

9. fejezet

Közönséges differenciálegyenletek numerikus megoldása

Az integrálás inverz művelete a differenciálás, így, ha

$$y(x) = c + \int_{x_0}^x f(t, y(t)) dt$$

ahol $c = \text{const}$, akkor ezt differenciálva kapjuk

$$\frac{dy}{dx} = f(x, y(x)),$$

mert az integrált felső határa szerint differenciálni úgy kell, hogy a felső határt behelyettesítjük az integranduszba.

A láncszabály szerint, ha $y = f(x)$ és $x = g(t)$, azaz $y = f(x(t))$, akkor

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt} = \frac{dy}{dx} \cdot \frac{dg}{dt} = f'(x) \cdot g'(t).$$

Ha pedig $y = f(x, z)$ és $z = g(x)$, akkor

$$\frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z} \frac{dg}{dx}.$$

9.1. Motiváció

Az időtől függő műszaki, természettudományos és gazdasági folyamatok matematikai leírását szolgáló modellekben fontos szerepe van a közönséges diffe-

renciálegyenleteknek – olyannyira, hogy a „szimuláció” szó sok szakember számára azt jelenti, hogy differenciálegyenleteket megoldani.

Tekintsünk egy ilyen esetet, amely szóbeli információ szétterjedését modellezi. Ehhez hasonló egy új termék eladása egy adott országban.

Ahhoz, hogy az új hír továbbterjedjen, kell, hogy egy ember, aki ismeri (ezeknek sűrűsége a teljes lakosságban legyen y), találkozzon olyannal, aki nem ismeri (értelemszerűen ezeknek sűrűsége $1 - y$). Annak valószínűsége, hogy ez történik, az $y(1 - y)$. A kontaktus eredményessége jelezhető egy (mondjuk: α) pozitív konstanssal: attól, hogy létrejött a híratadás, mennyivel növekedett azoknak száma (a teljes lakossághoz képest), akik tudják a hírt és tovább is adják? Ezenkívül arra számítunk, hogy minél több Δt időt hagyunk az információátadására, annál nagyobb lesz a sűrűség növekménye: első közelítésben ez a sűrűség lineárisan fog növekedni Δt -vel.

Ha tehát előbb egy t pillanatban $y(t)$ volt a sűrűség, akkor a kontaktus után, a $t + \Delta t$ pillanatban, már

$$y(t + \Delta t) = y(t) + \Delta \alpha y(t)(1 - y(t))$$

az új sűrűség – avagy

$$\frac{y(t + \Delta t) - y(t)}{\Delta t} = \alpha y(t)(1 - y(t)).$$

Ha itt egyre kisebb Δt -re gondolunk és y -t differenciálhatónak tekintjük, akkor a $\Delta t \rightarrow 0$ határátmenettel következik az

$$y'(t) = \alpha y(1 - y) \quad (9.1)$$

differenciálegyenlet.

Amennyiben a $t = 0$ kezdeti pillanatban $y(t) = 0$, akkor új hír nem is volt és így nem is tud szétterjedni: $y(t) = 0$ marad minden további t -re. Ha viszont a sűrűség $1 > y(0) > 0$, akkor arra számítunk, hogy $y(t)$ t -vel növekedni fog.

Ez igaz, ahogyan a differenciálegyenlet analitikus megoldásán lehet ellenőrizni, amelynek képlete:

$$y(t) = \frac{1}{1 + \beta e^{-\alpha t}}, \quad 1 > y(0) = \frac{1}{1 + \beta} > 0 \text{ vagyis } \beta = \frac{1 - y(0)}{y(0)} > 0. \quad (9.2)$$

Ellenőrizzük ezt a megoldást!

$$\frac{dy}{dt} = -\frac{1}{(1 + \beta e^{-\alpha t})^2} \frac{d(1 + \beta e^{-\alpha t})}{dt} = -\frac{1}{(1 + \beta e^{-\alpha t})^2} (-\alpha \beta e^{-\alpha t}).$$

De

$$1 - y(t) = \frac{1 + \beta e^{-\alpha t} - 1}{1 + \beta e^{-\alpha t}} = \frac{\beta e^{-\alpha t}}{1 + \beta e^{-\alpha t}},$$

és így valóban

$$\alpha y(1 - y) = \alpha \frac{\beta e^{-\alpha t}}{(1 + \beta e^{-\alpha t})^2}.$$

Eszerint a (9.2) képlet adja a (9.1) differenciálegyenlet megoldását, és az $y(0)$ kezdetiértékből a β konstanszt meg tudjuk határozni.

Következőnek egy mechanikai példát említünk, egy rugón felfüggesztett testet. A független változó a t idő, x a keresett függvény: a test kitérése a normálhelyzetből, és $\dot{x} := \frac{dx}{dt}$ annak t szerinti deriváltja (a test sebessége), valamint \ddot{x} a t szerinti második deriváltja (a test gyorsulása). Egy $r\dot{x}$ taggal a súrlódást is figyelembe vesszük, valamint az $f(t)$ külső erőt:

$$m\ddot{x} + r\dot{x} + kx = f(t).$$

Az m (tömeg) és k (rugóállandó) konstansokról feltesszük, hogy pozitívak.

Ezt a másodrendű egyenletet elsőrendű egyenletek rendszerébe alakítjuk át, a következő jelöléseket használva:

$$y_1(t) := x(t), \quad y_2(t) := \dot{x}(t),$$

úgyhogy

$$\begin{aligned} \dot{y}_1 &= \dot{x} = y_2, \\ \dot{y}_2 &= \ddot{x} = \frac{1}{m}(-r\dot{x} - kx + f) = \frac{1}{m}(-ry_2 - ky_1 + f). \end{aligned}$$

Ezt vektoralakban összefoglalva:

$$y := \begin{pmatrix} x \\ \dot{x} \end{pmatrix}, \quad A := -\frac{1}{m} \begin{pmatrix} 0 & -m \\ k & r \end{pmatrix}, \quad g := \begin{pmatrix} 0 \\ \frac{f}{m} \end{pmatrix},$$

kapjuk a differenciálegyenletek rendszerét:

$$\dot{y} = Ay + g.$$

Kémiai folyamatok is leírhatók differenciálegyenletekből álló rendszerek segítségével, de ezek — eltekintve a legegyszerűbb esetektől — nemlineárisak, például a környezetvédelem szempontjából fontos szmog-modellek, amelyekben kb. 60 reakciót és komponenst kell figyelembe venni.

Példáink a közönséges differenciálegyenleteknek két problémáját mutatják meg (amelyek megnehezítik, ill. gyakorlatilag lehetetlenné teszik az analitikus megoldást):

- 1) tipikusan rendszerekről van szó;
- 2) az egyenletek gyakran nemlineárisak a keresett változóknban.

9.2. Kezdetiérték feladatok

Ahogy példánk mutatják, a differenciálegyenlet információt ad a keresett $y(x)$ függvény deriváltjáról, de nem adja meg azt közvetlenül. A tipikus alak

$$y' = \frac{dy}{dx} = f(x, y), \quad (9.3)$$

ahol $f(x, y)$ adott függvény és keresett az $y(x)$ függvény, pl. a $[0, 1]$ intervallumban. Ha más intervallum érdekes, pl. az $[x_0, x_L]$ intervallum, akkor az $x = x_0 + (x_L - x_0)t$ transzformációval az előzőre tudjuk visszavezetni ($t \in [0, 1]$). Ekkor ugyanis

$$y(x) = y(x_0 + (x_L - x_0)t) =: z(t), \quad f(x, y) = f(x_0 + (x_L - x_0)t, z),$$

így

$$\frac{dz}{dt} = \frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt} = f(x, y) \cdot (x_L - x_0) = f(x_0 + (x_L - x_0)t, z) \cdot (x_L - x_0) =: g(t, z).$$

Ez lényegében ugyanolyan differenciálegyenlet, mint az eredeti.

