Kurssimateriaalia K3/P3-kursille syksyllä 2005.
23. marraskuuta 2006
Heikki Apiola

1. Ortogonaalisuus, matriisityyppejä ja spektrejä

Reaalinen	Kompleksinen
Symmetrinen: $A^T = A$	Hermiittinen: $\overline{A}^T = A$
Vinosymmetrinen: $A^T = -A$	Vinohermiittinen $\overline{A}^T = -A$
Ortogonaalinen: $A^T = A^{-1}$	Unitaarinen $\overline{A}^T = A^{-1}$.

Merkintöjä: $A^\star = \overline{A}^T$, joskus merkitään myös A^H ("hermitointi"). Myös A' on käytössä mm. Matlab:ssa.

Käyttämällä merkintää $A^\star$ (tai vastaavaa synonyymia), voidaan yllä olevat ehdot lausua yhtenäisesti: Hermiittinen tai symmetrisen ehto on: $A^\star = A$, jne.

Aloitamme lemmalla, joka on hyödyllinen monissa seuraavissa todistuksissa. Huomaa, että vaikka olisimme kiinnostuneita vain reaalisista matriiseista, on alla olevat laskut suoritettava kompleksiluvuilla, koska ominaisarvot ja ominaisvektorit saattavat sisältää kompleksilukuja.

Lemma 1.1. *Olkoon A $(m \times n)$- matriisi ja olkoot $\boldsymbol{u} \in \mathbb{C}^n$ ja $\boldsymbol{v} \in \mathbb{C}^m$. Tällöin*

$$\langle A\boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}, A^\star \boldsymbol{v}\rangle .$$

Tod.

$$\langle A\boldsymbol{u}, \boldsymbol{v}\rangle = (A\boldsymbol{u})^T \overline{v} = \boldsymbol{u}^T A^T \overline{v} \underbrace{=}_{A^T = \overline{A^\star}} \boldsymbol{u}^T \overline{A^\star} \overline{\boldsymbol{v}} = \boldsymbol{u}^T\, \overline{A^\star\, \boldsymbol{v}} = \langle \boldsymbol{u}, A^\star \boldsymbol{v}\rangle .$$

□

1.1. **Ortogonaaliset matriisit.**

Kyseessä on reaalinen neliömatriisi U, jolle $U^T U = U U^T = I$, ts. U on kääntyvä ja $U^{-1} = U^T$.

Ortogonaalisen matriisin prototyyppi on tason kierto. Toinen tyypillinen on heijastus jonkin origon kautta kulkevan suoran suhteen. Näiden matriisit ovat

$$\begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \text{ ja } \begin{bmatrix} \cos(\phi) & \sin(\phi) \\ \sin(\phi) & -\cos(\phi) \end{bmatrix}.$$

Edellisen determinantti = 1 ja jälkimmäisen -1. Determinanttien kertosäännön perusteella ortogonaaliselle matriisille pätee aina: $(\det A)^2 = \det(AA^T) = \det I = 1$, joten $\det A = \pm 1$.

Lause 1.2 (KRE Thm. 3)**.** *Reaalinen neliömatriisi A on ortogonaalinen, jos ja vain jos sen sarakkeet muodostavat ortonormaalin joukon, jos ja vain jos sen rivit muodostavat ortonormaalin joukon.*

Tod. Eipä muuta, kuin ajatellaan matriisitulon määritelmää yhtälössä $A^T A = I$. Jos merkitään rivivektoreita $a_{i.}$ ja sarakevektoreita $a_{.j}$, niin

$$(A^T A)_{ij} = a_{i.}^T a_{.j} = \langle a_{.i}, a_{.j}\rangle .$$

Jos merkitään δ_{ij}:llä ns. *Kroneckerin deltaa*, joka lyhyesti sanottuna tarkoittaa yksikkömatriisin alkiota $(I)_{ij}$, niin ehto $A^T A = I$ merkitsee samaa kuin $(A^T A)_{ij} = \delta_{ij}$, eli $\langle a_{\cdot i}, a_{\cdot j}\rangle = \delta_{ij}$, joka tarkoittaa juuri sarakevektorien ortonormaalisuutta.

Rivivektoreilla aivan vastaavasti tarkastelemalla tuloa $AA^T = I$.

Kääntäen muistamme "lineaarialgebran ihmeen": $r(A^T) = r(A)$, jonka perusteella jo toinen ehdoista $AA^T = I$ tai $A^T A = I$ takaa käänteismatriisin olemassaolon, eli sen, että molemmat ovat voimassa. Siten sarakkeiden (ja yhtä hyvin rivien) ortonormaalisuudesta seuraa matriisin ortogonaalisuus.

$\square$

Lause 1.3 (KRE s. 382, Lop. s. 730)**.** *Ortogonaalinen kuvaus U säilyttää sisätulon ja siten vektorin normin, ts.* $\langle U\boldsymbol{u}, U\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle$.

