorithm.

Solution: *The term $[S(a,c) + S(c,b) - S(a,b)]/15$ can be used as an error estimate. If this quantity is less than or equal to the user-provided tolerance, the integral is approximated by the given formula. If the quantity is greater than the tolerance, the intervals [a,c] and [c,b] are bisected and the test is redone for each subinterval [a,c] and [c,b], this time checking against one half of the user-provided tolerance.*

11. **(5p)** When an initial value problem was solved with a certain Runge-Kutta method using different step-sizes, the following errors were observed:

h	norm-1 error/N
1/1000	7.8013×10^{-9}
1/2000	1.9499×10^{-9}
1/4000	4.8743×10^{-10}
1/8000	1.2185×10^{-10}

What is the order of convergence of this method?

Solution:

$$\begin{aligned} 7.8013 \times 10^{-9}/1.9499 \times 10^{-9} &= 4.0009 \\ 1.9499 \times 10^{-9}/4.8743 \times 10^{-10} &= 4.0004 \\ 4.8743 \times 10^{-10}/1.2185 \times 10^{-10} &= 4.0002 \end{aligned}$$

If the error is of the form ch^p*, then* $\dfrac{ch^p}{c(h/2)^p} = 2^p$. As here $2^p \approx 4$, the method is of order $p = 2$.

12. **(5p)** Write the MATLAB `function yprime=myode(t,y)` you would need to solve the differential equation $y''' = yt - 3y'$.

    Solution:

$$\begin{aligned} y' &= v \\ v' &= z \\ z' &= yt - 3v \end{aligned}$$

```
function yprime=myode(t,y)
yprime(1,1)=y(2);
yprime(2,1)=y(3);
yprime(3,1)=y(1)*t-3*y(2);
```

13. **(5p)** Study the following algorithm:

$$\begin{aligned} y_{k-1} &= \frac{x_{k-1}}{\|x_{k-1}\|_2} \\ x_k &= Ay_{k-1} \\ \lambda_k &= y_{k-1}^T x_k \end{aligned}$$

    What method is this and what does it compute?

    Solution: *It is the normalized power method and it computes an approximation to the largest-modulus eigenvalue of* A *and its corresponding eigenvector.*

14. **(5p)** Suppose a 128×128 matrix A has a singular value decomposition $A = USV^T$, and the first two singular values s_1 and s_2 are much larger than the rest of the singular values. How can one compress the main part of the information by using the low-rank approximation property

$$A = \sum_{i=r}^{n} s_i u_i v_i^T$$

    and what would the compression ratio be?

    Solution: *One can use the first 2 terms of the sum,*

$$A \approx s_1 u_1 v_1^T + s_2 u_2 v_2^T$$

    to compute the best least squares approximation to A *of rank 2. As* A *contains* 128^2 *elements, and* $s_1 u_1 v_1^T + s_2 u_2 v_2^T$ *contains* $2 + 4 \cdot 128$ *elements, the compression ratio is approximately 32:1.*

Lycka till!
C.Arévalo