

Numerikus módszerek 2 (MSc)

Horváth Róbert

Budapesti Műszaki és Gazdaságtudományi Egyetem
Természettudományi Kar
Matematika Intézet
Analízis Tanszék

2019. tavasz

1.

Tantárgyismertető

Hasznos információk

- ▶ Elérhetőségeim: e-mail: rhorvath@math.bme.hu, Iroda: H.24/b
- ▶ Fogadóóra: Csütörtök 16-17 óra (az irodámban (H24b), e-mail).
- ▶ Gyakorlat: Beadandó feladatok, programok. A feladatok teljesítése és jelenlét szükséges az aláíráshoz.
- ▶ Előadás: követelmények ismertetése
- ▶ Vizsgák: vizsga módja, vizsgára bocsátás feltétele
- ▶ Jegyzet: órai jegyzet, *Horváth, Izsák, Karátson, Parciális differenciálegyenletek numerikus módszerei, elektronikus jegyzet, 2013*, az előadás diái, honlapon szereplő anyagok: <http://anal.math.bme.hu/nummod2>

Felhasznált irodalom

- J.W. Thomas, Numerical Partial Differential Equations, Finite Difference Methods, Texts in Applied Mathematics 22, Springer, 1995
- C. Johnson, Numerical solutions of PDEs by the finite element methods, Dover Publications 2009.
- Horváth, Izsák, Karátson, Parciális differenciálegyenletek numerikus módszerei, elektronikus jegyzet, 2013

Egyéb ajánlott irodalom:

- Morton, Mayers, Numerical solutions of partial differential equations, Cambridge University Press, 2005.
- Alfio Borzi, Introduction to multigrid methods, Universität Würzburg (<https://www.mathematik.uni-wuerzburg.de/~borzi/mgintro.pdf>)
- Stoyan Gisbert, Matlab, Frissített kiadás, Typotex, 2005.
- Stoyan Gisbert, Takó Galina, Numerikus módszerek II-III, Typotex.

2.

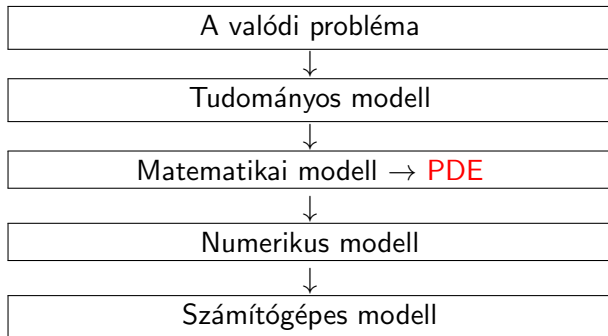
Bevezetés

Mivel foglalkozik a numerikus matematika?

A numerikus matematika a folytonos matematika problémáinak megoldásához konstruál algoritmusokat és elemzi azokat (pontosság, hatékonyság, viselkedés a számítógépes megvalósítások során).

A folytonos matematikai problémák általában valamilyen más tudományág problémáinak matematikai modelljei (fizika, biológia, közgazdaságtan, stb.).

Modellalkotás



Néhány példa (tud. és mat. modellek)

"A boldog családok hasonlók egymáshoz; minden boldogtalan család, a maga módján az." (Tolsztoj)

Transzportegyenletek:

Pl. A keresztmetszetű csőben folyadék áramlik, oldott anyag sűrűsége $u(t, x)$. Egy $[x, x + \Delta x]$ csődarabot tekintünk a $[t, t + \Delta t]$ időintervallumban.

Mérlegegyenlet (megmaradási törvény):

Tömegváltozás = Beáramló - Kiáramló + Forrás

$\psi(t, x)$: fluxus: egységnyi keresztmetszeten, egységnyi idő alatt átáramló anyag.

$f(t, x)$: forrás: időegység alatt egységnyi térfogatban keletkező anyag.

Néhány példa (tud. és mat. modellek)

$$\begin{aligned} & \int_x^{x+\Delta x} A(u(t+\Delta t, y) - u(t, y)) dy = \\ &= \int_t^{t+\Delta t} A(\psi(\tau, x) - \psi(\tau, x + \Delta x)) d\tau + \int_t^{t+\Delta t} \int_x^{x+\Delta x} Af(\tau, y) dy d\tau \end{aligned}$$

\Downarrow

$$u'_t(t, x) + \psi'_x(t, x) = f(t, x)$$

Speciális esetek:

$\psi(t, x) = cu(t, x) \rightarrow$ lineáris transzportegyenlet

$$u'_t(t, x) + cu'_x(t, x) = f(t, x),$$

$$u'_t(t, x) + cu'_x(t, x) - \lambda u(t, x) = 0$$

(advekción-reakció egyenlet)

Néhány példa (tud. és mat. modellek)

$\psi(t, x) = \alpha u(\beta - u) \rightarrow$ Burgers-egyenlet (közlekedési dugó modell). A $v(t, x) = \beta - 2u(t/\alpha, x)$ változócserevel

$$v'_t(t, x) + v(t, x)v'_x(t, x) = 0.$$

Diffúzió, hővezetés

Fourier-törvény, Fick-törvény: $\psi(t, x) = -ku'_x(t, x) \rightarrow$

$$u'_t(t, x) = ku''_{xx}(t, x) + f(t, x).$$

Hullámegyenlet

Newton 2. törvényéből:

$$u''_{tt} = c^2 u''_{xx} + f(t, x).$$

Néhány példa (tud. és mat. modellek)

Ugyanezek magasabb térdimenzióban:

- Transzportegyenlet: $u'_t(t, x) + \nabla \vec{\psi}(t, x) = 0$.
- $\vec{\psi}(t, x) = c\vec{b}u(t, x) \rightarrow u'_t(t, x) + c\vec{b}\nabla u(t, x) = 0$.
- Diffúzió, hővezetés

$$u'_t(t, x) = k\Delta u + f(t, x).$$

Spec.: ($\Delta u = f$ Poisson-egyenlet, $\Delta u = 0$ Laplace-egyenlet)

- Hullámegyenlet

$$u''_{tt}(t, x) = c^2\Delta u + f(t, x).$$

Néhány példa (tud. és mat. modellek)

Black–Scholes-egyenlet

Opcióárazás egyenlete:

$$V'_t + \frac{1}{2}\sigma^2 S^2 V''_{SS} + rSV'_S - rV = 0$$

Navier–Stokes-egyenlet

Nemösszenyomható folyadék áramlása:

$$\rho (\vec{v}'_t + \vec{v} \cdot \nabla \vec{v}) = -\nabla p + \mu \Delta \vec{v} + \vec{f}.$$

Néhány példa (tud. és mat. modellek)

Maxwell-egyenletek

Elektromágneses tér leírása:

$$\frac{\partial(\varepsilon\mathbf{E})}{\partial t} = \nabla \times \mathbf{H} - \sigma\mathbf{E} - \mathbf{J}_e,$$

$$\frac{\partial(\mu\mathbf{H})}{\partial t} = -\nabla \times \mathbf{E},$$

$$\nabla(\varepsilon\mathbf{E}) = \mathbf{0},$$

$$\nabla(\mu\mathbf{H}) = \mathbf{0}.$$

Schrödinger-egyenlet

$$i\hbar\Psi'_t = -\frac{\hbar^2}{2m}\Delta\Psi + V(t, x)\Psi.$$

A PDE-eket kezdeti- és peremfeltételek teszik korrekt kitűzésűvé.

Módszertípusok áttekintése

I. Véges differenciák módszere

II. Integrálmódszerek

- ▶ **Variációs módszer (Ritz-módszer)**
- ▶ **Súlyozott reziduum módszerek**
 - Kollokációs módszer
 - Véges térfogat módszer (vagy résztartomány módszer)
 - **Galjorkin-módszer**
 - Momentumok módszere
 - Legkisebb négyzetek módszere

Megj.: Ha az integrálmódszerekben a tesztfüggvényeket kis tartójú, szakaszonként polinomiális függvényeknek választjuk, akkor végeelem módszerről beszélünk.

3.

Elliptikus egyenletek numerikus megoldása (véges differencia)

A Poisson-egyenlet téglalapon

Poisson-egyenlet 2D-ban, téglalapon, Dirichlet-peremmel

$$-\Delta u = -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f(x, y), \quad (x, y) \in \Omega := (0, a) \times (0, b)$$
$$u|_{\Gamma} = g(x, y), \quad \Gamma := \partial\Omega.$$

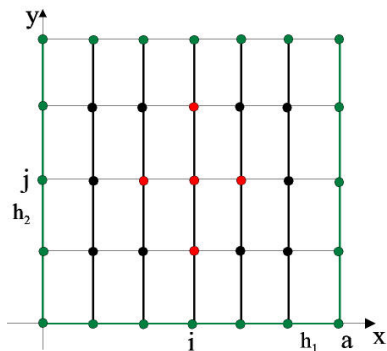
Véges differencia megoldáshoz rácsot definiálunk:

$$\bar{\Omega}_h := \underbrace{\{(x, y) \in \bar{\Omega} \mid x = ih_1, y = jh_2,}_{x_{ij}}$$
$$i = 0, \dots, n_1, j = 0, \dots, n_2\},$$

$$\bar{\Omega}_h \cap \Omega =: \Omega_h,$$

$$\bar{\Omega}_h \cap \Gamma =: \Gamma_h.$$

A véges differencia séma



Alkalmazzuk az

$$u_i^j \approx u(ih_1, jh_2), \quad f_{ij} = f(ih_1, jh_2), \quad g_{ij} = g(ih_1, jh_2)$$

közelítéseket, melyekkel a séma alakja a következő lesz.

A véges differencia séma

Véges differencia séma, Poisson-egyenlet, téglalap, Dirichlet-perem

$$-\frac{u_{i-1}^j - 2u_i^j + u_{i+1}^j}{h_1^2} - \frac{u_i^{j-1} - 2u_i^j + u_i^{j+1}}{h_2^2} = f_{ij}, \quad x_{ij} \in \Omega_h$$

$$u_i^j = g_{ij}, \quad x_{ij} \in \Gamma_h.$$

A fenti egyenletek egy $(n_1 + 1)(n_2 + 1)$ méretű lineáris egyenletrendszert eredményeznek

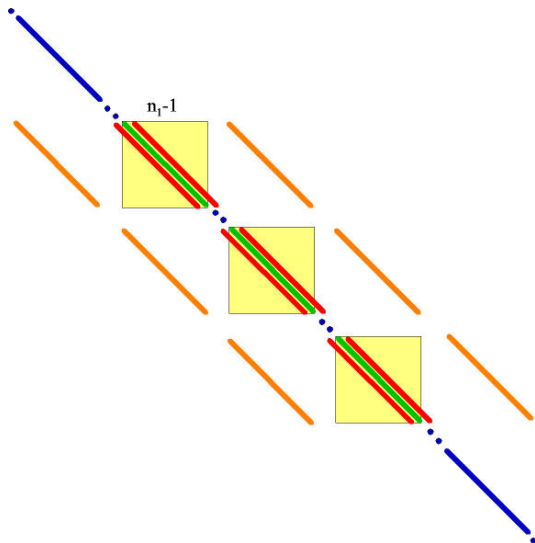
$$\mathbf{A}_h \bar{\mathbf{u}}_h = \bar{\boldsymbol{\phi}}_h,$$

ahol

$$(\bar{\boldsymbol{\phi}}_h)_{ij} = \begin{cases} f_{ij}, & x_{ij} \in \Omega_h, \\ g_{ij}, & x_{ij} \in \Gamma_h, \end{cases}, \quad \|\bar{\boldsymbol{\phi}}_h\|_\infty \leq \max\{\|g\|_{C(\Gamma)}, \|f\|_{C(\Omega)}\}$$

és a mátrix ill. vektorok alakja attól függ, hogy hogy rendeztük a rácspontokat. A továbbiakban a sorfolytonos sorszámozást alkalmazzuk (lentől felfelé, balról jobbra).

Az A_h mátrix szerkezete



kék:

$$1$$

zöld:

$$\frac{2}{h_1^2} + \frac{2}{h_2^2}$$

piros:

$$-\frac{1}{h_1^2}$$

narancs:

$$-\frac{1}{h_2^2}$$

Az \mathbf{A}_h mátrix szerkezete

Tétel. Az \mathbf{A}_h mátrix M-mátrix.

Biz. Előjelek stimmelnek. Tekintsük a

$$w(x, y) = 4 + x(a - x) + y(b - y)$$

függvényt és definiáljuk a $\bar{\mathbf{w}}$,

$$w_{ij} = w(ih_1, jh_2)$$

pozitív vektort, amit ugyanúgy rendezünk, mint az $\bar{\mathbf{u}}_h$ vektor elemeit. Erre a vektorra

$$(\mathbf{A}_h \bar{\mathbf{w}})_{ij} = \begin{cases} 4, & x_{ij} \in \Omega_h, \\ \geq 4, & x_{ij} \in \Gamma_h, \end{cases}$$

ami pozitív vektor. ■

Az A_h mátrix szerkezete

Megj. Ha $f_1 \geq f_2$ és $g_1 \geq g_2$, akkor az ezekkel előállított két numerikus megoldásra érvényes

$$(\bar{\mathbf{u}}_h)_1 \geq (\bar{\mathbf{u}}_h)_2.$$

Speciálisan ebből következik a nemnegativitás-megőrzési tulajdonság.

A numerikus séma konvergenciája

Legyen $E = \max_{i=0, \dots, n_1; j=0, \dots, n_2} |u_i^j - u(ih_1, jh_2)|$.

Def. A numerikus sémát *konvergensnek* hívjuk, ha $h := \max\{h_1, h_2\} \rightarrow 0$ esetén $E \rightarrow 0$. Ha $E = \mathcal{O}(h^r)$, akkor azt mondjuk, hogy a konvergencia rendje r .

Def. A numerikus séma *stabil*, ha $\|\mathbf{A}_h^{-1}\|$ felülről korlátos a rács finomságától függetlenül.

Def. A numerikus séma *konzisztens*, ha képlethibája minden pontban nullához tart, ha $h \rightarrow 0$. A konzisztenciarend r , ha a képlethiba minden pontban $\mathcal{O}(h^r)$ nagyságrendű (minden pontban ugyanazzal a konstanssal).

Tétel. Konzisztencia + stabilitás \Rightarrow konvergencia.

A numerikus séma konvergenciája

Tétel. A numerikus séma stabil, ugyanis

$$\|\mathbf{A}_h^{-1}\|_\infty \leq 1 + \frac{a^2 + b^2}{16}.$$

Biz. Az M-mátrixok inverzének maximumnormabeli becsléséből következik, hogy

$$\|\mathbf{A}_h^{-1}\|_\infty \leq \frac{4 + a^2/4 + b^2/4}{4} = 1 + \frac{a^2 + b^2}{16}. \blacksquare$$

A numerikus séma konvergenciája

Tétel. Amennyiben a vizsgált peremértékfeladat megoldása $C^4(\bar{\Omega})$ -beli, akkor a vizsgált numerikus séma másodrendben konzisztens.

Biz. Jelentse itt $(u)_i^j$ a pontos megoldás értékét az (i, j) rácspontban.

$$\begin{aligned} T_i^j &:= -\frac{(u)_{i-1}^j - 2(u)_i^j + (u)_{i+1}^j}{h_1^2} - \frac{(u)_i^{j-1} - 2(u)_i^j + (u)_i^{j+1}}{h_2^2} - f_{ij} \\ &= -(\partial_x^2 u)_i^j - \frac{(\partial_x^4 u)_{i^*}^j h_1^2}{12} - (\partial_y^2 u)_i^j - \frac{(\partial_y^4 u)_i^{j^*} h_2^2}{12} - f_{ij}, \end{aligned}$$

azaz

$$|T_i^j| \leq \frac{M_{4,0} h_1^2}{12} + \frac{M_{0,4} h_2^2}{12} \leq \frac{M h^2}{6}$$

a rácsponttól függetlenül. (Az M értékek a szokott módon a deriváltak felső becslései.) Ez mutatja a másodrendű konzisztenciát. ■

A numerikus séma konvergenciája

Tétel. Amennyiben a vizsgált peremértékfeladat megoldása $C^4(\overline{\Omega})$ -beli, akkor a vizsgált numerikus séma másodrendben konvergens.

Biz. Legyen a képlethibák vektora megfelelően rendezve $\overline{\mathbf{T}}_h$. Legyen $u_h(\overline{\mathbf{x}})$ a pontos megoldás értékeinek vektora a rácspontokban. Ekkor az

$$\mathbf{A}_h \overline{\mathbf{u}}_h - \overline{\boldsymbol{\phi}}_h = \mathbf{0}$$

és

$$\mathbf{A}_h u_h(\overline{\mathbf{x}}) - \overline{\boldsymbol{\phi}}_h = \overline{\mathbf{T}}_h$$

kivonásából adódik, hogy

$$\mathbf{A}_h (\overline{\mathbf{u}}_h - u_h(\overline{\mathbf{x}})) = -\overline{\mathbf{T}}_h,$$

melyből ...

A numerikus séma konvergenciája

...

$$\|\bar{\mathbf{u}}_h - u_h(\bar{\mathbf{x}})\|_\infty \leq \|\mathbf{A}_h^{-1}\|_\infty \|\bar{\mathbf{T}}_h\|_\infty$$

és a korábbi becsléseket használva

$$\|\bar{\mathbf{u}}_h - u_h(\bar{\mathbf{x}})\|_\infty \leq \left(1 + \frac{a^2 + b^2}{16}\right) \frac{Mh^2}{6},$$

amely mutatja a másodrendű konvergenciát. ■

Maximum-elv

A folytonos feladatra: Ha $f \leq 0$, akkor a megoldás a peremen veszi fel a maximumát. Igaz-e ez a numerikus megoldásra?

Rendezzük át az

$$\mathbf{A}_h \bar{\mathbf{u}}_h = \bar{\boldsymbol{\phi}}_h$$

egyenletrendszer sorait és oszlopait úgy, hogy előre kerüljenek Ω belső pontjai, és utána következzenek a határpontok. Ekkor a lineáris egyenletrendszer az alábbi alakú lesz:

$$\begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_\partial \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{u}}_0 \\ \bar{\mathbf{u}}_\partial \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{f}} \\ \bar{\mathbf{g}} \end{bmatrix}$$

Tétel. A numerikus séma kielégíti a maximum-elv diszkrét változatát, azaz, ha $\bar{\mathbf{f}} \leq 0$, akkor

$$\max(\bar{\mathbf{u}}_0) \leq \max(\bar{\mathbf{g}}).$$

Maximum-elv

Vegyük észre, hogy \mathbf{A}_0 M-mátrix ($\mathbf{A}_0^{-1} \geq \mathbf{0}$) és \mathbf{A}_∂ elemei nempozitívak, valamint

$$[\mathbf{A}_0 \mathbf{A}_\partial][1, \dots, 1]^T = \mathbf{0}, \quad [1, \dots, 1]^T =: [\bar{\mathbf{e}}_0^T, \bar{\mathbf{e}}_\partial^T]^T,$$

azaz $\bar{\mathbf{e}}_0 = -\mathbf{A}_0^{-1} \mathbf{A}_\partial \bar{\mathbf{e}}_\partial$.

Emiatt

$$\begin{aligned} \bar{\mathbf{u}}_0 &= \mathbf{A}_0^{-1} \bar{\mathbf{f}} - \mathbf{A}_0^{-1} \mathbf{A}_\partial \bar{\mathbf{u}}_\partial = \mathbf{A}_0^{-1} \bar{\mathbf{f}} - \mathbf{A}_0^{-1} \mathbf{A}_\partial \bar{\mathbf{g}} \\ &\leq \mathbf{A}_0^{-1} \bar{\mathbf{f}} - \mathbf{A}_0^{-1} \mathbf{A}_\partial \bar{\mathbf{e}}_\partial \max(\bar{\mathbf{g}}) \leq \bar{\mathbf{e}}_0 \max(\bar{\mathbf{g}}). \quad \blacksquare \end{aligned}$$

Poisson-egyenlet Robin-peremmel

Poisson-egyenlet 2D-ban, téglalapon, Robin-peremmel

$$-\Delta u = -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f(x, y), \quad (x, y) \in \Omega := (0, a) \times (0, b)$$
$$\frac{\partial u}{\partial n} = \sigma(x, y)(u_0(x, y) - u), \quad (x, y) \in \Gamma.$$

Tegyük fel, hogy f, σ, u_0 olyan függvények, melyekkel az egyenletnek van $C^4(\overline{\Omega})$ -beli megoldása.

A véges differenciás megoldás diszkretizációja nyilvánvalóan a belső pontokban ugyanaz, mint a Dirichlet-esetben, azaz csak a tartomány határpontjaiban kellene megadni (lehetőleg másodfokú) approximációkat.

Approximáció az oldaléleknél

$$\frac{(u)_{n_1}^j - (u)_{n_1-1}^j}{h_1} \\ = (\partial_x u)_{n_1}^j - \frac{h_1}{2} (\partial_x^2 u)_{n_1}^j + \frac{h_1^2}{6} (\partial_x^3 u)_{n_1}^{j*},$$

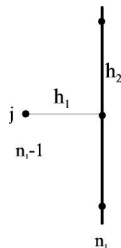
ahol $(\partial_x u)_{n_1}^j = \sigma_{n_1,j}((u_0)_{n_1}^j - u_{n_1}^j)$ a peremfeltétel miatt, és

$$(\partial_x^2 u)_{n_1}^j = -(\partial_y^2 u)_{n_1}^j - f_{n_1,j}$$

és

$$-(\partial_y^2 u)_{n_1}^j = -\frac{(u)_{n_1}^{j+1} - 2(u)_{n_1}^j + (u)_{n_1}^{j-1}}{h_2^2} - \frac{h_2^2}{12} (\partial_y^4 u)_{n_1}^{j*}$$

a differenciálegyenletből.



Approximáció az oldaléleknél

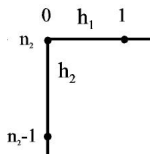
Így a séma

$$\frac{u_{n_1}^j - u_{n_1-1}^j}{h_1} = \sigma_{n_1,j}((u_0)_{n_1}^j - u_{n_1}^j) - \frac{h_1}{2} \left(-\frac{u_{n_1}^{j+1} - 2u_{n_1}^j + u_{n_1}^{j-1}}{h_2^2} - f_{n_1,j} \right).$$

Hasonlóan járunk el a többi oldalán a tartománynak.

A képlethiba $\mathcal{O}(h_1^2 + h_2^2)$.

Approximáció a sarkokban



$$\frac{(u)_0^{n_2} - (u)_0^{n_2-1}}{h_2} = (\partial_y u)_0^{n_2} - \frac{h_2}{2} (\partial_y^2 u)_0^{n_2} + \frac{h_2^2}{6} (\partial_y^3 u)_0^{n_2*} \quad / \cdot \frac{-2}{h_2},$$

$$\frac{(u)_1^{n_2} - (u)_0^{n_2}}{h_1} = (\partial_x u)_0^{n_2} + \frac{h_1}{2} (\partial_x^2 u)_0^{n_2} + \frac{h_1^2}{6} (\partial_x^3 u)_0^{n_2*} \quad / \cdot \frac{2}{h_1}.$$

A két egyenletet összeadva és kihasználva a peremfeltételt és a differenciálegyenletet, az alábbi sémát kapjuk ($H = h_1 h_2 / 2 / (h_1 + h_2)$):

$$H \left(-2 \frac{u_0^{n_2} - u_0^{n_2-1}}{h_2^2} + 2 \frac{u_1^{n_2} - u_0^{n_2}}{h_1^2} \right) = -\sigma_{0,n_2} ((u_0)_0^{n_2} - u_0^{n_2}) - H f_{0,n_2}$$

Approximáció a sarkokban

A többi sarokban is így járunk el. A képlethiba most is $\mathcal{O}(h_1^2 + h_2^2)$ nagyságrendű.

Tétel. Ha $\sigma(x, y) \geq \sigma_0 > 0$ és az f, σ, u_0 függvények elég simák ahhoz, hogy a feladat u megoldása $C^4(\bar{\Omega})$ -beli legyen, akkor a differenciasémánk másodrendben konvergál a pontos megoldáshoz.

Biz. A konzisztencia az approximációs képletekből látható. A stabilitás a korábbi M-mátrixos technikával igazolható az

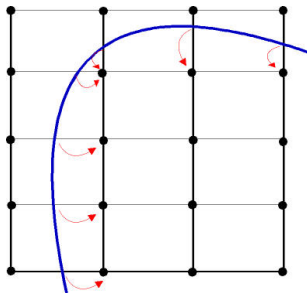
$$\|\bar{\mathbf{u}}_h\|_\infty \leq \left(\frac{a^2 + b^2}{16} + \frac{4 + a + b}{4\sigma_0} \right) \|\bar{\phi}_h\|$$

alakban.

Nem téglalap alakú tartományok

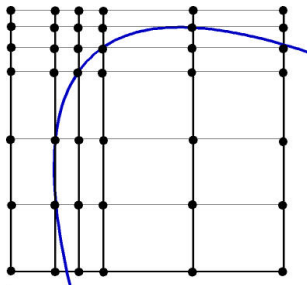
Mit lehet tenni, ha nem téglalap alakú az Ω tartomány?

- ▶ A peremértékeket a rácspontokban közelítjük (elveszik a másodrendű konvergencia).



Nem téglalap alakú tartományok

- ▶ Sűrítjük a rácsot.



- ▶ Hozzávesszük rácspontnak a rács és a perem metszéspontjait (Shortly–Weller-approximáció).
- ▶ Téglalapba transzformáljuk a tartományt.

Shortly–Weller-approximáció részletesebben

A belső pontokban ugyanaz a közelítés, mint korábban, a peremen pedig

$$\begin{aligned} & \frac{2}{h_1^- + h_1^+} \left(\frac{u_{i+1}^j - u_i^j}{h_1^+} - \frac{u_i^j - u_{i-1}^j}{h_1^-} \right) \\ & + \frac{2}{h_2^- + h_2^+} \left(\frac{u_i^{j+1} - u_i^j}{h_2^+} - \frac{u_i^j - u_i^{j-1}}{h_2^-} \right) = -f_{ij}. \end{aligned}$$

Ezen egyenletek egy egyenletrendszerre vezetnek: $\mathbf{A}_{S,h} \bar{\mathbf{u}}_h = \bar{\boldsymbol{\phi}}_h$,
érvényes a

$$\|\mathbf{A}_{S,h}^{-1}\|_{\infty} \leq 1 + \frac{1}{4}(\text{diam}(\Omega))^3$$

stabilitási becslés. Ha $u \in C^4(\bar{\Omega})$, akkor a módszer másodrendben konvergens lesz.

