

9 OMINAISARVOTEHTÄVÄ

Ch:EigSystem

9.1 Kompleksinen lineaariavaruus

Sec:CMatrix

9.1.1 Kompleksiluvut

Pian tulemme tarvitsemaan kompleksisen lineaariavaruuden alkeita. Tätä varten käymme läpi kompleksilukujen alkeet. Asiasta on oma kurssi, joten nyt toteamme asioita vain sen verran, että saamme tässä kurssissa tarvittavat asiat esitettyä.

Kompleksilukujen lukupariesitys. Kompleksiluvut voidaan esittää lukupareina

$$\mathbb{C} = \{(x, y) \mid x, y \in \mathbb{R}\}$$

joiden joukossa on määritelty yhteen- ja kertolasku seuraavasti

$$(x_1, y_1) \hat{+} (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \quad (9.1)$$

$$(x_1, y_1) \hat{*} (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) \quad (9.2)$$

Kolmikko $(\mathbb{C}, \hat{+}, \hat{*})$ on kunta (ks. sivu 102). Kunnassa laskeminen on helppoa, kunhan kompleksilukujen esittäminen saadaan helpoksi seuraavalla merkinnnällä.

Kompleksilukujen esitys imaginääriyksikön avulla. Sovimme seuraavan lyhennysmerkinnän

$$\begin{aligned} (0, 1) &=_{def} i = \text{imaginääriyksikkö} \\ (1, 0) &=_{def} 1 \\ \longrightarrow (x, y) &= x + iy \end{aligned}$$

Imaginääriyksikön tärkein ominaisuus, on

$$i^2 = -1$$

Tämän jälkeen kompleksiluvuilla laskeminen sujuu automaattisesti oikein. Laskusäännöt (9.1) ja (9.2) saavat nyt esityksen

$$\begin{aligned} (x_1 + iy_1) + (x_2 + iy_2) &= x_1 + x_2 + i(y_1 + y_2) \\ (x_1 + iy_1)(x_2 + iy_2) &= x_1 x_2 + ix_1 y_2 + ix_2 y_1 + i^2 y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1) \end{aligned}$$

Määritelmä 220 Kompleksiluvun $z = x + iy \in \mathbb{C}$ reaaliosa on

$$\operatorname{Re} z = x.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ imaginaariosa on

$$\operatorname{Im} z = y.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ liittoluku on

$$z^* = x - iy.$$

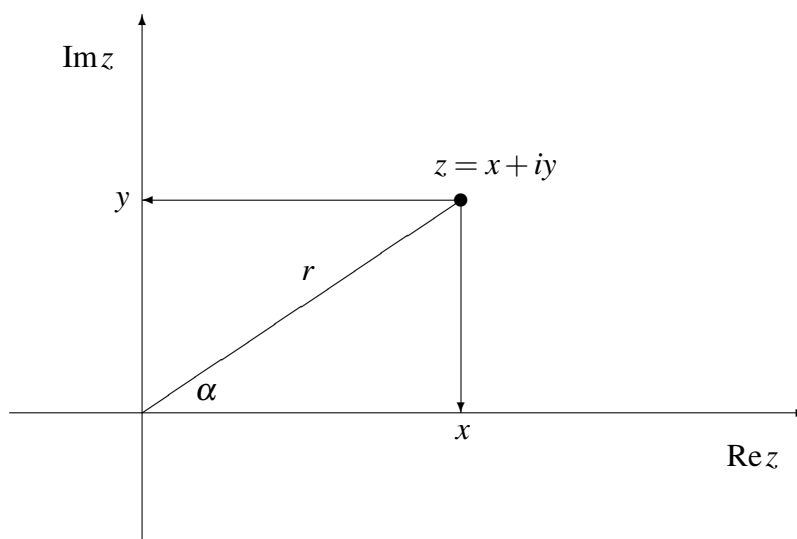
Kompleksiluvun $z = x + iy \in \mathbb{C}$ moduli on

$$|z| = \sqrt{x^2 + y^2} = \sqrt{zz^*}.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ argumentti on

$$\arg z = \arctan\left(\frac{y}{x}\right).$$

Kompleksitaso. Kompleksiluku $z = x + iy$ ”sijoitetaan” xy -tasoon pisteeseen (x, y) . Jos nyt $r = |z|$ ja $\alpha = \arg z$, niin suoraan kuvasta nähdään, että r on pisteen etäisyys origosta ja α on pisteen paikkavektorin ja x -akselin välinen kulma. Silloin



$$x = r \cos \alpha \quad \text{ja} \quad y = r \sin \alpha$$

9.1.2 Kompleksiset vektorit ja matriisit

Kompleksiset vektorit $\mathbb{C}^n$ ja matriisit määritellään nyt kaavioiksi, joihin on sijoitettu kompleksilukuja. Suora lasku osoittaa, että viisikko $(\mathbb{C}^n, \mathbb{C}, +, *, \circ)$ on kompleksiker-toiminen lineaariavaruus.

