TMA4329 Intro til vitensk. beregn. V2016

ving E01

[S]=T. Sauer, Numerical Analysis, Second International Edition, Pearson, 2014

1. Solution of equations Let $f(x) = ax + b$, where $a \neq 0$.

a) Let $x_1, x_2 \in \mathbb{R}$ be such that $f(x_1)f(x_2) < 0$. For which values of a, b, does the bisection algorithm applied to f converges starting from $[x_1, x_2]$.

Solution: The function f is continuous and satisfies the sign condition $f(x_1)f(x_2) < 0$, thus there is at least one root on the interval $[x_1, x_2]$. The bisection algorithm subdivides the interval into two subintervals and chooses the subinterval containing at least one root (i.e. satisfying the sign condition). Thus the root is bracketed by smaller and smaller subintervals regardless of the values of the constants a and b, and the algorithm converges unconditionally in this case.

Consider now a fixed point iteration $x_{k+1} = (a + 1)x_k + b$.

b) Show that any fixed point of this iteration is a root of the equation $f(x) = 0$, and vice versa.

Solution: Suppose that $\hat{x}$ is the root of our equation: $0 = f(\hat{x}) = a\hat{x} + b$. By adding $\hat{x}$ to both sides of this equation we obtain that it is also a fixed point: $\hat{x} = (a + 1)\hat{x} + b$.

Similarly, assuming that $\tilde{x}$ is a fixed point, that is, $\tilde{x} = (a + 1)\tilde{x} + b$ and by subtracting $\tilde{x}$ from both sides of the equation we obtain $0 = a\tilde{x} + b = f(\tilde{x})$. Therefore it is also a root of our equation.

c) For which values a, b does the fixed point iteration algorithm converge for any starting point?

Solution: Let $\hat{x} = -b/a$ be the root of our equation. Then we have $\hat{x} = (a + 1)\hat{x} + b$. For the fixed-point iteration we have $x_{k+1} = (a + 1)x_k + b$. Thus $x_{k+1} - \hat{x} = (a + 1)(x_k - \hat{x}) = (a + 1)^2(x_{k-1} - \hat{x}) = (a + 1)^{k+1}(x_0 - \hat{x})$. The latter quantity (which is the error of the fixed point iteration) goes to zero for an arbitrary starting point x_0 if and only if $|a + 1| < 1$, or $-2 < a < 0$.

d) For which values a, b does the fixed point iteration algorithm exhibits *faster* convergence rate than the bisection algorithm?

Solution: In the case of the bisection algorithm, at each iteration we are guaranteed to *divide* the bracketing interval by a factor of 2, whereas in the case of the fixed point iteration the error is *multiplied* by $a + 1$ (see the previous question). Thus we can expect FP iteration to converge faster than the bisection algorithm when $|a + 1| < 1/2$, or $-3/2 < a < -1/2$.

2. Interpolation of functions Throughout this question we put $f(x) = x^3 + x^2 + 1$.

a) Compute the lowest degree polynomial, which passes through the points $(0, f(0))$, $(1, f(1))$, $(2, f(2))$.

Solution: For example, we can use Lagrange form of the interpolation polynomial here:

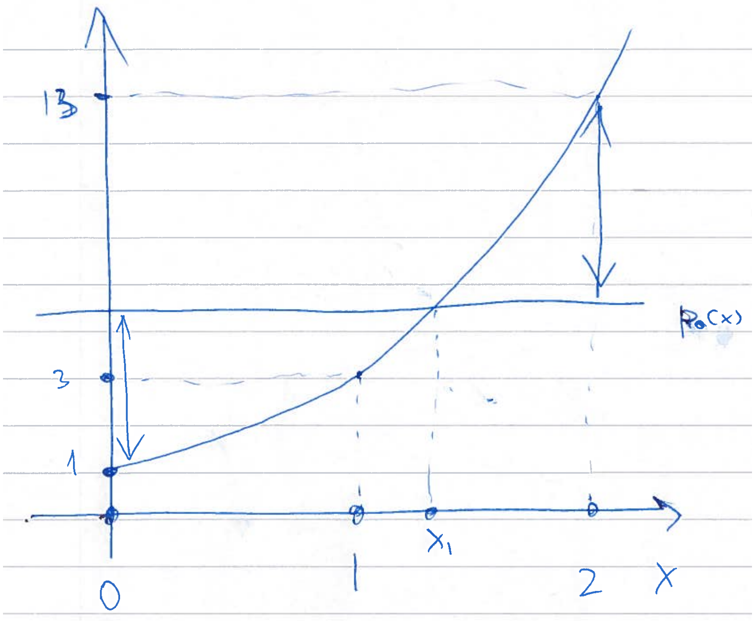
$$\begin{aligned}
L_0(x) &= \frac{(x-1)(x-2)}{(0-1)(0-2)} = \frac{x^2-3x+2}{2} \\
L_1(x) &= \frac{(x-0)(x-2)}{(1-0)(1-2)} = -x^2+2x \\
L_2(x) &= \frac{x(x-1)}{(2-0)(2-1)} = \frac{x^2-x}{2} \\
p(x) &= f(0)L_0(x) + f(1)L_1(x) + f(2)L_2(x) \\
&= 1\frac{x^2-3x+2}{2} + 3(-x^2+2x) + 13\frac{x^2-x}{2} \\
&= 4x^2 - 2x + 1.
\end{aligned}$$

