

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Karol Tarčák

Adaptivní metody pro řešení obyčejných diferenciálních rovnic

Katedra numerické matematiky

Vedúci bakalárskej práce: Prof. RNDr. Miloslav Feistauer,
DrSc., dr. h. c.

Studijní program: Numerická a výpočtová matematika

2007

Děkuji vedoucímu práce Prof. RNDr. Miloslavovi Feistauerovi, DrSc., dr. h. c. za cenné rady a návrhy na vylepšení práce a taky PaedDr. Zdenke Majcherovej, Mgr. Tatiane Tarčákovéj za cenní připomínky.

Prohlašuji, že jsem svou bakalářskou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 10.08.2007

Karol Tarčák
Tarčák

Obsah

Úvod	5
1 Numerické metódy riešenia ODR	6
1.1 Jednokrokové metódy	7
1.1.1 Odhad chyby	8
1.1.2 Konštrukcia niektorých jednokrokových metód . .	11
1.2 Viackrokové metódy	14
2 Adaptívne metódy riešenia ODR	17
2.1 Metóda cG(1)	17
2.2 Aposteriori odhad chyby cG(1)	18
2.2.1 Adaptívna kontrola časového kroku	21
2.3 Metóda dG(0)	22
2.4 Aposteriori odhad chyby dG(0)	22
2.4.1 Stiff úlohy (rovnice so silným tlmením)	23
2.5 Metóda s adaptívnou voľbou časového kroku	25
3 Deterministické dynamické systémy	26
3.1 Lorenzov systém	26
3.2 Ostatné systémy	28
Záver	31

Názov práce:

Adaptivní metody pro řešení obyčejných diferenciálních rovnic

Autor: Karol Tarčák

Katedra (ústav): Katedra numerické matematiky

Vedoucí bakalárskej práce:

Prof. RNDr. Miloslav Feistauer, DrSc., dr. h. c.

E-mail vedúceho: miloslav.feistauer@mff.cuni.cz

Abstrakt: Táto práca sa zaoberá konštrukciou aposteriori odhadov chýb u numerických metód pre riešenie obyčajných diferenciálnych rovnic. Popisuje základné metódy riešenia obyčajných diferenciálnych rovnic a príslušné odhady. Pre spojitú a nespojitú Galerkinovú metódu sa určia aposteriori odhady u lineárneho a všeobecného systému obyčajných diferenciálnych rovnic. Získané odhady umožňujú navrhnúť adaptívnu kontrolu časového kroku. Zároveň je skonštruovaná adaptívna metóda, ktorá je následne použitá na známy Lorenzov chaotický systém.

Kľúčové slová: adaptívna kontrola kroku, deterministický chaos, numerické riešenie, aposteriori odhad chyby

Title: Adaptive methods for ordinary differential equations

Author: Karol Tarčák

Department: Department of Numerical Mathematics

Supervisor: Prof. RNDr. Miloslav Feistauer, DrSc., dr. h. c.

Supervisor's e-mail address: miloslav.feistauer@mff.cuni.cz

Abstract: In the present work we study the construction of aposteriori error estimates for numerical methods of solving ordinary differential equations. We describe basic methods of solving differential equations and corresponding error estimates. We determine aposteriori error estimates for continuous and discontinuous Galerkin method in the linear and general system of ordinary differential equations. We suggest an adaptive time step control based on the aposteriori error estimates. An adaptive method is created and used on the well-known Lorenz chaotic system.

Keywords: adaptive step control, deterministic chaos, numerical solution, aposteriori error estimate

Úvod

Pri množstve matematických, fyzikálnych, ale i ekonomických problémov potrebujeme poznať riešenie nejakej sústavy diferenciálnych rovníc, či už obyčajných alebo parciálnych. Pri skúmaní toku elektrického prúdu v elektrických obvodoch, pohybe telies, termodynamických dejoch a pod. sa dopracujeme k sústave diferenciálnych rovníc. Málokedy dokážeme nájsť pre danú sústavu exaktné (analytické) riešenie (i to však často potrebujeme reprezentovať numericky), preto sa uchýľujeme k hľadaniu približného numerického riešenia. Hľadáme priblíženie exaktného riešenia na konečnom počte bodov, v konečnej aritmetike. Nájdene riešenie nám zväčša postačuje. Vznikli mnohé metódy výpočtu numerického riešenia. Niektoré boli postupom vývoja prekonané, iné sa však stále používajú kvôli svojej jednoduchosti a relatívnej nenáročnosti na výpočetné prostredie. Pri výpočte numerického riešenia nás vždy zaujíma chyba, ktorej sa dopúšťame, či už z dôvodu diskretizácie na konečný počet uzlov (kde určujeme hodnotu riešenia), alebo z dôsledku obmedzenia výpočetného prostredia (zaokrúhlenie hodnôt na konečný počet desatinných miest). Apriori odhady však charakterizujú najhorší možný scenár, preto sú z praktického hľadiska zaujímavejšie aposteriori odhady, ktoré nám umožňujú adaptívne meniť dĺžku kroku.

Hlavným zdrojom informácií pre túto prácu je kniha [1], zaoberajúca sa problémom adaptívnej zmeny kroku pri začiatočných úlohách a tiež Lorenzovým systémom. Základná teória o numerických metódach riešenia obyčajných diferenciálnych rovníc je čerpaná z [2].

Práca je členená do troch kapitol. Prvá kapitola popisuje hlavné myšlienky hľadania numerického riešenia obyčajných diferenciálnych rovníc, odvodenie základných metód a príslušných odhadov. Odvodenie metód $cG(1)$, $dG(0)$ sa nachádza v druhej kapitole spolu s aposteriori odhadmi a konštrukciou adaptívnej metódy. V tretej kapitole sú popísané známe chaotické systémy (obzvlášť Lorenzov systém) a problémy spojené s ich riešením.

Kapitola 1

Numerické metódy riešenia ODR

Newtonov pohybový zákon môžeme formulovať ako diferenciálnu rovnicu

$$m_p x'' = F(x', x, t),$$

čo je vlastne sústava troch obyčajných diferenciálnych rovníc druhého radu, kde t označuje čas, $x(t)$ je poloha častice hmotnosti m_p v čase t , F vyjadruje závislosť pôsobiacej sily na čase, polohe a rýchlosti častice. Táto sústava charakterizuje tak komplexnú závislosť, že s výnimkou triviálnych prípadov ju nedokážeme vyriešiť analyticky. Vybudujeme nasledujúcu teóriu (ktorá je hlbšie popísaná v [2]), aby sme boli schopní skonštruovať numerické riešenie.

Uvažujme problém

$$(1.1) \quad \begin{aligned} y' &= f(x, y) \quad x \in [a, b], \\ y(a) &= \eta. \end{aligned}$$

Nech $y : [a, b] \rightarrow R$ je riešením (1.1). Zavedieme delenie intervalu $[a, b]$ tak, aby platilo

$$a = x_0 < x_1 < \dots < x_N \leq b.$$

Chceme vypočítať v každom bode x_i približné hodnoty y_i ($i = 0, 1, \dots, N$). Budeme predpokladať, že problém (1.1) je obyčajná diferenciálna rovnica prvého radu, keďže každú diferenciálnu rovnicu n -tého radu môžeme transformovať na systém obyčajných diferenciálnych rovníc prvého radu. Teda

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$$

je ekvivalentné s

$$\begin{aligned}y &= y_1, \\y'_1 &= y_2, \\y'_2 &= y_3, \\&\vdots \\y'_{n-1} &= y_n, \\y'_n &= f(x, y_1, \dots, y_{n-1}).\end{aligned}$$

Numerické metódy riešenia delíme na jednokrokové a viackrokové podľa toho, či y_{n+1} vypočítame z jednej predchádzajúcej hodnoty alebo z viacerých.

1.1 Jednokrokové metódy

Majme pevne daný krok $h > 0$ a začiatočnú podmienku $y(a) = \eta$. Pre uzly delenia platí:

$$x_n = a + nh, \quad n \geq 0$$

Hodnoty približného riešenia dostaneme z rekurentného vzťahu

$$\begin{aligned}y_0 &= \eta, \\y_{n+1} &= y_n + h\Phi(x_n, y_n, h),\end{aligned}$$

kde Φ je prírastková funkcia (závisí na f).

Definícia 1.1. Akumulovanú diskretizačnú chybu e_n metódy v bode x_n definujeme vzťahom

$$e_n = y_n - y(x_n).$$

Budeme sa snažiť nájsť odhad e_n v závislosti na h a tiež ukázať, že $e_n \rightarrow 0$ pre $h \rightarrow 0$ (v nejakom zmysle). Predpokladajme, že $f : [a, b] \times \mathbb{R} \rightarrow R$ je spojitá a zároveň f je c -lipschitzovská vzhľadom k y ($c > 0$), t.j.

$$\forall_{x \in [a, b]} \forall_{y_1, y_2 \in R} |f(x, y_1) - f(x, y_2)| \leq c|y_1 - y_2|.$$

Z Picardovej vety vyplýva existencia a jednoznačnosť riešenia.

