

Kurssimateriaalia K3/P3-kursille syksyllä 2005.
27. lokakuuta 2005
Heikki Apiola

Kirjallisuutta, www-sivuja

[KRE] Kreyszig . *Advanced Engineering Mathematics*, 8th ed., Wiley ,1999.

[Lay] Lay . *Linear Algebra*, 3rd ed., Addison Wesley ,2003.

[TE] Timo Eirola . *Lineaarialgebra*, L3-kurssimoniste

KP3-kurssien www-sivuja

[wwwhome] <http://math.tkk.fi/teaching/k3/index.html.fi> (Kurssin pääsivu)

<http://math.tkk.fi/teaching/k3/05/L/ominaisarvot.pdf> (Tämä pruju)

[wwwmaple] <http://math.tkk.fi/teaching/k3/05/L/ominaisarvot.html> (Maple-ws:n html-versio: Ominaisarvojen havainnollistusta ja laskentaa, "pyörivät siniset ja punaiset nuolet")

1 Ominaisarvot ja -vektorit

Tarkasteltavana oleva matriisi on koko ajan **neliömatriisi**, $A(n \times n)$.

Johdanto

Ajatellaan neliömatriisia A ennen kaikkea sen määräämän lineaarikuvauksen kannalta.

Tähän asti staattisia ongelmia: lineaarisen yhtälöryhmän ratkaisu.

Nyt käsittelemme dynaamisia tehtäviä. Tyypillisesti tarkastelemme iteraatiota

$$\text{Alkupiste } \mathbf{x}_0, \quad \mathbf{x}_{k+1} = A\mathbf{x}_k, k = 0, 1, 2, \dots$$

Tällöin joudumme laskemaan matriisin potensseja (joita neliömatriisille voidaan laskea). Sovelluksina näemme mm. diskreettejä dynaamisia systeemejä (differenssiyhtälösystemeitä) ja differentiaaliyhtälösystemeitä.

Ominaisarvotehtäviä tulee luonnontieteellisissä sovelluksissa vastaan hyvin monessa paikassa. Kyseessä on myös insinöörimatematiikan kaikkein tärkeimpään ydinalueeseen kuuluva aihepiiri. Lisäksi: kaunis kappale lineaarialgebraa.

1.1 Määritelmä ja laskutekniikkaa

(A on $(n \times n)$ neliömatriisi.)

Tarkastellaan yhtälöä

$$A\mathbf{x} = \lambda\mathbf{x}, \lambda \in \mathbb{C} \tag{1.1}$$

Yhtälöllä on aina triviaaliratkaisu $\mathbf{x}=\mathbf{0}$, olipa λ mikä tahansa kompleksiluku.

Tehtävä:

(a) Määritä luku λ siten, että yhtälöllä (1.2) on ei-triviaaleja ratkaisuvektoreita \mathbf{x} .

(b) Määritä sitten kutakin ratkaisua λ kohti ao. ratkaisuvektorit \mathbf{x} .

Määritelmä 1.1 *Lukua λ sanotaan matriisin A ominaisarvoksi ja vektoria $\mathbf{x} \neq \mathbf{0}$ vastaavaksi ominaisvektoriksi, jos*

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Englanninkielä: "Eigenvalue, eigenvector"

Havainnollistusta

Geometrisesti ominaisvektori tarkoittaa suuntaa, joka säilyy samana (tai vastakkaisena) sovellettaessa lineaarikuvausta A . Ominaisarvo kuvaa venytys-/kutistussuhdetta, negatiivisessa tapauksessa lisäksi suunnan vaihtoa.

Viitteessä [wwmaple] on useita esimerkkejä lineaarikuvauksista \mathbb{R}^2 : ssa. Yksikköympyrän kehää kiertävä sininen lähtönuoli kuvataan matriisilla kertomalla punaiseksi kuvanuoleksi. Animaatiota askel kerrallaan ajettaessa näkyvät kohdat, joissa sininen ja punainen nuoli menevät päällekkäin. Siinä on ominaisvektori, ominaisarvo antaa venytys-/kutistuskertoimen, jolla sininen saadaan punaiseksi. (Animaatioita voi ajaa askel kerrallaan Maplessa, lataamalla `ominaisarvot.mws` Mapleen. Tällöin on mahdollisuus tehdä omia kokeiluja muillakin matriiseilla yms. Ilman Maplea voi katsella valmista esitystä pelkästään selaimella `ominaisarvot.html`.)

Tämä kuvailu pätee reaalisiin ominaisarvoihin (ja -vektoreihin) nähden. Kompleksisessa tapauksessa tilanne voi olla "epäintuitiivisempi", kannattaa muistaa, että esim. i :llä kertominen merkitsee kiertoa tasossa kulman $\pi/2$ verran, joten skalaarilla kertominen ei tarkoita (reaalisella) suoralla pysyttelemistä.

Tehtävän muokkaus ratkaistavaan muotoon

$$A\mathbf{x} = \lambda\mathbf{x}, \lambda \in \mathbb{C} \iff (A - \lambda I)\mathbf{x} = \mathbf{0} \quad (1.2)$$

Kyseessä on siis homogeeniyhtälö, jonka kerroinmatriisi sisältää parametrin λ . Se pitää valita siten, että HY:llä on ei-triv. ratkaisuja, ts. Gaussin eliminaation on tuotettava ainakin 1 nollarivi, joten on oltava $\det(A - \lambda I) = 0$.

Esimerkki 1.1 *Lasketaan aluksi wwmaple-työarokin ensimmäinen esimerkkitapaus:*

$$A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}, D(\lambda) = \begin{vmatrix} -\lambda + 3 & -2 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 3\lambda + 2.$$

Tästä saadaan nollakohdat, eli ominaisarvot: $\lambda_1 = 1, \lambda_2 = 2$.

Ominaisvektorit:

$$\lambda_1 = 1$$

$A - I = \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}$, mistä saadaan $x_1 = x_2$, joten ominaisvektori on $[1, 1]^T$. (Tämä vektori on ominaisavaruuden kanta, toki voidaan kertoa millä tahansa skalaarilla $c \neq 0$.)

$$\lambda_2 = 2$$

$A - 2I = \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix}$. Valitaan $x_2 = 1$, jolloin $x_1 = 2$. Saadaan siis $x_2 = [2, 1]^T$.

Muista verrata laskun tuloksia ww-työarkkiin, tai vielä mieluummin piirrä paperille.

Huomioita:

Jos A on 2×2 -matriisi, voidaan ominaisvektoreita laskettaessa aina jättää toinen yhtälö pois, sillä matriisin rivien on oltava LRV.

Kyseessä on nolla-avaruuden määräämistehtävä. Tässä tapauksessa oli **kaksi erisuurta ominaisarvoa**. Tällöin 2×2 -matriisin tilanteessa saadaan aina kumpaankin ominaisarvoon liittyen yksiulotteinen ominaisavaruus.

Esimerkki 1.2 $A = \begin{bmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{bmatrix}$

Ominaisarvot:

Muodostetaan $D(\lambda) = \det(A - \lambda I) = -(\lambda + 2)(\lambda - 6)^2$. Tämä saadaan kehittämällä 1. sarakkeen suhteen. Helpon tekijöihin jaon toivossa kannattaa säilyttää laskut tekijämuodossa, kiirehtimättä kertomaan auki, ennenkuin on pakko. Tässä tehtävässä pakkoa ei tule, vaan saadaan hyvin lyhyellä laskulla tuo tulos. (Yksityiskohdat kannattaa lukijan laskea itse.)

Koska tekijä $(\lambda + 2)$ esiintyy potenssissa 1, sanotaan, että ominaisarvon $\lambda_1 = -2$ **algebraalinen kertaluku** $M_{\lambda_1} = 1$. Vastaavasti $(\lambda - 6)$ esiintyy potenssissa 2, joten ominaisarvon λ_2 **algebraalisen kertaluvun** M_{λ_2} sanotaan olevan 2.

Lasketaan ominaisvektorit

Sijoitetaan karakteristiseen matriisiin $K_\lambda = A - \lambda I$ λ :n paikalle vuorollaan kukin ominaisarvo.

