

Katsaus konepellin alle - iteratiivisia menetelmiä ominaisarvotehtävälle

Saku Suuriniemi
`saku.suuriniemi@tut.fi`

TTY / Sähkötekniikka

Fysiikan seminaari 13.2.2014

Outline

Ominaisarvotehtävä

Matriisien spektrilause

Potenssiinkorotusmenetelmä

Rayleigh'n osamäärä

Inverssi-iteraatio

Ominaisarvojen siirto

Ominaisarvojen siirto ja kääntö

Useampi ominaisvektori kerralla

Deflaatio

Yhteenveto

Polkuja eteenpäin

Outline

Ominaisarvotehtävä

Matriisien spektrilause

Potenssiinkorotusmenetelmä

Rayleigh'n osamäärä

Inverssi-iteraatio

Ominaisarvojen siirto

Ominaisarvojen siirto ja kääntö

Useampi ominaisvektori kerralla

Deflaatio

Yhteenveto

Polkuja eteenpäin

Outline

Ominaisarvotehtävä

Matriisien spektrilause

Potenssiinkorotusmenetelmä

Rayleigh'n osamäärä

Inverssi-iteraatio

Ominaisarvojen siirto

Ominaisarvojen siirto ja kääntö

Useampi ominaisvektori kerralla

Deflaatio

Yhteenveto

Polkuja eteenpäin

Outline

- Ominaisarvotehtävä

 - Matriisien spektrilause

- Potenssiinkorotusmenetelmä

 - Rayleigh'n osamäärä

- Inverssi-iteraatio

 - Ominaisarvojen siirto

 - Ominaisarvojen siirto ja kääntö

- Useampi ominaisvektori kerralla

 - Deflaatio

- Yhteenveto

- Polkuja eteenpäin

Outline

Ominaisarvotehtävä

Matriisien spektrilause

Potenssiinkorotusmenetelmä

Rayleigh'n osamäärä

Inverssi-iteraatio

Ominaisarvojen siirto

Ominaisarvojen siirto ja kääntö

Useampi ominaisvektori kerralla

Deflaatio

Yhteenveto

Polkuja eteenpäin

Outline

- Ominaisarvotehtävä

 - Matriisien spektrilause

- Potenssiinkorotusmenetelmä

 - Rayleigh'n osamäärä

- Inverssi-iteraatio

 - Ominaisarvojen siirto

 - Ominaisarvojen siirto ja kääntö

- Useampi ominaisvektori kerralla

 - Deflaatio

- Yhteenveto

- Polkuja eteenpäin

Ominaisarvotehtävä

- ▶ Ominaisarvotehtävä määritellään lineaarioperaattorille

$$L : V \rightarrow V.$$

(Kuvaa vektoriavaruuden V alkiot takaisin avaruuteen V)

“Etsi $v \in V$, $v \neq 0$ ja luku $\lambda \in \mathbb{C}$, että yhtälö

$$Lv = \lambda v$$

toteutuu.”

- ▶ Kuva kollineaarinen alkukuvan kanssa (tai nolla).
- ▶ Ominaispareja (v, λ) usein paljon. Jos V :n dimensio ei ole äärellinen, ominaispareja yleensä ääretön määrä.

Ominaisarvotehtävä

Miksi sen pitäisi kiinnostaa?

Ominaisarvotehtävä

- ▶ Ajasta riippumattoman Schrödingerin yhtälön ratkaiseminen on ominaisarvotehtävä:

$$H\Psi = E\Psi.$$

(H Hamiltonin operaattori, kuvaa aaltofunktioiden avaruuden itselleen. Energialle E pätee $E \in \mathbb{R}$.)

- ▶ Ratkaisusoftat etsivät ratkaisun likiarvoja, koska aaltofunktioiden avaruuden dimensio ei ole äärellinen.
- ▶ Tehtävä palautetaan tavalla tai toisella koskemaan *matriisin* ominaisarvoja.

Ominaisarvotehtävä

Hieno homma! Käytetään MatLabin `eig`:iä.

