

Vizsgálatok a kezdeti érték problémák
numerikus megoldásával kapcsolatban

Vicsek Tamásné /Strehó Mária/

Budapest

1980

TARTALOMJEGYZÉK

	Oldal
BEVEZETÉS	5
I. A CAUCHY-FELADAT NUMERIKUS MEGOLDÁSA	7
1. A feladat kitűzése, alapfogalmak	7
2. A közelítő módszerek konvergenciája és stabilitása	12
3. A módszerek hibabecslése	15
3.1. Az egylépéses módszerek képlethibájának becslése	18
3.2. A k-lépéses prediktor-korrektor módszerek képlethibájának becslése	19
4. Döntési kritérium, lokális hibabecslési elv	20
5. Lépéshosszválasztás	22
6. Effektivitás, a módszerek alkalmazása	27
II. Módszerek összehasonlító vizsgálata	30
1. Általános szempontok	30
2. Kisérleti vizsgálatok	32
3. Elméleti összehasonlítási vizsgálat	38
III. ÁLTALÁNOSÍTOTT EGYLÉPÉSES MÓDSZEREK KONVERGENCIÁJA ÉS STABILITÁSA	46
1. Bevezetés	46
2. Általánosított egylépéses módszerek	48
3. Az általánosított egylépéses módszerek	52
1. MELLÉKLET	57
IRODALOMJEGYZÉK	58

Bevezetés

A környező világ számos jelenségét modellezhetjük differenciálegyenletek /röviden DE-k/ segítségével. Egy adott modellel vizsgálatához szükségünk van általában a DE megoldására. Ismeretes azonban, hogy az egyenletek nagy része nem oldható meg explicit alakban /azaz nem írható fel véges számú elemi művelet és elemi /esetleg speciális/ függvény segítségével. A reálisan megoldható feladatok osztálya a közelítő módszerek fejlődésével és az elektronikus számítógépek alkalmazásával bővült ki jelentősen.

A dolgozatban a közönséges DE-kre vonatkozó Cauchy-féle kezdeti értékprobléma közelítő módszereivel kapcsolatban végeztünk vizsgálatokat.

A dolgozat első fejezetében összefoglaljuk a DE közelítő megoldására vonatkozó klasszikus eredményeket /Dahlquist /1956/, Henrici /1962//, kibővitve azokat az irodalomban megjelent legfrissebb anyaggal, amely a változó lépéshosszú módszereket vizsgálja.

A II. fejezetben az egyes közelítő módszerek gyakorlati alkalmazásának kritikai leírását adjuk. Az irodalomban ismert eredmények alapján összehasonlítjuk őket különböző optimalitási szempontokból vizsgálva hatékonyságukat. A vizsgált módszerek közül számosak adaptáltuk a programját az Akadémia CDC 3300-as gépére. A programokat terjedelmük miatt nem mellékeljük, megtalálhatók a CDC 3300-as gép NUMCOSY programkönyvtárában [28].

Számítógépes tapasztalataink alapján a gyakorlatban a leghatékonyabbnak és legáltalánosabban használhatónak a lineáris többlelépéses módszerek közül a változó rendű és változó lépéshosszú formulákkal dolgozó Gear /1971/ módszer bizonyult. Ez motiválta, hogy vizsgáljuk egylépéses módszerek esetén a változó formulával és változó lépéshosszal kapcsolatos elméleti problémákat. A gyakorlatban sokszor előfordulnak /pl. biokémiában, számelméletben, epidemics stb./ olyan

DE-k. amelyekben a deriváltak a megoldás korábbi értékeitől is függenek. Ilyen általános típusu un. funkcionál DE-keket fogunk foglalkozni, amelyek közé tartoznak a közönséges DE-k is. A III. fejezetben ezek közelítő megoldására definiált általánosított egy lépéses módszer család konvergenciáját vizsgáljuk. Ez Tavernini /1972/ ekvidisztans felosztásra vonatkozó eredményének általánosítása. Definiáljuk és vizsgáljuk továbbá a módszer család stabilitását.

Végül ezen a helyen is szeretnék köszönetet mondani Móricz Ferencnek a kézirat átolvasásáért, valamint értékes megjegyzéseiért.

I. fejezet

A Cauchy - feladat numerikus megoldása

1. A feladat kitűzése, alapfogalmak

Tekintsük a közönséges differenciálegyenletekre /DE-kre/ vonatkozó

$$(1.1) \quad y' = f(x, y),$$

$$(1.2) \quad y(x_0) = y_0 \quad (x_0, y_0 \text{ adott valós konstans/}$$

kezdeti érték problémát /Cauchy-feladatot/, ahol az

$$f: I \times \mathbb{R}^m \longrightarrow \mathbb{R}^m \quad (I = [x_0, b] \subset \mathbb{R}^1)$$

leképezés folytonos és kielégíti az

$$\| f(x, y) - f(x, y^*) \| \leq L \| y - y^* \|$$

egyenletes Lipschitz feltételt minden $x \in I, y, y^* \in \mathbb{R}^m$ esetén /ahol $\mathbb{R}^m$ az m dimenziós euklideszi tér, b és L valós konstansok és $\| \cdot \|$ pl. a maximum norma/.

A továbbiakban mindig feltételezzük, hogy a fenti feltételek teljesülnek, ekkor a **Picard-Lindelöf** tétel szerint az (1.1), (1.2) DE-nek létezik egyértelmű megoldása /Henrici /1962//.

Keressük az (1.1) egyenletrendszernek az (1.2) kezdeti feltétel által meghatározott $y(x)$ partikuláris megoldását az $x_0 \leq x \leq b$ intervallumon. A feladat közelítő megoldásakor az a célunk, hogy az $\Delta_N = \{x_1, x_2, \dots, x_N\} \in I$ diszkrét ponthalmazon meghatározzuk az $y(x)$ partikuláris megoldás számértékeit.

Legyen Π az $[x_0, b]$ intervallum összes felosztásainak halmaza. Jelöljön $\Delta_N \in \Pi$ olyan felosztássorozatot, amelyre $x_N = x$ ($N = 0, 1, \dots$). Legyen $\|\Delta_N\| = \max_{0 \leq i \leq N-1} h_i$ ahol $h_i = x_{i+1} - x_i$ az i -edik lépéshossz

és $\{\Delta_N\}_{N=1}^{\infty} \subset \mathbb{T}$ mindig olyan, hogy teljesüljön a
 $\|\Delta_N\| \rightarrow 0 \quad /N \rightarrow +\infty/$ feltétel.

Az x_n pontok számos esetben ekvidisztáns elhelyezések, azaz $x_n = x_0 + nh$, $n = 1, 2, \dots, N$.

Jelölje y_n az (1.1), (1.2) feladat x_n pontbeli közelítő megoldását és $y(x_n)$ a pontos, elméleti megoldást.

Tekintsük a k-lépéses közelítő módszerek alábbi általános módszerosztályát:

$$(1.3) \begin{cases} y_r = s_r(\varphi, y_0, \{h_j\}_{j=0}^{r-1}), & (0 \leq r < k) \quad \text{kezdő értékek,} \\ \sum_{i=0}^k \alpha_i y_{n+i} / h_{n+k-1} = \phi_f(x_n, \dots, x_{n+k}; y_{n+k}, y_{n+k-1}, \dots, y_n) \\ 0 \leq n \leq N-k \end{cases}$$

ahol $\{\alpha_i\}_{i=0}^k \subset \mathbb{R}$ adott konstansok és

$\phi \in C([x_0, b]^{k+1} \times \mathbb{R}^{m(k+1)}, \mathbb{R}^m)$ adott függvény, mely az $y_n, \dots, y_{n+k}$ vektorváltozókban egyenletesen kielégíti a Lipschitz feltételt.

Ekvidisztáns esetben (1.3)-at más ekvivalens formákban is használjuk. Például ekkor a jobb oldalon

$\phi_f(x_n, y_{n+k}, y_{n+k-1}, \dots, y_n, h)$ -t írunk.

Egylépéses módszerek esetén az irodalomban elterjedtebb az

$$y_{n+1} - y_n = h_n \phi_f(x_n, y_n, y_{n+1}, h_n)$$

formula használata, ahol az előbbi feltételeknek az felel meg, hogy $\phi_f \in C([x_0, b] \times \mathbb{R}^m \times \mathbb{R}^m \times [0, h^*], \mathbb{R}^m)$ és az y_n, y_{n+1} változók szerint ϕ egyenletesen kielégíti a Lipschitz feltételt, ahol $h^* > 0$ konstans. Ha a ϕ nem függ az x_{n+k} értéktől, akkor a módszer explicit, különben implicit.

Definiáljuk a fenti módszerek /első/ karakterisztikus polinomját a következőképpen

$$(1.4) \quad g(x) = \sum_{i=0}^k \alpha_i x^i.$$

A továbbiakban az (1.3) módszereket (φ, ϕ_f) alakban is említjük majd.

Az (1.3) módszerosztály által tartalmazott ismertebb módszerek: pl.

a/ lineáris k-lépéses módszer /LMM/ esetén:

$$(1.5) \quad \phi_f = \sum_{i=0}^k \beta_i^{(n)} f(x_{n+i}, y_{n+i})$$

Ha a $\beta_i^{(n)}$ konstansok n-től függetlenek, akkor szokásos a második karakterisztikus polinomot is definiálni

$$\sigma(x) = \sum_{i=0}^k \beta_i x^i;$$

b/ prediktor-korrektor módszerek /P/EC/^mE/ /m=1/

$$(1.6) \quad \phi_f = \sum_{i=0}^{k-1} \beta_i f(x_{n+i}, y_{n+i}) + \beta_k f\left\{x_{n+k}, \frac{1}{\alpha_k^*} \sum_{i=0}^{k-1} [-\alpha_i^* y_{n+i} + h \beta_i^* f(x_{n+i}, y_{n+i})]\right\}$$

ahol α_i^*, β_i^* a prediktor formula együtthatói.

c/ Runge-Kutta módszerekre /m-pontos RK/

$$k=1$$

$$(1.7) \quad \phi_f = \sum_{r=1}^m c_r k_r$$

$$k_r = f\left\{x_n + h a_r, y_n + h \sum_{s=1}^m b_{rs} k_s\right\} \quad (1 \leq r \leq m).$$

Ha a k_r számításánál az összegzés csak $r-1$ -ig megy, a módszer explicit, különben implicit.

Az (1.5) módszerek speciális esetként tartalmazzák

a/ pl. az Adams-típusú módszereket:

$$(1.8) \quad y_{n+k} - y_{n+k-1} = h_{n+k-1} \sum_{i=0}^k \beta_i^{(n)} f(x_{n+i}, y_{n+i})$$

és

b/ az un. retrográd /BDF/ differencia formulákat:



$$(1.9) \quad \sum_{i=0}^k \alpha_i y_{n+i} = h_{n+k-1} \beta_k^{(n)} \phi(x_{n+k}, y_{n+k}).$$

Szükségünk lesz a következő fogalmakra.

1. Definíció Az $e_n = y_n - y(x_n)$ különbséget az (1.3) osztályba tartozó módszerek globális hibájának nevezzük az $x = x_n$ pontban ($0 \leq n \leq N$).

2. Definíció Az (1.3) osztályba tartozó k -lépéses módszerek $x_{n+k} \in [x_0, b]$ pontbeli képlethibájának a következőt nevezzük

$$(1.10) \quad \begin{cases} \tau(x_n, \{h_{n-j}\}_{j=1}^{k-1}) = y_n - y(x_n), & (0 \leq n < k, k > 1) \\ \tau(x_{n+k-1}, \{h_{n+k-j}\}_{j=1}^k) = \\ = \sum_{i=0}^k \alpha_i y(x_{n+i}) - h_{n+k-1} \phi_f(x_n, \dots, x_{n+k}, y(x_{n+k}), \dots, y(x_n)) \\ 0 \leq n \leq N-k \end{cases}$$

ahol $y(x)$ az (1.1), (1.2) probléma elméleti megoldása.

A $k=1$ esetben az x_{n+1} pontbeli képlethibára a $\tau(x_n, \{h_{n+1-j}\}_{j=1}^1$ jelölés helyett a hagyományos, ettől csak előjelben különböző

$$\tau(x_n, h_n) = -\tau(x_n, \{h_{n+1-j}\}_{j=1}^1)$$

definíciót használjuk és x_n pontbeli képlethibáról beszélünk.

Ekvidisztáns lineáris k -lépéses módszerek esetén ($k > 1$) pedig a $\hat{\tau}(x_n, h) = \tau(x_{n+k-1}, \{h_{n+k-j}\}_{j=1}^k)$ jelölést alkalmazzuk, ahol $h_n = h_{n+1} = \dots = h_{n+k-1} = h$.

Vizsgáljuk most az alábbi un. perturbált Cauchy problémát

$$(1.11) \quad \begin{cases} y' = f(x, y), \\ y_n(x_n) = y_n, \end{cases}$$

amelynek elméleti megoldását jelölje $y_n(x)$.

3. Definíció Az (1.3) alakú módszerek $x_{n+1} \in [x, b]$ pontbeli lokális hibájának nevezzük az

$$y_n(x_{n+1}) - y_{n+1} \quad (0 \leq n \leq N-1)$$

mennyiséget.

4. Definíció Az (1.3) osztályba tartozó közelítő módszer konvergens tetszőleges $\{\Delta_N\} \subset \Pi$ felosztássorozatra nézve, ha az (1.1), (1.2) Cauchy-problémára

$$\max_{0 \leq n < k} \|y_n - y_0\| \rightarrow 0 \quad (\|\Delta_N\| \rightarrow 0), \quad (1.12)$$

$$\max_{0 \leq n \leq N} \|y_n - y(x_n)\| \rightarrow 0 \quad (\|\Delta_N\| \rightarrow 0).$$

5. Definíció Az (1.3) osztályba tartozó közelítő módszer konzisztens a $\{\Delta_N\} \subset \Pi$ felosztássorozaton, ha

$$\max_{0 \leq n \leq N-1} \|\tau(x_n, \{h_{n-j}\}_{j=1}^k)\| \rightarrow 0 \quad (\|\Delta_N\| \rightarrow 0),$$

$$\sum_{n=0}^{k-1} \|\tau(x_n, \{h_{n-j}\}_{j=1}^{k-1})\| + \max_{0 \leq n \leq N-k} \|\tau(x_{n+k-1}, \{h_{n+k-j}\}_{j=1}^k)\| \rightarrow 0$$

$$(\|\Delta_N\| \rightarrow 0) \quad , \quad \text{ha} \quad k > 1$$

illetve

$$\max_{0 \leq n \leq N-1} \|\tau(x_n, h_n)\| \rightarrow 0 \quad (\|\Delta_N\| \rightarrow 0) \quad , \quad \text{ha} \quad k=1.$$

A konzisztencia p-ed rendű a $\{\Delta_N\} \subset \Pi$ felosztás sorozaton, ha p az a maximális pozitív szám, amelyre a fenti mennyiségek

$$O(\|\Delta_N\|^p)$$

nagyságrendűek.

Az (1.5) módszert formálisan konzisztensnek nevezzük, ha

$$\sum_{i=0}^k \alpha_i = 0, \quad \sum_{i=1}^k \alpha_i \sum_{j=1}^i h_{n+j-1} = h_{n+k-1} \sum_{i=0}^k \beta_i^{(n)}$$

teljesül $(0 \leq n \leq N-k)$.

A módszer rendjét a következőképpen definiáljuk. Jelölje $p/f/$ a módszer konzisztencia rendjét az (1.1), (1.2) problémán az euklidész-távolság felosztássorozatokra nézve, és legyen A : az $\mathbb{R}^{m+1}$ téren értelmezett analitikus függvények halmaza.

6. Definíció A módszer rendjének /globális rend/ a

$$p = \min \{ p(f) \mid f \in A \}$$

számot nevezzük.

2. A közelítő módszerek konvergenciája és stabilitása

A továbbiakban először az euklidész-távolság felosztássorozatokra kapott konvergencia és stabilitás eredményeket foglalkoztatjuk össze.

1. Tétel Ha egy (1.3) osztályba tartozó közelítő módszer konvergens, akkor konzisztens.

/Bizonyítás az LMM esetre Henrici /1962/, szélesebb osztályokra ld. Butcher /1966/ és Spijker /1966//.