b) Compute the lowest degree polynomial, which passes through the points $(0, f(0))$, $(1, f(1))$, $(2, f(2))$, $(\pi, f(\pi))$.

Solution: There is only one polynomial of degree 3 or less passing through 4 distinct points, namely $f(x)$.

c) Let $p_0(x) = f(x_1)$ be the zeroth degree polynomial passing through the point $(x_1, f(x_1))$. Consider the problem of finding the point $x_1 \in [0, 2]$ such that the quantity $\max_{x\in[0,2]} |f(x) - p_0(x)|$ is as small as possible. Graphically show that $x_1 \approx 1.5377\ldots$, the point that satisfies the equation $f(x_1) = 7$.

Solution:



↕ – this is where the distance between $f(x)$ and $p_0(x)$ is largest

From the picture we can see that

$$\begin{aligned}\min_{x_1\in[0,2]}\max_{x\in[0,2]}|f(x)-p_0(x)| &= \min_{x_1\in[0,2]}\max\{p_0(0)-f(0), f(2)-p_0(2)\}\\ &= \min_{x_1\in[0,2]}\max\{f(x_1)-1, 13-f(x_1)\}.\end{aligned}$$

Thus the minimum is attained when $f(x_1)$ is exactly in the middle between 1 and 13, that is, $(1+13)/2 = 7$. From this x_1 can be determined by solving the equation $f(x_1) = 7$ on $[0,1]$.

d) The Chebyshev node for $n = 1$ on the interval $[0,2]$ is $\hat{x}_1 = 1$ (the middle of the interval). In view of the optimality of Chebyshev interpolation, explain why it is possible for x_1 in the previous question *not* to coincide with $\hat{x}_1$?

Solution: Chebyshev interpolation nodes $\hat{x}_i$ are optimal in the sense that they minimize the worst case interpolation error for *all* possible (smooth enough) functions f over a given interval $[a,b]$, that is

$$\min_{\hat{x}_1,\dots,\hat{x}_n\in[a,b]}\max_{f:[a,b]\to\mathbb{R}}\max_{x\in[a,b]}|\hat{p}_n(x)-f(x)|,$$

where $\hat{p}_n$ is the interpolation polynomial for f at the nodes $\hat{x}_1,\dots,\hat{x}_n$. It is therefore possible, that for a *fixed* function one obtains a smaller worst case interpolation error over a given interval by selecting interpolation nodes *specifically for this function*, as we have done in the previous question:

$$\min_{x_1,\dots,x_n\in[a,b]}\max_{x\in[a,b]}|p_n(x)-f(x)|.$$

3. Numerical integration Midpoint quadrature rule is defined by/satisfies the equation

$$\int_{x_0}^{x_1} f(x)\,\mathrm{d}x = hf(w) + \frac{h^3}{24}f''(c), \tag{1}$$

where $c \in [x_0, x_1]$, $h = x_1 - x_0$, $w = (x_0 + x_1)/2$, assuming that f is twice continuously differentiable on $[a, b]$.

a) Provide an upper (pessimistic) estimate the error for the mid-point approximation of $\int_0^1 \exp(x^2)\,\mathrm{d}x$.

Solution: Let $f(x) = \exp(x^2)$, then $f''(x) = \exp(x^2)(4x^2 + 2)$. The latter function is a product of two positive and monotonically increasing functions on $[0, 1]$ and is therefore monotonically increasing funciton itself. Therefore, the pessimistic error estimate could be

$$\max_{c\in[0,1]} \frac{h^3}{24}|f''(c)| \le \frac{1^3}{24}f''(1) = \frac{6e}{24} = \frac{e}{4} \approx 0.67957.$$

b) Provide an estimate of the error term in (1) for the integral in **a)** using the adaptive quadrature idea, that is, by applying a composite midpoint quadrature on $[x_0, x_1]$ with two panels.

