

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

KÁLMÁN DÁVID SÁMUEL

Exponenciális Runge–Kutta
módszerek

BSc Szakdolgozat
Matematika BSc, Matematika elemző szakirány

Témavezető:
Dr. Csomós Petra
Egyetemi adjunktus
Alkalmazott Analízis és Számításmatematikai Tanszék



Budapest, 2018

Tartalomjegyzék

Köszönetnyilvánítás	2
Előszó	3
Bevezetés	5
1. Az exponenciális Runge–Kutta módszer levezetése	7
1.1. Levezetéshez alaprendszer felírása	7
1.2. Rácsfüggvény és rácsháló létrehozása	9
1.3. Konstans variációs formula felhasználása	10
1.4. Lagrange-interpoláció ismertetése	13
1.5. Közelítő polinom függvény létrehozása	15
1.6. Polinom függvény alkalmazása	18
2. Az exponenciális Runge–Kutta módszer átírása és az együtthetők tulajdonságai	23
2.1. Az exponenciális RKM felírása Taylor-sorral	23
2.2. A B_i tulajdonságai	27
2.3. Az $A_{i,j}$ tulajdonságai	28
3. Módszerek	29
3.1. A közöséges Runge–Kutta módszer	29
3.2. Explicit exponenciális Runge–Kutta módszer	30
3.3. Exponenciális Euler-módszer	32
4. Az exponenciális Euler-módszer tulajdonságai	34
4.1. A merevség ismertetése	34
4.2. A konvergencia ismertetése	35
4.3. A módszer konvergenciája	36
4.4. A konzisztencia ismertetése	42
4.5. A módszer konzisztenciája	42
4.6. Példák számolással és MATLAB kódokkal	44
Összefoglalás	55
Felhasznált jelölések	56
Irodalomjegyzék	57

Köszönetnyilvánítás

Elsősorban szeretnék köszönetet mondani Csomós Petra témavezetőmnek, aki segített a téma kiválasztásában és rengeteg segítséget nyújtott a szakdolgozatom megírásánál. Továbbá köszönöm, hogy körültekintően átnézte a szakdolgozatomat és észrevételeivel, tanácsaival, segítségével és türelmével segítette a megírását.

Köszönettel tartozom Anyukámnak és a barátaimnak az egyetemi évek és a szakdolgozat írása alatt kapott támogatásért és biztatásért.

Köszönöm Faragó István tanár úrnak az Alkalmazott analízis előadásokat, mert felkeltette a téma iránt az érdeklődésemet.

Köszönöm továbbá mindazoknak is, akik támogattak.

Előszó

A szakdolgozatom témaválasztását elsősorban az Alkalmazott analízis előadáson elhangzott Runge–Kutta típusú módszerek és ahhoz kapcsolódó tulajdonságaik motiválták.

A szakdolgozatom teljes egészében törekedtem az egységes jelölés használatára. Ez segíti az olvasót abban, hogy amit egyszer definiáltunk, az a szakdolgozat egészében igaz legyen. Ez a kevés helyen amikor eltér, azt külön jelöltem, hogy abban a részben más lesz. A felhasznált jelöléseket a könnyebb áttekinthetőség érdekében kigyűjtöttem egy külön fejezetbe. Ezáltal nem kellett minden egyes részben újra leírni a jelölések értelmét, ugyanakkor a „Felhasznált jelölések” fejezetben könnyen visszakereshetőek. Ezt a fejezetet több részre bontottam, ami miatt az adott jelölést még könnyebben meg lehet találni.

A szakdolgozatban nagy hangsúlyt fektettem a lépésenkénti levezetésre, ami az érthetőséget nagyban segíti. Az felhasznált irodalmakban ezek a lépések többnyire összevontak, amik egy rész megértését nehezebbé teszi. A kiírások által, az irodalmak is könnyebben érthetőek. Amikor egy új jelölést bevezettem, ott használtam a „ : ” jelet. A betűtípusuk is különböző jelölés-csoportokat jelentenek, amit a következő táblázat mutat:

TARTALOMJEGYZÉK

Típusok	Angol betűk	Görög betűk
1-D függvények	kisbetű	kisbetű
több-D függvények	Nagybetű	Nagybetű
változók	kisbetű	kisbetű
mátrixok/vektorok	Nagybetű	kisbetű
konstansok	Nagybetű	Nagybetű

A függvények többnyire több dimenziósak, ahogy a mátrixok/vektorok is, ezért a több dimenziós függvények könnyebb elképzelése érdekében, ahol szükségesnek éreztem, ott kiírtam a függvényeket és leképezéseiket koordinátánként is. A szakdolgozat céljai közt szerepelt, hogy a részletesség és egységes jelölés rendszere segítségével, a témában kevésbé jártas emberek számára is érthető legyen.

Az irodalom hivatkozásokat a fejezet elején megadtam.

A Runge–Kutta módszer széleskörben feldolgozott téma, ezért a szakdolgozatom témaválasztása az exponenciális Runge–Kutta módszerek. Ennek a módszernek egy aloszata az órán taglalt Runge–Kutta módszer.

Bevezetés

A Runge–Kutta módszert az 1900-as évek elején vezette be Carl Runge és Martin Kutta német matematikus. Ez a módszer még nem volt alkalmas olyan típusú problémák megoldására, amik tartalmaztak exponenciális tagot, mert a számolás hosszadalmas volt. Az exponenciális tagot is tartalmazó, exponenciális Runge–Kutta módszereket az 1950-es években fedezték fel. 1958-ban Hersch kezdett velük részletesebben foglalkozni, aki egy másik megoldási módszert készített erre, ami először tartalmazott exponenciális tagot. 1960-ban Certaine a nemlineáris részt közelítette a konstans variációs formulával. 1963-ban Pope lefektette a Rosenbrock módszerek alapját. 1967-ben Lawson a Runge–Kutta módszerrel az együtthatók megváltoztatásával megoldott egy exponenciális tagot tartalmazó differenciálegyenletet. Ez volt az első alkalom, hogy a Runge–Kutta módszerrel megoldottak egy ilyen problémát. 1978-ban Friedli magasabb rendű módszerekkel is megoldotta a nem merev differenciálegyenletet a Taylor-sor használatával. 1998-ban Hochbruck, Lubich és Selhofer bevezetett egy új integrálási módszert, ami javított a számítási sebességen. 2005-ben Hochbruck és Ostermann elkészítette a magasabb rendű exponenciális Runge–Kutta módszert merev egyenletekre is. Ebben a forrásban található részletesen a történet [7].

Az exponenciális hatványzásokor nagyon sok mátrix szorzást kellett elvégezni, és mivel nem állt rendelkezésre kellő számítási teljesítmény, ezért a módszert nem használták sokáig. A számítógépek fejlődésével, az algoritmusok javításával az e

szám mátrix hatványára emelése és az új módszerekkel való hatványozás feltalálásával felgyorsult, így ezt a módszert újra elővették és elkezdték használni.

Ezt a numerikus közelítési módszert használják például a kémiai modell alkotásnál, reakció-diffúzió egyenleteknél, a nemlineáris Schrödinger-egyenletnél, a merev kinetikus egyenletekre, stb. Ez a módszer a szemilineáris egyenleteket hatékonyan tudja megoldani. A szemilineáris egyenletek egy lineáris merev részből és egy nemlineáris nem merev részből állnak. A merev részek megoldása hatékonyan csak implicit módszerekkel lenne lehetséges, de az exponenciális Runge–Kutta módszerrel ennek a merev résznek az értékét pontosan kiszámoljuk, ezért az explicit módszereket használhatjuk.

A szakdolgozatban megismerjük a módszer levezetését, megvizsgáljuk egy speciális exponenciális Runge–Kutta módszert, az exponenciális Euler tulajdonságait, ezt példákkal és ábrákkal illusztráljuk, amiken jól látszódik a numerikus közelítő módszer hatékonysága. A példákat, ábrákat a MATLAB programban elkészített saját fejlesztésű kódokkal készítettük.

1. fejezet

Az exponenciális Runge–Kutta módszer levezetése

Ebben a fejezetben levezetjük az exponenciális Runge–Kutta módszert, amivel meg tudunk oldani elsőrendű közönséges r változós szemilineáris differenciálegyenlet rendszereket, egy kezdeti feltétel ismeretében. Ennek teljesítéséhez közelíteni fogjuk a pontos megoldást, majd a konstans variációs formulát felhasználjuk, végül a Lagrange-interpolációval az ismeretlen nemlineáris függvényt közelítve felírjuk a módszert. Az alábbi irodalmakban találhatóak a felhasznált információk [1] [2] [4] [6].

1.1. Levezetéshez alaprendszer felírása

Legyen $r \in \mathbb{N}^+$, 0 a kezdeti hely, aminek az értéke $U_0 \in \mathbb{R}^r$. Legyen $q \in \mathbb{R}^+$, $\mathbb{T} = [0, q]$, és $F : \mathbb{T} \times \mathbb{R}^r \rightarrow \mathbb{R}^r$ folytonosan differenciálható függvény, ahol $t \in \mathbb{T}$ és $U_1, U_2 \in \mathbb{R}^r$. Keressük azt az $U : \mathbb{T} \rightarrow \mathbb{R}^r$ folytonosan differenciálható ismeretlen r dimenziós lineáris függvényt, amire teljesül az alábbi feladat, ami egy elsőrendű közönséges r változós differenciálegyenlet rendszer kezdetiérték feladata,

avagy Cauchy-feladat:

$$\begin{cases} U'(t) = F(t, U(t)) , t > 0 \\ U(0) = U_0 . \end{cases} \quad (1.1.1)$$

Definíció. Azt mondjuk, hogy az F a második változójában Lipschitzes, amennyiben minden $(t, U_1), (t, U_2) \in \mathbb{T} \times \mathbb{R}^r$ pontpárra teljesül az alábbi, ahol $L \in \mathbb{R}^+$ konstans:

$$\|F(t, U_1) - F(t, U_2)\|_{\mathbb{R}^r} \leq L \|U_1 - U_2\|_{\mathbb{R}^r} .$$

Az F függvény második változójában való Lipschitzessége az egyik szükséges feltétele az (1.1.1) feladat egyértelmű megoldásának létezéséhez.

Legyen az $F(t, U(t))$ függvény a következő speciális alakú, ahol $\mathcal{A} \in \mathbb{R}^{r \times r}$ invertálható mátrix, azaz létezik \mathcal{A}^{-1} és $G : \mathbb{T} \times \mathbb{R}^r \rightarrow \mathbb{R}^r$ egy nemlineáris, folytonosan differenciálható függvény:

$$F(t, U(t)) = \mathcal{A}U(t) + G(t, U(t)) . \quad (1.1.2)$$

Ekkor az (1.1.1) átírható a következő alakra:

$$\begin{cases} U'(t) = \mathcal{A}U(t) + G(t, U(t)) , t > 0 \\ U(0) = U_0 . \end{cases} \quad (1.1.3)$$

Legyen $\iota, \zeta = 1, \dots, r$, majd $u_\iota : \mathbb{T} \rightarrow \mathbb{R}$ folytonosan differenciálható, $\mathbf{a}_{\iota, \zeta} \in \mathbb{R}$ és $g_\iota : \mathbb{T} \times \mathbb{R}^r \rightarrow \mathbb{R}$. Így az (1.1.3) egyenletben szereplő tagok mátrixos alakban a következők:

$$U'(t) = \begin{pmatrix} u'_1(t) \\ \vdots \\ u'_r(t) \end{pmatrix}, \quad \mathcal{A} = \begin{pmatrix} \mathbf{a}_{1,1} & \cdots & \mathbf{a}_{1,r} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{r,1} & \cdots & \mathbf{a}_{r,r} \end{pmatrix}, \quad U(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_r(t) \end{pmatrix},$$

$$G(t, U(t)) = \begin{pmatrix} g_1(t, u_1(t), \dots, u_r(t)) \\ \vdots \\ g_r(t, u_1(t), \dots, u_r(t)) \end{pmatrix} .$$

Az (1.1.1) egyenlet kiírva koordinátánként az alábbi alakú lesz:

$$\begin{pmatrix} u_1'(t) \\ \vdots \\ u_r'(t) \end{pmatrix} = \begin{pmatrix} \mathbf{a}_{1,1} \cdot u_1(t) + \cdots + \mathbf{a}_{1,r} \cdot u_r(t) + g_1(t, u_1(t), \dots, u_r(t)) \\ \vdots \\ \mathbf{a}_{r,1} \cdot u_1(t) + \cdots + \mathbf{a}_{r,r} \cdot u_r(t) + g_r(t, u_1(t), \dots, u_r(t)) \end{pmatrix} .$$

Ezzel felírtuk a kezdetiérték feladat általános alakját, amit meg akarunk oldani.

1.2. Rácsfüggvény és rácsháló létrehozása

Osszuk fel a \mathbb{T} intervallumot, $N \in \mathbb{N}^+$ egyenlő részre úgy, hogy $N > q$, majd a helyek nevei legyenek $t_0, \dots, t_N \in [0, q]$ és a következő legyen igaz: $0 = t_0 < t_1 < \dots < t_N = q$. Legyen $h \in \mathbb{R}^+$ a felosztott intervallum egy távolsága, azaz $h = \frac{q}{N}$. Legyen $n = 0, 1, \dots, N$ a továbbiakban, majd ezekből látszik, hogy

$$t_n - t_{n-1} = h \text{ és } t_n = n \cdot h, \text{ tehát } t_n + h = t_{n+1} . \quad (1.2.1)$$

Ezzel létrehoztunk egy egyenlő távolságú rácshálót.

Most gyártsuk le azt a közelítő függvényt, ami a megadott t_n pontokban jól közelíti az eredeti ismeretlen $U(t)$ függvényt, de a többi pontban nem feltétlen kell jól közelítenie. Ekkor a közelítő függvény legyen $Y : \mathbb{T} \rightarrow \mathbb{R}^r$ folytonosan differenciálható és a következők teljesüljenek:

$$U(t_0) = Y(t_0) , \quad (1.2.2)$$

$$U(t_n) \approx Y(t_n) , \text{ ahol } n = 1, 2, \dots, N ,$$

$$G(t_0, U(t_0)) = G(t_0, Y(t_0)) . \quad (1.2.3)$$

Ezzel létrehoztunk gyakorlatilag egy rácsfüggvényt, mert csak az adott pontokban közelíti jól az eredeti függvény értékeit.

