ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι

ΔΙΑΛΕΞΗ 16, 18/12/09

Ε. ΓΑΛΛΟΠΟΥΛΟΣ
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Παν/μιο Πατρών

Θέματα

- Πώς φτιάχνουμε το πλέγμα που διακριτοποιεί το χωρίο ορισμού;
- Πώς προσεγγίζουμε τις παραγώγους;
- Πώς λύνουμε τις προκύπτουσες εξισώσεις;
- Πόσο κοντά είναι η λύση των εξισώσεων από την πραγματική λύση;

ΠΡΟΣΟΧΗ Στις μεθόδους πεπερασμένων διαφορών:

- Η εξαρτημένη μεταβλητή, διακριτοποιείται και προσεγγίζεται το διάνυσμα των τιμών της στους κόμβους του επιλεγμένου πλέγματος. Το μέγεθος του διανύσματος είναι περίπου ίσο με τον αριθμό των εξαρτημένων μεταβλητών επί το πλήθος των κόμβων, άρα και από τη λεπτότητα του πλέγματος[1] Οι τιμές του διανύσματος υπολογίζονται από το διακριτό σύστημα με το οποίο αντικαθίσταται η διαφορική εξίσωση.
- Από τη διακριτή μορφή των παραγώγων και ολοκληρωμάτων (αν θυμάστε από την αριθμητική ολοκλήρωση, π.χ. κανόνας Simpson) μπορούμε να συμπεράνουμε ότι, συνήθως:
    1. Οι εξισώσεις που προέρχονται από την εφαρμογή της μεθόδου πεπερασμένων διαφορών σε διαφορικές εξισώσεις είναι αραιά συστήματα.
    2. Οι εξισώσεις που προέρχονται από διακριτοποίηση ολοκληρωματικών εξισώσεων συνήθως αντιστοιχούν σε πυκνά συστήματα.
- $(\Delta_+ u)_k = u_{k+1} - u_k$,
- $(\Delta_- u)_k = u_k - u_{k-1}$,
- $(\Delta_0 u)_k = u_{k+\frac{1}{2}} - u_{k-\frac{1}{2}}$,
- $(\mathcal{S} u)_k = u_{k+1}$,
- $(\mathcal{M} u)_k = \frac{1}{2}(u_{k+\frac{1}{2}} + u_{k-\frac{1}{2}})$.

[1] Εδώ θα θεωρηθεί στατικό, αλλά η πρόκληση είναι να αποφασίζεται δυναμικά και να υπάρχει αυτοπροσαρμογή ώστε να επιτευχθεί λύση με ικανοποιητικά μικρό σφάλμα.

λεςτ17 Πεπερασμένες διαφορές Κίνητρο: Ορισμός παραγώγων. Βασικό εργαλείο: Σειρές Taylor

$$
\begin{aligned}
u'(x) &= \lim_{h\to 0} \frac{u(x+h)-u(x)}{h} \\
u'(x) &= \lim_{h\to 0} \frac{u(x)-u(x-h)}{h}
\end{aligned}
$$

αλλά και

$$
u'(x) = \frac{1}{2}\lim_{h\to 0}\left(\frac{u(x+h)-u(x)}{h} + \frac{u(x)-u(x-h)}{h}\right)
$$

Πεπερασμένες διαφορές: 2 δοκιμές βασισμένες στους παραπάνω τρόπους να ορίσουμε την παράγωγο

$$
\begin{aligned}
1) D_h u(x) &:= \frac{u(x+h)-u(x)}{h} \\
2) D_h^2 u(x) &:= \frac{u(x+h)-u(x-h)}{2h}
\end{aligned}
$$

$u(x)$	$u'(x)$	$D_h u(x)$	$D_h^2 u(x)$
c	0	0	0
x	1	1	1
x^2	$2x$	$2x+h$	$2x$
$\sin x$	$\cos x$	$\frac{\sin(x+h)-\sin(x)}{h}$	$\cos x \frac{\sin h}{h}$

Παρατηρήσεις:

- Ο τελεστής πεπερασμένης διαφοράς D_h υπολογίζει σωστά την παράγωγο όταν η συνάρτηση είναι σταθερή ή γραμμική.
- Όταν η συνάρτηση είναι πολυώνυμο 2ου βαθμού, ο D_h προσεγγίζει την παράγωγο με σφάλμα που συμπεριφέρεται όπως το $O(h)$.
- Επιπλέον, ο τελεστής πεπερασμένης διαφοράς D_h^2 υπολογίζει σωστά την παράγωγο όταν η συνάρτηση είναι πολυώνυμο βαθμού 0, 1, ή 2.
- Η παράγωγος της ημιτόνου προσεγγίζεται α) με σφάλμα $O(h)$ με το D_h και σφάλμα $O(h^2)$ με το D_h^2.

ερρορ Το σφάλμα αν χρησιμοποιήσουμε τους παραπάνω τελεστές για παράγωγο 1ης τάξης

$u(x) = c \Rightarrow \; |u'(x) - D_h u(x)| = |u'(x) - D_h^2 u(x)| = 0$

$u(x) = x \Rightarrow \ |u'(x) - D_h u(x)| = |u'(x) - D_h^2 u(x)| = 0$

$u(x) = x^2 \Rightarrow \ |u'(x) - D_h u(x)| = h, \ |u'(x) - D_h^2 u(x)| = 0$

$u(x) = \sin x \Rightarrow \ D_h u(x) - \cos x$ οπότε ...

... μέσω Taylor

$$\sin(x+h) = \sin x + h\cos x - \sin x h^2/2 + O(h^3)$$

οπότε

$$(\sin(x+h) - \sin x)/h = \cos x + O(h)$$

ενώ

$$\frac{(\sin(x+h) - \sin(x-h))}{2h} \ = \ \cos x \, \frac{\sin h}{h}$$

και προσέξτε ότι (πάλι μέσω σειράς Taylor - ειδικότερα την εκδοχή Maclaurin)

$$\frac{\sin h}{h} = 1 + \frac{h^2}{6}\cos\theta, \quad \text{όπου } \theta \in [0, h]$$

επομένως

$$\begin{aligned} |\cos x - \frac{(\sin(x+h) - \sin(x-h))}{2h}| &= |\cos x \frac{h^2}{6}\cos\theta| \\ &= O(h^2), \end{aligned}$$

άρα η προσέγγιση είναι 2ης τάξης.
λεςτ18
Πεπερασμένες διαφορές

Προσεγγίσαμε τη συνάρτηση $u(x)$ από τις τιμές της σε ένα διακριτό υποσύνολο Ω_h του χωρίου ορισμού $\Omega \Rightarrow$ Δειγματοληψία!!!