Ominaisarvoon $\lambda_1 = -2$ liittyvät ominaisvektorit

$$K_{\lambda_1} = \begin{bmatrix} 4 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 4 \end{bmatrix}$$

Tässä ei rivioperaatioita tarvita. Alin rivi voidaan jättää pois (identtinen ylimmän kanssa). Voidaan haluttaessa jakaa 1. rivi 4:llä ja toinen 8:lla, muttei sekään ole tarpeen.

2. rivi $\implies x_3$ vapaa ja $x_2 = 0$.

1. rivi $\implies x_1 = -x_3$.

Jos valitaan: $x_3 = 1$, saadaan $\mathbf{v}_1 = [-1, 0, 1]^T$.

Saadaan 1-ulotteinen ominaisavaruus. Ominaisavaruuden λ dimensiolle käytetään merkintää m_λ ja nimitystä **geometrinen kertaluku**. Tässä tapauksessa algebraalinen ja geometrinen kertaluku yhtyvät $m_{\lambda_1} = M_{\lambda_1} = 1$.

Yleisesti voitaisiin muodostaa ref (tai rref).

$$\text{ref}(K_{\lambda_1}) = \begin{bmatrix} 4 & 0 & 4 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{Tästä nähdään heti, että nolla-avaruuden dimensio} = 1. \text{ (n-tukisarakkeiden lkm} = \text{nollarvien lkm, kun kerran on neliömatriisi.) Tässä}$$

tapauksessa laskut eivät edellisestä yksinkertaistu, koska matriisi oli jo heti yhtä rivioperaatiota ja normeerausta vaille jopa rref-muodossa.

Ominaisarvoon $\lambda_2 = 6$ liittyvät ominaisvektorit

$$K_{\lambda_2} = \begin{bmatrix} -4 & 0 & 4 \\ 0 & 0 & 0 \\ 4 & 0 & -4 \end{bmatrix}$$

Nyt jää vain yksi yhtälö, jossa voidaan kaksi muuttujaa valita vapaasti, vaikkapa x_3 ja x_2 , ja ratkaisemalla $x_1 = x_3$.

Saadaan kaksi LRT ominaisvektoria valitsemalla ensin $x_3 = 1, x_2 = 0$ ja sitten päinvastoin $x_3 = 0, x_2 = 1$

$\mathbf{u}_1 = [1, 0, 1]^T$, $\mathbf{u}_2 = [0, 1, 0]^T$. Kertaluvut ovat taas samat: $m_{\lambda_2} = M_{\lambda_2} = 2$.

Siis ominaisavaruudet ovat: $E_{-2} = \text{sp}\{[-1, 0, 1]\}$ ja $E_6 = \text{sp}\{[1, 0, 1], [0, 1, 0]\}$

(Yhtälösystemien yhteydessä opitun systematiikan mukaan toimien saataisiin lisäämällä 1. rivi kolmanteen ja jakamalla vielä -4 : llä, matriisi, jonka ensimmäinen rivi olisi $1, 0, -1$ ja muut nollarivejä. Siis $r = 1$, $n - r = 3 - 1 = 2$, joten nolla-avaruuden dimensio $= 2$.)

Jäljempänä esitettävän lauseen mukaan eri ominaisarvoihin liittyvät ominaisvektorit ovat LRT, joten voimme sen turvin suoraan päätellä, että laskemamme 3 ominaisvektoria ovat LRT ja siis \mathbb{R}^3 :n kanta. Ilman tuota lausetta täytyisi turvautua ref:iin.

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Kaikki sarakkeet ovat tukisarakkeita, joten sarakkeiksi latomamme ominaisvektorit ovat todellakin LRT.

Huomautus 1.1 Varo harhaluulemasta, että lineaarinen riippumattomuus voitaisiin todistaa osissa niin, että osoitettaisiin vaikka kaksi vektoria LRT:ksi ja kolmas kummastakin LRT:ksi. Tokihan kolmas voi sijaita kahden virittämässä tasossa (ja muodostaa siis ensinmainittujen kanssa LRV joukon) ja olla kummankin kanssa eri suuntainen.

Edellä mainitussa (jäljempänä esitettävässä) lauseessa esiintyvässä ”ominaisvektoriniputuksessa” on kyse siitä, että kaikki mahdolliset valinnat eri ominaisarvuuksista antavat LRT vektorijoukon. Mutta, kuten sanottu, palataan.

1.2 Karakteristinen polynomi, ominaisarvon kertaluvut

|

Jos determinantti $D(\lambda) = \det(A - \lambda I)$ kehitetään vaikkapa 1. sarakkeen mukaan, nähdään, että se on polynomi, joka on muotoa:

$$D(\lambda) = (-1)^n \lambda^n + c_{n-1} \lambda^{n-1} \dots + c_0.$$

Polynomi, jota sanotaan *karakteristiseksi polynomiksi*, on siis aina astetta n . Kertoimet c_0 ja c_{n-1} on helppo määrittää. Edellinen saadaan laskemalla

$$c_0 = p(0) = \det(A - 0I) = \det(A).$$

Jälkimmäinen saadaan kertomalla diagonaalitermit $(\lambda - a_{i,i})$ ja ottamalla mukaan ... Näin saadaan $c_{n-1} = -(a_{11} + a_{22} + \dots + a_{nn})$.

Algebran peruslause sanoo, että **jokaisella polynomilla, joka ei ole vakio, on kompleksitasossa nollakohta**. Lauseen todistus vaatii vähän enemmän kompleksianalyysin työkaluja, kuin olemme tällä kurssilla käsitelleet. Sitä voidaan pitää kurssin kannalta ”syvällisenä” ja toki muutenkin. Hieno tulos, jonka elegantti toditus pohjautuu ns. *Liouville’n* lauseeseen: Jokainen koko kompleksitasossa analyyttinen (kokonainen) funktio, joka on rajoitettu, on vakio. (Sovella tätä $1/p(z)$ -funktioon, jossa $p(z)$ on polynomi, jolla ei ole kompleksitasossa yhtään nollakohtaa Ristiriita on käden ulottuvilla.)

Lisäksi on käytössämme kaikkea muuta kuin syvällinen asia: **Jos polynomilla $p(z)$ on nollakohta z_0 , niin se on jaollinen $(z - z_0)$:lla.** Tämähän seuraa heti kirjoittamalla $p(z) - p(z_0)$ ja toteamalla, että koska vakiotermit kumoutuvat, saadaan jokaisesta termistä $z - z_0$ tekijäksi.

Saadaan siten:

Lause 1.1 (KRE8 Thm. 1 s. 373) *Neliömatriisilla A ($n \times n$) on ainakin yksi ja korkeintaan n erillistä ominaisarvoa. Ne voivat olla kompleksisia, vaikka matriisi olisi reaalinen.*

Karakteristinen polynomi voidaan kirjoittaa muotoon:

$$D(\lambda) = (\lambda - \lambda_1)^{k_1} \dots (\lambda - \lambda_n)^{k_n}$$

Ominaisarvon λ_k **algebraallinen kertaluku** M_λ tarkoittaa kyseisen polynomin nollakohdan kertalukua. Siten M_{λ_k} ilmaisee siis potenssin p_k , jossa tekijä $(\lambda - \lambda_k)$ esiintyy.

Voidaan sanoa, että $n \times n$ -matriisilla on n kappaletta ominaisarvoja, kun kukin otetaan niin monta kertaa kuin sen algebraallinen kertaluku ilmaisee.

Määritelmä 1.2 Ominaisarvoon λ liittyvä **ominaisavaruus** E_λ koostuu kaikista λ :aan liittyvistä ominaisvektoreista ja nollavektorista. E_λ on vektorialiavaruus, koska se on $(A - \lambda I)$:n nolla-avaruus. Ominaisarvon **geometrinen kertaluku** m_λ on ominaisavaruuden E_λ dimensio. Siis $m_\lambda = \dim(E_\lambda)$.

Ominaisarvojen ja -vektorien laskeminen, yhteenveto

Kootaan vielä kertaukseksi toimintaohje:

1. Muodostetaan karakteristinen polynomi

$$D(\lambda) = \det(A - \lambda I)$$

ja ratkaistaan sen nollakohdat, näin saadaan ominaisarvot.

2. Ratkaistaan homogeneeniyhtälö

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

ts. määrätään nolla-avaruus $N(A - \lambda I)$ kullakin ominaisarvolla λ .