1.3. Konstans variációs formula felhasználása

Tétel. Legyen $\Psi : \mathbb{T} \rightarrow \mathbb{R}^r$ folytonosan differenciálható, $\Gamma : \mathbb{T} \rightarrow \mathbb{R}^r$ aminek létezik az inverze is, $\Theta : \mathbb{T} \rightarrow \mathbb{R}^{r \times r}$ és $w \in \mathbb{T}$. A konstans variációs formula használható, ha az $\Psi'(t) = \mathcal{A}\Psi(t) + \Gamma(t)$, és $\int \mathcal{A}\Psi(t) = \Theta(t)$ teljesül, ekkor:

$$\Psi(t) = \Theta(t) \cdot \Theta(0)^{-1} \cdot \Psi(0) + \Theta(t) \cdot \int_0^t \Theta(w)^{-1} \Gamma(w) dw .$$

A konstans variációs formula akkor használható, ha a derivált egy homogén, azaz csak U függvénytől függő, és egy inhomogén, azaz U és t értékektől függő részek összegéből áll. Az (1.1.3) pontosan ilyen alakú, ezért használhatjuk most, így alkalmazhatjuk a következő megfeleltetéseket:

$$\begin{aligned} \Theta(t) &:= e^{t\mathcal{A}} , & (1.3.1) \\ \Psi(t) &:= U(t) , \\ \Theta(0)^{-1} = \Theta(0) &= e^{0 \cdot \mathcal{A}} = I , \\ \Psi(0) &:= U(0) = U_0 , \\ \Gamma(w) &:= G(w, U(w)) , \\ \Theta(w)^{-1} &:= e^{-w\mathcal{A}} . \end{aligned}$$

Erre felírva a konstans variációs formulát:

$$U(t) = e^{t\mathcal{A}} \cdot U_0 + e^{t\mathcal{A}} \int_0^t e^{-w\mathcal{A}} \cdot G(w, U(w)) dw . \quad (1.3.2)$$

Most az $U(t)$ függvény azon pontjaira nézzük, ahol a $t = t_n + h$, ahol az U értéke nem ismert, ekkor:

$$U(t_n + h) = e^{(t_n+h)\mathcal{A}} \cdot U_0 + e^{(t_n+h)\mathcal{A}} \int_0^{t_n+h} e^{-w\mathcal{A}} \cdot G(w, U(w)) dw .$$

Az integrál előtti exponenciális érték bevihető az integrál jel mögé, mert az nem függ w értéktől:

$$U(t_n + h) = e^{(t_n+h)\mathcal{A}} \cdot U_0 + \int_0^{t_n+h} e^{(t_n+h-w)\mathcal{A}} \cdot G(w, U(w)) dw . \quad (1.3.3)$$

Az (1.3.3) egyenletben, a $\Psi(0) = U_0$ értéket vettük, most legyen a $\Psi(t_n) = U(t_n)$ érték. Ekkor a $\Theta(0)^{-1}$ most $\Theta(t_n)^{-1}$ lesz, de így a $\Theta(0)^{-1} = \Theta(0) = e^{0 \cdot \mathcal{A}} = I$ helyett a következő lesz: $\Theta(t_n)^{-1} = e^{-t_n \mathcal{A}}$. Mivel nem a 0 helyen nézzük, hanem t_n helyen, ezért az integrál alsó határa a konstans variációs formula miatt t_n helytől indul. Ezzel módosítva a konstans variációs formulát:

$$U(t_n + h) = e^{(t_n+h)\mathcal{A}} \cdot U(t_n) \cdot e^{-t_n \mathcal{A}} + \int_{t_n}^{t_n+h} e^{(t_n+h-w)\mathcal{A}} \cdot G(w, U(w)) dw ,$$

$$U(t_n + h) = e^{h\mathcal{A}} \cdot U(t_n) + \int_{t_n}^{t_n+h} e^{(t_n+h-w)\mathcal{A}} \cdot G(w, U(w)) dw . \quad (1.3.4)$$

Az (1.3.4) megkapható az (1.3.3) levezetéséből a következőképpen is:

$$(1.3.3) = e^{(t_n+h)\mathcal{A}} \cdot U_0 +$$

$$+ e^{(t_n+h)\mathcal{A}} \left[\int_0^{t_n} e^{-w\mathcal{A}} \cdot G(w, U(w)) dw + \int_{t_n}^{t_n+h} e^{-w\mathcal{A}} \cdot G(w, U(w)) dw \right] . \quad (1.3.5)$$

Itt felhasználtuk az integrál additivitását és két részre osztottuk az integrálást a t_n köztes ponttal. Most emeljük ki az $e^{h\mathcal{A}}$ mátrixot és alakítsuk a következő formára az (1.3.5) egyenletet:

$$(1.3.5) = e^{h\mathcal{A}} \left[e^{t_n \mathcal{A}} \cdot U_0 + e^{t_n \mathcal{A}} \int_0^{t_n} e^{-w\mathcal{A}} \cdot G(w, U(w)) dw \right] +$$

$$+ \int_{t_n}^{t_n+h} e^{(t_n+h-w)\mathcal{A}} \cdot G(w, U(w)) dw . \quad (1.3.6)$$

Amennyiben az (1.3.2) egyenletben a $t = t_n$ behelyettesítést végezzük el, akkor a következőt kapjuk:

$$U(t_n) = e^{t_n \mathcal{A}} \cdot U_0 + e^{t_n \mathcal{A}} \int_0^{t_n} e^{-w \mathcal{A}} \cdot G(w, U(w)) dw . \quad (1.3.7)$$

Tehát az (1.3.6) szögletes zárójel közti része ugyanaz mint (1.3.7), így ha az (1.3.7) egyenletet behelyettesítjük, akkor megkapjuk az (1.3.4) egyenletet:

$$U(t_n + h) = e^{h \mathcal{A}} \cdot U(t_n) + \int_{t_n}^{t_n+h} e^{(t_n+h-w) \mathcal{A}} \cdot G(w, U(w)) dw .$$

Tétel. Legyen $v : \mathbb{T} \rightarrow \mathbb{R}$, $O : \mathbb{T} \rightarrow \mathbb{R}^r$, $z \in \mathbb{T}$, $\varkappa \in \mathbb{T}$ és $\varsigma \in \mathbb{T}$, ekkor a helyettesítéses integrálon a következőt értjük:

$$\int_{v(\varkappa)}^{v(\varsigma)} O(w) dw = \int_{\varkappa}^{\varsigma} O(v(z)) v'(z) dz .$$

Tegyük a következő megfeleltetést:

$$\begin{aligned} O(w) &:= e^{(t_n+h-w) \mathcal{A}} \cdot G(w, U(w)) , \\ O(v(z)) &:= e^{(t_n+h-v(z)) \mathcal{A}} \cdot G(v(z), U(v(z))) , \\ v(z) &:= t_n + z , \\ v(\varkappa) &= v(0) := t_n , \\ v(\varsigma) &= v(h) := t_n + h , \\ v'(z) &:= 1 . \end{aligned}$$

Ekkor a következőt kapjuk:

$$\int_{t_n}^{t_n+h} e^{(t_n+h-w) \mathcal{A}} \cdot G(w, U(w)) dw = \int_0^h e^{(t_n+h-v(z)) \mathcal{A}} \cdot G(v(z), U(v(z))) dz .$$

A $v(z)$ helyére behelyettesítve az értékét:

$$\int_{t_n}^{t_n+h} e^{(t_n+h-w)\mathcal{A}} \cdot G(w, U(w)) \, dw = \int_0^h e^{(h-z)\mathcal{A}} \cdot G(t_n + z, U(t_n + z)) \, dz . \quad (1.3.8)$$

Tehát az (1.3.8) egyenletet felhasználva az (1.3.4) integrál részére:

$$U(t_n + h) = e^{h\mathcal{A}} \cdot U(t_n) + \int_0^h e^{(h-z)\mathcal{A}} \cdot G(t_n + z, U(t_n + z)) \, dz . \quad (1.3.9)$$

Ezzel meghatároztunk egy módszert, amivel kiszámítható a $t_n + h$ helyen lévő érték a t_n helybeli értékből, de ez a módszer implicit, ugyanis ismernünk kellene a t_n és a $t_n + h$ közti értékeket is, az integrálon belüli $G(t_n + z, U(t_n + z))$ miatt. Ezt nem ismerjük, ezért a következőkben Lagrange-interpolációt fogunk alkalmazni, amivel közelíteni fogjuk a G függvényt.

1.4. Lagrange-interpoláció ismertetése

Az alábbiakban megvizsgáljuk a Lagrange-féle interpolációt. Legyen $s \in \mathbb{N}^+$. Az elkövetkezőkben legyen $i = 1, \dots, s$. Legyen $p : \mathbb{T} \rightarrow \mathbb{R}$ egy $s - 1$ fokú valós együtthatós polinom függvény. Definiáljunk egy $m : \mathbb{T} \rightarrow \mathbb{R}$ függvényt és olyan $z_i \in \mathbb{T}$ adott pontokat, amelyre a $p(z_i) = m(z_i)$, azaz a z_i pontokban egyezik az értékük. A $(z_i, m(z_i))$ pontpárookra gyártsuk le azt a polinom függvényt, ami illeszkedik, tehát m az ismeretlen függvény, amit p függvénnyel közelítjük, de a pontpárookban megegyezik.

Tétel. *Adott s pontpárra illeszkedő $s - 1$ fokú polinomból pontosan egy létezik.*

A $p(z)$ legyen a következő alakú, ahol a $d_i \in \mathbb{R}$ az együtthatók, de most s pontban ismerjük a z változót, ezért az „ismeretlenek” valójában a d_i számok:

$$p(z) = d_1 + d_2 z + d_3 z^2 + \dots + d_s z^{s-1} .$$

Mivel ismerjük a z értékeit, így írjuk fel a p függvényt a z_i helyeken:

$$\begin{aligned} p(z_1) &= d_1 + d_2 z_1 + d_3 z_1^2 + \dots + d_s z_1^{s-1} = m(z_1) , \\ &\vdots \\ p(z_i) &= d_1 + d_2 z_i + d_3 z_i^2 + \dots + d_s z_i^{s-1} = m(z_i) , \\ &\vdots \\ p(z_s) &= d_1 + d_2 z_s + d_3 z_s^2 + \dots + d_s z_s^{s-1} = m(z_s) . \end{aligned}$$

Definíció. Lagrange-interpolációs polinomnak nevezzük a következőt, ahol a továbbiakban $k = 1, \dots, s$, $\ell_i : \mathbb{T} \rightarrow \mathbb{R}$ valós együtthatós polinom függvény, amit Lagrange-polinomnak nevezünk, és ekkor $p(z)$ lesz az $s - 1$ -ed fokú Lagrange-interpolációs polinom:

$$\ell_i(z) = \prod_{\substack{k=1 \\ k \neq i}}^s \frac{z - z_k}{z_i - z_k} ,$$

$$\text{ahol } p(z) = \sum_{i=1}^s m(z_i) \cdot \ell_i(z) . \quad (1.4.1)$$

A továbbiakban legyen $j = 1, \dots, s$, ekkor a következő a Lagrange-polinom egy fontos tulajdonsága:

$$\ell_i(z_j) = 0, \text{ ha } j \neq i , \quad (1.4.2)$$

$$\ell_i(z_j) = 1, \text{ ha } j = i . \quad (1.4.3)$$

Most válasszuk meg a $(z_i, m(z_i))$ pontpárokat úgy, hogy minden i értékre az $m(z_i) = 1$ teljesüljön. Ekkor a következő lesz igaz:

$$p(z) = \sum_{i=1}^s m(z_i) \cdot \ell_i(z) = \sum_{i=1}^s \ell_i(z) = 1 . \quad (1.4.4)$$

Ez azért lesz igaz, mert a $(z_i, 1)$ pontpárokat vettük, de az erre illesztett interpolációs polinom a konstans 1 lesz, ezért az 1 értékkel egyenlő az (1.4.4).

1.5. Közelítő polinom függvény létrehozása

Legyenek c_i különböző valós számok a $[0, 1]$ intervallumon. Legyen $M : \mathbb{T} \rightarrow \mathbb{R}^r$ és $P : \mathbb{T} \rightarrow \mathbb{R}^r$ olyan vektor polinom függvény, aminek az elemei $m_\iota : \mathbb{T} \rightarrow \mathbb{R}$, illetve $p_\iota : \mathbb{T} \rightarrow \mathbb{R}$ függvények, és z_i pedig az i . függvényhely az interpolációból. Vektoros alakban felírva:

$$M(z_i) = \begin{pmatrix} m_1(z_i) \\ \vdots \\ m_r(z_i) \end{pmatrix}, \quad P(z) = \begin{pmatrix} p_1(z) \\ \vdots \\ p_r(z) \end{pmatrix}.$$

Ekkor az (1.4.1) képlettel felírt interpoláció a következőképpen néz ki a több dimenziós esetben:

$$P(z) = \sum_{i=1}^s M(z_i) \cdot \ell_i(z), \quad (1.5.1)$$

ahol $P(z)$ vektoros alakban kiírva:

$$P(z) = \begin{pmatrix} p_1(z) \\ \vdots \\ p_r(z) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^s m_1(z_i) \cdot \ell_i(z) \\ \vdots \\ \sum_{i=1}^s m_r(z_i) \cdot \ell_i(z) \end{pmatrix}.$$

Legyen ez a P egy olyan $s - 1$ fokú polinom függvény, amelynek a pontjai egybesnek az (1.3.9) egyenletben szereplő $G(t_n + z, U(t_n + z))$ függvény azon pontjaival, amelyre a $t_n + c_i h$ teljesül, ahol $z_i = c_i h$. Valójában a $t_n + c_i h$ helyeken felvett értékeket nem ismerjük, de egy Lagrange-polinommal közelítjük a pontokban. A $G(t_n + z, U(t_n + z))$ függvényt az (1.3.9) miatt 0 és h között kellene integrálni, de $G(t_n + z, U(t_n + z))$ nem integrálható, mert a megadott intervallumon nem ismerjük az értékeket, ezért egy polinom függvénnyel közelítjük a 0 és a h helyek között felvett értékekkel, de t_n értékkel eltolva a $G(t_n + z, U(t_n + z))$ függvényt. A polinom függvény könnyen integrálható. Ekkor az általunk definiált P függvényre

a következők igazak:

$$P(t_n + z) \approx G(t_n + z, U(t_n + z)) , \quad (1.5.2)$$

$$P(t_n + c_i h) := G(t_n + c_i h, U(t_n + c_i h)) . \quad (1.5.3)$$

A $G(t_n + c_i h, U(t_n + c_i h))$ vektoros alakja:

$$G(t_n + c_i h, U(t_n + c_i h)) = \begin{pmatrix} g_1(t_n + c_i h, u_1(t_n + c_i h), \dots, u_r(t_n + c_i h)) \\ \vdots \\ g_r(t_n + c_i h, u_1(t_n + c_i h), \dots, u_r(t_n + c_i h)) \end{pmatrix} .$$

Vizsgáljuk meg mit kapunk, ha az (1.5.3) egyenlet c_i értékébe a 0 értéket illetve az 1 számot helyettesítjük be, amik ezek a c_i értékének a szélső értékei:

$$P(t_n + 0 \cdot h) = G(t_n + 0 \cdot h, U(t_n + 0 \cdot h)) ,$$

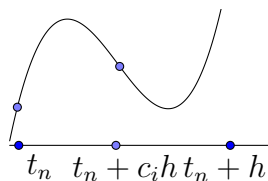
$$P(t_n) = G(t_n, U(t_n)) .$$

A 0 behelyettesítésével az eredeti t_n függvényértéket kapjuk meg, míg ha az 1 számot írjuk be, akkor a következő helyen lévő értéket kapjuk meg, amit keresünk:

$$P(t_n + 1 \cdot h) = G(t_n + 1 \cdot h, U(t_n + 1 \cdot h)) ,$$

$$P(t_n + h) = G(t_n + h, U(t_n + h)) .$$

A $t_n + c_i h$ helyek valójában az általunk ismert köztes helyek, és feltesszük hogy ismerjük a hozzá tartozó értékeket. Lásd (1.1) ábra. A $P(t_n + c_i h)$ függvénynek megfeleltettünk egy Lagrange-interpolációs függvényt, ami a $t_n + c_i h$ helyeken pontosan azt az értéket vette fel, mint a $G(t_n + c_i h, U(t_n + c_i h))$.