Maga a (9.3) egyenlet bizony nem elég ahhoz, hogy a keresett y függvényt egyértelműen megkapjuk, még abban a legegyszerűbb esetben sem, amikor $f(x, y)$ azonosan nulla: ilyenkor világos, hogy $y(x) = \text{const}$ az általános megoldás, mert egy (deriválható) függvény deriváltja pontosan akkor nulla, ha az a függvény konstans. De hogy mennyi a konstans, azt az $y' = 0$ differenciálegyenlet nem mondja meg. Tehát további információ kell még.

Egy ilyen adat, amellyel gyakran rendelkezünk, az $y(0)$ kezdetiérték. A (9.1) differenciálegyenlet (9.2) megoldásában ez az érték mint paraméter szerepel, amelynek figyelembevétel után a megoldás meg van határozva.

A mechanikai példában az $y(0)$ egy vektor, amely tartalmazza az $x(0)$ kezdeti helyzetet és a $\frac{dx}{dt}(0)$ kezdeti sebességet. Itt egy explicit megoldási képlet elég bonyolult lenne, de fizikailag világos, hogy a kezdeti pozíció és sebesség nélkül nincsen teljes információnk a felfüggesztett test állapotáról.

Matematikailag viszont azért következik, hogy egy kezdetiérték meghatározza a megoldást, mert a (9.3) egyenlet integrálja

$$y(x) - y(x_0) = \int_{x_0}^x f(t, y(t)) dt. \quad (9.4)$$

Ez a képlet mutatja, hogy a lehetséges megoldások közül egyet úgy választhatunk ki, hogyha valamelyik x_0 pontban a megoldás $y(x_0)$ értékét megadjuk, pl. $x_0 = 0$ -ban.

Ezzel a *kezdetiérték feladat* teljes:

Keresett a (9.3) differenciálegyenlet azon megoldása, amely 0-ban egy adott $y(0)$ értéket vesz fel:

$$\frac{dy}{dx} = f(x, y(x)), \quad y(0) \text{ adott, } y(x) \text{ keresett } x \in [0, 1]\text{-re.} \quad (9.5)$$

Az $y(0)$ megadása után, amennyiben f nem függ y -től, a (9.4) közvetlenül mutatja a megoldás képletét: $y(x) = y(0) + \int_0^x f(t) dt$. Itt persze lehetséges, hogy az integrandusz törzsfüggvénye ismeretlen – ami miatt egy numerikus közelítésre kell gondolnunk.

De az általános esetben is, amikor f függhet y -től, $y(0)$ megadása után a megoldás meg van határozva, és a megoldás közelítő meghatározását (9.4)-ből éppen az alábbiakban fogjuk tanulni.

Ahhoz viszont, hogy (9.4) matematikailag értelmes legyen, bizonyos feltételek kelljenek f -re nézve – ahogyan nem létezik minden egyváltozós függvény integrálja sem, hanem csak a Riemann-integralható függvényeké.

A közönséges differenciálegyenleteknél célszerű, a következő (elégséges) feltételekből kiindulni:

1) $f : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ legyen folytonos az első argumentumában (ez x), és

2) Lipschitz-folytonos a másodikban (ez y):

$$|f(x, y) - f(x, z)| \leq L|y - z| \quad \text{minden } x \in [0, 1] \text{ és } y, z \in \mathbb{R}\text{-re.}$$

Ebben a második feltételben y vagy z nem a megoldás, nem függvények, hanem bármely két valós szám. Az is elég, ha a Lipschitz-folytonosság feltételében y és z minden olyan érték, amelyet a megoldás fel tudna venni.

A fenti két elégséges feltétel – x szerinti folytonosság és y szerinti Lipschitz-folytonosság – nemcsak a (9.5) kezdetiérték feladat megoldásának létezését biztosítja, hanem az egyértelműségét is.

Példák: Elsőként vizsgáljuk a (9.1) egyenlet $f = f(y) = \alpha y(1 - y)$ jobboldalát. A fenti két feltétel közül az első teljesül: az $f(x, y) = \alpha y(1 - y)$ függvény x -re nézve konstans, tehát folytonos. Vizsgáljuk a függvény y szerinti Lipschitz-folytonosságát is:

$$f(y) - f(z) = \alpha(y(1 - y) - z(1 - z)) = \alpha(y - z - y^2 + z^2) = \alpha(y - z)(1 - y - z),$$

így

$$|f(y) - f(z)| \leq \alpha|y - z||1 - y - z|,$$

vagyis: f általában nem Lipschitz-folytonos, mert $|1 - y - z|$ tetszőlegesen nagy lehet. De amennyiben eleve tudjuk, hogy $0 < y, z \leq 1$ (mint ahogyan az explicit megoldási képletből következik $0 < y \leq 1$, ha $0 < y(0) \leq 1$), akkor viszont $|1 - y - z| \leq 1$, tehát

$$|f(y) - f(z)| \leq \alpha|y - z|,$$

vagyis ilyen argumentumokra f Lipschitz-folytonos és $\alpha =: L$ a Lipschitz-állandó. Ez azért megnyugtató a (9.2) pontos megoldás ismeretében, mert akkor ez az egyetlen megoldás, ha $0 < y(0) \leq 1$.

Második példaként nézzük azt az esetet, hogy $f(x, y) = \sqrt{y}$, az intervallum $[0, 1]$ és a kezdetiérték $y(0) = 0$. Tehát a kezdetiérték feladat:

$$y' = \sqrt{y}, \quad 0 \leq x \leq 1, \quad y(0) = 0.$$

Azt állítjuk, hogy ennek a feladatnak kettő megoldása van:

1) $y \equiv 0$, ami világos, hiszen ekkor $y' \equiv 0$ és $f(x, y) \equiv 0$.

2) $y = (\frac{x}{2})^2$. Most $y' = \frac{d}{dx} \frac{x^2}{4} = \frac{2x}{4} = \frac{x}{2}$, ami éppen \sqrt{y} .

Mindkét megoldás a kezdetiértéket is felveszi. A fenti két feltétel közül az első teljesül: az $f(x, y) = \sqrt{y}$ függvény x -re nézve megint konstans, tehát folytonos. A probléma most is az y szerinti Lipschitz-folytonosság: ha például $y > 0$ és $z \geq 0$, akkor

$$f(y) - f(z) = \sqrt{y} - \sqrt{z} = \frac{y - z}{\sqrt{y} + \sqrt{z}},$$

így

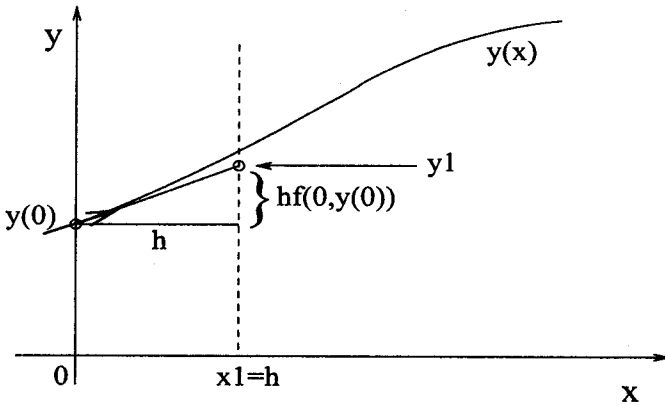
$$|f(y) - f(z)| = \frac{|y - z|}{\sqrt{y} + \sqrt{z}} \leq \frac{|y - z|}{\sqrt{y}},$$

és $\frac{1}{\sqrt{y}} \leq L = \text{const}$ kellene ahhoz, hogy ez az f Lipschitz-folytonos lehessen – ami viszont csak akkor igaz, ha $y \geq y_0 > 0$. De esetünkben y kezdetiértéke éppen 0.

A példa mutatja, hogy ha nincs Lipschitz-folytonosság, akkor lehet, hogy nincs unicitás.

9.3. Az Euler-módszer

Most fordulunk a differenciálegyenletek numerikus megoldásához. Ekkor azt fogjuk feltenni, hogy f második argumentumának Lipschitz-folytonos függvénye, mert egyébként az a probléma támadhat, hogy nincs unicitás. Ezenfelül



9.1. ábra. Az Euler-módszer első lépése

további differenciálhatósági feltételekre lesz szükségünk, hasonlóan mint a numerikus integrálásnál.