Ratkaisussa on ainakin yksi vapaa parametri, sanokaamme t , joka voidaan valita mielivaltaisesti (esim. $t=1$). Jos vapaita parametrejä on useita, täytyy pitää huoli siitä, että saadaan lineaarisesti riippumattomat ratkaisuvektorit. Varma valinta esim. tapauksessa ”3 vapaata parametria” on:

Yksi ominaisvektori : parametrit: 1, 0, 0
 Toinen ominaisvektori : parametrit: 0, 1, 0
 Kolmas ominaisvektori : parametrit: 0, 0, 1.

Jos kysytään vain ominaisavaruuden E_λ dimensiota, eli ominaisarvon λ geometrista kertalukua, riittää määrittää matriisin $A - \lambda I$ rangi r , ts. tukisarakkeiden lukumäärä. Kysytty dimensio on silloin $n - r$.

Kertaukseksi

1. Ominaisarvot ovat 1-käs, ominaisvektorit eivät.
2. Kuhunkin ominaisarvoon liittyy ominaisavaruus, päämääränä on löytää jokin kanta ominaisavaruudelle.

Käytännön (numeerisista) laskentamenetelmistä

Polynomiyhtälön ratkaisussa on useimmiten turvaututtava numeerisiin menetelmiin. Matlabissa on tähän tarkoitukseen funktio `roots`. Oikeissa numeerisissa menetelmissä lasketaan ensin ominaisvektorit (tai esim. muutama dominoiva ominaisvektori – sellaiset, jotka vastaavat muutamaa suurinta ominaisarvoa). Ominaisarvot lasketaan vasta sitten.

Itse asiassa Matlabin `roots`-komennon käyttö ominaisarvotehtävissä on lievää ”itsepetosta”. Se perustuu algoritmiin, joka laskee annettuun polynomiin liittyvän matriisin, ns. ”companion matrix”:n ominaisvektorit ja sitten ominaisarvot. Sivutuotteena, eräänlaisena kuriositeettina, siitä saadaan helposti kohtuullinen polynomiyhtälön ratkaisija.

Ominaisarvojen laskenta tätä kautta on siten hyödyllistä pelkästään perusasioiden opetteluun. Oikeaan laskentaan käytetään komentoa `eig`.

Matlabin "isä" *Cleve Moler*, joka on eräs nykyaikaisen tieteellisen laskennan huomattavia vaikuttajapersoonallisuuksia, pitää polynomiyhtälön ratkaisemista nykyaikana hieman numeerisen analyysin sivuraitteille kuuluvaksi. Siitä syystä hän ei luultavasti ole kiiruhtanut ottamaan Matlabiin parasta alan algoritmia käyttöön. Myös Matlabin historia on hyvin vahvasti juuri lineaarialgebrassa. Sillä alueella on erityisesti panostettu parhaisiin algoritmeihin, toki sitten myös mm. differentiaaliyhtälöissä.

1.3 Illuusiot romahtavat – esimerkkejä, defektiivisyys

Toinen illuusiomme voisi olla se, että jokaisella reaalisella matriisilla olisi reaalinen ominaisarvo. Koska kyseessä on polynomiyhtälön nollakohta, tuo on varsin epärealistinen toive. Myös kiertomatriisiin ajattelu antaa saman johtopäätöksen. Lasketaan joka tapauksessa esimerkki.

Vakavampi illuusio voisi olla se, että algebrallinen ja geometrinen kertaluku olisivat aina samat. Sitähän tukevat tähän mennessä esiintyneet esimerkit. Romutamme tämän illuusion heti esimerkillä ja kerromme, mitä tiedetään yleisesti kertalukujen suhtautumisesta toisiinsa.

Esimerkki 1.3 Olkoon $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Lineaarikuvausena tarkasteltuna $A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ 0 \end{bmatrix}$.

Sis kaikki x_1 -akselilla makaavat vektorit kuvautuvat 0-vektoriksi, joten ne ovat (**0**:aa lukuunottamatta) ominaisarvoon $\lambda = 0$ liittyviä ominaisvektoreita. Muut kuin x_1 - akselin suuntaiset vektorit kuvautuvat myös x_1 -akselille, joten ne eivät ole ominaisvektoreita. Näyttää siis siltä, että kaikenkaikkiaan matriisilla on vain yksi ominaisvektori (tarkemmin sanottuna yksi yksiuulotteinen ominaisavaruus).

No mutta lasketaan toki myös.

$D(\lambda) = \lambda^2$, joten $\lambda = 0$ on (algebrallisesti) kaksinkertainen ominaisarvo, eli $M_0 = 2$.

Ominaisarvoon $\lambda = 0$ liittyvä ominaisavaruus on sama kuin A :n nolla-avaruus. Koska A :lla on yksi tukisarake, tämän dimensio on $2 - 1 = 1$. Ominaisvektorit määräytyvät yhtälöstä

$0x_1 + x_2 = 0$, eli x_1 mielivaltainen (voidaan ottaa $x_1 = 1$), $x_2 = 0$.

Siten ominaisavaruuden virittää vektori $[1, 0]^T$.

Ominaisarvon 0 geometrinen kertaluku $m_0 = 1$. Sis $m_0 < M_0$. Sanotaan, että matriisi on defektiivinen. Ominaisvektoreita ei ole "tarpeeksi", jotta niistä voitaisiin muodostaa \mathbb{R}^2 :n kanta.

Missään esimerkissämme ei ole esiintynyt epäyhtälöä toisinpäin, eli geometrinen on ollut aina korkeintaan algebrallinen. Tämä onkin yleinen ominaisuus:

Lause 1.2 Yleisesti pätee: $m_\lambda \leq M_\lambda$. Aito $<$ on mahdollinen.

Edellinen on jossain määrin syvällinen asia, perustelu on pakko sivuuttaa. Jälkimmäisen perustelee edellinen esimerkki.

Esimerkki 1.4 KRE esim. 4 s. 375. Reaalinen matriisi, jolla on kompleksiset ominaisarvot (ja -vektorit)

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Kyseessä on kiertomatriisi $\begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$, kun kiertokulmana on $\alpha = \pi/2$.

$$\text{Nyt } \det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1, \text{ joten ominaisarvot ovat } \lambda = \pm i.$$

Lasketaan ominaisvektorit, ensin ominaisarvolle $\lambda = i$

Ratkaistavaksi tulee yhtälö $-ix_1 + x_2 = 0$. Jos valitaan $x_2 = 1$, on $x_1 = 1/i = -i$. Saadaan ominaisvektoriksi $[-i, 1]^T$.

Huomaa, että kompleksisella skalaarilla kertominen on nyt sallittua. Näin saadaan vektoreita, jotka eivät ensi silmäyksellä näytä lineaarisesti riippuvilta. Niinpä, jos kerromme vaikka i :llä, saamme yhtä hyvän kantavektorin: $[1, i]^T$.

Ominaisarvoa $\lambda = -i$ vastaava ominaisvektori saadaan samalla tavalla. Itse asiassa sitä ei tarvitse laskea, koska se on edellä lasketun ominaisvektorin liittovektori. Tällä tarkoitamme vektoria, jonka koordinaatit ovat edellisen liittolukuja.

Todistetaan viimeksi todettu oikein lauseena.

Lause 1.3 *Reaalisen matriisin kompleksiset ominaisarvot ovat pareittain liittolukuja. Liittolukuja vastaavat ominaisvektorit ovat toistensa liittovektoreita.*

Ominaisarvoja koskeva johtopäätös voidaan tehdä kompleksiprujussa olevasta lauseesta: reaali kertomisen polynomien kompleksiset juuret ovat toistensa liittolukuja.

Itse asiassa tätäkään ei tarvita, kun huomataan, että edellisen perustana olevasta, liittolukujen laskutoimituksia koskevasta lauseesta saadaan myös välittömästi sääntö $A\bar{v} = \overline{Av}$.

$$\text{Niinpä, jos } Av = \lambda v, \text{ niin } A\bar{v} \underbrace{=} \overline{Av} = \overline{\lambda v} \underbrace{=} \lambda \bar{v} = \overline{\lambda} \bar{v}.$$

A on reaalinen $Av = \lambda v$

Johtopäätös seuraa siten suoraan ominaisarvon/-vektorin määritelmästä.

