APPLIED MATHEMATICS – THE FUNDAMENTALS OF THE NUMERICAL MATHEMATICS

Lecture Notes

Mistakes

Numerical mathematics

- transformation the continuous task into discrete

(integral, derivation etc. -> ar. Operations +,-,*,/)

Input data -> Algorithm -> Output data

-> +,-,*,/

-> substitution

-> conditioned commands

-> …

- Input data = finite set of the numbers which we need for finding to the unique solution

- algorithm = finite set of the computational steps

- output data = finite set of the numbers, they are the solution of given task

- we are not able to solve problems in their original extent -> MODEL -> - the difference between the result obtained and the actual solution to the original problem (we must be able to estimate how much inaccuracy is)

Mistakes – division according to where they arise

Mistakes of the mathematical model – Neglect of some facts (difference between model and actual state)

Ex: Calendar as a model of the year (Julian: introduced by Julius Caesar

longest 365,25 days -> every 4th leap-year

-> model is delaying

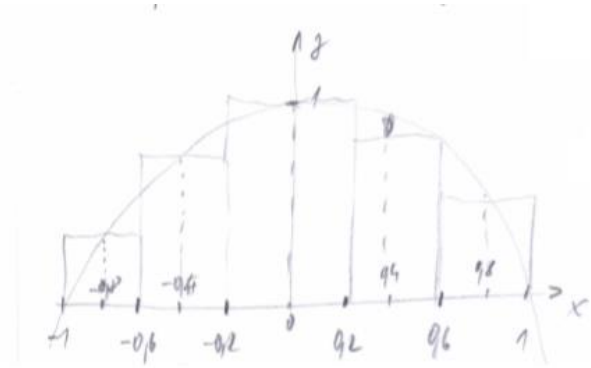
-> mistake is 1 day per 128 years

Mistakes of the input data – Input data is gained by measuring => mistakes of the measuring
=> multiple errors

Mistakes of the numerical method -> continuous task -> discret

Ex: Calculation of the definite integral $\int_{-1}^{1}(1-x^2)dx$ – the area under the graph

$$\int_{-1}^{1}(1-x^2) = \left[x - \frac{x^3}{3}\right]_{-1}^{1} = \left(1-\frac{1}{3}\right) - \left(-1+\frac{1}{3}\right) = \frac{4}{3} = 1,\overline{3}$$



Discretely: dividing to 0,4

$$\int_{-1}^{1}(1-x^2)dx = f(-0{,}3)\times 0{,}4 + f(-0{,}4)\times 0{,}4 + f(0)\times 0{,}4 + f(0{,}2)\times 0{,}4 + f(0{,}8)\times 0{,}4 =$$

$$= 1{,}36$$

$x = 1,\overline{3} \qquad |\tilde{x} - x| = 0{,}02\overline{6}$

$\tilde{x} = 1{,}36 \qquad \frac{|\tilde{x}-x|}{x} = 0{,}0196 \doteq 0{,}02 \rightarrow 2\% \quad Mistake\ of\ the\ method\ is\ almoust\ 2\%$

Rounding mistakes→ we are rounding the numbers → mistakes → cumulation

$\left.\begin{matrix} x \ldots real\ number \\ \tilde{x} \ldots aproximation \end{matrix}\right\rangle\ procces\ is\ called\ rounding$

$the\ mistake\ of\ the\ aproximation \quad \begin{matrix} Absolut\ mistake & \varepsilon(\tilde{x}) = |x-\tilde{x}| \\ Relativ\ mistake & \delta(\tilde{x}) = \frac{|x-\tilde{x}|}{\tilde{x}} \end{matrix}$

$\rightarrow estimate\ of\ the\ AM \qquad |x-\tilde{x}| \leq \varepsilon(\tilde{x})$

$\rightarrow estimate\ of\ the\ RM \qquad \frac{|x-\tilde{x}|}{\tilde{x}} \leq \delta(\tilde{x})$

$\rightarrow RM\ is\ independent\ to\ the\ choice\ of\ the\ unit$

$\rightarrow the\ rate\ of\ the\ accuracy\ of\ the\ computing \rightarrow \delta(\tilde{x})\times 100 \ldots . RCH\ v\ \%$

EX:

$\left.\begin{matrix} x_1 = 0{,}32 & \tilde{x}_1 = 0{,}3 \\ x_2 = 0{,}08 & \tilde{x}_2 = 0{,}1 \end{matrix}\right\rangle\ Are\ the\ approximations\ of\ the\ numbers\ x_1, x_2\ equally\ valuable?$

