



Norges teknisk–naturvitenskapelige  
universitet  
Institutt for matematiske fag

TMA4210 Numerisk  
løsning av part.  
diff.lign. med  
differansemetoder  
Vår 2005

Løsningsforslag — Øving 5

- 1 a) Vi skal undersøke stabilitet ved Fourier-metoden. Metodens karakteristiske polynom er gitt som

$$\tilde{a}(z, r) = 1 + \varepsilon r(z^{-1} - 2 + z) - \frac{\lambda h r}{2}(z - z^{-1}).$$

Fourier-metoden sier at metoden er stabil dersom kravet

$$|\tilde{a}(e^{i\theta}, r)| \leq 1 \quad \theta \in [0, 2\pi],$$

er oppfylt. Vi setter inn  $e^{i\theta}$  i  $\tilde{a}$  og får

$$\begin{aligned}\tilde{a}(e^{i\theta}, r) &= 1 - 2\varepsilon r(1 - \frac{1}{2}(e^{-i\theta} + e^{i\theta})) - \frac{\lambda h r}{2}(e^{i\theta} - e^{-i\theta}) \\ &= 1 - 4\varepsilon r \sin^2 \frac{\theta}{2} - 2i\alpha r \sin \theta,\end{aligned}$$

hvor  $\alpha = \frac{\lambda h}{2}$ . Siden  $|a(e^{i\theta}, r)| \leq 1$  er ekvivalent med  $|a(e^{i\theta}, r)|^2 \leq 1$ , får vi ved kvadrering

$$|a(e^{i\theta}, r)|^2 = 1 - 8\varepsilon r \sin^2 \frac{\theta}{2} + 16\varepsilon^2 r^2 \sin^4 \frac{\theta}{2} + 4\alpha^2 r^2 \sin^2 \theta.$$

Vi bruker  $\sin(2\frac{\theta}{2}) = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}$  og deler så på  $8 \sin^2 \frac{\theta}{2}$ . Stabilitetskravet blir dermed at

$$-\varepsilon r + 2\varepsilon^2 r^2 \sin^2 \frac{\theta}{2} + 2\alpha^2 r^2 \cos^2 \frac{\theta}{2} \leq 0$$

eller ekvivalent at

$$2\alpha^2 r - \varepsilon + 2r(\varepsilon^2 - \alpha^2) \sin^2 \frac{\theta}{2} \leq 0.$$

Dette skal holde for  $\theta \in [0, 2\pi]$ . Vi må skille mellom to tilfeller:

1. Anta at  $\alpha < \varepsilon$ . Da har vi stabilitet dersom

$$2\alpha^2 r - \varepsilon + 2r(\varepsilon^2 - \alpha^2) \leq 0,$$

dvs. dersom

$$r \leq \frac{1}{2\varepsilon}.$$

(sett inn  $\varepsilon = \alpha$ )



godt kan utvikle determinanten om en av søylene til matrisen. Vi velger derfor å utvikle problemmatrisen over ved siste søyle. Vi får dermed at

$$\begin{aligned}\det(Q - \lambda I) &= (2\varepsilon - \lambda)T_{N-1}(\lambda) + (0 + \dots + 0 - 2\varepsilon(\varepsilon - \alpha))T_{N-2}(\lambda) \\ &= (2\varepsilon - \lambda)T_{N-1}(\lambda) - 2\varepsilon(\varepsilon - \alpha)T_{N-2}(\lambda),\end{aligned}$$

som skulle vises. Videre ser vi at matrisene  $T_M(\lambda)$  er tridiagonale, så vi kjenner egenverdiene til disse fra notatet om tridiagonale matriser. Rekurrensformelen fås dog på akkurat samme måte som over, og startbetingelsene er trivielle.

- d) Matrise-metoden sier at metoden er stabil når  $\rho(A) \leq 1$ . Vi må finne  $\sigma(Q)$ , siden  $\sigma(A) = 1 - r\sigma(Q)$ . Dette gjøres ved å finne nullpunktene til  $G_2(\lambda)$ . Vi har at

$$G_2(\lambda) = (2\varepsilon - \lambda)T_1(\lambda) - 2\varepsilon(\varepsilon - \alpha)T_0(\lambda) = (2\varepsilon - \lambda)^2 - 2\varepsilon(\varepsilon - \alpha).$$

Vi har igjen to muligheter:

1. Anta at  $\alpha < \varepsilon$ . Da er  $G_2(\lambda) = 0$  hvis

$$\lambda = 2\varepsilon \pm \sqrt{2\varepsilon(\varepsilon - \alpha)}.$$

Stabilitetskravet er

$$r\lambda \geq 0 \quad \text{og} \quad r\lambda \leq 2.$$

Den første ulikheten er trivielt oppfylt, og den andre gir

$$r \leq \frac{2}{2\varepsilon + \sqrt{2\varepsilon(\varepsilon - \alpha)}}$$

2. Anta at  $\alpha \geq \varepsilon$ . Da er røttene gitt ved

$$\lambda = 2\varepsilon \pm i\sqrt{2\varepsilon(\alpha - \varepsilon)},$$

som gir

$$\sigma(A) = 1 - 2r\varepsilon \pm ir\sqrt{2\varepsilon(\alpha - \varepsilon)}.$$

For å få stabilitet krever vi at

$$|\sigma(A)|^2 = 1 - 4r\varepsilon + 4r^2\varepsilon^2 + 2\varepsilon\alpha r^2 - 2\varepsilon^2 r^2 \leq 1,$$

som er oppfylt når

$$r \leq \frac{2}{\varepsilon + \alpha}.$$

- e) Vi setter  $R = \sqrt{\varepsilon^2 - \alpha^2}$ , slik at  $T_M(x) = R^M U_M(x)$ . Rekurrensformelen for  $T^M$  gir da

$$R^M U_M(x) = 2xR R^{M-1} U_{M-1}(x) - R^2 R^{M-2} U_{M-2}(x),$$

som er ekvivalent med

$$U_M(x) = 2xU_{M-1}(x) - U_{M-2}(x).$$

Startsbetingelsene gir at

$$RU_1(x) = 2xR \Rightarrow U_1(x) = 2x$$

og

$$R^2 U_2(x) = 4x^2 R^2 - R^2 \Rightarrow U_2(x) = 4x^2 - 1.$$

Dermed kjenner vi igjen denne rekurrens-relasjonen som Chebyshev-polynomene av andre sort, som er nettopp  $U_M(x)$ .

- f) Vi har at  $\lambda = 2\varepsilon - 2xR$ , så rekurrensrelasjonen for  $G_N(\lambda)$  gir, når vi substituerer inn uttrykket for  $T_M$ , følgende uttrykk

$$\begin{aligned} H_N(x) &= G_N(2\varepsilon - 2xR) \\ &= 2xRR^{N-1}U_{N-1}(x) - 2\varepsilon(\varepsilon - \alpha)R^{N-2}U_{N-2}(x) \\ &= 2R^{N-2}[x(\varepsilon^2 - \alpha^2)U_{N-1}(x) - \varepsilon(\varepsilon - \alpha)U_{N-2}(x)] \\ &= 2(\varepsilon - \alpha)R^{N-2}[x(\varepsilon + \alpha)U_{N-1}(x) - \varepsilon U_{N-2}(x)]. \end{aligned}$$

Vi kan dermed konkludere med at røttene til  $H_N(x)$  ligger i intervallet  $(-1, 1)$ . La de korresponderende røttene til  $G_N(\lambda)$  være gitt ved  $\lambda_j$ . Vi har da at

$$\begin{aligned} \sigma_j(A) &= 1 - r\lambda_j \\ &= 1 - 2r(\varepsilon - x_jR), \end{aligned}$$

så kravet  $-1 \leq \sigma_j(A) \leq 1$  gir den trivielt oppfylte ulikheten

$$r(\varepsilon - x_jR) \geq 0$$

og ulikheten

$$r(\varepsilon - x_jR) \leq 1,$$

som er oppfylt når

$$r \leq \frac{1}{\varepsilon + \sqrt{\varepsilon^2 - \alpha^2}}$$

- g) Sammenligning av Fouriermetoden og matrisemetoden (“egenverditeknikker”) gir at vi ikke får de samme betingelsene:

	Fouriermetoden	Matrisemetoden
$\alpha \geq \varepsilon$	$r \leq \frac{\varepsilon}{2\alpha^2}$ (strengest)	$r \leq \frac{\varepsilon}{2\alpha^2}$
$\alpha < \varepsilon$	$r \leq \frac{1}{2\varepsilon}$ (strengest)	$r \leq \frac{1}{\varepsilon + \sqrt{\varepsilon^2 - \alpha^2}}$

Fouriermetoden støtter ikke annet enn periodiske randkrav, mens systemet vårt har Dirichlet i venste endepunkt og Neumann i høyre endepunkt.

På den andre side så krever matrisemetoden at iterasjonsmatrisa  $A$  er normal,  $AA^T = A^T A$  (symmetriske og skjev-symmetriske matriser har denne egenskapen).  $A$  er nesten symmetrisk så for en passende definisjon av nesten, så er den “nesten” normal.