Megjegyezzük, hogy a konzisztencia nem elégséges feltétele az (1.3) típusú módszerek konvergenciájának.

Tekintsük a következő perturbált kezdeti értékproblémát

$$z' = f(x, z) + \delta(x),$$

$$z(x_0) = y_0 + \delta_0, \quad x \in [x_0, b],$$

ahol $(\delta(x), \delta_0)$ a perturbáció és $z(x)$ a perturbált megoldás.

7. Definíció /Hahn /1967/, Stetter /1971//. Legyen

$(\delta(x), \delta_0)$, $(\delta^*(x), \delta_0^*)$, tetszőleges két perturbáció és legyen $z(x)$, $z^*(x)$ a kapott perturbált megoldások. Az (1.1),

(1.2) kezdeti értékproblémát totálisan stabilisnak nevezzük, ha létezik olyan pozitív S konstans, hogy minden $x \in [x_0, b]$ esetén teljesül

$$\|z(x) - z^*(x)\| \leq S\varepsilon,$$

ha csak $\|\delta(x) - \delta^*(x)\| \leq \varepsilon$ és $\|\delta_0 - \delta_0^*\| \leq \varepsilon$.

/Ha az f leképezés Lipschitz feltételnek tesz eleget valamely L konstanssal, ahogy azt feltételeztük, akkor ebből már következik, hogy a DE totálisan stabilis ld. Gear /1971//.

Vizsgáljuk most az (1.3) módszerek alábbi típusú perturbációit

$$z_r = s_r(\phi, y_0, \{h_j\}_{j=1}^{r-1}) + \delta_r, \quad (0 \leq r \leq k)$$

$$\sum_{i=0}^k \alpha_i z_{n+i}/h = \phi_j(x_n, z_{n+k}, \dots, z_n, h) + \delta_{n+k}, \quad (0 \leq n \leq N-k)$$

ahol $(\delta_n | n=0, 1, \dots, N)$ a perturbáció és $(z_n | n=0, 1, \dots, N)$ a kapott perturbált megoldás.

8. Definíció Legyen $(\delta_n | n=0, 1, \dots, N)$, $(\delta_n^* | n=0, 1, \dots, N)$ tetszőleges két perturbációsorozat és legyenek

$(z_n | n=0, 1, \dots, N)$, $(z_n^* | n=0, 1, \dots, N)$ a kapott perturbált megoldások. Ha létezik olyan h_0 és S pozitív konstans, hogy minden $h \in [0, h_0]$ és $\|\delta_n - \delta_n^*\| \leq \varepsilon$ $(0 \leq n \leq N)$ esetén

$$\|z_n - z_n^*\| \leq S\varepsilon \quad (0 \leq n \leq N),$$

akkor az (1.3) módszert D-stabilisnak /vagy zéró stabilisnak/ nevezzük /Dahlquist /1956//.

9. Definíció Az (1.3) módszer kielégíti a gyökfeltételt, ha a $g(\theta)$ un. karakterisztikus polinom gyökei az egységkörön vagy azon belül helyezkednek el, és a körön elhelyezkedő gyökök egyszeresek.

Ismertek a következő tételek:

2. Tétel Az (1.3) osztályba tartozó közelítő módszer D-stabilis akkor és csak akkor, ha kielégíti a gyökfeltételt.

3. Tétel Az (1.3) osztályba tartozó közelítő módszer akkor és csak akkor konvergens, ha konzisztens és D-stabilis.

4. Tétel Ha az (1.3) alakú módszer konzisztenciája p-ed rendű és a módszer D-stabilis, akkor a konvergencia is p-ed rendű lesz /Hall, Watt /1976//.

Megjegyezzük, hogy az (1.3) osztályba tartozó konzisztens egylépéses módszerek szükségképpen kielégítik a gyökfeltételt, minthogy a $\mathcal{P}(\Theta)$ karakterisztikus polinomnak egyetlen gyöke van a $\Theta_1 = +1$. Így igaz a

5. Tétel Az (1.3) osztályba tartozó egylépéses módszer /ahol $k=1$ / akkor és csak akkor konvergens, ha konzisztens.

Összefoglalva: a konzisztencia befolyásolja a lokális hiba nagyságát, a D-stabilitás pedig a hibaterjedést, ha $h \rightarrow 0$. Számos módszerosztályra bizonyított tény, hogy "stabilitás + konzisztencia = konvergencia".

Az LMM esetben a következő eredmény ismert.

6. Tétel A D-stabilis k-lépéses lineáris módszerek maximális rendje $k+1$, ha k páratlan és $k+2$, ha k páros. /Henrici /1962//.

A nem-ekvidisztans esetben /változó lépéshosszra/ ismert legáltalánosabb konvergencia és stabilitási tételek a következők.

Az (1.3) alakú egylépéses módszerekre tegyük fel, hogy az (1.1), (1.2) Cauchy probléma megoldásán folytonosak és kielégítik a Lipschitz feltételt az y_n és y_{n+1} vektorváltozók szerint.

7. Tétel Az (1.3) módszer családba tartozó egy lépéses módszer, melyre teljesülnek a fenti feltételek, akkor és csak akkor konvergens minden $x \in [x_0, b]$ és tetszőleges $\{\Delta_N\}_{N=1}^{\infty} \subset \Pi$ felosztássorozatra, ha konzisztens, azaz, ha $\Phi_f(x; y, y, 0) = f(x, y)$.

/Galántai /1978//.

Az (1.5) alakú változó lépéshosszú lineáris többlépéses módszerekről tegyük fel most, hogy kielégítik a gyökfeltételt; az $\alpha_k = 0$ és a $\{\beta_i^{(n)}\}_{i=0}^k$ együtthatók pedig a $h_n, \dots, h_{n+k-1}$ lépéshossz homogén függvényei. Minden esetben feltesszük, hogy $y_i \rightarrow y_0$ ($\|\Delta_N\| \rightarrow 0, 0 \leq i < k$).

8. Tétel Ha a $\{\Delta_N\}_{N=1}^{\infty} \subset \Pi$ ($x \in [x_0, b]$) felosztássorozat olyan, hogy $\sup_N \max_{n,i} |\beta_i^{(n)}| \leq C_1 < +\infty$, akkor az (1.5) formálisan konzisztens módszer konvergens.

/Galántai /1978//.

3. A módszerek hibabecslése

Itt csak rövid összefoglalást kívánunk adni a különböző típusú hibabecslő eljárásokról. Részletesebben például Hall, Watt /1976/ vagy Gear /1971/ könyvében találhatók meg.

Megjegyezzük, hogy általában p -ed rendű módszerre elég sima jobb oldalakkal rendelkező DE esetén a lokális hiba a következő alakú

$$(1.13) \quad \tau_n(x_n, h_n) = \Psi_f(x_n, h_n) h_n^{p+1} + O(h_n^{p+2})$$

ahol Ψ_f az ún. főhibatag függvény. A Ψ_f tagot szokás becsülni, lehetséges választások pl. /van der Houwen /1977//

$$\Psi(x_n, h_n) = C_n,$$

$$\Psi(x_n, h_n) = B_n x + C_n,$$

$$\Psi(x, h_n) = A_n x^2 + B_n x + C_n,$$

$$\Psi(x, h_n) = A_n x,$$

stb.

ahol C_n, B_n, A_n konstansok.

Vizsgáljuk meg röviden a skalár esetben a perturbációk hatását az (1.1), (1.2) problémára. Tegyük fel, hogy $z(x)$ kielégíti az alábbi egyenletet

$$(1.14) \quad z'(x) - f(x, z(x)) = \theta \delta(x), \quad (x \in [x_0, b]), \\ z(x_0) = y_0 + \theta \delta_0, \quad (\theta \text{ kicsi})$$

Legyen (1.15) $z(x) = y(x) + \theta e(x) + O(\theta^2)$ akkor felhasználva Taylor tételét az (1.14) egyenletből kapjuk
 $y'(x) + \theta e'(x) - f(x, y(x)) - f_y(x, y(x)) \theta e(x) = \theta \delta(x) + O(\theta^2)$
 $y(x_0) + \theta e(x_0) = y_0 + \theta \delta_0 + O(\theta^2),$
 ahol $f_y = \left\{ \frac{\partial f}{\partial y} \right\}.$

Igy az $e(x)$ függvénynek ki kell elégítenie az alábbi lineáris egyenletet:

$$(1.16) \quad e'(x) - f_y(x, y(x)) e(x) = \delta(x), \\ e(x_0) = \delta_0.$$

Azaz ha y az (1.1), (1.2) és a Z az (1.14) által definiált, valamint $e(x)$ (1.16) -ból, akkor (1.15) teljesül. Az $e(x)$ megoldás felírható zárt alakban:

$$e(x) = E(x_0, x) \delta_0 + \int_{x_0}^x E(u, x) \delta(u) du,$$

ahol

$$E(u, x) = \exp \left[\int_u^x f_y(t, y(t)) dt \right]$$

A fenti képletből könnyen belátható, hogy a $\delta(u)$ perturbáció hatása az u pontban az $E(u, x)$ -től függ, amely lehet csökkenő vagy növekvő függvény. A legegyszerűbb esetben

$y' = \lambda y$, így $f_y = \lambda$, akkor $E(u, x) = \exp(\lambda(x-u))$.
 Ha $\lambda > 0$, akkor egy u -hoz közeli pontban levő lokális

hiba hatása a globális hibára az x pontban exponenciálisan nő, ha $\lambda < 0$, akkor a fordítottja áll fenn.

Diszkrét esetben, ha ekvidisztans a lépéshossz ismert az alábbi egyszerű tétel.

9. Tétel Ha az (1.3) alakú módszer konzisztenciája p -ed rendű és a módszer D -stabilis, akkor a konvergencia is p -ed rendű lesz /Hall, Watt /1976//.

Könnyen belátható, hogy p -ed rendű módszerre

$$y_n = y(x_n) + h^p e(x_n) + o(h^{p+1}) \quad (0 \leq n \leq N)$$

ha

$$\tau_n = s_n(f, y_0) - y(x_n) = o(h^{p+1}), \quad (0 \leq n < k)$$

és

$$\tau_{n+k} = h^p \psi(x_{n+k}) + o(h^{p+1}) \quad (0 \leq n \leq N-k)$$

ahol

$$e'(x) - f_y(x, y(x))e(x) = -\psi(x), \quad e(x_0) = 0.$$

Ha az (1.3) alakú módszer D -stabilis, akkor y_n és $y(x_n) + h^p e(x_n)$ eltérése $O(h^{p+1})$ nagyságrendű.

Ez az eredmény azt jelenti, hogy ha egy módszer lokális hibájának főtagja $h^p \psi(x_n)$ az x_n pontban, akkor a globális hiba főtagja $h^p e(x_n)$ nagyságú és a pontos megoldástól való eltérés ugyanakkora mint ha $h^p \psi(x_n)$ taggal perturbáltuk volna az (1.1), (1.2) egyenletet feltéve, hogy a kezdeti feltételek $p+1$ rendig pontosak, /általánosabb megfogalmazásban lásd Stetter /1973//.



Néhány ismert módszer leírását megadjuk az 1. táblázatban, amely konkrét hibabecslési eljárásokat is tartalmaz.
/i. melléklet/

3.1 Az egylépéses módszerek képlethibájának becslése

A legismertebb hibabecslési módszerek a következő típusúak: Legyen p a módszer rendje.

a/ Extrapolációs módszer

$$\begin{aligned} \text{Legyen } x_{n+1} &= x_n + h, \\ x_{n+2} &= x_n + 2h \end{aligned}$$

$$\text{és } f \in C^{p+1}[x_0, b].$$

Jelölje továbbá $\bar{y}_{n+2}$ az x_n ponttól $2h$ lépéshosszal számított közelítő megoldást. Ekkor a képlethiba becsléseként

$$\tau(x_{n+2}, h) \approx \frac{\bar{y}_{n+2} - y_{n+2}}{2^{p+1} - 2}$$

/Henrici /1962// mennyiséget fogadjuk el, amely az $n=0$ esetben $O(h^{p+2})$ nagyságrendű becslését adja a képlethiba főtagjának.

b/ Beágyazási típusú módszerek

Minden integrálási lépést kétszer hajtunk végre, először egy p -ed rendű (1.7) alakú formulával, majd egy $p+1$ -ed rendű képlettel számolva, melyre a k_r értékek ugyanazok, csak új tagok is jönnek hozzá.

Ily módon a két közelítés különbsége segítségével becsülhető a p -ed rendű közelítés képlethibája.

A legismertebb beágyazási típusú formulákat pl. England /1967, 1969/; Shintani /1965/, Fehlberg /1968, 1969/; stb. konstruálták.

c/ Többlépéses hibabecslési módszerek

$p+1$ -ed rendű többlépéses formulával becsülhető a képlethiba az alábbi módon:

$$\tau(x_n, h) \approx \left\{ \sum_{i=-k+1}^1 \alpha_i y_{n+i} (h - \beta_i f(y_{n+i})) \right\} / \rho'(1).$$

/Ceschino, Kuntzman /1963//.

3.2 A k-lépéses prediktor-korrektor módszerek képlethibájának becslése

Milne módszere

Alkalmazható, ha a prediktor és korrektor formula azonos rendű.

A hiba alakja p -ed rendű módszerekre, feltéve, hogy az előző y_r értékek pontosak ($0 \leq r < k$) az x_{n+k} pontban prediktor $C^* h^{p+1} y^{(p+1)}(x_{n+k}) + O(h^{p+2})$,

korrektor $Ch^{p+1} y^{(p+1)}(x_{n+k}) + O(h^{p+2})$,

/C és C^* a formulák együtthatóitól függő konstansok/.

Feltesszük, hogy a kezdő értékekben a hiba h^{p+1} nagyságrendű és a lépéshossz konstans, /vagy szakaszosan konstans/, valamint a prediktor karakterisztikus polinomja ρ^* és a korrektor karakterisztikus polinomja ρ olyan, hogy $\rho^*(\theta) = 0$, ha $\|\theta\| = 1$ és $\rho(\theta) = 0$. A következő becslést kapjuk:

$$\begin{aligned} \hat{\tau}(x_n, h) &= h^p \psi(x_{n+k}) + O(h^{p+1}) = Ch^p y^{(p+1)}(x_{n+k}) + O(h^{p+1}) = \\ &= K (y_{n+k}^{(PR)} - y_{n+k}^{(C)}) / h, \end{aligned}$$

ahol

$$K = \frac{a_k^* C / \rho'(1)}{C / \rho'(1) - C^* / \rho^{*'}(1)}.$$

/lásd Henrici /1962//.

Gear módszere

Legyen $\underline{a} = [y, hy', \dots, h^p y^{(p)}/p!]^T$,

ha a korrektor p-ed rendű.

A főhibatag p-ed rendű módszere

$$C_{p+1} p! \|\nabla a_p / \omega\|_2,$$

alakú, ahol ω súlykomponens, $\|\cdot\|$ az l_2 norma és ∇a_p az $\underline{a}$ vektor: utolsó komponensének retrográd differenciája.
/lásd 1. táblázat/.

4. Döntési kritérium, lokális hibabecslési elv

Legyen az (1.1), (1.2) kezdeti érték feladat megoldására az x_n pontban ismert egy $y_n = Y(x_n)$ közelítés és az $x_{n+1} = x_n + h_{n+1}$ pontban szeretnénk kiszámítani az y_{n+1} értéket. Legyen $y_n(x)$ pontos megoldása az (1.11) Cauchy problémának és ε tetszőleges előre megadott pozitív hibakorlát. Az (1.1), (1.2) Cauchy-probléma ε pontossággal való numerikus megoldásán azt értjük, hogy a globális hiba $\|e_n\| \leq \varepsilon$. A gyakorlatban ennek biztosítására a következő két kritérium egyikét használják egy adott integrálási lépés elfogadásához.

Az un. lépésenkénti hibellenőrzés /error per step/ esetén vizsgáljuk a

$$(1.17) \quad \|y_n(x_{n+1}) - y_{n+1}\| \leq \varepsilon^* \quad (\varepsilon^* < \varepsilon)$$

feltétel teljesülését.