Predpoklad 1.1. Predpokladajme, že prírastková funkcia $\Phi(x, y, h) : [a, b] \times \mathbb{R} \times [0, h_0] \rightarrow R$ je spojitá a L -lipschitzovská vzhľadom k y ($L > 0$)

$$\forall_{x \in [a, b]} \forall_{y_1, y_2 \in R} \forall_{h \in [0, h_0]} |\Phi(x, y_1, h) - \Phi(x, y_2, h)| \leq L|y_1 - y_2|$$

Definícia 1.2. Jednokroková metóda s prírastkovou funkciou Φ konverguje práve vtedy, keď platí výrok: nech $y : [a, b] \rightarrow R$ je riešenie problému (1.1), potom

$$\forall x \in [a, b] \quad \lim_{\substack{x \rightarrow 0_+ \\ x_n = x}} y_n = y(x),$$

kde y_n je približné riešenie v bode x_n .

Označme $e(h)$ maximálnu chybu pre daný krok $h > 0$:

$$e(h) = \max_{\substack{x_n \in [a, b] \\ x_n = a + nh}} |e_n| = \max_{\substack{x_n \in [a, b] \\ x_n = a + nh}} |y_n - y(x_n)|$$

Definícia 1.3. Jednokroková metóda konverguje práve vtedy, keď platí výrok: nech $y : [a, b] \rightarrow R$ je riešenie problému (1.1), potom

$$\lim_{h \rightarrow 0_+} e(h) = 0.$$

Jednokroková metóda spĺňajúca podmienku konvergenie podľa definície (2.3), konverguje i v zmysle definície (2.2).

Definícia 1.4. Jednokroková metóda s prírastkovou funkciou Φ na riešenie rovnice $y' = f(x, y)$ je konzistentná práve vtedy, keď platí

$$\forall x \in [a, b] \quad \forall y \in R \quad \Phi(x, y, 0) = f(x, y).$$

Veta 1.1. *Jednokroková metóda s prírastkovou funkciou Φ , ktorá spĺňa predpoklad (2.1), konverguje práve vtedy, keď je konzistentná.*

1.1.1 Odhad chyby

Definícia 1.5. Lokálnu relatívnu diskretizačnú chybu δ v bode x definujeme ako výraz:

$$\delta(x) = \Delta(x, y(x), h) - \Phi(x, y(x), h),$$

kde $\Delta(x, y(x), h)$ je (presný) relatívny prírastok a platí

$$\Delta(x, y(x), h) = \frac{1}{h}(y(x+h) - y(x)), \quad x, x+h \in [a, b]$$

Lemma 1.1. *Nech $A, B \geq 0$, $M \in \mathbb{N}$ a platí*

$$|\xi_{n+1}| \leq A|\xi_n| + B, \quad n = 0, \dots, M-1.$$

Potom

$$|\xi_n| \leq A|\xi_0| + \begin{cases} \frac{A^n - 1}{A - 1} B & \text{pre } A \neq 1, \\ Bn & \text{pre } A = 1. \end{cases}$$

Dôkaz:

$$\begin{aligned} |\xi_n| &\leq A|\xi_{n-1}| + B \leq A^2|\xi_{n-2}| + B(1+A) \leq \\ &\leq \dots \leq A^n|\xi_0| + B \sum_{l=0}^{n-1} \end{aligned}$$

□

Položme $A = 1 + \delta$, $\delta > 0$. Keďže platí

$$1 + \delta < e^\delta,$$

potom podľa lemmy (2.1) dostávame

$$|\xi_n| \leq e^{n\delta}|\xi_0| + \frac{B}{\delta}(e^{n\delta} - 1).$$

Definícia 1.6. Pre $L > 0$, $x \in R$ definujeme Lipschitzovu funkciu nasledovne

$$E_L(x) = \frac{e^{Lx} - 1}{L}$$

Veta 1.2. Nech $y : [a, b] \rightarrow R$ je presné riešenie úlohy (1.1), y_n (pre $x_n \in [a, b]$) sú hodnoty približného riešenia získané pomocou jednokrokovvej metódy s prírastkovou funkciou Φ , ktorá spĺňa predpoklad (2.1). Nech existujú konštanty $N, p > 0$ tak, že platí

$$(1.2) \quad |\Delta(x, y(x), h) - \Phi(x, y(x), h)| \leq Nh^p, \quad x, x+h \in [a, b].$$

Potom pre akumulovanú diskretizačnú chybu $e_n = y_n - y(x_n)$ platí odhad

$$|e_n| \leq Nh^p E_L(x_n - a), \quad x_n \in [a, b].$$

Dôkaz: Máme vzťahy

$$\begin{aligned} e_0 &= y_0 - y(x_0) = 0, \\ y_{n+1} &= y_n + h\Phi(x_n, y_n, h), \\ y(x_{n+1}) &= y(x_n) + h\Phi(x_n, y(x_n), h) + h\delta_n, \end{aligned}$$

kde δ_n je lokálna relatívna diskretizačná chyba v bode x_n . Po odčítaní dostávame

$$e_{n+1} = e_n + h(\Phi(x_n, y_n, h) - \Phi(x_n, y(x_n), h)) - h\delta_n.$$

Použijeme predpoklad (1.2) a vlastnosť funkcie Φ (predpoklad (2.1)):

$$(1.3) \quad |e_{n+1}| \leq (1 + hL)|e_n| + Nh^{p+1}, \quad x_n, x_{n+1} \in [a, b]$$

Na nerovnosť (1.3) aplikujeme lemmu (2.1), čím dokážeme odhad

$$\begin{aligned} |e_n| &\leq \frac{(1 + hL)^n - 1}{hL} Nh^{p+1} \leq \\ &\leq \frac{e^{nhL} - 1}{L} Nh^p = Nh^p E_L(x_n - a). \end{aligned}$$

□

Odhad chyby metódou polovičného kroku

Pre dostatočne hladké f a Φ plynú z vety (2.2), existencia funkcie $e : [a, b] \rightarrow \mathbb{R}$ tak, že platí

$$|e_n| = h^p e(x_n) + O(h^{p+1}).$$

Na úlohu (1.1) použijeme postupne metódu s krokom $2h$ a h . Vypočítame uzly a hodnoty približného riešenia:

$$\begin{aligned}x_n^{(2h)} &= a + n2h, & y_n^{(2h)}, \\x_{2n}^{(h)} &= a + nh, & y_{2n}^{(h)}\end{aligned}$$

V bode $x_n^{(2h)} = x_{2n}^{(h)}$ porovnáme numerické riešenia:

$$\begin{aligned}e_n^{(2h)} &= y_n^{(2h)} - y(x_n^{2h}) = (2h)^p e(x_n^{2h}) + O(h^{p+1}), \\e_{2n}^{(h)} &= y_{2n}^{(h)} - y(x_{2n}^h) = h^p e(x_{2n}^h) + O(h^{p+1}).\end{aligned}$$

Rovnice odčítame:

$$\begin{aligned}y_n^{(2h)} - y_{2n}^{(h)} &= (2^p - 1)h^p e(x_n^{(2h)}) + O(h^{p+1}) \\h^p e(x_n^{(2h)}) &= \frac{y_n^{(2h)} - y_{2n}^{(h)}}{2^p - 1} + O(h^{p+1}) \\e_{2n}^{(h)} &= \frac{y_n^{(2h)} - y_{2n}^{(h)}}{2^p - 1} + O(h^{p+1}) \sim \frac{y_n^{(2h)} - y_{2n}^{(h)}}{2^p - 1}.\end{aligned}$$

Odhad chyby zaokrúhlenia

Doteraz sme uvažovali, že výpočet numerického riešenia prebieha v nekonečnej aritmetike, avšak v praxi vždy počítame v konečnej aritmetike, t.j. každému číslu $d \in \mathbb{R}$ priradíme jeho zaokrúhlenú hodnotu \hat{d} . Predpokladajme, že začiatočné hodnoty problému (1.1) sú už zaokrúhlené:

$$\begin{aligned}\hat{y}_0 &= \eta = y_0, \\ \hat{y}_{n+1} &= \hat{y}_n + h\Phi(x_n, \hat{y}_n, h) + \varepsilon_{n+1},\end{aligned}$$

kde ε_{n+1} nazývame lokálnou chybu zaokrúhlenia. Označme $r_n = \hat{y}_n - y_n$ ako akumulovanú chybu zaokrúhlenia v bode x_n .