A (9.5) alakú kezdetiérték feladat numerikus megoldását szolgáltató leg-egyszerűbb eljárás az Euler-módszer, amelynek bevezetéséhez felteszük, hogy $n = 1$.

Az alapvető ötlet az, hogy a differenciálegyenlet kezdetiérték feladatából kiszámítható a keresett $y(x)$ függvény deriváltja az $x_0 := 0$ pontban: $y'(0) = f(0, y(0))$. Ha ebben az irányban teszünk egy kis h lépést előre (ld. a 9.1. ábrát):

$$x_1 = x_0 + h = h, \quad y_1 = y_0 + hf(0, y_0), \quad y_0 := y(0),$$

akkor általában csak kis, mégpedig h -ban másodrendű hibát követünk el (erre lejjebb visszatérünk).

Ezután hasonló helyzetben vagyunk, mint az elején: rendelkezünk egy ponttal az x, y -síkon, amelyen keresztül (most már csak közelítőleg) vonul a keresett függvény, és a differenciálegyenlet segítségével kiszámíthatjuk a derivált értékét az adott pontban. Ennek segítségével újra közelíthetjük a megoldást egy h hosszú szakaszon:

$$x_2 = x_1 + h = 2h, \quad y_2 = y_1 + hf(x_1, y_1),$$

újra egy másodrendű hibát elkövetve. Általában az

$$x_{i+1} = x_i + h, \quad y_{i+1} = y_i + hf(x_i, y_i) \quad (9.6)$$

képletek szerint dolgozunk.

A másodrendű hibák a legrosszabb esetben összeadódnak, és ha már egy $1/h$ -val összemérhető lépésszámot megtettünk, akkor az összhiba már csak

elsőrendű lesz h -ban (erre lejjebb visszatérünk) – ami kicsi h esetén még elfogadható hiba lehet.

9.3.1. Az Euler-módszer algoritmus, tesztfeladatok

Pszudokód formájában megfogalmazva az egész eljárást, a következő lépések adódnak, mint az *explicit Euler-módszer algoritmus*, ha az egész $[0, 1]$ intervallumot N darab h hosszú intervallumra osztjuk fel és minden $x_i = ih$ osztópontokra szerzünk egy y_i értéket, amely reményünk szerint közelíti a pontos megoldás megfelelő $y(x_i)$ értékét:

Adott $x_0 = 0$, $y_0 = y(0)$ és a differenciálegyenlet $f(x, y)$ jobboldalát kiszámító eljárás, továbbá $N \geq 1$, az intervallumok száma.

1. $x_0 := 0$, $h := 1/N$
2. $i = 0, 1, \dots, N - 1$
3. $[x_{i+1} = x_i + h$
4. $y_{i+1} = y_i + hf(x_i, y_i)]_i$
5. [stop: eredmény $(x_0, y_0), (x_1, y_1), \dots, (x_N, y_N)]$

Mint **tesztelési példát** elsőnek olyan esetet említünk, amikor az Euler-módszer hiba nélkül adja a pontos megoldás értékeit az x_i pontokban, mert f konstans:

$$y' = 1, \quad x \in [0, 1], \quad y(0) = 0. \quad (9.7)$$

Ezen differenciálegyenlet általános megoldása $y = x + a$, ahol $a = \text{const}$. De a kezdetiérték miatt ez a konstans 0 kell, hogy legyen. Vagyis $y = x$ a kezdetiérték feladat megoldása, és ezt vissza kell kapnunk a programból – kerekítési hibákat nem számítva.

Második **tesztelési példának** tekintsük az

$$y' = Ly, \quad x \in [0, 1], \quad y(0) = 1 \quad (9.8)$$

kezdetiérték feladatot, amelyben $L = \text{const} > 0$. Ennek megoldása $y(x) = e^{Lx}$, úgyhogy nem nehéz a pontos megoldással összehasonlítani a numerikus megoldást. Legyen pl. $L = 10$, akkor 6 számjegyre

a) $h=0.1$:

x_i	$y(x_i)$	y_i
0.1	2.71828	2
0.2	7.38906	4
0.3	20.0855	8
\vdots	\vdots	\vdots
1.0	22026.5	1024

b) $h=0.05$:

x_i	$y(x_i)$	y_i
0.05	1.64872	1.50000
0.1	2.71828	2.25000
0.15	4.48169	3.37500
\vdots	\vdots	\vdots
1.0	22026.5	3325.26

Az eredmények tehát elég távol vannak a pontostól, viszont a program tesztelésére alkalmasak. De folytassuk a számítást kisebb $h = 1/N$ -értékekkel, mindig duplázva az intervallumok N számát! A következő táblázat szembeállítja az $x_N = 1 = Nh$ -ra kapott numerikus értékeket és $e_N := |y(x_N) - y_N|$ hibájukat. A pontos megoldási érték kerekítve $y(x_N) = 22026.5$ minden esetben. A táblázat utolsó oszlopában mutatjuk a jelenlegi és az előző hiba hányadosát.

N	h	y_N	e_N	$e_N/e_{N/2}$
10	0.1	1024	21002.5	-
20	0.05	3325.26	18701.24	0.8904
40	0.025	7523.16	14503.34	0.7755
80	0.0125	12365.2	9661.3	0.6661
160	0.00625	16316.6	5709.9	0.5910
320	0.003125	18900.3	3126.0	0.5475
640	0.0015625	20387.5	1639.0	0.5243

Ez a táblát azt mutatja, hogy bár a konvergencia lassú, de valóban elsőrendű: a jelenlegi és az előző hiba hányadosa közeledik 0.5 felé, vagyis: éppen úgy viselkedik, mint a h lépéstávolság, amelyet minden számításban feleztünk. A mindenkori h -val kifejezve az utolsó két hiba például

$$e_{320} \approx 1000384 \cdot h, \quad e_{640} \approx 1048960 \cdot h. \quad (9.9)$$

Az Euler-módszert rendszerekre is alkalmazhatjuk: ekkor $y \in \mathbb{R}^n$, $f \in \mathbb{R}^n$ és (9.6)-ban mind az y_i -k és $f(x_i, y_i)$ -k vektorok, tehát az $y_{i+1} = y_i + hf(x_i, y_i)$ képlet egy ciklust takar, amelyet mondjuk j indexszel fut 1-től n -ig.

Ekkor, amennyiben csak az utolsó x -értékre érdekel bennünket az eredmény, akkor az új x - és y -értékekkel felülírhatjuk a régiéket, és rendszerek esetén ez nagy különbség a tágirány szempontjából.

9.4. Az Euler-módszer hibaelemzése

Az Euler-módszert úgy is felfoghatjuk, hogy a (9.3) differenciálegyenletben az y deriváltját az ún. *haladó differenciahányadossal* helyettesítettük:

$$\frac{y(x_i + h) - y(x_i)}{h} \approx y'(x_i) = f(x_i, y(x_i)).$$

Az eltérés ezen reláció bal és jobb oldala között az *Euler-módszer képlethibája*:

$$g(x_i, h) = g_i := \frac{y(x_{i+1}) - y(x_i)}{h} - f(x_i, y(x_i)).$$

A (9.7) differenciálegyenlethez tartozó képlethiba pl. $g_i \equiv 0$, mert ekkor $y(x) = x$, $f(x, y) = 1$, és a haladó differenciahányados hiba nélkül adja a deriváltat egyszerűen azért, mert az konstans:

$$g_i = \frac{x_{i+1} - x_i}{h} - 1 = \frac{(i+1)h - ih}{h} - 1 = 0.$$

A (9.8) differenciálegyenlethez tartozó képlethiba viszont nem nulla. Ekkor $y(x) = e^{Lx}$, $f(x, y) = Ly = Le^{Lx}$ volt, és a haladó differenciahányados csak hibával adja a deriváltat:

$$\begin{aligned} g_i &= \frac{e^{Lx_{i+1}} - e^{Lx_i}}{h} - Le^{Lx_i} = e^{Lx_i} \left(\frac{e^{Lh} - 1}{h} - L \right) \\ &= e^{Lx_i} \left(\frac{1 + Lh + \frac{1}{2}(Lh)^2 + O(h^3) - 1}{h} - L \right) \\ &= e^{Lx_i} \left(L + \frac{1}{2}L^2h + O(h^2) - L \right) = O(h). \end{aligned}$$

A képlethibának az a jelentősége, hogy segítségével becsülhető a numerikus megoldás $e_i := y(x_i) - y_i$ hibája:

$$|e_{i+1}| \leq e^{Lx_{i+1}} \left[|e_0| + \sum_{k=0}^i |g_k| h \right]. \quad (9.10)$$

Ezen úgynevezett *stabilitási becslés* levezetéséhez elég mindössze azt feltenni, hogy f a második argumentumában Lipschitz-folytonos, L Lipschitz-állandóval.