4.

Többrácsos módszerek

A megoldandó lineáris egyenletrendszer

$$\begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_\partial \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{u}}_0 \\ \bar{\mathbf{u}}_\partial \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{f}} \\ \bar{\mathbf{g}} \end{bmatrix}$$

Olyan egyenletrendszert kell tudnunk megoldani, melynek együtthatómátrixa \mathbf{A}_0 . Ez egy szimmetrikus M -mátrix.

1. Tétel. A szimmetrikus M -mátrixok pozitív definiték.

Biz. Jelölje a mátrixot \mathbf{M} ! Szimmetrikus mátrix minden sajátértéke valós. Azt kell megmutatni, hogy mindegyik pozitív. A főátlóban álló elemek pozitívak. Legyen ezek maximuma μ , és írjuk fel \mathbf{M} -et $\mathbf{M} = \mu\mathbf{E} - \mathbf{H}$ alakban, ahol \mathbf{H} nyilvánvalóan nemnegatív mátrix. Így a fenti felírás egy nemnegatív inverzű mátrix (az M mátrixok ilyenek) egy reguláris felbontása. Azaz $\rho((1/\mu)\mathbf{H}) < 1$, ami azt jelenti, hogy $\rho(\mathbf{H}) < \mu$. Ha \mathbf{H} sajátértékeit λ_k jelöli, akkor \mathbf{M} $\mu - \lambda_k$ alakú sajátértékei nyilván pozitívak. ■

Szimmetrikus, pozitív definit mátrixú lineáris egyenletrendszerek megoldása

Az egyenletrendszer azon túl, hogy szimmetrikus, pozitív definit mátrixú, még ritka mátrixú is, így jól működnek az iterációs módszerek is rá.

Lehetséges megoldási módszerek ($\bar{\mathbf{u}}_0 \in \mathbb{R}^{N_0 \times N_0}$):

- ▶ Gauss-módszer ($2N_0^3/3$, nem praktikus)
- ▶ Cholesky-felbontás ($N_0^3/3$)
- ▶ Jacobi- és Gauss-Seidel ill. relaxált változatai ($N_0^2 \ln(1/\varepsilon)$, ε : elérni kívánt pontosság a leállási feltételben)
- ▶ Konjugált gradiens módszer ($N_0^{3/2} \ln(1/\varepsilon)$)
- ▶ **Többrácsos módszer.** 1960-as évek, Fedorenko (1962), Bakvalov (1966). Használjuk ki, hogy az egyenletrendszer egy elliptikus feladat diszkretizációjából keletkezett!

Iterációk simító tulajdonsága

Iterációk simító tulajdonsága

Egydimenziós peremértékfeladat

$$-u''(x) = f(x), \quad x \in (0, 1), \quad u(0) = u(1) = 0$$

A szokásos véges differenciás közelítés során megoldandó lineáris egyenletrendszer:

$$\begin{bmatrix} 2 & -1 & & & \\ -1 & \ddots & \ddots & & \\ & & \ddots & & \\ & & & -1 & \\ -1 & & & 2 & \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = h^2 \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{bmatrix}$$

vagy tömörebben

$$\mathbf{A}_h \bar{\mathbf{u}}_h = \bar{\phi}_h.$$

Iterációk simító tulajdonsága

Oldjuk meg az $\mathbf{A}_h \bar{\mathbf{u}}_h = \bar{\boldsymbol{\phi}}_h$ egyenletrendszert a relaxált Jacobi-iterációval. Ekkor az

$$\begin{aligned}\bar{\mathbf{u}}_h^{(k+1)} &= ((1 - \omega)\mathbf{E} + \omega\mathbf{D}^{-1}(\mathbf{L} + \mathbf{R}))\bar{\mathbf{u}}_h^{(k)} + \omega\mathbf{D}^{-1}\bar{\boldsymbol{\phi}}_h \\ &= \underbrace{(\mathbf{E} - \omega\mathbf{D}^{-1}(\mathbf{D} - \mathbf{L} - \mathbf{R}))}_{\mathbf{A}_h} \bar{\mathbf{u}}_h^{(k)} + \omega\mathbf{D}^{-1}\bar{\boldsymbol{\phi}}_h \\ &\quad \underbrace{\hspace{10em}}_{\mathbf{B}_J(\omega)}\end{aligned}$$

iterációt kapjuk.

A konvergenciához a $\rho(\mathbf{B}_J(\omega)) < 1$ feltételnek kell teljesülni.

\mathbf{A}_h sajátrendszere ($h = 1/(n + 1)$):

$$\bar{\mathbf{v}}_l = \sin(l\pi h \bar{\mathbf{j}}), \quad \bar{\mathbf{j}} = [1, \dots, n]^T, \quad \lambda_l(\mathbf{A}_h) = 2 - 2 \cos(l\pi h), \quad l = 1, \dots, n.$$

Így $\mathbf{B}_J(\omega)$ sajátrendszere:

$$\bar{\mathbf{v}}_l, \quad \lambda_l(\mathbf{B}_J(\omega)) = 1 - \omega(1 - \cos(l\pi h)),$$

azaz, ha $\omega \in (0, 1]$, akkor a JOR-módszer konvergens lesz.

Iterációk simító tulajdonsága

Megj. Vegyük észre, hogy ha pl. $\omega = 1$, akkor $\rho(\mathbf{B}_J(\omega)) = \cos(\pi h) \rightarrow 1$, ha $h \rightarrow 0$, azaz egyre lassabban konvergál a Jacobi-módszer. Az is látszik, hogy a magas frekvenciás összetevők (sajátvektorok) együtthatója hamar kicsivé válik, így ezek az összetevők hamar eltűnnek a hibából. Hogyan lehet ezeket a leggyorsabban eltüntetni?

Alacsony frekvenciás összetevők (LF): $1 \leq l < n/2$.

Magas frekvenciás összetevők (HF): $n/2 \leq l \leq n$.

2. Def. μ legyen az a legnagyobb tényező, amennyivel a HF összetevők csökkennek egy iterációban. μ -t simítási tényezőnek nevezzük.

Iterációk simító tulajdonsága

$$\mu = \max\{|\lambda_k(\mathbf{B}_J(\omega))| \mid n/2 \leq l \leq n\} = \max\{|1 - \omega|, |1 - 2\omega|\}$$

Ez akkor lesz minimum, ha $\omega = 2/3$, és a minimum értéke $\mu = 1/3$. Azaz a HF összetevők 2 iteráció alatt a tizedükre csökkennek, annak ellenére, hogy a konvergencia lassú. Ha a HF összetevők kicsik már, akkor a hibavektor jól közelíthető egy durvább rácson is.

A kétrácsos módszer

A maradékegyenlet

Térjünk vissza a kétdimenziós Poisson-egyenletből származtatott $\mathbf{A}_h \bar{\mathbf{u}}_h = \bar{\boldsymbol{\phi}}_h$ egyenletrendszer megoldásához! Végezzünk el néhány iterációs lépést valamilyen simító iterációval valamilyen kezdővektorral indulva. Kapjuk ekkor az $\tilde{\mathbf{u}}_h$ vektort.

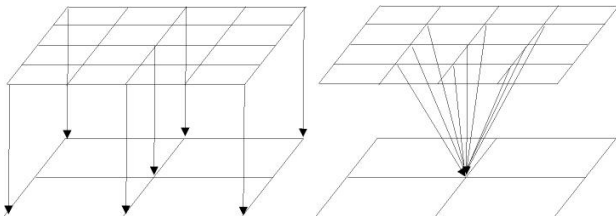
Ekkor az $\bar{\mathbf{e}}_h = \bar{\mathbf{u}}_h - \tilde{\mathbf{u}}_h$ hibavektor egy sima vektor lesz, és érvényes rá az

$$\mathbf{A}_h \bar{\mathbf{e}}_h = \mathbf{A}_h (\bar{\mathbf{u}}_h - \tilde{\mathbf{u}}_h) = \bar{\boldsymbol{\phi}}_h - \mathbf{A}_h \tilde{\mathbf{u}}_h =: \bar{\mathbf{r}}_h \text{ (maradékvektor),}$$

egyenlőség (ún. maradékegyenlet), ami megoldható $\bar{\mathbf{e}}_h$ -ra. De mivel a jobb oldala és a megoldása is sima az egyenletrendszernek, jó közelítést kapnánk akkor is, ha egy durvább rácson ($H = 2h$) oldanánk meg egy ugyanilyen egyenletrendszert.

Restrikció

Térjünk át a durvább rácásra: $\bar{\mathbf{r}}_H = R_h^H \bar{\mathbf{r}}_h$. Ezt egy ún. restriktációs operátorral valósíthatjuk meg: R_h^H . Ez általában valamilyen súlyozott vetítés.



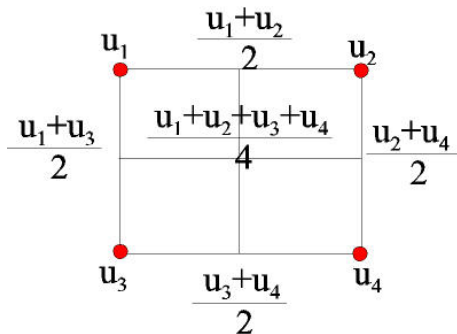
Teljes és részleges súlyozás:

$$\frac{1}{8} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Megoldjuk az $\mathbf{A}_H \bar{\mathbf{e}}_H = \bar{\mathbf{r}}_H$ egyenletet, ahol \mathbf{A}_H a differenciáloperátor közelítése a durvább rácson.

Prolongáció

Az \bar{e}_H hibavektort visszavetítjük a finomabb rácsra. Ezt egy prolongációs \mathbf{P}_H^h operátorral tesszük meg, amely általában valamilyen interpolációt jelent: $\bar{e}_h = \mathbf{P}_H^h \bar{e}_H$.



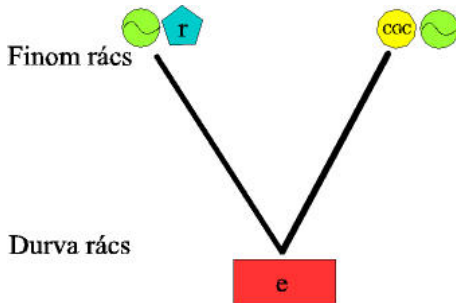
Course Grid Correction (CGC)

Adunk egy jobb közelítést $\bar{\mathbf{u}}_h$ -ra:

$$\bar{\mathbf{u}}_h^{new} = \tilde{\mathbf{u}}_h + \mathbf{P}_H^h \bar{\mathbf{e}}_H.$$

Megj.: Az interpoláció miatt megjelenhetnek HF összetevő, amiket egy újabb simítási ciklussal lehet eltüntetni.

A bemutatott eljárás egy V-alakú sémában foglalható össze (V-ciklus).



V-ciklus

Megj.: Természetesen a V-ciklusból többet kell végrehajtani ahhoz, hogy a pontos megoldást kellően megközelítsük.

3. Tétel. A V-ciklus iterációs mátrixa (\mathbf{S}_h a simítás):

$$\mathbf{S}_h^{k_2} (\mathbf{E} - \mathbf{P}_H^h \mathbf{A}_H^{-1} \mathbf{R}_h^H \mathbf{A}_h) \mathbf{S}_h^{k_1},$$

ezen mátrix spektrálsugara kisebb 1-nél, azaz konvergens.

4. Tétel. Csak önmagában a CGC nem konvergens.

Biz.: Mátrixa

$$\mathbf{E} - \mathbf{P}_H^h \mathbf{A}_H^{-1} \mathbf{R}_h^H \mathbf{A}_h,$$

ahol $\mathbf{P}_H^h \mathbf{A}_H^{-1} \mathbf{R}_h^H \mathbf{A}_h$ nem invertálható, így egyik sajátértéke nulla. Így a mátrixnak 1 sajátértéke lesz, azaz spektrálsugara legalább 1.



A többrácsos módszer (multigrid)

A többrácsos módszer alap gondolata

Ötlet: Az egyenlet durva rácson való megoldásához használjuk egy újabb V-ciklust egy még durvább ráccsal. Több rácsot definiálunk:

$$h_1 > h_2 = h_1/2 > h_3 = h_2/2 > \dots h_L > 0 \quad (L \text{ szintek száma}).$$

Legyen a k -edik szinten n_k a feladat mérete. A k -edik szinten az $\mathbf{A}_k \bar{\mathbf{u}}_k = \bar{\boldsymbol{\phi}}_k$ egyenletrendszert kell megoldanunk. A restriktív operátor legyen \mathbf{R}_k^{k-1} és a prolongáció \mathbf{P}_{k-1}^k .

A többrácsos módszer algoritmus

Oldjuk meg az $\mathbf{A}_k \bar{\mathbf{u}}_k = \bar{\boldsymbol{\phi}}_k$ egyenletrendszert.

1. Ha $k = 1$, akkor oldjuk meg direkt módon az egyenletrendszert.
2. Elősimítás, k_1 darab lépés $\rightarrow \tilde{\mathbf{u}}_k$.
3. Maradék kiszámítása: $\bar{\boldsymbol{\phi}}_k - \mathbf{A}_k \tilde{\mathbf{u}}_k = \bar{\mathbf{r}}_k$.
4. A maradék restriktója: $\bar{\mathbf{r}}_{k-1} = \mathbf{R}_k^{k-1} \bar{\mathbf{r}}_k$.
5. Legyen $\bar{\mathbf{u}}_{k-1} = \mathbf{0}$.
6. Hívjuk meg γ -szor a többrácsos módszert az $\mathbf{A}_{k-1} \bar{\mathbf{u}}_{k-1} = \bar{\mathbf{r}}_{k-1}$ egyenletre.
7. CGC: $\bar{\mathbf{u}}_k = \tilde{\mathbf{u}}_k + \mathbf{P}_{k-1}^k \bar{\mathbf{u}}_{k-1}$.
8. Utósimítás k_2 -szer.

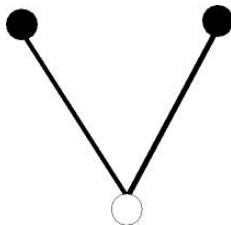
A többrácsos módszer algoritmus

5. Def. γ a ciklusindex, azt adja meg, hogy a durva rácson hányszor használjuk a többrácsos algoritmust. Általában $\gamma = 1$ (V-ciklus) vagy $\gamma = 2$ (W-ciklus).

6. Tétel. A többrácsos módszer iterációs mátrixa $k = 1$ esetén $\mathbf{0}$, különben $k = 2, \dots, L$ esetén

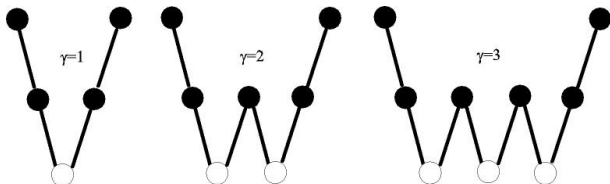
$$\mathbf{M}_k = \mathbf{S}_k^{k_2} (\mathbf{E} - \mathbf{P}_{k-1}^k (\mathbf{E} - \mathbf{M}_{k-1}^\gamma) \mathbf{A}_{k-1}^{-1} \mathbf{R}_k^{k-1} \mathbf{A}_k) \mathbf{S}_k^{k_1}.$$

Kétrácsos módszer: üres kör: pontos megoldás, teli kör: simítás, maradékszámítás, CGC

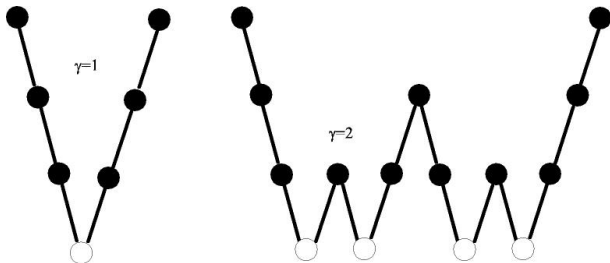


A többrácsos módszer algoritmus

Három rács esete:



Négy rács esete:



Műveletszámok vizsgálata

2D Poisson-egyenlet, max. 5 nemnulla elem soronként, simításhoz lineáris vektoriteráció

A k . szint műveletszáma:

1. $10k_1n_k$: elősimítás
2. $10n_k$: maradékszámítás
3. $10n_k/4$: teljes restrikció
4. $6n_k/4$: bilineáris interpoláció
5. $10k_2n_k$: utósimítás

Ez összesen $k_1 + k_2 = 2$ esetén $34n_k = \bar{Q}_k$.

\bar{Q}_1 a legdurvább rácson az egyenlet pontos megoldásának műveletszáma.

Műveletszámok vizsgálata

Q_k legyen a k . rács műveletszáma. Ekkor $Q_2 = \bar{Q}_2 + \bar{Q}_1$,
 $Q_3 = \gamma(\bar{Q}_2 + \bar{Q}_1) + \bar{Q}_3$, stb.

$$\begin{aligned}Q_k &= \sum_{l=0}^{k-2} \gamma^l \bar{Q}_{k-l} + \gamma^{k-2} \bar{Q}_1 \\&= \sum_{l=0}^{k-2} \gamma^l 34 n_{k-l} + \gamma^{k-2} \bar{Q}_1 \\&= \sum_{l=0}^{k-2} \gamma^l 34 \underbrace{\frac{n_k}{4^l}}_{\text{2D probléma}} + \gamma^{k-2} \bar{Q}_1 \\&= 34 n_k \sum_{l=0}^{k-2} \left(\frac{\gamma}{4}\right)^l + \gamma^{k-2} \bar{Q}_1 \\&\approx 34 n_k \frac{1}{1 - \gamma/4} = \mathcal{O}(n_k)\end{aligned}$$

Megjegyzések

Megj.: $\gamma = 1$ esetén $Q_k \approx 40n_k$, $\gamma = 2$ esetén $Q_k \approx 68n_k$ és $\gamma = 3$ esetén $Q_k \approx 136n_k$

Megj.: A 2D Poisson-feladatra csak a $\gamma \leq 3$ esetek praktikusak

Megj.: Kétrácsos módszer esetén, vagy több rácstra $\gamma = 1, 2$ esetén a konvergencia tényező $\mu^{k_1+k_2}$ a h rácstávolságtól függetlenül.

5.

Az egydimenziós hővezetési egyenlet véges differenciás megoldása

A hővezetési vagy diffúziós egyenlet

A hővezetési vagy diffúziós egyenlet

$$\frac{\partial u}{\partial t} = \Delta u + f$$

+ kezdeti- és peremfeltétel

1D-ban

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in (0, 1)$$
$$u(0, x) = u_0(x),$$
$$u(t, 0) = u(t, 1) = 0.$$

Más egyenletek, peremfeltételek → gyakorlaton.

A hővezetési vagy diffúziós egyenlet

A pontos megoldás előállítható Fourier-sor alakban:

$$u(t, x) = \sum_{l=1}^{\infty} a_l e^{-l^2 \pi^2 t} \sin(l \pi x),$$

ahol $a_l = 2 \int_0^1 u_0(x) \sin(l \pi x) dx$.

Problémák:

- ▶ a_l meghatározásához általában numerikus integrálás kell.
- ▶ Végtelen sort kell összegezni.

Explicit Euler-módszer (EE)

Rácshálót definiálunk a $[0, t_{\max}] \times [0, 1]$ intervallumon:

$$0 = x_0 < x_1 < \dots < x_{n+1} = 1, \quad x_{k+1} - x_k = \Delta x = \frac{1}{n+1},$$

legyen továbbá $\Delta t > 0$. Vezessük be az alábbi jelölést: u_j^k az $u(k\Delta t, x_j)$ közelítése.

Az egyenletben szereplő deriváltakat véges differenciákkal helyettesítve az alábbi sémát kapjuk.

$$\frac{u_j^{k+1} - u_j^k}{\Delta t} = \frac{u_{j-1}^k - 2u_j^k + u_{j+1}^k}{\Delta x^2}, \quad (1)$$

$u_0^k = u_{n+1}^k = 0$, u_j^0 adott u_0 segítségével.

Szokás a $\Delta t = \tau$, ill. $\Delta x = h$ jelölés is.

Az Explicit Euler-módszer konzisztenciája

Mennyire jogos a fenti közelítés?

7. Def. Lokális hiba: Mekkora téved a módszer a k . időrétegen lévő pontos megoldásértékekből indulva $u((k+1)\Delta t, j\Delta x)$ közelítésében? Jelölés: $\Delta t T_j^{k+1}$.

8. Def. Képlethiba: Mekkora az eltérés (1) két oldala között, ha az u_j^k értékek helyett a pontos $u(k\Delta t, j\Delta x)$ értékeket helyettesítjük? (A differenciálegyenlet két oldala között nyilván nulla az eltérés.) Jelölés: T_j^{k+1} .

Látható, hogy a lokális hiba Δt -szerese a képlethibának. Ezt tükrözi a jelölés is.

Az Explicit Euler-módszer konzisztenciája

Számítsuk ki a képlethibát!

$$\begin{aligned} T_j^{k+1} &= \frac{u(t_{k+1}, x_j) - u(t_k, x_j)}{\Delta t} \\ &\quad - \frac{u(t_k, x_{j-1}) - 2u(t_k, x_j) + u(t_k, x_{j+1}))}{\Delta x^2} \\ &= \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(\xi, x_j) - \frac{\Delta x^2}{12} \frac{\partial^4 u}{\partial x^4}(t_k, \eta). \end{aligned}$$

Ha a megoldás kellően sima függvény ($C^{2,4}$ -beli, peremfeltétellel konzisztens kezdeti feltétel is kell hozzá), akkor vannak olyan pozitív c_1, c_2 konstansok, hogy

$$|T_j^{k+1}| \leq c_1 \Delta t + c_2 \Delta x^2.$$

Az Explicit Euler-módszer konzisztenciája

9. Def. Konzisztencia: Azt mondjuk, hogy egy numerikus séma konzisztens (a differenciálegyenlettel), ha a T_j^{k+1} képlethiba nullához tart, ha $\Delta t, \Delta x \rightarrow 0$; azaz pontosabban, ha vannak olyan c_1 és c_2 j -től és k -tól független konstansok, hogy

$$|T_j^{k+1}| \leq c_1 \Delta t^r + c_2 \Delta x^s \quad (= \mathcal{O}(\Delta t^r + \Delta x^s))$$

($r, s \geq 1$ konstansok). Ilyenkor azt mondjuk, hogy a séma konzisztenciarendje Δt -ben r és Δx -ben s .

A fenti konzisztenciafogalom tulajdonképpen maximumnormabeli konzisztenciát jelent.

Az EE-séma konzisztenciarendje Δt -ben 1, és Δx -ben 2.

Az Explicit Euler-módszer konzisztenciája

Megj.:

$$\begin{aligned}T_j^{k+1} &= \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(\xi, x_j) - \frac{\Delta x^2}{12} \frac{\partial^4 u}{\partial x^4}(t_k, \eta) + \mathcal{O}(\Delta t^2 + \Delta x^4) \\ &= \frac{\Delta t}{2} \frac{\partial^4 u}{\partial x^4}(t_k, x_j) \left(1 - \frac{1}{6q}\right) + \mathcal{O}(\Delta t^2 + \Delta x^4).\end{aligned}$$

$q = 1/6$ választással magasabbrendű módszert kapunk (szuperkonvergencia).

EE-módszer vektoriterációs alakja

A $q = \Delta t / \Delta x^2$ rácshányados bevezetésével az alábbi vektoriterációval állítható elő a numerikus megoldás:

$$\bar{\mathbf{u}}^{k+1} = \text{tridiag} [q, 1 - 2q, q] \bar{\mathbf{u}}^k =: \mathbf{A}_{\Delta x} \bar{\mathbf{u}}^k.$$

Mérjük a közelítő és a pontos megoldás hibáját! Tegyük fel, hogy a feladatot a $[0, t_{\max}]$ időintervallumon kell megoldanunk. Legyen

$$E_j^k = u_j^k - u(k\Delta t, j\Delta x)$$

a numerikus módszer hibája egy adott rácspontban és

$$\bar{\mathbf{E}}^k = [E_1^k, \dots, E_n^k]^T.$$

Az EE-módszer konvergenciája

Az alapelvárásunk az előállított numerikus eljárástól, hogy "tartson a pontos megoldáshoz".

10. Def. Konvergencia: Azt mondjuk, hogy a numerikus séma konvergens, ha az $\bar{\mathbf{E}}^k$ hibavektor nullához tart, ha $\Delta x, \Delta t \rightarrow 0$; azaz pontosabban, ha vannak olyan d_1, d_2 t -től független konstansok, hogy tetszőleges rögzített $0 < t < t_{\max}$ értékre és $k\Delta t = t$ mellett

$$\|\bar{\mathbf{E}}^k\| \leq d_1 \Delta t^r + d_2 \Delta x^s \quad (= \mathcal{O}(\Delta t^r + \Delta x^s)).$$

Ilyenkor azt mondjuk, hogy a séma konvergenciarendje Δt -ben (vagy időben) r és Δx -ben (térben) s .

Látni fogjuk, hogy a konzisztencia (azaz, hogy a differenciálegyenletet jól közelítjük) nem elég a konvergencia biztosításához (azaz a megoldás jó közelítéséhez).

EE-módszer konvergenciája

11. Tétel. Ha az egydimenziós hővezetési egyenlet megoldása kellően sima (t -szerinti második és x -szerinti negyedik derivált is folytonos), akkor a $q \leq 1/2$ feltétel mellett (azaz a Δt és a Δx úgy tart a nullához, hogy közben a $q \leq 1/2$ feltétel teljesül) az explicit Euler-módszer konvergens lesz (maximumnormában). A konvergenciarendje Δt -ben elsőrendű és Δx -ben másodrendű.