$\left.\begin{matrix} \varepsilon(\tilde{x}_1)=0{,}02 \\ \varepsilon(\tilde{x}_2)=0{,}02 \end{matrix}\right\rangle\ the\ same\ value \qquad \left.\begin{matrix} \delta(\tilde{x}_1)=\frac{0{,}02}{0{,}3}=0{,}0\overline{6} \rightarrow 6\% \\ \delta(\tilde{x}_2)=\frac{0{,}02}{0{,}1}=0{,}2 \rightarrow 20\% \end{matrix}\right\rangle\ They\ are\ not\ equally\ valuable$

- When the rounding is OK: the absolute mistake never rise half unit of the last left digit

number of significant digits (NSD): number x (the decimal system, standardized shape) mantisa $\in (0; 1)$

$$x = 0,\alpha_1\alpha_2\alpha_3 \dots \alpha_k\alpha_{k+1} \dots \cdot 10^N$$

We say, the approximation $\tilde{x}$ of the number x has j-th digit significant, if it applies:

$$|x - \tilde{x}| \leq 0{,}5 \cdot 10^{n-j}$$

EX:

$x = \pi = 3{,}14159 \dots \tilde{x} = 3{,}1415$

Standardized shape $0{,}31415 \cdot 10^1$ ($n = 1$)

$|3{,}14159 - 3{,}1415| = 0{,}00009$

$j = 0$	$0{,}5 \cdot 10^1 = 5$	✓
$j = 1$	$0{,}5 \cdot 10^0 = 0{,}5$	✓
$j = 2$	$0{,}5 \cdot 10^{-1} = 0{,}05$	✓
$j = 3$	$0{,}5 \cdot 10^{-2} = 0{,}005$	✓
$j = 4$	$0{,}5 \cdot 10^{-3} = 0{,}0005$	✓
$j = 5$	$0{,}5 \cdot 10^{-4} = 0{,}00005$	✗

NSD $= j = 4$

3,141 → last digit 5 we cannot belive

→ it is not right round

Total calculation error

We suppose: $Y = F(x_1; \dots;\ x_N)$ theoretical dependency

- F we replace by the numerical method

$$y = f(x_1; \dots;\ x_n)$$

- Instead of the exact values $x_i; i = 1, \dots, n$ we use their approximations $\tilde{x}_i$

$$y' = f(\tilde{x}_1; \dots;\ \tilde{x}_n)$$

- Rounding during the calculation → $\tilde{y} = \tilde{f}(\tilde{x}_1; \dots;\ \tilde{x}_n)$
- Values y' and $\tilde{y}$ are different

Total mistake of the calculation is: $Y - \tilde{y} = F(x_1; \dots;\ x_n) - \tilde{f}(\tilde{x}_1; \dots;\ \tilde{x}_n)$

Primary mistake is: $y - y' = f(x_1; \dots; x_n) - f(\tilde{x}_1; \dots; \tilde{x}_n)$

Secondary mistake is: $y' - \tilde{y} = f(\tilde{x}_1; \dots; \tilde{x}_n) - \tilde{f}(\tilde{x}_1; \dots; \tilde{x}_n)$

The estimate of the primary mistake

$f(x_1; \dots; x_N)$ is continuous different in a set $G = \{x_i: |x_i - \tilde{x}_i| \leq \propto_i, \dots i = 1 \dots n\}$

$$\Rightarrow |f(x_1; \dots; x_n) - f(\tilde{x}_1; \dots; \tilde{x}_n)| \leq \sum_{i=1}^{N} A_i \cdot \propto_i \quad (1)$$

$$A_i = sup_G \left|\frac{\partial f}{\partial x_i}(x_1; \dots; x_n)\right| \; i = 1; \dots; N$$

EX: estimate the mistake of the operation $f(x; y) = y \cdot \ln x$ if we use the numbers $\tilde{x} = 1{,}25$ a $\tilde{y} = 0{,}3125$, they are all digits significant

$$|f(x; y) - f(\tilde{x}; \tilde{y})| \leq \sum_{i=1}^{2} A_i \cdot \propto_i =$$

$\propto_1 : |x - \tilde{x}| = |x - 1{,}25| \quad \propto_1 = 0{,}5 \cdot 10^{-2}$ (2 DP)

$\propto_2 : |y - \tilde{y}| = |y - 0{,}3125| \; \propto_1 = 0{,}5 \cdot 10^{-4}$ (4 DP)

$$A_1 : \left|\frac{\partial f}{\partial x}(\tilde{x}; \tilde{y})\right| = \frac{y}{x}(\tilde{x}; \tilde{y}) = 0{,}25$$