Προσεγγίσαμε την παράγωγο από γραμμικό συνδυασμό (διαφορές) των τιμών της u στο Ω_h

Προσεγγίζουμε το ολοκλήρωμα με γραμμικό συνδυασμό (άθροισμα) των τιμών της u στο Ω_h

Ανάλυση: με σειρές Taylor
Αν η u διαθέτει τουλάχιστον 4 παραγώγους:

$$u(x_j \pm h) = u(x_j) \pm hu^{(1)}(x_j) + \frac{h^2}{2}u^{(2)}(x_j) \pm \frac{h^3}{6}u^{(3)}(x_j) + \frac{h^4}{24}u^{(4)}(x_j + \theta_j^{\pm} h)$$

όπου $-1 < \theta_j^- < 0 < \theta_j^+ < 1$. Επομένως

$$\begin{aligned} u(x_j-h)+u(x_j+h)-2u(x_j) &= h^2u^{(2)}(x_j)+\frac{h^4}{24}\left(u^{(4)}(x_j+\theta_j^+h)+u^{(4)}(x_j+\theta_j^-h)\right)\\ u(x_j+h)-u(x_j-h) &= 2hu^{(1)}(x_j)+\frac{h^3}{6}\left(u^{(3)}(x_j+\eta_j^+h)+u^{(3)}(x_j+\eta_j^-h)\right) \end{aligned}$$

όπου $-1 < \eta_j^- < 0 < \eta_j^+ < 1$.

Θεώρημα ενδιάμεσης τιμής:

Αν f συνεχής στο $[a,b]$ και $f(a) < \lambda < f(b)$ τότε υπάρχει $x \in (a,b)$ ώστε $f(x) = \lambda$.

Μπορούμε να λάβουμε $\lambda := \frac{f(a)+f(b)}{2}$ και να εφαρμόσουμε το θεώρημα στις $u^{(3)}(x)$ και $u^{(4)}(x)$, οπότε

$$\frac{u^{(4)}(x_j+\theta_j^+h)+u^{(4)}(x_j+\theta_j^-h)}{2} = u^{(4)}(x_j+\theta_jh), |\theta_j| \leq \max\{\theta_j^+,\theta_j^-\} < 1,$$

$$\frac{u^{(3)}(x_j+\eta_j^+h)+u^{(3)}(x_j+\eta_j^-h)}{2} = u^{(3)}(x_j+\eta_jh), |\eta_j| \leq \max\{\eta_j^+,\eta_j^-\} < 1,$$

Τα παραπάνω δείχνουν ότι

$$\begin{aligned} \frac{u(x_j+h)-u(x_j-h)}{2h} &= u_j^{(1)}+\frac{h^2}{12}u^{(3)}(x_j+\eta_jh), |\eta_j| \leq \max\{\eta_j^+,\eta_j^-\} < 1,\\ \frac{u(x_j-h)+u(x_j+h)-2u(x_j)}{h^2} &= u_j^{(2)}+\frac{h^2}{24}u^{(4)}(x_j+\theta_jh), |\theta_j| \leq \max\{\theta_j^+,\theta_j^-\} < 1. \end{aligned}$$

Προσοχή: Γνωστό θεώρημα της Ανάλυσης λέει ότι αν μια συνάρτηση f είναι συνεχής σε ένα κλειστό σύνολο τότε θα είναι και φραγμένη εκεί.
Επομένως: Αν υποθέσουμε ότι οι $u^{(3)}, u^{(4)}$ είναι συνεχείς στο διάστημα ορισμού $[a,b]$, τα υπόλοιπα $h^2u^{(3)}(x_j+\eta_jh)$ και $h^2u^{(4)}(x_j+\theta_jh)$ θα τείνουν στο 0 όπως και το h^2.
Επομένως

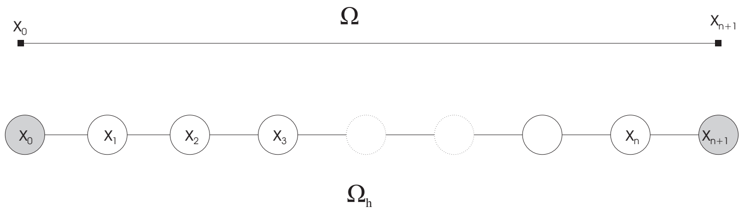
$$\begin{aligned} \frac{u(x_j+h)-u(x_j-h)}{2h} &= u_j^{(1)}+O(h^2)\\ \frac{u(x_j-h)+u(x_j+h)-2u(x_j)}{h^2} &= u_j^{(2)}+O(h^2) \end{aligned}$$

- ✓ το σφάλμα μειώνεται σαν το $O(h^2)$
- ✓ οι τύποι είναι «συμμετρικοί» ως προς τη θέση x

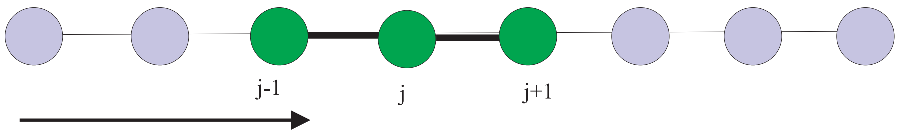
Οι παραπάνω σχέσεις αποτελούν προσεγγίσεις

2ης τάξης της 1ης και της 2ας παραγώγου αντίστοιχα με κεντρισμένες πεπερασμένες διαφορές

Υπολογιστικά διαγράμματα
«Φυσική αρίθμηση» κόμβων από αριστερά προς τα δεξιά.