Biz.: A $q \leq 1/2$ feltétel miatt tetszőleges méretű $\mathbf{A}_{\Delta x}$ mátrixra igaz, hogy $\|\mathbf{A}_{\Delta x}\|_{\infty} \leq 1$. Legyen $\bar{\mathbf{x}} = [x_1, \dots, x_n]^T$ azon osztópontok oszlopvektora, amelyekben keressük a megoldás közelítését. A k . időréteg osztópontjaiban lévő pontos megoldásértékek oszlopvektorát jelöljük $u(k\Delta t, \bar{\mathbf{x}})$ -vel (Matlabos jelölés).

EE-módszer konvergenciája

A séma vektoros alakja miatt

$$\bar{\mathbf{u}}^{k+1} = \mathbf{A}_{\Delta x} \bar{\mathbf{u}}^k.$$

A konzisztencia miatt

$$u((k+1)\Delta t, \bar{\mathbf{x}}) = \mathbf{A}_{\Delta x} u(k\Delta t, \bar{\mathbf{x}}) + \Delta t \bar{\mathbf{T}}^{k+1},$$

ahol $\bar{\mathbf{T}}^{k+1} = [T_1^{k+1}, \dots, T_n^{k+1}]^T$.

Vonjuk ki egymásból a két egyenletet!

$$\bar{\mathbf{E}}^{k+1} = \mathbf{A}_{\Delta x} \bar{\mathbf{E}}^k - \Delta t \bar{\mathbf{T}}^{k+1},$$

majd vegyük mindkét oldal normáját!

$$\|\bar{\mathbf{E}}^{k+1}\|_{\infty} \leq \|\bar{\mathbf{E}}^k\|_{\infty} + \Delta t \|\bar{\mathbf{T}}^{k+1}\|_{\infty}.$$

EE-módszer konvergenciája

Így azt kapjuk, hogy

$$\begin{aligned}\|\bar{\mathbf{E}}^k\|_\infty &\leq \|\bar{\mathbf{E}}^0\|_\infty + \Delta t(\|\bar{\mathbf{T}}^1\|_\infty + \dots + \|\bar{\mathbf{T}}^k\|_\infty) \\ &\leq \|\bar{\mathbf{E}}^0\|_\infty + \Delta t(k(c_1\Delta t + c_2\Delta x^2)) \\ &= t(c_1\Delta t + c_2\Delta x^2) \leq t_{\max}(c_1\Delta t + c_2\Delta x^2),\end{aligned}$$

ahol kihasználtuk a konzisztenciarendre kapott kifejezést, továbbá azt, hogy a kezdeti időpontban $\|\bar{\mathbf{E}}^0\|_\infty = 0$ és hogy $k\Delta t = t < t_{\max}$ rögzített.

Ezzel igazoltuk az állítást. ■

EE-módszer konvergenciája

Megj.: Mi történik akkor, ha a $q > 1/2$ választást használjuk, hiszen csak egy elégséges feltételt igazoltunk a konvergenciára? Mutassuk meg, hogy a $q \leq 1/2$ feltétel szükséges is!

Ismert, hogy a

$$\text{tridiag}(a, b, c) = \begin{bmatrix} b & c & 0 & \dots & & \\ a & b & c & 0 & \dots & \\ & & \dots & & & \\ \dots & 0 & a & b & c & \\ & \dots & 0 & a & b & \end{bmatrix} \in \mathbb{R}^{n \times n}$$

mátrix sajátértékei

$$\lambda_j = b + 2c\sqrt{\frac{a}{c}} \cos\left(\frac{j\pi}{n+1}\right), \quad (j = 1, \dots, n)$$

alakúak.

EE-módszer konvergenciája

A sajátvektorokat a

$$(\bar{\mathbf{v}}_j)_k = \left(\sqrt{\frac{a}{c}} \right)^k \sin \left(\frac{kj\pi}{n+1} \right), \quad (j, k = 1, \dots, n)$$

képlet adja.

Így $a = c = q$ és $b = 1 - 2q$ esetén

$$\begin{aligned} \lambda_j &= 1 - 2q \left(1 - \cos \left(\frac{j\pi}{n+1} \right) \right) \\ &= 1 - 4q \sin^2 \left(\frac{j\pi}{2(n+1)} \right). \end{aligned}$$

EE-módszer konvergenciája

Sajátvektoroknak pedig jók a

$$\bar{\mathbf{v}}_j = \sin \left(\frac{\bar{k}j\pi}{n+1} \right), \quad (j = 1, \dots, n)$$

vektorok ($\bar{k} = [1, \dots, n]^T$).

Ez mutatja, hogy ha $q > 1/2$, akkor $\Delta x = 1/(n+1) \rightarrow 0$ esetén lesz 1-nél nagyobb abszolút értékű sajátértéke az $\mathbf{A}_{\Delta x}$ mátrixnak. Így a numerikus megoldás nem tarthat a pontos megoldáshoz. A gépi számábrázolás és kerekítés miatt még akkor sem igaz a konvergencia, ha egy sajátvektorról indítjuk az iterációt.

Azaz a $q \leq 1/2$ feltétel a konvergencia szükséges és elégséges feltétele.

Implicit Euler-módszer (IE)

Az IE-módszer sémájához úgy jutunk, hogy a hely szerinti deriváltat a $(k + 1)$. időrétegen közelítjük.

$$\frac{u_j^{k+1} - u_j^k}{\Delta t} = \frac{u_{j-1}^{k+1} - 2u_j^{k+1} + u_{j+1}^{k+1}}{\Delta x^2},$$

A séma mátrixos alakja:

$$\text{tridiag}[-q, 1 + 2q, -q] \bar{\mathbf{u}}^{k+1} = \bar{\mathbf{u}}^k.$$

Implicit Euler-módszer (IE)

A sémával való egy lépéshez egyenletrendszert kell megoldani. Az együtthatómátrix tridiagonális, azaz az inga-módszerrel (Thomas-algoritmus) megoldható $8n$ flop művelettel (az EE-módszer műveletigénye $5n$ flop).

A módszer Δt -ben elsőrendű, Δx -ben másodrendű és tetszőleges q mellett konvergens lesz. Az ilyen sémákat feltétel nélkül konvergens sémáknak hívjuk.

θ -módszer

Súlyozzuk egy $\theta \in [0, 1]$ súllyal az EE és IE sémákat az alábbi módon:

$$\frac{u_j^{k+1} - u_j^k}{\Delta t} = (1 - \theta) \frac{u_{j-1}^k - 2u_j^k + u_{j+1}^k}{\Delta x^2} + \theta \frac{u_{j-1}^{k+1} - 2u_j^{k+1} + u_{j+1}^{k+1}}{\Delta x^2}.$$

Mátrixos alak:

$$\begin{aligned} & \text{tridiag} [-\theta q, 1 + 2\theta q, -\theta q] \bar{\mathbf{u}}^{k+1} \\ & = \text{tridiag} [(1 - \theta)q, 1 - 2(1 - \theta)q, (1 - \theta)q] \bar{\mathbf{u}}^k. \end{aligned}$$

- ▶ $\theta = 0$: EE-módszer,
- ▶ $\theta = 1$: IE-módszer,
- ▶ $\theta = 1/2$: Crank-Nicolson-módszer.

Műveletszáma általában $13n$ flop.

θ -módszer

Ha $\theta \geq 1/2$, akkor a konvergenciához nincs feltétel q megválasztására, $\theta < 1/2$ esetén pedig a konvergencia feltétele

$$q \leq \frac{1}{2(1 - 2\theta)}.$$

A képlethiba a

$$\begin{aligned} T_j^{k+1} &= (1/2 - \theta)\Delta t u_{xxt}''' - (1/12)\Delta x^2 u_{xxxx}'''' \\ &+ (1/24)\Delta t^2 u_{ttt}''' - (1/8)\Delta t^2 u_{xxtt}'''' + \dots \end{aligned}$$

alakban írható. Így a Crank–Nicolson-módszer ($\theta = 1/2$) esetén időben (Δt) is másodrendű módszert kapunk.

Nemnegativitás megőrzés

Mikor biztosított az, hogy nemnegatív kezdőfüggvény esetén nemnegatív lesz a numerikus megoldás?

$$\begin{aligned} & \text{tridiag} [-\theta q, 1 + 2\theta q, -\theta q] \bar{\mathbf{u}}^{k+1} \\ &= \text{tridiag} [(1 - \theta)q, 1 - 2(1 - \theta)q, (1 - \theta)q] \bar{\mathbf{u}}^k. \end{aligned}$$

A bal oldalon egy M-mátrix áll, aminek inverze nemnegatív, a jobb oldali mátrix pedig nemnegatív, ha

$$q \leq \frac{1}{2(1 - \theta)}.$$

Igazolható, hogy ez a feltétel kell is, és hogy a maximum-elv is pontosan ekkor érvényes csak (a térbeli felosztásszámtól függetlenül).

$\theta = 1/2$ esetén $q \leq 1$ a feltétel!

Kezdetiérték-feladat az 1D hővezetési egyenletre

Kezdetiérték-feladat

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x \in \mathbb{R}$$
$$u(0, x) = u_0(x).$$

A diszkrétizáció hasonlóan történik, mint korábban. Választunk egy Δt időlépést, és egy Δx rácstávolságot, és elkészítjük a rácsot. Mivel most a térbeli osztópontok

$$\bar{x} = \{x_j = j\Delta x \mid j = \dots, -2, -1, 0, 1, 2, \dots\}$$

(mindkét irányban végtelen sorozat), ezért a t_k időrétegen vett közelítő értékek is $\bar{\mathbf{u}}^k = (\dots, u_{-2}^k, u_{-1}^k, u_0^k, u_1^k, u_2^k, \dots)$ végtelen sorozatot alkotnak (az egyszerűség kedvéért ezeket is vektorként jelöljük, és vektornak hívjuk majd).

Kezdetiérték-feladat az 1D hővezetési egyenletre

Az EE-módszer szerint az $\bar{\mathbf{u}}^{k+1}$ vektor az (1) sémából határozható meg a

$$u_j^{k+1} = qu_{j-1}^k + (1 - 2q)u_j^k + qu_{j+1}^k$$

képlettel. Vezessük be az

$$l_\infty := \{ \bar{\mathbf{z}} = (\dots, z_{-2}, z_{-1}, z_0, z_1, z_2, \dots) \mid \sup_{j \in \mathbb{Z}} |z_j| < \infty \}$$

teret! Ez normált tér a $\|\bar{\mathbf{z}}\|_\infty = \sup_{j \in \mathbb{Z}} |z_j|$ normával (maximumnorma). Ekkor a fenti leképezés egy A lineáris operátort definiál l_∞ -ből l_∞ -be.

A konzisztencia és a konvergencia hasonlóan definiálható mint a kezdeti- és peremértékfeladat esetén, csak most a hibavektor egy végtelen elemszámú vektor lesz.

Kezdetiérték-feladat az 1D hővezetési egyenletre

12. Tétel. Ha az egydimenziós hővezetési egyenletre felírt kezdetiértékfeladat megoldása kellően sima (t -szerinti második és x -szerinti negyedik derivált is folytonos és korlátos), akkor a $q \leq 1/2$ feltétel mellett (azaz a Δt és a Δx úgy tart a nullához, hogy közben a $q \leq 1/2$ feltétel teljesül) az explicit Euler-módszer konvergens lesz (maximumnormában). A konvergenciarendje Δt -ben elsőrendű és Δx -ben másodrendű.

Biz.: Ha $q \leq 1/2$, akkor $\|A\|_\infty \leq 1$, hiszen

$$\begin{aligned} |(A\bar{z})_j| &\leq q|z_{j-1}| + (1 - 2q)|z_j| + q|z_{j+1}| \\ &\leq q\|\bar{z}\|_\infty + (1 - 2q)\|\bar{z}\|_\infty + q\|\bar{z}\|_\infty = \|\bar{z}\|_\infty, \end{aligned}$$

és így

$$\|A\bar{z}\|_\infty \leq \|\bar{z}\|_\infty.$$

Kezdetiérték-feladat az 1D hővezetési egyenletre

Innét a bizonyítás hasonló a kezdeti- és peremértékfeladatéhoz.

A séma alakja miatt

$$\bar{\mathbf{u}}^{k+1} = A\bar{\mathbf{u}}^k.$$

A konzisztencia miatt

$$u((k+1)\Delta t, \bar{\mathbf{x}}) = Au(k\Delta t, \bar{\mathbf{x}}) + \Delta t \bar{\mathbf{T}}^{k+1},$$

ahol $\bar{\mathbf{T}}^{k+1} = [\dots, T_{-1}^{k+1}, T_0^{k+1}, T_1^{k+1}, \dots]^T$.

Vonjuk ki egymásból a két egyenletet!

$$\bar{\mathbf{E}}^{k+1} = A\bar{\mathbf{E}}^k - \Delta t \bar{\mathbf{T}}^{k+1},$$

majd vegyük mindkét oldal normáját!

$$\|\bar{\mathbf{E}}^{k+1}\|_{\infty} \leq \|\bar{\mathbf{E}}^k\|_{\infty} + \Delta t \|\bar{\mathbf{T}}^{k+1}\|_{\infty}.$$

Kezdetiérték-feladat az 1D hővezetési egyenletre

Így azt kapjuk, hogy

$$\begin{aligned}\|\bar{\mathbf{E}}^k\|_\infty &\leq \|\bar{\mathbf{E}}^0\|_\infty + \Delta t(\|\bar{\mathbf{T}}^1\|_\infty + \dots + \|\bar{\mathbf{T}}^k\|_\infty) \\ &\leq \|\bar{\mathbf{E}}^0\|_\infty + \Delta t(k(c_1\Delta t + c_2\Delta x^2)) \\ &= t(c_1\Delta t + c_2\Delta x^2) \leq t_{\max}(c_1\Delta t + c_2\Delta x^2),\end{aligned}$$

ahol kihasználtuk a konzisztenciarendre kapott kifejezést, továbbá azt, hogy a kezdeti időpontban $\|\bar{\mathbf{E}}^0\|_\infty = 0$ és hogy $k\Delta t = t < t_{\max}$ rögzített.

Ezzel igazoltuk az állítást. ■

6.

Lax-elmélet

Speciális normák

Legyen \bar{z} egy Δx osztásközű rácson definiált rácsfüggvény értékeit tartalmazó vektor (\mathbb{R}^n -beli vektor, vagy mindkét irányban végtelen sorozat).

Az alábbi normák fontos szerepet játszanak majd a konvergenciavizsgálatnál:

Maximumnorma (l_∞ -téren):

$$\|\bar{z}\|_\infty = \sup_j |z_j|,$$

Speciális normák

l_2 norma az

$$l_2 := \{\bar{\mathbf{z}} = (\dots, z_{-2}, z_{-1}, z_0, z_1, z_2, \dots) \mid \sum_{j \in \mathbb{Z}} |z_j|^2 < \infty\}$$

téren:

$$\|\bar{\mathbf{z}}\|_2 = \sqrt{\sum_j |z_j|^2}, \quad \|\bar{\mathbf{z}}\|_{2, \Delta x} = \sqrt{\Delta x \cdot \sum_j |z_j|^2}.$$

Ezek könnyen általánosíthatók magasabb dimenzióra is.

Konvergencia és konzisztencia

Legyen a megoldandó differenciálegyenlet

$$Lu = f$$

alakú (+ kezdeti- és esetleg peremfeltételek), ahol L lineáris differenciáloperátor (t -szerint első deriválttal), f pedig adott függvény.

Egy adott rácssorozaton a feladat numerikus megoldására konstruált egylépéses véges differencia séma felírható az

$$\bar{\mathbf{u}}^{k+1} = Q\bar{\mathbf{u}}^k + \Delta t \bar{\mathbf{f}}^k \quad (2)$$

alakban (peremfeltétellel együtt, ha vannak ilyenek), ahol $\bar{\mathbf{u}}^k$ a megoldás közelítése a rácspontokban a t_k időrétegen, $\bar{\mathbf{f}}^k$ adott vektor minden időrétegen, és Q megfelelő lineáris operátor.

Konvergencia és konzisztencia

Legyen

$$E_j^k = u_j^k - u(k\Delta t, j\Delta x),$$

továbbá $\bar{\mathbf{E}}^k$ az E_j^k értékek vektora.

A séma legfontosabb tulajdonsága, hogy a vele generált rácsfüggvény jól közelíti-e a feladat pontos megoldását finomodó rácsfelosztás esetén. Ezt a tulajdonságot fogalmazzuk meg az alábbi definíció.

Konvergencia és konzisztencia

13. Def. Konvergencia: Azt mondjuk, hogy egy numerikus séma konvergens egy adott $\|\cdot\|$ normában, ha $\bar{\mathbf{E}}^k$ nullához tart ebben a normában $\Delta x, \Delta t \rightarrow 0$ esetén; azaz pontosabban, ha vannak olyan d_1, d_2 t -től független konstansok, hogy tetszőleges rögzített $0 < t < t_{\max}$ értékre és $k\Delta t = t$ mellett

$$\|\bar{\mathbf{E}}^k\| \leq d_1 \Delta t^r + d_2 \Delta x^s \quad (= \mathcal{O}(\Delta t^r + \Delta x^s)).$$

Ilyenkor azt mondjuk, hogy a séma konvergenciarendje Δt -ben (vagy időben) r és Δx -ben (térben) s .

Megj.: Fontos észrevenni, hogy ha $\bar{\mathbf{E}}^k$ elemszáma véges, akkor $\Delta x \rightarrow 0$ esetén $\bar{\mathbf{E}}^k$ mindig más-más térben van, így maga a norma is más-más téren van értelmezve.

Konvergencia és konzisztencia

A hővezetési egyenletnél segített a konvergencia bizonyításában a konzisztencia tulajdonság. Ezt most adott normában fogalmazzuk meg.

14. Def. Konzisztencia: Azt mondjuk, hogy egy numerikus séma konzisztens (a differenciálegyenlettel) egy adott $\|\cdot\|$ normában, ha a képlethibák vektorának normája nullához tart, ha $\Delta t, \Delta x \rightarrow 0$; azaz pontosabban, ha vannak olyan c_1 és c_2 k -tól független konstansok, hogy

$$\|\bar{\mathbf{T}}^{k+1}\| \leq c_1 \Delta t^r + c_2 \Delta x^s \quad (= \mathcal{O}(\Delta t^r + \Delta x^s))$$

($r, s \geq 1$ konstansok). Itt $\bar{\mathbf{T}}^{k+1}$ a $(k+1)$ -edik lépéshez tartozó képlethibák vektora. Ilyenkor azt mondjuk, hogy a séma konzisztenciarendje Δt -ben r és Δx -ben s .

Stabilitás

Láttuk, hogy maga a konzisztencia (az egyenlet jó közelítése) nem elég a konvergenciához (a megoldás jó közelítéséhez). Kellene egy másik könnyen ellenőrizhető feltétel, ami a konzisztenciával együtt már biztosítja majd a konvergenciát!

15. Def. Stabilitás (zéró-stabilitás exponenciális korláttal - exponenciális stabilitás): Azt mondjuk, hogy a (2) séma stabil, ha a Q operátor hatványainak normája az alábbi módon becsülhető (az $\bar{\mathbf{u}}^k$ vektor normájához tartozó indukált operátornormában): vannak olyan $K \geq 0$, $\beta \geq 0$, $\Delta t_0 > 0$ és $\Delta x_0 > 0$ számok, hogy minden $\Delta x < \Delta x_0$, $\Delta t < \Delta t_0$ rácsméretre és $t = k\Delta t$ esetén

$$\|Q^k\| \leq Ke^{\beta t}.$$

Megj.:

$$\|Q^k\| = \sup_{\bar{\mathbf{z}} \neq 0} \frac{\|Q^k \bar{\mathbf{z}}\|}{\|\bar{\mathbf{z}}\|}.$$

Lax-tétel

16. Tétel. (Lax-féle ekvivalencia-tétel) Az $Lu = f$ korrekt kitűzésű feladatra felírt (2) egylépéses konzisztens séma pontosan akkor konvergens, ha stabil.

17. Tétel. (Lax-tétel) Ha az $Lu = f$ korrekt kitűzésű feladatra felírt (2) egylépéses séma konzisztens és konzisztenciarendje Δt -ben r és Δx -ben s , valamint a séma stabil, akkor a séma konvergens és konvergenciarendje Δt -ben r és Δx -ben s (itt minden egy adott normában (ill. normasorozatban) értendő).



Biz.: Helyettesítsük a (2) sémába a pontos megoldást. A két oldal közti különbség a lokális hibát adja.

$$u(t_{k+1}, \bar{\mathbf{x}}) = Qu(t_k, \bar{\mathbf{x}}) + \Delta t \bar{\mathbf{f}}^k + \Delta t \bar{\mathbf{T}}^{k+1}.$$

Vonjuk ki ezt a (2) egyenletből:

$$\bar{\mathbf{E}}^{k+1} = Q\bar{\mathbf{E}}^k - \Delta t \bar{\mathbf{T}}^{k+1}.$$

Ebből, feltéve hogy $\|\bar{\mathbf{E}}^0\| = 0$, kapjuk hogy

$$\bar{\mathbf{E}}^k = -\Delta t (\bar{\mathbf{T}}^k + Q\bar{\mathbf{T}}^{k-1} + \dots + Q^{k-1}\bar{\mathbf{T}}^1),$$

majd becsüljük az adott normában a két oldalt.

$$\begin{aligned}\|\bar{\mathbf{E}}^k\| &\leq \Delta t(\|\bar{\mathbf{T}}^k\| + Ke^{\beta\Delta t}\|\bar{\mathbf{T}}^{k-1}\| + \dots + Ke^{\beta(k-1)\Delta t}\|\bar{\mathbf{T}}^1\|) \\ &= \Delta tkKe^{\beta k\Delta t}\mathcal{O}(\Delta t^r + \Delta x^s) = tKe^{\beta t}\mathcal{O}(\Delta t^r + \Delta x^s) \\ &\leq t_{\max}Ke^{\beta t_{\max}}\mathcal{O}(\Delta t^r + \Delta x^s)\end{aligned}$$

Ez pont a tétel állítását adja. ■

A konzisztencia megmutatható egyszerű sorfejtéssel elegendően sima megoldás esetén. Milyen módszerek vannak a stabilitás megmutatására?

Stabilitásvizsg. diszkrét Fourier-transzformációval

Kezdetiérték-feladatok esete:

$\bar{z} \in l_2$ -re definiáljuk az alábbi $L^2[-\pi, \pi]$ -beli függvényt:

$$\hat{z}(\xi) := \mathcal{F}(\bar{z}) := \frac{1}{\sqrt{2\pi}} \sum_{j=-\infty}^{\infty} e^{-ij\xi} z_j, \quad \xi \in [-\pi, \pi].$$

18. Tétel. (A transzformáció inverze)

$$z_j = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{ij\xi} \hat{z}(\xi) d\xi.$$

19. Tétel. A transzformáció lineáris, és normatartó:

$\|\hat{z}\|_2 = \|\bar{z}\|_2$. (Itt a bal oldali norma az L^2 -tér normája, a jobb oldali pedig az l_2 -téré. Parseval-egyenlőség.)

Stabilitásvizsg. diszkrét Fourier-transzformációval

Biz.: (formálisan)

A linearitás triviális.

$$\begin{aligned}\|\hat{z}\|_2^2 &= \int_{-\pi}^{\pi} |\hat{z}(\xi)|^2 d\xi = \int_{-\pi}^{\pi} \overline{\hat{z}(\xi)} \frac{1}{\sqrt{2\pi}} \sum_{j=-\infty}^{\infty} e^{-ij\xi} z_j d\xi \\ &= \frac{1}{\sqrt{2\pi}} \sum_{j=-\infty}^{\infty} z_j \int_{-\pi}^{\pi} \overline{\hat{z}(\xi)} e^{-ij\xi} d\xi \\ &= \sum_{j=-\infty}^{\infty} z_j \overline{\frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \hat{z}(\xi) e^{ij\xi} d\xi} = \sum_{j=-\infty}^{\infty} z_j \overline{z_j} = \|\bar{z}\|_2^2. \blacksquare\end{aligned}$$

Stabilitásvizsg. diszkrét Fourier-transzformációval

A sémák vizsgálatánál nagy segítségünkre lesz az alábbi tétel, amely a diszkrét Fourier-transzformáció definíciójának segítségével könnyen igazolható.

20. Tétel. Tekintsük az S_+ ill. S_- balra- ill. jobbra tolás operátorokat ($l_2 \rightarrow l_2$), melyek az alábbi módon vannak definiálva!

$$(S_+(\bar{\mathbf{z}}))_k = z_{k+1}, \quad (S_-(\bar{\mathbf{z}}))_k = z_{k-1}$$

Ekkor

$$\mathcal{F}(S_{\pm}(\bar{\mathbf{z}})) = e^{\pm i\xi} \mathcal{F}(\bar{\mathbf{z}}).$$

Stabilitásvizsg. diszkrét Fourier-transzformációval

Hogyan használhatók a fentiek a stabilitás vizsgálatára?

Legyen $\bar{z} \in l_2$ tetszőleges vektor, továbbá $\bar{w}^1 = Q\bar{z}$. Ekkor a diszkrét Fourier-transzformálás során kikerül a képletből a helykoordináta indexétől való függés:

$$\hat{w}^1(\xi) = g(\xi)\hat{z}(\xi),$$

ahol a $g(\xi)$ szorzót növekedési faktornak hívjuk (growth factor). $g(\xi)$ folytonos függvény. Emiatt a $\bar{w}^k = Q^k\bar{z}$ jelöléssel érvényes az

$$\hat{w}^k(\xi) = g^k(\xi)\hat{z}(\xi)$$

egyenlőség.

Megj.: A növekedési faktort általában a $\bar{z} = \bar{u}^k$ választással szoktuk a séma képletéből meghatározni.