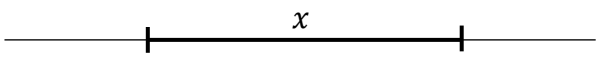
$$A_2 : \left|\frac{\partial f}{\partial y}(\tilde{x}; \tilde{y})\right| = \ln x\,(\tilde{x}; \tilde{y}) = \ln 1{,}25 \doteq 0{,}223143551$$

$$= A_1 \propto_1 + A_2 \propto_2 = 0{,}25 \cdot 0{,}5 \cdot 10^{-2} + 0{,}223143551 \cdot 0{,}5 \cdot 10^{-4} \doteq 0{,}12612 \cdot 10^{-2}$$

Mistakes of the sum, subtraction, product and division

$$x = \tilde{x} \pm \Delta x = \tilde{x} \pm \tilde{x}\delta(x) = \tilde{x}\big(1 \pm \delta(x)\big) \qquad \delta(x) = \frac{\Delta x}{\tilde{x}} \left(= \frac{\Delta x}{\tilde{x}} \cdot 100\ \%\right)$$

$\tilde{x}$ … central approximazion
Δx … absolut mistake $\varepsilon(x)$
x … right value
$\delta(x)$ … relativ mistake



$\tilde{x} - \Delta x$ Bootom approximation (estimation)

$\tilde{x} + \Delta x$ Upper aproximation (estimation)

For sum:

$x = \tilde{x} \pm \Delta x$
$y = \tilde{y} \pm \Delta y$

$x + y = (\tilde{x} + \tilde{y}) \pm (\Delta x + \Delta y)$
$0 \leq \delta(x+y) = \frac{\Delta x + \Delta y}{\tilde{x}+\tilde{y}} = \frac{\Delta x}{\tilde{x}+\tilde{y}} + \frac{\Delta y}{\tilde{x}+\tilde{y}} \leq \delta(x) + \delta(y)$ r. m. we sumarize

For substraction:

$x = \tilde{x} \pm \Delta x$
$y = \tilde{y} \pm \Delta y$

$x - y = (\tilde{x} - \tilde{y}) \pm (\Delta x + \Delta y)$
$\delta(x-y) = \frac{\Delta x + \Delta y}{\tilde{x}-\tilde{y}}$ can go to ∞, when $\tilde{x} - \tilde{y}$ is almost the same

r.m of the subtraction we cannot have under control

for product:

$x = \tilde{x} \pm \Delta x$
$y = \tilde{y} \pm \Delta y$

$x * y = (\tilde{x} \pm \Delta x)(\tilde{y} \pm \Delta y) = \tilde{x}\tilde{y} \pm (\Delta x * \tilde{y} + \Delta y * \tilde{x})$

$\delta(x;y) = \frac{\Delta x \tilde{y} + \Delta y \tilde{x}}{\tilde{x} * \tilde{y}} \leq \delta(x) + \delta(y)$

For division $\frac{1}{x}$:

$x = \tilde{x} \pm \Delta x$ $\tilde{x} - \Delta x \leq x \leq \tilde{x} \pm \Delta x$

$\frac{1}{x} = \frac{1}{\tilde{x} \pm \Delta x}$ $\frac{1}{\tilde{x}-\Delta x} \leq \frac{1}{x} \leq \frac{1}{\tilde{x}+\Delta x}$

$\rightarrow \left(\frac{\tilde{1}}{x}\right) = \frac{1}{2}\left(\frac{1}{\tilde{x}-\Delta x} + \frac{1}{\tilde{x}+\Delta x}\right) = \frac{\tilde{x}}{(\tilde{x}-\Delta x)(\tilde{x}+\Delta x)} = \frac{\tilde{x}}{(\tilde{x})^2} = \frac{1}{\tilde{x}}$

$\rightarrow \Delta\left(\frac{\tilde{1}}{x}\right) = \frac{1}{2}\left(\frac{1}{\tilde{x}-\Delta x} - \frac{1}{\tilde{x}+\Delta x}\right) = \frac{1}{2}\frac{\tilde{x}+\Delta x-(\tilde{x}-\Delta x)}{(\tilde{x})^2} = \frac{\Delta x}{(\tilde{x})^2}$

$\rightarrow \delta\left(\frac{\tilde{1}}{x}\right) = \frac{\Delta x}{(\tilde{x})^2} * x = \frac{\Delta x}{\tilde{x}} = \delta(x)$

For division $\frac{x}{y}$:

$\frac{x}{y} = x * \frac{2}{y} = \tilde{x} * \left(\frac{\tilde{1}}{y}\right) * \left(1 \pm \delta\left(\frac{x}{y}\right)\right) = \frac{\tilde{x}}{\tilde{y}}(1 \pm (\delta(x) + \delta(y)))$

EX: the estimation of the subtraction

Estimate the mistake of the subtraction $\widetilde{x_1} - \widetilde{x_2}$, if $\widetilde{x_1} = 97{,}132$ a $\widetilde{x_2} = 97{,}116$ and both numbers has the same number of the significant digits (5).