Begge teknikker brukes likevel for å indikere stabilitetskrav for metoder med litt ustandard randbetingelser fordi det er bedre med en viss indikasjon på hva som kreves for stabilitet enn ingen.

2] Matrisene  $B$  og  $C$  er gitt ved

$$B = \begin{bmatrix} D_M & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & D_M \end{bmatrix}, \quad C = \begin{bmatrix} -2I_M & I_M & & & \\ I_M & -2I_M & I_M & & \\ & & \ddots & \ddots & \ddots \\ & & & & I_M & -2I_M \end{bmatrix},$$

hvor matrisen  $D_M$  er gitt ved

$$D_M = \begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & -2 \end{bmatrix}.$$

Vi ønsker nå å løse ligningen hvor  $B$  og  $C$  inngår. Kall høyresiden  $f$ , og sett  $F = (I - rB)$  og  $H = (I - rC)$ . Vi omskriver ligningen til  $FHU^{n+1} = f$ . Først løser vi  $Fy = f$ , og deretter løser vi  $HU^{n+1} = y$ . Begge ligningene er tridiagonale og kan derfor løses med  $\mathcal{O}(M^2)$  operasjoner. NB: Husk at  $\mathcal{O}(M^2)$  operasjoner i denne sammenheng betyr " $\mathcal{O}(n)$  operasjoner", siden matrisen har  $M^2$  rader og dermed  $M^4$  elementer.

I Matlab kan disse matrisene enkelt konstrueres med et Kronecker tensorprodukt. Definer matrisa  $D$  i Matlab som  $D_M$  over. Vi kan da konstruere

```
>> M = 3
>> D = diag(ones(1,M-1),-1) + diag(-2*ones(1,M),0) + diag(ones(1,M-1),1)
D =
    -2     1     0
     1    -2     1
     0     1    -2
>> B = kron(eye(M), D)
>> C = kron(D, eye(M))
>> help kron
```

$BC$  er da det samme som  $\text{kron}(D, D)$ .

### 3 Oppgave 13.11 i læreboka.

We will use Fourier method to analyze the stability of the scheme

$$u_l^{n+1} = \frac{1}{2}(2 - 5\mu + 6\mu^2)u_l^n + \frac{2}{3}\mu(2 - 3\mu)(u_{l-1}^n + u_{l+1}^n) - \frac{1}{12}\mu(1 - 6\mu)(u_{l-2}^n + u_{l+2}^n).$$

The scheme has the following stability function  $\tilde{a}(z, \mu)$

$$\tilde{a}(z, \mu) = \frac{1}{2}(2 - 5\mu + 6\mu^2) + \frac{2}{3}\mu(2 - 3\mu)(z^{-1} + z) - \frac{1}{12}\mu(1 - 6\mu)(z^{-2} + z^2).$$

Substituting  $z = e^{i\theta}$ , we obtain

$$\tilde{a}(e^{i\theta}, \mu) = \frac{1}{2}(2 - 5\mu + 6\mu^2) + \frac{4}{3}\mu(2 - 3\mu)\cos\theta - \frac{1}{6}\mu(1 - 6\mu)\cos 2\theta.$$

Let us use  $\cos 2\theta = 2\cos^2\theta - 1$  to express  $\tilde{a}(e^{i\theta}, \mu)$  in terms of  $x = \cos\theta$ :

$$\tilde{a}(x, \mu) = \frac{1}{3}\mu(6\mu - 1)x^2 + \frac{4}{3}\mu(2 - 3\mu)x + \left(1 - \frac{7}{3}\mu + 2\mu^2\right). \quad (1)$$

We must find all values  $\mu$  such that

$$|\tilde{a}(e^{i\theta}, \mu)| \leq 1, \quad \theta \in [0, 2\pi],$$

or, equivalently,

$$|\tilde{a}(x, \mu)| \leq 1, \quad x \in [-1, 1],$$

Let us notice that according to the definition  $\mu \geq 0$ . We limit ourselves to this interval in the analysis. There are two cases to consider:

1. If  $\mu(6\mu - 1) = 0$ , i.e.  $\mu = 0$  or  $\mu = \frac{1}{6}$  the plot of the function  $\tilde{a}(x, \mu)$  is a line. The maximum value of  $|\tilde{a}(x, \mu)|$  for  $x \in [-1, 1]$  is either  $|\tilde{a}(-1, \mu)|$  or  $|\tilde{a}(1, \mu)|$ . We get

$$|\tilde{a}(-1, \mu)| = 1 - \frac{16}{3}\mu + 8\mu^2, \quad |\tilde{a}(1, \mu)| \equiv 1.$$

It is easy to see that  $|\tilde{a}(-1, \mu)| \leq 1$  for  $\mu = 0$  and  $\mu = \frac{1}{6}$ . Thus, the scheme is stable.