Stabilitásvizsg. diszkrét Fourier-transzformációval

21. Def. Von Neumann-feltétel a (2) sémára: Vannak olyan $\Delta t_0, \Delta x_0 > 0$ és $C \geq 0$ számok, hogy minden $\Delta t < \Delta t_0$, $\Delta x < \Delta x_0$, $\xi \in [-\pi, \pi]$ estén

$$|g(\xi)| \leq 1 + C\Delta t.$$

22. Tétel. A (2) séma pontosan akkor stabil $\|\cdot\|_{2,\Delta x}$ normában, ha teljesíti a von Neumann-feltételt.

Biz.: (\Leftarrow)

$$\begin{aligned} \|Q^k\|_{2,\Delta x} &= \sup_{\bar{\mathbf{z}} \neq 0} \frac{\|Q^k \bar{\mathbf{z}}\|_{2,\Delta x}}{\|\bar{\mathbf{z}}\|_{2,\Delta x}} = \sup_{\bar{\mathbf{z}} \neq 0} \frac{\|\hat{w}^k\|_2}{\|\hat{z}\|_2} \\ &= \sup_{\bar{\mathbf{z}} \neq 0} \frac{\sqrt{\int_{-\pi}^{\pi} |g^k(\xi)|^2 \cdot |\hat{z}(\xi)|^2 d\xi}}{\sqrt{\int_{-\pi}^{\pi} |\hat{z}(\xi)|^2 d\xi}} \leq \dots \end{aligned}$$

Stabilitásvizsg. diszkrét Fourier-transzformációval

$$\begin{aligned} &\leq \sup_{\bar{z} \neq 0} \frac{\sqrt{\int_{-\pi}^{\pi} (1 + C\Delta t)^{2k} \cdot |\hat{z}(\xi)|^2 d\xi}}{\sqrt{\int_{-\pi}^{\pi} |\hat{z}(\xi)|^2 d\xi}} \\ &= (1 + C\Delta t)^k \leq e^{C\Delta tk} = e^{Ct}, \end{aligned}$$

ami $K = 1$ és $\beta = C$ választással mutatja a stabilitást.

(\Rightarrow) Indirekt tegyük fel, hogy minden $C \geq 0$ esetén vannak olyan Δx_C (tetsz. kicsi), Δt_C (tetsz. kicsi) és ξ_C (C -től függő) számok, melyekkel

$$|g(\xi_C)| > 1 + C\Delta t_C.$$

A $g(\xi)$ függvény folytonossága miatt a fenti egyenlőség nem csak ξ_C -vel igaz, hanem ξ_C egy megfelelően kicsi $U(\xi_C)$ környezetében is. Speciálisan legyen $\bar{z} \in l_2$ olyan, hogy $\hat{z}(\xi) = 0$, ha $\xi \notin U(\xi_C)$.

Stabilitásvizsg. diszkrét Fourier-transzformációval

Ekkor ezzel a $\bar{\mathbf{z}}$ -vel

$$\begin{aligned} \frac{\|Q^k \bar{\mathbf{z}}\|_{2, \Delta x}}{\|\bar{\mathbf{z}}\|_{2, \Delta x}} &= \frac{\|\hat{w}^k\|_2}{\|\hat{z}\|_2} \\ &= \frac{\sqrt{\int_{-\pi}^{\pi} |g^k(\xi)|^2 \cdot |\hat{z}(\xi)|^2 d\xi}}{\sqrt{\int_{-\pi}^{\pi} |\hat{z}(\xi)|^2 d\xi}} \\ &= \frac{\sqrt{\int_{U(\xi_C)} |g^k(\xi)|^2 \cdot |\hat{z}(\xi)|^2 d\xi}}{\sqrt{\int_{U(\xi_C)} |\hat{z}(\xi)|^2 d\xi}} \\ &> (1 + C\Delta t_C)^k. \end{aligned}$$

Mivel ez minden rögzített C esetén tetszőlegesen kicsi Δt_C -re is igaz, így elég nagy k -ra

$$\frac{\|Q^k \bar{\mathbf{z}}\|_{2, \Delta x}}{\|\bar{\mathbf{z}}\|_{2, \Delta x}} > \left(1 + \frac{Ct}{k}\right)^k \rightarrow e^{Ct} \quad (k \rightarrow \infty).$$

Stabilitásvizsg. diszkrét Fourier-transzformációval

Így tetszőlegesen nagy C értékek esetén elég nagy k -ra igaz pl. $\|Q^k\|_{2,\Delta x} > e^{Ct/2}$. Így a $\|Q^k\|_{2,\Delta x} \leq Ke^{\beta t}$ feltétel nem teljesülhet minden k -ra, azaz a séma így nem lehetne stabil. Ez ellentmondás.



Megj.: A von Neumann-feltétel egy elégséges feltétele: $|g(\xi)| \leq 1$.

Megj.: A von Neumann-feltétel és a konzisztencia együtt a Lax-tétel miatt már garantálja a konvergenciát. Fontos viszont, hogy mivel a von Neumann-feltétel $l_{2,\Delta x}$ -normabeli konvergenciát garantál, így a konzisztenciát is ebben a normában kell igazolni.

Stabilitásvizsg. diszkrét Fourier-transzformációval

Az 1D hővezetési egyenletre felírt kezdetiérték-feladatok növekedési faktorai.

EE:

$$g(\xi) = 1 - 4q \sin^2(\xi/2)$$

Csak $q \leq 1/2$ esetén teljesíti a von Neumann-feltételt.

IE:

$$g(\xi) = \frac{1}{1 + 4q \sin^2(\xi/2)}$$

Minden q -ra teljesíti a von Neumann-feltételt.

Stabilitásvizsg. diszkrét Fourier-transzformációval

Kezdeti- és peremérték-feladatok esete:

23. Tétel. (Szükséges feltétel) Egy séma csak úgy lehet stabil egy kezdeti- és peremérték-feladatra, ha stabil a feladathoz tartozó kezdetiérték-feladatra is, azaz ha erre teljesíti a von Neumann-feltételt.

Emlékeztetőül: A stabilitáshoz a

$$\|Q^k\| \leq Ke^{\beta t}$$

feltételt kellene garantálni, kezdeti- és peremérték-feladatok esetén mátrixokra.

Ismert, hogy indukált mátrixnormában tetszőleges négyzetes mátrixra igaz, hogy $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$, ahol $\rho(\mathbf{A})$ a mátrix spektrálsugara (maximális abszolút értékű sajátérték abszolút értéke).

Stabilitásvizsg. spektrálsugárral

24. Tétel. Legyen a korábbi jelöléssel a kezdeti- és peremérték-feladatra felírt séma iterációs mátrixa \mathbf{Q} . Ekkor a $\|\cdot\|_{2,\Delta x}$ normabeli stabilitáshoz szükséges, hogy valamilyen $C \geq 0$ konstanssal teljesüljön a

$$\varrho(\mathbf{Q}) \leq 1 + C\Delta t$$

feltétel.

Biz. Indirekt, tegyük fel, hogy minden C -hez vannak tetszőlegesen kicsi Δt_C és Δx_C értékek, melyekkel $\varrho(\mathbf{Q}) > 1 + C\Delta t_C$. Ekkor

$$e^{Ct/2} \leq \underbrace{\left(1 + C\frac{t}{k}\right)^k}_{\rightarrow e^{tC}, k \rightarrow \infty} = (1 + C\Delta t_C)^k < (\varrho(\mathbf{Q}))^k \leq \varrho(\mathbf{Q}^k) \leq \|\mathbf{Q}^k\|_{2,\Delta x}$$

miatt nem lehet stabil a séma. ■

Stabilitásvizsg. spektrálsugárral

Az előző feltétel csak szükséges volt a stabilitáshoz. Most belátjuk, hogy szimmetrikus mátrixokra elégséges is.

25. Tétel. Legyen a korábbi jelöléssel a kezdeti- és peremérték-feladatra felírt séma *szimmetrikus* iterációs mátrixa \mathbf{Q} . Ekkor a $\|\cdot\|_{2,\Delta x}$ normabeli stabilitáshoz szükséges és elégséges, hogy valamilyen $C \geq 0$ konstanssal teljesüljön a

$$\varrho(\mathbf{Q}) \leq 1 + C\Delta t$$

feltétel.

Biz. A szükségességet az előző tételben láttuk. Az elégségesség következik az alábbi becslésből:

$$\|\mathbf{Q}^k\|_{2,\Delta x} = \varrho(\mathbf{Q}^k) = (\varrho(\mathbf{Q}))^k \leq (1 + C\Delta t)^k \leq e^{C\Delta tk} = e^{Ct}. \blacksquare$$

Stabilitásvizsg. spektrálsugárral

26. Tétel. Legyen a korábbi jelöléssel a kezdeti- és peremérték-feladatra felírt séma iterációs mátrixa \mathbf{Q} , és \mathbf{Q} legyen *hasonló egy szimmetrikus mátrixhoz* (a transzformáló mátrix és inverze is legyen egyenletesen korlátos). Ekkor a $\|\cdot\|_{2,\Delta x}$ normabeli stabilitáshoz elég (és szükséges is), hogy valamilyen $C \geq 0$ konstanssal teljesüljön a

$$\varrho(\mathbf{Q}) \leq 1 + C\Delta t$$

feltétel.

Biz. Legyen $\mathbf{V}^{-1}\mathbf{Q}\mathbf{V} = \mathbf{S}$, ahol \mathbf{S} szimmetrikus és \mathbf{V} a hasonlósági transzformáció mátrixa. Ekkor

$$\begin{aligned} \|\mathbf{Q}^k\|_{2,\Delta x} &= \|(\mathbf{V}\mathbf{S}\mathbf{V}^{-1})^k\|_{2,\Delta x} = \|\mathbf{V}\mathbf{S}^k\mathbf{V}^{-1}\|_{2,\Delta x} \\ &\leq \|\mathbf{V}\|_{2,\Delta x} \|\mathbf{S}^k\|_{2,\Delta x} \|\mathbf{V}^{-1}\|_{2,\Delta x} \\ &= \|\mathbf{V}\|_{2,\Delta x} \|\mathbf{V}^{-1}\|_{2,\Delta x} \underbrace{(\varrho(\mathbf{S}))^k}_{\varrho(\mathbf{Q})} \leq \|\mathbf{V}\|_{2,\Delta x} \|\mathbf{V}^{-1}\|_{2,\Delta x} e^{Ct}. \blacksquare \end{aligned}$$

Stabilitásvizsg. spektrálsugárral

Megj.: A fentiek alapján a stabilitáshoz nyilván elegendő a $\|\mathbf{Q}\|_{2,\Delta x} \leq 1$ feltétel, vagy maximumnormában a $\|\mathbf{Q}\|_{\infty} \leq 1$ feltétel.

27. Tétel. Ha $\|\mathbf{Q}\|_{\infty} \leq 1$ és $\|\mathbf{Q}\|_1 \leq 1$, akkor $\|\mathbf{Q}\|_{2,\Delta x} \leq 1$.

Biz.:

$$\begin{aligned}\|\mathbf{Q}\|_{2,\Delta x}^2 &= \varrho(\mathbf{Q}^T \mathbf{Q}) \\ &\leq \|\mathbf{Q}^T \mathbf{Q}\|_{\infty} \leq \|\mathbf{Q}^T\|_{\infty} \|\mathbf{Q}\|_{\infty} \\ &= \|\mathbf{Q}\|_1 \|\mathbf{Q}\|_{\infty},\end{aligned}$$

amiből már következik az állítás. ■

Pl., Neumann-perem másodrendű közelítése

EE-módszer esetén az iterációs mátrix alakja

$$\mathbf{Q} = \begin{bmatrix} 1 - 2q & 2q & 0 & \dots & & \\ q & 1 - 2q & q & 0 & \dots & \\ & & \dots & & & \\ \dots & 0 & q & 1 - 2q & q & \\ & \dots & 0 & q & 1 - 2q & \end{bmatrix}.$$

$$\lambda_j = 1 - 4q \sin^2 \left(\frac{(2j-1)\pi}{2n} \right), \quad j = 1, \dots, n.$$

Így a $q \leq 1/2$ feltétel szükséges az $\|\cdot\|_{2,\Delta x}$ -normabeli stabilitáshoz (ez a kezdetiérték-feladatra kapott eredményből is látszik).

A $q \leq 1/2$ feltétel elégségességét két módon is igazolhatjuk: 1)

$\|\mathbf{Q}\|_{\infty} = 1$ és $\|\mathbf{Q}\|_1 = 1$, emiatt $\|\mathbf{Q}\|_{2,\Delta x} \leq 1$ (azaz a séma maximum- és $\|\cdot\|_{2,\Delta x}$ normában is stabil).

2) A $\mathbf{V} = \text{diag}(\sqrt{2}, 1, \dots, 1)$ mátrixszal szimmetrizálható \mathbf{Q} .

Megjegyzés az implicit sémákról

Implicit sémák alakja:

$$Q_1 \bar{\mathbf{u}}^{k+1} = Q_2 \bar{\mathbf{u}}^k + \Delta t \bar{\mathbf{g}}^k.$$

A pontos megoldást helyettesítve kapjuk a $\bar{\mathbf{T}}^{k+1}$ képlethibát

$$Q_1 u(t_{k+1}, \bar{\mathbf{x}}) = Q_2 u(t_k, \bar{\mathbf{x}}) + \Delta t \bar{\mathbf{g}}^k + \Delta t \bar{\mathbf{T}}^{k+1}.$$

Q_1 invertálható, mert különben nem lenne kifejezhető $\bar{\mathbf{u}}^{k+1}$. Így

$$u(t_{k+1}, \bar{\mathbf{x}}) = Q_1^{-1} Q_2 u(t_k, \bar{\mathbf{x}}) + \Delta t Q_1^{-1} \bar{\mathbf{g}}^k + \Delta t Q_1^{-1} \bar{\mathbf{T}}^{k+1},$$

azaz a lokális hiba $\Delta t Q_1^{-1} \bar{\mathbf{T}}^{k+1}$ alakú, a séma explicit módon pedig a korábban vizsgált alakban írható:

$$\bar{\mathbf{u}}^{k+1} = \underbrace{Q_1^{-1} Q_2}_{Q} \bar{\mathbf{u}}^k + \Delta t \underbrace{Q_1^{-1} \bar{\mathbf{g}}^k}_{\bar{\mathbf{f}}^k}.$$

Megjegyzés az implicit sémákról

Emiatt a korábbi tételek használhatók implicit sémákra is, azzal a megkötéssel, hogy a $Q_1^{-1}\bar{\mathbf{T}}^{k+1}$ vektornak $\mathcal{O}(\Delta t^r + \Delta x^s)$ rendűnek kell lennie. Ez úgy garantálható, ha a $\|Q_1^{-1}\|$ norma egyenletesen korlátos, azaz van olyan K konstans, hogy minden elég kicsi Δt és Δx értékre $\|Q_1^{-1}\| \leq K$.

Pl.: 1D hővezetés, kezdeti- és peremérték-feladat, IE-módszer. M-mátrixos technikával igazoltuk, hogy $\|Q_1^{-1}\|_\infty \leq 1$. Ez a konzisztenciával és a nyilvánvaló stabilitással együtt garantálja a konvergenciát minden q rácshányadosra.

Megjegyzés az implicit sémákról

PI.: 1D, hővezetés, kezdetiérték-feladat, IE, maximumnormabeli konvergencia. A séma alakja:

$$(1 + 2q)u_j^{k+1} - qu_{j-1}^{k+1} - qu_{j+1}^{k+1} = u_j^k,$$

ami egy Q_1 lineáris operátorral a $Q_1 \bar{\mathbf{u}}^{k+1} = \bar{\mathbf{u}}^k$ alakban írható. A balra- és jobbra tolás operátorokkal tehát egy tetszőleges $\bar{\mathbf{z}}$ vektorra $Q_1 \bar{\mathbf{z}} = (1 + 2q)\bar{\mathbf{z}} - qS_- \bar{\mathbf{z}} - qS_+ \bar{\mathbf{z}}$. Így

$$\begin{aligned} \|Q_1 \bar{\mathbf{z}}\|_\infty &\geq |(1 + 2q)\|\bar{\mathbf{z}}\|_\infty - q\|S_- \bar{\mathbf{z}} + S_+ \bar{\mathbf{z}}\|_\infty| \\ &\geq |(1 + 2q)\|\bar{\mathbf{z}}\|_\infty - 2q\|\bar{\mathbf{z}}\|_\infty| = \|\bar{\mathbf{z}}\|_\infty, \end{aligned}$$

azaz Q_1 invertálható és $\|Q_1^{-1}\|_\infty \leq 1$. Azaz a séma stabil, és $\|Q_1^{-1}\|_\infty$ egyenletes korlátossága miatt konvergens is.

Megjegyzés az implicit sémákról

Pl.: Ugyanaz, mint az előző, csak $\|\cdot\|_{2,\Delta x}$ normában. Vegyük

$\bar{\mathbf{w}} := Q_1 \bar{\mathbf{z}} = (1 + 2q)\bar{\mathbf{z}} - qS_- \bar{\mathbf{z}} - qS_+ \bar{\mathbf{z}}$ mindkét oldalának diszkrét Fourier-transzformáltját!

$$(1 + 2q)\hat{z}(\xi) - qe^{i\xi}\hat{z}(\xi) - qe^{-i\xi}\hat{z}(\xi) = \hat{w}(\xi),$$

azaz

$$\hat{z}(\xi) = \frac{1}{1 + 4q \sin^2(\xi/2)} \hat{w}(\xi).$$

Mivel a

$$g(\xi) = \frac{1}{1 + 4q \sin^2(\xi/2)}$$

növekedési faktor kisebb 1-nél minden q -ra, és $\bar{\mathbf{w}}$ választható tetszőlegesen, így

$$\frac{\|Q_1 \bar{\mathbf{w}}\|_{2,\Delta x}}{\|\bar{\mathbf{w}}\|_{2,\Delta x}} = \frac{\|\hat{z}\|_2}{\|\hat{w}\|_2} \leq 1,$$

azaz $\|Q_1\|_{2,\Delta x} \leq 1$, ami elég a konvergenciához, ha a konzisztenciát is $\|\cdot\|_{2,\Delta x}$ normában tudjuk igazolni.

Megjegyzés a kétlépéses sémákról

Kétlépéses sémák diszkrét Fourier-transzformáció után felírhatók az alábbi alakban

$$\begin{bmatrix} \hat{u}^{k+1}(\xi) \\ \hat{u}^k(\xi) \end{bmatrix} = G(\xi) \begin{bmatrix} \hat{u}^k(\xi) \\ \hat{u}^{k-1}(\xi) \end{bmatrix},$$

ahol $G(\xi)$ az ún. növekedési mátrix. A növekedési mátrixra vonatkozó feltételekkel garantálható a módszer stabilitása.

Az $\|\cdot\|_{2,\Delta x}$ stabilitáshoz szükséges, hogy valamely $C \geq 0$ számra, minden $\xi \in [-\pi, \pi]$ értékre, és minden elég kicsi Δt és Δx lépéstávolság esetén

$$\rho(G(\xi)) \leq 1 + C\Delta t$$

teljesüljön. Ha $G(\xi)$ normális (pl. hermitikus), vagy minden sajátértéke az egységkörbe esik úgy, hogy maximum csak az egyik esik a körvonalra, akkor a feltétel elégséges is.

Megjegyzés a kétlépéses sémákról

Pl.: Az 1D hővezetési egyenletre a CTCS séma nem stabil.

$$G(\xi) = \begin{bmatrix} -8q \sin^2(\xi/2) & 1 \\ 1 & 0 \end{bmatrix}$$

Itt pl. $\xi = \pi$ esetén

$$\lambda_{\pm} = -4q \pm \sqrt{16q^2 + 1},$$

és $\lambda_- < -1$, azaz a szükséges feltétel nem teljesül.

Megjegyzés a kétlépéses sémákról

Pl.: Dufort-Frankel-séma

$$u_j^{k+1} = \frac{2q}{1+2q}(u_{j+1}^k + u_{j-1}^k) + \frac{1-2q}{1+2q}u_j^{k-1}$$

feltétel nélkül stabil.

$$G(\xi) = \begin{bmatrix} \frac{4q}{1+2q} \cos \xi & \frac{1-2q}{1+2q} \\ 1 & 0 \end{bmatrix}$$

Itt

$$\lambda_{\pm} = \frac{2q \cos \xi \pm \sqrt{1 - 4q^2 \sin^2 \xi}}{1 + 2q},$$

amik kielégítik a stabilitás elégséges feltételét.

7.

Az advekciós egyenlet megoldása (véges differencia)

Elsőrendű hiperbolikus egyenletek

Elsőrendű hiperbolikus differenciálegyenlet-rendszer

$$\bar{\mathbf{v}}'_t + \mathbf{A}\bar{\mathbf{v}}'_x = \bar{\mathbf{g}},$$

ahol az \mathbf{A} négyzetes mátrix diagonalizálható, és

$$\bar{\mathbf{v}} = \bar{\mathbf{v}}(t, x) = [v_1(t, x), \dots, v_n(t, x)]^T.$$

A diagonalizálhatóság miatt valamilyen \mathbf{V} invertálható és \mathbf{D} diagonális mátrixszal $\mathbf{A} = \mathbf{V}^{-1}\mathbf{D}\mathbf{V}$. Legyen $\bar{\mathbf{u}} = \mathbf{V}\bar{\mathbf{v}}$, amivel az egyenlet a \mathbf{V} -vel való balról szorzás után

$$\mathbf{V}\bar{\mathbf{v}}'_t + \mathbf{D}\mathbf{V}\bar{\mathbf{v}}'_x = \mathbf{V}\bar{\mathbf{g}},$$

majd

$$\bar{\mathbf{u}}'_t + \mathbf{D}\bar{\mathbf{u}}'_x = \bar{\mathbf{f}},$$

alakban írható, ahol $\bar{\mathbf{f}} = \mathbf{V}\bar{\mathbf{g}}$.

Így a feladat megoldható n darab skalár egyenlet megoldásával. Elég ezeket vizsgálni csak.

Az advekcións egyenlet 1D-ban

Advekcións egyenlet

$$u'_t(t, x) + cu'_x(t, x) = f(t, x),$$

kiegészítve a megfelelő kezdeti- és peremfeltételekkel; c az áramlás sebessége.

Egzakt megoldása a karakterisztikus görbék segítségével történik. A megoldás grafikonja összerakható a

$$\frac{d}{ds} \begin{bmatrix} t(s) \\ x(s) \\ u(t(s), x(s)) \end{bmatrix} = \begin{bmatrix} 1 \\ c \\ f(t(s), x(s)) \end{bmatrix}$$

KDER megoldásgörbéiből $([t(0), x(0), u(t(0), x(0))]^T$ adott vektor).

Karakterisztikák

Legyen $u(0, x) = u_0(x)$ adott kezdeti függvény.

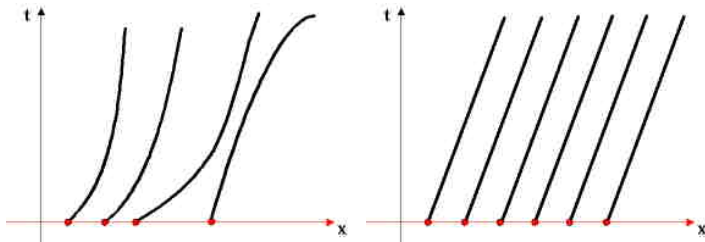
Ha $c = c(t, x)$ x -ben Lipschitz-folytonos és t -ben folytonos, akkor a karakterisztikák nem metszik egymást.

Ha $f \equiv 0$, akkor u konstans a karakterisztikákon. Ekkor, ha

$c = c(u)$, akkor egyenesek a karakterisztikák

($u(t, x) = u_0(x - c(u(t, x))t$), c konstans esetén párhuzamosak

($u(t, x) = u_0(x - ct)$).



Első próbálkozás

Vizsgáljuk az alábbi feladatot:

Konstans sebességű advekción, periodikus peremmel

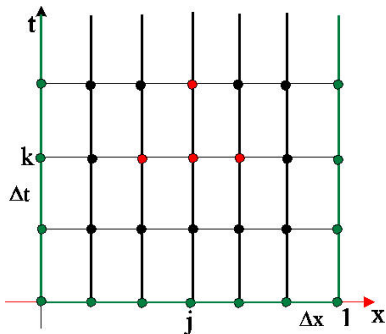
$$u'_t(t, x) + cu'_x(t, x) = 0, \quad u_0(x) \text{ adott, } u(t, 0) = u(t, 1). \quad (3)$$

A pontos megoldás ebben az esetben $u(t, x) = u_0(x - ct)$, azaz a kezdeti függvény tolódik c sebességgel jobbra ($c > 0$) ill. balra ($c < 0$).

Természetes elvárások a konvergencián kívül: maradjon meg a kezdeti függvény alakja, és a haladási sebessége.

Első próbálkozás

Rácshálót definiálunk a megoldási tartományon
($u_j^k \approx u(k\Delta t, j\Delta x)$).



FTCS-séma (EE-séma): ($k = 0, 1, \dots; j = 1, \dots, n$)

$$\frac{u_j^{k+1} - u_j^k}{\Delta t} + c \frac{u_{j+1}^k - u_{j-1}^k}{2\Delta x} = 0.$$

Első próbálkozás

Az $r = c\Delta t/\Delta x$ és $\bar{\mathbf{u}}^k = [u_1^k, \dots, u_n^k]^T$ választással a séma az alábbi vektoriterációs alakot ölti:

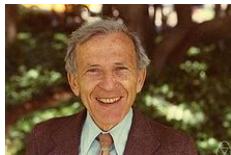
$$\bar{\mathbf{u}}^{k+1} = \begin{bmatrix} 1 & -r/2 & & & r/2 \\ r/2 & 1 & -r/2 & & \\ & \ddots & \ddots & \ddots & \\ & & r/2 & 1 & -r/2 \\ -r/2 & & & r/2 & 1 \end{bmatrix} \bar{\mathbf{u}}^k$$

Próbáljuk ki $u_0(x) = \sin(2\pi x)$ kezdeti feltétellel és $c = 1$, $n = 30$, $r = 1.5$ értékekkel!

A módszer $\mathcal{O}(\Delta t + \Delta x^2)$ -es konzisztenciájú. Sajnos az látható, hogy "valami nem stimmel" a módszerrel. Mi lehet a rossz viselkedés oka?

A stabilitásvizsgálat előtt bemutatunk egy könnyen ellenőrizhető szükséges feltételt a stabilitásra ill. konvergenciára (ez a CFL-feltétel).