Muutama asia pitää määritellä uudelleen.

Määritelmä 221 (1) Matriisin $\mathbf{A} = (a_{ij})$ kompleksikonjugaatti on matriisi

$$\mathbf{A}^* = (a_{ij}^*).$$

(Jokainen kaavion luku siis konjugoidaan.)

(2) $m \times n$ -matriisin $\mathbf{A}$ hermitoitu matriisi on

$$\mathbf{A}^H = (\mathbf{A}^*)^T$$

(Matriisi siis konjugoidaan ja transponoidaan.)

(3) Kahden kompleksisen vektorin $\vec{z} \in \mathbb{C}^n$ ja $\vec{w} \in \mathbb{C}^n$ sisätulo määritellään kaavalla

$$\langle \vec{z} | \vec{w} \rangle = z_1^* w_1 + z_2^* w_2 + \cdots + z_n^* w_n = \vec{z}^H \vec{w}$$

Tehdyt määritelmät eivät muuta aiemmin tehtyjä reaalisten vektoreiden ja matriisien vastaavia määritelmiä. Jos nimittäin kaaviossa on vain reaalilukuja, niin herminointi palautuu transponoinniksi.

Kompleksisten vektorien ja matriisien tapauksessa erään aiemmat kaavat saavat uuden asun. Seuraavat kaavat ovat suoria seuraksia tehdyistä määritelmistä, joten emme esitä niille erillisiä todistuksia.

$$(\mathbf{A}^H)^H = \mathbf{A} \quad (9.3)$$

$$(\mathbf{A} + \mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H \quad (9.4)$$

$$(\lambda \mathbf{A})^H = \lambda^* \mathbf{A}^H \quad (9.5)$$

$$(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H \quad (9.6)$$

$$\langle \vec{z} | \vec{w} \rangle = \langle \vec{w} | \vec{z} \rangle^* \quad (9.7)$$

$$\langle \lambda \vec{z} | \vec{w} \rangle = \lambda^* \langle \vec{z} | \vec{w} \rangle \quad (9.8)$$

$$\langle \vec{z} | \lambda \vec{w} \rangle = \lambda \langle \vec{z} | \vec{w} \rangle \quad (9.9)$$

$$\langle \lambda \vec{z} | \vec{w} \rangle = \langle \vec{z} | \lambda^* \vec{w} \rangle \quad (9.10)$$

$$\langle \mathbf{A} \vec{z} | \vec{w} \rangle = \langle \vec{z} | \mathbf{A}^H \vec{w} \rangle \quad (9.11)$$

$$\|\vec{z}\| = \sqrt{\langle \vec{z} | \vec{z} \rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2} \in \mathbb{R} \quad (9.12)$$

Määritelmä 222 Neliömatriisi $\mathbf{A}$ on

$$\begin{aligned} \text{symmetrinen, jos} \quad & \mathbf{A} = \mathbf{A}^T, \\ \text{hermiittinen, jos} \quad & \mathbf{A} = \mathbf{A}^H, \\ \text{ortogonaalinen, jos} \quad & \mathbf{A}^{-1} = \mathbf{A}^T, \\ \text{unitaarinen, jos} \quad & \mathbf{A}^{-1} = \mathbf{A}^H. \end{aligned}$$

9.2 Matriisin ominaisarvot

Sec:MatEig

9.2.1 Perusmääritelmät

Tässä kappaleessa oletamme skalaarit, vektorit ja matriisit kompleksisiksi. Vaikka lähtökohta olisikin reaalinen matriisi, niin laskujen aikana saattaa tulla esiin kompleksilukuja ja me nyt hyväksymme sen. Tämä helpottaa asian tarkastelua suuresti.

Seuraavaksi tarkastelemme teknistä ongelmaa, joka on yksinkertaisissa tapauksissaan mahdollista ratkaista suoraan laskemalla. Ensin opimme ratkaisemaan ominaisarvotetävän helpossa tapauksessa. Myöhemmin opimme lisää ratkaisun ominaisuuksista, ja niistä ehdoista joilla ratkaisu on olemassa.

Määritelmä 223 Olkoot $\mathbf{A}$ $n \times n$ -matriisi, $\vec{x} \in \mathbb{R}^n, \vec{x} \neq \vec{0}$ vektori ja $\lambda \in \mathbb{R}$ reaaliluku. Jos

$$\mathbf{A}\vec{x} = \lambda\vec{x}, \tag{9.13} \quad \text{eigveceq}$$

niin sanomme, että λ on matriisin $\mathbf{A}$ ominaisarvo (engl. *eigenvalue*) ja vektori $\vec{x}$ on ominaisarvoon λ liittyvä matriisin $\mathbf{A}$ ominaisvektori (engl. *eigenvector*).