1.1. ábra. Köztes t_n helyek

Vezessük be a következő jelölést, ahol $K \in \mathbb{R}^{r \times s}$ mátrix és $K_i \in \mathbb{R}^r$ a következő:

$$K_i := P(t_n + c_i h) . \quad (1.5.4)$$

Mátrixos alakban felírva, ahol $k_{\zeta,i} \in \mathbb{R}$ és $k_{\zeta,i} := p_{\zeta}(t_n + c_i h)$:

$$K = \begin{pmatrix} k_{1,1} & \cdots & k_{1,s} \\ \vdots & \ddots & \vdots \\ k_{r,1} & \cdots & k_{r,s} \end{pmatrix}, \quad K_i = \begin{pmatrix} k_{1,i} \\ \vdots \\ k_{r,i} \end{pmatrix} = \begin{pmatrix} p_1(t_n + c_i h) \\ \vdots \\ p_r(t_n + c_i h) \end{pmatrix} = P(t_n + c_i h) .$$

Mivel a $P(z)$ függvényt egy Lagrange-interpolációs polinomnak készítettük el, ezért az (1.5.1) egyenlet $M(z_i)$ függvény z_i helyei megegyeznek a $t_n + c_i h$ helyekkel és az $M(z_i) = P(t_n + c_i h)$ függvénnyel. Az (1.5.4) miatt

$$K_i = M(z_i) \quad (1.5.5)$$

összefüggés is teljesül, tehát a Lagrange-interpolációs polinomokkal fogjuk meghatározni a közelítő függvény értékeit. Vektoros alakban:

$$K_i = \begin{pmatrix} k_{1,i} \\ \vdots \\ k_{r,i} \end{pmatrix} = \begin{pmatrix} m_1(z_i) \\ \vdots \\ m_r(z_i) \end{pmatrix} = M(z_i) .$$

Az (1.5.5) miatt átírhatjuk az (1.5.1) Lagrange-polinomot a következő formára:

$$P(t_n + z) = \sum_{i=1}^s K_i \cdot \ell_i(z) . \quad (1.5.6)$$

Vegyük észre, hogy az (1.4.2) és az (1.4.3) miatt, a következő igaz:

$$P(t_n + z_j) = \sum_{i=1}^s K_i \cdot \ell_i(z_j) = K_j ,$$

ugyanis az $\ell_i(z_j)$ minden $j \neq i$ helyen 0, de amikor az i index értéke eléri j indexet, akkor az $\ell_i(z_j)$ értéke 1 lesz. Összegezve:

$$K_j = P(t_n + c_j h) = \sum_{i=1}^s K_i \cdot \ell_i(z_j) . \quad (1.5.7)$$

A Lagrange-interpolációt felhasználva, előállítottunk egy olyan $P(t_n + z)$ függvényt, ami a $t_n + z = t_n + c_i h$ pontokban egyezik a $G(t_n + z, U(t_n + z))$ függvénnyel. Ezt fogjuk a továbbiakban felhasználni.

1.6. Polinom függvény alkalmazása

Vegyük észre, hogy az (1.3.9) egyenletben szereplő $G(t_n + z, U(t_n + z))$ függvényt tudjuk közelíteni az (1.5.2) formula miatt a $P(t_n + z)$ függvénnyel, azaz:

$$U(t_n + h) \approx e^{hA} \cdot U(t_n) + \int_0^h e^{(h-z)A} \cdot P(t_n + z) dz .$$

Az (1.2) fejezetben legyártott Y közelítő függvényt felhasználva ez átírható az alábbi módon:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + \int_0^h e^{(h-z)A} \cdot P(t_n + z) dz .$$

Az (1.5.6) összefüggésből kifolyólag a $P(t_n + z)$ formulát átírhatjuk, így a következőt kapjuk:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + \int_0^h e^{(h-z)A} \cdot \sum_{i=1}^s K_i \cdot \ell_i(z) dz .$$

Az (1.5.4) összefüggést felhasználva:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + \int_0^h e^{(h-z)A} \cdot \sum_{i=1}^s P(t_n + c_i h) \cdot \ell_i(z) dz .$$

A $P(t_n + c_i h)$ megfeleltethető az (1.5.3) miatt $G(t_n + c_i h, U(t_n + c_i h))$ függvénynek, de mivel az Y függvénnyel közelítettük az U függvényt, ezért a $P(t_n + c_i h)$ jól közelíti a $G(t_n + c_i h, Y(t_n + c_i h))$ függvényt is:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + \int_0^h e^{(h-z)A} \cdot \sum_{i=1}^s G(t_n + c_i h, Y(t_n + c_i h)) \cdot \ell_i(z) dz ,$$

ezt exponenciális kvadratúra szabálynak is hívják. Vegyük észre, hogy a véges összeg nem függ az integrálban lévő z változótól, így azok kivihetők az integrálból:

$$Y(t_n+h) = e^{hA} \cdot Y(t_n) + \sum_{i=1}^s \left[\int_0^h e^{(h-z)A} \cdot \ell_i(z) dz \right] \cdot G(t_n+c_ih, Y(t_n+c_ih)) . \quad (1.6.1)$$

Az (1.6.1) integrál jelhez tartozó tagok már csak z változótól függenek, ezenkívül egy exponenciális és egy polinom szorzata tartozik az integrálhoz, ami könnyen integrálható 0 és h között. Ezt az integrált nevezzük el a következőképpen úgy, hogy kapjon egy $\frac{1}{h}$ szorzót, ahol $B_i \in \mathbb{R}^{r \times r}$:

$$B_i := \frac{1}{h} \int_0^h e^{(h-z)A} \cdot \ell_i(z) dz . \quad (1.6.2)$$

Ekkor az (1.6.1) és az (1.6.2) összefüggést felhasználva a következőt kapjuk, ahol a korábbi $\frac{1}{h}$ szorzót kompenzáljuk egy h szorzóval:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + h \sum_{i=1}^s B_i \cdot G(t_n + c_ih, Y(t_n + c_ih)) . \quad (1.6.3)$$

Az (1.6.3) formulában szereplő szorzatok sorrendje nem cserélhető meg ugyanis: $e^{hA} \in \mathbb{R}^{r \times r}$, $Y(t_n) \in \mathbb{R}^r$, ezek szorzata pedig csak az (1.6.3) egyenletben szereplő sorrendben lehetséges és a $B_i \in \mathbb{R}^{r \times r}$, ahol minden i indexre B_i egy $\mathbb{R}^{r \times r}$ mátrix, illetve a $G(t_n+c_ih, Y(t_n+c_ih)) \in \mathbb{R}^r$ miatt szintén csak az (1.6.3) egyenletben szereplő sorrendben lehetséges a szorzatuk. Ez a mátrix szorzás szabálya miatt van így.

Most megkaptuk az $Y(t_n + h)$ számítására vonatkozó képletet, de nem ismerjük még az $Y(t_n + c_ih)$ értékeket, most számoljuk ki ezt. Ennek meghatározására a konstans variációs formulát lehetne használni, és levezetni, de mi az egyszerűség kedvéért nézzük az (1.3.9) egyenletet a $h = c_ih$ helyen, és helyettesítsük be az újbóli levezetés helyett:

$$U(t_n + c_ih) = e^{c_ihA} \cdot U(t_n) + \int_0^{c_ih} e^{(c_ih-z)A} \cdot G(t_n + z, U(t_n + z)) dz . \quad (1.6.4)$$

Vegyük észre, hogy az (1.6.4) $G(t_n + z, Y(t_n + z))$ függvényét tudjuk közelíteni az (1.5.2) miatt a $P(t_n + z)$ függvénnyel.

$$U(t_n + c_i h) \approx e^{c_i h A} \cdot U(t_n) + \int_0^{c_i h} e^{(c_i h - z)A} \cdot P(t_n + z) dz .$$

Az (1.2) fejezetben legyártott Y közelítő függvényt felhasználva, a fenti átírható az alábbi módon:

$$Y(t_n + c_i h) = e^{c_i h A} \cdot Y(t_n) + \int_0^{c_i h} e^{(c_i h - z)A} \cdot P(t_n + z) dz . \quad (1.6.5)$$

Az (1.5.6) összefüggést átírhatjuk a következőképpen:

$$P(t_n + z) = \sum_{j=1}^s K_j \cdot \ell_j(z) . \quad (1.6.6)$$

Az (1.6.6) képletet felhasználva az (1.6.5) formulában a következőt kapjuk:

$$Y(t_n + c_i h) = e^{c_i h A} \cdot Y(t_n) + \int_0^{c_i h} e^{(c_i h - z)A} \cdot \sum_{j=1}^s K_j \cdot \ell_j(z) dz .$$

Az (1.6.6) egyenletben szereplő j indexre át kellett térni, mert az (1.6.5) képletben a tagok már függtek az i értéktől. Az (1.5.7) miatt a K_j átírható $P(t_n + c_j h)$ -ra:

$$Y(t_n + c_i h) = e^{c_i h A} \cdot Y(t_n) + \int_0^{c_i h} e^{(c_i h - z)A} \cdot \sum_{j=1}^s P(t_n + c_j h) \cdot \ell_j(z) dz .$$

A $P(t_n + c_j h)$ megfeleltethető $G(t_n + c_j h, Y(t_n + c_j h))$ függvénynek, ezzel behelyettesítve:

$$Y(t_n + c_i h) = e^{c_i h A} \cdot Y(t_n) + \int_0^{c_i h} e^{(c_i h - z)A} \cdot \sum_{j=1}^s G(t_n + c_j h, Y(t_n + c_j h)) \cdot \ell_j(z) dz .$$

Az egyenlet átrendezve:

$$Y(t_n + c_i h) = e^{c_i h A} \cdot Y(t_n) + \sum_{j=1}^s \left[\int_0^{c_i h} e^{(c_i h - z)A} \cdot \ell_j(z) dz \right] \cdot G(t_n + c_j h, Y(t_n + c_j h)) . \quad (1.6.7)$$

Alkalmazzuk a következő jelölést de az integrálnak adjunk egy $\frac{1}{h}$ szorzót, ahol $A_{i,j} \in \mathbb{R}^{r \times r}$:

$$A_{i,j} := \frac{1}{h} \int_0^{c_i h} e^{(c_i h - z)A} \cdot \ell_j(z) dz . \quad (1.6.8)$$

Az (1.6.8) egyenlet beírva az (1.6.7) képletbe és kompenzálva az $\frac{1}{h}$ szorzó:

$$Y(t_n + c_i h) = e^{c_i h A} \cdot Y(t_n) + h \sum_{j=1}^s A_{i,j} \cdot G(t_n + c_j h, Y(t_n + c_j h)) . \quad (1.6.9)$$

A szorzások itt is csak ebben a sorrendben végezhetőek el, a mátrix szorzás szabálya miatt. Megjegyezzük, hogy $A_{i,j} \in \mathbb{R}^{r \times r}$, ahol minden i és j indexre $A_{i,j}$ egy $\mathbb{R}^{r \times r}$ mátrix.

Most már kiszámítottuk az $Y(t_n + c_i h)$ értékét az előző tagokból, most az (1.6.9) egyenletet írjuk be az (1.6.3) összefüggésbe:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + h \sum_{i=1}^s B_i \cdot G(t_n + c_i h, e^{c_i h A} \cdot Y(t_n) + h \sum_{j=1}^s A_{i,j} \cdot G(t_n + c_j h, Y(t_n + c_j h))) .$$

A $G(t_n + c_j h, Y(t_n + c_j h))$ közelítettük a $P(t_n + c_j h)$ függvénnyel, és az (1.5.7) miatt a következő behelyettesítést végezhetjük:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + h \sum_{i=1}^s B_i \cdot G\left(t_n + c_i h, e^{c_i h A} \cdot Y(t_n) + h \sum_{j=1}^s A_{i,j} \cdot K_j\right) , \quad (1.6.10)$$

ahol

$$K_j = G(t_n + c_j h, Y(t_n + c_j h)) .$$

Alkalmazzuk a következő jelölést, ahol a $\mathcal{K}_i \in \mathbb{R}^r$:

$$\mathcal{K}_i := G\left(t_n + c_i h, e^{c_i h A} \cdot Y(t_n) + h \sum_{j=1}^s A_{i,j} \cdot K_j\right) .$$

Ekkor az (1.6.10) a következő lesz:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + h \sum_{i=1}^s B_i \cdot \mathcal{K}_i . \quad (1.6.11)$$

Ezzel megkaptuk az s lépcsős exponenciális Runge–Kutta módszert (RKM), minden $n = 0, 1, \dots, N - 1$ esetben:

$$\begin{aligned} A_{i,j} &= \frac{1}{h} \int_0^{c_i h} e^{(c_i h - z)A} \cdot \ell_j(z) \, dz , \\ B_i &= \frac{1}{h} \int_0^h e^{(h-z)A} \cdot \ell_i(z) \, dz , \\ K_j &= G(t_n + c_j h, Y(t_n + c_j h)) , \\ \mathcal{K}_i &= G(t_n + c_i h, e^{c_i h A} \cdot Y(t_n) + h \sum_{j=1}^s A_{i,j} \cdot K_j) , \\ Y(t_n + h) &= e^{hA} \cdot Y(t_n) + h \sum_{i=1}^s B_i \cdot \mathcal{K}_i . \end{aligned} \quad (1.6.12)$$

Ezzel a módszerrel megoldhatóvá vált az (1.1.3) elsőrendű közönséges r változós differenciálegyenlet rendszer, előre megadott kezdeti feltétellel. A fejezetben levezettük az exponenciális Runge–Kutta módszert.