Ίχνη/Άστρα
Σχηματική παρουσίαση των κόμβων που παρέχουν πληροφορίες για την προσέγγιση των παραγώγων σε ένα σημείο του πλέγματος:



Επίλυση ΣΔΕ με πεπερασμένες διαφορές Έστω ότι η άγνωστη μεταβλητή $u(x)$ έχει πεδίο ορισμού το $\Omega = [X_L, X_U]$ και ότι ικανοποιεί τη ΔΕ

$$-\frac{d^2u}{dx^2}(x) + b(x)\frac{du}{dx}(x) + c(x)u(x) = d(x)$$

Γράφουμε

$$\mathcal{L}(u, b, c, d, x) = 0$$

Το πρόβλημα είναι να υπολογίσουμε το $u(x)$ στο $X_L < x < X_U$.
Οριακές συνθήκες Για μοναδική λύση πρέπει να ορισθεί η συμπεριφορά του u στα «άκρα του χωρίου». Αναφερόμαστε σε «συνοριακές τιμές» ή «συνοριακές συνθήκες».

ΣΔΕ 2ης τάξης $\longrightarrow$ πρόβλημα συνοριακών τιμών

Σχετικά με τις συνοριακές συνθήκες Κάθε συνοριακή συνθήκη,

- μπορεί να αναφέρεται στην τιμή της συνάρτησης (Dirichlet),
- ή στην τιμή παραγώγου της συνάρτησης (Neumann),
- ή μείγμα των παραπάνω (Robin),

- … σε ολόκληρο ή σε τμήμα του συνόρου του χωρίου,
- ή περιοδικές.
- Κωδικοποιούμε ως $\mathcal{B}(u, x) = 0$.
- Προσοχή: Οι συνοριακές συνθήκες επιδρούν στην ύπαρξη και μοναδικότητα της λύσης της ΔΕ.

Παράδειγμα π.χ. για το παραπάνω πρόβλημα θα μπορούσαμε να έχουμε

- $u(X_L) = g_1, u(X_U) = g_2 \rightarrow$
- $\frac{du}{dx}|_{X_L} = g_1, \frac{du}{dx}|_{X_R} = g_1$ ή και
- $\gamma_1 u(x_L) + \gamma_2 \frac{du}{dx}|_{X_L} = g_1, \delta_1 u(x_U) + \delta_2 \frac{du}{dx}|_{X_U} = g_2,$
- $u(X_L) = u(X_U) = g.$

Ένα Θεώρημα Ύπαρξης Γνωρίζουμε ότι:
Θεώρημα 1. *Έστω το παραπάνω γραμμικό πρόβλημα συνοριακών τιμών 2 σημείων*

$$-\frac{d^2u}{dx^2}(x) + b(x)\frac{du}{dx}(x) + c(x)u(x) = d(x).$$

Αν οι συναρτήσεις $b(x), c(x), d(x)$ είναι συνεχείς και $c(x) \geq 0$ στο διάστημα ορισμού $[X_L, X_U]$ τότε το πρόβλημα έχει μία μοναδική λύση. □

Επίσης μπορεί να δειχτεί ότι αν $b(x) = 0$, η $u(x)$ έχει συνεχή 4η παράγωγο στο διάστημα και γίνει η διακριτοποίηση με κεντρισμένες διαφορές, τότε το συνολικό σφάλμα στην επίλυση είναι

$$\max_{0 \leq j \leq n} |u(x_j) - U_j| \leq \frac{1}{96} h^2 (X_U - X_L)^2 M_4$$

όπου M_4 είναι φράγμα για την 4η παράγωγο.
Σχετικά με το σφάλμα

- Προσέξτε ότι το παραπάνω φράγμα αναφέρεται στο σφάλμα προσέγγισης των τιμών της $u(x)$ από το διάνυσμα U στα σημεία του πλέγματος.
- Δηλαδή: Δεν έχει συνυπολογιστεί το σφάλμα στρογγύλευσης κατά τη λύση του συστήματος για το U
- Με α.κ.υ. το σύστημα $AU = F$ επιστρέφει $\hat{U}$ που περιέχει το εμπρός σφάλμα!
- Για μια πλήρη ανάλυση θάπρεπε να εξετάσουμε τις τιμές στα αριστερά της ανισότητας παρακάτω

$$\|u(x_j) - \hat{U}_j| \leq \|u(x_j) - U_j| + \|U_j - \hat{U}_j|$$

- Ο 1ος όρος δεξιά είναι το σφάλμα από την προσέγγιση του u με U,
- ο 2ος όρος είναι το (εμπρός) σφάλμα από την επίλυση του συστήματος
- το 1ο το φράσσουμε από το παραπάνω αποτέλεσμα, το 2ο από τη θεωρία που έχουμε αναπτύξει στα κεφ. 4, 5
- Οι επιλογές μας στην επίλυση (π.χ. πόσο μεγάλο n δηλ. μικρό h) πρέπει να λάβουν υπόψη τη σχέση μεταξύ αυτών των σφαλμάτων (καθώς μικραίνει το πρώτο μπορεί να μεγαλώνει το δεύτερο...)

Βήματα

διακριτοποίηση χωρίου $\Omega \to \Omega_h$

διακριτοποίηση τελεστών/εξίσωσης $\mathcal{L} \to \mathcal{L}_h$

διακριτοποίηση οριακών συνθηκών $\mathcal{B} \to \mathcal{B}_h$

Εύρεση διακριτής λύσης $u \to U$

Διακριτοποίηση χωρίου «Πλέγμα» Ω_h από $n+2$ ισαπέχοντες «κόμβους» στο $[X_L, X_U]$.

$$\Omega_h := \{x_j | x_j = X_L + jh, j = 0, ..., n+1\}$$

όπου $h = \frac{X_U - X_L}{n+1}$. $x_0 = X_L, x_{n+1} = X_U$ καλούνται ακραία ή συνοριακά σημεία.

$$\boxed{\text{Αντιστοιχία } u(x_j) \leftrightarrow U_j}$$

Παρατηρήσεις Η διακριτοποίηση του χωρίου με σύνολο κατάλληλα τοποθετημένων κόμβων είναι ένα σημαντικό και δύσκολο θέμα.