Innen kaphatjuk a módszer konvergenciáját, ha többet teszünk fel f -ről: legyen $f(x, y)$ mindkét argumentuma szerint folytonosan deriválható. Ez egyben biztosítja az analitikus megoldás létezését és unicitását is, mert ekkor f Lipschitz-folytonos y -ban, továbbá az y megoldás ekkor kétszer folytonosan differenciálható:

$$y' = f(x, y(x)), \quad y'' = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx}, \quad (9.11)$$

a láncszabály alapján. Innen látjuk, hogy valóban y'' folytonos. Ezt használjuk ki $y(x_{i+1})$ Taylor-sorfejtésében:

$$y(x_i + h) = y(x_i) + hy'(x_i) + \frac{h^2}{2}y''(x_i + \vartheta_i h),$$

ahol $0 < \vartheta_i < 1$. Ezután innen kapjuk a $g(x_i, h)$ képlethibának az x_i körüli Taylor-sorfejtését:

$$\begin{aligned} hg_i &= y(x_i + h) - y(x_i) - hf(x_i, y(x_i)) \\ &= [y(x_i) + hy'(x_i) + \frac{h^2}{2}y''(x_i + \vartheta_i h)] - y(x_i) - hy'(x_i), \end{aligned}$$

ahol figyelembe vettük, hogy a (9.3) differenciálegyenlet szerint $f(x, y(x)) = y'(x)$. Tehát

$$|g_i| = \frac{h}{2}|y''(x_i + \vartheta_i h)| \leq \frac{h}{2}M_2, \quad M_2 := \max_{[0,1]}|y''(x)|.$$

Amit itt le lehet olvasni: hogy az Euler-módszer képlethibája h -val együtt nullához tart, ezt a tulajdonságot úgy említjük, hogy az Euler-módszer *konzisztens*.

Most a g_i becslését (9.10)-be helyettesítjük be:

$$|e_{i+1}| \leq e^{Lx_{i+1}} \left[|e_0| + \sum_{k=0}^i |g_k| h \right] \leq e^L \left[|e_0| + \frac{h}{2} M_2 \sum_{k=0}^i h \right] \leq e^L \left[|e_0| + \frac{h}{2} M_2 \right], \quad (9.12)$$

mert $\sum_{k=0}^i h = (i+1)h = x_{i+1} \leq 1$.

Figyelmeztetés: $e_{i+1} = y(x_{i+1}) - y_{i+1}$ a hiba, $e^{Lx_{i+1}} \leq e^L$ az exponenciális függvény értéke.

(9.12)-ből azt olvassuk le, hogy $e_0 = y(0) - y_0 \neq 0$ esetén a konvergenciát nem tudjuk garantálni: ha y_0 eltér $y(0)$ -tól, akkor a későbbiekben ez az eltérés még nőhet is. Ez a megjegyzés akkor fontos, ha a pontos $y(0)$ érték nem áll rendelkezésünkre, mert az egy másik feladat eredménye és, mint ilyen, hibával terhelt.

Ha viszont $e_0 = 0$, akkor (9.12)-ből $|e_{i+1}| \leq e^L \frac{h}{2} M_2$ marad, és innen látjuk, hogy $e_{i+1} \rightarrow 0$, amikor $h \rightarrow 0$. Ekkor tehát a konvergencia következik.

Alkalmazzuk a (9.12) hibabecslésünket a (9.8) differenciálegyenlet numerikus megoldására!

Ekkor $L = 10$, $M_2 = y''(1) = L^2 e^L \approx 2202646.6$, $e^L \frac{M_2}{2} \approx 2.4 \cdot 10^{10}$.

Összehasonlítva a (9.9) számítási eredményünkkel a hibabecslés körülbelül $e^L \approx 22026.6$ -szerese az ott kapott értéknek, de a hiba h -ban elsőrendűségét bizonyítja.

Felhívjuk a figyelmet a következőre: a konvergencia nem függ y_0 konkrét értékétől, csak legyen $y_0 = y(0)$, ezenkívül a konvergencia nem függ a konkrét f -től: csak legyen az folytonosan differenciálható.

Továbbá, a numerikus eredmény az $\{y_i\}_{i=0}^N$ sorozat, és annak konvergenciáját várjuk az $y(x)$ függvényhez, pontosabban annak $\{y(x_i)\}_{i=0}^N$ értékeihez. A konvergencia ellenőrzéséhez a $[0, 1]$ intervallum valamelyik felosztásában az $x_i = ih = x_*$ pontot rögzítjük, ahol $i = i(h, x_*)$. A későbbiekben $i(h, x_*) \rightarrow \infty$, ha $h \rightarrow 0$. A gyakorlati számításnál a legcélszerűbb a felosztás intervallumszámát duplázni, a h lépésközű felosztás után $h/2$ -vel számítva stb.

Összefoglalás: az Euler-módszer tulajdonságai.

a) Ha az f függvény y -ban Lipschitz-folytonos L Lipschitz-állandóval, akkor az Euler-módszer *stabil*: érvényes

$$|e_i| \leq e^{Lx_i} \left(|e_0| + \sum_{k=0}^{i-1} |g_k| h \right).$$

b) Ha $y \in C^2$ (pl. amikor $f \in C^1\{[0, 1] \times \mathbb{R}\}$), akkor az Euler-módszer *konzisztens*: a g_k lokális hibáira igaz

$$|g_k| \leq \frac{h}{2} M_2 = O(h), \quad M_2 := \max_{[0,1]} |y''(x)|.$$

c) „konzisztencia + stabilitás = konvergencia”: Ha $e_0 = 0$ és $y \in C^2$, akkor a $[0, 1]$ intervallumon az Euler-módszer *konvergens*:

$$|e_i| = |y(x_i) - y_i| \leq e^{Lx_i} \frac{h}{2} M_2 = O(h).$$

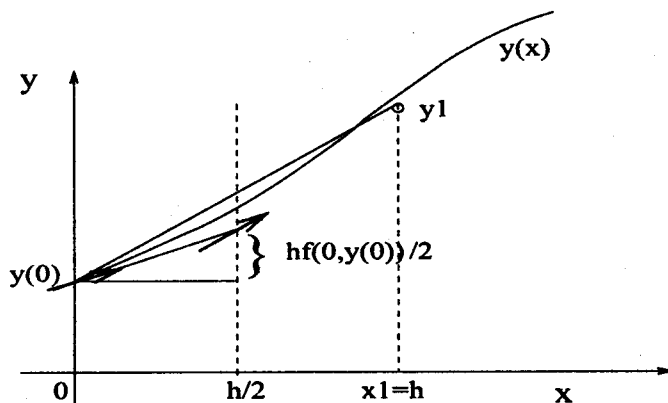
A módszer konvergenciarendje eszerint 1, amit „lassú de biztos”-nak lehet jellemezni.

9.5. A javított Euler-módszer

Lassú konvergenciája miatt (ld. fent a számítási eredményeket a második tesztfeladathoz) az Euler-módszert ritkán alkalmazzák, pl. akkor, amikor a feladat nemlineáris és olyan nagyméretű, hogy jobb (magasabbrendű) módszerek bevetése lehetetlen; a módszert viszont a következő módon javíthatjuk.