2. fejezet

Az exponenciális Runge–Kutta módszer átírása és az együtthatók tulajdonságai

Ebben a fejezetben a Taylor-sor segítségével átírjuk a módszerünket, majd megvizsgáljuk az együtthatók viselkedését, bizonyos feltételek mellett. A [4] irodalomban található a fejezet forrása.

2.1. Az exponenciális RKM felírása Taylor-sorral

Legyen $H : \mathbb{T} \rightarrow \mathbb{R}^r$, majd tegyük a következő megfeleltetést:

$$H(t) := G(t, Y(t)) . \quad (2.1.1)$$

Tekintsünk a Taylor-sor azonosságra, amit a H függvényre írunk fel:

$$H(t_n + z) = \sum_{l=1}^{\infty} \frac{z^{l-1} H^{(l-1)}(t_n)}{(l-1)!} , \quad (2.1.2)$$

ahol a $H^{(l-1)}$ függvény jelenti a H függvénynek az $(l-1)$ -edik deriváltját. Az (1.3.9) képletbe helyettesítsük be a (2.1.2) egyenletet:

$$Y(t_n + h) = e^{h\mathcal{A}} \cdot Y(t_n) + \int_0^h e^{(h-z)\mathcal{A}} \cdot \sum_{l=1}^{\infty} \frac{z^{l-1} H^{(l-1)}(t_n)}{(l-1)!} dz .$$

Mivel a H és a végtelen összeg l indextől függ és nem a z változótól, és a Taylor-sor egyenletesen konvergens, így ezek kivihetők az integrálból:

$$Y(t_n + h) = e^{h\mathcal{A}} \cdot Y(t_n) + \sum_{l=1}^{\infty} \int_0^h e^{(h-z)\mathcal{A}} \cdot \frac{z^{l-1}}{(l-1)!} dz \cdot H^{(l-1)}(t_n) . \quad (2.1.3)$$

Tegyük a következő elnevezéseket, ahol az $l = 1, 2, \dots$, és $\varphi_l \in \mathbb{R}^{r \times r}$ egy mátrix:

$$\varphi_l := \varphi_l(h\mathcal{A}) := \frac{1}{h^l} \int_0^h e^{(h-z)\mathcal{A}} \cdot \frac{z^{l-1}}{(l-1)!} dz , \quad (2.1.4)$$

$$\varphi_0 := \varphi_0(h\mathcal{A}) := e^{h\mathcal{A}} . \quad (2.1.5)$$

Ha a (2.1.4) egyenletben \mathcal{A} helyett a 0 mátrixot írjuk, a következőt kapjuk, ahol $I \in \mathbb{R}^{r \times r}$ identitás mátrix::

$$\varphi_l(0) = \frac{1}{h^l} \int_0^h I \frac{z^{l-1}}{(l-1)!} dz = \frac{I}{l!} .$$

Végezzük el a következő integrálást:

$$\begin{aligned}
 \varphi_{l+1} &= \frac{1}{h^{l+1}} \int_0^h e^{(h-z)\mathcal{A}} \cdot \frac{z^l}{l!} dz = \frac{1}{h^{l+1}} \left[-e^{(h-z)\mathcal{A}} \cdot \frac{z^l}{l!} \cdot \mathcal{A}^{-1} \right]_0^h - \\
 &\quad - \frac{1}{h^{l+1}} \int_0^h -l \cdot e^{(h-z)\mathcal{A}} \cdot \frac{z^{l-1}}{l!} \cdot \mathcal{A}^{-1} dz = \\
 &= -\frac{1}{l!} \cdot \frac{\mathcal{A}^{-1}}{h} + \frac{1}{h^l} \int_0^h e^{(h-z)\mathcal{A}} \cdot \frac{z^{l-1}}{(l-1)!} dz \cdot \frac{\mathcal{A}^{-1}}{h} = \\
 &= \left(\frac{1}{h^l} \int_0^h e^{(h-z)\mathcal{A}} \cdot \frac{z^{l-1}}{(l-1)!} dz - \frac{I}{l!} \right) \frac{\mathcal{A}^{-1}}{h} = \\
 &= (\varphi_l - \varphi_l(0)) \cdot \frac{\mathcal{A}^{-1}}{h} .
 \end{aligned}$$

Amennyiben ezt rekurzívan folytatjuk, a következőt kapjuk:

$$\varphi_l = e^{h\mathcal{A}} \cdot \frac{\mathcal{A}^{-l}}{h^l} - \sum_{\hat{l}=1}^l \frac{\mathcal{A}^{-\hat{l}}}{h^{\hat{l}}(l-\hat{l})!} . \quad (2.1.6)$$

Tekintsünk a következő azonosságra:

$$e^{h\mathcal{A}} = \sum_{\hat{l}=0}^{\infty} \frac{(h\mathcal{A})^{\hat{l}}}{\hat{l}!} = I + (h\mathcal{A}) + \frac{(h\mathcal{A})^2}{2} + \frac{(h\mathcal{A})^3}{6} + \dots . \quad (2.1.7)$$

A (2.1.7) behelyettesítve a (2.1.6) képletbe:

$$\varphi_l = \sum_{\hat{l}=0}^{\infty} \frac{(h\mathcal{A})^{\hat{l}}}{\hat{l}!} \cdot \frac{\mathcal{A}^{-l}}{h^l} - \sum_{\hat{l}=1}^l \frac{\mathcal{A}^{-\hat{l}}}{h^{\hat{l}}(l-\hat{l})!} .$$

Az első szumma kiírva és elvégezve a szorzás:

$$\sum_{\hat{l}=0}^{\infty} \frac{(h\mathcal{A})^{\hat{l}}}{\hat{l}!} \cdot \frac{\mathcal{A}^{-l}}{h^l} = \frac{\mathcal{A}^{-l}}{0!h^l} + \frac{\mathcal{A}^{-l+1}}{1!h^{l-1}} + \frac{\mathcal{A}^{-l+2}}{2!h^{l-2}} + \dots + \frac{\mathcal{A}^{-1}}{(l-1)!h} + \frac{I}{l!} + \frac{h\mathcal{A}}{(l+1)!} + \dots .$$

A második szumma kiírva:

$$\sum_{\hat{l}=1}^l \frac{\mathcal{A}^{-\hat{l}}}{h^{\hat{l}}(l-\hat{l})!} = \frac{\mathcal{A}^{-l}}{h^l} + \frac{\mathcal{A}^{-l+1}}{h^{l-1}} + \dots + \frac{\mathcal{A}^{-1}}{(l-1)!h} .$$

Elvégezve a kivonást:

$$\varphi_l = \frac{I}{l!} + \frac{h\mathcal{A}}{(l+1)!} + \frac{h^2\mathcal{A}^2}{(l+2)!} + \dots \quad (2.1.8)$$

Amennyiben $l \rightarrow \infty$, akkor a két szumma normában vett különbsége a 0 értékhez tart, ami így egy monoton csökkenő sorozat. Tehát $\|\varphi_l\|_{\mathbb{R}^{r \times r}} \rightarrow 0$. A (2.1.4) képletet helyettesítsük be a (2.1.3) egyenletbe. Behelyettesítve:

$$Y(t_n + h) = e^{h\mathcal{A}} \cdot Y(t_n) + \sum_{l=1}^{\infty} h^l \varphi_l \cdot H^{(l-1)}(t_n) \quad (2.1.9)$$

Végül az összeg kiírva:

$$Y(t_n + h) = e^{h\mathcal{A}} \cdot Y(t_n) + h\varphi_1 H(t_n) + h^2\varphi_2 H'(t_n) + h^3\varphi_3 H''(t_n) + \dots$$

Most az (1.6.4) azonosságra használjuk a (2.1.2) azonosságot:

$$Y(t_n + c_i h) = e^{c_i h \mathcal{A}} \cdot Y(t_n) + \int_0^{c_i h} e^{(c_i h - z)\mathcal{A}} \cdot \sum_{l=1}^{\infty} \frac{z^{l-1} H^{(l-1)}(t_n)}{(l-1)!} dz \quad .$$

Mivel H és az összeg független az integráltól, ezért azok kivihetőek:

$$Y(t_n + c_i h) = e^{c_i h \mathcal{A}} \cdot Y(t_n) + \sum_{l=1}^{\infty} \int_0^{c_i h} e^{(c_i h - z)\mathcal{A}} \cdot \frac{z^{l-1}}{(l-1)!} dz \cdot H^{(l-1)}(t_n) \quad .$$

Tegyük a következő megfeleltetést, ahol a $\varphi_{l,i} \in \mathbb{R}^{r \times r}$ egy mátrix:

$$\varphi_{l,i} := \frac{1}{h^l} \int_0^{c_i h} e^{(c_i h - z)\mathcal{A}} \cdot \frac{z^{l-1}}{(l-1)!} dz \quad (2.1.10)$$

Behelyettesítve:

$$Y(t_n + c_i h) = e^{c_i h \mathcal{A}} \cdot Y(t_n) + \sum_{l=1}^{\infty} h^l \varphi_{l,i} \cdot H^{(l-1)}(t_n) \quad , \quad (2.1.11)$$

minden $n = 0, 1, \dots, N - 1$ esetén. Kiírva:

$$Y(t_n + c_i h) = e^{c_i h \mathcal{A}} \cdot Y(t_n) + h\varphi_{1,i} H(t_n) + h^2\varphi_{2,i} H'(t_n) + \dots$$

Ezzel felírtuk a Runge–Kutta módszert a Taylor-sorral.

2.2. A B_i tulajdonságai

Most vizsgáljuk meg az (1.6.2) képlettel definiált mátrixokat. Vegyük a véges összeget az $i = 1, \dots, s$ értékek között:

$$\sum_{i=1}^s B_i = \sum_{i=1}^s \frac{1}{h} \int_0^h e^{(h-z)\mathcal{A}} \cdot \ell_i(z) dz .$$

A véges összeg i indextől függ, az integrál függvényei pedig z változótól, ezért a szumma bevihető az ℓ_i elé:

$$\sum_{i=1}^s B_i = \frac{1}{h} \int_0^h e^{(h-z)\mathcal{A}} \cdot \sum_{i=1}^s \ell_i(z) dz .$$

Legyen $B \in \mathbb{R}^{r \times r}$. Az (1.4.4) miatt a véges összeg értéke 1 lesz, így a következőt kapjuk:

$$B := \sum_{i=1}^s B_i = \frac{1}{h} \int_0^h e^{(h-z)\mathcal{A}} dz .$$

Az integrálás elvégezve:

$$B = \sum_{i=1}^s B_i = \frac{1}{h} e^{h\mathcal{A}} [-e^{-z\mathcal{A}} \mathcal{A}^{-1}]_0^h = \frac{1}{h} (e^{h\mathcal{A}} - I) \mathcal{A}^{-1} . \quad (2.2.1)$$

A (2.1.7) egyenletet behelyettesítve a (2.2.1) képletbe, és \mathcal{A}^{-1} mátrixszorzást végezve jobbról:

$$\sum_{i=1}^s B_i = I + \frac{h\mathcal{A}}{2} + \frac{h^2\mathcal{A}^2}{6} + \dots$$

A (2.1.4) képlet $l = 1$ esetben könnyen integrálható, ami a következő alakú:

$$\varphi_1 = \frac{1}{h} (e^{h\mathcal{A}} - I) \mathcal{A}^{-1} . \quad (2.2.2)$$

Ekkor a (2.2.1) képletre a (2.2.2) miatt teljesül a következő:

$$B = \sum_{i=1}^s B_i = \varphi_1 . \quad (2.2.3)$$

2.3. Az $A_{i,j}$ tulajdonságai

Az (1.6.8) képletre is alkalmazzunk összegzést j szerint:

$$\sum_{j=1}^s A_{i,j} = \sum_{j=1}^s \frac{1}{h} \int_0^{c_i h} e^{(c_i h - z)\mathcal{A}} \cdot \ell_j(z) dz .$$

A j indextől egyedül ℓ_i függ, így a szumma bevihető az integrál jel mögé:

$$\sum_{j=1}^s A_{i,j} = \frac{1}{h} \int_0^{c_i h} e^{(c_i h - z)\mathcal{A}} \cdot \sum_{j=1}^s \ell_j(z) dz .$$

Az (1.4.4) miatt a véges összeg 1 lesz, legyen $A_i \in \mathbb{R}^{r \times r}$, így a következőt kapjuk:

$$A_i := \sum_{j=1}^s A_{i,j} = \frac{1}{h} \int_0^{c_i h} e^{(c_i h - z)\mathcal{A}} dz .$$

Elvégezve az integrálást:

$$A_i := \sum_{j=1}^s A_{i,j} = \frac{1}{h} (e^{c_i h \mathcal{A}} - I) \mathcal{A}^{-1} . \quad (2.3.1)$$

A (2.1.10) képlet $l = 1$ esetben könnyen integrálható, ami a következő:

$$\varphi_{1,i} = \frac{1}{h} (e^{c_i h \mathcal{A}} - I) \mathcal{A}^{-1} . \quad (2.3.2)$$

Ekkor a (2.3.2) képletet felhasználva a (2.3.1) képletre, a következő teljesül:

$$A_i = \sum_{j=1}^s A_{i,j} = \varphi_{1,i} . \quad (2.3.3)$$

3. fejezet

Módszerek

Ebben a fejezetben felírunk néhány, az általános exponenciális Runge–Kutta módszerből levezethető további módszert. A [4] irodalomban található a fejezet forrása.

3.1. A közönséges Runge–Kutta módszer

A közönséges Runge–Kutta módszert megkapjuk, amennyiben az \mathcal{A} értékét a 0 mátrixnak választjuk. Ekkor az (1.6.12) módszerből kiesik az összes exponenciális tag, hiszen $e^{h0} = I$ így visszkapjuk a közönséges Runge–Kutta módszert. Azokat a módszereket, ahol az $\mathcal{A} = 0$ feltételt tesszük, underlying módszereknek nevezzük. Helyettesítsük be az (1.6.12) képletekben az \mathcal{A} helyére a 0 mátrixot, ekkor minden $n = 0, 1, \dots, N - 1$ esetén a következő módszert kapjuk:

$$A_{i,j} = \frac{1}{h} \int_0^{c_j h} \ell_j(z) dz ,$$

$$B_i = \frac{1}{h} \int_0^h \ell_i(z) dz ,$$

$$K_j = G(t_n + c_j h, Y(t_n + c_j h)) ,$$

$$\mathcal{K}_i = G\left(t_n + c_i h, Y(t_n) + h \sum_{j=1}^s A_{i,j} \cdot K_j\right),$$
$$Y(t_n + h) = Y(t_n) + h \sum_{i=1}^s B_i \cdot \mathcal{K}_i.$$

Ezzel megkaptuk az implicit Runge–Kutta módszert. Ezt a módszert akkor érdemes használni, amikor az (1.1.3) differenciálegyenlet rendszer nem tartalmazza az $\mathcal{A}U(t)$ merev részt, azaz (1.1.1) alakú. A tapasztalat azt mutatja, ha egy merev részt tartalmazó egyenletre használnánk, akkor a megoldása lényegesen lassabb lenne, és több számítást igényelne a kisebb lépés és nagyobb lépcsőszám miatt, ami szükséges lenne.