Solution: The idea of the estimate is as follows:

$$\begin{aligned} hf(w) + \frac{h^3}{24}f''(c) &= \int_{x_0}^{x_1} f(x)\,\mathrm{d}x = \int_{x_0}^{x_2} f(x)\,\mathrm{d}x + \int_{x_2}^{x_1} f(x)\,\mathrm{d}x \\ &= \frac{h}{2}f(w_1) + \frac{h^3}{8\cdot 24}f''(c_1) + \frac{h}{2}f(w_2) + \frac{h^3}{8\cdot 24}f''(c_2), \end{aligned}$$

where $x_2 = (x_0 + x_1)/2 = w$, $w_1 = (x_0 + x_2)/2$, $w_2 = (x_2 + x_1)/2$, $c_1 \in [x_0, x_2]$, and $c_2 \in [x_2, x_1]$. We then assume $c \approx c_1 \approx c_2$ to compute

$$\frac{3}{4}\frac{h^3}{24}f''(c) \approx \frac{h}{2}[f(w_1) + f(w_2)] - hf(w).$$

Let us substitute the numbers: $x_0 = 0$, $x_1 = 1$, $w = 1/2$, $w_1 = 1/4$, $w_2 = 3/4$, $h = 1$, $f(w) \approx 1.2840$, $f(w_1) = 1.0645$, $f(w_2) = 1.7551$, $h/2[f(w_1) + f(w_2)] - hf(w) \approx 0.12575$. Therefore the error term $\frac{h^3}{24}f''(c) \approx \frac{4}{3}0.1575 \approx 0.16767$.

4. Solution of ODEs Consider the initial value problem $y''(t) + [y'(t)]^2 = 1$, $y(0) = y'(0) = 1$.

a) Rewrite the problem as a system of first order ODEs.

Solution: We put $y_1(t) = y(t)$, $y_2(t) = y_1'(t)$. Then the sought system is

$$\begin{aligned} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}'(x) &= \begin{pmatrix} y_2(x) \\ 1 - y_2^2(x) \end{pmatrix}, \\ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(0) &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

b) Apply one step of the explicit trapezoid method with steplength $h = 0.1$.

Solution: Explicit trapezoid method is an explicit Runge–Kutta method with two stages. We put

$$k_1 = \begin{pmatrix} y_2(0) \\ 1 - y_2^2(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 - 1^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$k_2 = \begin{pmatrix} y_2(0) + hk_{1,2} \\ 1 - (y_2(0) + hk_{1,2})^2 \end{pmatrix} = \begin{pmatrix} 1 + 0.1 \cdot 0 \\ 1 - (1 + 0.1 \cdot 0)^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Finally

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} + \frac{h}{2}(k_1 + k_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0.05 \begin{pmatrix} 1 + 1 \\ 0 + 0 \end{pmatrix} = \begin{pmatrix} 1.1 \\ 1 \end{pmatrix}$$

5. DFT

a) Compute DFT of a sequence $[0, 1, 2, 3]$.

Solution:

$$y_0 = \frac{1}{\sqrt{4}} \sum_{j=0}^{3} x_j \exp 0 = \frac{1}{2} \sum_{j=1}^{3} x_j = 3,$$

$$y_1 = \frac{1}{\sqrt{4}} \sum_{j=0}^{3} x_j \exp(-i2\pi j/4) = \frac{1}{2}[0 \cdot 1 + 1 \cdot (-i) + 2 \cdot (-1) + 3 \cdot (i)] = -1 + i$$

$$y_2 = \frac{1}{\sqrt{4}} \sum_{j=0}^{3} x_j \exp(-i2\pi 2j/4) = \frac{1}{2}[0 \cdot 1 + 1 \cdot (-1) + 2 \cdot (1) + 3 \cdot (-1)] = -1.$$

Finally, since $x \in \mathbb{R}^4$ then $y_3 = \bar{y}_1 = -1 - i$.

b) If the input sequence x is *real* than its DFT y satisfies the properties: (i) $y_0 \in \mathbb{R}$, (ii) $y_{n-i} = \bar{y}_i$, $i = 1, \ldots, n-1$.
Suppose now that we know that the "output" sequence y (which is still DFT of x) is real. What can we say about x?

Solution: Let F_n be the $n \times n$ complex matrix describing the discrete Fourier transform. Then $y = F_n x$ and $\bar{y} = y = \bar{F}_n \bar{x}$. Thus $\bar{x} = F_n y$ (because $\bar{F}_n = F_n^{-1}$), and $\bar{x}$ is a discrete Fourier transform of a real sequence y. Therefore $\bar{x}_0 \in \mathbb{R}$ and $\bar{x}_{n-i} = \bar{\bar{x}}_i$, $i = 1, \ldots, n-1$, which is the same as $x_0 \in \mathbb{R}$ and $x_{n-i} = \bar{x}_i$, $i = 1, \ldots, n-1$.