Tehtävä 1.1 *Laske yleisen kiertomatriisin $A = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$ ominaisarvot ja -vektorit.*

Millä kiertokulman arvoilla ovat reaaliset?

Yhteenvedo käsitteistä ja määritelmistä

- Ominaisvaruus E_λ on ominaisarvoon λ kuuluvien ominaisvektorien joukko täydennettynä 0-vektorilla.
- Spektri = ominaisarvojen joukko (kompleksitason osajoukko).
- Spektraalisäde: itseisarvoltaan suurimman ominaisarvon itseisarvo. Se antaa ympyrän säteen, jonka sisällä kaikki ominaisarvot (eli spektrin pisteet) ovat.
- Karakteristinen polynomi : $D(\lambda) = \det(A - \lambda I)$, tasan astetta n .
- Ominaisarvon λ_k **algebraallinen kertaluku** = p_k , kun karakteristinen polynomi esitetään tekijöihin jaetussa muodossa:

$$D(\lambda) = (-1)^n (\lambda - \lambda_1)^{p_1} \cdots (\lambda - \lambda_n)^{p_n}$$

Ominaisarvon λ_k algebraallinen kertaluku M_{λ_k} on siis potenssi p_k , jossa tekijä $(\lambda - \lambda_k)$ esiintyy.

- Geometrinen kertaluku** m_λ tarkoittaa vastaavan ominaisvaruuden dimensiota.
- Pätee: $m_\lambda \leq M_\lambda$
- Matriisi A on **defektiivinen**, jos jollain ominaisarvolla λ pätee aito pienemmyys: $m_\lambda < M_\lambda$. Tällöin matriisin ominaisvektoreita ”puuttuu”, niitä ei ole tarpeeksi, jotta niistä saataisiin \mathbb{R}^n :n (tai \mathbb{C}^n :n) kanta.

1.4 Sovellutusesimerkkejä

1.4.1 Joustavan kalvon venytys, lineaarikuvauksen pääakselit

Yksikkökierokko olkoon joustava kalvo, jota muotoilee ”siirtymämatriisi”

$$A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}.$$

Määritä pääsuunnat, eli vektorit, joille venymä on mahdollisimman suuri/pieni. Tai: Suunnat, joissa siirtymä säilyy yhdensuuntaisena paikkavektorin kanssa. Matemaattisesti ilmaistuna: Määritä matriisin määräämän lineaarikuvauksen *pääakselit*.

Ratkaisu

Lasketaan ominaisarvot ja -vektorit, lukija tarkistakoon, että saadaan $\lambda_1 = 8, \lambda_2 = 2$ ja että vastaavat ominaisvektorit ovat $\mathbf{u}_1 = [1, 1]^T$, $\mathbf{u}_2 = [-1, 1]^T$

Ominaisvektorien käyttökelpoisuus näkyy nyt tässä: Saamme **ongelmaamme (matriisiin A) liittyen optimaalisen kannan** siinä mielessä, että A:n määräämä lineaarikuvaus saa mahdollisimman yksinkertaisen esityksen.

Ominaisvektorit ovat varmasti LRT (ne ovat jopa ortogonaaliset), joten ne muodostavat \mathbb{R}^2 :n kannan.

Esitetään nyt mielivaltainen tason vektori \mathbf{u} ominaisvektorikannassa: $\mathbf{u} = \xi_1 \mathbf{u}_1 + \xi_2 \mathbf{u}_2$.

Tällöin

$$\mathbf{z} = A\mathbf{u} = \xi_1 A\mathbf{u}_1 + \xi_2 A\mathbf{u}_2 = \xi_1 \lambda_1 \mathbf{u}_1 + \xi_2 \lambda_2 \mathbf{u}_2 = 8\xi_1 \mathbf{u}_1 + 2\xi_2 \mathbf{u}_2.$$

Koordinaattien välinen kuvaus on ”ihana”: $[\xi_1, \xi_2] \mapsto [8\xi_1, 2\xi_2]$. Kumpikin kuvakoordinaatti riippuu vain omasta lähtökoordinaatistaan, ts. se saadaan diagonaalimatriisilla kertomalla. Jos kuvapisteen koordinaatteja (ominaisvektorikannassa) merkitään (η_1, η_2) , saadaan

$$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}.$$

Tässä tapauksessa sattuu vielä niin onnellisesti, että ominaisvektorit ovat toisiaan vastaan kohtisuorassa ¹, jolloin meillä on uusi luonteva suorakulmainen koordinaatisto, johon voimme astua ja esittää pisteiden koordinaatit normaalissa napakoordinaatistossa $\mathbf{u}_1 \mathbf{u}_2$ -akselien suhteen. Otetaanpa siis jokin yksikköympyrän piste: $(\cos \phi) \mathbf{u}_1 + (\sin \phi) \mathbf{u}_2$. Sen kuvapiste on siis $(8 \cos \phi) \mathbf{u}_1 + (2 \sin \phi) \mathbf{u}_2$.

Kuvapisteeet piirtävät siten ellipsin, jonka esitys pääakselikoordinaatistossa $\mathbf{u}_1 \mathbf{u}_2$ on

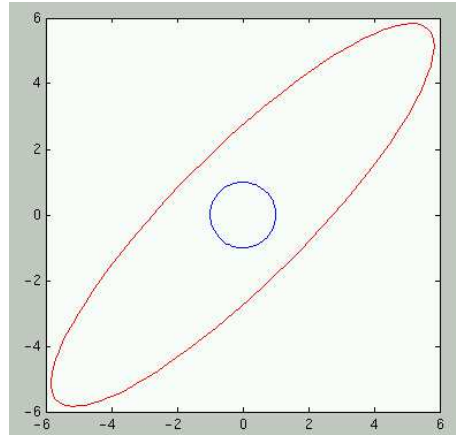
$$\begin{cases} \eta_1 = 8 \cos \phi \\ \eta_2 = 2 \sin \phi \end{cases},$$

Toisin sanoen

$$\left(\frac{\eta_1}{8}\right)^2 + \left(\frac{\eta_2}{2}\right)^2 = 1.$$

Esimerkkiin liittyvä **Matlab-ajo** on tiedostossa
<http://math.tkk.fi/teaching/k3/03/L/paakseli.m>

¹Jos osaisimme teoriaa enemmän, tietäisimme jo matriisin ulkonäöstä, että näin käy, koska matriisi on symmetrinen.



Kuva 1: Yksikköympyrä ja sen kuva

1.4.2 Stokastinen matriisi, Markovin prosessi

KRE8 s. 318 (“matrix multiplication”)

Oletetaan, että vuonna 2005 erään kaupungin maankäyttö jakaantuu näin:

I	asuntoalue	30%
II	kaupallinen käyttöalue	20%
III	teollisuusalue	50%

Oletetaan, että siirtymätodennäköisyydet 5-vuotisjakson aikana saadaan matriisista

$$P = \begin{array}{c|ccc|c} & \text{I:stä} & \text{II:sta} & \text{III:sta} & \\ \hline & 0.8 & 0.1 & 0 & \text{I:een} \\ \hline & 0.1 & 0.7 & 0.1 & \text{II:een} \\ \hline & 0.1 & 0.2 & 0.9 & \text{III:een} \\ \hline \end{array}$$

2

Tarkemmin sanottuna merkittäköön tyyppiä I, II ja III olevia maa-alueita 5-vuotisjakson lopussa

vektorilla $\mathbf{x}_n = [a_n, k_n, t_n]^T$. Seuraavan 5-vuotisjakson lopussa pätee
$$\begin{cases} a_{n+1} = 0.8 a_n + 0.1 k_n \\ k_{n+1} = 0.1 a_n + 0.7 k_n + 0.1 t_n \\ t_{n+1} = 0.1 a_n + 0.2 k_n + 0.9 t_n. \end{cases}$$

Sarakesummien on oltava ykkösiä, koska jokainen tyyppi muuttuu joksikin näistä kolmesta tyypistä, eikä miksikään muuksi. (Rivisummat eivät yleensä ole ykkösiä.) Matriisin lävistäjävaltaisuus kertoo sen, että kutakin tyyppiä olevasta alueesta suurin osa säilyy samaa tyyppiä olevana.