- ώστε να είναι ικανή να εφαρμοσθεί σε πολύπλοκα αντικείμενα.
- ώστε να αναπαραστεί με λεπτομέρεια περιοχές στις οποίες η λύση αναμένεται να μην είναι ομαλή, π.χ. περιοχές με ιδιαίτερα σημεία, γωνίες, κλπ.
- ώστε να γίνεται γρήγορα και φθηνά σε περιβάλλον φιλικής διεπαφής

Στο παρελθόν γινόταν στο χέρι και ήταν πολύ χρονοβόρα. Για παράδειγμα, η κατασκευή πλέγματος για την αναπαράσταση της ροής γύρω από ένα αεροσκάφος απαιτούσε περί τους 6 μήνες, ενώ σήμερα ο χρόνος έχει μειωθεί στις 2-3 εβδομάδες.
Τα πλέγματα που χρησιμοποιούν οι κατασκευαστές αντικειμένων κάθε είδους (από αεροπλάνα μέχρι μικροεργαλεία) θεωρούνται απόρρητα.
Διακριτοποίηση ΔΕ

Θα μπορούσαμε να χρησιμοποιήσουμε τις κεντρισμένες προσεγγίσεις

$$\begin{aligned} \frac{u_{j+1}-u_{j-1}}{2h} &\approx u_j^{(1)} \\ \frac{u_{j-1}+u_{j+1}-2u_j}{h^2} &\approx u_j^{(2)} \end{aligned}$$

Η ποιότητα κάθε προσέγγισης εξαρτάται από

1. Το μέγεθος του h
2. το μέγεθος των παραγώγων $u^{(3)}, u^{(4)}$ γύρω από το x_j.

Αν αγνοήσουμε τους όρους $O(h^2)$ της κάθε εξίσωσης λαμβάνουμε τις n σχέσεις

$$\frac{-U_{j-1}-U_{j+1}+2U_j}{h^2}+b_j\frac{U_{j+1}-U_{j-1}}{2h}+c_jU_j=d_j, j=1,...,n$$

για τους αγνώστους U_j.
Είναι n σχέσεις που συνδέουν τις τιμές της U στους κόμβους του πλέγματος.

- Αν έχουμε 2 συνοριακές συνθήκες Dirichlet, τις αξιοποιούμε στα σημεία x_1, x_n:

$$U_0=u(x_0), \qquad U_{n+1}=u(x_{n+1})$$

οπότε

$$\begin{aligned} \frac{-U_0-U_2+2U_1}{h^2}+b_1\frac{U_2-U_0}{2h}+c_1U_1 &= d_1 \\ \frac{-U_{n-1}-U_{n+1}+2U_n}{h^2}+b_n\frac{U_{n+1}-U_{n-1}}{2h}+c_nU_n &= d_n \end{aligned}$$

Όλες μαζί οι εξισώσεις για συνθήκες Dirichlet:

$$\begin{aligned} \frac{-U_2+2U_1}{h^2}+b_1\frac{U_2}{2h}+c_1U_1 &= d_1+\frac{u_0}{h^2}+b_1\frac{u_0}{2h} \\ \frac{-U_{j-1}-U_{j+1}+2U_j}{h^2}+b_j\frac{U_{j+1}-U_{j-1}}{2h}+c_jU_j &= d_j, j=2,\dots,n-1 \\ \frac{-U_{n-1}+2U_n}{h^2}+b_n\frac{-U_{n-1}}{2h}+c_nU_n &= d_n+\frac{u_{n+1}}{h^2}-b_n\frac{u_{n+1}}{2h} \end{aligned}$$

οπότε

$$\begin{aligned} (c_1+\frac{2}{h^2})U_1+(-\frac{1}{h^2}+\frac{b_1}{2h})U_2 &= d_1+\frac{u_0}{h^2}+b_1\frac{u_0}{2h} \\ -(\frac{1}{h^2}+\frac{b_j}{2h})U_{j-1}+(c_j+\frac{2}{h^2})U_j+(-\frac{1}{h^2}+\frac{b_j}{2h})U_{j+1} &= d_j, j=2,\dots,n-1 \\ (-\frac{1}{h^2}-\frac{b_n}{2h})U_{n-1}+(c_n+\frac{2}{h^2})U_n &= d_n+\frac{u_{n+1}}{h^2}-b_n\frac{u_{n+1}}{2h} \end{aligned}$$

Οι παραπάνω n εξισώσεις αποτελούν σύστημα ως προς $[U_1, \cdots, U_n]$:

$$AU = F$$

Το A είναι το τριδιαγώνιο με στοιχεία $\gamma_j, \alpha_j, \beta_j$ στις θέσεις $j-1, j, j+1$ της j γραμμής, δηλ.

$$A = \mathrm{trid}_n[\gamma_j, \alpha_j, \beta_j]$$

όπου

$$\alpha_j = (\frac{2}{h^2} + c_j), \gamma_j = (-\frac{1}{h^2} - \frac{b_j}{2h}), \beta_j = (-\frac{1}{h^2} + \frac{b_j}{2h})$$

Όπου έχουμε συνθήκες Neumann, π.χ. μπορεί να μην γνωρίζουμε τιμή του $u(x_0)$ αλλά μόνον ότι $u_x(x_0) = g$. Τι κάνουμε τότε;

1) Θεωρούμε ότι επιτρέπεται να γράψουμε μια ακόμα εξίσωση στο x_0:

$$\frac{-U_{-1} - U_1 + 2U_0}{h^2} + b_0 \frac{U_1 - U_{-1}}{2h} + c_0 U_0 = d_0$$

όπου $U_{-1} \approx u(x_{-1}) := u(x_0 - h)$.