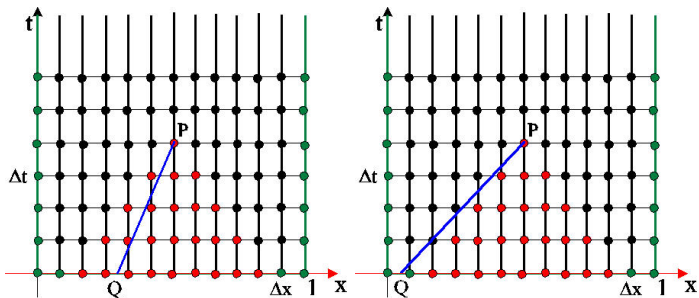
CFL-feltétel



- ▶ Richard Courant (1888-1972) német származású (amerikai) matematikus
- ▶ Kurt Otto Friedrichs (1901-1982) német származású (amerikai) matematikus
- ▶ Hans Lewy (1904-1988) német származású (amerikai) Wolf-díjas matematikus

Courant, R.; Friedrichs, K.; Lewy, H. (1928), "Über die partiellen Differenzengleichungen der mathematischen Physik" (in German), *Mathematische Annalen* 100 (1): 32-74.

CFL-feltétel



Piros pontok: A P pontbeli numerikus megoldás függőségi halmaza.
Kék vonal: A P pontbeli pontos megoldás függőségi halmaza.

CFL-feltétel: A numerikus séma, csak akkor lehet konvergens, ha numerikus függőségi halmaza tartalmazza a pontos megoldás függőségi halmazát (szükséges a konvergenciához).

CFL-feltétel

A bevezető példában:

A karakterisztika meredeksége: $1/c$. A numerikus függőségi tartomány oldalainak meredeksége $\Delta t/\Delta x$. A CFL-feltétel teljesül, ha

$$|r| = \left| \frac{c\Delta t}{\Delta x} \right| \leq 1.$$

(r elnevezése: CFL-szám (vagy rácshányados).)

A példában $r = 1.5$, azaz nem teljesül a feltétel.

Próbáljuk megoldani a feladatot akkor $r = 0.3$ -mal!

Most már teljesül a CFL-feltétel, de még mindig valami nem jó. A CFL-feltétel csak szükséges. Nézzük meg a stabilitás feltételét!

A FTCS-séma stabilitása

Könnyen ellenőrizhető, hogy az FTCS-séma növekedési faktora

$$g(\xi) = 1 - i r \sin(\xi).$$

Vegyük észre, hogy

$$|g(\xi)|^2 = 1 + r^2 \sin^2(\xi),$$

így pozitív r esetén nem lehet stabil a séma (nem teljesül a von Neumann-feltétel).

A FTCS-séma így a gyakorlatban nem használható, mert nem stabil, és így nem konvergens (konzisztensnek nyilván konzisztens).
A CFL-feltételt viszont kielégíti.

Upwind séma

Egy egyszerű és konvergens séma (előző időrétegen való lineáris interpolációból):

$$u_j^{k+1} = \begin{cases} u_j^k - r(u_{j+1}^k - u_j^k) = (1+r)u_j^k - ru_{j+1}^k & , c < 0, \\ u_j^k - r(u_j^k - u_{j-1}^k) = (1-r)u_j^k + ru_{j-1}^k & , c \geq 0. \end{cases}$$

CFL-feltétel: teljesül, ha $|r| \leq 1$.

Diszkrét Fourier-tr.-val a növekedési faktor:

$$g(\xi) = 1 - |r|(1 - e^{-i\xi}),$$

azaz

$$|g(\xi)|^2 = 1 - 4|r|(1 - |r|) \sin^2(\xi/2).$$

Azaz a CFL-feltétel mellett stabil is a séma (most ez elégséges is). A stabilitás mellett még a konzisztenciát kell igazolni, és ketten együtt biztosítják a konvergenciát már.

Az upwind séma konzisztenciája ($c > 0$ esete, másik hasonlóan)

A szokásos jelölések használva és feltéve, hogy a megoldás elegendően sima, a képlethibára az alábbi becslést kapjuk.

Képlethiba:

$$\begin{aligned} T_j^{k+1} &:= \frac{u_j^{k+1} - u_j^k}{\Delta t} + c \frac{u_j^k - u_{j-1}^k}{\Delta x} \\ &= [u'_t]_j^k + \frac{1}{2} \Delta t \underbrace{[u''_{tt}]_j^{k^*}}_{c^2 u''_{xx}} + c \left([u'_x]_j^k - \frac{1}{2} \Delta x [u''_{xx}]_{j^*}^k \right), \end{aligned}$$

tehát

$$|T_j^{k+1}| \leq \frac{1}{2} c \Delta x (1 - r) M_{xx} = \mathcal{O}(\Delta t + \Delta x),$$

ahol M_{xx} felső becslés a megoldás x szerinti második deriváltjára a tartományon.

Az upwind séma konvergenciája ($c > 0$ esete, másik hasonlóan)

Az előző oldali konzisztenciavizsgálat a pontonkénti konzisztenciát mutatja. A Lax-tételhez normabeli konzisztenciát kell igazolni. Ez teljesül, ha hasonló feltételeket szabunk, mint a hővezetési egyenlet esetén. Így igazoltuk az alábbi állítást.

28. Tétel. A (3) advekciós egyenletre (mint kezdetiérték-feladatra) alkalmazott konzisztens upwind-séma pontosan akkor konvergens $\|\cdot\|_{2,\Delta x}$ -normában, ha $|r| \leq 1$.

Az upwind séma periodikus peremmel ($c > 0$ esete, másik hasonlóan)

Tekintsük a (3) feladatot periodikus peremmel a $[0,1]$ intervallumon! Legyen $\Delta x = 1/n$, és közelítsük u értékét az $x_j = j\Delta x$ ($j = 1, \dots, n$) pontokban. x_0 -ban ugyanaz az érték, mint x_n -ben a periodikus perem miatt.

Ekkor a séma alakja:

$$\bar{\mathbf{u}}^{k+1} = \underbrace{\begin{bmatrix} 1-r & & & & r \\ & r & 1-r & & \\ & & \ddots & \ddots & \\ & & & & r & 1-r \end{bmatrix}}_{\mathbf{Q}} \bar{\mathbf{u}}^k.$$

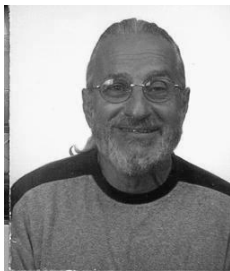
Az upwind séma periodikus peremmel ($c > 0$ esete, másik hasonlóan)

Az $|r| \leq 1$ feltétel nyilvánvalóan szükséges a stabilitáshoz. Az elégségesség pedig a következő módon látszik: A \mathbf{Q} iterációs mátrix 1-es ill. maximumnormája is 1, így $\|\mathbf{Q}\|_{2,\Delta x} \leq 1$. Ez biztosítja a stabilitást. Így igaz az alábbi tétel.

29. Tétel. A periodikus peremfeltétellel ellátott (3) advekcíós egyenletre alkalmazott konzisztens upwind-séma pontosan akkor konvergens $\|\cdot\|_\infty$ vagy $\|\cdot\|_{2,\Delta x}$ -normában, ha $|r| \leq 1$

Figyeljük meg a séma működését különböző r értékekre!

Lax–Wendroff-séma



Lax Péter (1926, Budapest), magyar származású amerikai matematikus.

Burton Wendroff (1930) amerikai matematikus.

Lax–Wendroff-séma

Kvadratikus interpoláció segítségével származtatható.

$$u_j^{k+1} = \frac{1}{2}r(1+r)u_{j-1}^k + (1-r^2)u_j^k - \frac{1}{2}r(1-r)u_{j+1}^k.$$

A séma konzisztenciarendje $\mathcal{O}(\Delta t^2 + \Delta x^2)$.

Próbáljuk ki periodikus perem esetén!

Növekedési faktor:

$$g(\xi) = 1 - ir \sin(\xi) - 2r^2 \sin^2(\xi/2),$$

ahonnan

$$|g(\xi)|^2 = 1 - 4r^2(1-r^2) \sin^4(\xi/2).$$

Ez mutatja, hogy a CFL-feltétel elégséges is a stabilitáshoz, azaz a konzisztenciával együtt a konvergenciához.

Leap-frog séma

Ez egy kétlépéses (három réteges séma).

$$\frac{u_j^{k+1} - u_j^{k-1}}{2\Delta t} + c \frac{u_{j+1}^k - u_{j-1}^k}{2\Delta x} = 0$$

(indítás pl. a Lax–Wendroff-sémával).

Konzisztenciarendje: $\mathcal{O}(\Delta t^2 + \Delta x^2)$.

Növekedési mátrix:

$$G(\xi) = \begin{bmatrix} -2ri \sin \xi & 1 \\ 1 & 0 \end{bmatrix},$$

melynek sajátértékei

$$\lambda_{\pm} = -ir \sin \xi \pm \sqrt{1 - r^2 \sin^2 \xi}.$$

Ha a CFL-feltétel teljesül, akkor a von Neumann-feltétel is teljesül ($|\lambda_{\pm}| = 1$). Megmutatható, hogy a CFL-feltétel elégséges is a stabilitáshoz.

BTCS-séma - Példa implicit sémára

Az

$$\frac{u_j^{k+1} - u_j^k}{\Delta t} + c \frac{u_{j+1}^{k+1} - u_{j-1}^{k+1}}{2\Delta x} = 0$$

alából kapjuk, hogy

$$-\frac{r}{2}u_{j-1}^{k+1} + u_j^{k+1} + \frac{r}{2}u_{j+1}^{k+1} = u_j^k.$$

Konzisztenciarendje: $\mathcal{O}(\Delta t + \Delta x^2)$.

Növekedési faktor:

$$g(\xi) = \frac{1}{1 + ri \sin \xi},$$

melyre

$$|g(\xi)| = \frac{1}{1 + r^2 \sin^2 \xi} \leq 1,$$

azaz a módszer feltétel nélkül stabil.

CN-séma - Másik példa implicit sémára

A Crank-Nicolson-séma az advekcíós egyenletre.

$$-\frac{r}{4}u_{j-1}^{k+1} + u_j^{k+1} + \frac{r}{4}u_{j+1}^{k+1} = \frac{r}{4}u_{j-1}^k + u_j^k - \frac{r}{4}u_{j+1}^k.$$

Konzisztenciarendje: $\mathcal{O}(\Delta t^2 + \Delta x^2)$.

Növekedési faktor:

$$g(\xi) = \frac{1 - \mathbf{i}r \sin(\xi)/2}{1 + \mathbf{i}r \sin(\xi)/2},$$

melyre

$$|g(\xi)| \equiv 1,$$

azaz a módszer feltétel nélkül stabil.

Egyenletmegoldás a Sherman–Morrison-formulával

Kezdeti- és peremérték-feladatok esetén használható az alábbi tétel.

Az \mathbf{B} mátrix inverzének ismeretében az $\mathbf{A} = \mathbf{B} - \overline{\mathbf{w}}\mathbf{z}^T$ mátrix inverzét szeretnénk kiszámolni praktikus módon.

30. Tétel. Tegyük fel, hogy létezik \mathbf{B}^{-1} és $\overline{\mathbf{z}}^T\mathbf{B}^{-1}\overline{\mathbf{w}} \neq 1$. Ekkor \mathbf{A}^{-1} is létezik, és

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \alpha\mathbf{B}^{-1}\overline{\mathbf{w}}\mathbf{z}^T\mathbf{B}^{-1},$$

ahol

$$\alpha = \frac{1}{1 - \overline{\mathbf{z}}^T\mathbf{B}^{-1}\overline{\mathbf{w}}}.$$

Egyenletmegoldás a Sherman–Morrison-formulával

Tegyük fel, hogy az $\mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{b}}$ egyenletrendszert szeretnénk megoldani az előző tételt használva. Ekkor

$$\bar{\mathbf{x}} = \mathbf{A}^{-1}\bar{\mathbf{b}} = \mathbf{B}^{-1}\bar{\mathbf{b}} + \alpha(\bar{\mathbf{z}}^T\mathbf{B}^{-1}\bar{\mathbf{b}})\mathbf{B}^{-1}\bar{\mathbf{w}}.$$

Így a megoldási algoritmus a következő:

- ▶ Oldjuk meg a $\mathbf{B}\bar{\mathbf{y}}_1 = \bar{\mathbf{b}}$ és $\mathbf{B}\bar{\mathbf{y}}_2 = \bar{\mathbf{w}}$ egyenletrendszereket!
- ▶ Legyen

$$\beta = \frac{\bar{\mathbf{z}}^T\bar{\mathbf{y}}_1}{1 - \bar{\mathbf{z}}^T\bar{\mathbf{y}}_2}.$$

- ▶ $\bar{\mathbf{x}} = \bar{\mathbf{y}}_1 + \beta\bar{\mathbf{y}}_2.$

Egyenletmegoldás a Sherman–Morrison-formulával

Pl. a Crank–Nicolson-módszerbeli egyenletrendszer megoldásánál a

$$\mathbf{Q}_1 = \begin{bmatrix} 1 & r/4 & & & -r/4 \\ -r/4 & 1 & r/4 & & \\ & \ddots & \ddots & \ddots & \\ & & & -r/4 & 1 & r/4 \\ r/4 & & & & -r/4 & 1 \end{bmatrix}$$

mátrix invertálható az alábbi választással: $\bar{\mathbf{w}} = [1, 0, \dots, 0, -1]^T$,

$\bar{\mathbf{z}} = [r/4, 0, \dots, 0, r/4]^T$ és

$$\mathbf{B} = \begin{bmatrix} 1 + r/4 & r/4 & & & 0 \\ -r/4 & 1 & r/4 & & \\ & \ddots & \ddots & \ddots & \\ & & & -r/4 & 1 & r/4 \\ 0 & & & & -r/4 & 1 - r/4 \end{bmatrix}.$$

Példa a Dirichlet-peremfeltétel kezelésére

Legyen $c > 0$. Ekkor csak az $x = 0$ pontban kell peremfeltételt megadni, az $x = 1$ -ben nem. Legyen a peremfeltétel $u(t, 0) = g(t)$.

Pl. implicit upwind-séma:

$$\begin{bmatrix} 1+r & 0 & & & \\ -r & 1+r & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & -r & 1+r & 0 \\ & & & -r & 1+r \end{bmatrix} \mathbf{u}^{k+1} = \mathbf{u}^k + \begin{bmatrix} rg^{k+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Ekkor

$$g(\xi) = \frac{1}{1+r-re^{-i\xi}},$$

azaz

$$|g(\xi)|^2 = \frac{1}{1+4r\sin^2(\xi/2)+4r^2\sin^2(\xi/2)}.$$

Példa a Dirichlet-peremfeltétel kezelésére

A séma kezdetiérték-feladatokra minden $r > 0$ esetén stabil.

A Dirichlet-peremes esetben \mathbf{Q}_1 és transzponáltja is M-mátrix. Így a szokásos becsléssel

$$\|\mathbf{Q}_1^{-1}\|_{\infty} \leq 1, \quad \|\mathbf{Q}_1^{-1}\|_{2,\Delta x} \leq 1.$$

Azaz Dirichlet-perem esetén is stabil a séma tetszőleges $r > 0$ mellett maximum- és $\|\cdot\|_{2,\Delta x}$ normában is.

Az advekciós egyenlet disszipációja és diszperziója

Fourier-transzformáció (hely-hullámszám transzformáció):

$$\hat{v}(\beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v(x) e^{-ix\beta} dx$$

Vegyük az advekciós egyenlet mindkét oldalának Fourier-transzformáltját a helyváltozó szerint!

$$\hat{u}'_t(t, \beta) + ci\beta\hat{u}(t, \beta) = 0.$$

Ennek megoldása

$$\hat{u}(t, \beta) = e^{-ic\beta t} \hat{u}(0, \beta).$$

Visszatranszformálva kapjuk, hogy

$$u(t, x) = u_0(x - ct).$$

Az advekciónak egyenlet disszipációja és diszperziója

Azaz az adja a megoldást, hogy a kezdeti függvény c sebességgel tolódik el jobbra.

Δt idő alatt tehát nem változik a hullám alakja (minden összetevő amplitúdója állandó marad), és $c\Delta t$ -vel kerül jobbra (minden összetevő).

Ezeket leolvashatjuk az egyenlet

$$g_{eq}(\beta) = e^{-ic\beta\Delta t}$$

növekedési faktoráról. A növekedési faktor $R_{eq}(\beta) := |g_{eq}(\beta)|$ abszolút értéke az amplitúdóváltozást, míg a

$$\Phi_{eq}(\beta) := \frac{-\arg(g_{eq})}{\beta} = c\Delta t = r\Delta x$$

hányados az eltolódás mértékét adja meg.

Az advekciós egyenlet disszipációja és diszperziója

31. Def. Disszipáció: A differenciálegyenlet akkor disszipatív, ha $R_{eq}(\beta) \leq 1$, és legalább egy β -ra szigorú egyenlőtlenség van.

32. Def. Diszperzió: A differenciálegyenlet akkor diszperzív, ha $\Phi_{eq}(\beta)$ függ β -tól. Különben diszperziómentes.

Így tehát a vizsgált advekciós egyenlet nem disszipatív és nem is diszperzív.

Vizsgáljuk meg, hogy a numerikus megoldás hogyan viselkedik ebből a szempontból!

Az advekción egyenlet disszipációja és diszperziója

Tekintsük az alábbi közelítést:

$$\begin{aligned}\hat{v}(\beta) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v(x) e^{-ix\beta} dx \\ &\approx \frac{1}{\sqrt{2\pi}} \sum_{j=-\infty}^{\infty} v(j\Delta x) e^{-ij \overbrace{\Delta x}^{\xi} \beta} \Delta x = \hat{v}(\xi) \Delta x.\end{aligned}$$

Így a numerikus séma növekedési faktora közelíti az egyenlet növekedési faktorát, és

$$R(\beta) := |g(\Delta x \beta)|$$

megadja az egyes összetevők amplitúdójának növekedését, és

$$\Phi(\beta) := \frac{-\arg(g(\Delta x \beta))}{\beta}$$

megadja az eltolódásuk mértékét.

Az upwind séma disszipációja és diszperziója

Gyakran használhatjuk az alábbi közelítést, melyet sorfejtéssel igazolhatunk.

Lemma.

$$\begin{aligned} & \operatorname{arctg}(c_1 p + c_2 p^2 + c_3 p^3 + c_4 p^4 + \dots) \\ &= c_1 p + c_2 p^2 + (c_3 - c_1^3/3) p^3 + (c_4 - c_1^2 c_2) p^4 + \dots \end{aligned}$$

Az upwind-séma növekedési faktora:

$$g(\xi) = 1 - |r|(1 - e^{-i\xi}),$$

így

$$R(\beta) = \sqrt{1 - 4|r|(1 - |r|) \sin^2(\Delta x \beta/2)}.$$

A $\sqrt{1 - x} = 1 - x/2 - x^2/8 - x^3/16 - \dots$ sorfejtés miatt ez azt mutatja, hogy az amplitúdó hibája $(\beta \Delta x)^2$ -tel arányos.

Ez mutatja, hogy általában van disszipáció, sőt a képletből az is látszik, hogy ez mikor (r -től és β -től függően) mekkora.

Az upwind-séma disszipációja és diszperziója

Mivel $g(\beta\Delta x) = (1 - r) + r \cos(\beta\Delta x) - ir \sin(\beta\Delta x)$, emiatt az argumentum

$$\arg(g) = -\operatorname{arctg} \frac{r \sin(\beta\Delta x)}{(1 - r) + r \cos(\beta\Delta x)}$$

$$= -r\beta\Delta x(1 - (1 - r)(1 - 2r)(\beta\Delta x)^2/6 + \dots).$$

Így az eltolódás mértéke

$$\Phi(\beta) = c\Delta t(1 - (1 - r)(1 - 2r)(\beta\Delta x)^2/6 + \dots).$$

Ez függ β -tól, azaz más hullámszámú hullámok eltolódása más és más lesz. $r > 1/2$ esetén sietni fog, $r < 1/2$ esetén késni fog a numerikus megoldás a pontos megoldáshoz képest.

A Lax–Wendroff-séma disszipációja és diszperziója

Növekedési faktor:

$$g(\xi) = 1 - ir \sin(\xi) - 2r^2 \sin^2(\xi/2),$$

ahonnét

$$R(\beta)^2 = 1 - 4r^2(1 - r^2) \sin^4(\beta\Delta x/2),$$

ill.

$$\arg(g(\beta\Delta x)) = -r\beta\Delta x(1 - (1 - r^2)(\beta\Delta x)^2/6 + \dots),$$

így az eltolódás

$$\Phi(\beta) = c\Delta t(1 - (1 - r^2)(\beta\Delta x)^2/6 + \dots).$$

Azaz disszipáció (nem olyan erős, mint az upwindnél) és diszperzió is van (minden összetevő késik).

Inicializáció

Már több kétlépéses sémával is találkoztunk eddig. Kétlépéses (három réteges) sémák esetén csak a 0. réteg értékei következnek a kezdeti feltételből. Az első réteg értékeit nekünk kell közelíteni (inicializálás).

Nézzük meg, hogy milyen lehetőségek vannak az indításra (pl., ha maga a séma $\mathcal{O}(\Delta t^2 + \Delta x^2)$ -es, pl. a leap-frog-séma az advekciós egyenletre):

- ▶ Használjunk egy ugyanolyan rendű egylépéses sémát, mint maga a séma rendje ($\mathcal{O}(\Delta t^2 + \Delta x^2)$).
- ▶ Használjunk egy időben elsőrendű és térben másodrendű sémát az alábbi módon: $\delta t = \Delta t^2$, $\delta x = \Delta x$. Ekkor a konvergenciarend $\mathcal{O}(\delta t + \delta x^2) = \mathcal{O}(\Delta t^2 + \Delta x^2)$.

Inicializálás

- ▶ Nem is kell a fenti két megszorítás, hiszen az első időrétegen a hibát nem a konvergencia, hanem a konzisztencia (a lokális hiba) határozza meg. Így elegendő olyan sémát használni az első lépésben, amelynek lokális hibája $\mathcal{O}(\Delta t^2 + \Delta x^2)$, azaz $\bar{\mathbf{T}}^1$ képlethibája olyan, hogy $\Delta t \|\bar{\mathbf{T}}^1\| = \mathcal{O}(\Delta t^2 + \Delta x^2)$. (Azaz pl. nem stabil sémával is lehet inicializálni, ha konzisztenciarendje megfelelő.)

Végezzünk numerikus kísérleteket pl. az leap-frog sémára az advekciós egyenlet esetén!

Egy elsőrendű differenciálegyenlet-rendszer

$$u'_t + cv'_x = 0$$

$$v'_t + cu'_x = 0$$

A feladatra jól alkalmazható az eltolt rácson leap-frog séma.

Rács u -nak (egész sorszámú időrétegeken):

$$x_j = j\Delta x, \quad j = \dots, -1, 1, 1, \dots$$

Rács v -nek (feles sorszámú időrétegeken):

$$x_j = j\Delta x, \quad j = \dots, -1/2, 1/2, 3/2, \dots$$

Az eltolt rácsos leap-frog séma

$$\frac{u_j^{k+1} - u_j^k}{\Delta t} + c \frac{v_{j+1/2}^{k+1/2} - v_{j-1/2}^{k+1/2}}{\Delta x} = 0$$
$$\frac{v_{j+1/2}^{k+1/2} - v_{j+1/2}^{k-1/2}}{\Delta t} + c \frac{u_{j+1}^k - u_j^k}{\Delta x} = 0$$

A v inicializálása az $1/2$ -es időrétegen valamilyen alkalmas módszerrel történik.

A séma $\mathcal{O}(\Delta t^2 + \Delta x^2)$ -es és stabil, ha $|r| \leq 1$ (CFL-feltétel).

8.

A másodrendű hullámegyenlet megoldása (véges differencia)

A másodrendű hullámeqyenlet

A másodrendű hullámeqyenlet

$$\begin{aligned}\frac{\partial^2 u}{\partial t^2} &= c^2 \frac{\partial^2 u}{\partial x^2}, \\ u(0, x) &= u_0(x), \\ u'_t(0, x) &= v_0(x)\end{aligned}$$

(ha véges intervallumon tűzzük ki, akkor még peremfeltétel is kell).

Pontos megoldás:

$$u(t, x) = \frac{1}{2}(u_0(x - ct) + u_0(x + ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} v_0(y) dy.$$

A folytonos feladat megoldásának függőségi halmaza: $u(t, x)$ függ a kezdeti feltételek $[x - ct, x + ct]$ halmazon vett értékeitől.

CTCS-séma a hullámegyenletre

Tekintsük az alábbi sémát a hullámegyenletre:

$$\frac{u_j^{k-1} - 2u_j^k + u_j^{k+1}}{\Delta t^2} = c^2 \frac{u_{j-1}^k - 2u_j^k + u_{j+1}^k}{\Delta x^2}.$$

A séma növekedési mátrixa:

$$G(\xi) = \begin{bmatrix} 2 + 2r^2(\cos \xi - 1) & -1 \\ 1 & 0 \end{bmatrix}.$$

Megj.: A korábbi stabilitás-definíció teljesüléséhez pl. teljesülni kellene a $\|G^k(0)\|_2 \leq Ke^{\beta t}$ feltételnek valamilyen megfelelő konstansokkal. Mivel

$$G(0) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix},$$

így

$$G^k(0) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}.$$

CTCS-séma a hullámegyenletre

Ezek szerint

$$\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} G^k(0) \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix},$$

azaz

$$\left\| \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \right\|_2 \left\| G^k(0) \right\|_2 \left\| \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\|_2 \geq \left\| \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \right\|_2 \geq k,$$

ami mutatja, hogy a korábbi értelemben nem stabil a séma.

A numerikus kísérletek mégis azt mutatják, hogy a séma jól használható.

Megj.: Az $|r| \leq 1$ CFL-feltétel mellett $\varrho(G(\xi)) \equiv 1$.

Stabilitás

Az egyenlet másodrendű, így gyengébb stabilitás is garantálhatja a konvergenciát.