Kysymys siitä, mikä on tilanne m :n 5-vuotisjakson jälkeen, ratkeaa pelkästään matriisikertolaskulla.

$$\mathbf{x}_0 = \begin{bmatrix} 30 \\ 20 \\ 50 \end{bmatrix}, \quad \mathbf{x}_1 = P \mathbf{x}_0 = \begin{bmatrix} 26.0 \\ 22.0 \\ 52.0 \end{bmatrix}, \quad \mathbf{x}_2 = P \mathbf{x}_1 = \begin{bmatrix} 23.0 \\ 23.2 \\ 53.8 \end{bmatrix}, \quad \mathbf{x}_3 = P \mathbf{x}_2 = \begin{bmatrix} 20.72 \\ 23.92 \\ 55.36 \end{bmatrix}, \dots$$

² P on sama kuin KRE-kirjan A^T , kirjassa kerrotaan matriisilla oikealta rivivektoria, kun taas me kerromme vasemmalta sarakevektoria.

Matlabilla voitaisiin iteroida vaikka tähän tapaan.

```
>> P=[0.8 0.1 0;0.1 0.7 0.1;0.1 0.2 0.9]
>> x0=[30;20;50];
>> x0'
ans =
    30    20    50           % v. 2005
>> x=x0           % x olkoon muuttuja, johon iterointi kohdistuu.
>> x=P*x; x'     % Tilan säästämiseksi katsotaan vaakasuorassa.
ans =
    26    22    52           % v. 2010
>> x=P*x; x'
ans =
    23.0000    23.2000    53.8000 % v. 2015
>> x=P*x; x'
ans =
    20.7200    23.9200    55.3600 % v. 2020
```

Tämä oli siis pelkkää matriisikertolaskuharjoittelua.

Ominaisarvoteoria tulee mukaan, kun kysellään, mitä tapahtuu pitkällä aikavälillä. Voidaan tietysti harrastaa matriisikertolaskua raakaan voimaan turvautuen, mutta se ei anna eväitä analysoida tilannetta.

Lasketaanpa ominaisarvot ja -vektorit. Emme uppoudu tässä ominaisarvolaskujen yksityiskohtiin, niitä on jo harjoiteltu. Toisaalta on hyvä nähdä, miten laadukkaita numeerisia ohjelmia käytetään ominaisarvotehtävissä. Siksi näytetään tämä MATLAB:lla laskettuna.

```
>> P=[0.8 0.1 0;0.1 0.7 0.1;0.1 0.2 0.9]
P =

    0.8000    0.1000         0
    0.1000    0.7000    0.1000
    0.1000    0.2000    0.9000
>> [V,D]=eig(P)

V =
   -0.1826   -0.7071    0.4082   Ominaisvektorit ovat
   -0.3651    0.0000   -0.8165   sarakkeina.
   -0.9129    0.7071    0.4082
D =
    1.0000         0         0   Ominaisarvot ovat
         0    0.8000         0   diagonaalilla.
         0         0    0.6000
>> V*diag(1./V(1,:))   % Ominaisvektorien skaalaus.
ans =
    1.0000    1.0000    1.0000
    2.0000    0.0000   -2.0000
    5.0000   -1.0000    1.0000
```

MATLAB-komento $[V,D]=\text{eig}(P)$ paluttaa kaksi matriisia, V :n sarakkeina ovat ominaisvektorit normeerattuna yksikkövektoreiksi. Matriisi D on diagonaalimatriisi, jonka lävistäjällä ovat vastaavat ominaisarvot. Viimeinen matriisi on tehty jakamalla kukin ominaisvektori ensimmäisellä komponentillaan, jolloin päästään (tässä tehtävässä) mukaviin kokonaislukukomponentteihin. (Kommento

sisältää hiukan MATLAB-ajattelutavan sisäistäneen käyttäjän tarjoilemaa eleganssia, toki voi tehdä (salaa) kömpelömminkin.)

Nähdään, että kaikki ominaisarvot ovat positiivisia, suurin on 1 ja muut ovat pienempiä. Myöhemmin osoitetaan, että eri ominaisarvoihin liittyvät ominaisvektorit ovat LRT. Kun ei sitä vielä ole esitetty, niin ei muuta kuin "gaussataan".

```
>> rref(V)
ans =
     1     0     0
     0     1     0
     0     0     1
```

Kaikki sarakkeet ovat tukisarakkeita, joten V :n sarakkeet ovat varmasti LRT.

Siispä ne muodostavat \mathbb{R}^3 :n kannan. Aivan, kuten äskeisessä esimerkissä, saadaan mahdollisimman yksinkertainen tilanne operoimalla ominaisvektorikannassa.

Merkitään ominaisvektoreita: $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, ja esitetään:

$\mathbf{x}_0 = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + c_3 \mathbf{v}_3$. Koska $P \mathbf{v}_k = \lambda_k \mathbf{v}_k$, $k = 1, 2, 3$, on $\mathbf{x}_1 = P \mathbf{x}_0 = c_1 \lambda_1 \mathbf{v}_1 + c_2 \lambda_2 \mathbf{v}_2 + c_3 \lambda_3 \mathbf{v}_3$, missä $\lambda_1 = 1$.

Jatkamalla iterointia, saadaan:

$$\mathbf{x}_k = c_1 \mathbf{v}_1 + c_2 \lambda_2^k \mathbf{v}_2 + c_3 \lambda_3^k \mathbf{v}_3.$$

Koska $0 < \lambda_j < 1$, kun $j = 2, 3$, lähenevät 2 viimeistä termiä kohti 0:aa, kun k kasvaa. Siten prosessia dominoi ykköstä (suurinta) ominaisarvoa vastaava ominaisvektori.

Rajalla päädytään vektoriin $\mathbf{x} = c_1 \mathbf{v}_1$. Kerroin c_1 voidaan määrätä lausumalla alkupistevektori \mathbf{x}_0 ominaisvektorikannassa. Helpompaa on kuitenkin todeta, että kaikki vektorit \mathbf{x}_k ovat prosenttilukuja (johtuen ehdosta: sarakesummat = 1), joten myös rajavektorin \mathbf{x} komponenttien summan on oltava 100.

Siispä $c_1 = \frac{100}{1+2+5} = 12.5$ ja siis $\mathbf{x} = [12.5, 25, 62.5]^T$.

Merkillepantavaa on, että tulos ei riipu alkujakaumasta lainkaan, vaan määräytyy pelkästään matriisin P dominoivasta ominaisvektorista.

1.4.3 Populaatiodynamiikkaa

Esimerkki 1.5 (Täpläpöllöjen henkinjäämistaiselu Kaliforniassa) Vuonna 1990 oli vastakkainasettelu puuteollisuuden ja ympäristönsuojelijoiden välillä. Vanhojen ikimetsien hakkuut ja sen seurauksena mm. täpläpöllöjen tuhoutuminen, vastaan 30000...100000 työpaikan menetys puuteollisuudessa.

Matemaatikot ryhtyivät viileän objektiivisesti tutkimaan.

Täpläpöllön elinkaari voidaan jakaa kolmeen luokkaan:

1. Lapsuus, 0 – 1 v.,
 2. esi aikuisuus, 1 – 2 v.,
 3. aikuisuus, yli 2 v.
- Keskimääräinen elinikä n. 20 v.

Kriittinen ajanjakso on silloin, kun lapsipöllö lähtee pesästään, jolloin sen on löydettävä uusi kotialue ja pari.

Mallinnetaan tilannetta tarkastelemalla populaatiota vuoden välein ja jakamalla se näihin kolmeen luokkaan. Oletetaan, että naaraita ja uroita on saman verran, jolloin tarvitsee tarkastella vain naisväkeä.

Populaatiovektori vuonna k olkoon $\mathbf{x}_k = [l_k, e_k, a_k]$ ("lapsi", "esi aikuinen", "aikuinen").

$$\begin{bmatrix} l_{k+1} \\ e_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.18 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{bmatrix} \begin{bmatrix} l_k \\ e_k \\ a_k \end{bmatrix}$$

1. rivi: Kukin aikuinen naaras synnyttää keskimäärin 0.33 tyttölapsipöllöä vuonna k (vuoden $k+1$ alkuun mennessä).

2. rivi: Keskimäärin 0.18 tyttölasta selviytyy esiaikuisiksi vuoden $k+1$ alkuun.

3. rivi: Vuoden k esiaikuisista 0.71 selviytyy aikuisuuteen ja aikuisista 0.94 jatkaa aikuiselämää vuonna $k+1$.