Először csak egy fél Euler-lépést teszünk meg, majd a kapott pontban újra számítjuk az f értékét, a megoldás deriváltját – és ezzel a meredekséggel tesszük meg az egész lépést (x_i, y_i) -ből. A fél lépés elvisz az $[x_i, x_{i+1}]$ intervallum közepére, és a szimmetria alapján remélhetjük, hogy az ott kiszámított meredekség másodrendű pontosságot hoz.

Képletekkel (ld. a 9.2. ábrát):



9.2. ábra. A javított Euler-módszer első lépése

a) a fél Euler-lépés:

$$x_{i+1/2} = x_i + \frac{h}{2}, \quad y_{i+1/2} = y_i + \frac{h}{2} f(x_i, y_i),$$

b) az egész lépés új meredekséggel:

$$x_{i+1} = x_i + h, \quad y_{i+1} = y_i + h f(x_{i+1/2}, y_{i+1/2}).$$

Ezen módszer lokális hibáját, a 9.4. pontban megtárgyalt Euler-módszerhez hasonlóan, a következőképpen definiáljuk:

$$g(x_i, h) = g_i := \frac{y(x_{i+1}) - y(x_i)}{h} - k_2, \quad (9.13)$$

$$k_1 := f(x_i, y(x_i)), \quad k_2 := f\left(x_i + \frac{h}{2}, y(x_i) + \frac{h}{2} k_1\right).$$

A javított Euler-módszer képlethibájának vizsgálatához feltesszük, hogy f legyen kétszer folytonosan differenciálható mindkét változója szerint, és így y háromszor folytonosan differenciálható x szerint.

A továbbiakban kellene tudni a kétváltozós f függvény Taylor-sorfejtését. Ezt viszont vissza lehet vezetni az egyváltozós esetre. Ami minket érdekelni fog, az $f(x + \delta x, y + \delta y)$ sorfejtése az (x, y) helyen. Vezessük be a következő $F(t)$ segédfüggvényt:

$$F(t) := f(x + t\delta x, y + t\delta y),$$

ahol az egyszerűség kedvéért x_i helyett x -et írtuk, $y(x_i)$ helyett y -t, és

$$\frac{h}{2} =: \delta x, \quad \frac{h}{2} k_1 = \frac{h}{2} f(x_i, y(x_i)) =: \delta y.$$

Ekkor $k_2 = F(1) = f(x + \delta x, y + \delta y)$ sorfejtése a $t = 0$ helyen hasznos lesz. Ez a következő alakot nyer, ha a második deriváltakkal írjuk fel a hibát:

$$\begin{aligned} F(t) &= F(0) + tF'(0) + \frac{t^2}{2}F''(\vartheta t) \quad (\text{ahol } 0 < \vartheta < 1) \\ &= f(x, y) + t[f_x\delta x + f_y\delta y](x, y) \\ &\quad + \frac{t^2}{2}[f_{xx}\delta x^2 + 2f_{xy}\delta x\delta y + f_{yy}\delta y^2](x + \vartheta t\delta x, y + \vartheta t\delta y). \end{aligned}$$

Itt a parciális deriváltakat alsó indexszel jelöltük, pl. $f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$, valamint felhasználtuk, hogy

$$\frac{d(x + t\delta x)}{dt} = \delta x, \quad \frac{d(y + t\delta y)}{dt} = \delta y$$

és így a láncszabály szerint

$$\begin{aligned} F'(t) &= \left[f_x \frac{d(x + t\delta x)}{dt} + f_y \frac{d(y + t\delta y)}{dt} \right] (x + t\delta x, y + t\delta y) \\ &= [f_x\delta x + f_y\delta y](x + t\delta x, y + t\delta y), \\ F''(t) &= \frac{dF'}{dt} = [f_{xx}\delta x^2 + 2f_{xy}\delta x\delta y + f_{yy}\delta y^2](x + t\delta x, y + t\delta y). \end{aligned}$$

Az előző relációkba helyettesítve a

$$t = 1, \quad x = x_i, \quad y = y(x_i), \quad \delta x = \frac{h}{2}, \quad \delta y = \frac{h}{2}k_1 = \frac{h}{2}f(x_i, y(x_i))$$

kifejezéseket, kapjuk

$$\begin{aligned} k_2 &= F(1) = F(0) + F'(0) + \frac{1}{2}F''(\vartheta) \\ &= f(x_i, y(x_i)) + \frac{h}{2}[f_x + f_y f](x_i, y(x_i)) + O(h^2), \end{aligned}$$

mert

$$F''(\vartheta) = \frac{h^2}{4}[f_{xx} + 2f_{xy}f + f_{yy}f^2](x + \vartheta\delta x, y + \vartheta\delta y) = O(h^2)$$

amikor feltevésünk szerint f kétszer folytonosan differenciálható. Figyelembe véve (9.11)-et, végül is azt kapjuk, hogy

$$k_2 = f(x_i, y(x_i)) + \frac{h}{2}[f_x + f_y f](x_i, y(x_i)) + O(h^2) = y'(x_i) + \frac{h}{2}y''(x_i) + O(h^2).$$

Most a (9.13)-ban definiált g_i sorfejtését az $(x_i, y(x_i))$ helyen össze tudjuk állítani:

$$g_i = \frac{1}{h} \left(y(x_i) + hy'(x_i) + \frac{h^2}{2}y''(x_i) + O(h^3) - y(x_i) \right) - \left(y'(x_i) + \frac{h}{2}y''(x_i) + O(h^2) \right) = O(h^2).$$

Ez azt jelenti, hogy a javított Euler-módszer konzisztens.

Mivel a módszerre hasonló stabilitási becslés érvényes, mint az Euler-módszerre, a másodrendű konvergencia következik.

Ugyanazzal a (9.8) feladatra vonatkozó számítási példával, mint amelyet az Euler-módszernél használtuk, a javított Euler-módszer pontosságát is szemléltetjük:

a) $h=0.1$:

x_i	$y(x_i)$	y_i
0.1	2.71828	2.50000
0.2	7.38906	6.25000
0.3	20.0855	15.6250
\vdots	\vdots	\vdots
1.0	22026.5	9536.74

b) $h=0.05$:

x_i	$y(x_i)$	y_i
0.05	1.64872	1.62500
0.1	2.71828	2.64063
0.15	4.48169	4.29102
\vdots	\vdots	\vdots
1.0	22026.5	16484.2

Ezenkívül olyan táblázatot is mutatunk (mint az Euler-módszer esetén), amely az $e_N := |y(x_N) - y_N|$ hiba viselkedését mutatja N függvényében:

N	h	y_N	e_N	$e_N/e_{N/2}$
10	0.1	9536.74	12489.8	-
20	0.05	16484.2	5542.29	0.4437
40	0.025	20200.2	1826.29	0.3295
80	0.0125	21510.1	516.357	0.2827
160	0.00625	21890.0	136.423	0.2642
320	0.003125	21991.5	34.9928	0.2565
640	0.0015625	22017.6	8.85644	0.2531

A táblázat mutatja a másodrendű konvergenciát: a mindenkori és az előző hiba hányadosa közeledik 0.25 felé, tehát úgy viselkedik, mint a h^2 . Amilyen hibát ér el a javított Euler-módszer $N = 40$ -nel, azt az Euler-módszer (közelítőleg) $N = 640$ -nel éri el. Közben ehhez a javított Euler-módszernek 80 függvénykiértékelésre van szüksége, az Euler-módszernek 640-re, vagyis elkönnyelhetjük: ezen a példán a javított Euler-módszer kb. 8-szor hatékonyabb.

A javított Euler-módszer példájából azt látjuk, hogy egynél magasabbrendű módszer is konstruálható anélkül, hogy az f deriváltjainak képletére szükség lenne.