Jos λ on matriisin $\mathbf{A}$ ominaisarvo, niin homogeenisella yhtälöryhmällä

$$(\mathbf{A} - \lambda \mathbf{I})\vec{x} = \vec{0}$$

on ei-triviaali ratkaisu. Tämä on mahdollista vain jos

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \tag{9.14}$$

$$\Leftrightarrow \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{9.15}$$

Ominaisarvotehtävän ratkaisu tehdään $(n + 1)$:ssä vaiheessa:

(1) Ensin ratkaistaan ominaisarvot karakteristisesta yhtälöstä

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0. \quad (9.16)$$

Karakteristinen yhtälö on λ :n suhteen astetta n oleva yhtälö, joten sillä on korkeintaan n juurta (ominaisarvoa).

(2) Jokaista karakteristisen yhtälön juurta (ominaisarvoa) λ kohti ratkaistaan vastaava ominaisvektori homogeenisesta yhtälöryhmästä

$$\mathbf{A}\vec{x} = \lambda\vec{x}. \quad (9.17)$$

c4s15es1

Esimerkki 224 Ratkaise matriisin

$$\mathbf{A} = \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}$$

ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

$$\Leftrightarrow \begin{vmatrix} 2-\lambda & 2 & 0 \\ -1 & 3-\lambda & 2 \\ 0 & 2 & 2-\lambda \end{vmatrix} = 0$$

$$\Leftrightarrow (2-\lambda) \begin{vmatrix} 3-\lambda & 2 \\ 2 & 2-\lambda \end{vmatrix} - 2 \begin{vmatrix} -1 & 2 \\ 0 & 2-\lambda \end{vmatrix} = 0$$

$$\Leftrightarrow (2-\lambda)[(3-\lambda)(2-\lambda) - 2 \cdot 2] - 2[-1 \cdot (2-\lambda) - 0 \cdot 2] = 0$$

$$\Leftrightarrow (2-\lambda)[\lambda^2 - 5\lambda + 2] + 2(2-\lambda) = 0$$

$$\Leftrightarrow (2-\lambda)(\lambda^2 - 5\lambda + 4) = 0$$

$$\Leftrightarrow (2-\lambda) = 0 \text{ tai } (\lambda^2 - 5\lambda + 4) = 0$$

$$\Leftrightarrow \lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 4$$

(2) Ominaisarvoa $\lambda_1 = 1$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned}
 & \mathbf{A}\vec{x} = \lambda_1 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 1 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 + 2x_2 = x_1 \\ -x_1 + 3x_2 + 2x_3 = x_2 \\ 2x_2 + 2x_3 = x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} [x_1] + 2x_2 = 0 \\ -x_1 + 2x_2 + 2x_3 = 0 \\ 2x_2 + x_3 = 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \end{array} \\
 \Leftrightarrow & \begin{cases} (x_1) + 2x_2 = 0 \\ + [4x_2] + 2x_3 = 0 \\ 2x_2 + x_3 = 0 \end{cases} \begin{array}{l} \\ -1/2 \\ \leftarrow \end{array} \\
 \Leftrightarrow & \begin{cases} (x_1) + 2x_2 = 0 \\ + (4x_2) + 2x_3 = 0 \\ 0 = 0 \end{cases} \begin{array}{l} \rightarrow x_1 = -2x_2 \\ \rightarrow x_3 = -2x_2 \end{array}
 \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} -2a \\ a \\ -2a \end{pmatrix}, a \in \mathbb{R}.$$

Yleensä vektori valitaan niin, että sen pituus on yksi. Siis $(-2a)^2 + a^2 + (-2a)^2 = 1$, josta ratkeaa $a = 1/3$. Siis lopullinen ominaisarvoon $\lambda_1 = 1$ liittyvä ominaisvektori on

$$\vec{x}_1 = \begin{pmatrix} -2/3 \\ 1/3 \\ -2/3 \end{pmatrix}.$$

(3) Ominaisarvoa $\lambda_2 = 2$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned}
 & \mathbf{A}\vec{x} = \lambda_2 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 + 2x_2 = 2x_1 \\ -x_1 + 3x_2 + 2x_3 = 2x_2 \\ 2x_2 + 2x_3 = 2x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} 2x_2 = 0 \\ -x_1 + x_2 + 2x_3 = 0 \\ 2x_2 = 0 \end{cases} \begin{array}{l} \rightarrow x_2 = 0 \\ \rightarrow x_1 = 2x_3 \end{array}
 \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} 2b \\ 0 \\ b \end{pmatrix}, b \in \mathbb{R}.$$