Yllä olevaa matriisia A voidaan kutsua ”vaihematriisiksi”(”stage matrix”). Tämä on peräisin oikeasta tutkimusaineistosta:

Lamberson et. al: *Dynamic analysis of the viability of the Northern spotted owl in a fragmented forest environment*, julkaisussa: *Conservation biology* vuodelta 1992.

Tässä matriisissa alkio $a_{2,1} = 0.18$ on kriittisin ympäristötekijöiden vaikutuksille. Jos ikimetsät hakataan aukeiksi, niin tämä arvo käy liian pieneksi. Arvo 0.18 on jo liian pieni, kuten nähdään, jos analyysi viedään päätökseen.

Muotoa $\mathbf{x}_{k+1} = A\mathbf{x}_k$ olevaa systeemiä sanotaan **differenssiyhtälöksi**, usein puhutaan myös diskreetistä dynaamisesta systeemistä.

Yllä oleva esimerkki on peräisin kirjasta [Lay] KRE-kirjassa [KRE] vastaavanlainen esimerkki esiintyy nimellä *Leslie model* s. 378, Example 3

Palattaneen tähän esimerkkiin myöhemmin.

Maailman suurin ominaisarvotehtävä – Googlen ”Page rank”

Vuonna 2002 toukokuussa 2.7 miljardia verkkosivua. The World’s Largest Matrix Computation Google’s PageRank is an eigenvector of a matrix of order 2.7 billion. (Huom: amerikkalaisille biljoona on sama kuin meille miljardi (10^9).)

www.mathworks.com/company/newsletters/news_notes/clevescorner/oct02_cleve.html

1.5 Matriisin diagonalisointi, ominaisvektorien LRT

(KRE 7.5)

Edellisissä esimerkeissä näimme, miten riemukasta oli, kun LRT ominaisvektoreita oli riittävästi, avaruuden kannaksi saakka. Tässä kohdassa todetaan, että LRT ominaisvektoreita on ainakin yhtä monta kuin on erillisiä ominaisarvoja.

Esitämme myös yleisessä matriisimuodossa saman asian, jonka esimerkeissämme laskimme tapauksessa, jossa ominaisvektoreita on riittävästi. Tätä sanotaan matriisin diagonalisoinniksi.

Määritelmä 1.3 *Neliömatriisit A ja B ovat similaariset, jos on olemassa säännöllinen (kääntyvä) P siten, että*

$$B = P^{-1}AP. \text{ Merkitään } A \sim B.$$

Lause 1.4 [Similaarisuus säilyttää ominaisarvot, KRE Thm 1]

Olkoon annettu A ja olkoon $B = P^{-1}AP$.

Tapa 1) Olkoon \mathbf{x} matriisin A ominaisarvoon λ liittyvä ominaisvektori. Siiis $A\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{x} \neq \mathbf{0}$, joten $P^{-1}A\mathbf{x} = \lambda P^{-1}\mathbf{x}$. Kirjoitetaan A :n ja \mathbf{x} :n väliin $I = P P^{-1}$ ja merkitään $\mathbf{y} = P^{-1}\mathbf{x}$. Siiis

$\underbrace{P^{-1}AP}_B \underbrace{P^{-1}\mathbf{x}}_y = P^{-1}A\mathbf{x} = \lambda \underbrace{P^{-1}\mathbf{x}}_y$, eli $B\mathbf{y} = \lambda\mathbf{y}$. Kaiken lisäksi $\mathbf{y} \neq \{0\}0$, sillä muutenhan olisi $\mathbf{x} = P\mathbf{y} = P\{0\}0 = \{0\}0$, vastoin \mathbf{x} :n ominaisvektoriutta.

Niinpä λ on myös B :n ominaisarvo (ja \mathbf{y} on vastaava B :n ominaisvektori).

Tapaa 2) $D_B(\lambda) = \det(B - \lambda I) = \det(P^{-1}(A - \lambda I)P) = \det(P^{-1}) \det(A - \lambda I) \det(P) = \det(A - \lambda I) = D_A(\lambda)$. (Tässä käytettiin determinanttien kertosääntöä.)

Nyt tulemme mainittuun ominaisvektorien LRT-lauseeseen.

Lause 1.5 [KRE s. 392] Olkoot $\lambda_1, \dots, \lambda_k$ matriisin A erillisiä ominaisarvoja. Tällöin vastaavat ominaisvektorit ovat LRT

1) Todistetaan ensin 2:n ominaisvektorin tapauksessa. Olk. $\lambda_1 \neq \lambda_2$, ja olkoot \mathbf{v}_1 ja \mathbf{v}_2 vastaavat ominaisvektorit.

Kirjoitetaan vektori yhtälö $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$.

Sovelletaan siihen A :ta (eli kerrotaan vasemmalta A :lla) ja käytetään ominaisuutta $A\mathbf{v}_j = \lambda_j\mathbf{v}_j, j = 1, 2$.

Näin saadaan vektori yhtälöpari:
$$\begin{cases} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0} \\ \lambda_1 c_1\mathbf{v}_1 + \lambda_2 c_2\mathbf{v}_2 = \mathbf{0} \end{cases}$$

Kerrotaan 1. yhtälö λ_1 :llä ja vähennetään alemmasta ylempi, niin saadaan: $c_2(\lambda_2 - \lambda_1)\mathbf{v}_2 = \mathbf{0}$. koska $\lambda_2 - \lambda_1 \neq 0$, on oltava $c_2 = 0$ ja siten myös $c_1 = 0$.

Siten myös $c_1 = 0$.

2) Askel tapauksesta $k = 2$ tapaukseen $k = 3$ saadaan lähtemällä vektori yhtälöstä

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}.$$

Aivan samantyyppisellä eliminaatioaskeleella kuin edellä, päästään jo todistettuun tapaukseen $k = 2$. Tarkemmin sanottuna kerrotaan yhtälö A :lla ja otetaan huomioon $A\mathbf{v}_k = \lambda_k\mathbf{v}_k, k = 1, 2, 3$. Saadaan yhtälöpari:

$$\begin{cases} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0} \\ \lambda_1 c_1\mathbf{v}_1 + \lambda_2 c_2\mathbf{v}_2 + \lambda_3 c_3\mathbf{v}_3 = \mathbf{0} \end{cases}$$

Kerrotaan 1. yhtälö λ_1 :llä ja vähennetään toisesta $\implies c_2(\lambda_2 - \lambda_1)\mathbf{v}_2 + c_3(\lambda_3 - \lambda_1)\mathbf{v}_3 = \mathbf{0}$.

Edellä todistetun perusteella \mathbf{v}_2 ja \mathbf{v}_3 ovat LRT, joten $c_2 = c_3 = 0$, koska $\lambda_2 - \lambda_1 \neq 0$ ja $\lambda_3 - \lambda_1 \neq 0$. Tästä seuraa yhtälön 1. perusteella: $c_1 = 0$. Niinpä $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ on LRT.

Yleinen induktioaskel $k \rightarrow k + 1$ on aivan samanlainen.

Tästä seuraa välittömästi:

Lause 1.6 Jos $n \times n$ -matriisilla A on n erillistä ominaisarvoa, niin \mathbb{R}^n :llä (kompl. tap. \mathbb{C}^n :llä) on A :n ominaisvektoreista koostuva kanta.

Tätä käyttäen oltaisiin Markov-prosessi-tehtävässä voitu välttyä ei-dominoivien ominaisvektoreiden laskemiselta.

Huomautus 1.2 Ominaisvektorikanta voi olla, vaikka esiintyisi moninkertaisia ominaisarvoja, kuten on nähty. Toisaalta olemme nähneet esimerkkejä (ainakin yhden), jossa käy huonosti: 2×2 -matriisilla on vain yksi ominaissuora.

Lause 1.7 \mathbb{R}^n :llä (\mathbb{C}^n :llä) on A :n ominaisvektorikanta, jos ja vain jos A :n jokaisen ominaisarvon algebrallinen ja geometrinen kertaluku yhtyvät, ts. kaikilla ominaisarvoilla λ on $m_\lambda = M_\lambda$.

Muistamme, että aina pätee $m_\lambda \leq M_\lambda$. Olkoot erilliset ominaisarvot $\lambda_1, \dots, \lambda_p$, ja niiden algebralliset kertaluvut k_1, \dots, k_p . Siis $k_1 + \dots + k_p = n$.