Ezt az utat lehet általánosítani. Így kapjuk az egyik legismertebb eljárást közönséges differenciálegyenletek numerikus megoldására, a *negyedrendű Runge-Kutta-módszert*:

Először definiáljuk az alábbi k -számokat (rendszer esetén vektorokat):

$$\begin{aligned} k_1 &:= f(x_i, y_i), \\ k_2 &:= f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_1\right), \\ k_3 &:= f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_2\right), \\ k_4 &:= f(x_i + h, y_i + hk_3). \end{aligned}$$

Ezeket a k_j számokat lineárisan kombinálva az új y -értéket így számítjuk ki:

$$y_{i+1} = y_i + \frac{h}{6} (k_1 + 2k_2 + 2k_3 + k_4).$$

Nem véletlen, hogy az $x_{i+1/2}$ ponthoz tartozó f -értékek összesen négyszeres súllyal szerepelnek, míg az $[x_i, x_{i+1}]$ intervallum végpontjaihoz tartozó f -értékek csak egyszeres súllyal: ez a Simpson-féle kvadratúráképletnek felel meg.

A negyedrendű Runge-Kutta-módszer esetén is mutatjuk a (9.8) feladat számításánál kapott eredményeket:

a) $h=0.1$:

x_i	$y(x_i)$	y_i
0.1	2.71828	<u>2.70833</u>
0.2	7.38906	<u>7.33507</u>
0.3	20.0855	19.8658
\vdots	\vdots	\vdots
1.0	22026.5	<u>21233.5</u>

b) $h=0.05$:

x_i	$y(x_i)$	y_i
0.05	1.64872	<u>1.64844</u>
0.1	2.71828	<u>2.71735</u>
0.15	4.48169	<u>4.47938</u>
\vdots	\vdots	\vdots
1.0	22026.5	<u>21950.8</u>

Itt már érdemes a numerikus eredmények helyes számjegyeit aláhúzni.

9.6. Az implicit Euler-módszer

Tekintsük most az $y' = f(x, y)$ egyenlet közelítő megoldását az úgynevezett *implicit Euler-módszerrel*, vagyis az

$$y_{i+1} = y_i + hf(x_{i+1}, y_{i+1}) \quad (9.14)$$

képlet segítségével, amelynek jobb oldalán a keresett y_{i+1} érték is előfordul. Úgy, mint az eredeti Euler módszer (amit a megkülönböztetés miatt *explicit*nek is hívunk), az implicit Euler-módszer is konzisztens és konvergencia rendje 1.

Az explicit Euler-módszerrel szemben az a hátránya, hogy használata egy (általában: nemlineáris) egyenlet megoldásával jár. Ez a hátrány csökken, amikor lineáris differenciálegyenletek rendszeréről van szó.

9.6.1. Az implicit Euler-módszer lineáris rendszerekre

Legyen $y \in \mathbb{R}^n$ és $f: [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, mégpedig $f(x, y) = Ay + g(x)$, $A \in \mathbb{R}^{n \times n}$:

$$y' = Ay + g(x).$$

Ekkor az implicit Euler-módszer igen fontos abban az esetben, amikor az A mátrix rosszul kondicionált és sajátértékei negatívak.

A módszer most a következő alakot veszi fel:

$$y_{i+1} = y_i + h [Ay_{i+1} + g(x_{i+1})],$$

és a még ismeretlen y_{i+1} vektort átrendezve kapjuk

$$[I - hA]y_{i+1} = y_i + hg(x_{i+1}), \quad (9.15)$$

ahol I az $n \times n$ -es egységmátrix.

Az itt balra álló mátrix LU-felbontását számítjuk ki, majd megoldjuk a rendszert. Mivel A nem függ x -től, ezt a felbontást elegendő egyszer az elején elvégezni.

A módszer előnyét az explicit Euler-módszerrel szemben abban a legegyszerűbb esetben fogjuk elmagyarázni, amikor $n = 1$, tehát csak egy differenciálegyenlet van és g sem függ x -től, vagyis A és g számok.

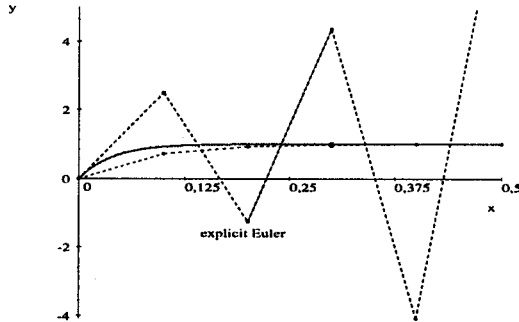
Legyen $g > 0$, és A -ról feltesszük, hogy negatív: $A = -q$, $q > 0$. A differenciálegyenletünk tehát ($n = 1$ miatt $I = 1$):

$$y' = g - qy, \text{ megoldása } y(0) = 0 \text{ esetén: } y(x) = \frac{g}{q} (1 - e^{-qx}). \quad (9.16)$$

Ez a függvény monoton módon növekszik és $x \rightarrow \infty$ -re $\frac{g}{q}$ -hoz tart. Esetünkben (9.15) alakja:

$$[1 + hq]y_{i+1} = y_i + hg, \text{ vagyis } y_{i+1} = \frac{y_i + hg}{1 + hq}.$$

Ebből azt látjuk, hogy y_{i+1} kiszámításával semmi probléma nincs, és $y_i \geq 0$ -val és $g > 0$ -val együtt y_{i+1} is pozitív.



9.3. ábra. Pontos megoldás, explicit és implicit Euler-módszer eredménye: $y' = g - qy$, $g = q = 25$, $h = 0.1$

Most ugyanazt a differenciálegyenletet az explicit Euler-módszerrel próbáljuk megoldani:

$$y_{i+1} = y_i + h(g - qy_i) = y_i(1 - qh) + hg.$$

Korábbi eredményeink szerint az elsőrendű konvergencia itt garantált, de a képlet mutatja, hogy $1 - qh < -1$ esetén növekvő oszcillációk keletkeznek. Ez nem cáfolja a konvergenciát, mert $h \rightarrow 0$ -kor bizonyára $1 - qh \rightarrow +1$ lesz, de a számítógépen véges h -val dolgozunk, és ekkor könnyen megtörténhet, hogy $hq > 2$, különösen, ha q nagy, ld. a 9.3 ábrát.

A numerikus megoldás itt két okból is elfogadhatatlan: nem marad korlátos $i \rightarrow \infty$ esetén, ezenkívül oszcillációkat mutat – holott a pontos megoldás monoton módon közeledik egy konstanshoz, amikor $x \rightarrow \infty$.

Amikor, mint a fenti példában, csak egy egyenletről van szó, egy transzformációval a nagy q -t el tudjuk tüntetni: legyen $t = qx$ az új független változó. Ekkor

$$x = \frac{t}{q}, \quad y(x(t)) =: z(t), \quad \frac{dz}{dt} = \frac{dy}{dx} \frac{dx}{dt} = y'(x) \frac{1}{q} = \frac{1}{q}(g - qy) = \frac{g}{q} - z.$$

Az ilyen transzformációnak viszont nincs a kívánt hatása rendszerek esetén. Nézzünk erre egy példát! A kezdetiérték feladat legyen

$$y' = \begin{pmatrix} -501 & 500 \\ 500 & -501 \end{pmatrix} y + \begin{pmatrix} 6 \\ -7 \end{pmatrix}, \quad y(0) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Az itteni mátrix megfelel annak, amit az elején mint az explicit Euler-módszer számára problémás esetet említettünk: sajátértékei negatívak (konkrétan

sajátértékei -1 és -1001), és a mátrix mérsékelten rosszul kondicionált, ugyanis $\text{cond } A = 1001$. A fenti transzformációval az A mátrix elé lehet egy szorzót vinni – de ettől nem változik a kondíciószáma.

Az Euler-módszer esetén akkor nem lesz gond, ha h elég kicsi:

- a numerikus megoldás korlátos marad, ha $h \leq \frac{2}{|\lambda_{\max}|} = \frac{2}{1001}$, de ekkor még lehetnek oszcillációk;

- nincsenek oszcillációk a numerikus megoldásban, ha $h \leq \frac{1}{|\lambda_{\max}|} = \frac{1}{1001}$,

Ez már egy elég kellemetlen korlátozás, ami könnyen fokozható (ha -501 és 500 helyett a mátrixban -5001 és 5000 áll, akkor az utóbbi korlát $h \leq \frac{1}{10001}$ -re változik, stb). Mivel a pontos megoldás

$$y(x) = \frac{38}{77} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-1001x} + 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-x} - \frac{1}{77} \begin{pmatrix} 38 \\ 39 \end{pmatrix},$$

abban lehetne reménykedni, hogy miután az a rész a megoldásban, amely kapcsolatos a -1001 sajátértékkel, jelentéktelenné vált, nagyobb lépéssel lehetne haladni az explicit Euler-módszerrel is. De ez tévedés, mert a mátrix nem változott és a növekvő oszcillációk most is beindulnak.