Vektori valitaan taas niin, että sen pituus on yksi. Siis $(2b)^2 + 0^2 + (b)^2 = 1$, josta ratkeaa $b = 1/\sqrt{5}$. Siis lopullinen ominaisarvoon $\lambda_2 = 2$ liittyvä ominaisvektori on

$$\vec{x}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix}.$$

(4) Ominaisarvoa $\lambda_3 = 4$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned} \mathbf{A}\vec{x} &= \lambda_3\vec{x} \\ \Leftrightarrow \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 4 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ \Leftrightarrow \begin{cases} 2x_1 + 2x_2 &= 4x_1 \\ -x_1 + 3x_2 + 2x_3 &= 4x_2 \\ 2x_2 + 2x_3 &= 4x_3 \end{cases} \\ \Leftrightarrow \begin{cases} [-2x_1] + 2x_2 &= 0 \\ -x_1 - x_2 + 2x_3 &= 0 \\ 2x_2 - 2x_3 &= 0 \end{cases} \begin{array}{l} -1/2 \\ \leftarrow \\ \end{array} \\ \Leftrightarrow \begin{cases} (-2x_1) + 2x_2 &= 0 \\ [-2x_2] + 2x_3 &= 0 \\ 2x_2 - 2x_3 &= 0 \end{cases} \begin{array}{l} \\ +1 \\ \leftarrow \end{array} \\ \Leftrightarrow \begin{cases} (-2x_1) + 2x_2 &= 0 \\ (-2x_2) + 2x_3 &= 0 \\ 0 &= 0 \end{cases} \begin{array}{l} \rightarrow x_1 = x_2 \\ \rightarrow x_3 = x_2 \\ \end{array} \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} c \\ c \\ c \end{pmatrix}, c \in \mathbb{R}.$$

Vektori valitaan taas niin, että sen pituus on yksi. Siis $c^2 + c^2 + c^2 = 1$, josta ratkeaa $c = 1/\sqrt{3}$. Siis lopullinen ominaisarvoon $\lambda_3 = 4$ liittyvä ominaisvektori on

$$\vec{x}_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$$

Vastauksena ongelmaan voimme siis luetella ominaisarvot ja niihin liittyvät ominaisvektorit

$$\lambda_1 = 1, \vec{x}_1 = \begin{pmatrix} -2/3 \\ 1/3 \\ -2/3 \end{pmatrix}, \text{ ja } \lambda_2 = 2, \vec{x}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix}, \text{ ja } \lambda_3 = 4, \vec{x}_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$$

Saman matriisin ominaisarvot ja ominaisvektorit saadaan Matlabin komennolla

$[V,D] = \text{eigs}(A)$ seuraavasti:

```
>> A = [2 2 0 ; -1 3 2 ; 0 2 2]
```

```
A =
```

```
    2    2    0
   -1    3    2
    0    2    2
```

```
>> [V,D] = eigs(A)
```

```
V =
```

```
    0.5774   -0.8944    0.6667
    0.5774   -0.0000   -0.3333
    0.5774   -0.4472    0.6667
```

```
D =
```

```
    4.0000         0         0
         0    2.0000         0
         0         0    1.0000
```

```
>>
```

Siis matriisin V sarakkeet ovat ominaisvektoreita ja diagonaalimatriisin D diagonaalialkiot ovat vastaavat ominaisarvot.

On melko tavallista, että karakteristisella yhtälöllä on kompleksisia juuria. Nämä juuret ovat myös ominaisarvoja ja niitä vastaavat ominaisvektorit ovat yleensä myös kompleksisia. Kompleksisen esimerkin laskeminen käsin on melko työläs tehtävä, joten tyydymme nyt tietokoneella laskettuun esimerkkiin.

c4s15es2

Esimerkki 225 >> A = [2 2 0 ; -1 3 2 ; 0 2 -1]

```
A =
```

```
    2    2    0
   -1    3    2
    0    2   -1
```

```
>> [V,D] = eigs(A)
```

```
V =
```

```
    0.8165         0.8165         0.1851
    0.3585 - 0.3711i    0.3585 + 0.3711i   -0.3476
    0.2178 - 0.1403i    0.2178 + 0.1403i    0.9192
```



```

D =
    2.8782 - 0.9089i    0    0
         0    2.8782 + 0.9089i    0
         0    0    -1.7563
>>

```

Edellisen esimerkin matriisi ei näyttänyt hankalalta, mutta silti ominaisarvot ja ominaisvektorit ovat kompleksiset. Tässä vaiheessa helposti tulee epätoivoinen tunne siitä, että asia on liian mutkikas.

Jotta epätoivo ei ahdistaisi toteamme kaksi tulosta, jotka perustelemme pian:

(1) Symmetrisen matriisin ($\mathbf{A}^T = \mathbf{A}$) ominaisarvot ovat reaaliset.

(2) Ns. normaalin matriisin ($\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T$) ominaisvektoreista voidaan muodostaa $\mathbb{R}^n$:n kanta.