Tod. Olkoon $\boldsymbol{u} = U\boldsymbol{a}$, $\boldsymbol{v} = U\boldsymbol{b}$.

$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle U\boldsymbol{a}, U\boldsymbol{b}\rangle = \langle U^\star U\boldsymbol{a}, \boldsymbol{b}\rangle = \langle \boldsymbol{a}, \boldsymbol{b}\rangle$, koska $U^\star U = I$.

Erityisesti $||\boldsymbol{u}||^2 = \langle \boldsymbol{u}, \boldsymbol{u}\rangle = \langle \boldsymbol{a}, \boldsymbol{a}\rangle = ||\boldsymbol{a}||^2$.

$\square$

Listataan myös edellä laskettu det −ominaisuus lauseeksi:

Lause 1.4. *Ortogonaalisen matriisin* det $= \pm 1$.

Lause 1.5. *Ortogonaalisen matriisin U ominaisarvot ovat yksikköympyrällä, ts, $|\lambda| = 1$ kaikille ominaisarvoille λ.*

Tod. Kuten edellä totesimme, $||U\boldsymbol{u}|| = ||\boldsymbol{u}||$.

Jos $\boldsymbol{u}, \lambda$ on omisarvo/-vektoripari, niin $U\boldsymbol{u} = \lambda\boldsymbol{u}$, $\boldsymbol{u} \neq \boldsymbol{0}$.

$||\boldsymbol{u}|| = ||U\boldsymbol{u}|| = |\lambda|||\boldsymbol{u}||$. Koska $||\boldsymbol{u}|| \neq 0$, voidaan sillä jakaa, ja saadaan $|\lambda| = 1$.

$\square$

Huomautus 1.1. *Yllä oleva todistus menee sanasta sanaan myös unitaariselle.*

1.2. **Symmetrinen (hermiittinen) ja vinosymmetrinen matriisi.**

Lause 1.6 (Lop s. 732, KRE8 s. 387)**.** *(Hiukan eri muodot) Olkoon A symmetrinen reaalinen matriisi (kohdassa (1b) vinosymmetrinen).*

(1) A:n ominaisarvot ovat reaaliset.
(1 b) Vinosymmtrisen A :n ominaisarvot ovat puhtaasti imaginaariset
(2) A:n eri ominaisarvoja vastaavat ominaisvektorit ovat ortogonaaliset.
(3) A:n ominaisvektoreista voidaan muodostaa jopa ortonormaali kanta $\mathbb{R}^n$:lle.
(4) A on diagonalisoituva, diagonalisoiva matriisi voidaan valita ortogonaaliseksi, ts. voidaan kirjoittaa

$A = SDS^T$*, missä S on ortogonaalinen.*

Tod. (1) Olkoon $A\boldsymbol{x} = \lambda\boldsymbol{x}$, $\boldsymbol{x} \neq \boldsymbol{0}$.

Huomaa, että seuraavat laskut on tehtävä kompleksiluvuilla, koska emmehän voi todistuksessa käyttää väitettä hyväksi! :-)

Turvaudumme tuohon mainioon lemmaan, jonka jälkeen kaikki risuaidat saadaan jättää taakse.

Aloitetaan kertomalla ominaisarvon/vektorin määrittelevä yhtälö sisätulon mielessä $\boldsymbol{x}$:llä:

$$\langle A\boldsymbol{x}, \boldsymbol{x} \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \lambda \|\boldsymbol{x}\|^2.$$

Tuon mainitun mainion mukaan vasen puoli on:

$$\langle \boldsymbol{x}, A^T \boldsymbol{x} \rangle \underbrace{=}_{A^T=A} \langle \boldsymbol{x}, A\boldsymbol{x} \rangle = \langle \boldsymbol{x}, \lambda \boldsymbol{x} \rangle = \overline{\lambda} \langle \boldsymbol{x}, \boldsymbol{x} \rangle .$$

Siis $\overline{\lambda}||\boldsymbol{x}||^2 = \lambda||\boldsymbol{x}||^2$, josta jakamalla $||\boldsymbol{x}||^2$: lla ($\neq 0$) saadaan $\lambda = \overline{\lambda}$, eli $\lambda \in \mathbb{R}$.

Vinosymmtrinen tapaus menee aivan samoin, paitsi $-$ -merkki, joka antaa johtopäätöksen: $\overline{\lambda} = -\lambda$, joka merkitsee sitä, että reaaliosa on nolla, kuten väitettiin.

(2) Olkoon $A\boldsymbol{x} = \lambda \boldsymbol{x}$, $A\boldsymbol{y} = \mu \boldsymbol{y}$, $\quad \lambda \neq \mu$.

$\langle A\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, A\boldsymbol{y} \rangle$

$\langle A\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \lambda \boldsymbol{x}, \boldsymbol{y} \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{y} \rangle$

$\langle \boldsymbol{x}, A\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mu \boldsymbol{y} \rangle = \mu \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ ($\overline{\mu} = \mu$). Koska $\lambda \neq \mu$, on oltava $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.