Veta 1.3. *Nech $\Phi(x, y, h) : [a, b] \times \mathbb{R} \times [0, h_0] \rightarrow \mathbb{R}$ je L -lipschitzovská v y a nech platí*

$$|\varepsilon_h| \leq \varepsilon,$$

potom

$$|r_n| \leq \frac{\varepsilon}{h} E_L(x_n - a).$$

Dôkaz: Máme $r_0 = 0$ a

$$\begin{aligned}\hat{y}_{n+1} &= \hat{y}_n + h\Phi(x_n, \hat{y}_n, h) + \varepsilon_{n+1} \\ y_{n+1} &= y_n + h\Phi(x_n, y_n, h) \\ |r_{n+1}| &\leq |r_n + hLr_n| + |\varepsilon_{n+1}| \leq |r_n|(1 + hL) + \varepsilon\end{aligned}$$

Aplikujeme lemmu (2.1) a dostaneme odhad

$$|r_n| \leq \frac{(1 + hL)^n - 1}{hL} \varepsilon \leq \frac{\varepsilon e^{nL} - 1}{hL} = \frac{\varepsilon}{h} E_L(x_n - a).$$

□

1.1.2 Konštrukcia niektorých jednokrokových metód

Eulerova metóda

Predpokladajme, že riešenie $y : [a, b] \rightarrow \mathbb{R}$ úlohy (1.1) je triedy \mathcal{C}^2 . Aplikujeme vetu o Taylorovom rozvoji na y v bode x_{n+1}

$$y(x_{n+1}) = y(x_n) + hy'(x_n) + O(h^2),$$

z čoho dostávame

$$y'(x_n) = \frac{1}{h}(y(x_{n+1}) - y(x_n)) + O(h).$$

Prejdeme k približnému riešeniu, čím získame

$$f(x_n, y_n) = \frac{1}{h}(y_{n+1} - y_n).$$

Potom rekurentný vzťah

$$(1.4) \quad \begin{aligned}y_{n+1} &= y_n + hf(x_n, y_n), \\ y_0 &= \eta,\end{aligned}$$

definuje Eulerovu metódu pre úlohu (1.1).

Metódy typu Runge-Kutta

Odvodíme metódu p -tého radu za predpokladu, že riešenie y je triedy \mathcal{C}^{p+1} ($p \in \mathbb{N}$). Definujme diferenciálny operátor D pre $g \in \mathcal{C}^1(G)$, $G \subseteq \mathbb{R}^2$:

$$Dg = \frac{\partial g}{\partial x} + f \frac{\partial g}{\partial y}$$

a zároveň

$$\begin{aligned}D^{(k+1)} &= D(D^{(k)}g), \quad k \geq 0, g \in \mathcal{C}^{k+1}, \\ D^{(0)}g &= g.\end{aligned}$$

Z rovnice úlohy (1.1) spočítame niekoľko derivácií y :

$$\begin{aligned}
 y'(x) &= f(x, y(x)) = (D^{(0)}f)(x, y(x)), \\
 y''(x) &= \frac{d}{dx}f(x, y(x)) = \frac{\partial f}{\partial x}(x, y(x)) + y'(x)\frac{\partial f}{\partial y}(x, y(x)) = \\
 &= \frac{\partial f}{\partial x}(x, y(x)) + f\frac{\partial f}{\partial y}(x, y(x)) = (D^{(1)}f)(x, y(x)), \\
 y'''(x) &= \frac{d}{dx}\left(\frac{\partial f}{\partial x} + f\frac{\partial f}{\partial y}\right)(x, y(x)) = \frac{\partial}{\partial x}(\dots) + f\frac{\partial}{\partial y}(\dots) = \\
 &= (D^{(2)}f)(x, y(x)), \\
 &\vdots \\
 y^{(p)} &= (D^{(p-1)}f)(x, y(x)).
 \end{aligned}$$

Aplikujeme vetu o Taylorovom rozvoji na y v bode $x + h$ a upravíme do tvaru

$$\begin{aligned}
 \frac{1}{h}(y(x+h) - y(x)) &= \sum_{l=1}^p \frac{y^{(l)}(x)}{l!} h^{l-1} + \frac{y^{(p+1)}(\tilde{x})}{(p+1)!} h^p = \\
 &= \sum_{l=1}^p \frac{(D^{(l-1)}f)(x, y(x))}{l!} h^{l-1} + \frac{y^{(p+1)}(\tilde{x})}{(p+1)!} h^p.
 \end{aligned}$$

Ak definujeme prírastkovú funkciu rovnosťou

$$\Phi_p(x, y, h) = \sum_{l=1}^p \frac{D^{(l-1)}f}{l!} h^{l-1},$$

tak získame nasledujúci odhad príslušnej diskretizačnej chyby

$$\begin{aligned}
 |\delta(x)| &= |\Delta(x, y(x), h) - \Phi(x, y(x), h)| \leq \\
 &\leq \left| \frac{y^{(p+1)}}{(p+1)!} h^p \right| \leq N h^p,
 \end{aligned}$$

kde

$$N = \max_{x \in [a, b]} \left| \frac{y^{(p+1)}}{(p+1)!} \right|.$$

Funkcia Φ_p nám umožňuje získať metódu ľubovoľného radu, avšak v definícii Φ_p potrebujeme vypočítať komplikované výrazy obsahujúce parciálne derivácie funkcie f , čo je nepraktické. Preto sa budeme snažiť nájsť vhodnejšie vyjadrenú prírastkovú funkciu Φ tak, aby výsledná metóda bola radu p . Teda chceme Φ v tvare

$$\Phi(x, y, h) = \sum_{i=1}^p Pw_i k_i(x, y, h),$$

kde

$$\begin{aligned} k_1(x, y, h) &= f(x, y), \\ k_2(x, y, h) &= f(x + h\alpha_2, y + h\beta_{21}k_1(x, y, h)), \\ &\vdots \\ k_P(x, y, h) &= f\left(x + h\alpha_P, y + h\sum_{j=1}^{P-1} \beta_{Pj}k_j(x, y, h)\right) \end{aligned}$$

a $P, w_i, \alpha_i, \beta_{ij}$ sú vhodne zvolené konštanty tak, aby platilo

$$|\Phi(x, y, h) - \Phi_p(x, y, h)| = O(h^p).$$

Položme $p = 2$, odvodíme Runge-Kuttovu metódu druhého radu. Chceli by sme tiež $P = 2$:

$$\begin{aligned} \Phi(x, y, h) &= w_1 f(x, y) + w_2 f(x + h\alpha_2, y + h\beta_{21}f(x, y)), \\ \Phi_2(x, y, h) &= f(x, y) + \frac{h}{2}(Df)(x, y) = \\ &= f(x, y) + \frac{h}{2} \left(\frac{\partial f}{\partial x}(x, y) + f \frac{\partial f}{\partial y}(x, y) \right). \end{aligned}$$

Použijeme vetu o Taylorovom rozvoji (pre funkcie n premenných) na f v bode $(x + h\alpha_2, y + h\beta_{21}f(x, y))$:

$$\begin{aligned} f(x + h\alpha_2, y + h\beta_{21}f(x, y)) &= f(x, y) + \left(h\alpha_2 \frac{\partial}{\partial x} + h\beta_{21}f \frac{\partial}{\partial y} \right) f(x, y) + \\ &+ \frac{1}{2} \left(h\alpha_2 \frac{\partial}{\partial x} + h\beta_{21}f \frac{\partial}{\partial y} \right)^2 f(x, y) = \\ &= f(x, y) + h\alpha_2 \frac{\partial f}{\partial x}(x, y) + h\beta_{21}f(x, y) \frac{\partial f}{\partial y} + \\ &+ \frac{1}{2} \left(h^2 \alpha_2^2 \frac{\partial^2 f}{\partial x^2} + 2h^2 \alpha_2 \beta_{21} f(x, y) \frac{\partial^2 f}{\partial x \partial y} + h^2 \beta_{21}^2 f^2(x, y) \frac{\partial^2 f}{\partial y^2} \right) \cdot \\ &\cdot (x + h\alpha_2, y + h\beta_{21}f(x, y)). \end{aligned}$$

Keďže posledný sčítanec je triedy $O(h^2)$, môžeme vyjadriť Φ nasledovne

$$\Phi(x, y, h) = (w_1 + w_2)f(x, y) + hw_2\alpha_2 \frac{\partial f}{\partial x}(x, y) + hw_2\beta_{21}f(x, y) \frac{\partial f}{\partial y} + O(h^2).$$