Useissa sovelluksissa tarkasteltava matriisi on symmetrinen, ja tiedämme siis, että ominaisarvot eivät ole kompleksisia.

Kun karakteristisen yhtälön, $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$, kaikki termit viedään vasemmalle puolelle yhtälöä, niin sinne syntyy karakteristinen polynomi $P(\lambda)$. Karakteristinen polynomi voidaan aina myös saattaa tulomuotoon

$$\begin{aligned}
 P(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{I}) \\
 &= (-1)^n \lambda^n + \alpha_{n-1} \lambda^{n-1} + \cdots + \alpha_0
 \end{aligned} \tag{9.18}$$

$$= (-1)^n (\lambda - \lambda_1)^{k_1} (\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_q)^{k_q} \tag{9.19}$$

Jos matriisin karakteristinen polynomi on muotoa ^{karpoltulo} (9.19), niin sanomme, että ominaisarvon λ_j algebraalinen kertaluku on k_j . Kompleksianalyysin kurssilla opimme, että kertalukujen summa on $k_1 + k_2 + \cdots + k_q = n$.

$\mathbb{C}^n$:n lineaarinen aliavaruus $\ker(\mathbf{A} - \lambda_j \mathbf{I})$ sisältää kaikki ominaisarvoon λ_j liittyvät ominaisvektorit ja nollavektorin (joka ei ole ominaisvektori). Tämän aliavaruuden dimensiota

$$\rho(\lambda_j) = \dim \ker(\mathbf{A} - \lambda_j \mathbf{I}) \geq 1$$

sanomme ominaisarvon λ_j geometriseksi kertaluvuksi. Geometrinen kertaluku kertoo, miten monta lineaarisesti riippumatonta ominaisvektoria voidaan liittää ominaisarvoon λ_j .

Normaalisti matriisin ominaisarvojen algebralliset- ja geometriset kertaluvut yhtyvät. Jos näin ei tapahdu, niin sanomme, että matriisi on defektiivinen. Defektiivisten matriisien tutkiminen ylittää tämän peruskurssin tavoitteet. Termi mainitaan tässä, jotta opiskelija osaa alkaa etsiä lisätietoa kirjallisuudesta, jos laskentaohjelma antaa termin osana virheilmoitusta.

c4s15es3

Esimerkki 226 Tutkitaan matriisin

$$\mathbf{M} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

Ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö saa muodon

$$\begin{aligned} \det(\mathbf{M} - \lambda \mathbf{I}) = 0 &\Leftrightarrow \begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 3-\lambda & 1 \\ 0 & 1 & 3-\lambda \end{vmatrix} = 0 \\ &\Leftrightarrow (2-\lambda) \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = 0 \\ &\Leftrightarrow (2-\lambda)((3-\lambda)(3-\lambda) - 1) = 0 \\ &\Leftrightarrow (\lambda - 2)^2(\lambda - 4) = 0 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ algebrallinen kertaluku on 2, ja ominaisarvon $\lambda_2 = 4$ algebrallinen kertaluku on 1.

(2) Seuraavaksi määritämme ominaisarvoon $\lambda_1 = 2$ liittyvät ominaisvektorit.

$$\begin{aligned}
 & \mathbf{M}\vec{x} = \lambda_1 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 & = 2x_1 \\ & 3x_2 + x_3 = 2x_2 \\ & x_2 + 3x_3 = 2x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} & 0 = 0 & \rightarrow x_1 = \text{vapaa} \\ x_2 + x_3 & = 0 & \rightarrow x_2 = -x_3 \\ x_2 + x_3 & = 0 \end{cases} \\
 \Leftrightarrow & \begin{cases} x_1 & = x_1 \\ x_2 & = -x_3 \\ x_3 & = x_3 \end{cases} \\
 \Leftrightarrow & \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}
 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ geometrinen kertaluku on 2 ja ominaisvektorien virittäjiksi voidaan valita

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{ja} \quad \vec{x}_2 = \begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

(3) Seuraavaksi määritämme ominaisarvoon $\lambda_2 = 4$ liittyvät ominaisvektorit.

$$\begin{aligned}
 & \mathbf{M}\vec{x} = \lambda_2 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 & = 4x_1 \\ & 3x_2 + x_3 = 4x_2 \\ & x_2 + 3x_3 = 4x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} -2x_1 & = 0 & \rightarrow x_1 = 0 \\ & -x_2 + x_3 = 0 & \rightarrow x_2 = x_3 \\ & x_2 + -x_3 = 0 \end{cases}
 \end{aligned}$$