Az a probléma, hogy a numerikus megoldás nem marad korlátos $i \rightarrow \infty$ esetén, az nem tűnik el más explicit módszer használatakor sem. Vegyük pl. a javított Euler-módszert. Ha ezt a (9.16)-re alkalmazzuk, akkor a képlet

$$y_{i+1} = y_i + h \left[g - q(y_i + \frac{h}{2}(g - qy_i)) \right] = y_i \left(1 - qh + \frac{q^2 h^2}{2} \right) + g \left(1 - \frac{qh}{2} \right).$$

Itt y_i mellett áll $z := qh$ egy polinomja: $p(z) = 1 - z + \frac{z^2}{2} = \frac{1}{2}(1 + (1 - z)^2)$, amelynek értékei nagyobbak egynél, ha $z = qh > 2$. Így y_i gyorsan fog növekedni i -vel, kivéve, ha $h \leq \frac{2}{q}$. Ez ugyanaz a korlátozás, mint az explicit Euler-módszer esetén.

Az ilyen egyenleteket, amelyek Jacobi-mátrixa rosszul kondicionált vagy szinguláris, sajátértékei nem pozitívak (ill. valós részei nem pozitívak), amelyekben az implicit Euler-módszer lényegesen jobban működik, mint az explicit módszerek (akár magasabbrendűek is), *merev differenciálegyenleteknek* hívjuk.

A merev rendszerek megoldásakor az explicit Euler-módszer oszcillációmentességét biztosító h -korlát erősebben hátráltatja a számítást, mint az, hogy az implicit Euler-módszer minden lépése algebrai egyenletek megoldásával jár.

9.6.2. Nemlineáris rendszerek

Ha $f : [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ és nem feltétlenül lineáris y -ban, akkor a (9.14) implicit Euler-módszer minden lépésében szembenállunk az

$$F_i(y) := y - hf(x_{i+1}, y) - y_i = 0 \quad (9.17)$$

egyenlettel, amelynek megoldása $y =: y_{i+1}$ keresett. Erre alkalmazzuk a Newton-módszert. Kézenfekvő nulladik közelítésként y_{i+1} -hez az előző y_i vektort használni: $y_{i+1,0} := y_i$.

A Newton-módszer képlete ekkor (ld. a 7.3.1. pontot)

$$0 = F_i(y_i) + F_i'(y_i)\delta y,$$

avagy

$$F_i'(y_i)\delta y = -F_i(y_i) = -(y_i - hf(x_{i+1}, y_i) - y_i) = hf(x_{i+1}, y_i), \quad (9.18)$$

ahol $\delta y = y_{i+1,1} - y_{i+1,0} = y_{i+1,1} - y_i$ a nulladik közelítés javítása. Az F_i' alakja itt

$$F_i'(y_i) = I - hJ(x_{i+1}, y_i), \quad (9.19)$$

az $f(x, y)$ vektorfüggvény $J(x, y)$ Jacobi-mátrixszával. A (9.19) képlet megértéséhez vizsgáljuk az $n = 2$ esetet, amikor (9.17)-ben $y_i = \begin{pmatrix} y_i^{(1)} \\ y_i^{(2)} \end{pmatrix}$ és $y = \begin{pmatrix} y^{(1)} \\ y^{(2)} \end{pmatrix}$, tehát

$$F_i(y) = \begin{pmatrix} y^{(1)} \\ y^{(2)} \end{pmatrix} - h \begin{pmatrix} f_1(x_{i+1}, y) \\ f_2(x_{i+1}, y) \end{pmatrix} - \begin{pmatrix} y_i^{(1)} \\ y_i^{(2)} \end{pmatrix}.$$

Ezen összefüggés első sorának $y^{(1)}$ szerinti parciális deriváltja két részből tevődik össze, mert a harmadik jobboldali tag konstans, deriváltja 0. Az első tag deriváltja viszont 1, a másodiké $-h \frac{\partial f_1}{\partial y^{(1)}}$. Továbbá, az első sor $y^{(2)}$ szerinti deriváltja $-h \frac{\partial f_1}{\partial y^{(2)}}$, mert ennek kiszámításakor az első tag $y^{(1)}$ változóját kell konstansnak tekinteni.

Hasonlóan a második sor $y^{(1)}$ szerinti parciális deriváltja $-h \frac{\partial f_2}{\partial y^{(1)}}$, de az első tag $y^{(2)}$ szerinti deriváltja 1, a másodiké $-h \frac{\partial f_2}{\partial y^{(2)}}$. Összegezve valóban

$$F_i' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - h \begin{pmatrix} \frac{\partial f_1}{\partial y^{(1)}} & \frac{\partial f_1}{\partial y^{(2)}} \\ \frac{\partial f_2}{\partial y^{(1)}} & \frac{\partial f_2}{\partial y^{(2)}} \end{pmatrix} = I - hJ.$$

Most a (9.19) képletet vesszük figyelembe (9.18)-ban, továbbá az ottani $y_{i+1,1}$ első közelítést (9.17) megoldásához máris y_{i+1} -nek, tehát végsőnek fogadjuk

el. Ugyanis utána lehet számolni, hogy $y_{i+1,1}$ kiszámításával a differenciálegyenletet h szerint elsőrendben megoldó eljárást kaptunk és további Newton-iterációk ezt a rendet már nem javítják.

Ezenkívül, a merev differenciálegyenletek szempontjából a lényeg már megtörtént: megjelent az approximációban a keresett új érték szorzójaként az f Jacobi-mátrixa.

Ezzel a következő lineáris egyenletrendszert javasoljuk számítási képletnek:

$$[I - hJ(x_{i+1}, y_i)]\delta y = hf(x_{i+1}, y_i), \quad y_{i+1} = y_i + \delta y. \quad (9.20)$$

Abban az esetben, amikor $f(x, y) = Ay + g(x)$, a Jacobi-mátrix éppen $J = A$, és ekkor (9.20) azonos (9.15)-tel. Itt, a nemlineáris rendszereknél, az $I - hJ(x_{i+1}, y_i)$ mátrix változik i -vel, tehát minden lépésben LU-felbontást kell készítenünk.

9.6.3. Az implicit Euler-módszer algoritmus, tesztfeladatok

Az *implicit Euler-módszer algoritmusánál* is a $[0, 1]$ intervallumot N darab h hosszú intervallumra osztjuk fel, de $y(x)$ most n -dimenziós vektorfüggvény és minden $x_i = ih$ osztópontához szerzünk egy y_i vektort, amely közelíti a pontos megoldás megfelelő $y(x_i)$ vektorát.

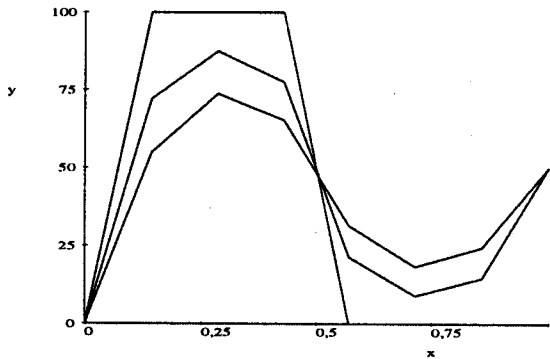
Adott az $y_0 = y(0)$ kezdetvektor, a differenciálegyenlet $f(x, y)$ jobboldalát és $J(x, y)$ Jacobi-mátrixát kiszámító eljárás, továbbá $N \geq 1$, az intervallumok száma.