A (2.1.4) $l = 1$ és $\mathcal{A} = 0$ esetben a következő lesz:

$$\varphi_1 = I.$$

Ekkor a (2.2.3) a következőképpen fog kinézni:

$$B = \sum_{i=1}^s B_i = I.$$

A (2.1.10) $l = 1$ és $\mathcal{A} = 0$ esetben a következő lesz:

$$\varphi_{1,i} = c_i I.$$

Ekkor a (2.3.3) a következőképpen fog kinézni:

$$A_i = \sum_{j=1}^s A_{i,j} = c_i I.$$

Ezek valójában az underlying Runge–Kutta módszernek az elsőrend eléréséhez szükséges rendszabályok.

3.2. Explicit exponenciális Runge–Kutta módszer

A módszert explicitnek nevezzük, ha a j index nem s értékig megy, hanem $i - 1$ indexig, ugyanis ekkor a következő tag kiszámításához csak az előző ismert tagokat

használjuk fel. Az implicit esetben a nem ismert jövőbeli értékeket is felhasználjuk. A továbbiakban az explicit exponenciális Runge–Kutta módszerekkel foglalkozunk. Tegyük a következő kikötést: $c_1 = 0$. Ekkor az (1.6.8) miatt az $A_{1,j}$ mátrixok is 0 értékűek, ugyanis az integrál 0 és 0 közötti tartományon megy, ahol így 0 lesz értéke. Ekkor a \mathcal{K}_i mátrix alakja:

$$\mathcal{K}_i = G\left(t_n + c_i h, e^{c_i h A} \cdot Y(t_n) + h \sum_{j=1}^{i-1} A_{i,j} \cdot K_j\right).$$

Értékenként kiírva:

$$\begin{aligned} \mathcal{K}_1 &= G(t_n, Y(t_n)) , \\ \mathcal{K}_2 &= G(t_n + c_2 h, e^{c_2 h A} \cdot Y(t_n) + h A_{2,1} \cdot G(t_n, Y(t_n)) , \\ &\vdots \\ \mathcal{K}_s &= G(t_n + c_s h, e^{c_s h A} \cdot Y(t_n) + \\ &\quad + h A_{s,1} \cdot G(t_n, Y(t_n)) + \dots + h A_{s,s-1} \cdot G(t_n + c_{s-1} h, Y(t_n + c_{s-1} h))) . \end{aligned}$$

A módszert implicitnek nevezzük, ha a j indexet s indexig nézzük.

Az $A_{i,j}$ és B_i mátrixokat, valamint a c_i értékeket felhasználva felírhatjuk az exponenciális Runge–Kutta módszerekre vonatkozó Butcher-táblázatokat:

0				
c_2	$A_{2,1}$			
\vdots	\vdots	\ddots		
c_s	$A_{s,1}$	\dots	$A_{s,s-1}$	
	B_1	B_2	\dots	B_s

Explicit exponenciális RKM Butcher-tábla

c_1	$A_{1,1}$	$A_{1,2}$	\dots	$A_{1,s}$	A_1
c_2	$A_{2,1}$	$A_{2,2}$	\dots	$A_{2,s}$	A_2
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
c_s	$A_{s,1}$	$A_{s,2}$	\dots	$A_{s,s}$	A_s
	B_1	B_2	\dots	B_s	B

Implicit exponenciális RKM Butcher-tábla

Az utolsó oszlopban szereplő A_i és B tagok, a sorösszeget jelentik, amiket a (2.3) és (2.2) fejezetekben vezettünk be.

3.3. Exponenciális Euler-módszer

Legyen $\mathcal{H} : \mathbb{T} \rightarrow \mathbb{R}^r$, és tegyük a következő megfeleltetést:

$$\mathcal{H}(t) := G(t, U(t)) . \quad (3.3.1)$$

A (2.1.2) összefüggést felírhatjuk az $\mathcal{H}(t_n + z)$ tagra is a $H(t_n + z)$ tag helyett, ahol az $R : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{R}^r$ a Taylor-sor maradék tagját jelenti:

$$\mathcal{H}(t_n + z) = \mathcal{H}(t_n) + R(t_n, z) . \quad (3.3.2)$$

Most az (1.3.9) képletbe helyettesítsük be a (3.3.2) azonosságot, így a következőt kapjuk:

$$U(t_n + h) = e^{hA} \cdot U(t_n) + \int_0^h e^{(h-z)A} \cdot (\mathcal{H}(t_n) + R(t_n, z)) \, dz .$$

A zárójel felbontás után:

$$U(t_n + h) = e^{hA} \cdot U(t_n) + \int_0^h e^{(h-z)A} \cdot \mathcal{H}(t_n) \, dz + \int_0^h e^{(h-z)A} \cdot R(t_n, z) \, dz .$$

A maradék tagokat elhagyva megkapjuk a módszerünket. Mivel tagokat hagyunk el, így az eredeti U függvény helyett a közelítő Y függvényre kell felírnunk.

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + \int_0^h e^{(h-z)A} \cdot H(t_n) dz . \quad (3.3.3)$$

A fenti képlet átírható a (2.1.4) képlet $l = 1$ eset segítségével a következő alakra:

$$Y(t_n + h) = e^{hA} \cdot Y(t_n) + h\varphi_1 H(t_n) , \quad (3.3.4)$$

minden $n = 0, 1, \dots, N - 1$ érték esetén. Ezzel előállítottuk az explicit exponenciális Euler-módszert. A módszer az (1.6.12) módszerből is megkapható, ha az együtthatókat a következőnek választjuk:

$$\begin{aligned} s &= 1 , \\ c_1 &= 0 , \\ A_{1,1} &= 0 , \\ \ell_1(z) &= 1 , \\ B_1 &= \frac{1}{h} \int_0^h e^{(h-z)A} dz . \end{aligned}$$

Ekkor felírható a következő Butcher-tábla:

$$\begin{array}{c|c} 0 & 0 \\ \hline & B_1 \end{array}$$

Explicit exponenciális Euler-módszer Butcher-táblája

4. fejezet

Az exponenciális Euler-módszer tulajdonságai

Ebben a fejezetben bevezetjük a merevség, a konzisztencia, és a konvergencia fogalmát. Ellenőrizzük a módszer konvergenciáját, konzisztenciáját, és a konvergenciához szükséges feltételeket. A [3] forrásban található a konvergencia bizonyítása, amit felhasználtunk, a [4] számú forrásban van a konvergencia bizonyítása Banach-terekre és a merevség ismertetése, és az [5] irodalomban van a konzisztencia bizonyítása az általános elsőrendű esetben.

4.1. A merevség ismertetése

A differenciálegyenletek a merevségük alapján két csoportra oszthatóak: merev és nem merev.

A merevségre nincs általánosan elfogadott definíció, ezt a numerikus módszerünk használata után állapíthatjuk meg. Merevség megállapítására szokták használni a differenciálegyenlet rendszerre felírt Jacobi-mátrix normában vett legnagyobb és legkisebb sajátértékének hányadosát. Amennyiben ez a hányados nagy,

akkor merev az egyenlet, de sajnos vannak esetek, amikor az egyenlet merev, de a hányados kicsi. Egy egyenletet merevnek mondunk, ha az explicit módszernek nagy (h) lépésközt választva, a numerikus megoldás a pontos megoldás körül oszcillál vagy nem tart hozzá. Amennyiben ezt a lépésközt kellően kicsinek választjuk, akkor már jól és gyorsan fogja közelíteni, de ez nagy mennyiségű számítást igényelne. Az explicit módszer numerikusan instabil módszer egy merev egyenlet megoldásánál, mert létezik egy kritikus lépésköz, ami fölött a módszer csak sok lépés után közelít jól vagy egyáltalán nem közelíti, csak oszcillál. A merev egyenleteket implicit módszerekkel tudunk jól megoldani.

Nem merevnek nevezünk egy egyenletet, ha az explicit és az implicit módszerek is jól közelítik a pontos megoldást a megválasztott akár nagy h érték esetén is.

Az exponenciális Runge–Kutta típusú módszereknél az $U'(t)$ függvényt két komponensre bontottuk: egy lineáris merev részre, valamint egy nemlineáris nem merev részre. A merev részt pontosan kiszámítjuk mindig az $e^{hA} \cdot Y(t_n)$ képlettel, így a merev rész nem okoz problémát. Ebből következik, hogy nem szükséges implicit megoldást használnunk, mert az explicit exponenciális Runge–Kutta módszer is jól fogja közelíteni a merev és a nem merev egyenleteket is.

4.2. A konvergencia ismertetése

Definíció. Egy módszert konvergensenek nevezünk, ha a közelítő megoldás tart a pontos megoldáshoz, tehát a pontos és a közelítő megoldás különbsége a 0 értékhez tart normában minden $t = nh = t_n \in \mathbb{T}$ pontban, amennyiben a h távolsággal is 0 számhoz tartunk. Ezzel egy finomodó rácsháló fog létrejönni:

$$\lim_{\substack{h \rightarrow 0 \\ nh = t_n}} \|Y(t_n) - U(t_n)\|_{\mathbb{R}^r} = 0 . \quad (4.2.1)$$

A pontos és a közelítő megoldás különbsége a globális hiba, ami minden egyes lépésben növekedik, de a lépésköz csökkentésével ez a hiba a 0 számhoz tart.

Definíció. Egy módszert ρ rendben konvergensenek nevezünk, ha létezik olyan $\mathbf{L} \in \mathbb{T}$ konstans, ami nem függ n és h értékeitől úgy, hogy az alábbi teljesül:

$$\|Y(t_n) - U(t_n)\|_{\mathbb{R}^r} = \mathbf{L}h^\rho . \quad (4.2.2)$$

4.3. A módszer konvergenciája

Tétel. Legyen $\mathcal{L} \in \mathbb{R}^+$ konstans, majd tegyük fel, hogy a G függvény is Lipschitzes a második változójában, azaz:

$$\|G(t_n, Y(t_n)) - G(t_n, U(t_n))\|_{\mathbb{R}^r} \leq \mathcal{L} \|Y(t_n) - U(t_n)\|_{\mathbb{R}^r} ,$$

és létezik szuprénuma a $G'(t, U(t))$ függvénynek. Legyen $S \in \mathbb{R}^+$ véges értékű, és legyen

$$S := \sup_{0 \leq t \leq t_n} \|G'(t, U(t))\|_{\mathbb{R}^r} .$$

Ekkor a (3.3.3), azaz az exponenciális Euler-módszer első rendben konvergens.

Bizonyítás. Tekintsünk a \mathcal{H} függvény alábbi Taylor-sor alakjára, ahol a maradék tagokat integrál formában írjuk fel, ahol a $\sigma \in \mathbb{T}$:

$$\mathcal{H}(t_n + z) = \mathcal{H}(t_n) + \int_0^z \mathcal{H}'(t_n + \sigma) d\sigma . \quad (4.3.1)$$

Helyettesítsük be az (1.3.9) képletbe a (4.3.1) Taylor-sor fejtést, ekkor a következőt kapjuk:

$$\begin{aligned} U(t_n + h) &= e^{h\mathcal{A}} \cdot U(t_n) + \int_0^h e^{(h-z)\mathcal{A}} \cdot \mathcal{H}(t_n) dz + \\ &+ \int_0^h e^{(h-z)\mathcal{A}} \cdot \left[\int_0^z \mathcal{H}'(t_n + \sigma) d\sigma \right] dz . \end{aligned} \quad (4.3.2)$$

Tegyük a következő megfeleltetést, ahol $\delta_n \in \mathbb{R}^r$ és $n = 0, 1, \dots, N - 1$ esetén:

$$\delta_{n+1} := \int_0^h e^{(h-z)\mathcal{A}} \cdot \left[\int_0^z \mathcal{H}'(t_n + \sigma) d\sigma \right] dz . \quad (4.3.3)$$

Ekkor a (4.3.2) a következőképpen írható fel a (4.3.3) segítségével:

$$U(t_n + h) = e^{hA} \cdot U(t_n) + \int_0^h e^{(h-z)A} dz \cdot \mathcal{H}(t_n) + \delta_{n+1} .$$

A (2.1.4) képletet $l = 1$ esetben nézve, majd behelyettesítve a fentibe:

$$U(t_n + h) = e^{hA} \cdot U(t_n) + h\varphi_1 \mathcal{H}(t_n) + \delta_{n+1} . \quad (4.3.4)$$

Ezzel felírtuk az exponenciális Euler-módszert az (1.1.3) probléma pontos megoldásának segítségével. Tegyük a következő megfeleltetést, ahol $\mathcal{E}_n \in \mathbb{R}^r$ hibamátrix, amivel mérhető az eredeti függvény és a közelítő közti eltérés, majd felhasználjuk a (4.3.4) és a (3.3.4) képletet:

$$\mathcal{E}_n := Y(t_n) - U(t_n) , \quad (4.3.5)$$

$$\begin{aligned} \mathcal{E}_{n+1} := Y(t_n + h) - U(t_n + h) &= e^{hA} Y(t_n) + h\varphi_1 H(t_n) - \\ &- e^{hA} U(t_n) - h\varphi_1 \mathcal{H}(t_n) - \delta_{n+1} , \end{aligned}$$

minden $n = 0, 1, \dots, N - 1$ esetén.