Ominaisvektoriksi voidaan valita

$$\vec{x}_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Sama matlabilla:

```
>> M = [2 0 0 ; 0 3 1 ; 0 1 3]
```

```
M =
```

```
      2      0      0
      0      3      1
      0      1      3
```

```
>> [V,D] = eigs(M)
```

```
V =
```

```
      0      0      1.0000
  0.7071  -0.7071      0
  0.7071   0.7071      0
```

```
D =
```

```
      4      0      0
      0      2      0
      0      0      2
```

```
>>
```

c4s15es4

Esimerkki 227 Tutkitaan matriisin

$$\mathbf{Q} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$$

Ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö saa muodon

$$\begin{aligned} \det(\mathbf{Q} - \lambda \mathbf{I}) = 0 &\Leftrightarrow \begin{vmatrix} 2-\lambda & 1 & 0 \\ 0 & 1-\lambda & -1 \\ 0 & 2 & 4-\lambda \end{vmatrix} = 0 \\ &\Leftrightarrow (2-\lambda) \begin{vmatrix} 1-\lambda & -1 \\ 2 & 4-\lambda \end{vmatrix} - 1 \cdot \begin{vmatrix} 0 & 1-\lambda \\ 0 & 2 \end{vmatrix} = 0 \\ &\Leftrightarrow (2-\lambda)((1-\lambda)(4-\lambda) + 2) = 0 \\ &\Leftrightarrow (\lambda-2)^2(\lambda-3) = 0 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ algebrallinen kertaluku on 2, ja ominaisarvon $\lambda_2 = 3$ geometrinen kertaluku on 1.

(2) Seuraavaksi määritämme ominaisarvoon $\lambda_1 = 2$ liittyvät ominaisvektorit.

$$\begin{aligned}
 & \mathbf{Q}\vec{x} = \lambda_1 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 + x_2 = 2x_1 \\ x_2 - x_3 = 2x_2 \\ 2x_2 + 4x_3 = 2x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} x_2 = 0 \\ -x_2 - x_3 = 0 \\ 2x_2 + 2x_3 = 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \end{array} \\
 \Leftrightarrow & \begin{cases} x_2 = 0 \\ -x_3 = 0 \\ 2x_2 + 2x_3 = 0 \end{cases} \begin{array}{l} \rightarrow x_1 = \text{vapaa} \\ \rightarrow x_2 = x_3 = 0 \end{array} \\
 \Leftrightarrow & \vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}
 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ geometrinen kertaluku on 1.

(3) Seuraavaksi määritämme ominaisarvoon $\lambda_2 = 3$ liittyvät ominaisvektorit.

$$\begin{aligned}
 & \mathbf{Q}\vec{x} = \lambda_2 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 3 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 + x_2 = 3x_1 \\ x_2 - x_3 = 3x_2 \\ 2x_2 + 4x_3 = 3x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} -x_1 + x_2 = 0 \\ -2x_2 - x_3 = 0 \\ 2x_2 + x_3 = 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \end{array} \\
 \Leftrightarrow & \begin{cases} -x_1 + x_2 = 0 \\ -2x_2 - x_3 = 0 \\ 0 = 0 \end{cases} \begin{array}{l} \rightarrow x_1 = x_2 \\ \rightarrow x_3 = -2x_2 \end{array}
 \end{aligned}$$

Ominaisvektoriksi voidaan valita

$$\vec{x}_2 = \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix}$$

Koska ensimmäisen ominaisarvon algebrallinen kertaluku on 2 ja geometrinen kertaluku on 1, on matriisi defektiivinen.

Sama matlabilla:

```
>> Q = [2 1 0 ; 0 1 -1 ; 0 2 4]
```

```
Q =
```

```
    2    1    0
    0    1   -1
    0    2    4
```

```
>> [V,D] = eigs(Q)
```

```
V =
```

```
    0.4082    1.0000    1.0000
    0.4082   -0.0000         0
   -0.8165    0.0000         0
```

```
D =
```

```
    3    0    0
    0    2    0
    0    0    2
```

```
>>
```

Defektiivisyys paljastuu nyt siitä, että tuloksena saadut ominisvektorit eivät muodosta vapaata systeemiä, sillä

```
>> det(V'*V)
```

```
ans =
```

```
    4.6259e-017
```

```
>>
```

Ominaisarvoihin ja ominaisvektoreihin palaamme takaisin, mutta ensin toteamme joi-takin uusia asioita.

9.2.2 Geometrinen tulkinta helpossa tapauksessa

Edellä on annettu vain tekninen määritelmä, joka saattaa joskus johtaa kompleksisiin ominaisarvoihin ja ominaisvektoreihin. Jotta asiasta saisi nyt konkreettisen hyödyllisen mielikuvan, oletamme tässä alakappaleessa, että $n \times n$ -matriisilla $\mathbf{A}$ on n erisuurta reaalista ominaisarvoa $\lambda_1, \lambda_2, \dots, \lambda_n$. Tämä yksinkertaistaa kovasti tarkastelua ja antaa mahdollisuuden nähdä, mitä merkitystä ominaisarvoilla on.

c4s15theigbase

Lause 228 Olkoon $\mathbf{A}$ $n \times n$ -matriisi, jonka ominaisarvot ovat reaaliset. Matriisin $\mathbf{A}$ ominaisvektorit muodostavat $\mathbb{R}^n$:n kannan, jos ja vain jos $\mathbf{A}$ ei ole defektiivinen.