Keď porovnáme získane koeficienty pri $f, \frac{\partial f}{\partial x}, f \frac{\partial f}{\partial y}$ vo vyjadreniach Φ a Φ_2 , získame tri rovnice so štyrmi neznámymi (jednu neznámu si teda môžeme voliť). Rovnice

$$\begin{aligned} w_1 + w_2 &= 1, \\ w_2\alpha_2 &= \frac{1}{2}, \\ w_2\beta_{21} &= \frac{1}{2} \end{aligned}$$

popisujú celú triedu metód druhého radu. Pre presnosť takto získanej metódy s prírastkovou funkciou platí odhad

$$(1.5) \quad |\Delta(x, y(x), h) - \Phi(x, y(x), h)| \leq Nh^p, \quad N > 0.$$

V praxi vyberáme metódu, buď aby boli koeficienty najjednoduchšie, alebo aby konštanta N v odhade presnosti metódy (1.5) bola najmenšia. Položme napr. $w_2 = \frac{1}{2}$, získame jednu z Runge-Kuttových metód druhého radu pre úlohu (1.1):

$$(1.6) \quad \begin{aligned} y_{n+1} &= y_n + \frac{h}{2}(f(x_n, y_n) + f(x_n + hf(x_n, y_n))), \\ y_0 &= \eta. \end{aligned}$$

1.2 Viackrokové metódy

Pod pojmom (všeobecná) k -kroková metóda chápeme rekurentný predpis typu

$$(1.7) \quad \sum_{l=0}^k \alpha_l y_{n+l} = h \sum_{l=0}^k \beta_l f_{n+l},$$

kde $f_j = f(x_j, y_j)$, $\alpha_k \neq 0$ a $|\alpha_0| + |\beta_0| > 0$, hodnoty y_0, y_1, \dots, y_{k-1} nazývame začiatkové podmienky ($y_0 = \eta$ poznáme a y_1, \dots, y_{k-1} vypočítame pomocou jednokrokovej metódy, $y_i = \eta_i(h)$, $i = 1, \dots, k-1$). Položme $\beta_k = 0$, dostaneme explicitné vyjadrenie y_{n+k} :

$$\begin{aligned} \sum_{l=0}^k \alpha_l y_{n+l} &= h \sum_{l=0}^{k-1} \beta_l f_{n+l}, \\ y_{n+k} &= \frac{1}{\alpha_k} \sum_{l=0}^{k-1} (h\beta_l f_{n+l} - \alpha_l y_{n+l}). \end{aligned}$$

Takúto metódu nazývame explicitnou. Položme teraz naopak $\beta_k \neq 0$, získavame implicitnú metódu

$$(1.8) \quad y_{n+k} = h \frac{\beta_k}{\alpha_k} f(x_{n+k}, y_{n+k}) + \frac{1}{\alpha_k} \sum_{l=0}^{k-1} (h\beta_l f_{n+l} - \alpha_l y_{n+l}).$$

Rovnica (1.8) je v tvare $y_{n+k} = F(y_{n+k})$, y_{n+k} počítame po iteráciách, t.j. zvolíme začiatkové priblíženie y_{n+k}^0 (napr. y_{n+k-1} alebo $2y_{n+k-1} - y_{n+k-2}$) a pre $s \geq 0$ počítame $y_{n+k}^{s+1} = F(y_{n+k}^s)$. Ak f je L -lipschitzovská v y , tak pre dostatočne malé h je F kontrakciou a podľa Banachovej vety o kontrakcii platí, že

$$\lim_{s \rightarrow \infty} y_{n+k}^s = y_{n+k}$$

pre ľubovoľné začiatkové priblíženie y_{n+k}^0 . Výhodou implicitných metód je ich numerická stabilita pre väčšie časové kroky oproti explicitným metódam, nevýhodou je samozrejme potrebné riešiť nelineárnu rovnicu $y_{n+k} = F(y_{n+k})$.

Definícia 1.7. Viackroková metóda (1.7) je stabilná práve vtedy, keď všetky korene polynómu $\rho(\zeta) = \sum_{l=0}^k \alpha_l \zeta^l$ sú v absolútnej hodnote menšie alebo rovné 1 a všetky korene, ktorých absolútna hodnota je rovná 1, sú jednoduché.

Definícia 1.8. Viackroková metóda (1.7) je silne stabilná, ak má tvar

$$y_{n+k} - y_{n+k-1} = h \sum_{l=0}^k \beta_l f_{n+l}.$$

Definícia 1.9. Viackroková metóda (1.7) je radu $p \geq 0$, ak platí

$$\sum_{l=0}^k \alpha_l = 0$$

a navyac pre $p > 0$

$$\sum_{l=0}^k \alpha_l \frac{l^s}{s!} = \sum_{l=0}^k \beta_l \frac{l^{s-1}}{(s-1)!}, \quad s = 1, \dots, p.$$

Metóda je konzistentná, ak je radu $p \geq 1$.

Definícia 1.10. Viackroková metóda (1.7) konverguje, ak pre riešenie $y : [a, b] \rightarrow \mathbb{R}$ úlohy (1.1) pre ľubovoľné η a pre ľubovoľnú f , ktorá je spojitá a lipschitzovská v y , platí výrok: nech $\{y_n\}_{x_n \in [a, b]}$ je približné riešenie získané pomocou metódy (1.7) s takými začiatkovými podmienkami, že

$$\lim_{h \searrow 0} \eta_i(h) = \eta \quad \text{pre } i = 0, 1, \dots, k-1,$$

potom platí

$$\forall x \in [a, b] \quad \lim_{\substack{h \searrow 0 \\ x_n \in [a, b]}} y_n = y(x).$$

Nasledujúce vety sú dokázané v [2].

Veta 1.4. Každá konvergentná metóda je stabilná.

Veta 1.5. Ak viackroková metóda (1.7) konverguje, potom je konzistentná, t.j.

$$\sum_{l=0}^k \alpha_l = 0, \quad \sum_{l=0}^k \alpha_l l = \sum_{l=0}^k \beta_l.$$

Veta 1.6. Nech viackroková metóda (1.7) je radu p , nech $y : [a, b] \rightarrow \mathbb{R}$ je presné riešenie úlohy (1.1) a zároveň $y \in C^{p+1}([a, b])$. Potom existujú konštanty $h_0, k > 0$ tak, že

$$|\delta_n| \leq kh^p \quad \text{pre } \forall_n x_n, x_n + h \in [a, b], \forall h \leq h_0,$$

kde δ_n je lokálna diskretizačná chyba:

$$\delta_n = \frac{1}{h} \sum_{l=0}^k \alpha_l y(x_{n+l}) - \sum_{l=0}^k \beta_l f(x_{n+l}, y(x_{n+l})).$$

Veta 1.7. Viackroková metóda konverguje práve vtedy, keď je stabilná a konzistentná.

Veta 1.8. Nech $y \in C^{p+1}([a, b])$ je presné riešenie úlohy (1.1), $f \in C([a, b] \times \mathbb{R})$ je L -lipschitzovská ($L > 0$), $\{y_n\}_{x_n \in [a, b]}$ je približné riešenie vypočítané viackrokovou stabilnou metódou radu $p \geq 1$. Označme

$$\delta = \max_{i=0, \dots, b-1} |y(x_i - y_i)|.$$

Potom existujú konštanty $h_0, N, \theta > 0$ tak, že pre chybu metódy $e_n = y_n - y(x_n)$ platí odhad

$$|e_n| \leq e^{(x_n - a)\theta} \delta + E_\theta(x_n - a)Nh^p \quad \forall_{x_n \in [a, b]} \quad \forall_{h \in (0, h_0)}.$$

Metóda prediktor-korektor

Táto metóda zlučuje explicitnú a implicitnú metódu, čím získame výhody oboch metód a znížime ich nedostatky. Pomocou explicitnej metódy (je vhodné, aby bola výpočetne nenáročná) určíme počiatočné priblíženie y_{n+1}^0 :

$$(1.9) \quad y_{n+k}^0 = \frac{1}{\alpha_k^{(P)}} \sum_{l=0}^{k-1} (h\beta_l^{(P)} f_{n+l} - \alpha_l^{(P)} y_{n+l}).$$

Implicitnou metódou získame y_{n+1}^s :

$$(1.10) \quad y_{n+k}^s = h \frac{\beta_k^{(K)}}{\alpha_k^{(K)}} f(x_{n+k}, y_{n+k}^{s-1}) + \frac{1}{\alpha_k^{(K)}} \sum_{l=0}^{k-1} (h\beta_l^{(K)} f_{n+l} - \alpha_l^{(K)} y_{n+l}).$$

Postačuje zväčša $s \in 1, 2, 3$. Teda po s iteráciách položíme $y_{n+k} = y_{n+k}^s$. Vzťah (1.9) nazývame prediktorová formula a (1.10) nazývame korektorová formula.

Kapitola 2

Adaptívne metódy riešenia ODR

Nasledujúci výklad je založený na knihe [1].