A képlet összevonva:

$$\mathcal{E}_{n+1} = e^{hA} (Y(t_n) - U(t_n)) + h\varphi_1 (H(t_n) - \mathcal{H}(t_n)) - \delta_{n+1} . \quad (4.3.6)$$

Használjuk a következő elnevezést, ahol $E_n \in \mathbb{R}^r$:

$$E_n := H(t_n) - \mathcal{H}(t_n) = G(t_n, Y(t_n)) - G(t_n, U(t_n)) . \quad (4.3.7)$$

Felhasználjuk a (4.3.5) és a (4.3.7) megfeleltetést a (4.3.6) hibában:

$$\mathcal{E}_{n+1} = e^{hA} \mathcal{E}_n + h\varphi_1 E_n - \delta_{n+1} .$$

Rekurzívan folytatva, ahol $\kappa = 0, 1, \dots, n - 1$, a következőt kapjuk:

$$\mathcal{E}_n = h \sum_{\kappa=0}^{n-1} e^{(n-\kappa-1)hA} \varphi_1 E_\kappa - \sum_{\kappa=0}^{n-1} e^{\kappa hA} \delta_{n-\kappa} . \quad (4.3.8)$$

Írjuk ki a (4.3.8) második véges összegében lévő értéket a (4.3.3) segítségével:

$$e^{\kappa h \mathcal{A}} \delta_{n-\kappa} = e^{\kappa h \mathcal{A}} \int_0^h e^{(h-z)\mathcal{A}} \cdot \left[\int_0^z \mathcal{H}'(t_{n-\kappa-1} + \sigma) d\sigma \right] dz. \quad (4.3.9)$$

A (4.3.9) hiba felülről becsülhető, amennyiben a tételben szereplő felső becslést, az S konstans behelyettesítjük. A behelyettesítés azért lehetséges, mert a belső integrál bár 0 és z között mozog, de a külső viszont csak 0 és h között. Ennek következményeként a belső integrál is 0 és h között mozog, ami miatt a belső \mathcal{H}' függvény maximum a t_n helyig megy el, ami viszont a felső becslésünkben is figyelve van.

$$\|e^{\kappa h \mathcal{A}} \delta_{n-\kappa}\|_{\mathbb{R}^r} \leq \left\| e^{\kappa h \mathcal{A}} \int_0^h e^{(h-z)\mathcal{A}} \cdot \left[\int_0^z S d\sigma \right] dz \right\|_{\mathbb{R}^{r \times r}}.$$

A belső integrálás elvégezve:

$$\begin{aligned} \|e^{\kappa h \mathcal{A}} \delta_n\|_{\mathbb{R}^r} &\leq \left\| e^{\kappa h \mathcal{A}} \int_0^h e^{(h-z)\mathcal{A}} \cdot \left[S\sigma \right]_0^z dz \right\|_{\mathbb{R}^{r \times r}}, \\ \|e^{\kappa h \mathcal{A}} \delta_n\|_{\mathbb{R}^r} &\leq \left\| e^{\kappa h \mathcal{A}} \int_0^h z e^{(h-z)\mathcal{A}} dz \cdot S \right\|_{\mathbb{R}^{r \times r}}. \end{aligned} \quad (4.3.10)$$

A (2.1.4) $l = 2$ esetben használva, és behelyettesítve a (4.3.10) képletbe:

$$\|e^{\kappa h \mathcal{A}} \delta_n\|_{\mathbb{R}^r} \leq \left\| e^{\kappa h \mathcal{A}} \int_0^h z e^{(h-z)\mathcal{A}} dz \cdot S \right\|_{\mathbb{R}^{r \times r}} = \|e^{\kappa h \mathcal{A}} h^2 \varphi_2 S\|_{\mathbb{R}^{r \times r}}. \quad (4.3.11)$$

A (4.3.8) második szummája felülről becsülve a (4.3.11) segítségével:

$$\sum_{\kappa=0}^{n-1} \|e^{\kappa h \mathcal{A}} \delta_{n-\kappa}\|_{\mathbb{R}^r} \leq \sum_{\kappa=0}^{n-1} \|e^{\kappa h \mathcal{A}} h^2 \varphi_2 S\|_{\mathbb{R}^{r \times r}}. \quad (4.3.12)$$

Tekintsünk a következő felső becslés levezetésére, mivel ez egy abszolút konvergens sor, ezért lehetséges a levezetés:

$$\|e^{t\mathcal{A}}\|_{\mathbb{R}^{r \times r}} = \left\| \sum_{l=0}^{\infty} \frac{t^l \mathcal{A}^l}{l!} \right\|_{\mathbb{R}^{r \times r}} \leq \sum_{l=0}^{\infty} \frac{t^l \|\mathcal{A}^l\|_{\mathbb{R}^{r \times r}}}{l!} \leq \sum_{l=0}^{\infty} \frac{t^l \|\mathcal{A}\|_{\mathbb{R}^{r \times r}}^l}{l!} \leq e^{\|\mathcal{A}\|_{\mathbb{R}^{r \times r}} t}.$$

Legyen $\Omega \in \mathbb{R}^+$ és $\Omega := \|\mathcal{A}\|_{\mathbb{R}^{r \times r}}$, ekkor a fenti összefüggés a következőképpen írható fel:

$$\|e^{t\mathcal{A}}\|_{\mathbb{R}^{r \times r}} \leq e^{\Omega t} . \quad (4.3.13)$$

Írjuk ki, és alakítsuk át a felülről becsülő véges összeget a (4.3.12) egyenlőtlenségben:

$$\begin{aligned} \sum_{\kappa=0}^{n-1} \|e^{\kappa h\mathcal{A}} h^2 \varphi_2 S\|_{\mathbb{R}^{r \times r}} &= \|(e^{0h\mathcal{A}} + e^{1h\mathcal{A}} + \dots + e^{(n-1)h\mathcal{A}}) S h^2 \varphi_2\|_{\mathbb{R}^{r \times r}} = \\ &= \|(e^{t_0\mathcal{A}} + e^{t_1\mathcal{A}} + \dots + e^{t_{n-1}\mathcal{A}}) h^2 \varphi_2 S\|_{\mathbb{R}^{r \times r}} . \end{aligned} \quad (4.3.14)$$

A (4.3.14) felülről becsülhető a (4.3.13) használatával, mert csak pozitív kitevők vannak:

$$\begin{aligned} \|e^{t_0\mathcal{A}}\|_{\mathbb{R}^{r \times r}} &\leq e^{\Omega t_0} \leq e^{\Omega t_n} , \\ &\vdots \\ \|e^{t_{n-1}\mathcal{A}}\|_{\mathbb{R}^{r \times r}} &\leq e^{\Omega t_{n-1}} \leq e^{\Omega t_n} , \\ \|(e^{t_0\mathcal{A}} + e^{t_1\mathcal{A}} + \dots + e^{t_{n-1}\mathcal{A}}) h^2 \varphi_2 S\|_{\mathbb{R}^{r \times r}} &\leq \|n e^{\Omega t_n} h^2 \varphi_2 S\|_{\mathbb{R}^{r \times r}} . \end{aligned}$$

Legyen $L \in \mathbb{R}^+$ konstans, ami függ Ω és t_n értékektől, tehát $L := L(\Omega, t_n)$:

$$L := e^{\Omega t_n} . \quad (4.3.15)$$

Behelyettesítve:

$$\sum_{\kappa=0}^{n-1} \|e^{\kappa h\mathcal{A}} h^2 \varphi_2 S\|_{\mathbb{R}^{r \times r}} \leq \|n L h^2 \varphi_2 S\|_{\mathbb{R}^{r \times r}} .$$

A (2.1.8) miatt és a (2.1.5) valamint mert a $h < t_n$ miatt a következő igaz:

$$\|\varphi_2\|_{\mathbb{R}^{r \times r}} \leq \|\varphi_0\|_{\mathbb{R}^{r \times r}} = \|e^{h\mathcal{A}}\|_{\mathbb{R}^{r \times r}} < \|e^{t_n\mathcal{A}}\|_{\mathbb{R}^{r \times r}} .$$

Tegyük a következő elnevezést, ahol a $\Phi \in \mathbb{R}^+$:

$$\Phi := \|e^{t_n\mathcal{A}}\|_{\mathbb{R}^{r \times r}}$$

Mivel S, L, h, n és Φ eleve pozitívak, így a következő felső becslés is igaz:

$$\sum_{\kappa=0}^{n-1} \|e^{\kappa h \mathcal{A}} h^2 \varphi_2 S\|_{\mathbb{R}^r \times r} \leq \|n L h^2 \varphi_2 S\|_{\mathbb{R}^r \times r} \leq n L h^2 \Phi S = n h L h \Phi S = L h t_n \Phi S .$$

Tehát a következő lesz a fenti alapján:

$$\sum_{\kappa=0}^{n-1} \|e^{\kappa h \mathcal{A}} h^2 \varphi_2 S\|_{\mathbb{R}^r \times r} \leq L h t_n \Phi S . \quad (4.3.16)$$

A (4.3.12) és a (4.3.16) miatt a következő is igaz:

$$\sum_{\kappa=0}^{n-1} \|e^{\kappa h \mathcal{A}} \delta_{n-\kappa}\|_{\mathbb{R}^r \times r} \leq L h t_n \Phi S . \quad (4.3.17)$$

A (4.3.8) átírható a (4.3.17) segítségével a következő alakra:

$$\|\mathcal{E}_n\|_{\mathbb{R}^r} \leq \left\| h \sum_{\kappa=0}^{n-1} e^{(n-\kappa-1)h \mathcal{A}} \varphi_1 E_\kappa \right\|_{\mathbb{R}^r} + L h t_n \Phi S . \quad (4.3.18)$$

A (4.3.13) és a (4.3.15) miatt a következő felső becslés is igaz:

$$\|e^{(n-\kappa-1)h \mathcal{A}}\|_{\mathbb{R}^r \times r} = \|e^{(t_n - \kappa - 1) \mathcal{A}}\|_{\mathbb{R}^r \times r} \leq e^{(t_n - \kappa - 1) \Omega} \leq e^{\Omega t_n} = L .$$

Ezt a becslést behelyettesítve a (4.3.18) képletbe:

$$(4.3.18) \leq \left\| h L \sum_{\kappa=0}^{n-1} \varphi_1 E_\kappa \right\|_{\mathbb{R}^r} + L h t_n \Phi S . \quad (4.3.19)$$

A (4.3.7) és a (4.3.5) jelölést helyettesítsük be a tételben szereplő Lipschitz kritériumba:

$$\|E_n\|_{\mathbb{R}^r} \leq \mathcal{L} \|\mathcal{E}_n\|_{\mathbb{R}^r} .$$

Ezt helyettesítsük be a (4.3.19) egyenlőtlenségbe:

$$\|\mathcal{E}_n\|_{\mathbb{R}^r} \leq \left\| h L \mathcal{L} \varphi_1 \sum_{\kappa=0}^{n-1} \|\mathcal{E}_\kappa\|_{\mathbb{R}^r} \right\|_{\mathbb{R}^r} + L h t_n \Phi S \leq \|h L \mathcal{L} \varphi_1\|_{\mathbb{R}^r \times r} \sum_{\kappa=0}^{n-1} \|\mathcal{E}_\kappa\|_{\mathbb{R}^r} + L h t_n \Phi S .$$

A (2.1.8) és a (2.1.5) miatt, valamint mert a $h < t_n$ miatt a következő igaz:

$$\|\varphi_1\|_{\mathbb{R}^r \times r} \leq \|\varphi_0\|_{\mathbb{R}^r \times r} = \|e^{h \mathcal{A}}\|_{\mathbb{R}^r \times r} < \|e^{t_n \mathcal{A}}\|_{\mathbb{R}^r \times r} = \Phi .$$

Így a fentibe behelyettesíthetjük. A (2.1.8) miatt és a (2.1.5) valamint mert a $h < t_n$ miatt a következő igaz:

$$\|\mathcal{E}_n\|_{\mathbb{R}^r} \leq hL\mathcal{L}\Phi \sum_{\kappa=0}^{n-1} \|\mathcal{E}_\kappa\|_{\mathbb{R}^r} + Lht_n\Phi S \leq hL\mathcal{L}\Phi \sum_{\kappa=0}^{n-1} \|\mathcal{E}_\kappa\|_{\mathbb{R}^r} + Lht_n\Phi S .$$

Lemma. Most tekintsünk a diszkrét Gronwall lemmára, ahol $\epsilon_n \in \mathbb{R}^+$ egy nem negatív számsorozat és $\Lambda, \Xi \in \mathbb{R}^+$:

$$\epsilon_n \leq \Lambda h \sum_{\kappa=0}^{n-1} \epsilon_\kappa + \Xi , \quad (4.3.20)$$

ekkor létezik olyan $\mathfrak{L} \in \mathbb{R}^+$ amely Λ konstanstól függ. Ekkor az $\epsilon_n \leq \mathfrak{L}\Xi$ teljesül. Tegyük a következő megfeleltetéseket:

$$\begin{aligned} \epsilon_n &:= \|\mathcal{E}_n\|_{\mathbb{R}^r} , \\ \Lambda &:= L\mathcal{L}\Phi , \\ \Xi &:= Lht_n\Phi S . \end{aligned} \quad (4.3.21)$$

Ekkor a (4.3.20) átírva a (4.3.21) megfeleltetésekkel és átírva a diszkrét Gronwall lemma $\epsilon_n \leq \mathfrak{L}\Xi$ felső becslését a (4.3.21) behelyettesítéssel, a következőt kapjuk:

$$\|\mathcal{E}_n\|_{\mathbb{R}^r} \leq \mathfrak{L}Lht_n\Phi S .$$

Helyettesítsük vissza a (4.3.5) hibát a fentibe:

$$\|Y(t_n) - U(t_n)\|_{\mathbb{R}^r} \leq \mathfrak{L}Lht_n\Phi S .$$

Most a (4.2.2) definícióban szereplő \mathbf{L} konstansnak feleltessük meg a $\mathfrak{L}Lht_n\Phi S$ konstansok szorzatát. Ekkor a következőt kapjuk:

$$\|Y(t_n) - U(t_n)\|_{\mathbb{R}^r} \leq \mathbf{L}h .$$

A fenti egyenlőtlenség bal oldalán a (4.2.1) konvergencia definíciójában szereplő hiba van a t_n helyen nézve, míg a jobb oldalon a h távolságot leszámítva konstans.

A konvergenciában h értékkel tartunk a 0 számhoz, tehát a jobb oldal a 0 értékhez fog tartani, tehát a hiba is 0 felé tart. Az exponenciális Euler-módszer tehát első rendben konvergens, mert $h = h^1$ miatt a $\rho = 1$ teljesül. \square

4.4. A konzisztencia ismertetése

Definíció. Lokális hibának nevezzük a pontos és a közelítő megoldás helyenként vett különbségét normában, tehát az egy lépés alatt bekövetkező hibát feltéve, hogy az előző lépésben a hiba értéke 0 volt.

Definíció. Egy módszert konzisztensnek nevezünk, amennyiben a lokális hiba normában tart a 0 értékhez az $n \rightarrow \infty$ esetén.

Definíció. Egy módszer ρ rendben konzisztens, ha a lokális hiba $\mathcal{O}(h^{\rho+1})$ nagyságrendű.

4.5. A módszer konzisztenciája

Tétel. Tegyük fel, hogy a (2.2.3) teljesül $s = 1$ esetben, ekkor a (3.3.3) azaz az exponenciális Euler-módszer elsőrendben konzisztens.