Todistus. $\Rightarrow$ Jos $\mathbf{A}$ on defektiivinen, niin lineaarisesti riippumattomien ominaisvektorien lukumäärä on pienempi kuin n , joten ominaisvektorit eivät voi virittää $\mathbb{R}^n$:ää.

$\Leftarrow$ Oletetaan seuraavaksi, että $\mathbf{A}$ ei ole defektiivinen. Luetellaan matriisin ominaisarvot suuruusjärjestyksessä $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ siten, että jokainen ominaisarvo toistuu luettelossa kertalukunsa mukaisesti. Olkoot $\{\vec{x}_1, \dots, \vec{x}_n\}$, vastaavat ominaisvektorit. Koska $\mathbf{A}$ ei ole defektiivinen, voidaan samaan ominaisarvoon liittyvät ominaisvektorit valita vapaiksi.

Ominaisarvoon λ_1 liittyvä ominaisvektori $\{\vec{x}_1\}$ muodostaa vapaan systeemin. Oletamme seuraavaksi, että ominaisarvoihin $(\lambda_1, \lambda_2, \dots, \lambda_{p-1})$ liittyvät ominaisvektorit $\{\vec{x}_1, \dots, \vec{x}_{p-1}\}$ muodostavat vapaan systeemin. Osoitamme seuraavaksi, että ominaisarvoihin $(\lambda_1, \lambda_2, \dots, \lambda_p)$ liittyvät ominaisvektorit $\{\vec{x}_1, \dots, \vec{x}_{p-1}, \vec{x}_p\}$ muodostavat vapaan systeemin. oletamme sitä varten, että

$$\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_p \vec{x}_p = \vec{0}. \quad (9.20) \quad \text{eigbase0}$$

Oletamme, että k viimeistä ominaisvektoria liittyvät samaan ominaisarvoon ($\lambda_{p-k+1} = \lambda_{p-k+2} = \dots = \lambda_p$). Nyt

$$\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_{p-k} \vec{x}_{p-k} = -\mu_{p-k+1} \vec{x}_{p-k+1} - \dots - \mu_p \vec{x}_p = \vec{a}. \quad (9.21) \quad \text{eigbase1}$$

missä vektori $\vec{a}$ on ominaisarvoon λ_p liittyvä ominaisvektori. Kertomalla yhtälö (9.21) puolittain vasemmalta $\mathbf{A}$:lla saadaan

$$\mu_1 \lambda_1 \vec{x}_1 + \mu_2 \lambda_2 \vec{x}_2 + \dots + \mu_{p-k} \lambda_{p-k} \vec{x}_{p-k} = \lambda_p \vec{a}, \quad (9.22) \quad \text{eigbase2}$$

Yhdistämällä (9.21) ja (9.22) saadaan

$$\mu_1 \left(1 - \frac{\lambda_1}{\lambda_p}\right) \vec{x}_1 + \mu_2 \left(1 - \frac{\lambda_2}{\lambda_p}\right) \vec{x}_2 + \dots + \mu_{p-k} \left(1 - \frac{\lambda_{p-k}}{\lambda_p}\right) \vec{x}_{p-k} = \vec{0}.$$

Tämä on mahdollista vain, jos $\mu_1 = \mu_2 = \dots = \mu_{p-k} = 0$. Siis $\vec{a} = \vec{0}$, mikä edelleen on mahdollista vain, jos $\mu_{p-k+1} = \mu_{p-k+2} = \dots = \mu_p = 0$. Olemme siis osoittaneet, että

$$(\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_p \vec{x}_p = \vec{0}) \Rightarrow (\mu_1 = \mu_2 = \dots = \mu_p = 0).$$

Induktioaskel on valmis, ja lauseen todistus saadaan induktiolla p :n suhteen. $\square$

Lause 229 Jos $n \times n$ -matriisilla $\mathbf{A}$ on n erisuurta reaalista ominaisarvoa $\lambda_1, \lambda_2, \dots, \lambda_n$, niin vastaavista ominaisvektoreista muodostettu jono $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ on vapaa ja siis $\mathbb{R}^n$:n kanta.