1) Jos kertaluvut eivät yhdy, niin jollain ominaisarvolla λ_j on $m_{\lambda_j} < M_{\lambda_j}$. Koska jokainen ominaisvektori kuuluu johonkin ominaisavaruuteen, ei LRT ominaisvektoreita voi olla enempää kuin ominaisavaruuksien dimensioiden summa, joka nyt on aidosti pienempi kuin algebrallisten kertalukujen summa $= n$. Siten ominaisvektorikantaa varten ei ole tarpeeksi LRT vektoreita. (Tässä emme tarvitse edes lausetta 1.5(s. 14).)

2) Oletetaan, että $m_{\lambda_j} = M_{\lambda_j}$ kaikilla $j = 1, \dots, p$.

Nyt lause 1.5 on kaiken avain. Kyse on yksinkertaisesti siitä, että jos niputetaan kutakin erillistä ominaisarvoa vastaavien ominaisavaruuksien kannat omiksi nipuikseen, saadaan yhteensä n vektoria, jotka ovat LRT.

Otetaan havainnollisuuden vuoksi vaikka kolme ominaisarvoa $\lambda_1, \lambda_2, \lambda_3$, joiden kertaluvut olkoot 1, 2, 3 vastaavasti. Olkoot "vektoriniput" $\{\mathbf{u}\}$, $\{\mathbf{v}_1, \mathbf{v}_2\}$, $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

Olkoon $(c_1 \mathbf{u}) + (c_2 \mathbf{v}_1 + c_3 \mathbf{v}_2) + (c_4 \mathbf{w}_1 + c_5 \mathbf{w}_2 + c_6 \mathbf{w}_3) = \mathbf{0}$.

Suluissa olevat vektorilausekkeet ovat kaikki nollavektoreita. Jos nimittäin joku ei olisi, niin otettaisiin summaan kaikki nollasta erilliset sulkulausekkeet. Nämä ovat eri ominaisarvoihin liittyviä ominaisvektoreita, ja siis LRT. Niiden summa $= \mathbf{0}$, joten ristiriita LRT:n kanssa on valmis.

Kussakin sulkulausekkeessa esiintyvät vektorit ovat ao. ominaisavaruuden kantavektoreita (siis LRT), joten on oltava: $(c_1 = 0)$, $(c_2 = c_3 = 0)$, $(c_4 = c_5 = c_6 = 0)$.

Yleinen "niputus" menee pariaatteessa aivan vastaavasti.

Huomautus 1.3 Olennainen asia tuossa "niputuspäätelyssä" on tieto, että otettiinpa mitkä tahansa vektorit eri nipuista, niin ne ovat LRT.

Erilaisten matriisityyppien (kuten symmetristen, ortogonaalisten ym.) ominaisvektorikäytöstä selvitetään jatkossa.

Ominaisvektorikannan merkitys, diagonalisointi

Olkoon $n \times n$ -matriisilla A n LRT ominaisvektoria $\mathbf{u}_1, \dots, \mathbf{u}_n$, jotka siis muodostavat \mathbb{R}^n :n (tai \mathbb{C}^n :) kannan. Suoritetaan yleisesti sama lasku, joka tehtiin sovellusesimerkkien tapauksessa.

Olkoon $\mathbf{x} \in \mathbb{R}^n$, esitetään se tämän kannan avulla: $\mathbf{x} = c_1 \mathbf{u}_1 + \dots + c_n \mathbf{u}_n$. Samoin kuin edellä olleissa esimerkissä (pääakseliesitys, Markovin prosessi) voidaan silloin laskea: $A \mathbf{x} = c_1 A \mathbf{u}_1 + \dots + c_n A \mathbf{u}_n = \lambda_1 c_1 \mathbf{u}_1 + \dots + \lambda_n c_n \mathbf{u}_n$.

Jos ajatellaan matriisia A lineaarikuvauksena $T = T_A$, niin vektorin $\mathbf{x} = [x_1, \dots, x_n]$ kuvan $\mathbf{y} = A \mathbf{x}$ koordinaatit avaruuden peruskannassa saadaan matriisikertolaskulla: $y_j = \sum_{i=1}^n a_{ij} x_i$. Siis kukin kuvapisteen koordinaatti riippuu "mutkikkaan" kaavan välityksellä kaikista lähtöpisteen koordinaateista.

Jos sensijaan esitetään \mathbf{x} ominaisvektorikannassa, niin kuvapisteen j :s koordinaatti saadaan kertomalla lähtöpisteen sama j :s koordinaatti luvulla λ_j . Siten kuvan koordinaatti j riippu pelkästään lähtöpisteen koordinaatista j .

Matriisimuodossa kuvavektorin $\mathbf{y} = A \mathbf{x}$ esitys ominaisvektorikannassa saadaan kertomalla \mathbf{x} :n

koordinaattiesitys $[c_1, \dots, c_n]^T$ diagonaalimatriisilla $\begin{bmatrix} \lambda_1 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \lambda_n \end{bmatrix}$

Todetaan siis, että A :n määräämän lineaarikuvauksen matriisi ominaisvektorikannan suhteen on diagonaalimatriisi.

Asia voidaan esittää puhtaana matriisiyhtälönä, matriisin diagonalisointiesityksenä.

Lause 1.8 (Diagonalisointi) Olkoon $n \times n$ -matriisilla A n kpl LRT ominaisvektoreita (eli \mathbb{R}^n :n tai \mathbb{C}^n :n kanta). Tällöin A voidaan esittää muodossa $A = XDX^{-1}$, missä D on diagonaalimatriisi. Tässä esityksessä X :n sarakkeet ovat A :n ominaisvektorit ja D :n diagonaalilla ovat A :n ominaisarvot samassa järjestyksessä.

Olkoon $\mathbf{x}_1, \dots, \mathbf{x}_n$ A :n ominaisvektorikanta ja $\lambda_1, \dots, \lambda_n$ A :n vastaavat ominaisarvot. Siis $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1, \dots, A\mathbf{x}_n = \lambda_n\mathbf{x}_n$. Nämä voidaan kirjoittaa matriisiyhtälönä:

$$A[\mathbf{x}_1|\mathbf{x}_2|\dots|\mathbf{x}_n] = [\lambda_1\mathbf{x}_1|\lambda_2\mathbf{x}_2|\dots|\lambda_n\mathbf{x}_n],$$

missä $X = [\mathbf{x}_1|\mathbf{x}_2|\dots|\mathbf{x}_n]$ on matriisi, jossa ominaisvektorit \mathbf{x}_j ovat sarakkeina.

Jälkimmäinen matriisi, jossa j :s sarake on kerrottu luvulla λ_j , voidaan esittää diagonaalimatriisina kertomisena oikealta:

$$[\lambda_1\mathbf{x}_1|\lambda_2\mathbf{x}_2|\dots|\lambda_n\mathbf{x}_n] = XD, \quad \text{missä } D = \begin{bmatrix} \lambda_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \lambda_n \end{bmatrix}$$

Näin saadaan yhtälö $AX = XD$. Tässä ei tehty mitään muuta kuin kirjoitettiin ominaisarvojen ja -vektorien määritelmät yhdeksi matriisiyhtälöksi. (Tähän saakka olisimme päässeet olettamatta ominaisvektoreita lineaarisesti riippumattomiksi.)

No nyt, koska ne ovat LRT, niin matriisin X rangi on täysi n (eli kaikki sarakkeet ovat tukisarakkeita). Niinpä on olemassa käänteismatriisi X^{-1} . Kun kerrotaan tuo matriisiyhtälö oikealta X^{-1} :llä, saadaan väitetty esitys: $A = XDX^{-1}$.

Määritelmä 1.4 Matriisia A sanotaan diagonalisoituvaksi, jos edellisen lauseen ehto on voimassa.

Huomautus 1.4 Esityksen yksikäsitteisyydestä:

1) Alkuperäiset yhtälöt voidaan kirjoittaa eri järjestyksessä, mikä aiheuttaa sarakkeiden järjestyksen vaihdon. Tärkeää on, että ominaisvektorit matriisissa X ja ominaisarvot matriisissa D otetaan samassa järjestyksessä.