2) Διακριτοποιούμε τη συνοριακή συνθήκη Neumann:

$$\begin{aligned} g &\approx \frac{U_1 - U_{-1}}{2h} \\ &\Downarrow \\ U_{-1} &= U_1 - 2hg. \end{aligned}$$

3) Συνδυάζουμε τα παραπάνω για να απαλείψουμε τον όρο U_{-1} από την εξίσωση

$$\frac{-(U_1 - 2hg) - U_1 + 2U_0}{h^2} + b_0 \frac{U_1 - (U_1 - 2hg)}{2h} + c_0 U_0 = d_0$$

Όλες οι εξισώσεις για συνθήκη Neumann στο ένα άκρο (π.χ. στον κόμβο υπ' αριθμ. 0) και Dirichlet στον κόμβο υπ' αριθμ. $n+1$:

$$\frac{-2U_1 + 2U_0}{h^2} + c_0 U_0 = d_0 - b_0 g - \frac{2g}{h}$$

$$\begin{aligned}
\frac{-U_2+2U_1}{h^2}+b_1\frac{U_2}{2h}+c_1U_1 &= d_1+\frac{u_0}{h^2}+b_1\frac{u_0}{2h}\\
\frac{-U_{j-1}-U_{j+1}+2U_j}{h^2}+b_j\frac{U_{j+1}-U_{j-1}}{2h}+c_jU_j &= d_j, j=1,\ldots,n-1\\
\frac{-U_{n-1}+2U_n}{h^2}+b_n\frac{-U_{n-1}}{2h}+c_nU_n &= d_n+\frac{u_{n+1}}{h^2}-b_n\frac{u_{n+1}}{2h}
\end{aligned}$$

οπότε

$$\begin{aligned}
(c_0+\frac{2}{h^2})U_0-\frac{2}{h^2}U_1 &= d_0-b_0g-\frac{2g}{h}\\
-(\frac{1}{h^2}+\frac{b_j}{2h})U_{j-1}+(c_j+\frac{2}{h^2})U_j+(-\frac{1}{h^2}+\frac{b_j}{2h})U_{j+1} &= d_j, j=1,\ldots,n-1\\
(-\frac{1}{h^2}-\frac{b_n}{2h})U_{n-1}+(c_n+\frac{2}{h^2})U_n &= d_n+\frac{u_{n+1}}{h^2}-b_n\frac{u_{n+1}}{2h}
\end{aligned}$$

Οι παραπάνω $n+1$ εξισώσεις αποτελούν σύστημα ως προς $[U_0, U_1, \cdots, U_n]$:

$$A_{ND}U = F$$

Το A_{ND} είναι το τριδιαγώνιο με στοιχεία $\gamma_j, \alpha_j, \beta_j$ στις θέσεις $j-1, j, j+1$ της j γραμμής, δηλ.

$$A = \mathbf{trid}_n[\gamma_j, \alpha_j, \beta_j]$$

όπου

$$\alpha_j = (\frac{2}{h^2}+c_j), \text{για } j=1,...,n+1, \gamma_j = (-\frac{1}{h^2}-\frac{b_j}{2h}) \text{για } j=2,...,n+1,$$

και

$$\beta_1 = -\frac{2}{h^2}, \beta_j = (-\frac{1}{h^2}+\frac{b_j}{2h}), j=2,...,n.$$

<u>ΠΡΟΣΟΧΗ:</u> Για μια συνοριακή συνθήκη με παράγωγο, εισαγάγαμε μια παραπάνω εξίσωση!!! Για Neumann και στα 2 άκρα θα είχαμε $n+2$ εξισώσεις. Για περιοδικές συνθήκες, θα είχαμε $n-1$ εξιώσεις.

Έτσι κατασκευάζουμε το μητρώο των συντελεστών που αντιστοιχεί σε κάθε επιλογή συνοριακών συνθηκών. Αξίζει να σημειωθεί ότι μια «καλή» διακριτοποίηση πρέπει να «κληρονομεί» τις ιδιότητες της διαφορικής εξίσωσης στο διακριτό πρόβλημα. Επίσης, να ελαχιστοποιεί τη δημιουργία και μετάδοση «τεχνητών» προβλημάτων στο διακριτό πρόβλημα γιατί μετά αυτά θα εμφανιστούν και θα εισάγουν θόρυβο στη λύση. Ένα παράδειγμα: Αν το πρόβλημα έχει μοναδική λύση, θέλουμε το μητρώο που κωδικοποιεί τη διακριτοποίηση να είναι αντιστρέψιμο. Αντίστροφα, αν το μητρώο δεν είναι αντιστρέψιμο, αυτό θέλουμε να οφείλεται στο ότι το συνεχές πρόβλημα δεν έχει μοναδική λύση.

Αν, π.χ. διακριτοποιήσετε σύμφωνα με τα παραπάνω το πρόβλημα $-u''(x) = f(x)$, με συνοριακές συνθήκες Neumann και στα δυο άκρα, τότε το μητρώο είναι τριδιαγώνιο, αλλά έχει στοιχεία

$$\frac{1}{h^2}\begin{pmatrix} 2 & -2 & 0 & & \\ -1 & 2 & -1 & & \\ 0 & \ddots & \ddots & \ddots & \\ & \ddots & -1 & 2 & -1 \\ & & & -2 & 2 \end{pmatrix}$$

Μπορείτε να εξακριβώσετε ότι δεν είναι αντιστρέψιμο, όπως και το συνεχές πρόβλημα δεν έχει μοναδική λύση (αν το $\hat{U}$ είναι λύση, για κάθε σταθερή τιμή C το $\hat{U} + C$ θα είναι και αυτό λύση της ΔΕ.

> Προσέξτε: Το μητρώο ανέδειξε την ιδιότητα του συνεχούς προβλήματος! Αυτό είναι κάτι καλό - το διακριτό πρόβλημα να «κληρονομεί» τις ιδιότητες του συνεχούς!

Θέματα

{ Είναι το μητρώο των συντελεστών αντιστρέψιμο;

{ Αν ναι, πώς να λύσουμε το σύστημα;

{ Ποιά είναι η σχέση μεταξύ της λύσης του διακριτού συστήματος $A\hat{U}(h) = f$ και της λύσης της ΔΕ u;
Πρέπει να συγκλίνει η μέθοδος

Αυτό συμβαίνει αν το σφάλμα $\max_j |U_j(h) - u(x_j)| \rightarrow 0$ καθώς $h \rightarrow 0$. Πόσο γρήγορα συγκλίνει;

{ Οι παραπάνω σχέσεις προέκυψαν αφού αγνοήσαμε τους όρους $O(h^2)$, θα είναι προσεγγίσεις των αρχικών εξισώσεων ενώ η λύση U θα είναι μόνο προσέγγιση της λύσης του άρχικού προβλήματος.