Követeljük meg a $\|Q^k\| \leq kKe^{\beta t}$ ún. k -rendű stabilitást valamilyen megfelelő konstansokkal.

Nézzük meg tehát, hogy hogyan módosul a Lax-tétel bizonyítása időben másodrendű feladatra és a fenti stabilitási definícióval.

Most a séma kétlépéses, így az egylépésessé átírt séma léptetőoperátorát jelöljük Q -val. Az iterációs vektor is az $[\bar{\mathbf{u}}^k; \bar{\mathbf{u}}^{k-1}]$ vektor lesz ($k = 1, 2, 3, \dots$ esetén értelmes).

A pontos megoldás sémába helyettesítése és az így kapott egyenlet sémából való kivonása után kapjuk, hogy

$$\bar{\mathbf{E}}^{k+1} = Q\bar{\mathbf{E}}^k - \Delta t^2 \bar{\mathbf{T}}^{k+1}$$

($\bar{\mathbf{T}}^{k+1}$ a képlethibavektor).

Stabilitás

Ebből kapjuk, hogy

$$\bar{\mathbf{E}}^k = Q^{k-1}\bar{\mathbf{E}}^1 - \Delta t^2(\bar{\mathbf{T}}^k + Q\bar{\mathbf{T}}^{k-1} + \dots + Q^{k-2}\bar{\mathbf{T}}^2),$$

majd becsüljük az adott normában a két oldalt a k -rendű stabilitást ill. az $\mathcal{O}(\Delta t^r + \Delta x^s)$ normabeli konzisztenciát felhasználva.

$$\begin{aligned}\|\bar{\mathbf{E}}^k\| &\leq kKe^{\beta k\Delta t}\|\bar{\mathbf{E}}^1\| + \Delta t^2 kKe^{\beta k\Delta t} k\mathcal{O}(\Delta t^r + \Delta x^s) \\ &\leq Ke^{\beta t_{\max}} k\|\bar{\mathbf{E}}^1\| + t_{\max}^2 Ke^{\beta t_{\max}} \mathcal{O}(\Delta t^r + \Delta x^s).\end{aligned}$$

Mivel $\bar{\mathbf{u}}^0$ hibája nulla, így ha $\bar{\mathbf{u}}^1$ hibája $\mathcal{O}(\Delta t(\Delta t^r + \Delta x^s))$ (feltétel az inicializálásra), akkor a módszer konvergens lesz, és konvergenciarendje megegyezik a konzisztenciarenddel, azaz $\mathcal{O}(\Delta t^r + \Delta x^s)$

A CTCS-séma inicializálása

Egy egyszerűen adódó közelítés:

$$v_0(\bar{\mathbf{x}}) = \frac{\bar{\mathbf{u}}^1 - \bar{\mathbf{u}}^0}{\Delta t},$$

amiből

$$\bar{\mathbf{u}}^1 = \bar{\mathbf{u}}^0 + \Delta t v_0(\bar{\mathbf{x}}).$$

Mivel

$$u(\Delta t, \bar{\mathbf{x}}) = u(0, \bar{\mathbf{x}}) + u'_t(0, \bar{\mathbf{x}})\Delta t + u''_{tt}(0, \bar{\mathbf{x}}^*)\Delta t^2/2,$$

így a két egyenlet kivonásából látható, hogy $\|\bar{\mathbf{u}}^1 - u(\Delta t, \bar{\mathbf{x}})\| = \mathcal{O}(\Delta t^2)$, ami nem elég ahhoz, hogy megmaradjon a séma $\mathcal{O}(\Delta t^2 + \Delta x^2)$ rendje.

Próbáljuk ki egy tesztfeladaton!

A CTCS-séma inicializálása

Magasabbrendű közelítést kapunk, ha tovább használjuk a sorfejtést:

$$\begin{aligned}u(\Delta t, \bar{\mathbf{x}}) &= u(0, \bar{\mathbf{x}}) + u'_t(0, \bar{\mathbf{x}})\Delta t + u''_{tt}(0, \bar{\mathbf{x}})\Delta t^2/2 + \mathcal{O}(\Delta t^3) \\&= u(0, \bar{\mathbf{x}}) + u'_t(0, \bar{\mathbf{x}})\Delta t + c^2 u''_{xx}(0, \bar{\mathbf{x}})\Delta t^2/2 + \mathcal{O}(\Delta t^3) \\&= u(0, \bar{\mathbf{x}}) + u'_t(0, \bar{\mathbf{x}})\Delta t + r^2 \text{tridiag}(1, -2, 1)u(0, \bar{\mathbf{x}})/2 + \mathcal{O}(\Delta t^3 + \Delta t^2 \Delta x^2)\end{aligned}$$

Ez mutatja, hogy az

$$\bar{\mathbf{u}}^1 = \bar{\mathbf{u}}^0 + v_0(\bar{\mathbf{x}})\Delta t + r^2 \text{tridiag}(1, -2, 1)\bar{\mathbf{u}}^0/2$$

közelítés meg fogja őrizni a séma másodrendjét időben és térben is.

Próbáljuk ki egy tesztfeladaton!

9.

Másodrendű, kétdimenziós, időfüggő,
lineáris PDE-ek megoldása a véges
differencia módszerrel

A hővezetési egyenlet két dimenzióban

$$\frac{\partial u}{\partial t} = \Delta u + f, \text{ kezdeti- és peremfeltételekkel}$$

A Laplace-operátor approximációjához a Poisson-egyenletnél tanultakat használjuk (\mathbf{A}_h mátrix). $\bar{\mathbf{u}}^k$ a megoldás közelítéseinek vektora a k . időrétegen, $\bar{\mathbf{f}}^k$ a forrásfüggvény közelítéseinek vektora a k . időrétegen.

EE-módszer:

$$\bar{\mathbf{u}}^{k+1} = (\mathbf{E} - \Delta t \mathbf{A}_h) \bar{\mathbf{u}}^k + \Delta t \bar{\mathbf{f}}^k.$$

ADI módszer a kétdimenziós hővezetési egyenletre

CN-módszer:

$$(\mathbf{E} + (1/2)\Delta t \mathbf{A}_h) \bar{\mathbf{u}}^{k+1} = (\mathbf{E} - (1/2)\Delta t \mathbf{A}_h) \bar{\mathbf{u}}^k + \frac{\Delta t}{2} (\bar{\mathbf{f}}^{k+1} + \bar{\mathbf{f}}^k).$$

A megoldás műveletigényes, mert nem lehet kis sávszélességű mátrixként felírni $\mathbf{A}_h - t$. Egy lehetséges megoldás lehet az **ADI (alternating directions implicit)** módszer. Írjuk fel $-\Delta t \mathbf{A}_h$ -t

$$-\Delta t \mathbf{A}_h = q_x \mathbf{X}_h + q_y \mathbf{Y}_h$$

alakban. Ezzel a CN-módszer:

$$\begin{aligned} & (\mathbf{E} - (q_x/2)\mathbf{X}_h - (q_y/2)\mathbf{Y}_h) \bar{\mathbf{u}}^{k+1} \\ &= (\mathbf{E} + (q_x/2)\mathbf{X}_h + (q_y/2)\mathbf{Y}_h) \bar{\mathbf{u}}^k + \frac{\Delta t}{2} (\bar{\mathbf{f}}^{k+1} + \bar{\mathbf{f}}^k) \end{aligned}$$

alakú lesz.

ADI módszer a kétdimenziós hővezetési egyenletre

Ez az iteráció rendvesztés nélkül közelíthető az alábbi módon (Peaceman–Rachford-séma):

$$(\mathbf{E} - (q_x/2)\mathbf{X}_h)\bar{\mathbf{u}}^{k+1/2} = (\mathbf{E} + (q_y/2)\mathbf{Y}_h)\bar{\mathbf{u}}^k + \frac{\Delta t}{2}\bar{\mathbf{f}}^k$$

$$(\mathbf{E} - (q_y/2)\mathbf{Y}_h)\bar{\mathbf{u}}^{k+1} = (\mathbf{E} + (q_x/2)\mathbf{X}_h)\bar{\mathbf{u}}^{k+1/2} + \frac{\Delta t}{2}\bar{\mathbf{f}}^{k+1}.$$

A módszer feltétel nélkül stabil és a maximumelv teljesülésének feltétele $q_x, q_y \leq 1$.

A másodrendű hullámeqyenlet két dimenzióban

$$\frac{\partial^2 u}{\partial t^2} = \Delta u, \text{ kezdeti- és peremfeltételekkel}$$

Pl. homogén Dirichlet eset:

$$\frac{\bar{\mathbf{u}}^{k+1} - 2\bar{\mathbf{u}}^k + \bar{\mathbf{u}}^{k-1}}{\Delta t^2} = -\mathbf{A}_h \bar{\mathbf{u}}^k$$

Inicializálás:

$$\bar{\mathbf{u}}^1 = \bar{\mathbf{u}}^0 + \Delta t \bar{\mathbf{v}} - \frac{\Delta t^2}{2} \mathbf{A}_h \bar{\mathbf{u}}^0$$

10.

Módszertípusok

Módszertípusok áttekintése

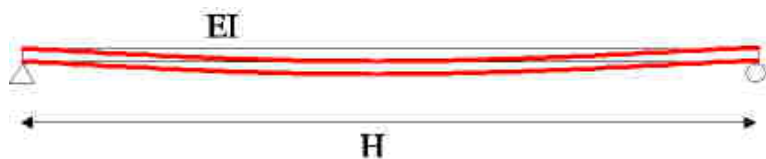
I. Véges differenciák módszere

II. Integrálmódszerek

- ▶ **Variációs módszer (Ritz-módszer)**
- ▶ **Súlyozott reziduum módszerek**
 - Kollokációs módszer
 - Véges térfogat módszer (vagy résztartomány módszer)
 - **Galjorkin-módszer**
 - Momentumok módszere
 - Legkisebb négyzetek módszere

Megj.: Ha az integrálmódszerekben a tesztfüggvényeket kis tartójú, szakaszonként polinomiális függvényeknek választjuk, akkor végeelem módszerről beszélünk.

Egy egydimenziós peremértékfeladat



Lehajló rúd egyenlete:

$$-u'' = \frac{-M(x)}{EI} = f(x) \equiv -M_0, \quad u(0) = u(H) = 0.$$

A pontos megoldás:

$$u(x) = \frac{M_0}{2}(x^2 - Hx),$$

a rúd közepén a lehajlás

$$u(H/2) = -\frac{M_0 H^2}{8} = -0.125 M_0 H^2.$$

Véges differenciák módszere

Az ismeretlen függvényt rácsfüggvénnyel közelítjük, a deriváltakat pedig véges differenciákkal. Az így nyert egyenletrendszert megoldjuk az ismeretlen rácsfüggvényértékekre.

Legyenek $\mathbf{Q} = \text{tridiag}[-1, 2, -1] \in \mathbb{R}^{n \times n}$, $h = H/(n+1)$, $n \in \mathbb{N}$, $x_i = ih$ ($i = 1, \dots, n$). u_i közelítse az $u(x_i)$ függvényértéket.

Ekkor a

$$-\frac{1}{h^2} \mathbf{Q} \bar{\mathbf{u}} = M_0 \bar{\mathbf{e}}$$

egyenletrendszert kell megoldani.

$n=101$ esetén a megoldásvektor középső eleme:

$$u_{51} = -0.124999999999999 M_0 H^2 \approx -M_0 H^2 / 8.$$

Integrálmódszerek

Tesztfüggvények V_h N dim. vektortereből választjuk ki valamilyen szempontból a legmegfelelőbbet a tesztfüggvényeket megfelelő integrálokba helyettesítve.

Variációs módszer: A differenciálegyenletet egy variációs feladat Euler–Lagrange-egyenleteként fogjuk fel, majd a kapott funkcionált minimalizáljuk a tesztfüggvények halmazán.

Integrálmódszerek

Súlyozott reziduum módszer: A $-\mathcal{L}u^* = f$ (u^* a pontos megoldás) általános alakú feladat esetén a reziduum: $r = f + \mathcal{L}u$. Így az

$$\int_{\Omega} w(x)r(x) \, dx = 0$$

egyenlőség minden $w(x)$ ún. súlyfüggvényre fennáll, ha a reziduumot a pontos megoldással számítjuk. A súlyozott reziduum módszerben a fenti egyenlőséget követeljük meg N adott súlyra ($w_i(x)$) a tesztfüggvényekre. Olyan $\phi(x)$ tesztfüggvényt keresünk tehát, amire

$$I_i(\phi(x)) = \int_{\Omega} w_i(x)r(\phi(x)) \, dx = 0, \quad i = 1, \dots, N.$$

Variációs módszer



Walter Ritz (1878-1909) svájci elméleti fizikus.

$$I(\phi(x)) = \int_0^H \left(\frac{1}{2}(\phi'(x))^2 + M_0\phi(x) \right) dx$$

$\phi(x) = u(x)$ -re minimum a $V = \{\phi : C^2[0, H], \phi(0) = \phi(H) = 0\}$ halmazon



$u(x)$ megoldása a $\phi'' = M_0$, $y(0) = y(H) = 0$ peremértékfeladatnak (Euler–Lagrange-egyenlet).

Variációs módszer

Legyen a próbafüggvények halmaza most

$$V_h = \{ \phi : [0, H] \rightarrow \mathbb{R} \mid \phi(x) = A \sin(\pi x/H), A \in \mathbb{R} \}.$$

Minimalizáljuk az I funkcionált V_h -n. Legyen $\phi_A \in V_h$. Ekkor

$$\begin{aligned} I(\phi_A(x)) &= \int_0^H \left(\frac{1}{2} \left(A \frac{\pi}{H} \cos \left(\frac{\pi x}{H} \right) \right)^2 + M_0 A \sin \left(\frac{\pi x}{H} \right) \right) dx \\ &= \frac{\pi^2}{4H} A^2 + \frac{2M_0 H}{\pi} A. \end{aligned}$$

Minimális, ha

$$A = -\frac{4M_0 H^2}{\pi^3},$$

azaz a

$$\phi(x) = -\frac{4M_0 H^2}{\pi^3} \sin \left(\frac{\pi x}{H} \right)$$

függvényre. Maximális lehajlás

$$-\frac{4M_0 H^2}{\pi^3} = -0.129006 M_0 H^2.$$

Kollokációs módszer

N tetszőleges pontban a tartományon belül legyen a reziduum nulla.

Integrálalakban is felírható a $w_i(x) = \delta_{x_i}(x)$ Dirac-féle függvény segítségével. (x_i -ben 1, máshol nulla az értéke, és egy adott függvénnyel vett szorzatának integrálja a függvény x_i -beli értéke.)

$$I_i(\phi(x)) = \int_{\Omega} \delta_{x_i}(x)r(x) dx = r(x_i) = 0, \quad i = 1, \dots, N.$$

Kollokációs módszer

A példabeli reziduum: $r(x) = M_0 - u''(x)$. A tesztfüggvényeken az alakja

$$r(x) = M_0 + A \left(\frac{\pi}{H} \right)^2 \sin \left(\frac{\pi x}{H} \right).$$

Az $x = H/2$ felezőpontban akkor ad ez nullát, ha

$$A = -\frac{H^2 M_0}{\pi^2}.$$

Ekkor a közelítő megoldás tehát

$$\phi(x) = -\frac{H^2 M_0}{\pi^2} \sin \left(\frac{\pi x}{H} \right),$$

Maximális lehajlás

$$-0.1013 M_0 H^2.$$

Véges térfogat módszer (vagy résztartomány módszer)

A tartományt diszjunkt résztartományokra osztjuk:

$\Omega = \Omega_1 \cup \dots \cup \Omega_N$ és a $w_i(x) = \chi_{\Omega_i}(x)$ ($i = 1, \dots, N$)

karakterisztikus függvényeket választjuk súlyoknak:

$$I_i(\phi(x)) = \int_{\Omega} \chi_{\Omega_i}(x)r(x) dx = \int_{\Omega_i} r(x) dx = 0, \quad i = 1, \dots, N.$$

A példában csak egy résztartomány van:

$$\int_0^H r(x) dx = \int_0^H M_0 + A \left(\frac{\pi}{H}\right)^2 \sin\left(\frac{\pi x}{H}\right) dx = M_0H + \frac{2A\pi}{H} = 0.$$

A értéke tehát $A = -M_0H^2/(2\pi) = -0.1592M_0H^2$, ez a max. lehajlás. Közelítő megoldás

$$\phi(x) = -\frac{M_0H^2}{2\pi} \sin\left(\frac{\pi x}{H}\right).$$

Galjorkin-módszer



Borisz Grigorjevics Galjorkin (1871-1945) orosz mérnök / matematikus.

V_h bázisa legyen ϕ_1, \dots, ϕ_N , és a közelítő megoldást keressük $\phi = \sum_k c_k \phi_k(x)$ alakban. A súlyfüggvényeket válasszuk a bázisfüggvényeknek!

$$I_i\left(\sum_k c_k \phi_k(x)\right) = \int_{\Omega} \phi_i(x) r(x) dx = 0, \quad i = 1, \dots, N.$$

Ez a módszer a végeelem módszer alapja.

Galjorkin-módszer

A példában:

$$\phi_1(x) = \sin\left(\frac{\pi x}{H}\right).$$

$$\begin{aligned} I_1(A\phi_1(x)) &= \int_0^H \phi_1(x)r(x) \, dx \\ &= \int_0^H \sin\left(\frac{\pi x}{H}\right) \left(M_0 + A \left(\frac{\pi}{H}\right)^2 \sin\left(\frac{\pi x}{H}\right) \right) \, dx \\ &= \frac{A\pi^2}{2H} + \frac{2M_0H}{\pi} = 0. \end{aligned}$$

A értéke tehát $A = -4M_0H^2/\pi^3$ (mint a Ritz-módszernél), ez a max. lehajlás. Közelítő megoldás

$$\phi(x) = \frac{-4M_0H^2}{\pi^3} \sin\left(\frac{\pi x}{H}\right).$$

Momentumok módszere

$$w_i(x) = x^{i-1}, \quad i = 1, 2, 3, \dots$$

A példában most $w_i(x) = 1$ és egy intervallum van csak, így ugyanazt kapjuk, mint a véges térfogat módszer esetén.

A értéke tehát $A = -M_0H^2/(2\pi) = -0.1592M_0H^2$, ez a max. lehajlás. Közelítő megoldás

$$\phi(x) = -\frac{M_0H^2}{2\pi} \sin\left(\frac{\pi x}{H}\right).$$

Legkisebb négyzetek módszere

Az

$$I(\phi(x)) = \int_{\Omega} (r(x))^2 dx \rightarrow \min.$$

feladat megoldását keressük. Ha V_h a ϕ_i bázisfüggvények lineáris kombinációja és a megoldást $\phi(x) = \sum_k c_k \phi_k(x)$ alakban keressük, akkor a minimumhelyen a c_k ismeretlenek szerinti deriváltaknak nullának kell lenniük. Innét kapjuk a

$$2 \int_{\Omega} r(x) \frac{\partial r(x)}{\partial c_i} dx = 0, \quad (i = 1, \dots, N)$$

egyenlőségeket. Azaz a reziduum deriváltjai a súlyok ebben a módszerben ($w_i(x) = \partial r / \partial c_i$).

Legkisebb négyzetek módszere

A feladatban:

$$\begin{aligned} I(\phi_A(x)) &= \int_0^H \left(M_0 + A \left(\frac{\pi}{H} \right)^2 \sin \left(\frac{\pi x}{H} \right) \right)^2 dx \\ &= \frac{A^2 \pi^4}{2H^3} + \frac{4M_0 \pi A}{H} + M_0^2 H \rightarrow \min. \end{aligned}$$

Minimum

$$A = -\frac{4M_0 H^2}{\pi^3}$$

esetén. Ez a max. lehajlás (mint a Ritz-módszernél), a közelítő megoldás pedig

$$\phi(x) = -\frac{4M_0 H^2}{\pi^3} \sin \left(\frac{\pi x}{H} \right).$$

Megjegyzések

- ▶ Önadjunált operátorok esetén a Ritz- és Galjorkin módszerek ugyanazt az eredményt adják.
- ▶ Végeelem módszernek hívunk egy integrálmódszert, ha a tesztfüggvények szakaszonként folytonos függvények (a bázis tesztfüggvények tartója kicsi az egész tartomány méretéhez képest). Általában a Galjorkin-módszert szokták használni ebben az esetben.
- ▶ Természetesen eddig csak közelítő megoldásokat állítottunk elő, de nem foglalkoztunk azzal, hogy mekkora a közelítő megoldás hibája. Ezt a kérdést majd később a konkrét feladatoknál fogjuk vizsgálni.

11.

A végeselem módszer elmélete

A minimalizációs és variációs feladatok

Legyen V valós vektortér és $a : V \times V \rightarrow \mathbb{R}$ szimmetrikus ($a(u, v) = a(v, u)$, $\forall u, v \in V$), pozitív ($a(u, u) > 0$, ha $u \neq 0$) és bilineáris (mindkét változóban lineáris) forma, továbbá legyen $l : V \rightarrow \mathbb{R}$ lineáris funkcionál.

Tekintsük az alábbi ún. minimalizációs feladatot: Keressük meg azt az $u \in V$ elemet, ami minimalizálja a

$$(M) \quad I(v) = \frac{1}{2}a(v, v) - l(v)$$

függvényt V -n!

Tekintsük az alábbi ún. variációs feladatot: Keressük meg azt az $u \in V$ elemet, mellyel

$$(V) \quad a(u, v) = l(v), \quad \forall v \in V!$$

(M) és (V) ekvivalenciája

33. Tétel. u pontosan akkor megoldása (V)-nek, ha (M)-nek. Az (M) feladatnak maximum egy megoldása lehet csak.

Biz.: (V) \Rightarrow (M) Legyen u a variációs feladat megoldása, továbbá $0 \neq v \in V$ tetszőleges elem és t tetszőleges valós szám. Ekkor

$$\begin{aligned} I(u + tv) &= \frac{1}{2}a(u + tv, u + tv) - l(u + tv) \\ &= \frac{1}{2}(a(u, u) + 2ta(u, v) + t^2a(v, v)) - l(u) - tl(v) \\ &= I(u) + t \underbrace{(a(u, v) - l(v))}_{(V)\Rightarrow 0} + \frac{t^2}{2} \underbrace{a(v, v)}_{>0}. \end{aligned}$$

Azaz u az egyetlen elem, ami minimalizálja J -t.

(M) és (V) ekvivalenciája

Biz. (folyt.): (M) \Rightarrow (V) Legyen u a minimalizációs feladat megoldása. Az előzőek alapján a

$$I(u + tv) = I(u) + t(a(u, v) - l(v)) + \frac{t^2}{2}a(v, v)$$

t -től függő egyváltozós függvénynek minimuma van $t = 0$ -ban, azaz a derivált itt nulla:

$$(a(u, v) - l(v)) + 0 \cdot a(v, v) = a(u, v) - l(v) = 0.$$

Ezt akartuk igazolni. ■

Megj.: A bizonyításból kiderült, hogy

$$I(u + v) = I(u) + \frac{1}{2}a(v, v) > I(u), \quad \forall \mathbf{0} \neq v \in V.$$

Megj.: Hasonló tételt igazoltunk LER-re NM1-ből.

A megoldás létezése

Mi biztosítja, hogy van megoldása (M)-nek vagy (V)-nek? Több feltétel kell V -ről, a -ról és l -ről.

- ▶ Legyen V Hilbert-tér a $\langle \cdot, \cdot \rangle$ skaláris szorzattal és a vele indukált $\|\cdot\|$ normával.
- ▶ a legyen szimmetrikus, folytonos ($|a(u, v)| \leq \gamma \|u\| \|v\|$, $\forall u, v \in V$) és koercív (V -elliptikus) ($\exists \alpha > 0$, $a(v, v) \geq \alpha \|v\|^2$, $\forall v \in V$).
- ▶ l legyen korlátos lineáris funkcionál ($\exists \Lambda > 0$, $|l(v)| \leq \Lambda \|v\|$).

Megj.: A fenti feltételek mellett $\langle u, v \rangle_a = a(u, v)$ skaláris szorzat V -n, és a $\|v\|_a = \sqrt{a(v, v)}$ kifejezés a $\|\cdot\|$ normával ekvivalens normát definiál V -n (energianorma).

Lax-Milgram-tétel

34. Tétel. Legyen V Hilbert-tér, legyen továbbá $a : V \times V \rightarrow \mathbb{R}$ szimmetrikus, folytonos és koercív bilineáris forma, és $l : V \rightarrow \mathbb{R}$ korlátos lineáris funkcionál. Ekkor az (M) minimalizációs feladatnak (és a korábbiak szerint (V)-nek is) egyetlen megoldása van H -ban.

Biz.: Az $I(v)$ értékek alulról korlátosak V -ben, ugyanis

$$\begin{aligned} I(v) &= \frac{1}{2}a(v, v) - l(v) \geq \frac{\alpha}{2}\|v\|^2 - \Lambda\|v\| \\ &= \frac{1}{2\alpha}(\alpha\|v\| - \Lambda)^2 - \frac{\Lambda^2}{2\alpha} \geq -\frac{\Lambda^2}{2\alpha}. \end{aligned}$$

Legyen $c_1 = \inf\{I(v) \mid v \in V\}$, és legyen $\{v_k\} \subset V$ olyan sorozat, melyre $I(v_k) \rightarrow c_1$. Megmutatjuk, hogy $\{v_k\}$ Cauchy-sorozat.

Lax-Milgram-tétel

$$\begin{aligned} \alpha \|v_n - v_m\|^2 &\leq a(v_n - v_m, v_n - v_m) \\ &= 2a(v_n, v_n) + 2a(v_m, v_m) - a(v_n + v_m, v_n + v_m) \\ &\quad - 4l(v_n) - 4l(v_m) + 8l((v_n + v_m)/2) \\ &= 4I(v_n) + 4I(v_m) - 8I((v_n + v_m)/2) \\ &\leq 4I(v_n) + 4I(v_m) - 8c_1. \end{aligned}$$

Mivel $v_k \rightarrow c_1$, ezért minden $\varepsilon > 0$ számhoz van olyan n_0 , hogy minden $n, m \geq n_0$ esetén

$$\alpha \|v_n - v_m\|^2 \leq 4I(v_n) + 4I(v_m) - 8c_1 < \varepsilon.$$

Így a sorozat valóban Cauchy. Mivel V Hilbert-tér, így van $u \in V$, hogy $v_k \rightarrow u$. A folytonosság miatt pedig $I(u) = c_1$. u tehát megoldása az (M) feladatnak (a korábbiak szerint (V)-nek is).