(3) Tämä on syvällisempi tulos, todistetaan kylläkin L3:ssa (kts. [TE]), perustuu ns. Schur'n hajoitelmaan.

(4) Kanta voidaan valita ortonormaaliksi, koska eri ominaisarvoja vastaavat ominaisavaruudet ovat keskenään ortogonaalisia. Kussakin ominaisavaruudessa voidaan suorittaa edellä esitetty *Gram-Schmidt'n* ortonormalisointi, jolloin saadaan koko avaruuden ON kanta. [1]

A voidaan siis diagonalisointilauseen mukaan kirjoittaa muotoon $A = VDV^{-1}$, missä V:n sarakkeet ovat A :n ominaisvektoreita (samassa järjestyksessä kuin vastaavat ominaisarvot D :ssä.

Voidaan siis valita $V : n$ sarakkeet ortonormaaleiksi, eli V ortogonaaliseksi, jolloin $V^{-1} = V^T$.

□

Huomautus 1.2. *Symmetrisen matriisin diagonaaliesitys ei automaattisesti ole muotoa VDV^T, vaan sitä varten on huolehdittava, että*
(a) V:n sarakkeet on normeerattu ykkösiksi.
(b) Useampikertaisia ominaisarvoja vastaavat ominaisvektorit on ortonormeerattu.

Huomautus 1.3. *Jos käytämme jotain ominaisarvot laskevaa ohjelmaa, emme voi (ilman dokumentointiin perehtymistä) tietää, saadaanko automaattisesti tällainen muoto.* MATLAB*:n tapauksessa help-tekstikään ei kerro. Esimerkki-istunto:*

```
>> A=rand(10,10);     % 10 x 10-satunnaismatriisi.
>> A=A*A';            % Eräs tapa tehdä symmetrinen (toinen: A+A')
>> [V,D]=eig(A);

>> V*V'
Columns 1 through 8

    1.0000   -0.0000   -0.0000   -0.0000    0.0000   -0.0000    0.0000   -0.0000
   -0.0000    1.0000   -0.0000    0.0000    0.0000    0.0000   -0.0000    0.0000
   -0.0000   -0.0000    1.0000   -0.0000    0.0000    0.0000    0.0000   -0.0000
   -0.0000    0.0000   -0.0000    1.0000    0.0000   -0.0000    0.0000   -0.0000
    0.0000    0.0000    0.0000    0.0000    1.0000   -0.0000   -0.0000    0.0000
   -0.0000    0.0000    0.0000   -0.0000   -0.0000    1.0000    0.0000    0.0000
```

[1] ON: ortonormaali, OG: ortogonaalinen

```
    0.0000   -0.0000    0.0000    0.0000   -0.0000    0.0000    1.0000    0.0000
   -0.0000    0.0000   -0.0000   -0.0000    0.0000    0.0000    0.0000    1.0000
    0.0000    0.0000    0.0000    0.0000    0.0000    0.0000    0.0000   -0.0000
    0.0000    0.0000   -0.0000   -0.0000   -0.0000    0.0000   -0.0000   -0.0000

>> format long   % Suurempi tulostustarkkuus.
>> V*V'; ans(1:5,1:5)

ans =

   1.00000000000000  -0.00000000000000  -0.00000000000000  -0.00000000000000   0.00000000000000
  -0.00000000000000   1.00000000000000  -0.00000000000000   0.00000000000000   0.00000000000000
  -0.00000000000000  -0.00000000000000   1.00000000000000  -0.00000000000000   0.00000000000000
  -0.00000000000000   0.00000000000000  -0.00000000000000   1.00000000000000   0.00000000000000
   0.00000000000000   0.00000000000000   0.00000000000000   0.00000000000000   1.00000000000000

           % Laskentatarkkuuden puitteissa V on ortogonaalinen.
>> format  % Palataan alkuperäiseen
>> diag(D)'; ans(1:9)  % Katsotaan vain 9 ekaa, että mahtuu sivulle.

ans =

    0.0042    0.0385    0.1929    0.2121    0.3841    0.5493    0.9745    1.5073    1.5950
   % Ominaisarvot ovat erillisiä.
```

Tämä koe oli vain osittain paljastava. On luonnollista, että satunnaisesti generoidun matriisin ominaisarvot ovat erilliset, niinpä ominaisvektorit ovat automaattisesti ortogonaaliset (koska teimme matriisista symmetrisen), ja kun MATLAB *aina normeeraa, niin ne ovat ortonormaalit.*

Yleisesti voidaan asia hoitaa varmasti oikein soveltamalla $V-$matriisiin funktiota `orth`.

```
>> help orth

 ORTH   Orthogonalization.
    Q = ORTH(A) is an orthonormal basis for the range of A.
    That is, Q'*Q = I, the columns of Q span the same space as
    the columns of A, and the number of columns of Q is the
    rank of A.

    See also SVD, RANK, NULL.
```