{ Οι σχέσεις μετετράπηκαν σε εξισώσεις με πεπερασμένο αριθμό όρων που ελπίζουμε να μπορούμε να λύσουμε με γνωστές ή νέες τεχνικές. Οι εξισώσεις αυτές λέγονται εξισώσεις διαφορών και αποτελούν ένα γραμμικό σύστημα n εξισώσεων για n αγνώστους.

{ Οι άγνωστοι είναι ένα διάνυσμα $U = [U_1, ..., U_n]$ που λαμβάνουμε ως προσεγγίσεις του u στους κόμβους του πλέγματος. Η ακρίβεια της προσέγγισης ελέγχεται π.χ. μέσω κάποιας νόρμας του διανύσματος $\|U - u\|$, όπου τώρα το u συμβολίζει τις τιμές της u στους κόμβους.

- Τα σφάλματα που προκύπτουν από τη διακριτοποίηση και την χρήση των εξισώσεων διαφορών για να προσεγγίσουμε το αρχικό πρόβλημα ονομάζονται *τοπικά σφάλματα διακριτοποίησης* ή *τοπικά σφάλματα αποκοπής* ενώ το $\|U-u\|$ ονομάζεται *ολικό σφάλμα*.

Χαρακτηριστικά

$$A = \mathrm{trid}_n[\gamma_j, \alpha_j, \beta_j]$$

όπου

$$\gamma_j = (-\frac{1}{h^2} - \frac{b_j}{2h}), \alpha_j = (\frac{2}{h^2} + c_j), \beta_j = (-\frac{1}{h^2} + \frac{b_j}{2h})$$

Είναι το μητρώο αντιστρέψιμο;

- Η απόδειξη δεν είναι απλή στη γενική περίπτωση
- Έξετάζουμε μία ειδική περίπτωση: Ειδικά όταν $b_j = 0$, $c_j > 0 \qquad \forall j$:

$$A = \frac{1}{h^2}\mathrm{trid}_n[-1, 2 + c_j h^2, -1]$$

- Η αντιστρεψιμότητα πορεί να αποδειχθεί με διάφορους τρόπους:

Gerschgorin Οι ιδιοτιμές του $[a_{ij}] = A \in \mathbb{C}^{n\times n}$ βρίσκονται στην κλειστή περιοχή του μιγαδικού επιπέδου που ορίζεται από την ένωση των n δίσκων

$$|z - \alpha_{ii}| \leq \sum_{j=1, j\neq i}^{n} |\alpha_{ij}|, \qquad i = 1:n.$$

Αρχή μεγίστου Δεν την εξετάζουμε εδώ

Μέσω θεωρήματος Gerschgorin Θεώρημα G.: *Οι ιδιοτιμές του* $[a_{ij}] = A \in \mathbb{C}^{n\times n}$ *βρίσκονται στην κλειστή περιοχή του μιγαδικού επιπέδου που ορίζεται από την ένωση των* n *δίσκων*

$$|z - \alpha_{ii}| \leq \sum_{j=1, j\neq i}^{n} |\alpha_{ij}|, \qquad i = 1:n.$$

Εφαρμόζοντας το θεώρημα G. στην περίπτωση του τριδιαγώνιου που προέκυψε από τη διακριτοποίηση,

$$\text{εφόσον } A = A^\top \Rightarrow \text{όλες οι ιδιοτιμές } \lambda(A) \in \mathbb{R}$$
$$\Downarrow$$
$$-2 \le \lambda(A) - (2 + c_i h^2) < 2, i = 2, ..., n-1$$
$$\Downarrow$$
$$c_i h^2 \le \lambda(A) < 4 + c_i h^2$$
$$\Downarrow$$
$$\text{αν } c_i > 0 \text{ έπεται ότι } \lambda(A) > 0.$$

Επίσης, στις δυο ακραίες περιπτώσεις $-1 \le \lambda(A) - (2 + c_1 h^2) < 1$ και $-1 \le \lambda(A) - (2 + c_n h^2) < 1$ επομένως

$$1 + c_1 h^2 \le \lambda(A) \le 3 + c_1 h^2, \quad 1 + c_n h^2 \le \lambda(A) \le 3 + c_n h^2,$$

και πάλι έπεται ότι αν $c_i > 0$, τότε $\lambda(A) > 0$.

αν ένα μητρώο έχει όλες του τις ιδιοτιμές μη μηδενικές τότε είναι αντιστρέψιμο
$\rightarrow$ π.χ. LU με μερική οδήγηση

Επίσης

αν ένα μητρώο έχει όλες του τις ιδιοτιμές θετικές πραγματικές τότε είναι ΣΘΟ
$\rightarrow$ π.χ. Cholesky

Πώς λύνεται το σύστημα;

$-\frac{d^2u}{dx^2}(x) + b(x)\frac{du}{dx}(x) + c(x)u(x) = d(x),\ u(0) = g_1, u(1) = g_2$

$$\boxed{A = \text{trid}_n[-\frac{1}{h^2} - \frac{b_j}{2h}, \frac{2}{h^2} + c_j, -\frac{1}{h^2} + \frac{b_j}{2h}]}$$

Πρέπει να εκμεταλλευτούμε τη μορφή του μητρώου:

{ Κεντρισμένες πεπερασμένες διαφορές 2ης τάξης

$\Rightarrow$ τριδιαγώνιο

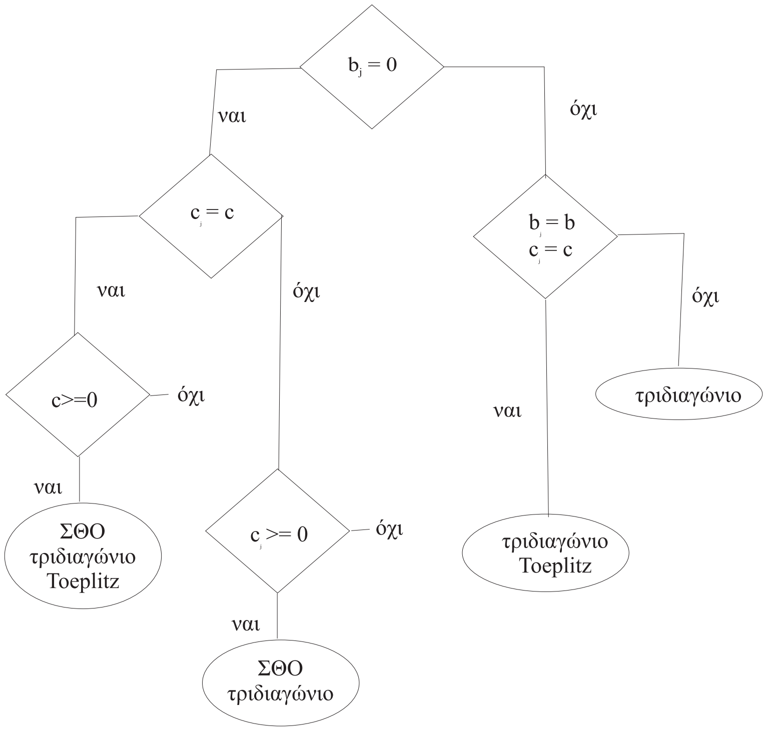
{ Συντελεστές και συνοριακές συνθήκες

$\rightarrow$ συμμετρικό,

$\rightarrow$ θετικά ορισμένο

$\rightarrow$ Toeplitz

δένδρο απόφασης $\rightarrow$ έμπειρο/έξυπνο σύστημα, recommender



Παράδειγμα

$\mathcal{L}$: $-\frac{d^2u}{dx^2}(x) + xu(x) = (9+x)\sin 3x$

$\mathcal{B}$: $u(0) = 0$, $u(\pi/6) = 1$

Ω: $[0, \pi/6]$

{ $h = \frac{\pi}{6(n+1)}$, $\Omega_h = \{x_j = jh | j = 0, ..., n+1\}$

{

$$\begin{aligned}(x_1 + \frac{2}{h^2})U_1 - \frac{1}{h^2}U_2 &= (9+x_1)\sin 3x_1 \\ -\frac{1}{h^2}U_{j-1} + (x_j + \frac{2}{h^2})U_j - \frac{1}{h^2}U_{j+1} &= (9+x_j)\sin 3x_j, j = 2, \dots, n-1 \\ -\frac{1}{h^2}U_{n-1} + (x_n + \frac{2}{h^2})U_n &= (9+x_n)\sin 3x_n + \frac{1}{h^2}\end{aligned}$$

$\hookrightarrow AU = F$ όπου

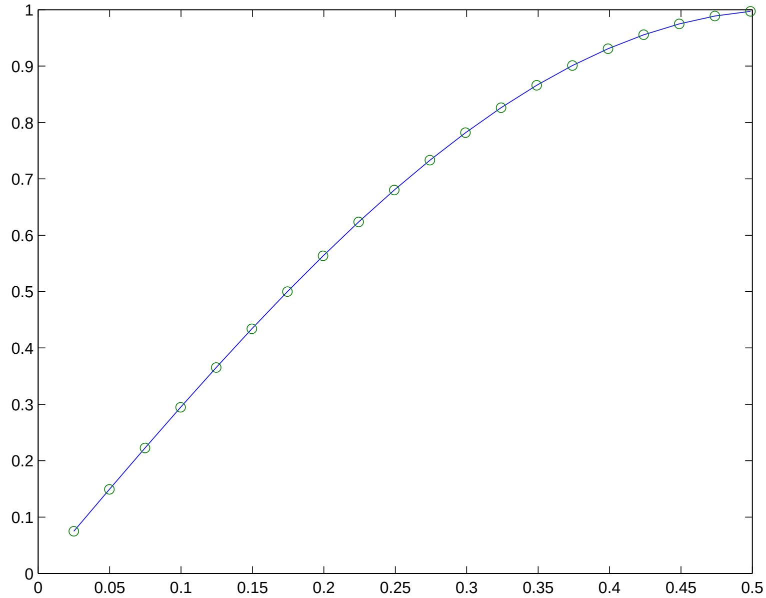
$$A = \textbf{trid}_n[\gamma_j, \alpha_j, \beta_j]$$

$$\alpha_j = (\frac{2}{h^2} + jh), \gamma_j = -\frac{1}{h^2}, \beta_j = -\frac{1}{h^2}$$

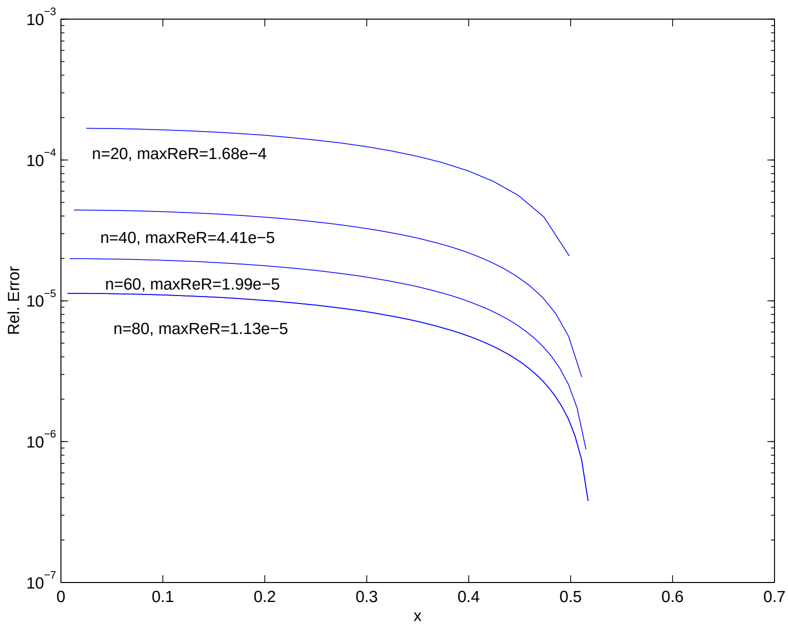
ενώ

$$\begin{aligned} F_j &= (9 + jh) \sin 3x_j, \qquad j = 1, ..., n-1 \\ F_n &= (9 + nh) \sin 3x_n + \frac{1}{h^2} \end{aligned}$$

Λύνουμε $AU = F$ χρησιμοποιώντας Cholesky για τριδιαγώνια συστήματα (κόστος $O(n)$) και συγκρίνουμε τις υπολογισμένες τιμές του U με τις τιμές της ακριβούς λύσης u.