2.1 Metóda cG(1)

Majme začiatočnú úlohu

$$(2.1) \quad \begin{aligned} \dot{u}(t) &= f(u(t)), \quad t \in (0, T], \\ u(0) &= u^0, \end{aligned}$$

kde $f \in C^0$, $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Zostrojíme spojitú Galerkinovu metódu po častiach polynomiálnu stupňa q pre úlohu (2.1). Keďže metóda musí splniť začiatočnú podmienku, zvolíme ako testovací priestor polynómy stupňa $q-1$. Prenásobíme (2.1) testovacou funkciou $v \in P_{q-1}$, kde $q \in \mathbb{N}$:

$$(2.2) \quad u(t)v(t) = f(u(t))v(t).$$

Stupeň q je rovný 1, teda $v \in \mathbb{R}^d$. Integrujeme (2.2) na intervale $(t_{n-1}, t_n) \subset [0, T]$:

$$\int_{t_{n-1}}^{t_n} : \quad u(t_n) - u(t_{n-1}) = \int_{t_{n-1}}^{t_n} f(u(t)) dt, \quad n = 1, 2, 3, \dots$$

Prejdeme k približnému riešeniu $U \in C^0([0, T])$ po častiach lineárnemu ($q = 1$):

$$(2.3) \quad U(t_n) - U(t_{n-1}) = \int_{t_{n-1}}^{t_n} f(U(t)) dt, \quad n = 1, 2, 3, \dots$$

Použijeme kvadratúrny vzorec [3] (4.23):

$$(2.4) \quad U(t_n) - U(t_{n-1}) \approx k_n f \left(\frac{U(t_{n-1}) + U(t_n)}{2} \right), \quad n = 1, 2, 3, \dots,$$

kde $k_n = t_n - t_{n-1}$ je dĺžka kroku. Člen $f\left(\frac{U(t_{n-1})+U(t_n)}{2}\right)$ aproximujeme $f(U(t_{n-1}))$:

$$U(t_n) - U(t_{n-1}) = k_n f(U(t_{n-1})), \quad n = 1, 2, 3, \dots,$$

čím dostávame hľadajú metódu (explicitná, dopredná Eulerova metóda). Zo vzťahu (2.2) a (2.3) vyplýva tzv. Galerkinova ortogonalita:

$$\int_{t_{n-1}}^{t_n} (\dot{U}(t) - f(U(t)))v dt = 0, \quad v \in \mathbb{R}^d.$$

2.2 Aposteriori odhad chyby cG(1)

Lineárny prípad

Budeme sa zaoberať prípadom, kde α je skalár a $f(t) \in C^0([0, T])$ daná funkcia:

$$(2.5) \quad \begin{aligned} \dot{u}(t) &= \alpha u(t) + f(t), \quad t \in (0, T], \\ u(0) &= u^0. \end{aligned}$$

Predpokladajme, že $u \in C^1([0, T])$. Vynásobme rovnicu (2.5) ľubovoľnou funkciou $v \in C^1([0, T])$ a integrujeme, dostaneme vzťah

$$\int_0^T (\dot{u} - \alpha u)v dx = \int_0^T f v dt$$

Použijeme metódu per partes, získame identitu

$$u(T)v(T) - u(0)v(0) + \int_0^T u(t)(-\dot{v} - \alpha v) dt = \int_0^T f v dt.$$

Uvažujme tzv. duálny problém:

$$\begin{aligned} -\dot{\varphi} &= \alpha \varphi, \quad t \in (0, T], \\ -\varphi(T) &= e(T), \end{aligned}$$

kde $e(t) = u(t) - U(t)$. Riešenie duálneho problému môžeme vyjadriť explicitne:

$$\varphi(t) = e(T) \exp(\alpha(T - t))$$

Ak zvolíme testovaciu funkciu v ako riešenie duálneho problému, dostaneme

$$u(T)v(T) = u(0)v(0) + \int_0^T f v dt.$$

Teda získame informáciu o hodnote u v čase T s pomocou počiatočnej hodnoty $u(0)$ a funkcie f .

Predpokladajme, že $e(0) = u(0) - U(0) = 0$. Potom z rovností

$$\begin{aligned} e^2(T) &= e^2(T) + \int_0^T e(t)(-\dot{\varphi} - \alpha\varphi) dt = \\ &= e^2(T) - e(T)\varphi(T) + e(0)\varphi(0) + \int_0^T \varphi(t)(\dot{e} - \alpha e) dt = \\ &= \int_0^T \varphi(t)(\dot{e} - \alpha e) dt, \\ \dot{e} - \alpha e &= \dot{u} - \dot{U} - \alpha u + \alpha U = f + \alpha U = \dot{U}, \end{aligned}$$

zo vzťahu (2.5) vyplýva identita

$$(2.6) \quad e^2(T) = \int_0^T (f(t) + \alpha U(t) - \dot{U}(t)) dt = - \int_0^T r(U)\varphi dt,$$

kde označíme rezíduum $r(U) = \dot{U} - \alpha U - f$.

Označme

$$(2.7) \quad \bar{\varphi}(t) = \frac{1}{k_n} \int_{t_{n-1}}^{t_n} \varphi(s) ds, \quad t \in [t_{n-1}, t_n].$$

Za použitia Galerkinovej ortogonalnosti metódy, t.j.

$$\int_{t_{n-1}}^{t_n} r(U) dt = 0, \quad n = 1, 2, 3, \dots,$$

prepíšeme (2.6)

$$e^2(T) = - \int_0^T r(U)(\varphi - \bar{\varphi}) dt.$$

Vieme, že pre $\tau_1, \tau_2 \in [t_{n-1}, t_n]$ platí

$$|\varphi(\tau_1) - \varphi(\tau_2)| \leq \int_{\tau_1}^{\tau_2} |\dot{\varphi}| d\tau$$

a

$$\varphi(t) - \bar{\varphi}(t) = \frac{1}{k_n} \int_{t_{n-1}}^{t_n} (\varphi(t) - \varphi(s)) ds,$$

z čoho získame nerovnosť

$$\int_{t_{n-1}}^{t_n} |\varphi - \bar{\varphi}| dt \leq k_n \int_{t_{n-1}}^{t_n} |\dot{\varphi}| dt.$$

Ohraničíme zhora $|e^2(T)|$:

$$|e^2(T)| \leq \sum_{n=1}^N r_n \int_{t_{n-1}}^{t_n} |\varphi - \bar{\varphi}| dt \leq \sum_{n=1}^N r_n k_n \int_{t_{n-1}}^{t_n} |\dot{\varphi}| dt,$$

kde $r_n = \max_{t \in [t_{n-1}, t_n]} |r(U(t))|$. Získané odhady upravíme použitím nerovnosti

$$r_n k_n \leq \max_{1 \leq n \leq N} |r_n k_n|,$$

takže

$$(2.8) \quad |e^2(T)| \leq \max_{1 \leq n \leq N} r_n k_n \int_{t_{n-1}}^{t_n} |\dot{\varphi}| dt.$$

Definujme faktor stability $S(T)$:

$$(2.9) \quad S(T) = \frac{\int_{t_{n-1}}^{t_n} |\dot{\varphi}| dt}{|\varphi|}.$$

Keďže sme požadovali $e(T) = \varphi(T)$, môžeme odhad zapísať v tvare

$$(2.10) \quad |e(T)| \leq \max_{t \in [0, T]} |k(t)r(U(t))| S(T).$$

Odhad faktoru stability $S(T)$ v závislosti na a popisuje nasledujúcu lemmu.

Lemma 2.1. *Platí*

$$S(T) \leq \begin{cases} \exp(\alpha T) & \text{pre } \alpha > 1, \\ 1 & \text{pre } \alpha \leq 0. \end{cases}$$

Dôkaz: Máme duálny problém (2.6). Vyjadríme jeho riešenie:

$$\begin{aligned} \varphi(t) &= C e^{-\alpha t}, \\ \varphi^0 &= \varphi(T) = C e^{-\alpha T} \\ \Rightarrow C &= \varphi e^{\alpha T}. \end{aligned}$$

Derivujeme φ :

$$\dot{\varphi}(t) = (-\alpha) \varphi^0 e^{\alpha(T-t)}.$$

Môžeme vyjadriť $S(t)$:

$$\begin{aligned} S(t) &= \frac{\int_0^t |\varphi(s)| ds}{|\varphi^0|} = \frac{|\alpha| |\varphi^0| e^{\alpha T} \int_0^t e^{-\alpha s} ds}{|\varphi^0|} = \\ &= (-\operatorname{sgn} \alpha) e^{\alpha T} (e^{-\alpha t} - 1). \end{aligned}$$