A másik módszer az egységlépésenkénti hibellenőrzés /error per unit step/, ekkor a

$$(1.18) \quad \|y_n(x_{n+1}) - y_{n+1}\| \leq \varepsilon^* h_n \quad (\varepsilon^* < \varepsilon)$$

feltételt kötjük ki a lépés elfogadására.

Az (1.17), (1.18) feltételek ellenőrzését lényegében a következő megfontolások alapján végezzük. Explicit egylépéses módszerek esetén a lokális hiba megegyezik a módszer

(1.11) perturbált Cauchy problémára vonatkozó képlethibájával, azaz

$$y_n(x_{n+1}) - y_{n+1} = \tau_n(x_n, h) \quad , \text{ ahol}$$

$$\tau_n(x, h) = y_n(x) + h \phi_f(x, y_n(x), h) - y_n(x+h).$$

Az $n=0$ esetet jelölje röviden $\tau(x, h)$. Implicit módszerek esetén belátható, hogy

$$y_n(x_{n+1}) - y_n = \tau_n(x_n, h) + O(h^{p+2}).$$

A $\tau_n(x_n, h)$ perturbált képlethiba /"lokális hiba"/ becslésére általában valamely az előzőekben ismertetett képlethibabecslő eljárást használunk. Ismert eredmény, hogy ha az eredeti (1.1), (1.2) problémára vonatkozó $\tau(x, h)$ képlethiba kielégíti a

$$\|\tau(x, h)\| \leq Dh^{p+1} \quad (x, x+h \in [x_0, b], \quad p > 0)$$

feltételt, akkor $\Delta_N \in \Pi$, $\|\Delta_N\| \leq h^*$ esetén létezik olyan $C > 0$ konstans, hogy

$$\|e_n\| \leq ch^p.$$

A fenti eredmény alapján alakult ki az a gyakorlat, hogy az n -edik lépésben vett képlethibabecslést vetjük össze az előírt pontossági feltételekkel. ($0 \leq n \leq N$).

Valójában azonban nem a $\tau(x_n, h_n)$ képlethibát hanem a már említett $\tau_n(x_n, h_n)$ perturbált képlethibát /lokális hibát/ becsüljük. Minthogy a $\tau_n(x_n, h_n)$ képlethiba az eredeti $y(x_n)$ megoldástól $O(h^p)$ nagyságrenddel különböző $y_n(x_n) = y_n$ kezdeti feltételhez rendelt Cauchy problémára vonatkozik, elvileg lehetséges, hogy a $O(h^{p+2})$ pontossággal becsült perturbált képlethiba ellenére a globális hiba eggyel kisebbrendű lesz, mint az elméletileg garantált. Ez azonban nincs így. Ennek a problémának az elvi megalapozása két részből áll; egyrészt a képlethiba és képlethibabecslő eljárás viszonyának vizsgálatából, másrészt az ún. lokális hibabecslési elv igazolásából. De ezekkel itt nem foglalkozunk behatóbban /ld. pl. Galántai /1978//.

10. Tétel Ha $\Delta_N \in \Pi$ ($\|\Delta_N\| \leq h^*$) olyan, hogy az (1.3) alakú egy lépéses módszer ($k=1$) hibájára

$$(1.19) \quad \|\tau_n(x_n, h_n)\| \leq h_n \varepsilon^* \quad (\|\tau_n(x_n, h_n)\| \leq \varepsilon^*)$$

fennáll, akkor létezik $c = c(\phi, \phi) > 0$ konstans [$c = c(N)$] szám, amelyre

$$(1.20) \quad \|y_n - y(x_n)\| \leq c \varepsilon^* \quad (0 \leq n \leq N).$$

/Galántai /1978//.

Ez a tétel indokolja az un. lokális hibabecslési elvet, azaz, hogy a $\|e_n\| < \varepsilon$ feltétel teljesülése helyett az (1.17) vagy (1.18) feltétel teljesülését vizsgáljuk, amely a lokális hibára vonatkozik.

5. Lépéshosszválasztás

Egy adott integrálási lépést sikeresnek fogadnak el, ha teljesül rá az (1.17) vagy (1.18) feltétel a megadott ε pontossággal. Jelölje p -ed rendű módszerre az $\|y_n(x_{n+1}) - y_{n+1}\|$ eltérésre adott becslést "est". Ha a lépés nem elfogadható pontosságú, akkor csökkentenünk kell a lépéshosszt általában úgy, hogy a feltétel teljesüljön, de a lépéshossz a lehető legnagyobb legyen. Az

$$(1.21) \quad |est| = |h_n^{p+1} \Psi(x_n, y_n)|$$

formulából látható, hogy a lépést a következő h_n' , új lépéshosszal kell megismételni:

$$h_n' = |\varepsilon / est|^{\frac{1}{p+1}} h_n$$

mert

$$|h_n'^{p+1} \Psi(x_n, y_n)| = \frac{\varepsilon}{|est|} h_n^{p+1} |\Psi(x_n, y_n)| = \varepsilon.$$

Ha a lépést elfogadtuk, akkor meg akarjuk határozni azt a legnagyobb lépéshosszt, amellyel sikeresen számolhatunk a következő lépésben.

A következő megfigyelésből

$$\begin{aligned} y_{n+1}(x_{n+2}) - y_{n+2} &= h_{n+1}^{p+1} \psi(x_{n+1}, y_{n+1}) + O(h_{n+1}^{p+2}) = \\ &= h_{n+1}^{p+1} \psi(x_n, y_n) + O(h_{n+1}^{p+2}), \end{aligned}$$

látjuk, hogy a megfelelő érték

$$h_{n+1} = |\varepsilon|^{1/(p+1)} h_n$$

Igy ugyanaz a technika használható sikeres és sikertelen lépés esetén is az új lépéshossz kiválasztására. Az ilyen módon meghatározott mennyiségeket "lokálisan optimális" lépéshossznak fogjuk nevezni. /lásd pl. Jackson /1978//.

Óvatosnak kell azonban lennünk a lokálisan optimális lépéshossz gyakorlati felhasználásakor. Az (1.21) közelítő egyenlőség nagyon pontatlan lehet, ha pl. a lokális hibabecslés pontatlan, vagy mert az (1.13) h-szerinti sorfejtésének z-nél magasabb rendű tagjai nem elhanyagolhatók. A tapasztalatok szerint ez különösen sikertelen lépés esetén állhat fenn. Ha a DE nem elég sima, akkor (1.13) nem igaz, a h_n hatványai redukálódnak annak megfelelően, mennyire sima a DE megoldása. Ha a megoldás sima, még akkor is előfordulhat, hogy a főhibatag $\psi(x, y)$ elég gyorsan változik, és a fent leírt módon felhasználva, egy sikeres lépés után gyenge közelítést kaphatunk. Általában a lokálisan optimális lépéshossz egy tört-részt használják a gyakorlatban. Pl. az irodalomban DIFSUB név alatt ismert programban /Gear /1971// a lokálisan optimális lépéshossz 0.99-ed részét használják; de általában ez 0.5; vagy a 0.8 értéktől 0.9-ig változik.

Megjegyezzük még, hogy a globális hiba nem reguláris viselkedésű, ha nem a közelítőleg optimális lépéshosszt használjuk. Pl. Ebben az esetben előfordul, hogy a hibakorlát csökkentésével a globális hiba kicsit változik, ha egyáltalán változik.

Példa: az irodalomból RKGS (CDC programkönyvtár) néven

ismert Runge-Kutta-Gill módszer programját alkalmazzuk a következő DE-re

$$y_1' = y_2, \quad y_1(0) = 0,$$

$$y_2' = -y_1, \quad y_2(0) = 1,$$

a $[0, 8\pi]$ intervallumon. Az eljárás lépéshosszfelezéssel ill. duplázással változtatja a lépéshosszt. Bemutatjuk a 2. táblázatban a megoldás két komponensének hibáját a 8π pontban különböző ε hibakorlátok esetén.

$-\log_{10} \varepsilon$	y_1 hibája	y_2 hibája	függvénybehívások száma
1	-8.2/-4/	-1.7/-4/	562
2	-8.2/-4/	-1.7/-4/	562
3	-8.2/-4/	-1.7/-4/	562
4	-8.2/-4/	-1.7/-4/	562
5	-6.6/-5/	-8.7/-6/	1108
6	-5.0/-5/	-5.4/-6/	1119
7	-1.6/-6/	-1.3/-7/	2226
8	1.4/-6/	1.5/-8/	4411
9	1.5/-6/	1.7/-8/	4433
10	1.7/-6/	1.1/-8/	8840
11	1.7/-6/	5.5/-9/	17559
12	1.7/-6/	5.5/-9/	17647

2. táblázat Példa

Az alábbiakban Gear, Tu /1974/ Gear, Watanabe /1974/ eredményeit összegezzük az (1.3) módszerosztályra ill. annak egyes speciális módszereire vonatkozóan.

10. Definíció. Lépéshosszválasztási sémának nevezzük az olyan függvényt, amelyre

$$h_n = h \Theta(x_n, h),$$

ahol tetszőleges $h > 0$, $0 \leq x \leq b$: $1 \geq \theta(x, h) \geq c > 0$.
A legismertebb a következő két lépéshosszváltoztatási technika.

Az un. interpolációs technika a h_n lépéshosszal való számoláskor az új lépéshosszra ugyanolyan rendű formula beinterpolálásából áll, azaz

a/ interpolálás az ismert közelítésekből, hogy az

$\tilde{x}_{n-i} = x_n - ih_n$ ($k > i \geq 0$) pontokban megkapjuk a megoldás közelítését.

b/ ezeken az $\tilde{x}_{n-i}$ pontokon alapuló fix-lépéses formula segítségével megkeressük a közelítő megoldást az $x_{n+1} = x_n + h_n$ pontban.

A változó lépéshosszú technika alkalmazásakor k darab nem egyenletesen elhelyezett ($k > i \geq 0$) pontból kiindulva kiszámítjuk az új lépést és a k -lépéses formula együtthatóit, hogy az megfelelő rendű legyen.

Ismert /Dahlquist /1956//, hogy a k -lépéses módszer rendje nem érheti el a $2[k/2] + 2$ értéket, ha azt akarjuk, hogy a módszer rögzített lépéshossz esetén stabilis legyen; ezért néhány $\alpha_{i,n}$ és $\beta_{i,n}$ együttható értékét előírjuk. Például, ha azt adjuk meg, hogy $\alpha_{j,n} = 0$ ($j \geq 2$) esetén és a módszer $(k+1)$ -ed rendű legyen, akkor a változó lépéshosszú Adams-Moulton módszer lesz a megfelelő rendű formula.

Ismertek a következő tételek:

11. Tétel A változó lépéshosszú (1.5) módszerek közé tartozó Adams-Basforth /AB/, A-B-Moulton /ABM/ prediktor-korrektor és A-M módszerek stabilisak¹¹ és konvergensek tetszőleges lépéshosszválasztási sémára.

¹¹ A stabilitás a 11. 12. 13. tételben kicsit eltérő az ismertett definíciótól, de az f függvényre vonatkozó alkalmas feltételek mellett ekvivalensek.

12. Tétel Ha legalább k konstans lépést teszünk a lépéshossz változtatások között, akkor a k -lépéses interpolációs AB, ABM és AM módszerek stabilisak és konvergenssek.

13. Tétel Ha egy módszer fix-lépéshosszra kielégíti a stabilitási feltételt, akkor tetszőleges olyan lépéshossz választási sémára stabilis /változó lépéshosszú ill. interpolációs sémára/, amely csak kis változást eredményez, azaz amelyre

$$\frac{h_{n+1}}{h_n} = 1 + O(h) \quad (h \rightarrow 0).$$

Az interpolációs technika előnye a rugalmas használhatóságában rejlik, / az áttérés az egyik lépéshosszról a másikra általában egy mátrix-vektor szorzás segítségével megoldható /Nordsieck /1962//. De ez a technika nem megfelelő ha magasabb rendű formulákat használunk /pl. Krogh /1973//.

A változó lépéshosszú technika esetén nehezkesebb a lépéshosszváltoztatás utáni számolás.

11. Definíció. Formula választási sémának nevezzük azt az $I(h, x)$ függvényt, amely definiálja az $[x_n, x_{n+1}]$ intervallumon használt $F_I(h, x_n)$ formulát, ahol h a lépéshosszválasztási séma által definiált paraméter.

12. Definíció. Változó formulát használó módszer a következőkből áll: az $\{F_i\}$ többlelépéses formulák halmazából, valamint a $\{\Theta_{ji}\}$ formula változtatási operátorok halmazából, a lépéshosszváltoztatási technikából, a lépéshossz- és formulaválasztási sémából.

14. Tétel A változó lépéshosszú technikát felhasználó változó rendű (1.10) Adams módszer stabilis tetszőleges formula választási séma esetén.

6. Effektivitás, a módszerek alkalmazása

Végül szükséges megemlítenünk a következő az irodalomban eddig kevésbé vizsgált problémát. A módszerek konvergenciájára vonatkozó tételek egy fix módszer adott felosztássorozaton való konvergenciáját biztosítják adott $\|\Delta_N\| \rightarrow 0$ felosztássorozaton. Jelölje $M = \langle (\varphi, \Phi), B, D, H \rangle$ azokat az eljárásokat ahol (φ, Φ) az adott (1.3) alakú módszer, B a képlethiba valamely becslése, D adott döntési kritérium és H adott lépéshosszválasztási stratégia. Tekintsük az (1.1) (1.2) feladat $x \in [x_0, b]$ pontbeli megoldását az (1.17) feltétel mellett, azaz a

$$P(\varepsilon) = \langle \varphi, x_0, y_0, x; \begin{aligned} &\|\tau_n(x_n, h_n)\| \leq \varepsilon, x_{n+1} \in \Delta_N \in \pi \\ &\|B(x_n, h_n)\| \leq \varepsilon, \varphi \in F \end{aligned} \rangle$$

problémaosztályt /ahol F adott függvényosztály/.

13. Definíció. A $\langle (\varphi, \Phi), B, D, H \rangle$ eljárást effektívnek nevezzük a $P(\varepsilon)$ problémaosztályra nézve, ha vele az osztálynak minden eleme megoldható /Hull /1969//.

Ha a (φ, Φ) módszer konvergens bizonyos típusu $\{\Delta_N\} \subset \pi$ sorozatokra, akkor az eljárás effektivitása olyan B, D, H kritériumokat követel meg, amely ilyen típusu $\{\Delta_{N(\varepsilon)}\} \subset \pi$ ($\varepsilon \rightarrow 0$) felosztássorozatot állít elő és

$$(1.22) \quad \|y_{N(\varepsilon)} - y(x_N)\| \leq \varepsilon \quad (0 < \varepsilon < \varepsilon_0).$$

Az effektivitás vizsgálata azért lényeges, mert a (φ, Φ) módszer konvergenciájából még nem következik, hogy adott $p(\varepsilon) \in P(\varepsilon)$ probléma esetén a $\langle (\varphi, \Phi), B, D, H \rangle$ módszer által számított közelítés globális hibájára teljesül az (1.22) feltétel.