Ominaisarvotehtävä

- ▶ Matriisit voivat olla aika isoja, ja *harvoja*:
 - ▶ Tarkkaa likiarvoa etsittävä “laajasta” avaruudesta.
 - ▶ Muistin ja laskentatehon kannalta tärkeää, ettei matriisin jokainen alkio poikkea nolasta.
 - ▶ Tavallisesti kiinnostavia ominaisarvoja vain muutama.
- ▶ Yleisalgoritmi kaikille ominaispareille epäkäytännöllinen.
Halutaan
 - ▶ Harvoille matriiseille tehokas tapa.
 - ▶ Ominaisparit yksi kerrallaan tai muutamia (satoja) kerrallaan *halutulta alueelta*.
- ▶ Ominaisarvotehtävien ratkaisu tiheille ja pienillekin matriiseille iteraatio. (Yli 5. asteen polynomin juuritehtävä).
 - ▶ MatLabin `roots` etsii polynomin juuret ominaisarvotehtävän ratkaisijalla, eikä Newtonilla.
- ▶ Tämä esitys ei edes pyri olemaan tyhjentävä, vaan intro!

Matriisien spektrilause

- ▶ n -kertaan- n -matriisille A löytyy n ominaisparia. Ne voidaan koota matriisimuotoon

$$AH = H\Lambda,$$

missä matriisin H sarakkeina ominaisvektorit ja diagonaalimatriisin Λ halkaisijalla ominaisarvot.

- ▶ H :n sarakkeiden järjestystä voi vaihdella (kunhan päivittää myös Λ :n).

Matriisien spektrilause

- ▶ Tärkeää: H :n *rangi on täysi*, eli se virittää määrittely- ja kuva-avaruuden.
- ▶ Siksi mille tahansa neliömatriisille on *spektrihajotelma*

$$A = H\Lambda H^{-1}.$$

- ▶ Hamiltonin operaattorista tuotettu matriisi on hermiittinen, eli $A^H = A$ pätee.
- ▶ Siksi
 - ▶ H unitaari ($HH^H = H^H H = I$).
 - ▶ Ominaisarvot reaalisia.

Potenssiinkorotusmenetelmä

Harvojen matriisien muuttaminen kallista,
niillä kertominen halpaa.

Potenssiinkorotusmenetelmä

- ▶ Satunnaisesti valittu vektori x kerrotaan k kertaa matriisilla A

$$v_k = A^k x.$$

- ▶ Esitetään x ominaisvektorien avulla:

$$x = \sum_{i=1}^n c_i h_i.$$

- ▶ Nyt siis pätee

$$v_k = A^k \sum_{i=1}^n c_i h_i = \sum_{i=1}^n c_i A^k h_i = \sum_{i=1}^n c_i \lambda_i^k h_i.$$

Potenssiinkorotusmenetelmä

- ▶ Moduliltaan suurin ominaisarvo on *dominanttti ominaisarvo*.
- ▶ Olkoon λ_1 dominantti. Jaetaan

$$v_k = \sum_{i=1}^n c_i \lambda_i^k h_i.$$

$|\lambda_1|^k$:lla ja erotetaan ensimmäinen termi:

$$\frac{v_k}{|\lambda_1|^k} = c_1 \left(\frac{\lambda_1}{|\lambda_1|} \right)^k h_1 + \sum_{i=2}^n c_i \left(\frac{\lambda_i}{|\lambda_1|} \right)^k h_i.$$

- ▶ Mitä pienempiä suhteet $\frac{|\lambda_i|}{|\lambda_1|}$, sitä nopeammin v_k suppenee dominantin ominaisvektorin suuntaiseksi.
- ▶ Käytännön algoritmi: $v \leftarrow Av$, $v \leftarrow v/|v|$, toistetaan.
 - ▶ Tällä Googlekin porskuttaa. Ns. Krylov-menetelmien teho perustuu myös osin tähän.

Rayleigh'n osamäärä

- ▶ Jos ominaisvektorin likiarvo tiedossa, miten estimoidaan ominaisarvoa?
- ▶ Ominaisarvoyhtälö voimaan mahdollisimman tarkasti:

$$Av - \lambda v \approx 0.$$

- ▶ $\min_{\lambda \in \mathbb{C}} \|Av - \lambda v\|_2.$
- ▶ Ratkaisu: Rayleigh'n osamäärä $\lambda = \frac{v^H Av}{v^H v}.$

Rayleigh'n osamäärä

Pärjätäänkö näillä?