Najprv položíme $\alpha > 0$:

$$S(T) = e^{\alpha T} - 1 \leq e^{\alpha T}.$$

Naopak nech $\alpha \leq 0$, potom $e^{\alpha T} \in (0, 1]$ a teda:

$$S(T) = 1 - e^{\alpha T} \leq 1.$$

□

Všeobecný prípad

Uvažujme úlohu (2.1). Približné riešenie označíme $U(t)$. Predpokladajme, že $e(0) = u(0) - U(0) = 0$. Úlohe (2.1) priradíme linearizovaný duálny problém

$$(2.11) \quad \begin{aligned} -\dot{\varphi}(t) &= A^T(t)\varphi(t), \quad t \in (0, T], \\ \varphi(T) &= e(T), \end{aligned}$$

kde

$$(2.12) \quad A(t) = \int_0^1 f'(su(t) + (1-s)U(t)) ds.$$

Platí

$$(2.13) \quad \begin{aligned} A(t)e(t) &= \int_0^1 f'(su(t) + (1-s)U(t))e(t) ds = \\ &= \int_0^1 \frac{d}{ds} f(su(t) + (1-s)U(t)) ds = f(u(t)) - f(U(t)). \end{aligned}$$

Podobne ako v lineárnom prípade identitu

$$\|e(T)\|^2 = \|e(T)\|^2 + \int_0^T e(-\dot{\varphi} - A^T\varphi) dt$$

upravíme (metóda per partes) na tvar

$$\|e(T)\|^2 = \int_0^T (\dot{e} - Ae)\varphi dt.$$

Zo zadania úlohy (2.1) plynie, že

$$\dot{e} - Ae = \dot{u} - f(u) - \dot{U} + f(U) = -\dot{U} + f(U)$$

a teda

$$\|e(T)\|^2 = - \int_0^T r(U)\varphi dt.$$

Ďalej pokračujeme analogicky ako v lineárnom prípade. Získame odhad:

$$(2.14) \quad \|e(T)\| \leq S(T) \max_{t \in [0, T]} \|k(t)r(U(t))\|.$$

2.2.1 Adaptívna kontrola časového kroku

Poznáme už odhad chyby pre cG(1):

$$\|u(T) - U(T)\| \leq S(T) \max_{t \in [0, T]} \|k(t)r(t)\|,$$

kde $r(t) = \dot{U}(t) - f(\dot{U}(t))$. Predpokladajme, že faktor stability poznáme (buď ho máme vypočítaný, alebo odhadnutý). Chceme

$$\|u(T) - U(T)\| \leq K,$$

kde $K > 0$ je daná tolerancia. Voľme veľkosť časového kroku $k_n = t_n - t_{n-1}$ tak, aby platilo:

$$k_n \approx \frac{K}{S(T) r_n}.$$

Namiesto r_n je vhodné použiť r_{n-1} .

2.3 Metóda dG(0)

Chceme pre úlohu (2.1) zostrojiť nespojitú Galerkinovu metódu po častiach konštantnú. Postupujeme podobne ako pri konštrukcii cG(1), avšak testovaciu funkciu v zvolíme z priestoru P_q , kde $q = 0$ (teda $P_q = \mathbb{R}^d$). Zároveň aproximujeme člen $f\left(\frac{U(t_{n-1})+U(t_n)}{2}\right)$ vo vzťahu (2.4) výrazom $f(U_n)$, čím dostávame žiadanú (implicitnú, spätnú Eulerovu) metódu:

$$U(t_n) - U(t_{n-1}) = k_n f(U(t_n)), \quad n = 1, 2, 3, \dots,$$

Metóda spĺňa podmienku Galerkinovej ortogonality:

$$(2.15) \quad U(t_n)v = U(t_{n-1})v + \int_{t_{n-1}}^{t_n} f(U(t))v dt, \quad \forall v \in P_q, \quad n = 1, 2, 3, \dots,$$

$$\int_{t_{n-1}}^{t_n} (\dot{U}(t) - f(U(t)))v dt = 0, \quad \forall v \in P_q, \quad n = 1, 2, 3, \dots,$$

2.4 Aposteriori odhad chyby dG(0)

K odhadu sa dopracujeme podobne ako pre cG(1). Začneme identitou:

$$\|e(T)\|^2 = \|e(T)\|^2 + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} e(-\dot{\varphi} - A^T \varphi) dt,$$

kde $A(t)$ definujeme ako v (2.12). Použijeme metódu per partes na jednotlivých intervaloch (t_{n-1}, t_n) :

$$\|e(T)\|^2 = \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (-\dot{e} - Ae)\varphi dt - \sum_{n=2}^{N-1} (U(t_n) - U(t_{n-1}))\varphi(t_{n-1}).$$

Posledný člen je dôsledkom skokov približného riešenia $U(t)$ v uzlových bodoch t_n . Keďže platí (2.13) a

$$\dot{e} - Ae = \dot{u} - f(u) - \dot{U} + f(U) = -\dot{U} + f(U) = f(U) \quad \text{na } (t_{n-1}, t_n) \quad \forall 1 \leq n \leq N,$$

tak dostávame

$$\|e(T)\|^2 = \int_0^T f(U(t))\varphi dt - \sum_{n=2}^{N-1} (U(t_n) - U(t_{n-1}))\varphi(t_{n-1}).$$

Položme $v = \bar{\varphi}$, kde $\bar{\varphi}$ definujeme ako v (2.7). Potom z (2.15) získame

$$\|e(T)\|^2 = \sum_{n=2}^{N-1} \int_{t_{n-1}}^{t_n} (\varphi - \bar{\varphi}) dt - \sum_{n=2}^{N-1} (U(t_n) - U(t_{n-1}))(\varphi(t_{n-1}) - \bar{\varphi}(t_{n-1})).$$

Vieme, že $f(U(t))$ je konštantná na (t_{n-1}, t_n) :

$$\int_{t_{n-1}}^{t_n} f(U(t))(\varphi - \bar{\varphi}) dt = 0 \quad \forall 1 \leq n \leq N.$$

Teda

$$\|e(T)\|^2 = - \sum_{n=2}^{N-1} (U(t_n) - U(t_{n-1}))(\varphi(t_{n-1}) - \bar{\varphi}(t_{n-1})).$$

Platí

$$\|\varphi(t_{n-1}) - \bar{\varphi}(t_{n-1})\| \leq \int_{t_{n-1}}^{t_n} \|\dot{\varphi}(t)\| dt \quad \forall 1 \leq n \leq N,$$

čiže dostávame odhad

$$\|e(T)\| \leq S(T) \max_{1 \leq n \leq N} \|U(t_n) - U(t_{n-1})\|.$$

2.4.1 Stiff úlohy (rovnice so silným tlmením)

Pod pojmom stiff úlohy chápeme začiatkové úlohy (1.1), kde veľkosť faktoru stability $S(T)$ je nevelká i pre veľké T , ale norma linearizovaného operátora $f'(u(t))$ je veľká. Prototyp stiff úlohy má tvar

$$(2.16) \quad \begin{aligned} \dot{u} &= f(u) \equiv -Au, \quad t > 0, \\ u(0) &= u^0 = (u_i^0)_{i=1}^d, \end{aligned}$$

kde $u : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $A \in \mathbb{R}^{d \times d}$ je konštantná symetrická pozitívne semidefinitná matica, ktorej vlastné čísla λ_i sú nezáporné. Predpokladajme, že A je diagonálna. Potom vieme určiť riešenie $u(t) = (u_i(t))_{i=1}^d$:

$$(2.17) \quad u_i(t) = \exp(-\lambda_i t) u_i^0, \quad t > 0,$$

Pre veľké kladné vlastné číslo λ_i i -tá zložka riešenia rýchlo klesá k 0, pre malé vlastné číslo λ_i zostáva nejaký čas relatívne nemenná, avšak nakoniec tiež poklesne k 0. V prípade, že λ_i by bolo záporné, zložka riešenia u_i by rástla exponenciálne.

Použijeme doprednú Eulerovu metódu na úlohu (2.16):

$$U^n = U^{n-1} - k_n A U^{n-1},$$

kde $U^n = U(t_n)$ je hodnota približného riešenia v bode t_n a $k_n = t_n - t_{n-1}$ je veľkosť časového kroku. Keďže A je diagonálna ($A = \text{diag}(\lambda_i)$), platí

$$U_i^n = (1 - k_n \lambda_i) U_i^{n-1}.$$

Pre veľké λ_i musí byť časový krok menší než $\frac{2}{\lambda_i}$, aby platilo

$$|1 - k_n \lambda_i| \leq 1,$$

inak by približné riešenie rástlo exponenciálne, čo nezodpovedá exaktnému riešeniu (2.17), ktoré pre rastúci čas konverguje k 0. Teda metóda zlyháva, keďže nemôžeme zmenšovať kroky neustále (zaokrúhľovacie chyby).