Effektivitási tételeket egyszerűbb esetekben ekvidisz-tans felosztásokon Hull /1969/ /lásd alább/ és Sacks-Davis

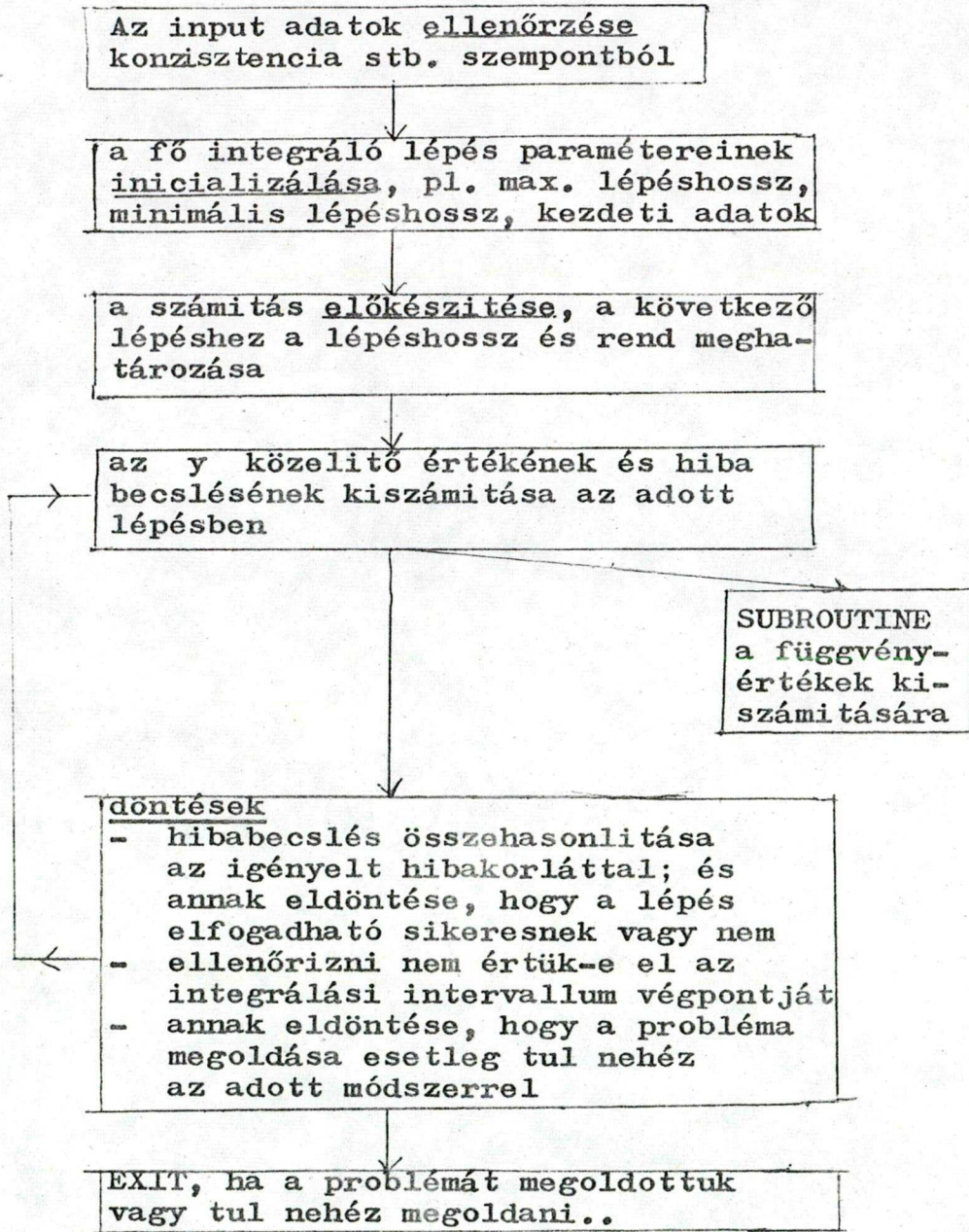
/1977/ igazoltak. Megjegyezzük, hogy a 10. tétel is interpretálható általános effektivitási tételként.

15. Tétel /Hull /1969//. Tekintsük az $y' = Ax, y(x_0) = y_0$ feladatot, (y_0, x_0) adott), A $m \times m$ dimenziós négyzetes mátrix. Legyen továbbá b adott végpont és egy lépés elfogadhatóságának kritériuma az, hogy a lokális hiba egy adott ε korlát alatt van. Tegyük fel, hogy explicit negyedrendű Runge-Kutta módszert alkalmazunk lépésfelezéses hibabecsléssel.

Akkor a módszer effektív a fenti problémaosztályon, ha

- 1/ egylépés eredményét csak akkor fogadjuk el, ha a lokális hibára kapott becslés nem éri el a $3/4 h \varepsilon$ értéket és
- 2/ a h lépéshossz kiválasztási stratégiája olyan, hogy $h \leq h_{\max}$, ahol $h_{\max} \leq 1/4 \|A\|$ és a stratégia olyan, hogy a módszer segítségével a b végpont elérhető legyen.

Az 1. ábrán bemutatjuk egy DE közelítő megoldásának gyakorlati folyamatát, amelynek segítségével például egy módszer számítógépre vihető.



1. ábra Az implementálás folyamata

II. fejezet

Módszerek összehasonlító vizsgálata

1. Általános szempontok

Számos módszer létezik az (1.1) - (1.2) Cauchy-féle kezdeti érték probléma közelítő megoldására, de nem létezik olyan optimális módszer, amely tetszőleges differenciálegyenlet esetén mindig a legjobb eredményt adja.

A különböző szempontokból optimális módszerek kiválasztására számos technika ismeretes. Ezek lényegét világítjuk meg az alábbiakban.

Általában ha numerikus módszereket akarunk összehasonlítani, definiálnunk kell a következőket:

1. A megoldandó problémák osztálya
(Jelölje ezt P , elemeit $p \in P$)
2. A módszerek osztálya: M ($m \in M$)
3. Összehasonlítási kritérium: $A(p, m)$
módszer és problémapárnak megfeleltetünk egy $C(p, m) > 0$ értéket, amely a p probléma m módszerrel történő megoldása jóságának mértéke.

Ha ezeket definiáltuk, akkor világos, hogy mit értünk azon, hogy az egyik módszer jobb mint a másik. Azt mondjuk, hogy az m módszer jobb, mint az m' módszer a P problémaosztályra a C kritériumnak megfelelően, ha $C(p, m) < C(p, m')$.

Azt is mondhatjuk, hogy m a legjobb módszere az M módszerosztálynak a C kritériumra nézve, ha a P problémaosztályra alkalmazva $C(p, m) \leq C(p, m')$; $\forall m' \in M$ esetén.

A problémák osztályát esetünkben az (1.1), (1.2) Cauchy-feladatok alkotják. Ezek leírhatók a következőképpen:

$$p = \langle f, x_0, y_0, b, \varepsilon, \dots, stb. \rangle$$
 ahol f, x_0, y_0 definiálja a problémát matematikailag, ε a lokális hiba felső korlátja és b a végpont, amelyben a megoldást keressük.

A közelítő módszereket a korábbiaknak megfelelően definiálhatjuk /algoritmus, hibabecslés, döntés, lépéshosszválasztás/.

Az összehasonlítási kritérium lehet az algoritmussal való számoláshoz és a hibabecsléshez szükséges függvénybehelyettesítések száma.

2. Kísérleti vizsgálatok

Hull és társai /1972/ dolgozták ki az első komolyan megalapozott kísérleti /számítógépes/ összehasonlító vizsgálatokat.

Vizsgálataikban ε az egységnyi lépésre megkövetelt lokális hibakorlát. A probléma további jellemzésére bevezetik a $h_{\max}$ maximálisan megengedett lépéshosszt. A problémák nem tartalmazznak sem szakadásos jobboldallal rendelkező egyenletet, sem stiff típusu vagy más, az átlagnál rosszabb egyenletet. Az egyes problémaosztályok a következők

- a/ egy egyenlet,
- b/ kisméretű egyenletrendszer
- c/ közepes méretű egyenletrendszer
- d/ orbitális egyenletek
- e/ magasabbrendű egyenletek.

Minden osztály öt feladatból áll, amelyek mindegyikét vizsgálják $10^{-3}, 10^{-6}, 10^{-9}$ nagyságú ε hibakorlát esetén.

A felhasznált algoritmusok a következők: két Runge-Kutta módszer /Butcher nyolcadrendű módszere /1965/ és Shanks módszere /1966//, az Adams-módszer /Gear féle változat /1971/, lásd még az 1. táblázatot/, Bulirsch, Stoer /1966/ extrapolációs módszere. Összehasonlítási kritériumnak a felhasznált függvénybehelyettesítések számához hozzáveszik az ún. "overhead" időt, amely alatt a gépidő azon részét értik, amely fennmarad, ha kivonjuk a függvénybehelyettesítésekre elhasznált gépidőt.

Egy statisztikai program segítségével gyűjtik össze ezeket az adatokat, valamint bizonyos megbízhatósági vizsgálatához szükséges számolásokat /a program kiszámítja a feladat pontos megoldását is és az attól való eltérést/.

Az alábbi következtetésekre jutnak:

a/ A változó rendű lineáris többlépéses módszerek általában jobbak, mint a rögzített rendűek. Általános esetben a leghatékonyabban használhatók, ezért ezek a legelterjedtebbek ma.

b/ A Bulirsch-Stoer módszer kevesebb "overhead" időt igényel, mint az Adams módszerek /azaz, ha a függvénybehelyettesítések viszonylag költségesek /kb. ≥ 25 aritmetikai műveletet igényelnek komponensenként/, akkor az Adams módszer használata célszerűbb.

c/ A Runge-Kutta módszerek általában nem versenyképesek a többi vizsgált módszerrel. Kivétel az az eset, ha nincs szükségünk nagy pontosságú megoldásra; akkor ui, az alacsonyabbrendű Runge-Kutta módszereket érdemes használni, ha a függvényértékek számítása nem túlságosan bonyolult.

A továbbiakban egy lépéses módszerek összehasonlításával foglalkozunk.

Shampine és Watts /1976/ széleskörű szempontok figyelembevételével elvégzett kísérleti /számítógépes/ módszerekkel hasonlítják össze a Runge-Kutta formulák alkalmazását egy adott feladatosztályon.

A problémaosztály a következő:

$$(2.1) \quad y'(x) = \sum_{i,j=0}^{\infty} a_{ij} x^i y^j, \quad \text{/konvergens hatványsor/}$$

$$y(0) = 0,$$

ahol az a_{ij} értékek véletlen számok, amelyek a $[-1, 1]$ szakaszon egyenletes eloszlásúak /mintegy 500 egyenletet generálnak a teszteléshez/.

Az (1.7) alakú m -pontos p -ed rendű / $m \leq p \leq 5$ / explicit Runge-Kutta módszerek osztályából vizsgálja az alábbiakat:

- a/ a klasszikus negyedrendű módszert lépésfelezéssel
- b/ Gill /1951/ módszerét lépésfelezéssel
- c/ England /1969/ módszerét, amelynek hibabecslése beágyazási típusu
- d/ Merson negyedrendű módszerét /amely csak lineáris egyenletekre ad aszimptotikusan pontos hibabecslést/ harmadrendű módszerként interpretálva, amely lokális extrapolációt használ /igy aszimptotikusan pontos hibabecslés adható/
- e/ Zonneveld /1964/ ötödrendű módszerét, amelyet lokális extrapolációt használó negyedrendű módszerként interpretálnak
- f/ Shintani /1966/ negyed-ötödrendű beágyazási típusu módszerét
- g/ Fehlberg /1969/ negyed-ötödrendű beágyazási típusu módszerét két változatban, amelyben csak a módszer paraméterei térnek el egymástól.

Összehasonlitásokat végeznek az adott problémaosztályon, amelyek alapja a pontosság, a hibabecslés minősége, a stabilitás és az általános effektivitás /hatékonyság/.

Az adott módszerek pontosságát a főhibatag szerinti összehasonlítással vizsgálják /három különböző normában/.

Ahhoz, hogy értelmes összehasonlítást lehessen tenni különböző számú pontból álló formulák esetén, szükségünk van arra, hogy skálázzuk a lépéshosszakat, így véve figyelembe a nem egyenlő mennyiségű munkát. Így tekintjük a $h^* = mh$ lépéshosszt, azaz m -pontos módszer esetén a h^q -ad rendű tagokhoz tartozó hibaegyütthatókat beszorozzuk az m^q mennyiséggel. Ez utóbbit fogjuk "skálázott"-nak nevezni a továbbiakban. A 3. táblázatban a hibakonstansokat mutatjuk be "skálázott" és nem skálázott esetben.

A hibabecslés minőségének vizsgálatára végeznek aszimptotikus és nemaszimptotikus összehasonlitásokat. Felteszik, hogy a módszer lokális hibája és annak becslése az alábbi alakú:

$$le = ah^{p+1} + bh^{p+2} + O(h^{p+3})$$

/a lokális hiba/

$$est = \alpha h^{q+1} + \beta h^{q+2} + O(h^{q+3})$$

/a lokális hiba becslése/

Aszimptotikusan pontos becslés esetén $q = p$ és $\alpha = a$.

Megkapják közelítőleg a

$$(2.2) \quad D(f) := \lim_{h \rightarrow 0} \frac{le - est}{h^{p+2}} = b - \beta$$

aszimptotikusan pontos becslést úgy, hogy a lokális hibát és becslését a $h = 2^{-N}$ ($N = 3, 4, \dots$) sorozatra számítják ki a vizsgált feladatosztályon. Ezt addig végzik, amíg a $D(f)$ szukcessziv közelítései egy előírt mennyiségnél kevésbé térnek el. Végül a $|D(f)|$ értékek átlagát és

maximális értékét számítják ki minden vizsgált módszerre a

(2.1) alakú véletlen módon előállított feladatokra. A $D(f)$

mennyiség numerikus konvergenciájának elérése után a hibabecslés egy másik mértékét, a

$$(2.3) \quad R(f) = \frac{D(f)}{\beta} \quad \text{mennyiséget}$$

számítják ki közelítőleg.

A 4. táblázatban bemutatjuk a $|D(f)|$ és $|R(f)|$ közelítő értékek átlagát és maximumát.

Az 5. táblázatban a felhasznált módszerek rangsorolását próbálják megadni a hibabecslések szempontjából. Megjegyzik, hogy bizonyos esetekben nem lehet egyértelműen kiválasztani az eljárások közül a legjobbat, legalábbis az extrapoláció nélküli esetben mindegyik hibabecslés megfelelő.

3. táblázat

A pontosság mértéke, a hibatag együtthatói

Módszerek	p	Skálázatlan			Skálázott		
		h^{p+1}	h^{p+2}	h^{p+3}	h^{p+1}	h^{p+2}	h^{p+3}
Merson	3	.010	.021	.028	6.4	6.4/+1/	4.4/+2/
Klasszikus	4	.0022	.0035	.0054	3.5/+2/ ^{II}	6.3/+3/	1.1/+5/
Gill	4	.0019	.0031	.0048	3.1/+2/	5.5/+3/	9.4/+4/
Zonneveld	4	.20	.51	.70	3.4/+3/	6.0/+4/	6.5/+5/
England	4	.0019	.0031	.0048	1.1/+2/	1.6/+3/	2.3/+4/
Shintani	4	.00089	.0039	.0087	1.5/+1/	4.6/+2/	7.1/+3/
Fehlberg/1/--		.0061	.015	.025	4.8/+1/	7.0/+2/	7.0/+3/
Fehlberg/2/--		.0033	.018	.039	2.5/+1/	8.3/+2/	1.1/+4/

^{II} a zárójelben lévő szám a 10 megfelelő hatványával történő szorzást jelöli

4. táblázat

A hibabecslések aszimptotikus összehasonlítása

Módszerek	D/f/ átlaga skálázott	D/f/ max. skálázott	R/f/ átlaga	R/f/max.
Shintani	77.	4.1/+2/	5.4	166.
Fehlberg/2/	102.	4.1/+2/	2.4	85.
Merson	20.	6.9/+2/	1.1	37.
England	78.	7.0/+2/	.72	33.
Fehlberg/1/	47.	2.1/+2/	.69	15.
Gill	202.	1.4/+3/	.43	19.
Klasszikus	271.	1.8/+3/	.39	11.
Zonneveld	151.	9.8/+2/	.036	2.8

5. táblázat

A hibabecslések rangsorolása legkevésbé hatékonytól a legjobbig

Nem extrapolációs formulákra	Extrapolációs formulákra
Shintani, Fehlberg/2/	Zonneveld
Merson, Fehlberg/1/	Fehlberg/1/
England, Klasszikus, Gill	Shintani
Zonneveld	Fehlberg/2/

Shampine és Watts /1976/ ismerttetett munkájukban effektivitási kérdéseket is érintenek. Összehasonlítják a módszerek hatékonyságát, amikor azok már összemérhető pontosságot értek el. Az egyszerűség kedvéért a magasabbrendű tagokat elhanyagolják a hiba sorfejtésben, így csak a következő tagot veszik:

$$\text{error} = \epsilon h^{p+1}$$

Felteszik, hogy a k-adik módszer hibatagja $\epsilon_k h^{p+1}$, ahol az ϵ_k értékeket a 3. táblázat tartalmazza. A k-adik módszer

h_k lépéshossza, amely mellett a hibatag ugyanaz lesz

$$h_k = \left(\frac{\varepsilon}{\varepsilon_k} \right)^{\frac{1}{p+1}} h$$

Ha N darab h hosszúságú lépésre van szükség, egy adott pont eléréséhez, akkor

$$N_k = \left(\frac{\varepsilon_k}{\varepsilon} \right)^{\frac{1}{p+1}} N$$

darab h_k hosszúságú lépés szükséges ugyanazon pont eléréséhez. A 7. táblázat az egyes módszerek teljes költségét mutatja a függvénybehelyettesítések számával és az ún. "overhead" idővel mérve.