2. If  $\mu(6\mu - 1) \neq 0$ , the plot of the function  $\tilde{a}(x, \mu)$  is a parabola so that there is one extremum point

$$x_0 = -2 \frac{(2 - 3\mu)}{(1 - 6\mu)}.$$

The maximum value of  $|\tilde{a}(x, \mu)|$  for  $x \in [-1, 1]$  can be in  $x = -1$ ,  $x = 1$  or  $x = x_0$ . It is important to notice that we must consider the point  $x = x_0$  only if  $x_0 \in [-1, 1]$ . For the stability of the scheme we require

$$|\tilde{a}(-1, \mu)| \leq 1, \quad |\tilde{a}(1, \mu)| \leq 1$$

and

$$|\tilde{a}(x_0, \mu)| \leq 1, \quad \text{if } x_0 \in [-1, 1].$$

We obtain  $|\tilde{a}(1, \mu)| \equiv 1$ ,  $|\tilde{a}(-1, \mu)| \leq 1$  if  $\mu \leq \frac{2}{3}$ . The point  $x_0$  is located in the interval  $[-1, 1]$  if  $x \geq \frac{5}{12}$ . Because for  $\mu \in [\frac{5}{12}, \frac{2}{3}]$  parabola (1) is has minimum in  $x_0$  (why?) and it is entirely located in the upper half plane the value  $|\tilde{a}(x_0, \mu)|$  can not have maximum at  $x_0$ .

Finally, the scheme is stable if  $\mu \in [0, \frac{2}{3}]$ .

#### 4 Oppgave 13.12 i læreboka.

Vi skal drøfte stabiliteten til FM-skjemaet

$$u_l^{n+1} - \frac{1}{2}(\mu - \zeta)(u_{l-1}^{n+1} - 2u_l^{n+1} + u_{l+1}^{n+1}) = u_l^n + \frac{1}{2}(\mu + \zeta)(u_{l-1}^n - 2u_l^n + u_{l+1}^n)$$

for forskjellige valg av  $\zeta$ . Først benytter vi matrisemetoden. La  $\Delta x = \frac{1}{d+1}$ . Vi setter som alltid

$$u^n = [u_1^n, \dots, u_d^n]^T.$$

Vi må først bestemme systemmatrisen  $A_{\Delta x}$ . Vi har at skjemaet kan skrives som

$$A_{\Delta x}^+ u^{n+1} = A_{\Delta x}^- u^n,$$

hvor

$$A_{\Delta x}^+ = \begin{bmatrix} 1 + (\mu - \zeta) & -\frac{1}{2}(\mu - \zeta) & & & \\ -\frac{1}{2}(\mu - \zeta) & 1 + (\mu - \zeta) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & -\frac{1}{2}(\mu - \zeta) & 1 + (\mu - \zeta) \end{bmatrix}$$

og

$$A_{\Delta x}^- = \begin{bmatrix} 1 - (\mu + \zeta) & \frac{1}{2}(\mu + \zeta) & & & \\ \frac{1}{2}(\mu + \zeta) & 1 - (\mu - \zeta) & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & \frac{1}{2}(\mu + \zeta) & 1 - (\mu + \zeta) \end{bmatrix}.$$

Begge matrisene er TST (tridiagonal, symmetrisk og Toeplitz) og egenverdiene er gitt ved

$$\begin{aligned} \lambda_j^+ &= 1 + (\mu - \zeta) - 2\frac{1}{2}(\mu - \zeta) \cos\left(\frac{\pi j}{d+1}\right) = 1 + 2(\mu - \zeta) \sin^2\left(\frac{\pi j \Delta x}{2}\right) & j = 1, \dots, d \\ \lambda_j^- &= 1 - (\mu + \zeta) + 2\frac{1}{2}(\mu + \zeta) \cos\left(\frac{\pi j}{d+1}\right) = 1 - 2(\mu + \zeta) \sin^2\left(\frac{\pi j \Delta x}{2}\right) & j = 1, \dots, d. \end{aligned}$$