Bizonyítás. Az elsőrendű konzisztencia bizonyításához, belátjuk az elsőrendű konzisztencia eléréshez szükséges rendszabályt. Az (1.3.9) egyenletet írjuk fel az alábbi alakban a t_n helyett a t_0 helyen nézve:

$$U(t_0 + h) = e^{hA} \cdot U(t_0) + \int_0^h e^{(h-z)A} \cdot G(t_0 + z, U(t_0 + z)) dz . \quad (4.5.1)$$

Ezzel a pontos megoldást írtuk fel. A (4.5.1) képletre, ha a (2.1.2) képletet a H függvény helyett a \mathcal{H} függvényre írjuk fel, és a (2.1) fejezetben megfelelően

levezetnénk, akkor a következőt kapnánk, ami a (2.1.9) lenne az U függvényre felírva a $t_n = t_0$ helyen nézve:

$$U(t_0 + h) = e^{hA} \cdot U(t_0) + \sum_{l=1}^{\infty} h^l \varphi_l \cdot \mathcal{H}^{(l-1)}(t_0) . \quad (4.5.2)$$

Az elsőrendű konzisztencia feltétel megállapításához a pontos megoldásnak és a közelítő megoldásnak egyeznie kell, legalább elsőrendű hibával. Ehhez tegyük egyenlővé a pontos megoldás (4.5.2) Taylor-sorral felírt alakját és a közelítő (1.6.11) megoldást a $t_0 + h$ helyen, valamilyen $\mathcal{O}(h^{\rho+1}) \in \mathbb{R}^r$ hibával:

$$\begin{aligned} U(t_0 + h) &= Y(t_0 + h) + \mathcal{O}(h^{\rho+1}) , \\ e^{hA} \cdot U(t_0) + \sum_{l=1}^{\infty} h^l \varphi_l \cdot \mathcal{H}^{(l-1)}(t_0) &= e^{hA} \cdot Y(t_0) + h \sum_{i=1}^s B_i \cdot \mathcal{K}_i + \mathcal{O}(h^{\rho+1}) . \end{aligned}$$

Az (1.2.2) miatt egyszerűsíthetünk, ekkor a fenti a következőképpen néz ki:

$$\sum_{l=1}^{\infty} h^l \varphi_l \cdot \mathcal{H}^{(l-1)}(t_0) = h \sum_{i=1}^s B_i \cdot \mathcal{K}_i + \mathcal{O}(h^{\rho+1}) .$$

Mivel az exponenciális Euler-módszer egy lépcsős módszer, így emiatt az $s = 1$ teljesül. Az (1.6.12), a $c_1 = 0$ és mivel most t_0 helyen nézzük, ezért a \mathcal{K}_1 átírható. Így a fenti a következő alakú:

$$\sum_{l=1}^{\infty} h^l \varphi_l \mathcal{H}^{(l-1)}(t_0) = h B_1 \mathcal{K}_1 + \mathcal{O}(h^{\rho+1}) = h B_1 G(t_0, Y(t_0)) + \mathcal{O}(h^{\rho+1}) . \quad (4.5.3)$$

Most ha a végtelen összeg $l = 1$ esetét külön kiírjuk, akkor a következő lesz:

$$\sum_{l=1}^{\infty} h^l \varphi_l \cdot \mathcal{H}^{(l-1)}(t_0) = h \varphi_1 \mathcal{H}(t_0) + \sum_{l=2}^{\infty} h^l \varphi_l \cdot \mathcal{H}^{(l-1)}(t_0) .$$

Használjuk fel a (3.3.1) képletet $t = t_0$ esetre, ekkor a fenti a következő lesz:

$$\sum_{l=1}^{\infty} h^l \varphi_l \cdot G^{(l-1)}(t_0, U(t_0)) = h \varphi_1 G(t_0, U(t_0)) + \sum_{l=2}^{\infty} h^l \varphi_l \cdot G^{(l-1)}(t_0, U(t_0)) .$$

Ez felhasználva a (4.5.3) képletre:

$$h\varphi_1 G(t_0, U(t_0)) + \sum_{l=2}^{\infty} h^l \varphi_l \cdot G^{(l-1)}(t_0, U(t_0)) = hB_1 G(t_0, Y(t_0)) + \mathcal{O}(h^{\rho+1}) .$$

Tegyük a következő megfeleltetést:

$$\mathcal{O}(h^{\rho+1}) = \sum_{l=2}^{\infty} h^l \varphi_l \cdot G^{(l-1)}(t_0, U(t_0)) . \quad (4.5.4)$$

Ekkor a $h^{\rho+1} = h^2$ tehát $\rho = 1$, így ez lesz az elsőrendű hiba. A következő marad:

$$h\varphi_1 G(t_0, U(t_0)) = hB_1 G(t_0, Y(t_0)) .$$

Az (1.2.3) miatt egyszerűsíthetünk a G függvénnyel és a h értékkel, ekkor a következő lesz:

$$\varphi_1 = B_1 . \quad (4.5.5)$$

Ez lesz az exponenciális Euler-módszer konzisztencia szabálya, ha a (4.5.4) képletben szereplőt vesszük a hibának, és ekkor elsőrendben konzisztens. Ez új információt nem ad, mert így alkottuk meg eredetileg. Ezzel megkaptuk, hogy az exponenciális Euler-módszer elsőrendben konzisztens, amennyiben a (4.5.5) teljesül. \square

4.6. Példák számolással és MATLAB kódokkal

Példa. Tekintsünk az alábbi differenciálegyenletre, ahol $u : \mathbb{T} \rightarrow \mathbb{R}$ és az $a \in \mathbb{R}$, amit mi most 5 értéknek választunk, valamint az $u_0 = y_0 \in \mathbb{R}$ pedig a 2 számnak:

$$\begin{cases} u'(t) = 5 \cdot u(t) + \sin(u(t)) , t > 0 \\ u(0) = 2 . \end{cases}$$

Ezt a példát most megoldjuk az explicit exponenciális Euler-módszerrel. Legyen most $h = 0.5$ és $n = 2$. Ekkor a (3.3.3) képlet felhasználva $y : \mathbb{T} \rightarrow \mathbb{R}$ függvénnyel:

$$y(0 + 0.5) = e^{0.5 \cdot 5} \cdot y(0) + \int_0^{0.5} e^{(0.5-z) \cdot 5} \cdot \sin(y(0)) dz .$$

Mivel $u(0) = y(0) = 2$, a fenti egyenlet átírható az alábbira:

$$\begin{aligned}
 y(0.5) &= e^{2.5} \cdot 2 + \int_0^{0.5} e^{(0.5-z) \cdot 5} \cdot \sin(2) \, dz = 24.36 + 11.0775 \cdot \int_0^{0.5} e^{-5z} \, dz = \\
 &= 24.36 + 11.0775 \cdot \left[\frac{-e^{-5z}}{5} \right]_0^{0.5} = 24.36 + 11.0775 \cdot \left(\frac{1}{5} - \frac{e^{-2.5}}{5} \right) = 26.39 , \\
 y(0.5) &= 26.39 .
 \end{aligned}$$

Most tegyünk még egy lépést:

$$\begin{aligned}
 y(0.5 + 0.5) &= e^{0.5 \cdot 5} \cdot y(0.5) + \int_0^{0.5} e^{(0.5-z) \cdot 5} \cdot \sin(y(0.5)) \, dz , \\
 y(1) &= e^{2.5} \cdot 26.39 + \int_0^{0.5} e^{(0.5-z) \cdot 5} \cdot \sin(26.39) \, dz = \\
 &= 321.5 + 11.5886 \cdot \left[\frac{-e^{-5z}}{5} \right]_0^{0.5} = 321.5 + 11.5886 \cdot \left(\frac{1}{5} - \frac{e^{-2.5}}{5} \right) = 323.63 , \\
 y(1) &= 323.63 .
 \end{aligned}$$

A MATLAB programkódja a módszernek, ahol a függvény bemeneti adatnak várja a lépésközt, a lépésszámot, egy $a \in \mathbb{R}$ számot, illetve azt, hogy melyik lépésnél írassa ki az eredményt és egy $y_0 \in \mathbb{R}$ kezdeti értéket:

```

function [ t y ]=EEEM(h, n, a, t, y_0)
syms z;
y(1)=y_0;
for i=1:n
y(i+1)=exp(h*a)*y(i)+int(exp((h-z)*a)*sin(y(i)), z, 0, h);
x(i+1)=i*h;
end
t=y(t);
plot(x,y);

```

A program az EEM függvény meghívására az alábbi bemeneti adatokkal a következő eredményeket adja:

$$EEM(0.5, 2, 5, 1, 2) = 2$$

$$EEM(0.5, 2, 5, 2, 2) = 26.3986$$

$$EEM(0.5, 2, 5, 3, 2) = 323.7345$$

Ez a számolásunkhoz viszonyítva csak kerekítésbeli hibában tér el.

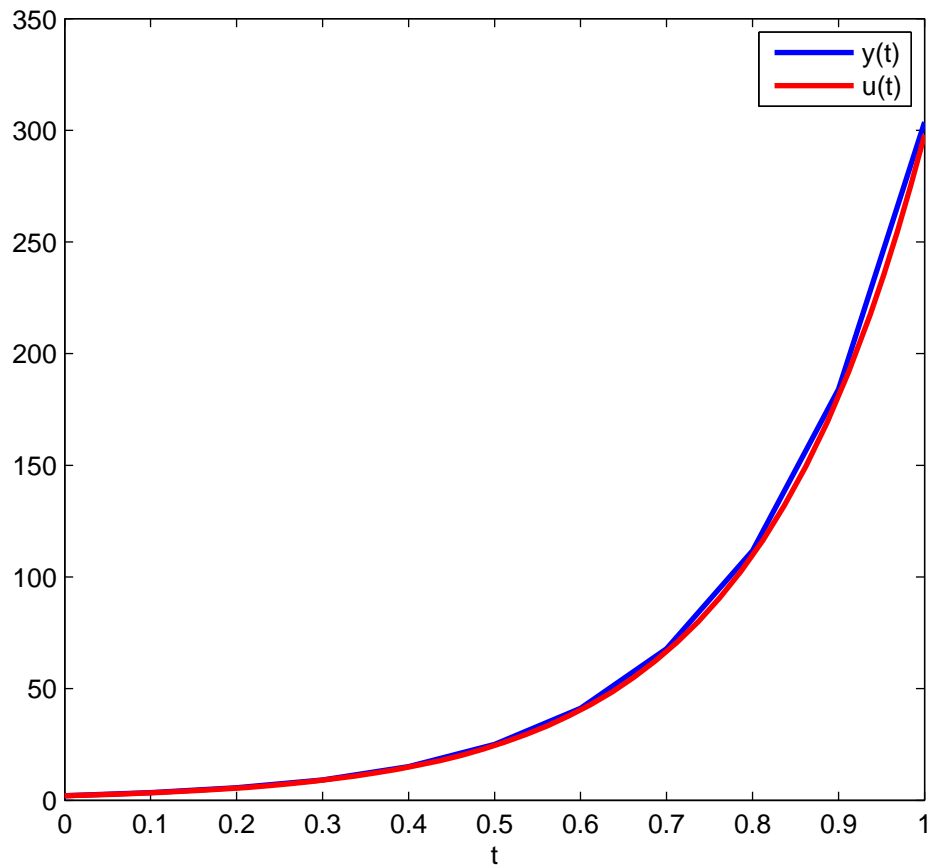
Most ugyanezt a feladatot oldjuk meg a MATLAB beépített ODE45 megoldókészletével is. Mi most ennek az ODE45 megoldókészlet megoldását tekintjük a pontos megoldásnak, tehát az u függvénynek. Az ODE45 egy negyedrendű, változó lépésszámú módszer, és bemeneti adat a megoldandó függvény, az intervallum, és a kezdeti érték:

```
[t, u]=ode45(@(t, u) 5*u+sin(u), [0 1], 2);  
plot(t, u)
```

Tekintsük a két függvényt egy ábrában, az alábbi MATLAB kóddal, ahol a numerikus módszer bemenete a következő:

EEEMD(0.1, 10, 5, 11, 2)

```
function [t y]=EEEMD(h, n, a, t, y_0)
syms z;
y(1)=y_0;
for i=1:n
y(i+1)=exp(h*a)*y(i)+int(exp((h-z)*a)*sin(y(i)), z, 0, h);
x(i+1)=i*h;
end
t=y(t);
plot(x,y);
hold on
[t,u]=ode45(@(t,u) 5*u+sin(u), [0 1], 2);
plot(t,u)
hold off
```

Pontos (ODE45) és exponenciális Euler megoldás

Amennyiben a lépésközt $h = 0.1$ értéknek vesszük, az exponenciális Euler már nagyon jól fogja közelíteni az ODE45 általi megoldást.

Példa. Oldjuk meg a következő példát, ahol $r = 2$ az U függvény dimenziója:

$$\begin{cases} U'(t) = \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} U(t) + \sqrt{U(t)}, t > 0 \\ U(0) = \begin{pmatrix} 11 \\ 9 \end{pmatrix}. \end{cases}$$

Ezt exponenciális Euler-módszerrel fogjuk megoldani, ahol a $h = 0.5$ és az $n = 2$ teljesül. Ekkor a (3.3.3) képlet felhasználva és az értékek behelyettesítve:

$$Y(0.5) = \exp \left(0.5 \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} \right) \begin{pmatrix} 11 \\ 9 \end{pmatrix} + \int_0^{0.5} \exp \left((0.5 - z) \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} \right) \sqrt{\begin{pmatrix} 11 \\ 9 \end{pmatrix}} dz .$$

Kiszámítva:

$$Y(0.5) = \begin{pmatrix} 422.2 \\ 1098.9 \end{pmatrix} + \left(\exp \left(0.5 \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} \right) - I \right) \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix}^{-1} \sqrt{\begin{pmatrix} 11 \\ 9 \end{pmatrix}},$$

$$Y(0.5) = \begin{pmatrix} 437.7 \\ 1137.5 \end{pmatrix} .$$

A következő lépés:

$$Y(1) = \exp \left(0.5 \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} \right) \begin{pmatrix} 437.7 \\ 1137.5 \end{pmatrix} + \int_0^{0.5} \exp \left((0.5 - z) \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} \right) \sqrt{\begin{pmatrix} 437.7 \\ 1137.5 \end{pmatrix}} dz .$$

Elvégezve:

$$Y(1) = \begin{pmatrix} 37210 \\ 97955.4 \end{pmatrix} .$$

A MATLAB programkódja a módszernek, ahol a függvény bemenetként várja a lépésközt, a lépésszámot, egy $A \in \mathbb{R}^{2 \times 2}$ mátrixot, és azt, hogy melyik lépésnél írassa ki az eredmény vektort és egy $Y_0 \in \mathbb{R}^2$ kezdeti vektort:

```
function [t Y]=EEEMM(h, n, A, t, Y_0)
syms z;
Y(:,1)=Y_0;
for i=1:n
Y(:,i+1)=expm(h*A)*Y(:,i)+int(expm((h-z)*A)*sqrt(Y(:,i)), z
    , 0, h);
x(i+1)=i*h;
end
t=Y(:,t);
plot(x,Y(1,:));
figure;
plot(x,Y(2,:));
```

A program meghívása az alábbi bemeneti adatokkal a következő eredményeket adja:

$$EEEMM(0.5, 2, [13; 57], 1, [11; 9]) = [11; 9]$$

$$EEEMM(0.5, 2, [13; 57], 2, [11; 9]) = [437.7459; 1137.5379]$$

$$EEEMM(0.5, 2, [13; 57], 3, [11; 9]) = [37211.7736; 97960.1452]$$

Ez a számolásunkhoz viszonyítva csak kerekítésbeli hibában tér el.