Harjoitusvastustaja

- ▶ Harjoitusvastustaja: Vetyatomi, ei-relativistinen radiaalinen aaltoyhtälö ($l = 0$) atomiyksiköissä

$$\left(-\frac{1}{2} \frac{d^2}{dr^2} - \frac{1}{r} \right) u = Eu.$$

- ▶ Sidottujen tilojen ominaisarvot $-1/(2n^2)$, $n = 1, 2, \dots$
- ▶ Dimensio äärelliseksi differenssimenetelmällä:

$$\begin{array}{ccccccc} & & & & h & & \\ & & & & | & & \\ \times & & \times & & \times & & \times \\ | & & | & & | & & | \\ u_{k-1} & & u_k & & u_{k+1} & & \end{array}$$
$$u' \approx \frac{u_k - u_{k-1}}{h}$$
$$u''_k \approx \frac{u_{k+1} - 2u_k + u_{k-1}}{h^2}$$

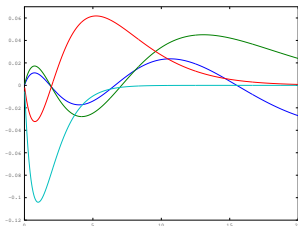
- ▶ Tridiagonaalimatriisi, ominaisvektorissa u :n pisteittäiset arvot.

Harjoitusvastustaja

- ▶ $r \in [0, 4]$ Bohrin sädettä, päädyissä ehto $u = 0$, sata pistettä.
- ▶ Pienimmät ominaisarvot:
6.53473 3.87331 1.82540 0.40390 -0.48403 .
- ▶ Yksi vaivainen negatiivinen ja paljon isompia positiivisia arvoja. Syy: “väkisin katkaistu” määrittelyalue.
- ▶ Maksimisäde 20, sata pistettä:
0.108730 0.014892 -0.050198 -0.124678 -0.495098
- ▶ Marginaalisesti parempi tilanne. Suurin ominaisarvo 49.901.

Harjoitusvastustaja

- ▶ Tarvitaan paljon isompi maksimisäde — siksi myös paljon enemmän kuin 100 pistettä.
- ▶ Säde 200, 10000 pistettä: negatiivisia ominaisarvoja 12
-0.0011057 -0.0031202 -0.0047611 -0.0061573
-0.0078124 -0.0102041 -0.0138889 -0.0199999
-0.0312498 -0.0555549 -0.1249969 -0.4999500
- ▶ Suurin ominaisarvo 5000.



Kuva : Neljä ensimmäistä aaltofunktiota

Potenssiinkorotusmenetelmä ei löydä pieniä negatiivisia ominaisarvoja. Suppenee kohti dominantin ominaisarvon 5000 vektoria.

Inverssi-iteraatio

- ▶ A :n pienimmät ominaisarvot kuvautuvat A^{-1} :n suurimmiksi.
- ▶ Ominaisvektorit säilyvät päikseen:

$$AA^{-1} = H\Lambda H^{-1}(H\Lambda H^{-1})^{-1} = H\Lambda H^{-1}H\Lambda^{-1}H^{-1} = I.$$

- ▶ Idea: Jos halutaan moduliltaan pienin ominaisarvo, ajetaan potenssiinkorotusmenetelmää käänteismatriisilla!

Ominaisarvojen siirto

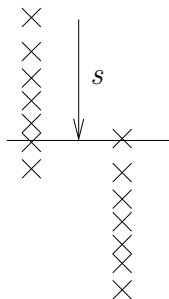
Muita ominaisvektoreita säilyttäviä jekkuja?

Ominaisarvojen siirto

- ▶ Matriisin $A - sI$ spektrihajotelma:

$$A - sI = H\Lambda H^{-1} - sI = H(\Lambda - sI)H^{-1}.$$

- ▶ Jokainen ominaisarvo λ_i siirtyy arvoon $\lambda_i - s$.
- ▶ Tällä pääsee käsiksi esimerkin negatiivisiin ominaisarvoihin: tehdään negatiivisimmasta ominaisarvosta dominantti.



Ominaisarvojen siirto

- ▶ Ongelma: Pelkällä siirrolla moni ominaisarvo lähellä dominanttia.
 - ▶ *Suppenee* hitaasti.

Koeajo 1

- ▶ Siirretään ominaisarvot luvulla -5000, jolloin saadaan dominantti ominaisarvo -5000,5. Montako kierrosta tarvitaan, että virhe ominaisarvoestimaatissa on N :nessä numerossa?
- ▶ 1:825:(3), 12:95.4:(2), 123:9.42:(1),
2315:0.499:(0), 28978:-0.40001:(-1),
40852:-0.490:(-2)
- ▶ $\sim 40\,000$ iteraatiota kahteen oikeaan numeroon!

Ominaisarvojen siirto ja kääntö

- ▶ Jos matriisin ominaisarvot ensin siirretään ja sitten vielä käännetään, saadaan ominaisarvot

$$\frac{1}{\lambda_i - s}.$$

- ▶ Miten saadaan λ_i :n ominaisvektori huikean dominantiksi?
- ▶ Millä matriisilla pitää tehdä ja mitä, jotta vektori selviää nopeasti potenssiinkorotusmenetelmällä?
- ▶ Jos siirtoa s vielä päivitetään sopivasti, estimaatti suppenee hirmuista vauhtia.