7. táblázat

A módszerek költsége

	Függvénybehelyettesítések / lépés száma
Merson	5
Fehlberg	6
Zonneveld, Shintani	7
England	9
Klasszikus, Gill	11

	műveletek / lépés száma	műveletek / pont száma
Klasszikus	31	2.8
Gill	37	3.4
Merson	19	3.8
England	35	3.9
Zonneveld	33	4.7
Fehlberg	30	5.0
Shintani	37	5.3

A különböző szempontok alapján elvégzett vizsgálatok azt mutatják, hogy nem létezik minden kategóriában legjobb módszer.

A fenti vizsgálatok eredményei alapján a szerzők a Fehlberg módszert találták a legjobban és a legáltalánosabban alkalmazhatónak.

3. Elméleti összehasonlítási vizsgálat

Elméleti módszerekkel vizsgálja Jackson /1978/ a Runge-Kutta formulák költségének nagyságát bizonyos feladatosztályon.

Az (1.7) alakú m -pontos p -ed rendű explicit Runge-Kutta módszerek osztályát tekinti a lineáris nem stiff homogén konstans együtthatós problémákra. A lokális hiba becslésére felhasználja a korábban említett módszerek közül a lépésfelzési technikát, Ceschino-Kuntzman többlépéses hibabecslését és néhány beágyazási típusú módszert. Az adott integrálási lépés elfogadhatóságának kritériuma az ún. retrográd hibaanalízis elvén alapszik. Adott ε hibakorlátra megkövetelik, hogy az eredeti (1.1), (1.2) probléma közelítő megoldása az $[x_0, b]$ integrálási intervallumon a perturbált probléma /mely az eredeti feladat ε -nal korlátos perturbációja/ pontos megoldása legyen. Az integrálási lépés elfogadására használt stratégiák közül számos követeli meg az alábbi egyenlőtlenség teljesülését

$$(2.4) \quad a(h) \|E(h)\| \leq \varepsilon$$

ahol $E(h)$ a lokális hiba becslése egy adott lépésben, $a(h)$ pedig tetszőleges függvény /sok esetben 1 vagy h /.

Modellmódszerük olyan, hogy az $a(h)$ függvényt helyettesítik egy $b(h)$ ún. effektivitási függvénnyel.

Egy problémaosztályra vonatkozó effektivitási függvény-nek $b(h)$ nevezik azt a legkisebb /pontoszerű/ függvényt,

amely garantálja, hogy minden $p \in P$ esetén, ha a (2.4) egyenlőtlenség az integrálás minden lépésénél teljesül, akkor a módszer által generált approximáció kielégíti a problémához tartozó elfogadhatósági kritériumot.

A lépéshossz kiválasztásának stratégiája olyan, hogy kiválasztja a (2.4) feltételt kielégítő legnagyobb lépéshosszt, azaz $H(\varepsilon)$ a következő egyenlet egyértelmű gyöke

$$(2.5) \quad b(h) \max_{\|A\| \|y_0\| \leq 1} \|E(h)\| = \varepsilon.$$

Az adott módszer költségét a következőképpen definiálja

$$(2.6) \quad c(\varepsilon) = \max_{\substack{\|A\| \leq 1 \\ \|y_0\| \leq 1}} \left\{ \begin{array}{l} \text{az } y' = Ay, y(0) = y_0 \text{ Cauchy} \\ \text{probléma megoldásához egy egységnyi} \\ \text{lépésre felhasznált azon függvény-} \\ \text{számolások minimuma, amelyekre szük-} \\ \text{ség van adott } \varepsilon \text{ hibakorlát mellett} \end{array} \right\}$$

A lineáris konstans együtthatós homogén problémák (ahol $\|A\| \leq 1$ és $\|y_0\| \leq 1$) osztályán egy lépés költsége

$$(2.7) \quad c(\varepsilon) = F/H(\varepsilon)$$

ahol F a formula és a hozzá tartozó hibabecslés által lépésenként elhasznált függvénybehelyettesítések száma.

Szigorú hibakorlátokra a költséggörbék közel lineárisak

$$\log c(\varepsilon) \approx K - \frac{1}{p} \log \varepsilon,$$

ahol K egy konstans, p a módszer rendje.

A $b(h)$ effektivitási függvény közelítését is sikerült előállítaniuk három lépésben. Először összekapcsolják az $[x_{k-1}, x_k]$ intervallumra vonatkozó retrográd hibát és az ugyanerre az intervallumra vonatkozó egységnyi lépéshez tartozó τ_k hibát egy $z(x)$ függvény segítségével. Az egységnyi lépésre vonatkozó hiba a k -adik lépésben

$$(2.8) \quad \tau_k(h_k) = [y_k(x_k) - y_k] / h_k,$$

ahol $y_k(x)$ az $y'(x) = f(x, y(x))$, $y_k(x_{k-1}) = y_{k-1}$
 "lokális" kezdeti érték probléma megoldása.

A $z(x)$ függvényre a következő követelményeket írják elő:

a/ $z(x)$ legyen folytonos az $[x_0, b]$ intervallumon;

b/ $z(x_k) = y_k$ ($k=0, 1, \dots, n$), ahol $y_n = \tilde{y}_n$ és

c/ $z'(x) - Az(x) = u_k(x)$, $x \in [x_k, x_{k+1}]$ ($k=1, 2, \dots, n$)

ahol $u_k(x)$ az ismeretlen. A szabályozás elméletből ismert, /lásd pl. Pontrjagin /1968/, hogy tetszőleges olyan $u_k(x)$ függvény, amely kielégíti a fenti feltételt, és minimalizálja a

$\max \|u_k(x)\|$, $x \in [x_{k-1}, x_k]$ mennyiséget, az $[x_{k-1}, x_k]$

szakaszon konstans normájú, ezért $u_k(x)$ helyett itt vehető egy u_k konstans vektor /lásd pl. Sedgwick /1973//. Így $x \in [x_{k-1}, x_k]$ esetén

$$(2.9) \quad z(x) = e^{(x-x_{k-1})A} y_{k-1} + (x-x_{k-1}) e_1 ((x-x_{k-1})A) u_k$$

ahol

$$e_1(x) = \begin{cases} \frac{e^x - 1}{x} & , \quad \text{ha } x \neq 0, \\ 1 & , \quad \text{ha } x = 0. \end{cases}$$

Az $e_1(x)$ függvény kiterjeszthető négyzetes mátrix függvényekre, létezik a reciproka és analitikus mindenütt a C téren kivéve az $e_1(x)$ gyökeinél $\{2\pi i k; k=\pm 1, \pm 2, \dots\}$. Így az $e_1^{-1}(x)$ szintén kiterjeszthető mátrixfüggvényekre, feltéve, hogy a mátrixargumentum egyetlen sajátértéke sem gyöke az $e_1(x)$ -nek. $e_1^{-1}(hA)$ az $e_1(hA)$ mátrix inverze, ha $\|hA\| \leq 2\pi$. A (2.4) formulába $x = x_k$ értéket helyettesítve, valamint felhasználva, hogy $h_k = x_k - x_{k-1}$ és $\tau_k = (e^{h_k A} y_{k-1} - y_k) / h_k$, kapjuk



$$(2.10) \quad u_k = -e_1^{-1}(h_k A) \tau_k.$$

A p -ed rendű explicit Runge-Kutta formula esetén τ_k kielégíti a következő egyenletet

$$\tau_k = S(h_k A) \frac{(h_k A)^{p+1}}{h_k} y_{k-1},$$

ahol $S(x)$ egy végtelen konvergencia sugarú hatványsor. A formulához tartozó hibabecslés kielégíti a következő egyenletet

$$E_k = R(h_k A) \frac{(h_k A)^{p+1}}{h_k} y_{k-1},$$

ahol $R(x)$ egy polinom. Ha $R(h_k A)$ nem szinguláris, akkor E_k és τ_k összekapcsolható a következő egyenlettel:

$$(2.11) \quad \tau_k = S(h_k A) R^{-1}(h_k A) E_k$$

Felteszik, hogy az $S(x)$ és $R(x)$ polinomoknak nincs közös gyökük.

A (2.10) és (2.11) egyenletből következik, hogy u_k és E_k között a következő összefüggés

$$u_k = -e_1^{-1}(h_k A) S(h_k A) R^{-1}(h_k A) E_k$$

feltéve, hogy $R(h_k A)$ nem szinguláris. Így az effektivitási függvény közelítésére az alábbi kapták:

$$(2.12) \quad b(h) = \max_{\|A\| \leq 1} \|e_1^{-1}(hA) S(hA) R^{-1}(hA)\|,$$

$$0 \leq h < \min(h_{\max}, 2\pi).$$

A $b(h)$ kiszámításához $R^{-1}(x)$ -et $h_{\max}$ konvergenciasugarú hatványsorba fejtjük. Ekkor a (2.12) formulába a $\| \cdot \|$ kifejezésben szereplő mennyiségek reprezentálhatók egy $P(x)$ hatványsorral, amelynek a konvergenciasugara nagyobb vagy egyenlő, mint $\min(h_{\max}, 2\pi)$. Általánosabb feltételek mellett is közelíthető az effektivitási függvény, ha $\|A\| \leq a$. A költségfüggvény közelítését is előállítják. A $H(\varepsilon)$ lépéshossz az alábbi egyenlet egyetlen gyöke /lásd (2.5) formula/.

$$b(h) \max_{\|A\| \|y_0\| \leq 1} \|R(hA)(hA)^P(Ay_0)\| = \varepsilon.$$

Sedgwick lemmáját felhasználva kapjuk, hogy

$$\max_{\|A\| \|y_0\| \leq 1} \|R(hA)(hA)^P(Ay_0)\| = h^P R^*(h),$$

ahol $R(x) = \sum_{i=0}^N r_i x^i$ esetén $R^*(x) = \sum_{i=0}^N |r_i| x^i.$

Igy $H(\varepsilon)$ az alábbi egyenlet egyetlen gyöke

$$h^P b(h) R^*(h) = \varepsilon.$$

Végül a költségfüggvényre a (2.7) formulát kapja. A költségfüggvény szigorúan csökkenő.

A 3. ábrán közöljük a vizsgált módszerek költségfüggvényeit, amelyek nem használnak lokális extrapolációt, a 4. ábra pedig a lokális extrapolációt használó módszerek költségfüggvényeinek görbéit ábrázolja.

Néhány megfigyelésüket megemlítjük.

1. Ha a hibakorlát viszonylag erős, a magasabbrendű formulát használó módszer kevésbé költséges, mint az alacsonyabb rendű formulán alapuló. Gyakran előfordul viszont, hogy az alacsonyabbrendű formulák enyhébb hibakorlátok esetén kevésbé költségesek.

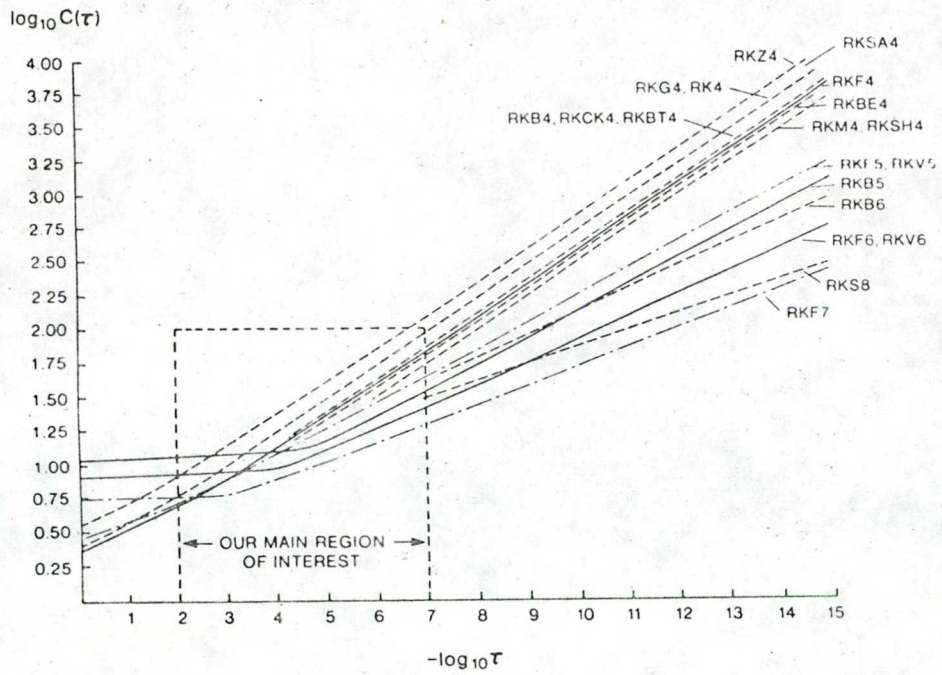
2. A lokális extrapoláció használata lehetőséget nyújt egy módszer viselkedésének javítására. A lokális extrapoláció megnöveli a formula rendjét. Így elég erős hibakorlátok esetén az extrapoláció bizonyára segít. Ugyanakkor az általuk vizsgált módszerek esetén azok, amelyek lokális extrapolációt használtak, kevésbé költségesek a különböző hibakorlátokra.

3. Általában nem érvényes az a feltételezés, hogy ha az egyik módszer a lokális extrapoláció használata nélkül jobb, mint a másik, akkor abban az esetben is jobb lesz, ha lokális extrapolációt használnak.

4. Általánosan elfogadott nézet, hogy a formula páron alapuló módszerek jobbak, mint az egy formulából és lépésfelezési hibabecslésből álló módszerek. Az elméleti eredmények ezt igazolni látszanak.

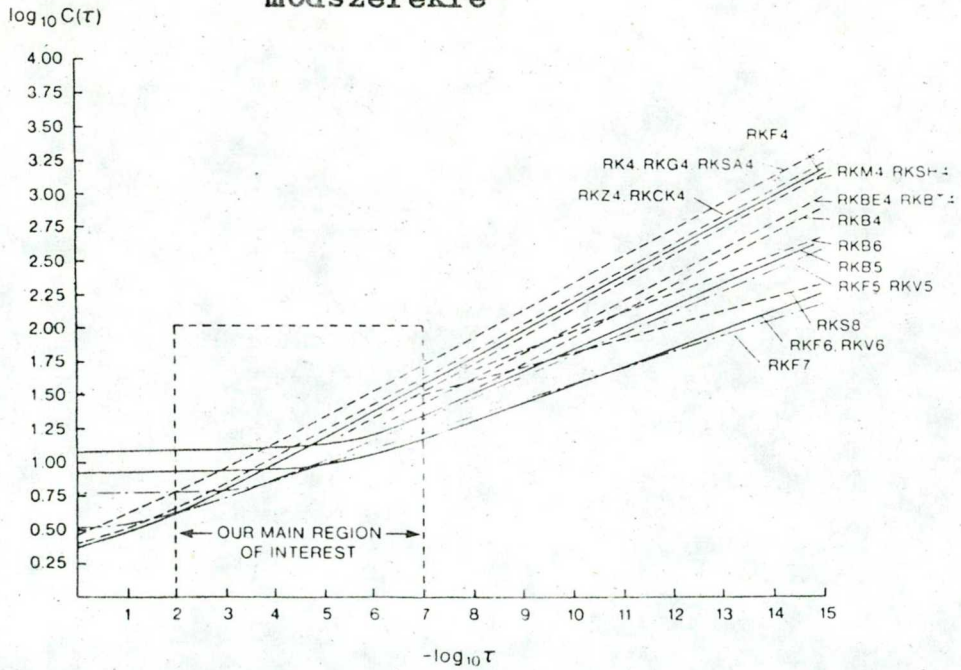
Az adott problémaosztályra a legjobbnak az ötöd- és hatodrendű Verner formulák bizonyultak. /Verner /1978//.