1. $x_0 := 0, \quad h := 1/N$
2. $i := 0, 1, \dots, N - 1$
3. $[x_{i+1} := x_i + h, \quad J_i := J(x_{i+1}, y_i)$
4. $I - h * J_i \implies [L, U, \text{sing}]$
5. $? \text{ sing} ? \text{ [stop: „szinguláris mátrix”, } x = x_{i+1}, y = y_i]$
6. $LU\delta y = h * f(x_{i+1}, y_i) \implies \delta y$
7. $y_{i+1} := y_i + \delta y]_i$
8. [stop: eredmény $(x_0, y_0), (x_1, y_1), \dots, (x_N, y_N)$]

Amennyiben szingularitást tapasztalunk, újból indítsuk be a számítást nagyobb N -nel, mert a 2.3.3. pontbeli perturbációs lemma garantálja, hogy elég kicsi h -ra $I - h * J$ reguláris lesz.

Mint első tesztelési példát az egyszerű (9.7) egyenletet rendszerré bővíthetjük ki:

$$\frac{d}{dx} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix},$$



9.4. ábra. A kiindulási és az első két időlépés utáni hőeloszlás: $N = 6, h = 0.01$

mert ennek pontos megoldását, $y_1 = 3 + x$, $y_2 = 4 + 2x$, hiba nélkül kell, hogy adja a program – kerekítési hibáktól eltekintve. Itt tehát

$$f(x, y) = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad J(x, y) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Második tesztfeladatunk egy hővezetési példa: egy 100 fokos és egy 0 fokos rudat összerakunk a $t = 0$ időpontban, majd a meleg rúd külső oldalán nulla hőmérsékletet állítunk be, a hideg rúd külső oldalán 50 fokos hőmérsékletet, aztán N belső pontban figyeljük az $y = (y_1(t), \dots, y_N(t))^T$ hőmérséklet kiegyenlítődését. Ennek diszkrét modellje a következő, lineáris differenciálegyenletekből álló rendszer:

$$\frac{dy}{dt} = -(N+1)^2 * \begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & 0 & 0 & -1 & 2 \end{pmatrix} y + \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 50(N+1)^2 \end{pmatrix},$$

a kezdetiérték $y(0) = (100, \dots, 100, 0, \dots, 0)^T$. A mátrix szimmetrikus, tehát minden sajátértéke valós, továbbá, a Gersgorin-körök segítségével ellenőrizhető, hogy a sajátértékek a $[0, 4(N+1)^2]$ intervallumban vannak – vagyis a rendszer merev lesz, ha N nagy. (Nagyobb N -re jobb pontossággal kapjuk a hőeloszlást.) A Jacobi-mátrix megegyezik az egyenletrendszerben látható tridiagonális mátrixszal, beleértve a $-(N+1)^2$ szorzóját.

Az alábbiakban mutatjuk $N = 6$ -ra az első és második időlépés (6 számjegyre kerekített) eredményeit akkor, amikor $h = 0.01$ az időlépés, ld. a (9.4

ábrát is, ahol az eredményeket a szélén felvett 0 ill. 50 értékkel egészítettük ki.

$$y^{(1)} = (0.721672, 0.875330, 0.774557, 0.213697, 0.088954, 0.145751),$$

$$y^{(2)} = (0.547329, 0.738856, 0.651865, 0.314483, 0.182787, 0.242584).$$

Mint **harmadik tesztfeladatot** egy konkrét kémiai folyamatot leíró rendszert tekintünk, amelyben y_1, y_2 , ill. y_3 egy A, B , ill. C anyag t időben változó koncentrációja. Összefoglalva az $y(t) = (y_1(t), y_2(t), y_3(t))^T$ megoldásvektorba, a következő nemlineáris egyenletrendszer modellezi a jelenséget:

$$\frac{dy}{dt} = \begin{pmatrix} -k_1 & 0 & k_2 y_2 \\ k_1 & -k_3 y_2 & -k_2 y_2 \\ 0 & k_3 y_2 & 0 \end{pmatrix} y, \quad y(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

A paraméterek értékei:

$$k_1 = 0.04, \quad k_2 = 10^4, \quad k_3 = 3 \cdot 10^7.$$

Ezen nemlineáris rendszer Jacobi-mátrixa

$$J = J(y) = \begin{pmatrix} -k_1 & k_2 y_3 & k_2 y_2 \\ k_1 & -2k_3 y_2 - k_2 y_3 & -k_2 y_2 \\ 0 & 2k_3 y_2 & 0 \end{pmatrix}.$$

J sajátértékei (ha $y_i = y_i(0) - t$, $i = 1, 2, 3$, helyettesítünk be)

$$\lambda_1 = \lambda_2 = 0, \quad \lambda_3 = -k_1,$$

tehát induláskor a rendszer nem merev – de már kis idő elteltével azzá válik. A nagy k_3 miatt azzal számolhatunk, hogy hamarosan érvényes $k_3 y_2 \gg k_1$ és $k_3 y_2 \gg k_2 y_3$, és ekkor

$$\lambda_1 = 0, \quad \lambda_2 \underset{<}{\approx} 0, \quad \lambda_3 \approx -2k_3 y_2.$$

A $h = 0.1$ lépéstávolsággal beindítva a számítást, kapjuk a következő eredményeket (hat számjegyre kerekítve):

x_i	y_{1i}	y_{2i}	y_{3i}
0.1	0.996016	0.003984	0.0
0.2	0.996808	0.001992	0.001200
0.3	0.996538	$0.9961 \cdot 10^{-3}$	0.002465
1.0	0.978334	$0.3270 \cdot 10^{-4}$	0.021633

Ekkor J sajátértékei már

$$\lambda_1 = 0, \lambda_2 \approx -0.3306, \lambda_3 \approx -2178.07.$$

Ezen $h = 0.1$ -gyel kapott számok ellenőrzése végett még olyan eredményeket is mutatunk, amelyek a $h = 0.01$ időlépéshez tartoznak. Ekkor kiderül, hogy a fenti számítás stabil volt, de nem különösen pontos:

x_i	y_{1i}	y_{2i}	y_{3i}
0.1	0.996122	$0.3581 \cdot 10^{-4}$	0.003842
0.2	0.992356	$0.3513 \cdot 10^{-4}$	0.007609
0.3	0.988729	$0.3449 \cdot 10^{-4}$	0.011237
1.0	0.966536	$0.3076 \cdot 10^{-4}$	0.033434

Ekkor J sajátértékei

$$\lambda_1 = 0, \lambda_2 \approx -0.2943, \lambda_3 \approx -2179.58.$$

9.7. Feladatok

1. Lipschitz-folytonosak-e a következő függvények, amikor $y \in [0, 1]$?

a) $f(y) = e^y$;

b) $f(y) = \sqrt[3]{y}$;

c) $f(y) = \frac{1}{1-y}$.

2. Írjuk fel az

$$y' = f(x, y), \quad y(0) \text{ adott}$$

kezdetiérték feladat megoldását szolgáló explicit Euler-módszer képletét és magyarázzuk meg az egyes abban előforduló tagok értelmét egy rajzzal az x, y -síkon, az origó közelében.

3. Becsüljük a negyedrendű Runge-Kutta-módszer relatív hatékonyságát az explicit Euler-módszerhez képest a következőképpen:

Legyen e_h az Euler-módszer hibája és h a mindenkori lépéstávolság. Számítsuk ki az

$$e_h := c_1 h + c_2 h^2$$

képlet c_1, c_2 együtthatóit úgy, hogy ez a (9.9)-ben közölt 2 ($h = 1/320$ -hoz és $h = 1/640$ -hoz tartozó) adatra illeszkedjen.

Ebből kapjuk meg azt a h_E lépéstávolságot, amely az explicit Euler-módszert használva ugyanazt a hibát adja, mint a negyedrendű Runge-Kutta-módszer a $h_{RK4} = 0.1$ esetén $x = 1$ -nél (ld. 196. oldalt).

Legyen N_E az Euler-módszernek h_E -hez tartozó és N_{RK4} a negyedrendű Runge-Kutta-módszernek h_{RK4} -hez tartozó függvénykiértékelések száma.

Ekkor a kívánt relatív hatékonyság N_E/N_{RK4} , mert ennyiszor gyorsabban fogja produkálni a negyedrendű módszer ugyanazt a pontosságot, mint az elsőrendű.