Todistus. Suora seuraus edellisestä. □

Tarkastellaan kuvausta $f: \mathbb{R}^n \rightarrow \mathbb{R}^n, \vec{u} \mapsto \mathbf{A}\vec{u}$. Kun käytämme $\mathbb{R}^n$:n kantana ominaisvektoreista muodostettua kantaa, niin jokainen vektori voidaan lausua kannassa

$$\vec{u} = u_1 \vec{x}_1 + u_2 \vec{x}_2 + \dots + u_n \vec{x}_n$$

ja vektorin kuva on vastaavasti

$$\begin{aligned} f(\vec{u}) &= \mathbf{A}\vec{u} = \mathbf{A}(u_1 \vec{x}_1 + u_2 \vec{x}_2 + \dots + u_n \vec{x}_n) \\ &= u_1 \lambda_1 \vec{x}_1 + u_2 \lambda_2 \vec{x}_2 + \dots + u_n \lambda_n \vec{x}_n \end{aligned}$$

joten kuvauksen matriisi on

$$\mathbf{M}_f = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Geometrinen tulkinta on nyt siis se, että kun vektori lausutaan kantavektoreiden lineaarikombinaationa, se jaetaan kantavektoreiden suuntaisiin komponentteihin. Kerrottaessa matriisilla $\mathbf{A}$ jokainen komponentti tulee kerrottua vastaavalla ominaisarvolla. Jos $|\lambda_j| > 1$, niin komponentti venyy, jos $|\lambda_j| < 1$, niin komponentti kutistuu.

9.2.3 Ominaisarvohajoitelma

Merkitsemme $\mathbf{X}$:llä matriisia, jonka sarakkeina ovat matriisin $\mathbf{A}$ ominaisvektorit ja $\mathbf{D}$:llä disgonaalimatriisia, jonka diagonaalilla ovat matriisin $\mathbf{A}$ ominaisarvot. Jos nyt matriisin $\mathbf{A}$ ominaisvektorit muodostavat ortonormitetun systeemin, niin

$$\mathbf{A} = \mathbf{X} \mathbf{D} \mathbf{X}^T. \tag{9.23} \quad \text{eigsys haj}$$

Kaavan (9.23) ^{eigsvshaj} perusteluksi riittää osoittaa, että $\mathbf{XDX}^T \vec{x}_j = \lambda_j \vec{x}_j = \mathbf{A} \vec{x}_j$ kaikilla $j = 1, 2, \dots, n$. Tämä taas on suora lasku, sillä

$$\begin{aligned} \mathbf{X}^T \vec{x}_j &= \vec{e}_j \\ \Rightarrow \mathbf{DX}^T \vec{x}_j &= \mathbf{D} \vec{e}_j = \lambda_j \vec{e}_j \\ \Rightarrow \mathbf{XDX}^T \vec{x}_j &= \lambda_j \mathbf{X} \vec{e}_j = \lambda_j \vec{x}_j \end{aligned}$$

Ominaisarvohajoittelman avulla on helppo perustella joitakin tuloksia.

Lause 230 Jos neliömatriiseilla $\mathbf{A}$ ja $\mathbf{B}$ ovat samat ominaisvektorit ja ominaisvektorit muodostavat ortonormitetun jonon, niin

$$\mathbf{AB} = \mathbf{BA}$$

eli matriisit kommutoivat.

Todistus. Merkitään $\mathbf{A}$:n ominaisarvoista muodostettua diagonaalimatriisia $\mathbf{D}_A$:lla ja vastaavasti $\mathbf{B}$:n ominaisarvoista muodostettua diagonaalimatriisia $\mathbf{D}_B$:lla. Diagonaalimatriisit kommutoivat, joten

$$\begin{aligned} \mathbf{AB} &= \mathbf{XD}_A \mathbf{X}^T \mathbf{XD}_B \mathbf{X}^T = \mathbf{XD}_A \mathbf{D}_B \mathbf{X}^T = \mathbf{XD}_B \mathbf{D}_A \mathbf{X}^T \\ &= \mathbf{XD}_B \mathbf{X}^T \mathbf{XD}_A \mathbf{X}^T = \mathbf{BA} \end{aligned}$$

□

Matriisin potenssit on nyt helppo laskea käyttäen hajoitelmaa (9.23): ^{eigsvshaj}

$$\begin{aligned} \mathbf{A}^2 &= \mathbf{AA} = \mathbf{XDX}^T \mathbf{XDX}^T \\ &= \mathbf{XD}^2 \mathbf{X}^T = \mathbf{X} \begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix} \mathbf{X}^T \\ \mathbf{A}^n &= \mathbf{XD}^n \mathbf{X}^T = \mathbf{X} \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^n \end{pmatrix} \mathbf{X}^T \end{aligned}$$

Lause 231 (Cayley-Hamilton) Olkoon neliömatriisilla $\mathbf{A}$ ominaisarvohajoitelma $\mathbf{A} = \mathbf{XDX}^T$, missä ominaisvektoreista muodostettu matriisi on ortogonaalinen. Matriisi $\mathbf{A}$ toteuttaa oman karakteristisen yhtälönsä.