Selve systemmatrisen er gitt ved  $A_{\Delta x} = (A_{\Delta x}^+)^{-1} A_{\Delta x}^-$ . Siden komponentmatrisene er tridiagonale og derfor normale, er systemmatrisen selv en normal matrise. Fordi matriser  $A_{\Delta x}^+$  og  $A_{\Delta x}^-$  er TST kan vi skrive

$$A_{\Delta x}^+ = QD^+Q^T \quad \text{og} \quad A_{\Delta x}^- = QD^-Q^T,$$

der  $Q$  er den ortonormale matrisen med egenvektorene (egenvektorene er like for  $A_{\Delta x}^+$  og  $A_{\Delta x}^-$ ) og  $D^+$  og  $D^-$  diagonale matriser med de respektive egenverdiene (Lemma 10.5 i boka). Dermed følger det umiddelbart at

$$\begin{aligned} A_{\Delta x} &= (QD^+Q^T)^{-1}QD^-Q^T \\ &= Q(D^+)^{-1}Q^TQD^-Q^T \\ &= Q(D^+)^{-1}D^-Q^T. \end{aligned}$$

Dette viser at egenverdiene til  $A_{\Delta x}$  er gitt som  $(\lambda_j^+)^{-1}\lambda_j^-$ , dvs egenverdiene er gitt som

$$\lambda_j = \frac{1 - 2(\mu + \zeta) \sin^2\left(\frac{\pi j \Delta x}{2}\right)}{1 + 2(\mu - \zeta) \sin^2\left(\frac{\pi j \Delta x}{2}\right)} \quad j = 1, \dots, d.$$

Vi vet at  $0 \leq \sin^2\left(\frac{\pi j \Delta x}{2}\right) \leq 1$ . Anta at verdien er  $s$  for å slippe å skrive så mye.  $|\lambda_j| \leq 1$  betyr da for våre reelle tall:

$$-1 \leq \frac{1 - 2(\mu + \zeta)s}{1 + 2(\mu - \zeta)s} \leq 1$$

og vi skiller i to tilfeller:

1. Nevneren er positiv,  $1 + 2s\mu - 2s\zeta > 0 \iff \zeta < \mu + 1/(2s)$ . Da ganger vi opp nevneren og venstre ulikhet gir

$$-1 - 2s\mu + 2s\zeta \leq 1 - 2s\mu - 2s\zeta$$

som betyr at vi må kreve

$$\zeta \leq \frac{1}{2s}$$

for stabilitet. Høyre ulikhet blir

$$1 - 2s\mu - 2s\zeta \leq 1 + 2s\mu - 2s\zeta$$

som kun krever  $\mu > 0$  for stabilitet, og dette er alltid tilfredstilt i denne oppgaven (noe annet ville innebære negativ tid).

2. Nevneren er negativ,  $\zeta > \mu + 1/(2s)$ . Når nevneren er negativ må vi endre på fortegnet til ulikheten når vi ganger opp. Venstre ulikhet gir

$$-1 - 2s\mu + 2s\zeta \geq 1 - 2s\mu - 2s\zeta$$

som betyr at  $\zeta \geq 1/(2s)$ . Men høyre ulikhet gir nå  $\mu \leq 0$  som ikke er sant i denne oppgaven, dermed er aldri skjemaet stabilt når nevneren er negativ.

Siden  $0 \leq s \leq 1$  har vi at strengeste stabilitetsgrense kommer når  $s = 1$ , dermed har vi at skjemaet kun er stabilt når

$$\zeta \leq \frac{1}{2}.$$

Vi skal så benytte Fouriermetoden på samme skjema. Stabilitetsfunksjonen er gitt ved

$$\tilde{a}(z, \mu) = \frac{1 + \frac{1}{2}(\mu + \zeta)(z^{-1} - 2 + z)}{1 - \frac{1}{2}(\mu - \zeta)(z^{-1} - 2 + z)}.$$

Vi setter  $z = e^{i\theta}$  og får

$$\begin{aligned} \tilde{a}(e^{i\theta}, \mu) &= \frac{1 - (\mu + \zeta) + \frac{1}{2}(\mu + \zeta)(e^{i\theta} + e^{-i\theta})}{1 + (\mu - \zeta) - \frac{1}{2}(\mu - \zeta)(e^{i\theta} + e^{-i\theta})} \\ &= \frac{1 - 2(\mu + \zeta) \sin^2(\frac{\theta}{2})}{1 + 2(\mu - \zeta) \sin^2(\frac{\theta}{2})} \quad \theta \in [0, 2\pi]. \end{aligned}$$

Vi ser at stabilitetsfunksjonen sammensvarer med uttrykket for egenverdiene fra matrisemetoden, så vi kan konkludere med at Fouriermetoden gir samme stabilitetskrav. Legg merke til at hvis vi setter  $\zeta = 0$ , så får vi Crank–Nicolsons metode, så vi kan forvente at  $\zeta$  er en slags relaksasjonsparameter, dvs at den skal øke konvergenstigheten.