Számítógépes kísérleti eredményeik igazolni látszanak, azt a megállapítást, hogy az általuk konstruált elméleti költségfüggvények az adott problémaosztályra az egyes Runge-Kutta formuláknak és a formulák hibabecsléseinek jó jellemzését adják. Az elméleti módszer ugyanakkor gyorsabb és kevésbé költséges, mint a kísérleti tesztelések.



3. ábra

Költségfüggvénygörbék lokális extrapolációt nem használó módszerekre



4. ábra

Költségfüggvénygörbék lokális extrapolációt használó módszerekre

4. Összefoglalás

Az ismertetett elméleti és kísérleti vizsgálatok a módszerek objektív kiválasztásának lehetséges közelítéseit adják. Érdemük, hogy a numerikus matematikában meghonosodott korábbi gyakorlattal szemben a vizsgálatok feladat és módszer osztályokra vonatkoznak, valamint komplex jellegűek, amennyiben sokféle hatást /hibabecslés, lépéshosszválasztás stb./ vesznek figyelembe. Nyilvánvaló hiányosságuk azonban, hogy a vizsgált feladatosztályok rendkívül speciálisak, és előfordulhat az is, hogy nem azonos rendű módszereket hasonlítanak össze.

Igy általánosnak mondott következtetések csak a vizsgált osztályok vonatkozásában fogadhatók el. Eredményeik azonban a korábbi állapothoz képest nagymértékben hozzájárulnak a módszerek megalapozottabb kiválasztásához, noha vizsgálataik következtetései számos esetben ellentmondanak az aszimptotikus vizsgálatok eredményeinek. Ennek oka az, hogy a vizsgálatokat olyan pontossági tartományokra végezték, amelyek az adott konkrét feladattípusok gyakorlati megoldásában előfordulnak.

Megjegyezzük, hogy a fejezetben ismertetett módszerek közül számosnak adaptáltuk a programját az Akadémia CDC 3300-as számítógépére. Többek között ezek a következők:

1. a negyedrendű klasszikus Runge-Kutta módszer
2. Runge-Kutta-Gill módszer
3. Zonneveld módszer
4. Merson módszer
5. Bulirsch-Stoer féle extrapolációs eljárás
6. Fehlberg negyed-ötödrendű módszere
7. Gear módszer stb.

Az algoritmusok és programok rövid leírása a dolgozat 2. táblázatában ill. a CDC 3300-as számítógép programkönyvtári anyagában található /lásd. Numerikus módszerek programgyűjtemény/.



III. fejezet

Általánosított egylépéses módszerek konvergenciája és stabilitása

1. Bevezetés

A fejezetben változó lépéshosszú és változó formulájúnak tekinthető általánosított egylépéses módszerekkel foglalkozunk.

Az eredmények egyszerűen adódnak az (1.1), (1.2) közösleges DE-knél szélesebb differenciálegyenletosztályra, funkcionál DE-kre, ezért azokat ilyen általánosabb alakban írjuk le.

Tekintsük az alábbi funkcionál DE-re vonatkozó kezdeti-érték problémát. Legyen

$$(3.1) \quad \begin{aligned} z'(x) &= f(x, z_x), & (0 < x < X) \\ z_0(x) &= \varphi(x), & (-\bar{x} \leq x \leq 0) \end{aligned}$$

$\varphi(x) \in C([-\bar{x}, 0], \mathbb{R}^n)$, ahol Hale /1971/ jelöléseit használtuk, azaz $z_x(\theta) = z(x+\theta)$ ($-\bar{x} \leq \theta \leq 0$) és $z_x \in C([-\bar{x}, 0], \mathbb{R}^n)$.

A továbbiakban vezessük be a $C_n[x_1, x_2]$ jelölést, amely valós x_1, x_2 értékekre, ha $x_1 \leq x_2$, akkor az $[x_1, x_2]$ intervallumon folytonos függvények leképezése $\mathbb{R}^n$ -be.

Vizsgáljuk meg most az alábbi kezdeti érték problémát

$$(3.2) \quad \begin{aligned} y'(x) &= F(x, y), & (a < x \leq b) \\ y(x) &= q(x), & (\alpha \leq x \leq a) \end{aligned}$$

ahol $q \in C_n[\alpha, a]$ kezdő függvény és

$F: [a, b] \times C_n[\alpha, b] \rightarrow \mathbb{R}^n$ olyan, hogy

a/ rögzített y esetén $F(x, y)$ folytonos minden $x \in [a, b]$

b/ és kielégíti a Lipschitz feltételt y szerint, azaz

$$\|F(x, y_1) - F(x, y_2)\| \leq L \|y_1 - y_2\|^{[\alpha, x]}$$

minden $y_1, y_2 \in C_n[\alpha, b]$ és $x \in [\alpha, b]$ esetén,

ahol $\|\cdot\|$ bármelyik természetes R^n norma és

$$\|y_1 - y_2\|^{[\alpha, x]} = \max_{\alpha \leq s \leq x} \|y_1(s) - y_2(s)\|.$$

Fenti feltételek mellett a (3.2) funkcionálegyenletnek létezik egyértelmű megoldása /Driver /1962//. A későbbiekben mindig feltesszük, hogy ezek a feltételek teljesülnek.

Könnyen belátható, hogy az $X = b - \alpha$, $\bar{x} = b - \alpha$ helyettesítést alkalmazva

$$\varphi(x) = \begin{cases} g(a+x), & (-\alpha \leq x \leq 0) \\ g(a-\alpha), & (-\bar{x} \leq x < -\alpha) \end{cases}$$

$$z(x) = \begin{cases} \varphi(x), & (-\bar{x} \leq x \leq 0) \\ y(x+a), & (0 < x \leq X) \end{cases}$$

és $\phi(x, z_x) = F(x, y)$,

akkor z kielégíti a (3.1) egyenletet és $\phi: [0, X] \times C_n[-\bar{x}, 0]$ egy olyan folytonos leképezés, amelyre

$$\|\phi(x, u) - \phi(x, w)\| \leq L \|u - w\|$$

minden $x \in [0, X]$ és $u, w \in C_n[-\bar{x}, 0]$.

Igy beláttuk, hogy a (3.1) és (3.2) egymással ekvivalens feladatok. A továbbiakban a (3.2) jelölési módot fogjuk használni.

Minthogy az $F(x, y)$ funkcionál y szerint kielégíti a Lipschitz feltételt, így $F(x, y)$ független lesz az $y(s)$ jövőbeli értékeitől, azaz az $s > x$ értékektől. Így F un. Volterra funkcionál.

A gyakorlatban számos helyen fordulnak elő ilyen funkcionál DE-k. Röviden megemlítjük néhány speciális esetüket. Pl.

a/ a közönséges DE

b/ a retardált DE:

$$1/ \left. \begin{aligned} z'(x) &= f(x, z(x), z(x-\tau(x))), \quad (0 < x < X) \\ z_0 &= \varphi \in C_n[-\bar{x}, 0]. \end{aligned} \right\}$$

2/ vagy ennek egy rész esete, ha a késleltetés konstans

$$\left. \begin{aligned} z'(x) &= f(x, z(x), z(x-\bar{x})), \quad (0 < x < X) \\ z_0 &= \varphi \in C_n[-\bar{x}, 0], \end{aligned} \right\}$$

ahol $\bar{x} > 0$.

3/ egy másik részesete, ha f független $z(x)$ -től

$$\left. \begin{aligned} z'(x) &= f(x, z(x-\bar{x})), \quad (0 < x < X) \\ z_0 &= \varphi \in C_n[-\bar{x}, 0], \end{aligned} \right\} \text{ stb.}$$

ahol $\bar{x} > 0$.

c/ a Volterra típusu integro-differenciálegyenletek

$$\left. \begin{aligned} z'(x) &= f(z(x), \int_a^x K[y(s), s, x] ds, x) \\ z(0) &= \varphi(x) \in \mathbb{R}^n \end{aligned} \right\}$$

ahol a K magfüggvény folytonos és egyenletesen kielégíti a Lipschitz feltételt az s és az $y(s)$ szerint.

2. Általánosított egylépéses módszerek

A (3.2) típusu Volterra funkcionál differenciálegyenletek numerikus megoldási problémáival többen foglalkoztak. Itt most csak az egylépéses módszerekre vonatkozó eddigi eredményeket említjük meg.

Az Euler módszer konvergenciáját funkcionál DE-kre Cryer, Tavernini /1972/ vizsgálta. Tavernini /1971/ általánosított egylépéses módsezrekre bizonyított konvergencia tételt állandó h lépéshossz mellett.

A továbbiakban a fenti eredményt általánosítjuk változó lépéshossz esetén.

Definiáljunk a (3.2) egyenletre az $[(1.3), k=1]$ egy-
lépéses módszerek természetes általánosításával kapott közeli-
tő módszert, vezessük be az ilyen módszerek konzisztenciájá-
nak, konvergenciájának és képlethibájának fogalmát.

Tekintsük először a legegyszerűbb esetet az általánosí-
tott Euler módszert. Jelölje $\tilde{y}$ a (3.2) közelítő megoldását.
Ekkor

$$\begin{aligned}\tilde{y}(x_i + rh_i) &= \tilde{y}(x_i) + rh_i F(x_i, \tilde{y}), \\ (x_i \in \{\Delta_N\}, r \in [0, 1], 0 \leq i \leq N) \\ \tilde{y}(x) &= \tilde{g}(x), \quad (x \in [\alpha, a]).\end{aligned}$$

Ezt átírhatjuk a következő ekvivalens alakra

$$\left. \begin{aligned}\tilde{y}(x_{i+1}) &= \tilde{y}(x_i) + h_i F(x_i, \tilde{y}), \quad (x_i \in \{\Delta_N\}, r=1), \\ \tilde{y}(x_i + rh_i) &= r\tilde{y}(x_{i+1}) + (1-r)\tilde{y}(x_i), \quad (0 < r < 1).\end{aligned}\right\}$$

Igy látható, hogy az $\tilde{y}$ közelítő megoldás értékét az
osztópontokban a szokásos Euler módszerrel kapjuk és a többi
pontban pedig lineáris interpolációval.

Megjegyezzük, hogy mivel az F értelmezési tartománya
 $[a, b] \times C_n[\alpha, b]$, így a közelítő megoldást az $[\alpha, b]$
intervallum minden pontjában ki kell számítanunk, nemcsak az
osztópontokban.

Vezessünkbe most az Euler módszernél általánosabb össze-
függést.

3.1 Definíció Általánosított egylépéses módszernek
nevezzük a következőt

$$\begin{aligned}(3.3) \quad \tilde{y}(x_i + rh_i) &= \tilde{y}(x_i) + rh_i \phi(x_i, \tilde{y}, h_i, r), \\ (x_i \in \{\Delta_N\}, r \in [0, 1], 0 < i \leq N) \\ \tilde{y}(x) &= \tilde{g}(x), \quad (x \in [\alpha, a]),\end{aligned}$$

ahol a $\phi: S \times C_n[\alpha, b] \times [0, 1] \rightarrow \mathbb{R}^n$ funkcionál a mód-
szer növekményfüggvénye, és

$S = \{(x, h_i), \quad a \leq x \leq b \quad \text{és} \quad 0 < h_i \leq b - x \leq h^*\}$
a $\tilde{g}$ folytonos függvény pedig legyen g egy közelítése.

Továbbá mindig feltesszük, hogy

- a/ $\phi(x, y(x), h, r)$ folytonos r szerint ($r \in [0, 1]$)
minden rögzített x, y, h esetén
b/ és kielégíti a Lipschitz feltételt, azaz

$$(3.4) \|\phi(x, y_1, h, r) - \phi(x, y_2, h, r)\| \leq K \|y_1 - y_2\| [\alpha, \beta]$$

minden $y_1, y_2 \in C_n[\alpha, \beta]$, $(x, h \in S, r \in [0, 1])$ esetén,
 $K > 0$ konstans.

A közelítő megoldást itt a korábbiaktól eltérően szükségesnek láttuk megkülönböztetni és az $\tilde{y}$ jelölést bevezetni. Ezzel is hangsúlyozni kívánjuk, hogy ez más, függvényjellegű közelítés lesz, nemcsak diszkrét pontokban /az osztópontokban/ definiálja a (3.2) DE közelítő megoldását, hanem a köztük lévő pontokban is.

A (3.2) funkcionál DE-k (3.3) módszerrel történő megoldásakor definiáljuk az alábbi fogalmakat. Jelölje a (3.2) DE pontos megoldását y .

3.2 Definíció

A (3.3) általánosított egylépéses módszert konvergensnek nevezzük, ha tetszőleges $\{\Delta_N\}_{N=1}^{\infty} \subset \pi$ felosztás-sorozat esetén

$$\|y - \tilde{y}\| \rightarrow 0, \quad (\|\Delta_N\| \rightarrow 0, \|q - \tilde{q}\| \rightarrow 0).$$

3.3 Definíció

A (3.3) módszer képlethibája az $x_i + rh_i$ helyen

$$\tau(x_i, h_i, r) := y(x_i) + h_i \phi(x_i, y, h_i, r) - y(x_i + rh_i).$$

Tegyük most fel, hogy a ϕ olyan, amelyre a képlethiba kielégíti az alábbi feltételt

$$|\tau(x_i, h_i, r)| \leq h_i \varepsilon(x, r, h),$$

ahol $\varepsilon(x, r, h)$ egy megadott hibakorlát függvény, amely függ az x, r és h értékétől.

3.4 Definíció

A (3.3) módszert konzisztensnek nevezzük, ha

$$\phi(x, y, h, 1) \rightarrow F(x, y) \quad (\|\Delta_N\| \rightarrow 0)$$

egyenletesen X szerint, vagy /ekvivalens/

$$\varepsilon(x, 1, h) \rightarrow 0 \quad (\|\Delta_N\| \rightarrow 0)$$

minden $x \in [a, b]$ esetén.

A konzisztencia fogalmának talán természetesebb általánosításának tűnne, ha azt tennénk fel, hogy

$$\varepsilon(x, r, h) \rightarrow 0 \quad (\|\Delta_N\| \rightarrow 0)$$

minden $r \in [0, 1]$ értékre; de a (3.3) formulából nyilvánvaló, hogy az $\tilde{y}$ kiszámításában elkövetett hiba az x_i osztópontoknál egyszerűen összegződik, de az intervallum többi pontjában előbb beszorzódik az rh_i tényezővel.

3.5 Definíció

A (3.3) módszer globális hibájának nevezzük az

$$\|y - \tilde{y}\|_{[a, b]}$$

különbséget.

Vezessük be a következő jelölést

$$\nabla(x, \sigma) = \begin{cases} \frac{1}{\sigma}(y(x+\sigma) - y(x)), & \sigma > 0, \\ y'(x), & \sigma = 0, \end{cases}$$

minden $(x, \sigma) \in S$.

3.6 Definíció

A (3.3) módszert stabilisnak nevezzük, ha léteznek olyan $c, h^* > 0$ konstansok, hogy

$$\|y^* - \tilde{y}\|_{[a, b]} \leq C \{q^* - \tilde{q}\| + \sum_{i=1}^N \|\sigma_i\|\}$$

hacsak $\|\Delta_N\| \leq h^*$,

y^* pedig az alábbi perturbált rekurzió megoldása

$$(3.3') \quad y^*(x_i + rh_i) = y^*(x_i) + rh_i [\phi(x_i, y^*, h_i, r) + \sigma_{i+1}].$$

3. Az általánosított egy lépéses módszerek konvergenciája és stabilitása

Szükségünk lesz a következő tételre.