Lax-Milgram-tétel

Az egyértelműség igazolása: Tegyük fel, hogy az $u_1 \neq u_2$ elemek is megoldásai (M)-nek. Ekkor $I(u_1) = I(u_2) = c_1$, és az $u_1, u_2, u_1, u_2, \text{ stb.}$ sorozatra nyilván $I(u_k) \rightarrow c_1$, azaz a sorozat a korábbiak szerint Cauchy, de ez csak úgy lehet, ha $\|u_1 - u_2\| = 0$, ami ellentmondást ad. ■

Megj.: A tétel általánosabban kimondható úgy, hogy V helyett V egy zárt és konvex részhalmazát tekintjük.

Köv.: Legyen a V Hilbert-téren $a(u, v) = \langle u, v \rangle$. Ez valóban szimmetrikus, bilineáris, folytonos (C-B-Sch) és koercív. Tehát minden l korlátos lineáris funkcionálhoz egyértelműen létezik olyan $u \in V$ elem, mellyel

$$\langle u, v \rangle = l(v), \quad \forall v \in V.$$

Riesz reprezentációs tétele.

Lax-Milgram-tétel

Köv.: Igaz az

$$\|u\| \leq \frac{\Lambda}{\alpha}$$

stabilitási becslés, hiszen $v = u$ választással

$$\alpha\|u\|^2 \leq a(u, u) = l(u) \leq \Lambda\|u\|,$$

melyből $\|u\|$ -val és α -val való osztás után adódik az állítás.

A Galjorkin-módszer

Legyen V_N a V Hilbert-tér egy N dimenziós altere, legyen ennek egy bázisa ϕ_1, \dots, ϕ_N . A Lax–Milgram-tétel szerint egyértelműen létezik olyan $u_N \in \mathbf{V}_N$ elem, melyre

$$(V) \quad a(u_N, v) = l(v), \quad \forall v \in V_N,$$

ill. minimalizálja az

$$(M) \quad I(v) = \frac{1}{2}a(v, v) - l(v)$$

funkcionált.

Azt a módszert, amikor a (V) feladat megoldásával adunk közelítést u -ra Galjorkin-módszernek nevezzük. (Nemszimmetrikus bilineáris formák esetén is használható.)

A Galjorkin-módszer

Keressük u_N -t $u_N = \sum_{i=1}^N c_i \phi_i$ alakban, a tetszőleges V_N -beli elemet pedig írjuk $v = \sum_{i=1}^N d_i \phi_i$ alakban. Ekkor a variációs feladat alakja

$$a \left(\sum_{i=1}^N c_i \phi_i, \sum_{i=1}^N d_i \phi_i \right) = l \left(\sum_{i=1}^N d_i \phi_i \right), \quad \forall \bar{\mathbf{d}}_N \in \mathbb{R}^N,$$

amely az

$$\bar{\mathbf{d}}_N^T \mathbf{A}_N \bar{\mathbf{c}}_N = \bar{\mathbf{d}}_N^T \bar{\mathbf{b}}_N$$

egyenletre vezet, ahol $(\mathbf{A}_N)_{ij} = a(\phi_j, \phi_i)$ és $(\mathbf{b}_N)_i = l(\phi_i)$. Mivel ennek minden $\bar{\mathbf{d}}_N \in \mathbb{R}^N$ vektorra teljesülni kell, ezért az ismeretlen együtthetők az

$$\mathbf{A}_N \bar{\mathbf{c}}_N = \bar{\mathbf{b}}_N$$

egyenletrendszer megoldásával nyerhetők. (\mathbf{A}_N elnevezése merevségi (stiffness) mátrix, $\bar{\mathbf{b}}_N$ elnevezése terhelési (load) vektor.)

A Galjorkin-módszer

Az

$$\bar{\mathbf{d}}_N^T \mathbf{A}_N \bar{\mathbf{d}}_N = a \left(\sum_{k=1}^N d_k \phi_k, \sum_{k=1}^N d_k \phi_k \right) \geq \alpha \left\| \sum_{k=1}^N d_k \phi_k \right\|^2$$

egyenlőség miatt \mathbf{A}_N pozitív definit, a szimmetrikusság a szimmetriájából következik. Így a korábbi egyenletrendszer egyértelműen megoldható.

Megj.: A korábbiakhoz hasonlóan igazolható, hogy az u_N megoldásra ugyanolyan stabilitási becslés igaz, mint az u -ra, azaz

$$\|u_N\| \leq \frac{\Lambda}{\alpha}.$$

Ritz-módszer

Azt a módszert, ahol az (M) feladat megoldásával adunk becslést u -ra, Ritz-módszernek nevezzük.

$$\begin{aligned} I \left(\sum_{k=1}^N c_k \phi_k \right) &= \frac{1}{2} a \left(\sum_{k=1}^N c_k \phi_k, \sum_{k=1}^N c_k \phi_k \right) - l \left(\sum_{k=1}^N c_k \phi_k \right) \\ &= \frac{1}{2} \bar{\mathbf{c}}_N^T \mathbf{A}_N \bar{\mathbf{c}}_N - \bar{\mathbf{c}}_N^T \bar{\mathbf{b}}_N, \end{aligned}$$

melynek minimuma szintén az $\mathbf{A}_N \bar{\mathbf{c}}_N = \bar{\mathbf{b}}_N$ egyenletrendszer megoldásánál van.

Megj.: Szimmetrikus formák esetén a Galjorkin és Ritz-módszerek ugyanarra az egyenletrendszer megoldására vezetnek. Emiatt a módszert ebben az esetben Ritz–Galjorkin-módszernek nevezzük.

Céa-lemma

35. Tétel. Legyen a szimmetrikus, folytonos és koercív bilineáris forma és l korlátos lineáris funkcionál a V Hilbert-téren. Legyen a variációs egyenletek megoldása V -n u , és V_N -en u_N . Ekkor

$$\|u - u_N\| \leq \frac{\gamma}{\alpha} \inf_{v \in V_N} \|u - v\|.$$

Biz.: Vonjuk ki egymásból a

$$a(u, v) = l(v), \quad \forall v \in V,$$

$$a(u_N, v) = l(v), \quad \forall v \in V_N,$$

variációs egyenleteket V_N -beli függvényekre alkalmazva:

$$a(u - u_N, v) = 0, \quad \forall v \in V_N.$$

Írjuk fel v -t $v_N - u_N$ alakban, ahol v_N tetszőleges V_N -beli elem.

Céa-lemma

Ekkor kapjuk az ún. Galjorkin-ortogonalitás képletét:

$$a(u - u_N, v_N - u_N) = 0, \quad \forall v_N \in V_N.$$

Továbbá

$$\begin{aligned} \alpha \|u - u_N\|^2 &\leq a(u - u_N, u - u_N) \\ &= a(u - u_N, u - v_N) + \underbrace{a(u - u_N, v_N - u_N)}_{=0} \leq \gamma \|u - u_N\| \|u - v_N\|, \end{aligned}$$

amiből következik a lemma állítása. ■

Köv.: $u_N \rightarrow u$, ha $\forall u \in V$ -re $\text{dist}(u, V_N) \rightarrow 0$ $N \rightarrow \infty$ esetén.

Megj.: A végeelem módszernél v -t az u függvény interpolációs közelítésének választva a Céa-lemma segítségével becsülhető a hiba.

12.

A végeelem módszer konstrukciója

A megoldáshalmaz kiterjesztése, gyenge egyenlet

Olyan alakra kell átírni a megoldandó feladatot, melyre a Lax–Milgram-tétel elmélete alkalmazható. Ehhez

- ▶ ki kell bővíteni a megoldás terét Hilbert-térre,
- ▶ az egyenletet át kell írni $a(u, v) = l(v)$ alakra,
- ▶ és garantálni kell tudni a Lax–Milgram-tétel feltételeit.

A megoldáshalmaz kiterjesztése, gyenge egyenlet

Tekintsük az alábbi peremérték-feladatot mintafeladatként:

Reakció-diffúziós egyenlet homogén Dirichlet peremmel

$$-\Delta u + u = f, \quad (x, y) \in \Omega, \quad u|_{\partial\Omega} = 0$$

Keressük az $u \in C^2(\Omega) \cap C(\bar{\Omega})$ megoldásfüggvényt.

Tegyük fel, hogy $u \in C^1(\bar{\Omega})$ is igaz. Vegyünk egy tetszőleges $v \in C_0^1(\bar{\Omega})$ függvényt (peremen nulla), szorozzunk vele, és integráljuk a két oldalt Ω -n!

$$-\int_{\Omega} \Delta u v + \int_{\Omega} u v = \int_{\Omega} f v,$$

alkalmazva a Green-formulát az alábbi ún. variációs egyenletet nyerjük:

$$\int_{\Omega} \nabla u \nabla v + \int_{\Omega} u v = \int_{\Omega} f v, \quad \forall v \in C_0^1(\bar{\Omega}).$$

Ennek teljesüléséhez elég a függvényektől gyengébb simasági feltételeket megkövetelni.

36. Def. Tekintsük a $C^k(\overline{\Omega})$ halmazt, és lássuk el az

$$\langle u, v \rangle_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \int_{\Omega} (\partial^{\alpha} u)(\partial^{\alpha} v)$$

skaláris szorzattal (ez valóban skaláris szorzat). Ez normát is definiál:

$$\|u\|_{H^k(\Omega)} = \sqrt{\sum_{|\alpha| \leq k} \int_{\Omega} (\partial^{\alpha} u)^2}.$$

A fenti teret a fenti normában teljessé téve Hilbert-teret kapunk. Jelölése $H^k(\Omega)$ (Szoboljev-tér).

37. Tétel. $H^k(\Omega)$ izomorf azon $L_2(\Omega)$ -beli (mérhető, négyzetesen integrálható függvények, $\int f^2 < \infty$) u függvények terével, melyek minden legfeljebb k -adrendű parciális deriváltja is négyzetesen integrálható, a skaláris szorzat a fenti módon van értelmezve, és van olyan $C^k(\overline{\Omega})$ -beli függvénysorozat, ami a fenti normában az u függvényhez tart.

Hasonlóan értelmezhető a $H_0^k(\Omega)$ tér is a $C_0^k(\overline{\Omega})$ teret teljessé téve a korábbi normában.

38. Tétel. $H_0^k(\Omega)$ izomorf azon $L^2(\Omega)$ -beli u függvények terével, melyek minden legfeljebb k -adrendű parciális deriváltja is négyzetesen integrálható, a skaláris szorzat a fenti módon van értelmezve, és van olyan $C_0^k(\overline{\Omega})$ -beli függvénysorozat, ami a korábbi normában az u függvényhez tart.

Szoboljev-terek

$H^1(\Omega)$ fontos altere a

$$H_D^1 = \{u \in H^1(\Omega) \mid u|_{\partial\Omega_D} = 0\}$$

tér, ahol $\partial\Omega_D \subset \partial\Omega$ az ún. Dirichlet-perem (Dirichlet-peremfeltétel lesz rajta előírva), ami a perem egy pozitív mértékű, szakaszonként sima részhalmaza.

39. Tétel. (Beágyazási tételek) (Def.: $V_1 \hookrightarrow V_2$, ha $V_1 \subset V_2$, és van olyan $C > 0$, hogy $\|v\|_{V_2} \leq C\|v\|_{V_1} \forall v \in V_1$.)

Legyen $\Omega \subset \mathbb{R}^d$ korlátos, nyílt tartomány, nemüres Lipschitz-peremmel.

- ▶ ha $k < d/2$, akkor $H^k(\Omega) \hookrightarrow L^q(\Omega)$ ($q \leq 2d/(d - 2k)$),
- ▶ ha $k = d/2$, akkor $H^k(\Omega) \hookrightarrow L^q(\Omega)$ ($q \in [2, \infty)$),
- ▶ ha $k > d/2$, akkor $H^k(\Omega) \hookrightarrow C^0(\overline{\Omega})$

Köv.: Ha Ω egydimenziós, akkor $H^1(\Omega)$ -ban csak folytonos függvények (osztályai) vannak.

40. Tétel. (Poincaré–Friedrichs-egyenlőtlenség) Van olyan $C_\Omega > 0$ csak az Ω tartománytól függő konstans, hogy

$$\int_{\Omega} u^2 \leq C_\Omega^2 \int_{\Omega} |\nabla u|^2 \quad (\forall u \in H_D^1(\Omega)).$$

Bevezetjük az alábbi ún. fél- vagy szeminormákat:

$$|u|_{H^k(\Omega)} = \sqrt{\sum_{|\alpha|=k} \int_{\Omega} (\partial^\alpha u)^2}.$$

Szoboljev-terek

Hasznos kiírni a nevezetes normákat néhány speciális tér esetén (két térdimenzióban):

$$\|u\|_{H^1(\Omega)} = \sqrt{\int_{\Omega} \left(u^2 + \underbrace{\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2}_{|\nabla u|^2} \right) dx dy,}$$

$$|u|_{H^1(\Omega)} = \sqrt{\int_{\Omega} \left(\underbrace{\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2}_{|\nabla u|^2} \right) dx dy,}$$

$$|u|_{H^2(\Omega)} = \sqrt{\int_{\Omega} \left(\left(\frac{\partial^2 u}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 u}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 u}{\partial y^2} \right)^2 \right) dx dy.}$$

Szoboljev-terek

Így pl.

$$\|u\|_{H^1(\Omega)}^2 = \|u\|_{L_2(\Omega)}^2 + |u|_{H^1(\Omega)}^2,$$

a Poincaré–Friedrichs-egyenlőtlenség alakja pedig:

$$\|u\|_{L_2(\Omega)} \leq C_\Omega |u|_{H^1(\Omega)} \quad (\forall u \in H_D^1(\Omega)).$$

A megoldáshalmaz kiterjesztése, gyenge egyenlet

Térjünk vissza a korábbi variációs feladathoz!

$$\int_{\Omega} \nabla u \nabla v + \int_{\Omega} uv = \int_{\Omega} fv, \quad \forall v \in C_0^1(\bar{\Omega}).$$

A korábbiak alapján ez a feladat bővebb függvénytéren is értelmezhető: Keressük azt az $u \in H_0^1(\Omega)$ -beli függvényt, amelyre a fenti egyenlőség igaz tetszőleges $v \in H_0^1(\Omega)$ függvény esetén (f -ről is elég, hogy $f \in L_2(\Omega)$). Ez a differenciálegyenlet ún. gyenge alakja (gyenge egyenlet).

A minimalizációs feladat megfogalmazása: Keressük azt az $u \in H_0^1(\Omega)$ -beli függvényt, ami minimalizálja az

$$I(u) = \int_{\Omega} \left(\frac{1}{2} ((\nabla u)^2 + u^2) - uf \right)$$

funkcionált.

A megoldáshalmaz kiterjesztése, gyenge egyenlet

Legyen $V = H_0^1(\Omega)$, és ellenőrizzük le a Lax–Milgram-tétel feltételeit az

$$a(u, v) = \int_{\Omega} (\nabla u \nabla v + uv),$$

és

$$l(v) = \int_{\Omega} f v$$

leképezésekre!

$a(u, v)$ bilineáris (szimmetrikus is), valamint folytonos, mert

$$|a(u, v)| = \left| \int_{\Omega} (\nabla u \nabla v + uv) \right| = |\langle u, v \rangle| \leq \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)},$$

azaz $\gamma = 1$ megfelelő konstans, ill. koercív, mert

$$|a(u, u)| = \left| \int_{\Omega} (\nabla u \nabla u + uu) \right| = \|u\|_{H^1(\Omega)}^2 \geq \|u\|_{H^1(\Omega)}^2$$

azaz $\alpha = 1$ megfelelő konstans.

A megoldáshalmaz kiterjesztése, gyenge egyenlet

$l(v)$ lineáris, és korlátos is, mert

$$|l(v)| = \left| \int_{\Omega} f v \right| \leq \|f\|_{L_2(\Omega)} \|v\|_{L_2(\Omega)} \leq \|f\|_{L_2(\Omega)} \|v\|_{H^1(\Omega)},$$

azaz $\Lambda = \|f\|_{L_2(\Omega)}$ megfelelő konstans.

A fentiek szerint az $a(u, v) = l(v)$ ($\forall v \in V$) variációs feladatnak egyértelműen létezik u megoldása V -n, azaz $H_0^1(\Omega)$ -n.

A Galjorkin-módszer alkalmazása

Választunk egy $V_h \subset V = H_0^1(\Omega)$ alteret, melyet generáljanak a ϕ_1, \dots, ϕ_N bázisfüggvények! Ezen a téren keressük a

$$\int_{\Omega} \nabla u_h \nabla v_h + \int_{\Omega} u_h v_h = \int_{\Omega} f v_h, \quad \forall v_h \in V_h$$

feladat (ún. vetületi egyenlet) u_h megoldását (Lax–Milgram miatt egyértelműen létezik megoldás). Ezt a megoldást $u_h = \sum_{i=1}^N c_i \phi_i$ alakban keresve a c_i együtthatók meghatározhatók az alábbi egyenletrendszer megoldásával:

$$\mathbf{K}_h \bar{\mathbf{c}}_h = \bar{\mathbf{b}}_h,$$

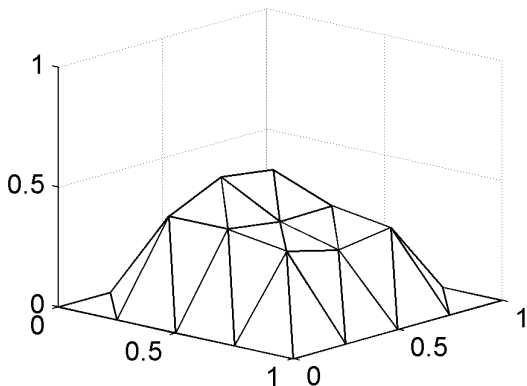
ahol \mathbf{K}_h a merevségi mátrix, $\bar{\mathbf{b}}_h$ pedig a terhelési vektor.

$$(\mathbf{K}_h)_{ij} = \int_{\Omega} (\nabla \phi_j \nabla \phi_i + \phi_j \phi_i),$$

$$(\bar{\mathbf{b}}_h)_i = \int_{\Omega} f \phi_i$$

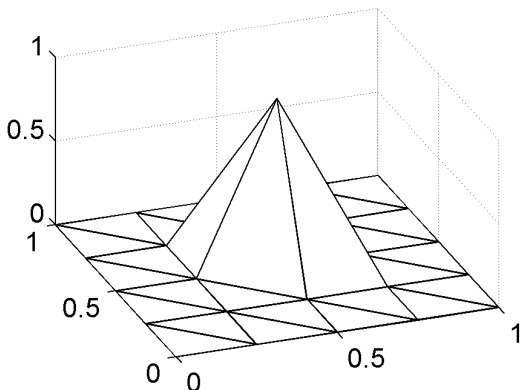
A Galjorkin-módszer alkalmazása

\mathbf{K}_h szimmetrikus, pozitív definit mátrix. Ahhoz, hogy ritka mátrix legyen, alkalmas módon kell választani a ϕ_i bázisfüggvényeket.



A Galjorkin-módszer alkalmazása

Pl.: Courant-elemek: Háromszögekre bontjuk a tartományt, és minden tartományon lineáris függvénnyel közelítünk. Ahhoz, hogy $H_0^1(\Omega)$ -beli függvényt kapjunk, meg kell követelni a folytonosságot, és hogy a függvények a peremen nullák legyenek. Így pl. a belső pontokhoz tartozó alábbi ún. sátorfüggvények jók lesznek bázisfüggvényeknek.



A Galjorkin-módszer alkalmazása

Szabályos háromszögrácsot választunk, és célunk a merevségi mátrix meghatározása.

Ehhez meghatározzuk az ún. elemi merevségi mátrixot egy $P_1(0, 1)$, $P_2(h, 0)$, $P_3(0, h)$ csúcsú R referenciaháromszögön.

A referenciaelemhez tartozó bázisfüggvények:

$$\Psi_1(x, y) = -\frac{x}{h} - \frac{y}{h} + 1, \quad \Psi_2(x, y) = \frac{x}{h}, \quad \Psi_3(x, y) = \frac{y}{h},$$

Ψ_i P_i -ben 1, a többi csúcsban 0, és lineáris.

Az elemi gradienstorozat-integrálok mátrixa ($\int_R \nabla \Psi_j \nabla \Psi_i$):

$$\begin{bmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1/2 & 0 \\ -1/2 & 0 & 1/2 \end{bmatrix}$$

A Galjorkin-módszer alkalmazása

Az elemi szorzat-integrálok mátrixa ($\int_R \Psi_j \Psi_i$):

$$\frac{h^2}{24} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

A referenciaelemhez tartozó elemi merevségi mátrix a fenti mátrixok segítségével már könnyen összeállítható.

A terhelési vektorhoz szükséges integrálokat pedig közelíthetjük az alábbi módon (elegendően magas rendű közelítés lesz):

$$\int_{\Omega} f \phi_i \approx h^2 f(P_i)$$

($f(P_i)$, az f függvény P_i pontbeli értéke).

A fenti módon megkonstruált közelítés hibájáról később lesz majd szó.

A Galjorkin-módszer alkalmazása

Megj.: Ha a referenciaelem egy tetszőleges háromszög, akkor az elemi gradiensszorzat-integrálok és szorzatintegrálok mátrixai a következők lesznek (lineáris elemek esetén):

$$\int_R \nabla \Psi_j \nabla \Psi_i = \begin{cases} \frac{1}{m_i} \text{meas}(R), & \text{ha } i = j, \\ -\frac{\text{ctg} \alpha_{ij}}{2}, & \text{ha } i \neq j, \end{cases}$$

ahol m_i a P_i csúcsból induló magasságvonal hossza, α_{ij} pedig a $P_i P_j$ oldallal szembeni szög, valamint

$$\int_R \Psi_j \Psi_i = \frac{1 + \delta_{ij}}{12} \text{meas}(R).$$

(d -dimenzióban

$$\int_R \Psi_j \Psi_i = \frac{(1 + \delta_{ij})d!}{(d + 2)!} \text{meas}(R).)$$

A Galjorkin-módszer alkalmazása

A gradiensek szorzatainak integrálja az alábbi módon számolható:

Könnyen igazolható, hogy ha $P_i(x_i, y_i)$ ($i = 1, 2, 3$) az R referenciaháromszög csúcsai, akkor az i -edik elemi bázisfüggvény (azaz, az a lineáris függvény, ami P_i -ben 1-et vesz fel, a másik két pontban pedig 0-t) az alábbi alakban írható fel:

$\Psi_i(x, y) = a_i + b_i x + c_i y$, ahol

$$a_i = \frac{x_j y_k - x_k y_j}{2 \text{area}(R)}, \quad b_i = \frac{y_j - y_k}{2 \text{area}(R)}, \quad c_i = \frac{x_k - x_j}{2 \text{area}(R)}.$$

Innét látható, hogy $\nabla \Psi_i(x, y) = (b_i, c_i)$. Ennek segítségével az integrálok már számíthatók.

Összefoglalás

Az eddig tárgyalt mintafeladat alapján felírhatjuk a végeelem módszer megvalósításának lépéseit.

- ▶ Az egyenletből megkonstruáljuk a variációs egyenleteket (gyenge alak) vagy a minimalizálandó funkcionált.
- ▶ A megoldási tartományt szakaszokra, háromszögekre, stb. bontjuk.
- ▶ Választunk egy végeelem teret a közelítésre (kis tartójú, szakaszonként sima függvények). Hibabecslést adunk.
- ▶ Meghatározzuk az elemi merevségi mátrixot.
- ▶ Összeállítjuk a merevségi mátrixot és a terhelési vektort.
- ▶ Megoldjuk az egyenletrendszert.
- ▶ Szemléltetjük a megoldást.

Poisson-egyenlet inhomogén Dirichlet-peremmel

Reakció-diffúziós egyenlet inhomogén Dirichlet-peremmel

$$-\Delta u + u = f, \quad (x, y) \in \Omega, \quad u|_{\partial\Omega} = g.$$

Terjesszük be a teljes tartományra a g függvényt, és írjuk fel a megoldást $u = u^0 + g$ alakban. Ekkora nyilván u^0 a peremen nullát vesz fel. Így az egyenlet alakja (a homogén esetnél tárgyaltak szerint)

$$-\Delta u^0 + u^0 = f + \Delta g - g, \quad (x, y) \in \Omega, \quad u^0|_{\partial\Omega} = 0.$$

Ez gyenge alakban:

$$\int_{\Omega} (\nabla u^0 \nabla v + u^0 v) = \int_{\Omega} (f + \Delta g - g)v, \quad \forall v \in H_0^1(\Omega).$$

Poisson-egyenlet inhomogén Dirichlet-peremmel

A g -t tartalmazó tagokat visszavisszük a bal oldalra.

$$\int_{\Omega} (\nabla(u^0 + g)\nabla v + (u^0 + g)v) = \int_{\Omega} f v, \quad \forall v \in H_0^1(\Omega).$$

Így a gyenge alak: Keressük azt az $u \in H^1(\Omega)$, $u|_{\partial\Omega} = g$ függvényt, amelyre

$$\int_{\Omega} (\nabla u \nabla v + uv) = \int_{\Omega} f v, \quad \forall v \in H_0^1(\Omega).$$

A Lax–Milgram-tétel miatt egyértelmű megoldás van.

A Galjorkin-módszeres megoldásnál nem csak a belső pontokhoz, hanem a perempontokhoz is kell bázisfüggvényt rendelni. Ezekkel közelítjük tulajdonképpen a peremfeltételt.