Na úlohu (2.16) použijeme spätnú Eulerovu metódu:

$$U^n = U^{n-1} - k_n A U^n,$$

prepíšeme po zložkách:

$$U_i^n = (1 - k_n \lambda_i)^{-1} U_i^{n-1}.$$

Metóda je stabilná bez obmedzenia časového kroku, lebo pre $\lambda_i \geq 0$ platí

$$1 + k_n \lambda_i \geq 1.$$

Voľba časového kroku

Zistili sme, že na zaručenie stability potrebujeme veľmi malé kroky, čo je však neefektívne. Pokúsime sa skonštruovať metódu, ktorá bude postupovať jedným veľkým a niekoľkými menšími krokmi a zároveň bude stabilná. Uvažujme úlohu

$$\begin{aligned} \dot{u}(t) &= -\alpha u(t), \quad t > 0, \\ u(0) &= u^0, \end{aligned}$$

kde $\alpha > 0$. Aplikujeme explicitnú Eulerovu metódu tak, že vykonáme ako prvý väčší časový krok $K > 0$, aby platilo $\alpha K > 2$ a potom m menších krokov, aby platilo $\alpha k < 2$, t.j.:

$$U^n = (1 - \alpha k)^m (1 - \alpha K) U^{n-1}.$$

Aby sme dosiahli stabilitu, potrebujeme splniť nerovnosť

$$|(1 - \alpha k)^m (1 - \alpha K)| \leq 1,$$

teda pre počet menších krokov musí platiť

$$m \geq -\frac{\log(\alpha K - 1)}{\log|1 - \alpha k|}.$$

2.5 Metóda s adaptívnou voľbou časového kroku

V predchádzajúcej kapitole sme definovali metódu prediktor-korektor. V tejto kapitole sme navrhli adaptívnu kontrolu časového kroku. Použijeme ako prediktor explicitnú Eulerovú metódu s adaptívnou voľbou časového kroku. Keďže $f(U(t_{n+1}))$ môžeme aproximovať nasledovne:

$$f(U(t_{n+1})) \approx f(U(t_n)) + (U(t_{n+1}) - U(t_n)) \frac{\partial f}{\partial U}(U(t_{n+1})).$$

kde $\frac{\partial f}{\partial U}$ je Jakobián f , zavedieme nasledujúcu modifikáciu implicitnej Eulerovej metódy ako korektor:

$$U(t_{n+1}) = U(t_n) + k_n \left[\mathcal{I} - k_n \frac{\partial f}{\partial U}(U(t_{n+1})) \right]^{-1} f(U(t_{n+1})),$$

kde $[\dots]^{-1}$ je inverzná matica. Nevýhodou tejto metódy je nutnosť poznať Jakobián a zároveň potreba počítať riešenie sústavy lineárnych rovníc.

Kapitola 3

Deterministické dynamické systémy

Systém, ktorý sa mení s plynúcim časom (prechádza zo stavu do stavu) môžeme označiť za dynamický. Dynamické systémy nachádzame všade okolo seba. Na základe dejov, prebiehajúcich v prírode, boli odvodené mnohé systémy vo forme diferenciálnych rovníc. Častokrát sú tieto systémy nelineárne, kvôli tomu ich nedokážeme vyriešiť analyticky. Niektoré nelineárne systémy vykazujú chaotické chovanie, avšak systémy sú z princípu deterministické, týmto charakterizujeme pojem deterministický chaos. Nasledujúci výklad vychádza z knihy [1].

3.1 Lorenzov systém

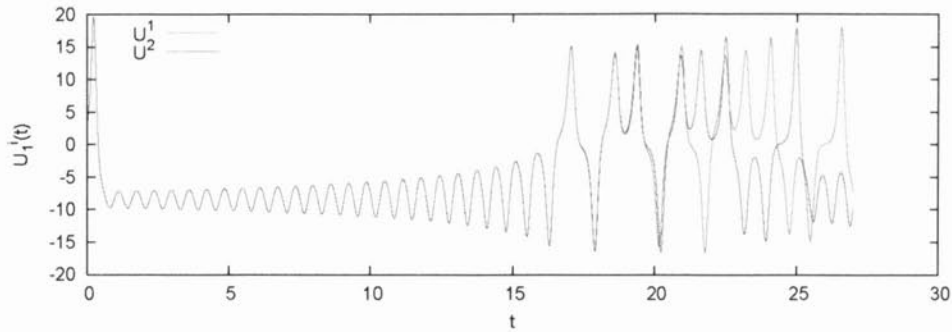
Lorenzov systém vznikol ako zjednodušenie rovníc vynúteného prúdenia v atmosfére, charakterizuje ho vzťah

$$(3.1) \quad \begin{aligned} \dot{u} &= f(u) = \begin{pmatrix} \sigma(u_2 - u_1) \\ u_1(\rho - u_3) - u_2 \\ u_1u_2 - \beta u_3 \end{pmatrix}, \\ u(0) &= u^0, \end{aligned}$$

kde σ (Prandtlovo číslo), ρ (Raykeighovo číslo), $\beta > 0$ sú konštanty (zvyčajne $\sigma = 10$, $\rho = 28$, $\beta = \frac{8}{3}$).

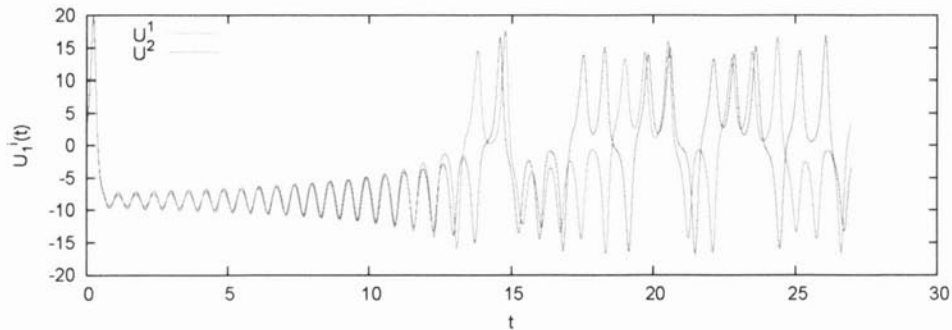
Položme začiatočné podmienky $u^{01} = (3, 5, 1)$ a $u^{02} = (3, 5.1.0001)$. Spočítame riešenia U^1, U^2 pre tieto začiatočné podmienky pomocou adaptívnej metódy z predchádzajúcej kapitoly. Obrázok 3.1 zobrazuje zložku U_1^i riešení. Všimnime si, že po určitom čase ($t = 26$) sa riešenia od seba podstatne vzdialia, túto vlastnosť dynamického systému nazývame citlivá závislosť na začiatočnú podmienku (populárny názov je efekt motýľich krídiel).

Vypočítajme riešenie pre pevne dané časové kroky $k^1 = 10^{-3}$, $k^2 = 10^{-5}$ pre rovnakú začiatočnú podmienku $u^0 = (3, 5, 1)$ pomocou dopred-



Obrázok 3.1: Zložky riešení Lorenzovho systému pre nepatrne odlišné začiatočné hodnoty

nej Eulerovej metódy. Na obrázku 3.1 vidíme, že riešenia sa počínajúc časom $t = 14$ od seba vzdialujú. Graf 3.1 noriem aproximovaných rezíduí $r(U^i)$ poukazuje, že chyba pre prvé riešenie rástla rýchlejšie. Avšak po krátkom časovom intervale sú normy oboch riešení v norme rovnako veľké. Zmenšili sme veľkosť kroku o dve desatinné miesta, ale získali sme o málo stabilnejšie riešenie.