3.1 Tétel Ha

a/ a Φ növekményfüggvény olyan, hogy a képlethibára teljesül

$$(3.5) \quad \|\Phi(x, y, h, r) - \nabla(x, r, h)\| \leq \varepsilon(x, h, r)$$

minden $x \in \{\Delta_N\}$, $r \in [0, 1]$ és $0 < h \leq h^*$ esetén,

b/ ahol az ε h -ban monoton függvény és

$$(3.6) \quad \varepsilon(x, h, r) \leq \bar{\varepsilon}(x, h^*, r) \leq \varepsilon_1 \quad (\varepsilon_1 \text{ konstans})$$

c/ valamint

$$(3.7) \quad \bar{\varepsilon}(x, h^*, 1) \rightarrow 0, \quad (C \|\Delta_N\| \rightarrow 0)$$

akkor igaz az $[\alpha, b]$ intervallumon a következő becslés:

$$(3.8) \quad \begin{aligned} & \|y - \tilde{y}\|_{[\alpha, b]} \leq [\|g - \tilde{g}\|_{[\alpha, q]} + h^* \varepsilon_1 + \int_a^b \bar{\varepsilon}(x, h^*, 1) dx] \\ & (1 + (b-a)Ke^{K(b-a)}) \end{aligned}$$

Bizonyítás A (3.3) formula szerint $r = 1$ esetén

$$\tilde{y}(x_i) = \tilde{g}(a) + \sum_{j=0}^{i-1} h_j \Phi(x_j, \tilde{y}, h_j, 1)$$

$$\text{Az } y(x+h) = y(x) + h \nabla(x, h)$$

azonosságból

$$y(x_i) = g(a) + \sum_{j=0}^{i-1} h_j \nabla(x_j, h_j).$$

Legyen $\eta = y - \tilde{y}$. A fenti két összefüggést kivonva

és a háromszög egyenlőtlenséget alkalmazva kapjuk

$$\begin{aligned} & \|y(x_i) - \tilde{y}(x_i)\| \leq \|g(a) - \tilde{g}(a)\| + \\ & + \left\| \sum_{j=0}^{i-1} h_j [\Phi(x_j, \tilde{y}, h_j, 1) - \nabla(x_j, h_j)] \right\| \leq \\ & \leq \|\eta\|_{[\alpha, a]} + \sum_{j=0}^{i-1} \|\Phi(x_j, y, h_j, 1) - \nabla(x_j, h_j)\| + \\ & + \sum_{j=0}^{i-1} h_j \|\Phi(x_j, \tilde{y}, h_j, 1) - \Phi(x_j, y, h_j, 1)\| \end{aligned}$$

A (3.4) Lipschitz feltétel és a (3.5) feltétel miatt:

$$\begin{aligned} & \|\eta(x_i)\| \leq \|\eta\|_{[\alpha, a]} + \sum_{j=0}^{i-1} h_j \varepsilon(x_j, h_j, 1) + \\ & + k \sum_{j=0}^{i-1} h_j \|\eta\|_{[\alpha, x_j]} \end{aligned}$$

A tétel bizonyításához szükségünk lesz Gronwall-Bellman lemma alábbi diszkrét alakjára /Babuska /1966/, Schmidt /1976// Gronwall-Bellman lemma. Legyenek $x(i), y(i), z(i)$ az $i = 0, 1, \dots, N$ számokon értelmezett valós függvények és $z(i) \geq 0$ ($i = 0, 1, \dots, N$). Ha az $x(i), y(i), z(i)$ függvények kielégítik az

$$(3.9) \quad x(i) \leq y(i) + \sum_{j=0}^{i-1} z(j) x(j) \quad (i = 0, 1, \dots, n)$$

egyenlőtlenséget ($n \leq N$) akkor igaz, hogy

$$(3.10) \quad x(n) \leq y(n) + \sum_{i=0}^{n-1} z(i) y(i) \prod_{j=i+1}^{n-1} [1 + z(j)].$$

Vonjuk ki most a (3.3) formulát az $y(x+rh) = y(x) + rh \nabla(x, rh)$ azonosságból az $x = x_i$ pontban, hasonlóképpen kapjuk, hogy

$$\|\eta(x_i + rh_i)\| \leq \|\eta(x_i)\| + h_i K \|\eta\|^{[\alpha, x_i]} + h_i \varepsilon(x_i, h_i, r)$$

Irjuk be ide az $\eta(x_i)$ helyébe az előbb kapott korlátot

$$\begin{aligned} \|\eta\|^{[\alpha, x_{i+1}]} &\leq \|\eta\|^{[\alpha, a]} + h_i \varepsilon(x_i, h_i, r) + \sum_{j=0}^{i-1} h_j \varepsilon(x_j, h_j, 1) + \\ &+ K \sum_{j=0}^i h_j \|\eta\|^{[\alpha, x_j]} = P_{i+1} \end{aligned}$$

Ugyanakkor

$$\begin{aligned} \|\eta\|^{[\alpha, x_i]} &\leq \max \{ \|\eta\|^{[\alpha, a]}, P_1, \dots, P_i \} \quad \text{és mivel} \\ \|\eta\|^{[\alpha, a]} &\leq P_1 \leq \dots, P_i, \quad \text{így} \quad \|\eta\|^{[\alpha, x_i]} \leq P_i \end{aligned}$$

Alkalmazzuk a Gronwall-Bellman lemmát a következő szereposztással

$$\begin{aligned} x(i) &= \|\eta\|^{[\alpha, x_i]}, \quad z(i) = Kh_i, \\ y(i) &= \|\eta\|^{[\alpha, a]} + h_{i-1} \varepsilon(x_{i-1}, h_{i-1}, r) + \sum_{j=0}^{i-2} h_j \varepsilon(x_j, h_j, 1), \end{aligned}$$

akkor kapjuk, hogy

$$\|\eta\|^{[\alpha, x_N]} \leq \|\eta\|^{[\alpha, a]} + h_{N-1} \varepsilon(x_{N-1}, h_{N-1}, r) + \sum_{j=0}^{N-2} h_j \varepsilon(x_j, h_j, 1) +$$

$$+ \sum_{i=0}^{N-1} Kh_i \left[\|\eta\|^{[\alpha, a]} + h_{N-1} \varepsilon(x_{N-1}, h_{N-1}, r) + \sum_{j=0}^{N-2} h_j \varepsilon(x_j, h_j, 1) \right] \times$$

$$\times \prod_{j=i+1}^{N-1} (1 + Kh_j)$$



Az $1+x \leq e^x$ egyenlőtlenség miatt

$$\prod_{j=i+1}^{N-1} (1+kh_j) \leq \prod_{j=i+1}^{N-1} e^{kh_j} \leq e^{K(b-a)}.$$

Továbbá a (3.6) monotonitási feltételt felhasználva és azt, hogy létezik $h^* = \max_i h_i$

$$\begin{aligned} \|\eta\|_{[\alpha, b]} &\leq \|\eta\|_{[\alpha, a]} + h^* \varepsilon_1 + \sum_{j=0}^{N-2} h^* \bar{E}(x, h^*, 1) + \\ &+ \sum_{i=0}^{N-1} kh_i e^{K(b-a)} [\|\eta\|_{[\alpha, a]} + h^* \varepsilon_1 + \sum_{j=0}^{N-2} h^* \bar{E}(x, h^*, 1)] \end{aligned}$$

Összevonva a megfelelő tagokat és kihasználva, hogy létezik és korlátos az $\int_a^b \bar{E}(x, h^*, 1) dx$ Lebesgue integrál, kapjuk az alábbi becslést

$$\|\eta\|_{[\alpha, b]} \leq (1 + (b-a)Ke^{K(b-a)}) [\|\eta\|_{[\alpha, a]} + h^* \varepsilon_1 + \int_a^b \bar{E}(x, h^*, 1) dx] \quad Q.E.D.$$

A fenti becslés segítségével egyszerűen belátható a következő tétel.

3.2 Tétel

Ha a (3.3) módszer

a/ konzisztens,

b/ a ϕ növekmény függvény egyenletesen korlátos t és r szerint, minden $h^* \rightarrow 0$ esetén, azaz $\|\phi(x, y, h^*, r)\| \leq M$,

/ahol M konstans/

akkor a (3.3) módszer konvergens minden $\{\Delta_N\}_{N=1}^{\infty}$ felosztássorozatra, ha $\|\Delta_N\| = \max_{0 \leq i \leq N-1} h_i = h^* \rightarrow 0$ és $\|\tilde{g} - g\| \rightarrow 0$.

Bizonyítás

Vizsgáljuk meg, hogy a módszer konzisztenciájából és a ϕ egyenletes korlátosságából hogyan következnek az előző tétel feltételei. Azaz becsüljük az alábbi mennyiséget

$$\begin{aligned} &\|\phi(x, y, h, r) - \nabla(x, rh)\| \leq \|\phi(x, y, h, r) - \\ &- \phi(x, y, h, 1)\| + \|\phi(x, y, h, 1) - \nabla(x, rh)\| \\ &\leq \begin{cases} 2M & r \in [0, 1) \\ \bar{E}(h^*) \rightarrow 0 & r = 1 \end{cases} \quad (h^* \rightarrow 0) \end{aligned}$$

Ezért alkalmazható a (3.8) becslés, azaz

$$\|y - \tilde{y}\|^{[\alpha, b]} \leq [\|q - \tilde{q}\|^{[\alpha, a]} + h^* 2M + \int_a^b \bar{E}(x, h^*, 1) dx] * \\ * (1 + (b-a) K e^{K(b-a)})$$

Igy

$$\|y - \tilde{y}\|^{[\alpha, b]} \rightarrow 0, \text{ amikor } h^* \rightarrow 0 \text{ és } q \rightarrow \tilde{q} \quad \text{Q.E.D.}$$

3.3 Tétel A (3.3) alakú egy lépéses módszerek stabilisek.

Bizonyítás

A (3.3) és (3.3') formula szerint $r=1$ esetén kapjuk, hogy

$$y^*(x_i) = q^*(a) + \sum_{j=0}^{i-1} h_j [\phi(x_j, y^*, h_j, 1) + \delta_{j+1}]$$

$$\tilde{y}(x_i) = \tilde{q}(a) + \sum_{j=0}^{i-1} h_j \phi(x_j, \tilde{y}, h_j, 1).$$

Legyen $\eta = \tilde{y} - y^*$. Vonjuk ki egymásból a fenti két egyenlőséget és használjuk fel a háromszög egyenlőtlenséget és azt, hogy a ϕ kielégíti a Lipschitz feltételt y szerint.

$$\text{Akkor } \|\eta(x_i)\| = \|y^*(x_i) - \tilde{y}(x_i)\| \leq \|q^*(a) - \tilde{q}(a)\| + \sum_{j=0}^{i-1} h_j \|\delta_{j+1}\| + \\ + \sum_{j=0}^{i-1} h_j \|\phi(x_j, y^*, h_j, 1) - \phi(x_j, \tilde{y}, h_j, 1)\| \leq$$

$$\leq \|\eta\|^{[\alpha, a]} + \sum_{j=0}^{i-1} h_j \|\delta_{j+1}\| + \sum_{j=0}^{i-1} h_j K \|y^* - \tilde{y}\|^{[\alpha, x_j]}$$

Vonjuk ki most a (3.3) és (3.3') formulát egymásból

$$\|\eta\|^{[x_i, x_{i+1}]} = \|y^*(x_i + rh_i) - \tilde{y}(x_i + rh_i)\| \leq \|y^*(x_i) - \tilde{y}(x_i)\| +$$

$$+ h_i \|\phi(x_i, \tilde{y}, h_i, r) - \phi(x_i, y^*, h_i, r)\| + h_i \|\delta_{i+1}\| \leq$$

$$\leq \|\eta\|^{[\alpha, a]} + K \sum_{j=0}^i h_j \|\tilde{y} - y^*\|^{[\alpha, x_j]} + \sum_{j=0}^i h_j \|\delta_{j+1}\| = P_{i+1}$$

$$\|\eta\|^{[\alpha, x_i]} \leq \max\{\|\eta\|^{[\alpha, a]}, P_i\} \quad \text{és mivel}$$

$$\|\eta\|^{[\alpha, a]} \leq P_1 \leq \dots \leq P_i, \text{ így } \|\eta\|^{[\alpha, x_i]} \leq P_i$$

A Gronwall-Bellman lemmát alkalmazva az alábbi szereposztásban

$$x(i) = \|\eta\|^{[\alpha, x_i]}, \quad z(i) = Kh_i,$$

$$y(i) = \|\eta\|^{[\alpha, a]} + \sum_{j=0}^i h_j \|\delta_{j+1}\|,$$

kapjuk, hogy

$$\|\eta\|^{[\alpha, b]} \leq \|\eta\|^{[\alpha, a]} + \sum_{j=0}^N h_j \|\delta_{j+1}\| + \sum_{i=0}^{N-1} Kh_i \left[\|\eta\|^{[\alpha, a]} + \sum_{j=0}^i h_j \|\delta_{j+1}\| \right].$$

$$\prod_{j=i+1}^{N-1} (1 + Kh_j).$$

$$\text{Az } 1+x \leq e^x$$

összefüggés miatt:

$$\prod_{j=i+1}^{N-1} (1 + Kh_j) \leq \prod_{j=i+1}^{N-1} e^{Kh_j} \leq e^{K(b-a)}.$$

Igy

$$\|\eta\|^{[\alpha, b]} \leq \|\eta\|^{[\alpha, a]} + h^* \sum_{j=0}^N \|\delta_{j+1}\| + Ke^{K(b-a)} \sum_{i=0}^{N-1} \left[h_i \|\eta\|^{[\alpha, a]} + \sum_{j=0}^i h_j \|\delta_{j+1}\| \right]$$

$$\leq \|\eta\|^{[\alpha, a]} (1 + K(b-a)e^{K(b-a)}) + \sum_{j=0}^N \|\delta_{j+1}\| (h^* + Kh^* e^{K(b-a)})$$

$$\leq (C_1 + K(b-a)e^{K(b-a)}) (\|\eta\|^{[\alpha, a]} + \sum_{j=0}^N \|\delta_{j+1}\|),$$

$$\text{ahol } C_1 = \max(1, h^*).$$

Q.E.D.