Ugyanezt a feladatot oldjuk meg a beépített ODE45 megoldókészlettel is:

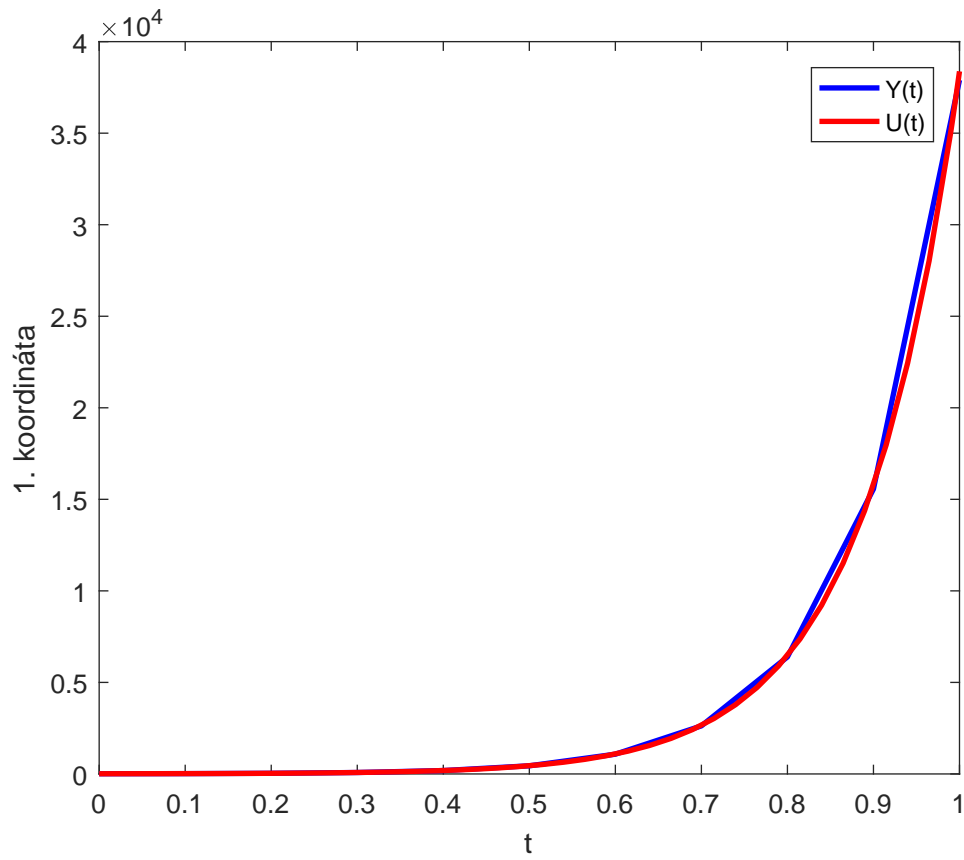
```
[t, U]=ode45(@(t, U) [1 3; 5 7]*U+sqrt(U), [0 1], [11; 9]);
plot(t, U(:,1));
figure;
plot(t, U(:,2));
```

Tekintsük a numerikusan és a beépített megoldóval megoldott függvényt egy ábrában az alábbi programmal az első koordinátában, ahol a numerikus módszer bemenete legyen az alábbi:

EEEMMD1(0.1, 10, [13; 57], 10, [11; 9])

```
function [t Y]=EEEMMD1(h, n, A, t, Y_0)
syms z;
Y(:,1)=Y_0;
for i=1:n
Y(:,i+1)=expm(h*A)*Y(:,i)+int(expm((h-z)*A)*sqrt(Y(:,i)), z
    , 0, h);
x(i+1)=i*h;
end
t=Y(:,t);
plot(x,Y(1,:));
hold on
[t, U]=ode45(@(t, U) [1 3; 5 7]*U+sqrt(U), [0 1], [11; 9]);
plot(t, U(:,1));
hold off
```

A következő ábrán látszik, hogy a két függvény eltérése az első koordinátában minimális.



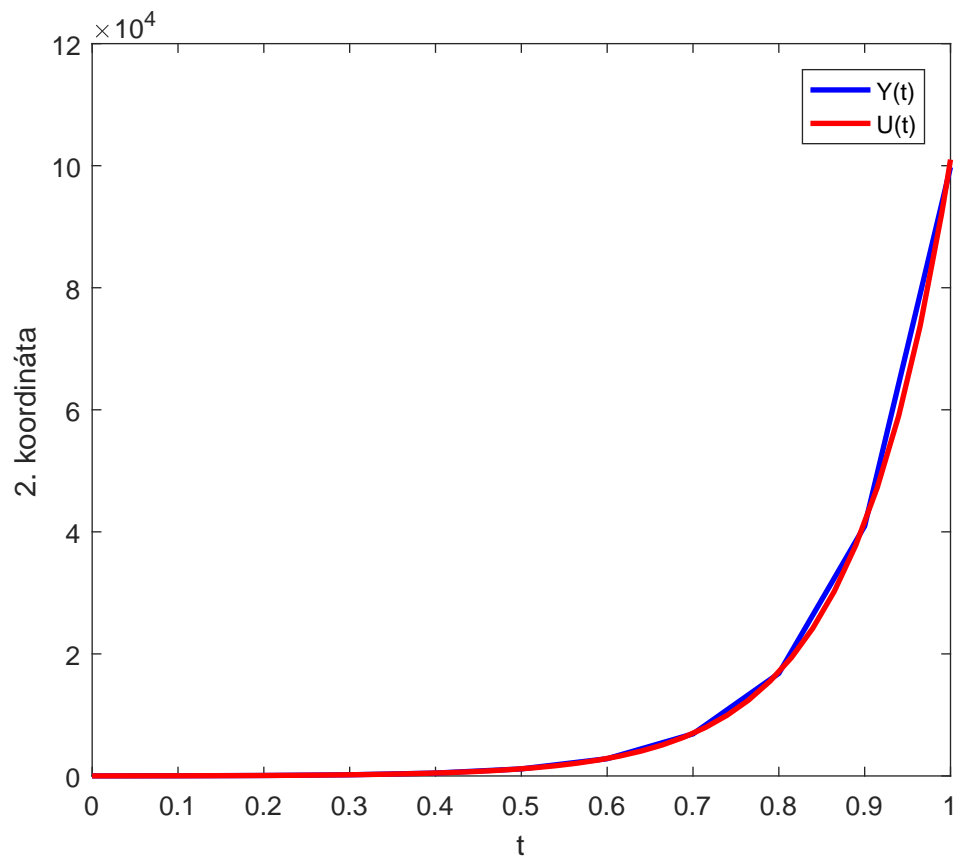
Pontos (ODE45) és exponenciális Euler megoldás az első koordinátában

Most az alábbi programmal ábrázoljuk a második koordinátában:

EEEMMD2(0.1, 10, [13; 57], 10, [11; 9])

```
function [t Y]=EEEMMD2(h, n, A, t, Y_0)
syms z;
Y(:,1)=Y_0;
for i=1:n
Y(:,i+1)=expm(h*A)*Y(:,i)+int(expm((h-z)*A)*sqrt(Y(:,i)), z
, 0, h);
x(i+1)=i*h;
end
t=Y(:,t);
plot(x,Y(2,:));
hold on
[t, U]=ode45(@(t, U) [1 3; 5 7]*U+sqrt(U), [0 1], [11; 9]);
plot(t, U(:,2));
hold off
```

A következő ábrán látszik, hogy a második koordinátában is jól közelít.



Pontos (ODE45) és exponenciális Euler megoldás a második koordinátában

Amennyiben a lépésközt 0.1 értéknek vesszük, az exponenciális Euler-módszer jól közelíti a pontos megoldást.

Összefoglalás

A szakdolgozatban levezettük az exponenciális Runge–Kutta módszert, és felírtuk a Taylor-soros alakját. Bevezettük a φ jelölést az együtthatók elemzéséből, amit felhasználtunk utána. Néhány módszert levezettünk, majd egy speciális módszer, az exponenciális Euler-módszer elsőrendű konvergenciáját és konzisztenciáját részletesen elemeztük. Erre a módszerre tekintettünk példákat, amik lefedték az egy és a két dimenziós esetet. A több dimenziós esetben hasonló a megoldási menet, mint a két dimenziós esetben.

Saját készítésű MATLAB programmal illusztráltuk a közelítő és a pontos megoldást egy ábrában. Egy és két dimenziós esetre is néztünk feladatot programmal, ábrával.

Ebben a témában még sok mindent lehetne tárgyalni. Mi csak elsőrendű módszert néztünk, de vannak magasabb rendűek is. Egy módszernek minél magasabb a rendje, annál több rendszabály lesz. Minden egyes rendszabálynak megfeleltethető egy két színezésű fa. A rendszabályok és a lépcsőszámok közt is van kapcsolat, minél magasabb rendű módszert használunk, a rend eléréséhez szükséges lépcsőszám is magasabb. Adott lépcsőszám esetén van maximálisan elérhető konzisztencia rend. A lépcsőszám és rendszabályok számára jelenleg nem létezik képlet. A fa megfeleltetéseket elsősorban az elemi differenciáknak feleltették meg, és mivel a Runge–Kutta módszer rendszabályai is előállíthatóak annak megfelelően, így a rendszabályok felírhatóak az exponenciális Runge–Kutta módszerek esetében is.

Felhasznált jelölések

Terek:

$$\mathbb{N}^+, \mathbb{R}, \mathbb{R}^+, \mathbb{R}^r, \\ \mathbb{R}^{r \times r}, \mathbb{R}^{r \times s}, \\ \mathbb{T} = [0, q], [0, 1]$$

Változók:

$$r, s, \rho \in \mathbb{N}^+ \\ h, q, \epsilon_n \in \mathbb{R}^+ \\ a, \mathbf{a}_{\iota, \zeta}, d_i, k_{\zeta, i}, \\ y_0, u_0 \in \mathbb{R} \\ t, t_n, w, z, z_i, \\ \sigma, \varkappa, \varsigma \in \mathbb{T} \\ c_i \in [0, 1] \\ i, j, k = 1, \dots, s \\ l = 1, 2, \dots \\ \check{l} = 0, 1, \dots \\ \hat{l} = 1, \dots, l \\ n = 0, 1, \dots, N \\ \zeta, \iota = 1, \dots, r \\ \kappa = 0, 1, \dots, n - 1$$

Mátrixok:

$$E_n, \mathcal{E}_n, K_i, \mathcal{K}_i, K_j, \\ \mathcal{O}, U_0, U_1, U_2, \delta_n \in \mathbb{R}^r \\ \mathcal{A}, A_i, A_{i,j}, B, B_i, I, \\ \varphi_l, \varphi_{l,i} \in \mathbb{R}^{r \times r} \\ K \in \mathbb{R}^{r \times s}$$

1-D függvények:

$$g_\iota : \mathbb{T} \times \mathbb{R}^r \rightarrow \mathbb{R} \\ \ell_i : \mathbb{T} \rightarrow \mathbb{R} \\ m : \mathbb{T} \rightarrow \mathbb{R} \\ m_\iota : \mathbb{T} \rightarrow \mathbb{R} \\ p : \mathbb{T} \rightarrow \mathbb{R} \\ p_\iota : \mathbb{T} \rightarrow \mathbb{R} \\ u : \mathbb{T} \rightarrow \mathbb{R} \\ u_\iota : \mathbb{T} \rightarrow \mathbb{R} \\ y : \mathbb{T} \rightarrow \mathbb{R} \\ v : \mathbb{T} \rightarrow \mathbb{R}$$

Állandók:

$$N \in \mathbb{N}^+ \\ L, \mathbf{L}, \mathbf{L}, \mathcal{L}, \mathcal{L}, S, \\ \Lambda, \Phi, \Xi, \Omega \in \mathbb{R}^+$$

Több-D függvények:

$$F : \mathbb{T} \times \mathbb{R}^r \rightarrow \mathbb{R}^r \\ G : \mathbb{T} \times \mathbb{R}^r \rightarrow \mathbb{R}^r \\ H : \mathbb{T} \rightarrow \mathbb{R}^r \\ \mathcal{H} : \mathbb{T} \rightarrow \mathbb{R}^r \\ M : \mathbb{T} \rightarrow \mathbb{R}^r \\ O : \mathbb{T} \rightarrow \mathbb{R}^r \\ P : \mathbb{T} \rightarrow \mathbb{R}^r \\ R : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{R}^r \\ U : \mathbb{T} \rightarrow \mathbb{R}^r \\ Y : \mathbb{T} \rightarrow \mathbb{R}^r \\ \Gamma : \mathbb{T} \rightarrow \mathbb{R}^r \\ \Psi : \mathbb{T} \rightarrow \mathbb{R}^r \\ \Theta : \mathbb{T} \rightarrow \mathbb{R}^{r \times r}$$

Irodalomjegyzék

- [1] Faragó: *Numerikus modellezés és közönséges differenciálegyenletek numerikus megoldási módszerei*, Online elérhető egyetemi jegyzet 2013
http://web.cs.elte.hu/~faragois/jegyzet_Szeged.pdf
- [2] Faragó–Horváth: *Numerikus Módszerek*, Typotex 2013
<http://tankonyvtar.ttk.bme.hu/pdf/30.pdf>
- [3] Hochbruck: *A short course on exponential integrators*, Gene Golub SIAM summer school in Shanghai 2013
http://g2s3.cs.ucdavis.edu/lecturers/Hochbruck/Summary_Hochbruck.pdf
- [4] Hochbruck–Ostermann: *Exponential integrators*, Acta Numerica, Cambridge University Press 2010
<http://www.math.kit.edu/ianm3/lehre/semjahnke2012w/media/exp-int-acta-numerica.pdf>
- [5] O’Callaghan: *The Analysis and Implementation of Exponential Almost Runge-Kutta Methods for Semilinear Problems*, A Dissertation Submitted for the Degree of Doctor of Philosophy 2011
<http://www.binarynotions.com/public/publications/EoinOCallaghan-EiThesis.pdf>
- [6] Simon–Tóth: *Differenciálegyenletek*, Typotex 2005
<https://books.google.hu/books?isbn=9639326461>
- [7] Sofyalioglu: *Numerical Solutions of the Reaction-Diffusion Equations by Exponential Integrators*, A Thesis Submitted to the Graduate School of Engineering and Sciences of Izmir Institute of Technology in Partial Fulfillment of the Requirements for the Degree of Master of Science in Mathematics 2014
<http://library.iyte.edu.tr/tezler/master/matematik/T001241.pdf>