Ominaisarvojen siirto ja kääntö

Koeajo 2

- ▶ Siirretään -1:llä ja käännetään.
- ▶ 3:-0.22834, 4:-0.40948, 5:-0.47591, 6:-0.49385, 8:-0.49954:(-3), 10:-0.49992:(-4), 11:-0.49994:(-5), 13:-0.49995:(-6) 15:(-7)
- ▶ No niin! 7 numeroa 15 iteraatiolla. Noin 2 iteraatiota uutta oikeaa numeroa kohti.

Koeajo 3

- ▶ Aloitetaan siirrosta -1, ja päivitetään siirtoa varovasti Rayleigh'n osamäärän r suuntaan, eli $s = (10s + r)/11$.
- ▶ Suppenee arvoon -0.49995 7:llä askeleella.

Uusi oikea numero melkein joka iteraatiolla.

Useampi ominaisvektori kerralla

- ▶ Yleensä yritetään ratkaista useampi kuin yksi ominaispari.
- ▶ Tarpeet:
 1. Moninkertainen ominaisarvo
 2. Siirron vaihto kallista. Sama siirto kiihdyttää useamman vektorin suppenemista.
- ▶ Haasteet:
 1. Kaikki vektorit pyrkivät kohti dominanttia ominasvektoria.
 2. Jokin vektoreista supennut. Mitä tehdään?

Deflaatio

- ▶ Ominaisvektori h_1 supennut: Rayleigh'n osamäärä r toteuttaa

$$\frac{||Ah_1 - rh_1||}{|r|} \leq \delta.$$

Aika jättää tämä pari rauhaan.

- ▶ Tehtävä hajotettuna tunnettuun ja tuntemattomaan:

$$A \begin{bmatrix} h_1 & H_2 \end{bmatrix} = \begin{bmatrix} h_1 & H_2 \end{bmatrix} \begin{bmatrix} r & 0 \\ 0 & \Lambda_2 \end{bmatrix}$$

- ▶ Hermittinen A , eli $H^{-1} = H^H$ ja

$$A = \begin{bmatrix} h_1 r & H_2 \Lambda_2 \end{bmatrix} \begin{bmatrix} h_1^H \\ H_2^H \end{bmatrix} = h_1 r h_1^H + H_2 \Lambda_2 H_2^H.$$

- ▶ Matriisilla $A - h_1 r h_1^H$ yksi nollasta poikkeava ominaisarvo vähemmän: ominaispariksi $(h_1, 0)$.

Deflaatio

- ▶ Uudet ominaisvektorit h_1 :lle ortogonaalisia: $h_1^H H_2 = 0$.
⇒ Tiedetään mistä etsiä.
- ▶ Deflaatio nollaa kerralla useamman tunnetun ominaisarvon.
- ▶ Matriisia A ei oikeasti tarvitse muokata: deflaatio voidaan toteuttaa algoritmeissa implisiittisenä.

Yhteenveto

- ▶ Redusointi likiarvottehtäväksi matriisimuotoon.
- ▶ Iteraatio: parannetaan likiarvoa askelittain.
- ▶ Avainoperaatio matriisilla tai sen inverssillä kertominen.
- ▶ Jahdataan *dominanttia* ominaisparia.
- ▶ Toteutus ammattilaisten koodaamia moduleja käyttäen.
(Numeerinen lineaarialgebra on sudenkuoppia täynnä.)

Polkuja eteenpäin

- ▶ Erinomaisia kirjoja:
 - ▶ Lloyd N. Trefethen, David Bau: Numerical Linear Algebra, SIAM, Philadelphia, 1997.
 - ▶ Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst: Templates for the solution of algebraic eigenvalue problems: a practical guide, SIAM, Philadelphia, 2000.
- ▶ Avainsanoja:
 - ▶ Lanczosin menetelmä, Arnoldin menetelmä
 - ▶ Ritzin parit, Ritzin projektiot
 - ▶ Esikäsittely (preconditioning)
 - ▶ QR-iteraatio (tai aliavaruusiteraatio)
 - ▶ Yleistetty ominaisarvotehtävä
 - ▶ Jacobi-Davidsonin menetelmä
- ▶ Luotettavia lohkoja lineaarialgebraan
 - ▶ Matlab, Octave, SciPy
 - ▶ LAPACK
 - ▶ PETSc