zer am neve	FORMULA	Hibabecslés	Döntési kritérium /lépésfogadásának feltétele/ lépéshossz /rend/ változtatási mechanizmus
eveld 64/	Ötödrendű RK: $y_{n+1}=y_n+(35k_0+162k_2+125k_4+14k_5)/336$; $k_0=hf(y_n)$, $k_1=hf(y_n+2/9 k_0)$, $k_3=hf(y_n+(k_0+3k_2)/8)$, $k_4=hf(y_n+(53k_0-135k_1+126k_2+56k_3)/125)$, $k_5=hf(y_n+(-63k_0+189k_1-36k_2-112k_3+50k_4)/28)$.	az y_n h szerinti Taylor sorfejtésének 5. tag $th^5 dy = (21k_0 - 162k_2 + 224k_3 - 125k_4 + 42k_5)/14$; $k_6 = hf(y_n + (133k_0 - 378k_1 + 276k_2 + 112k_3 + 25k_4)/168)$	$th^5 dy (abs(f \cdot \text{relativ hibakorlát} + absz.h.k.) \cdot abs(h)/(b-a))$ 1. számolunk h lépéshosszal; 2. kiszámítjuk a $h_{uj} =$ $= 0.95 h \sqrt[5]{\epsilon/abs(th^5 dy)}$ 3. kétlépés elfogadásáig 2./ alapján számolunk; 4. kezdetben vagy elutasítás után $h_{uj} = mu_1 \cdot h$, $mu_1 = \min(x_i) / (\min(x_i) + abs(th^5 dy(x_i)) +$ $= 0.45$; 5. különben $h_{uj} = h / (h/h_{előző} + 1) \cdot mu_1 - mu_0$
/	nem stiff DE-re: Adams-Basforth-Moulton /1.10/ azonos rendű prediktor-korrektor pár /1-13-ad rendig/. stiff DE-re: /BDF/ retrográd differenciálás formula /1-6-ad rendig/	főhibatag súlyozott becslése osztott differen ciákkal $/p-1/-ed$ rendre: $c_p \cdot p! \cdot a_p/\omega _2 = \alpha_{p-1}$, mivel $a_p \approx \frac{h}{p!} y^{(p)}$ $p-ed$ rendre: $C_{p+1} \cdot \nabla a_p \omega _2$ $/p+1/-ed$ rendre: $C_{p+2} \cdot p! \cdot \nabla^2 a_p \omega _2 = \alpha_{p+1}$, mivel $\nabla a_p = \frac{h}{p!} y^{(p+1)}$, $a = [y, y', \dots, h^p y^{(p)} / p!]^T$, ∇a_p az utolsó komponens retrográd differenciája; ω súlykomponens, $ \cdot _2$ norma	$\alpha_p \leq \epsilon$ /megadott hibakorlát/. lok.opt.lépéshossz: $C_{p+1} \cdot p! \cdot \alpha_{p+1} \cdot \nabla a_p \omega _2 = \epsilon$; $h_{uj} = 0.99 \cdot h$, ahol $\alpha = \frac{1}{1.2} \left[\frac{\epsilon}{\alpha_p} \right]^{1/p+1}$ $p-ed$ rendre; $\alpha = \frac{1}{1.4} \left[\frac{\epsilon}{\alpha_{p-1}} \right]^{1/p}$ $/p-1/-ed$ rendre, $\alpha = \frac{1}{1.3} \left[\frac{\epsilon}{\alpha_{p+1}} \right]^{1/p+2}$ $/p+1/-ed$ rendre; α -t számoljuk, ha: 1. a lépés sikertelen volt, kivéve, ha éppen a rendet akartuk növelni; 2. $/p+1/$ lépéssel az utolsó lépéshossz, ill. rendváltoztatás után; 3. tíz lépés sel az utolsó α becslés után, ha csak közben nem növeltük a lépéshosszt, a legnagyobb α -nak megfelelő rendű módszert választjuk, h is változik
berg 72/	Implicit érintő formula: $y_{n+1} = y_n + hf((y_n + y_{n+1})/2)$, és passzív simítás az $x_i = x_0 + ih_0$ $/i=1,2,\dots/$ pontokban: $(h=h_0, h_0/2)$ $\hat{y}_n(h) = [y_{n-1}(h) +$ $2y_n(h) + y_{n+1}(h)]/4$; passzív extrapoláció $\hat{y}_n(h_0) = y_n(h_0/2) + [y_n(h_0/2) - \hat{y}_n(h)]/3$ /lépéshosszváltásnál így indul/	globális hiba becslése; a lokális hiba becslés lépésfelezéssel és extrapolációval $est = c_n \cdot h^5 \approx 1/36 \cdot \Delta_n^3 [\hat{y}_n(h/2) - \hat{y}_n(h)] $.	$est \leq \epsilon_{Lok}$, $\epsilon = 2 \cdot \frac{\sqrt{12\epsilon}}{ x } \cdot e^{-\frac{2\sqrt{12\epsilon}}{ x }}$ est/ϵ_{Lok} $< 1/5110$ $\frac{h_{uj}}{(5\sqrt{\epsilon_{Lok}}/5 \cdot est) \cdot h}$ $1/5110, 1/80$ $2 \cdot h$ $1/20, 1/2$ változatlan > 1 az utolsó lépés újra
mpine, ts 76/	4-ed és 5-öd rendű Fehlberg form. /1969/ $y_{n+1} = y_n + h(16/135 k_1 + 6656/12825 k_3 +$ $+ 28561/56430 k_4 - 9/50 k_5 + 2/55 k_6)$ $5-öd$ $y_{n+1} = y_n +$ $+ h(25/216 k_1 + 1408/2565 k_3 + 2197/4104 k_4 - 1/5 k_5)$; $k_1 = f(x_n, y_n)$; $k_2 = f(x_n + 1/4 h, y_n + 1/4 k_1)$; $k_3 = f(x_n + 3/8 h,$ $y_n + 3/32 k_1 + 9/32 k_2)$ $k_4 = f(x_n + 12/13 h, y_n + 1932/2197$ $* k_1 - 7200/2197 k_2 + 7296/2197 k_3)$ $k_5 = f(x_n + h, y_n + 439/216$ $* k_1 - 8k_2 + 3680/513 k_3 - 845/4104 k_4)$; $k_6 = f(x_n + 1/2 h; y_n -$ $8/27 k_1 + 2k_2 - 3544/2565 k_3 + 1859/4104 k_4 - 11/40 k_5)$	lokális hiba becslése (beágyazási típusú módszer) $est = h(-1/360 k_1 + 128/4275 k_3 + 2197/75240 k_4$ $- 1/50 k_5 - 2/55 k_6)$.	$est \leq y \cdot \text{rel.hibakorlát} + absz.h.k. = \epsilon_{összes}$ /relativ h.k.a lépés kezdő és végpontjában/. $h_{uj} = 0.9 \sqrt[5]{\frac{\epsilon}{est}} h$; $h_{min} = 0.26 \cdot h$ /ahol u az adott számítógépen legkisebb pozitív szám/, sikertelen lépés után: a lépéshossz nem növelhető; az $1/10$ -nél nem jobban lecsök- kentett lépéshosszal megismétli a lépést; teljesülnie kell a $h_{uj} \geq h_{min}$ feltételnek
ner 76/ FI	negyedrendű Hermite-Obreschkoff formula $y_{n+1} - h/2 y'_{n+1} + h^2/6 y''_{n+1}/2! =$ $= y_n + h/2 y'_n + h^2/6 y''_n/2!$	főhibatag becslése $est = h^5/6 y^{(5)}/5!$ lépésfelezéssel osztott differenciák segítségével.	$est = \epsilon / \text{adott hibakorlát}$; $h_{uj} = \sqrt[5]{6\epsilon / \max_{i,n} y_i[x_{n-2}, x_{n-2}, x_{n-1}, x_{n-1}, x_n, x_n] }$, ha $est \geq 20 \epsilon$, akkor a h_{uj} lépéshosszal megismétli a lépést; különben fenti formula szerint változtatja h -t, de nem engedi meg, hogy öt- szörösénél nagyobbra nőjön.
der	diagonálisan implicit RK-módszer /együtthatómátrix alsó háromszög ti- pusú azonos diagonális elemekkel/ pl. 2-pontos 2-od rendű: $y_{n+1} = y_n + (1-\alpha)k_1 + \alpha k_2$; ahol $k_1 = hf(x_n + \alpha h, y_n + \alpha k_1)$, $k_2 = hf(x_n + h, y_n + (1-\alpha)k_1 + \alpha k_2)$; $(\alpha = 1 \pm \frac{1}{2} \sqrt{2})$.	főhibatag súlyozott becslése az RMS normában ($ y = 1/m \sum_{i=1}^m (y^{(i)})^2 / y_{max}^2$) m-pontos formulára). $ y_{n+1} - y_{n+2}/2 / (2^p - 1)$	ha $est > \epsilon$, a lépést elutasítjuk és h -t csökkentjük úgy, hogy a várt hiba $\approx \epsilon/5$ legyen $3\epsilon/4 < est \leq \epsilon$, a lépést elfogadjuk, de h -t fenti módon lecsökkentjük; $\epsilon/10 < est \leq 3\epsilon/4$, a lépést elfogadjuk és ugyanazzal szá- mol tovább; ha $est < 1/10$, a lépést elfogadjuk és h -t megnöveljük úgy, hogy a várt hiba $\epsilon/2$ legyen; feltéve, hogy a/ legalább $p+1$ sikeres lépést tettünk a h utolsó csökkentése óta; b/ csökkentés után h legfeljebb kétszeresére nőtt, mialatt a legnagyobb megengedett növe- kedés 10-szeres; c/ a növekedés legalább 1.3-szoros.
	m-pontos semi implicit Rosenbrock formula /1963/: $y_{n+1} = y_n + h \sum_{r=1}^m c_r k_r$; $k_1 = f(y_n) + \alpha_1 h \frac{\partial f}{\partial y}(y_n) k_1$; $k_r = f(y_n + h \sum_{s=1}^{r-1} b_{rs} k_s + \alpha_r h \frac{\partial f}{\partial y}(y_n + h \sum_{s=1}^{r-1} b_{rs} k_s) k_r$ $/r=2,3,\dots,m/$ $c_i, \alpha_i, \beta_{ij}, b_{ij}$ konstans; 3-ad rendű 3-pontos eset.	a lokális hiba főtagjának becslése lé- pésfelezéssel /est/	a/ $\max est_{n+2} _j > \epsilon \cdot h_{uj} = h/2$ és megismétli a lépést $/x_n, y_n/$ -től; $/j=1,2,\dots,m$ egyenletek száma/; b/ $\max est_{n+2} _j < \epsilon/25$ $h_{uj} = 2h$ lépés- hosszal folytatja az $/x_{n+2}, y_{n+2}/$ ponttól; c/ különben h rögzített és folytatjuk tovább az $/x_{n+2}, y_{n+2}/$ ponttól
	/BDF/ retrográd differenciálási formula /harmadrendű/	lokális hiba főtagjának becslése /est/	$125 est < \epsilon \cdot h_{uj} = 4h$; $16 est < \epsilon \leq 125 est \cdot h_{uj} = 2h$; $1/16 est < \epsilon \leq 16 est \cdot h_{uj} = h$; $\epsilon \leq 1/16 \cdot est \cdot h_{uj} = h/2$;

1. táblázat

Néhány ismert módszer leírása

IRODALOM

1. Alexander R. /1977/ Diagonally implicit Runge-Kutta methods for stiff O.D.E.'s
SIAM J. Numer. Anal. V.14. N.6. Dec. 1977-1978
2. Babuska I., Práger M., Vitasek E. /1966/ Numerical processes in DE's
Wiley /Interscience/ London
3. Bulirsch R. and Stoer J. /1966/ Numerical treatment of O.D.E.'s by extrapolation methods
Num. Math. 8. 1-13.
4. Butcher J.C. /1966/ On the convergence of numerical solutions to O.D.E.'s
Math. of Comput. 20. 1-10.
5. Cash J.R., Liem C.B. /1977 / On the computational aspects of semi-implicit Runge-Kutta methods
The Computer Journal V.24. N.4. 363-365
6. Ceschino F. and Kuntzman J. /1966/ Numerical solution of Initial Value Problems
Prentice Hall, Englewood Cliffs
7. Coddington, Levinson /1955/ Theory of O.D.E.'s
Mc Graw-Hill Book Comp. Inc. New York
8. Dahlquist G. /1956/ Convergence and stability in the numerical integration of O.D.E.'s
Math. Scandinavica 4. 33-53
9. Driver R.D. /1962/ Existence and stability of solutions of a delay-differential system
Arch. Rational Mech. Anal. 10. 401-426.

10. England, R. /1969/ Error estimates for Runge-Kutta type solutions to systems of O.D.E's
Computer Journal 12. 166-169.
11. Fehlberg, E. /1969/ Klassische Runge-Kutta Formeln funfter und siebenter Ordnung mit Schrittweiten Kontrolle Computing, 4, 93-106.
12. Galántai A. /1978/ Vizsgálatok a közönséges differenciálegyenletek közelítő módszereinek konvergencia és hibaanalízisének körében
Kandidátusi disszertáció
13. Galántai A. - Strehó M. /1977/ Stiff problems I. ELTE Numerikus és Gépi Matematikai Tanszék kiadványa
Budapest
14. Gear, C.W. /1971/ Numerical Initial Value Problems in O.D.E's
Prentice Hall, Englewood Cliffs, New Jersey
15. Gear C.W. Tu K.W. /1974/ The effect of variable mesh size on the stability of multistep methods
SIAM J. Numer. Anal. V. 11. N.5. October 1025-1043.
16. Gear C.W., Watanabe D.S. /1974/ Stability and convergence of variable order multistep methods
SIAM J. Numer. Anal. V.11. N.5. October 1044-1058.
17. Gill, S. /1951/ A process for the step-by-step integration of DE's in an automatic computing machine
Proc. Cambridge Philos. Soc. 47. 96-108.
18. Hahn, W. /1967/ Stability of Motion
Springer Verlag

19. Hale J. /1971/ Functional Differential Equations
New York, Springer
20. Hall, G., Watt, J.M. /1976/ Modern numerical methods
for O.D.E's
Clarendon Press, Oxford 1976.
21. Henrici, P. /1962/ Discrete Variable Methods in
Ordinary Differential Equations Wiley New York
22. Hull, T.E. /1969/ The numerical integration of O.D.E's
Information Processing 68 - North-Holland Publ.Comp.
Amsterdam 40-53.
23. Hull, T.E., Enright, W.H., Fellen, B.M. and Sedgwick
A.E. /1972/ Comparing numerical methods for O.D.E's
SIAM J. Numer. Anal. 11. 965-978
24. Jackson, K.R., Enright W.H., Hull T.E. /1978/
A theoretical criterion for comparing Runge-Kutta
formulas
SIAM J. Numer. Anal. V.15. N.3. June 618-641.
25. Krogh, F.T. /1973/ Algorithms for changing the step
size
SIAM J. Numer. Anal. 10. 949-965.
26. Lindberg, B. /1972/ IMPEX-A program package for
solution of systems of stiff differential equations
Report NA 72.50, The Royal Institute of Technology
Stockholm Sweden
27. Nordsieck A. /1962/ On the numerical integration of
O.D.E's
Math. Comp. 16. 22-49.

28. Numerikus módszerek programgyűjteménye I.II.III.
MTA SZTAKI 1971. 1973. 1976.
29. Pontrjagin, Boltyanskij /1968/ Optimális folyamatok
elmélete
Közgazdasági és Jogi Könyvkiadó, Budapest
30. Rosenbrock H.H. /1963/ Some General Implicit
Processes for the Numerical Solution of D.E.'s
Comp. J. 5. 329-330.
31. Sacks-Davis R. /1977/ Error estimates for a stiff
Differential Equation Procedure
Mathematics of Computation V.31. N.140. Oct. 939-953.
32. Shintani, H. /1966/ Two step processes by one-step
methods of order 3 and of order 4
Ibid 183-195.
33. Sarafyan D. /1968/ Estimation of errors for the
approximate solution of differential equations and
their systems, Riv. Mat. Univ. Parma 9. 109-127.
34. Schmidt, J.W. /1976/ Über lineare Ungleichungen vom
Gronwallischen Typ
Beiträge zur Num. Math. 5. 171-189.
35. Shampine, L.F. /1973/ Local extrapolation in the
solution of O.D.E's
Mathematics of Computation 27. 91-97.
36. Shampine L.F. and Watts H.A. /1976/ Practical
Solution of O.D.E's by Runge-Kutta methods
SAND 76-0585 U.S.A.

37. Spijker, M.N. /1966/ Convergence and stability of step-by-step methods for the numerical colution of initial value problems
Numerische Mathematics 8. 161-177.
38. Stetter, H.J. /1971/ Stability of discretizations on infinite intervals
Lecture Notes in Mathematics 228. Springer Verlag 207-222.
39. Stetter, H.J. /1973/ Analysis of Discretization Methods for O.D.E's
Springer Verlag
40. Strehó M. /1976/ Stiff tipusu közönséges differenciál-egyenletek megoldásáról
SZTAKI tanulmányok 48.
41. Tavernini L. /1971/ One step methods for the numerical solution of Volterra funct. diff. equations
SIAM J. Numer. Anal. 8. /1971/ 766-795.
42. Пухан А. Ф. (1960) Теория продолжения функций действительного переменного
Гос. издат. физ. мат. лит. Москва
43. Varga R.S. /1961/ On higher order stable implicit methods for solving parabolic partial D.E.'s
Journ. Math. and Phys. 40, 220-231.
44. Varga R.S. /1971/ Functional analysis and approximation theory in numerical analysis
Regional conference Series in applied Mathematics, SIAM

45. Wanner G. /1976/ STIFFI, A program for O.D.E's
Report, Universite de Geneve Oct.
46. Verner J.H. /1978/ Explicit Runge-Kutta methods with
estimates of the local truncation error
SIAM J. Numer. Anal. V.15. N.4. August 772-790.
47. Zlatev Z. /1979/ Application of backward differentia-
tion methods to the finite element solution of time-
dependent problems
Intern. J. for Num. Meth. in Engin. V.14. 1051-1061.
48. Zonneveld J.A. /1964/ Automatic Numerical Integ-
ration
Mathematisch Centrum, Amsterdam.
49. Cryer, C.W., L. Tavernini /1972/ The numerical
solution of Volterra functional differential
equations by Euler's method
SIAM J. Numer. Anal. 9. 105-129.
50. Sedgwick A.E. /1973/ An effective variable order
variable step Adams method
Ph. D. thesis, Dept. of Computer Science Tech. Rep.
N. 53, University of Toronto, Toronto.