-> subtraction $\widetilde{x_1} - \widetilde{x_2} = 97{,}132 - 97{,}116 = 0{,}016 = 0{,}16 * 10^{-1}$ n=-1

$\varepsilon(\widetilde{x_1} - \widetilde{x_2}) = 0{,}5 * 10^{-3} + 0{,}5 * 10^{-3} = 0{,}001$

↓

3 DP

$0{,}001 \leq 0{,}5 * 10^{-1-j}$

j=0 : $0{,}001 \leq 0{,}5 * 10^{-1} = 0{,}05$

j=1 : $0{,}001 \leq 0{,}5 * 10^{-2} = 0{,}005$

j=2 : $0{,}001 \leq 0{,}5 * 10^{-3} = 0{,}0005$

j=1 -> the subtraction has only 1 significant digit

-> rel. mistakes: $\delta(\widetilde{x_1}) = \frac{\varepsilon\widetilde{x_1}}{\widetilde{x_1}} = \frac{0{,}5*10^{-3}}{97{,}132} = 5 * 10^{-6}$

$\delta(\widetilde{x_2}) = \frac{\varepsilon\widetilde{x_2}}{\widetilde{x_2}} = 5 * 10^{-6}$

$\delta(\widetilde{x_1} - \widetilde{x_2}) = \frac{0{,}001}{0{,}016} = 0{,}0625$ $\frac{0{,}0625}{5*10^{6}} = 12500$

R.M. of the subtraction is 12500 times bigger then R.M. of the numbers $\widetilde{x_1}; \widetilde{x_2}$

Well and badly conditioned tasks

-> Correct tasks = continuous

-> Uncorrect (well, badly) conditioned

-> Uncorrect task is well condicioned if the small changing of the input data corresponds a small changing of the output data

-> the number of the conditionality $C_p = \frac{relative\ mistake\ of\ the\ output\ data}{relative\ mistake\ of\ the\ input\ data}$

-> the border C_p is not given uniquely C_p-> 1 well-conditioned

C_p->100 badly conditioned

-> conditionality is not related to the solution algorithm used, it is a property of the task itself

EX: conditionality of the system of the equations

Decide how the system LR (Ax=b) $x_1 + 1{,}01x_2 = 2{,}01$

$1{,}01x_1 + 1{,}02x_2 = 2{,}03$

Is conditioned, the changing of the right sides is 0,01

$x_1 + 1{,}01x_2 = 2$ $\qquad$ $x_1 + 1{,}01x_2 = 2{,}01$

$1{,}01x_1 + 1{,}02x_2 = 2{,}02$ $\qquad$ $1{,}01x_1 + 1{,}02x_2 = 2{,}03$

$x_1 = 1$ $\qquad$ $x_1 = 2$

$x_2 = 1$ $\qquad$ $x_2 = 0$

VSTUP $\quad b = \begin{pmatrix} 2{,}01 \\ 2{,}03 \end{pmatrix} \quad \tilde{b} = \begin{pmatrix} 2 \\ 2{,}02 \end{pmatrix}$

VÝSTUP $\quad x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \tilde{x} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

$$C_p = \frac{\left|\frac{x-\tilde{x}}{\tilde{x}}\right|}{\left|\frac{b-\tilde{b}}{\tilde{b}}\right|} = \frac{\left|\frac{\binom{1}{1}-\binom{2}{0}}{\binom{2}{0}}\right|}{\left|\frac{\binom{2{,}01}{2{,}03}-\binom{2}{2{,}02}}{\binom{2}{2{,}02}}\right|} = \frac{\left|\frac{\binom{-1}{1}}{\binom{2}{0}}\right|}{\left|\frac{\binom{0{,}01}{0{,}01}}{\binom{2}{2{,}02}}\right|} = \frac{\left|\binom{-1}{1}\right|}{\left|\binom{1}{1}\right|} * \frac{\left|\binom{2}{2{,}02}\right|}{\left|\binom{0{,}01}{0{,}01}\right|} = \frac{\sqrt{1+1}}{\sqrt{1+1}} * \frac{\sqrt{4+2{,}02^2}}{\sqrt{0{,}01^2+0{,}01^2}}$$

$= 201$ badly conditioned

Stable algorithm

- little sensitive to input data change - well conditioned algorithm

- little sensitive to the effect of rounding errors - stable algorithm