Συμπεριφορά της υπολογισμένης λύσης του διακριτοποιημένου συστήματος του παραδείγματος σε σχέση με την πραγματική λύση $u(x) = \sin(3x)$ για $n = 20$



Το σχήμα δείχνει τις υπολογισμένες τιμές του U όταν διακριτοποιήσουμε με πλέγμα $n = 20$ εσωτερικών κόμβων ενώ το δεξιό σχήμα δείχνει τις καμπύλες του σχετικού σφάλματος $|\frac{u_j - U_j}{u_j}|$ καθώς και το μέγιστο σχετικό σφάλμα για τις περιπτώσεις που $n = 20, 40, 60, 80$. Αν τώρα εξετάσετε τις τιμές του μέγιστου σφάλματος για $n = 20, 40, 80$ αυτές είναι $[0.1680, 0.0441, 0.0113] \times 10^{-3}$ και βλέπετε ότι όταν διπλασιάζουμε τους κόμβους του πλέγματος, το σφάλμα υποτετραπλασιάζεται. Οι γραφικές παραστάσεις υποδεικνύουν ότι το μέγιστο σχετικό σφάλμα συμπεριφέρεται ως $O(h^2)$. Η πρόκληση βέβαια είναι να αποδειχθεί αυτή η συμπεριφορά και αναλυτικά - κάτι που αναβάλλουμε για πιο προχωρημένα μαθήματα.

Μία κοινή περίπτωση

$$-u_{xx} = d(x) \Rightarrow$$

$$\frac{-U_{j-1} - U_{j+1} + (2)U_j}{h^2} = d_j, j = 2, \ldots, n-1$$

Σύστημα $AU = F$ όπου

$$\frac{1}{h^2}\begin{pmatrix} 2 & -1 & \ddots & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & & \\ & & & -1 & 2 \end{pmatrix}\begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ \\ U_n \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ \\ F_n \end{pmatrix}$$

Απρόσμενη ιδιότητα (ΠΡΟΑΙΡΕΤΙΚΟ)

- Οι ιδιοτιμές και ιδιάζουσες τιμές του A είναι διαθέσιμες αναλυτικά!!!

$$\lambda_k(A) = 2 - 2\cos(\frac{\pi k}{n+1}), k = 1, ..., n$$

Από τις αναλυτικές τιμές ακολουθεί αμέσως ότι

$$\min_k \lambda_k(A) = \lambda_1(A) = 2 - 2\cos(\frac{\pi}{n+1}) > 0$$

άρα όλες οι ιδιοτιμές είναι θετικές.

Ένα συμμετρικό μητρώο με όλες τις ιδιοτιμές θετικές είναι ΣΘΟ

- Οι ιδιοτιμές είναι διαθέσιμες αναλυτικά για όλα τα τριδιαγώνια μητρώα Toeplitz$A =$ trid$[\gamma, \underline{\alpha}, \beta]$

Επιπλέον ... από τις ιδιοτιμές συμμετρικού A υπολογίζεται και ο δείκτης κατάστασης

$$\kappa_2(A) = \frac{|\lambda_{\max}(A)|}{|\lambda_{\min}(A)|}$$

αλλά αφού θετικές

$$\begin{aligned}
\lambda_{\min}(A) &= \min_k \lambda_k(A) = \lambda_1(A) = 2 - 2\cos(\frac{\pi}{n+1}) \\
&= 2 - 2(1 - 2\sin^2\frac{\pi}{2(n+1)}) = 4\sin^2\frac{\pi}{2(n+1)} \\
&\approx 4(\frac{\pi}{2(n+1)})^2 = \frac{\pi^2}{(n+1)^2}
\end{aligned}$$

$$\lambda_{\max}(A) = \max_k(2 - 2\cos(\frac{\pi k}{n+1})) < 4$$

$$\Rightarrow \boxed{\kappa_2(A) \simeq \frac{4(n+1)^2}{\pi^2}}$$

Παράδειγμα Εξετάζουμε την ίδια ΔΕ με διαφορετικές συνθήκες:

$$-u'' - \pi^2 u = 1, \quad u(0) = u(1) = 0 \quad \text{ΔΕΝ ΕΧΕΙ ΚΑΜΙΑ ΛΥΣΗ}$$

$$-u'' - \pi^2 u = 1, \quad u(0) = 0, u'(1) = 1 \text{ ΕΧΕΙ ΜΙΑ ΜΟΝΑΔΙΚΗ ΛΥΣΗ}$$

$$-u'' - \pi^2 u = 1, \quad u(0) = 0, u(1) = \frac{2}{\pi^2} \text{ ΕΧΕΙ ΑΠΕΙΡΕΣ ΛΥΣΕΙΣ}$$

Η αναλυτική λύση:

$$u(x) = A\sin(\pi x) + B\cos(\pi x) + \frac{1}{\pi^2}$$

οπότε εύκολα φαίνεται τί ισχύει για τις λύσεις.

Παρατηρήσεις Στη 2η περίπτωση (μια μοναδική λύση) αντιστοιχεί αντιστρέψιμο μητρώο και το σφάλμα διακριτοποίησης υποτετραπλασιάζεται καθώς υποδιπλασιάζουμε το βήμα

Η 1η περίπτωση όμως πώς εμφανίζεται στο διακριτό πρόβλημα; Μπορείτε να δείξετε ότι καθώς μικραίνει το h, ο δείκτης κατάστασης του A μεγαλώνει πάρα πολύ, δείχνοντας ότι οριακά το μητρώο δεν είναι αντιστρέψιμο και επομένως δεν υπάρχει λύση.