Neumann-feladat esete

Reakció-diffúziós egyenlet Neumann-peremmel

$$-\Delta u + u = f, \quad (x, y) \in \Omega, \quad \frac{\partial u}{\partial n} = g, \quad (x, y) \in \partial\Omega$$

Tegyük fel, hogy $u \in C^1(\bar{\Omega})$. Vegyünk egy tetszőleges $v \in C^1(\bar{\Omega})$ függvényt, szorozzunk vele, és integráljuk a két oldalt Ω -n.

$$-\int_{\Omega} \Delta uv + \int_{\Omega} uv = \int_{\Omega} fv,$$

alkalmazva a Green-formulát az alábbi variációs egyenletet nyerjük:

$$-\int_{\partial\Omega} \nabla uv n \, ds + \int_{\Omega} \nabla u \nabla v + \int_{\Omega} uv = \int_{\Omega} fv, \quad \forall v \in C^1(\bar{\Omega}),$$

azaz

$$\int_{\Omega} \nabla u \nabla v + \int_{\Omega} uv = \int_{\Omega} fv + \int_{\partial\Omega} gv \, ds, \quad \forall v \in C^1(\bar{\Omega}).$$

Neumann-feladat esete

Gyenge alak tehát: Keressük azt az $u \in H^1(\Omega)$ függvényt, mellyel

$$\int_{\Omega} \nabla u \nabla v + \int_{\Omega} uv = \int_{\Omega} fv + \int_{\partial\Omega} gv \, ds, \quad \forall v \in H^1(\Omega).$$

A minimalizációs feladat: Keressük azt az $u \in H^1(\Omega)$ függvényt, ami minimalizálja az

$$I(u) = \int_{\Omega} \left(\frac{1}{2} ((\nabla u)^2 + u^2) - uf \right) - \int_{\partial\Omega} gu$$

funkcionált.

A Galjorkin-módszeres közelítésnél a perempontokhoz tartozó rácspontokban is veszünk fel bázisfüggvényeket, és a megoldás előállításában ezek szorzói is meghatározandók.

A fenti feladatok könnyen átfogalmazhatók arra az esetre is, ha a tartomány peremének csak egy részén van Dirichlet- vagy Neumann-feltétel.

Végelem példa

Oldjuk meg az alábbi feladatokat a végelem-módszer segítségével!

Homogén Dirichlet-perem:

$$-\Delta u = -4, \quad (x, y) \in \Omega, \quad u = 0, \quad (x, y) \in \partial\Omega$$

Ω az egységnyégyzet, melyből el van távolítva az $(1/2, 1/2)$ középpű, $1/4$ sugarú kör.

Inhomogén Dirichlet-perem:

$$-\Delta u = -4, \quad (x, y) \in \Omega,$$
$$u = (x - 1/2)^2 + (y - 1/2)^2, \quad (x, y) \in \partial\Omega$$

Végeselem példa

Gyenge alakok.

Keressük azt az $u \in H_0^1(\Omega)$ függvényt, melyre

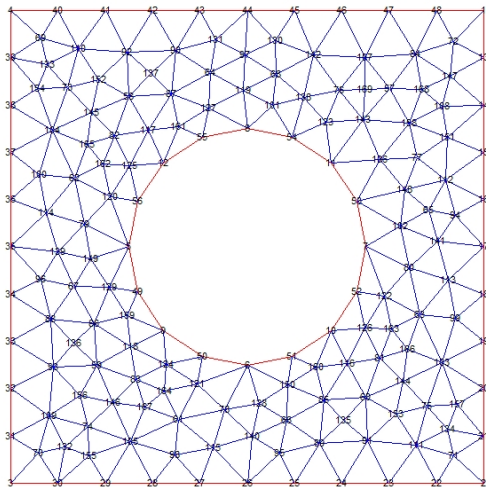
$$\int_{\Omega} \nabla u \nabla v = - \int_{\Omega} 4v, \quad \forall v \in H_0^1(\Omega)$$

Keressük azt az $u \in H^1(\Omega)$, $u|_{\partial\Omega} = (x - 1/2)^2 + (y - 1/2)^2$ függvényt, melyre

$$\int_{\Omega} \nabla u \nabla v = - \int_{\Omega} 4v, \quad \forall v \in H_0^1(\Omega)$$

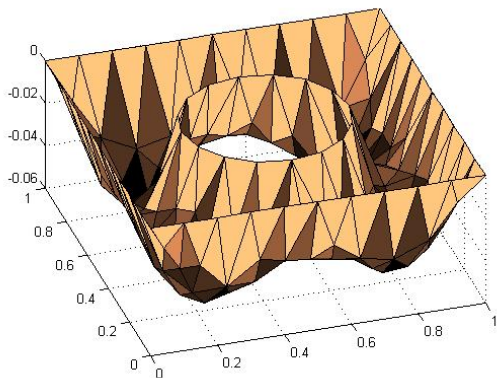
Végelem példa

Megkonstruáljuk a tartomány háromszögelését, amiből (szakaszonként lineáris bázisfüggvényeket választva) nyerjük a bázisfüggvényeket.

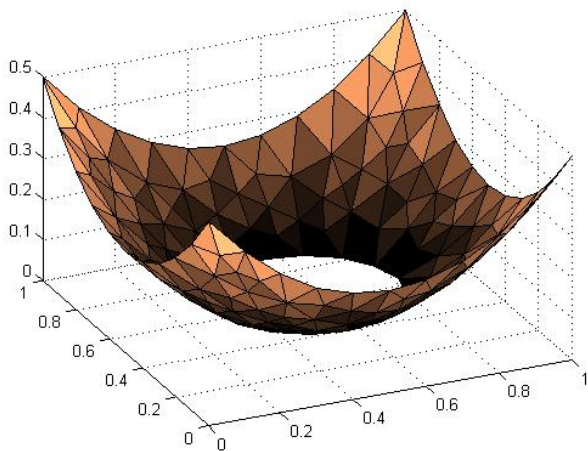


Végeselem példa

A bázisfüggvényekből meghatározható a merevségi mátrix ill. a terhelési vektor (háromszögenként haladva töltjük fel ezeket). Megoldjuk az egyenletrendszert a bázisfüggvények együtthatóira, majd ábrázoljuk a megoldást.



Végeselem példa



Lényeges és természetes peremfeltételek

Def.: Egy peremfeltétel lényeges, ha azt a tesztfüggvényeknek is ki kell elégíteniük, különben természetes peremfeltételről beszélünk.

Vegyük észre, hogy a Neumann-feltétel nem ad megkötést u megválasztására $H^1(\Omega)$ -ból, a peremfeltételt csak implicit módon tartalmazza a variációs egyenlet. A korábbi Dirichlet-peremfeltétel lényeges peremfeltétel volt.

Megj.: Általában k -adrendű differenciálegyenlet esetén a $k/2$ -nél kisebb rendű deriváltak lényeges peremfeltételt adnak, az ennél nagyobbak pedig természetest.

13.

Néhány végeselem tér

A végeelem tér definíciója

Def.: A (K, P_K, Σ) rendezett hármast végeelem térnek hívjuk, ha K valamilyen geometriai alakzat, P_K egy K -n értelmezett függvényekből álló véges dimenziós vektortér, és ΣP_K függvényeinek szabadságfokainak halmaza (P_K minden eleme egyértelműen meghatározott a Σ -beli szabadságfokokkal).

Megj.:

1. K általában szakasz, háromszög, téglalap, tetraéder, téglatest.
2. P_K általában polinomokat tartalmaz ($P_r(K)$: legfeljebb r -edfokú polinomok halmaza K -n).
3. A szabadságfokok a polinomok és deriváltjainak értékei bizonyos pontokban.

Regularitási követelmények

A V_h tér olyan v függvények vektortere, melyekre $v|_K \in P_r(K)$. A K halmazok belsejeinek nincs közös pontja, lezártjaik uniója Ω (korlátos). Egyik K halmaz csúcsa sem eshet egy másik élére vagy lapjára.

Könnyen látható, hogy

$$V_h \subset H^1(\Omega) \Leftrightarrow V_h \subset C^0(\bar{\Omega})$$

és

$$V_h \subset H^2(\Omega) \Leftrightarrow V_h \subset C^1(\bar{\Omega}).$$

Néhány konkrét végelem tér

- ▶ (1D szakaszok, P_1 , fv.-értékek a végpontokban)
- ▶ (1D szakaszok, P_2 , fv.-értékek a végpontokban és középén)
- ▶ (1D szakaszok, P_3 , fv. értékek és deriváltak a két szélső pontban)
- ▶ (2D háromszögek, P_1 , fv. értékek a csúcspontokban), Courant-elemek.
- ▶ (2D háromszögek, P_2 , fv. értékek a csúcspontokban, oldalfelezőkön)
- ▶ (2D téglalapok, Q_1 , fv. értékek a csúcspontokban)
- ▶ (3D tetraéderek, P_1 , fv. értékek a csúcspontokban)
- ▶ stb.

(Gondoljuk át, hogy mely pontokban kell megadni a függvényértékeket ill. a deriváltakat a bázisfüggvények megadásához!)

14.

Végelelem hibabecslések

Előzmények

Gyakorlaton láttuk az egydimenziós $-u'' = f$, $u(0) = u(L) = 0$ peremértékfeladat esetén, hogy ha a megoldás $H^2(0, L)$ -beli, akkor igaz, hogy

$$\|u - u_h\|_{H^1(0,L)} \leq ch|u|_{H^2(0,L)},$$

ahol a c konstans $c = \gamma\sqrt{1 + L^2}/\alpha$. Ez $H^1(0, L)$ normában elsőrendű konvergenciát mutat.

L_2 normában másodrendű a konvergencia (Nitsche-trükk):

$$\|u - u_h\|_{L_2(0,L)} \leq \gamma c^2 c_1 |u|_{H^2(0,L)} h^2.$$

(A kiegészítő egyenletről felteendő feltételek: megoldása $z \in H^2(0, L)$, folytonos függés a jobboldaltól $|z|_{H^2(0,L)} \leq c_1 \|u - u_h\|_{L_2(0,L)} \cdot$)

Hibabecslés általában

41. Tétel. Ha az interpolációra szakaszonként r -edfokú polinomokat használunk egy poligonális tartomány egy reguláris háromszögrácsán (szakasz, háromszög, tetraéder), és a rácsfinomítás során a

$$(\text{maximális élhossz}) / (\text{beírható gömb sugara})$$

korlátos marad, akkor minden $v \in H^{r+1}(\Omega)$ esetén

$$\|v - \Pi v\|_{L_2(\Omega)} \leq Ch^{r+1} |v|_{H^{r+1}(\Omega)},$$

$$|v - \Pi v|_{H^1(\Omega)} \leq Ch^r |v|_{H^{r+1}(\Omega)},$$

ahol Π az interpolációs operátor, amely a v függvényhez a rácpontokbeli függvényértékekre felírt interpolációs polinomját rendeli. C minden esetben (h -tól és a rácstól független) valamilyen konstans értéket jelöl (nem mindig ugyanazt). Ha v csak $H^s(\Omega)$ -beli, akkor a fenti képletekben $r + 1$ helyett s , r helyett $s - 1$ írandó.

Hibabecslés általában

Köv.: A Céa-lemma miatt az alábbi becslést nyerjük a Galjorkin-módszeres numerikus megoldásra:

$$\begin{aligned}\|u - u_h\|_{H^1(\Omega)} &\leq \frac{\gamma}{\alpha} \inf_{v \in V_h} \|u - v\|_{H^1(\Omega)} \\ &\leq C \|u - \Pi u\|_{H^1(\Omega)} \leq Ch^r |u|_{H^{r+1}(\Omega)},\end{aligned}$$

továbbá nyilvánvalóan

$$\|u - u_h\|_{L_2(\Omega)} \leq Ch^r |u|_{H^{r+1}(\Omega)}.$$

(Már 1D-ben is láttuk, hogy a Nitsche-trükkal eggyel nagyobb rendű konvergencia is igazolható. Mutassuk ezt meg most is!)

Pl.: 1D, folytonos, szakaszonként elsőfokú bázisfüggvények esetén, $r = 1$ -gyel érvényesek a korábbi becslések.

Pl.: 1D, folytonos, szakaszonként másodfokú bázisfüggvények esetén, $r = 2$ -vel érvényesek a korábbi becslések.

Hibabecslés L_2 -ben

Tekintsük az alábbi ún. kiegészítő (vagy adjungált) egyenletet!

$$a(z_\psi, v) = \int_{\Omega} \psi v, \quad \forall v \in V,$$

ahol $\psi \in L_2(\Omega)$ tetszőleges adott függvény! Tegyük fel, hogy az eredeti feladat megoldására

- a) $u \in H^{r+1}(\Omega)$, valamint a fenti kiegészítő egyenlet megoldására
- b) $z_\psi \in H^2(\Omega)$ és
- c) $\|z_\psi\|_{H^2(\Omega)} \leq C\|\psi\|_{L_2(\Omega)}$, $\forall \psi \in L_2(\Omega)$ (folytonos függés a jobb oldaltól; megfelelően szép tartomány és peremfeltétel kell hozzá).

42. Tétel. (Nitsche-trükk) A korábbi és a fenti feltételek mellett

$$\|u - u_h\|_{L_2(\Omega)} \leq Ch^{r+1}|u|_{H^{r+1}(\Omega)}.$$

Hibabecslés L_2 -ben

Biz.: Legyen $e = u - u_h$ és tekintsük a kiegészítő egyenletet $\psi = e$ és $v = e$ esetén: $a(z_e, e) = \int_{\Omega} ee$.

Így tehát a kiegészítő egyenlet v_h Galjorkin-módszeres megoldásával:

$$\|u - u_h\|_{L_2(\Omega)}^2 = \int_{\Omega} ee = a(z_e, e) = a(z_e - \underbrace{v_h}_{\substack{\text{tetsz. } v_h \in V_h \\ \text{Galj. ort.}}}, e)$$

$$\begin{aligned} &\leq \gamma \|z_e - v_h\|_{H^1(\Omega)} \|e\|_{H^1(\Omega)} \\ &\leq \gamma Ch |z_e|_{H^2(\Omega)} h^r |u|_{H^{r+1}(\Omega)} \\ &\leq \gamma Ch \|e\|_{L_2(\Omega)} h^r |u|_{H^{r+1}(\Omega)} \\ &= \gamma Ch^{r+1} \|u - u_h\|_{L_2(\Omega)} |u|_{H^{r+1}(\Omega)}, \end{aligned}$$

ebből pedig osztással már következik az állítás. ■

15.

Időfüggő másodrendű lineáris differenciálegyenletek végesesemes megoldása

Hővezetési egyenlet

Hővezetési egyenlet, homogén Dirichlet-peremfeltétel

$$\partial_t u - \Delta u = f, \quad (t, x) \in (0, T) \times \Omega,$$

$$u|_{(0, T) \times \partial\Omega} = 0,$$

$$u(0, \bar{x}) = u_0(\bar{x}) \text{ adott.}$$

Az elliptikus esethez hasonlóan szorozzuk az egyenlet mindkét oldalát egy tetszőleges, megfelelően sima $v(\bar{x})$ tesztfüggvénnyel (peremen nulla), majd integráljuk mindkét oldalt Ω -n!

$$\int_{\Omega} \partial_t uv - \int_{\Omega} \Delta uv = \int_{\Omega} f v,$$

$$\int_{\Omega} \partial_t uv + \int_{\Omega} \nabla u \nabla v = \int_{\Omega} f v.$$

Gyenge alak

Így tehát az alábbi gyenge alakhoz jutunk: Keressük minden $t \in (0, T)$ esetén azt az $u(t) \in H_0^1(\Omega)$ függvényt, melyre

$$\int_{\Omega} \partial_t u v + \int_{\Omega} \nabla u \nabla v = \int_{\Omega} f v$$

minden $v \in H_0^1(\Omega)$ tesztfüggvény esetén, továbbá $u(0) = u_0(\bar{x})$.

Ebből a gyenge alakból származtathatjuk az ún. szemidiszkrét (a térbeli diszkrétizációt a Galjorkin végeelem módszer adja) alakját az egyenletnek:

Legyen V_h a szokott módon egy véges dimenziós altere $H_0^1(\Omega)$ -nak a ϕ_1, \dots, ϕ_N bázisfüggvényekkel. Keressük minden $t \in (0, T)$ esetén azt az $u_h(t) \in V_h$ függvényt, melyre

$$\int_{\Omega} \partial_t u_h v + \int_{\Omega} \nabla u_h \nabla v = \int_{\Omega} f v,$$

továbbá $\int_{\Omega} u_h(0) v = \int_{\Omega} u_0 v$ minden $v \in V_h$ tesztfüggvény esetén.

A szemidiszkrét megoldás előállítás

Keressük $u_h(t)$ -t $u_h(t) = \sum_{i=1}^N c_i(t)\phi_i(\bar{\mathbf{x}})$ alakban! Ezt a szemidiszkrét gyenge egyenletbe egyenletbe helyettesítve, és kihasználva, hogy az egyenletnek minden V_h -beli v -re teljesülnie kell, kapjuk, hogy

$$\int_{\Omega} \partial_t \left(\sum_i c_i(t)\phi_i \right) \phi_j + \int_{\Omega} \nabla \left(\sum_i c_i(t)\phi_i \right) \nabla \phi_j = \int_{\Omega} f \phi_j$$

($j = 1, \dots, N$). Ezt mátrixos alakban felírt KDE-rendszert nyerjük:

$$\mathbf{B}_h \bar{\mathbf{c}}_h'(t) + \mathbf{A}_h \bar{\mathbf{c}}_h(t) = \bar{\mathbf{b}}_h,$$

ahol $(\mathbf{A}_h)_{ij} = a(\phi_j, \phi_i)$ a korábbi merevségi (stiffness) mátrix, $(\bar{\mathbf{b}}_h)_i = \int_{\Omega} f \phi_i$ a terhelési (load) vektor és $(\mathbf{B}_h)_{ij} = \int_{\Omega} \phi_j \phi_i$ az ún. tömeg (mass) mátrix.

A szemidiszkrét megoldás előállítás

A \mathbf{B}_h tömeg mátrix szimmetrikus és pozitív definit. Így az alábbi KDE-rendszert kell megoldanunk:

$$\bar{\mathbf{c}}'_h(t) = -\mathbf{B}_h^{-1} \mathbf{A}_h \bar{\mathbf{c}}_h(t) + \mathbf{B}_h^{-1} \bar{\mathbf{b}}_h.$$

Ezt a differenciálegyenlet-rendszert megoldhatjuk bármely korábban tanult numerikus módszerrel is egy alkalmas Δt időlépést választva (teljes diszkretizáció). Általában ez egy merev rendszer, így ennek megfelelően implicit módszereket kell a megoldáshoz használni.

43. Tétel. Ha u megoldása a kétdimenziós hővezetési egyenletnek egy Ω konvex, poligonális tartományon és $u_h(t)$ megoldása a szemidiszkrét feladatnak lineáris végeeselemes térbeli diszkretizáció esetén, akkor van olyan C pozitív konstans, mellyel

$$\begin{aligned} & \max_{t \in (0, T)} \|u(t) - u_h(t)\|_{L_2(\Omega)} \\ & \leq C \left(1 + |\ln(T/h^2)|\right) \max_{t \in (0, T)} h^2 \|u(t)\|_{H^2(\Omega)}. \end{aligned}$$

Teljes (tér-idő) diszkrétizáció

Alkalmazzuk pl. a θ -módszert a szemidiszkrét feladat megoldására!
Ekkor $\theta \in [0, 1]$ mellett a $\bar{\mathbf{c}}_h^k = [c_1^k, \dots, c_N^k]^T$ közelítő vektorokra
($c_i^k \approx c_i(k\Delta t)$) az alábbi egyenlőségrendszereket kapjuk
($k = 1, \dots, k_{\max}$):

$$\begin{aligned} & \frac{1}{\Delta t} \mathbf{B}_h (\bar{\mathbf{c}}_h^k - \bar{\mathbf{c}}_h^{k-1}) + \theta \mathbf{A}_h \bar{\mathbf{c}}_h^k + (1 - \theta) \mathbf{A}_h \bar{\mathbf{c}}_h^{k-1} \\ & = \theta \bar{\mathbf{b}}_h(k\Delta t) + (1 - \theta) \bar{\mathbf{b}}_h((k-1)\Delta t) =: \bar{\mathbf{b}}_h^{\theta, k}. \end{aligned}$$

Innét a $\bar{\mathbf{c}}_h^k$ vektorok iterációval meghatározhatók. Ezen vektorok elemei pedig megadják az egyes időrétegeken a bázisfüggvények szorzóit.

Nemnegativitás megőrzése

Mi biztosítja, hogy ha $f \geq 0$, $u_0(\bar{\mathbf{x}}) \geq 0$, akkor a megoldásra is a fizikából elvárt $u_h^k = \sum_{i=1}^N c_i^k \phi_i \geq 0$ teljesül?

Elég ehhez, hogy

$$(\mathbf{B}_h + \theta \Delta t \mathbf{A}_h)^{-1} \geq \mathbf{0},$$

és

$$(\mathbf{B}_h + \theta \Delta t \mathbf{A}_h)^{-1} (\mathbf{B}_h - (1 - \theta) \Delta t \mathbf{A}_h) \geq \mathbf{0}$$

teljesüljön.

Nemnegativitás megőrzése

Pl. az egydimenziós feladat lineáris végeselemes megoldása esetén, amikor ekvidisztáns a felosztás:

$$\mathbf{B}_h = \frac{h}{6} \text{tridiag}[1, 4, 1], \quad \mathbf{A}_h = \frac{1}{h} \text{tridiag}[-1, 2, -1].$$

Így az előbbi egyszerű elégséges feltétellel kapjuk, hogy az 1D lineáris végeselemes megoldás nemnegativitásmegőrző lesz, ha $\theta \in [1/3, 1]$ és

$$\frac{1}{6\theta} \leq q \leq \frac{1}{3(1-\theta)}.$$

(Igazolható (Faragó, 1996), hogy ez a feltétel szükséges is tetszőleges felosztás esetén.)

Tanulság

- ▶ A másodrendű Crank–Nicolson-módszer "hiába" feltétel nélkül stabil, a nemnegativitásmegőrzéshez q nem lehet tetsz. nagy (2D és 3D-ben sem!)
- ▶ Alsó korlát is van q -ra! (mass lumping esetén nincs)
- ▶ Implicit Euler-módszer úgy nemnegativitásmegőrző, hogy csak alsó korlát van q -ra.

Megj.: Más, a homogén Dirichlet-peremfeltételtől eltérő peremfeltétel esetén hasonlóan járunk el, mint az elliptikus feladatok esetén.

Hullámegyenlet megoldása

Tekintsük az alábbi hullámegyenletet homogén peremfeltétellel:

Hullámegyenlet

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = f,$$

$u|_{\partial\Omega} = 0$, $u_0(\bar{\mathbf{x}}) = u(0, \bar{\mathbf{x}})$ és $u_1(\bar{\mathbf{x}}) = \partial_t u(0, \bar{\mathbf{x}})$ adottak.

Keressük az u megoldást a $(0, T)$ intervallumon!

Gyenge alak: Keressük minden $t \in (0, T)$ esetén azt az $u(t) \in H_0^1(\Omega)$ függvényt, melyre

$$\int_{\Omega} \partial_t^2 u v + \int_{\Omega} \nabla u \nabla v = \int_{\Omega} f v$$

minden $v \in H_0^1(\Omega)$ tesztfüggvény esetén, továbbá $u(0) = u_0(\bar{\mathbf{x}})$,
 $\partial_t u(0) = u_1(\bar{\mathbf{x}})$.

Szemidiszkrét megoldás

Keressük minden $t \in (0, T)$ esetén azt az $u_h(t) \in V_h$ függvényt, melyre

$$\int_{\Omega} \partial_t^2 u_h v + \int_{\Omega} \nabla u_h \nabla v = \int_{\Omega} f v,$$

továbbá $\int_{\Omega} u_h(0)v = \int_{\Omega} u_0 v$, $\int_{\Omega} \partial_t u_h(t)v = \int_{\Omega} u_1 v$ minden $v \in V_h$ tesztfüggvény esetén.

A szemidiszkrét megoldás előállítás

Keressük $u_h(t)$ -t $u_h(t) = \sum_{i=1}^N c_i(t)\phi_i(\bar{\mathbf{x}})$ alakban! Helyettesítsük ezt a szemidiszkrétizált egyenletbe! Az alábbi egyenletrendszert kapjuk:

$$\int_{\Omega} \partial_t^2 \left(\sum_i c_i(t)\phi_i \right) \phi_j + \int_{\Omega} \nabla \left(\sum_i c_i(t)\phi_i \right) \nabla \phi_j = \int_{\Omega} f \phi_j$$

($j = 1, \dots, N$), ahonnan az alábbi mátrixos alakban felírt KDE-rendszert nyerjük:

$$\mathbf{B}_h \bar{\mathbf{c}}_h''(t) + \mathbf{A}_h \bar{\mathbf{c}}_h(t) = \bar{\mathbf{b}}_h.$$

(A kezdeti feltételt az eredeti kezdeti feltételekből nyerhetünk.)

A szemidiszkrét megoldás előállítás

Meg kell oldani a másodrendű KDE-rendszert az adott kezdeti feltételekkel. Ez hatékonyan megtehető az alábbi módszerrel:

$$\begin{aligned} & \frac{1}{\Delta t} \mathbf{B}_h (\bar{\mathbf{d}}_h^k - \bar{\mathbf{d}}_h^{k-1}) + \theta \mathbf{A}_h \bar{\mathbf{c}}_h^k + (1 - \theta) \mathbf{A}_h \bar{\mathbf{c}}_h^{k-1} \\ &= \theta \bar{\mathbf{b}}_h(k\Delta t) + (1 - \theta) \bar{\mathbf{b}}_h((k-1)\Delta t) =: \bar{\mathbf{b}}_h^{\theta, k}, \\ & \frac{\bar{\mathbf{c}}^k - \bar{\mathbf{c}}^{k-1}}{\Delta t} - (\alpha \bar{\mathbf{d}}^k + (1 - \alpha) \bar{\mathbf{d}}^{k-1}) = 0, \end{aligned}$$

ahol $\bar{\mathbf{c}}^0$ és $\bar{\mathbf{d}}^0$ a kezdeti feltételekből ismert, és $\alpha \in [0, 1]$ valós paraméter.

A módszer $\theta, \alpha \geq 1/2$ esetén feltétel nélkül stabil, és pl. ha mindkettő $1/2$, akkor másodrendű.