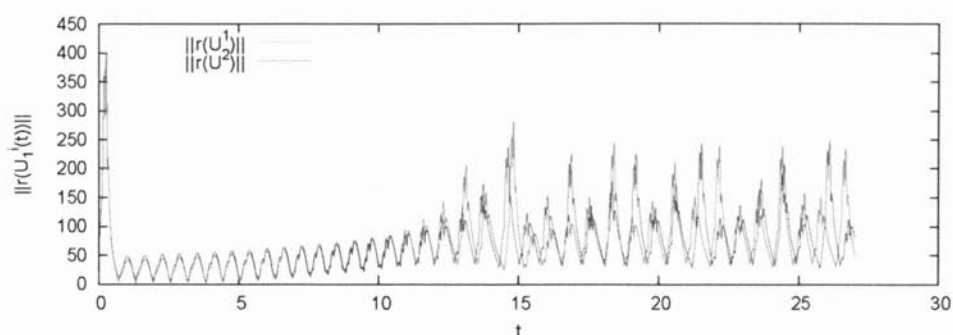
Obrázok 3.2: Zložky riešení Lorenzovho systému s rozdielnymi pevnými dĺžkami kroku

Vieme, že pre chybu doprednej Eulerovej metódy platí odhad (2.14):

$$\|e(T)\| \leq S(T) \max_{t \in [0, T]} \|k(t)r(U(t))\|.$$

Na prvý pohľad sa môže zdať, že pre dostatočne malý časový krok $k(t)$ získame stabilné riešenie do ľubovoľnej časovej hodnoty T . Avšak nesmieme zabúdať, že pracujeme v konečnej aritmetike. Teda pri každej iterácii je nevyhnutne vnesená zaokrúhľovacia chyba veľkosti ε , kde $\varepsilon > 0$ je najmenšia konštanta, pre ktorú platí $1 + \varepsilon \neq 1$ v konečnej aritmetike. Doplňme predchádzajúci odhad o ďalší člen:

$$\|e(T)\| \leq S(T) \max_{t \in [0, T]} \|k(t)r(U(t))\| + S_0(T) \max_{t \in [0, T]} \frac{\varepsilon}{k(t)},$$



Obrázok 3.3: Normy rezíduí pre rozdielne pevné dĺžky kroku

kde faktor stability $S_0(T)$ je v zmysle duálneho problému definovaný nasledovne:

$$S(T) = \frac{\int_0^T \|\varphi\| dt}{\|\varphi\|}.$$

Predpokladajme, že $\varepsilon = 10^{-16}$ (dvojnásobná presnosť v dvojkovej aritmetike s pohyblivou desatinnou čiarkou) a aproximujeme $S_0(T) \approx 10^{\frac{T}{3}}$. Potom zaokrúhľovacia chyba

$$E = \frac{10^{\frac{T}{3}-16}}{k}$$

rastie pre klesajúci časový krok k . Jednou z možností ako získať stabilné riešenie i pre dlhšie časové úseky je použiť metódy vyššieho radu (napr. $q = 11, 12, \dots$).

3.2 Ostatné systémy

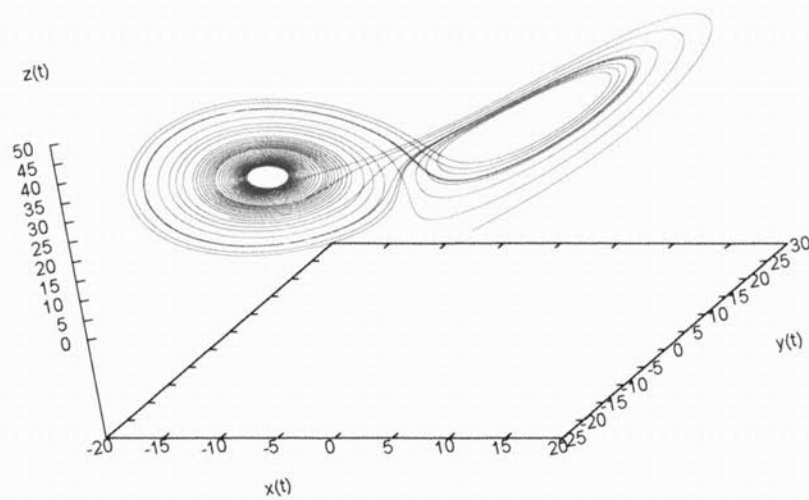
Nasledujúce dynamické systémy boli pomenované podľa objaviteľov a sú prebraté z ich práce [4], [5].

Rössler

Systém je určený vzťahom:

$$(3.2) \quad \begin{aligned} \dot{u} &= f(u) = \begin{pmatrix} -u_2 - u_3 \\ u_1 + \alpha u_2 \\ u_3(u_1 + \gamma) + \beta \end{pmatrix}, \\ u(0) &= u^0, \end{aligned}$$

kde konštanty obvykle $\alpha = 0.2, \beta = 0.2, \gamma = 5.7$.



Obrázok 3.4: Lorenzov systém

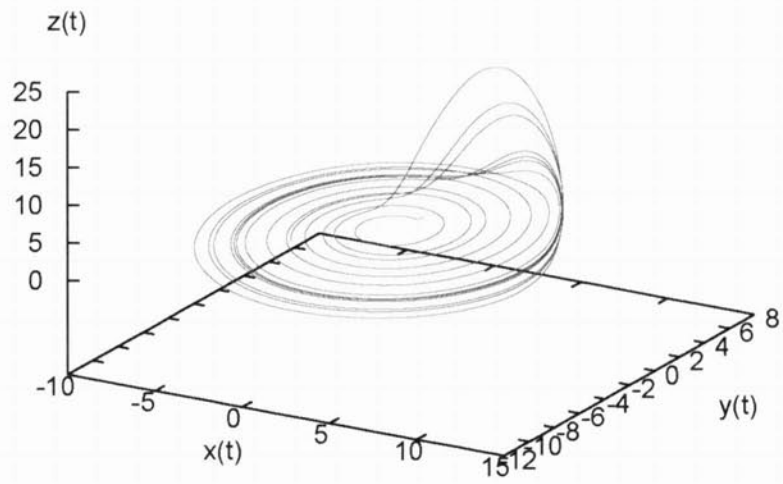
Rabinovich-Fabrikant

Systém je určený vztahom:

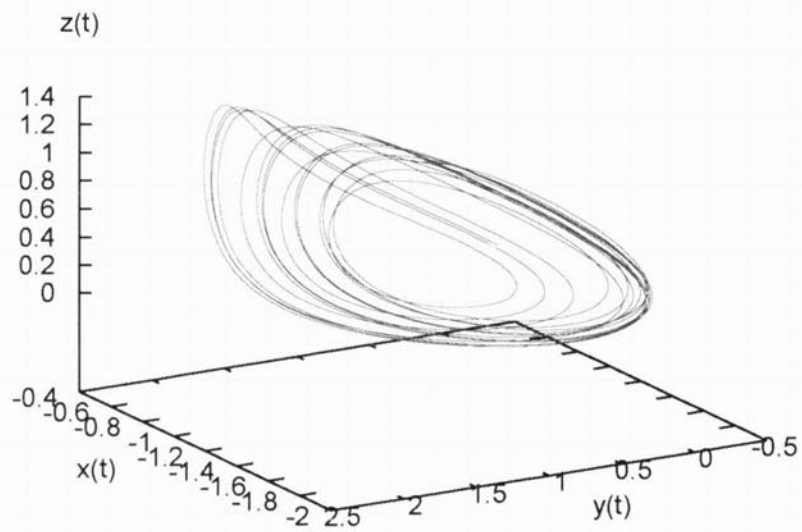
$$(3.3) \quad \dot{u} = f(u) = \begin{pmatrix} u_2(u_3 - 1 + u_1^2) + \gamma u_1 \\ u_1(3u_3 + 1 - u_1^2) + \gamma u_2 \\ -2u_3(\alpha + u_1 u_2) \end{pmatrix},$$

$$u(0) = u^0 = (-1.05, 0.9, 1.01),$$

kde konštanty obvykle $\alpha = 1.1, \gamma = 0.87$.



Obrázok 3.5: Rössler



Obrázok 3.6: Rabinovich-Fabrikant

PŘIJATO K OBHAJOBĚ 10.8.04


PŘEDSEDA KOMISE PRO BSZZ
STUDIJNÍ PROGRAM MATEMATIKA

Záver

Základnou témou v tejto práci boli aposteriorne odhady a adaptívna zmena časového kroku. Ukázali sme, že adaptívna zmena časového kroku umožňuje zlepšiť a zrýchliť výpočet riešenia. Nemôžeme však zmenšovať časový krok neustále, keďže pre malý časový krok začnú prevažovať zaokrúhľovacie chyby, je to iba jeden z ďalších užitočných nástrojov. Možný spôsob zlepšenia implementovanej metódy by bolo použiť ako prediktor a korektor explicitné a implicitné metódy vyššieho radu. Podobne i apriorne odhady majú svoj význam v analýze problémov, sú silným teoretickým nástrojom, i keď (a práve preto že) zahrňujú i tie nepriaznivejšie možnosti. Lorenzov systém neustále zostáva výzvou pri zlepšovaní numerických metód.

Literatúra

- [1] Eriksson, K., Estep D., Johnson C.: *Applied Mathematics: Body and Soul, Volume 3.*, str. 831-857, Springer, 2004.
- [2] Feistauer, M.: *Diskrétní metody řešení diferenciálních rovnic*, SPN, Praha, 1981.
- [3] Segethová, J.: *Základy numerické matematiky*, Karolinum, str. 63, Praha, 2002.
- [4] Rössler, O.E.: *An Equation for Continuous Chaos*, Physics Letters Vol. 57A no 5, str. 397-398, 1976.
- [5] Rabinovich, M. I., Fabrikant, A. L.: *Stochastic Self-Modulation of Waves in Nonequilibrium Media*, Sov. Phys. JETP 50, str. 311-317, 1979.