2) Ominaisvektorit voidaan skaalata millä tahansa nollasta poikkeavalla kertoimella. (Mitään "luonnollisinta" skaalausta ei edes voida esittää.) Tämä kompensoituu käänteismatriisin muodostamisessa, jossa ao. sarake tulee jaetuksi vastaavalla skaalauskertoimella.

Matriisin potenssit

Differenssiyhtälön $\mathbf{x}_{n+1} = A\mathbf{x}_n$ ratkaisu on $\mathbf{x}_n = A^n\mathbf{x}_0$. Tästä syystä on usein tarvetta laskea matriisin potensseja tehokkaasti. Samalla voi paljastua seikkoja, jotka mahdollistavat prosessin analysoinnin.

Jos matriisi A on diagonalisoituva, voidaan kirjoittaa $A = XDX^{-1}$. Nyt

$$A^2 = XD \underbrace{X^{-1}X}_I DX^{-1} = XD^2X^{-1}.$$

Jatkamalla samoin nähdään, että $A^k = XD^kX^{-1}$.

Koska samankokoisten diagonaalimatriisien tulo on diagonaalimatriisi, jonka lävistäjä koostuu tekijämatriisien diagonaalialkioiden tuloista, pätee $(\text{diag}([\lambda_1, \dots, \lambda_n])^k = \text{diag}([\lambda_1^k, \dots, \lambda_n^k]))$.

(Käytämme MATLAB:n tyylistä notaatiota diagonaalimatriisille, se on helppo kirjoittaa, säästää tilaa ja on itsensä selittävä.)

Esimerkki 1.6 Laske lauseke A^k :lle, kun $A = \begin{bmatrix} -2 & 12 \\ -1 & 5 \end{bmatrix}$.

Ratkaisu Diagonalisoidaan:

$$A = VDV^{-1},$$

$$V = \begin{bmatrix} 4 & 3 \\ 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, V^{-1} = \begin{bmatrix} 1 & -3 \\ -1 & 4 \end{bmatrix}.$$

$$\text{Nyt } A^k = VD^kV^{-1} = \begin{bmatrix} 4 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2^k \end{bmatrix} \begin{bmatrix} 1 & -3 \\ -1 & 4 \end{bmatrix}.$$

Kun suoritetaan kertolaskut, saadaan:

$$A^k = \begin{bmatrix} -3 \cdot 2^k + 4 & 12 \cdot 2^k - 12 \\ -2^k + 1 & 4 \cdot 2^k - 3 \end{bmatrix}$$

Tässä tapauksessa saatiin jopa yksinkertainen kaava A^k :lle.

Yleensä nyt ei ihan näin pitkälle päästä. Joka tapauksessa saadaan valtava säästö aritmetiikassa ja usein saadaan selkeä näkemys alunperin sotkuiseen asiaan.

Esimerkki 1.7 Lasketaan yksi esimerkki MATLAB:lla. Tässä näkyy samalla MATLAB:n kätevyys matriisioperaatioissa. Alla oleva eig-komento on hyvä esimerkki tarkoituksenmukaisesta suunnittelusta (“intelligent design”).

Esimerkin pitäisi selittää hyvin itsensä myös henkilölle, joka ei ole MATLAB:ia koskaan käyttänyt.

```
>> A=[0 -4 -6;-1 0 -3;1 2 5] % Muodostetaan matriisi A.
A =

     0     -4     -6
    -1     0     -3
     1     2     5
>> [V,D]=eig(A) % Ominaisvektorit V:n sarakkeiksi,
V = % ominaisarvot D:n diagonaalille
    -0.8165    0.5774   -0.9269
    -0.4082    0.5774   -0.0850
     0.4082   -0.5774    0.3656
D =
     1.0000         0         0
         0     2.0000         0
         0         0     2.0000

% Koska 2 on kaksinkertainen ominaisarvo,
% emme tiedä, onko V kääntävä.
>> VI=inv(V) % Voidaan käyttää inv-komentoa käänteismatriisiin
VI = % muodostamiseen.(Jos sitä ei ole, tulee ‘herja’)

    -2.4495   -4.8990   -7.3485
    -1.7321   -1.2075   -4.6716
         0     3.5634    3.5634
```

Tarkistetaan, laitetaan A ja VDV^{-1} vierekkäin (ja lisätään rajaviiva editoimalla).


```
>> [A V*D*VI]
```

```
ans =
```

```
      0   -4.0000   -6.0000 | -0.0000   -4.0000   -6.0000
-1.0000      0   -3.0000 | -1.0000   -0.0000   -3.0000
 1.0000   2.0000   5.0000 |  1.0000   2.0000   5.0000
```

Tästäpä nähdään, että A ja $V D V^{-1}$ ovat samat.

Matriisin A potensseja

Lasketaan vaikkapa A^{16} .

```
>> A^16
```

```
ans =
```

```
 -65534   -262140   -393210
 -65535   -65534    -196605
  65535    131070    262141
```

Lasketaan diagonalisoimalla:

```
>> D16=D^16
```

```
D16 =
```

```
1.0e+04 *
 0.0001      0      0
 0   6.5536      0
 0      0   6.5536
```

Näin ei oikeastaan kannattaisi laskea. Nyt voimme nimittäin laskea pisteittäisen potenssin, kun on kyse diagonaalimatriisista. Silloin aritmetiikkaa on paljon vähemmän.

```
>> D16=D.^16 % diag(diag(D).^16) % olisi vielä tehokkaampi (help diag)
```

```
D16 =
```

```
1.0e+04 *
 0.0001      0      0
 0   6.5536      0
 0      0   6.5536
```

```
>> format bank
```

```
>> V*D16*VI
```

```
ans =
```

```
 -65534.00   -262140.00   -393210.00
 -65535.00   -65534.00   -196605.00
  65535.00    131070.00    262141.00
```

```
>> A^16 % Tässä taas vertailuksi raakaa voimaa:
```

```
ans =
```

```
 -65534.00   -262140.00   -393210.00
 -65535.00   -65534.00   -196605.00
  65535.00    131070.00    262141.00
```

Maankäyttöesimerkki diagonalisoimalla

Edellä esitettiin lähtövektori ominaisvektorikannan avulla. Katsotaan, miltä homma näyttäisi diagonalisoimalla. Esitetään matriisi P muodossa $P = VDV^{-1}$. Iteraatiojonon k :s termi on $x_k = P^k x_0$. Mutta $P^k = VD^kV^{-1}$.

$$D^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.8^k & 0 \\ 0 & 0 & 0.6^k \end{bmatrix} \quad \text{Kun } k \rightarrow \infty, \text{ niin } Dlim = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

```
>> P=[0.8 0.1 0;0.1 0.7 0.1;0.1 0.2 0.9]
```

```
>> [V,D]=eig(P)
```

```
V =
```

```
   -0.1826   -0.7071    0.4082  
   -0.3651   -0.0000   -0.8165  
   -0.9129    0.7071    0.4082
```

```
D =
```

```
   1.0000    0    0  
    0   0.8000    0  
    0    0   0.6000
```

```
>> VI=inv(V)
```

```
VI =
```

```
   -0.6847   -0.6847   -0.6847  
   -1.0607   -0.3536    0.3536  
    0.3062   -0.9186    0.3062
```

```
>> Dlim=diag([1,0,0])
```

```
Dlim =
```

```
    1    0    0  
    0    0    0  
    0    0    0
```

```
>> V*Dlim*VI
```

```
ans =
```

```
   0.1250   0.1250   0.1250  
   0.2500   0.2500   0.2500  
   0.6250   0.6250   0.6250
```

```
>> Plim=V*Dlim*VI
```

```
Plim =
```

```
0.1250    0.1250    0.1250
0.2500    0.2500    0.2500
0.6250    0.6250    0.6250
```

```
>> Plim*[30 20 50]'
```

```
ans =
```

```
12.5000
25.0000
62.5000
```

```
>> Plim*[99 .5 .5]'
```

```
ans =
```

```
12.5000
25.0000
62.5000
```

Samaan päädyttiin kuin ennen.

Ominaisvektorikanta vai diagonalisointi

Kysehän on samasta asiasta esitettynä eri muodossa. Tämäntyyppisissä tehtävissä näyttäisi aiemmin esitetty ominaisvektorikannalla operointi matriisiin diagonalisointiesitystä lyhyemmältä tavalta.