Ominaisarvot

Lineaaristen yhtälöryhmien ratkaisemisen ohella toinen tärkeä lineaarialgebran tehtävä on ominaisarvojen etsiminen.

Olkoon F jokin operaattori ja $\mathbf{x}$ vektori. Jos F ei muuta vektorin $\mathbf{x}$ suuntaa, $\mathbf{x}$ on operaattorin ıominaisvektori. Sille pätee

$$F(\mathbf{x}) = \lambda \mathbf{x}, \tag{1}$$

missä λ on reaali- tai kompleksiluku, ominaisvektoria vastaava ıominaisarvo. Operaattori muuttaa siis vain vektorin pituutta ominaisarvon ilmoittamalla kertoimella.

Jos F on lineaarinen operaattori, on $F(a\mathbf{x}) = aF(\mathbf{x}) = a\lambda\mathbf{x}$, joten myös $a\mathbf{x}$ on ominaisvektori. Ominaisvektorit eivät siten määräydy yksikäsitteisesti, vaan ne voidaan kertoa millä tahansa vakiolla.

Seuraavassa tarkastellaan vain matriisien ominaisarvoja. Matriisia voi pitää operaattorina, joka muuntaa vektoria jollakin tavalla. Ominaisvektorin tapauksessa tämä muunnos on pelkkä skaalaus.

Olkoon siis $\mathbf{A}$ neliömatriisi. Jos on olemassa reaali- tai kompleksiluku λ ja vektori $\mathbf{x}$ siten, että

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}, \tag{2}$$

λ on matriisin ominaisarvo ja $\mathbf{x}$ ominaisvektori. Yhtälö (2) voidaan kirjoittaa myös muotoon

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

Jotta yhtälöllä olisi ei-triviaali ratkaisu, on oltava

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0. \tag{3}$$

Kun determinentti kirjoitetaan auki, saadaan ıkarakteristinen polynomi, jonka nollakohtia ominaisarvot ovat.

Jos matriisi on reaaliarvoinen ja symmetrinen, ominaisarvot ovat reaalisia. Muuten ainakin osa ominaisarvoista voi olla kompleksisia. Kompleksiset ominaisarvot esiintyvät aina konjugaattipareina.

Esimerkiksi matriisin

$$\begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$$

karakteristinen polynomi on

$$\det \begin{pmatrix} 1-\lambda & 4 \\ 1 & 1-\lambda \end{pmatrix} = \lambda^2 - 2\lambda - 3,$$

jonka nollakohdista saadaan ominaisarvot $\lambda_1 = 3$ ja $\lambda_2 = -1$.

Ominaisarvojen laskeminen karakteristisen polynomin avulla on hyvin työlästä, jos matriisi on vähänkään isompi. Jo yhtälön kertoimien määrittäminen on hankalaa. Menetelmä soveltuu muutenkin huonosti numeeriseen laskentaan.

Ominaisarvojen laskemiseksi matriisi on muunnettava sopivampaan muotoon. Gaussin eliminointi muuttaisi sen kolmiomatriisiksi, mutta valitettavasti ominaisarvot eivät säily muunnoksessa.

Sen sijaan ısimilariteettimuunnoksissa ominaisarvot eivät muutu. Muunnos on muotoa

$$\mathbf{A} \to \mathbf{A}' = \mathbf{S}^{-1}\mathbf{A}\mathbf{S},$$

missä $\mathbf{S}$ on mikä tahansa ei-singulaarinen matriisi. Matriiseja $\mathbf{A}$ ja $\mathbf{A}'$ sanotaan ısimilaarisiksi.

QR-hajotelma

Paljon käytetty menetelmä ominaisarvojen laskemiseksi tunnetaan QR-menetelmänä. Menetelmä on iteraatio, jossa lasketaan toistuvasti matriisin tietynlainen hajotelma, QR-hajotelma. Tämä hajotelma voidaan laskea äärellisellä askelmäärällä, ja sillä on muutakin käyttöä. Tutustumme aluksi tämän hajotelman laskemiseen.

QR-hajotelmassa matriisi kirjoitetaan muotoon

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

missä $\mathbf{Q}$ on ortogonaalinen matriisi ja $\mathbf{R}$ yläkolmiomatriisi. Tällainen hajotelma voidaan muodostaa onpa matriisi minkä muotoinen tahansa.

Hajotelma voidaan laskea useilla eri tavoilla:

1) Householderin muunnos

2) Givensin rotaatiot

3) Gram-Schmidt-ortogonalisointi

Seuraavassa käsitellään kahta ensimmäistä esimerkkien avulla. Menetelmät selvinnevät niistä helpommin kuin formaalista algoritmista.

Householderin muunnos

Tarkastellaan aluksi QR-hajotelman laskemista Householderin muunnoksilla. Tarkoituksena on etsiä joukko muunnoksia, joista kullakin nollataan yhdeltä sarakkeelta lävistäjän alapuolella olevat alkiot.

Oletetaan, että hajotettavana on matriisi

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

Poimitaan aluksi tämän ensimmäinen pystyrivi

$$\mathbf{x}_1 = \mathtt{a}(:,1) = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} 2$$

ja lasketaan siitä vektori

$$\mathbf{u}_1 = \mathbf{x}_1 - \parallel \mathbf{x}_1 \parallel \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} 0 = \begin{pmatrix} -0.7416574 \\ 1 \\ 1 \end{pmatrix} 2$$

Tämän avulla muodostetaan Householderin muunnosmatriisi

$$\mathbf{P}_1 = \mathbf{I} - 2\frac{\mathbf{u}_1\mathbf{u}_1^T}{\parallel u_1 \parallel^2} = \begin{pmatrix} 0.8017837 & 0.2672612 & 0.5345225 \\ 0.2672612 & 0.6396433 & -0.7207135 \\ 0.5345225 & -0.7207135 & -0.4414270 \end{pmatrix}$$

Voidaan osoittaa, että tämä on ortogonaalinen matriisi. Asia on on helppo todeta laskemalla minkä tahansa kahden sarakkeen skalaaritulo. Tulot ovat nollia, joten matriisin pystyvektorit ovat keskenään ortogonalisia.

Kun alkuperäinen matriisi kerrotaan tällä muunnoksella, ensimmäiseltä sarakkeelta nollautuvat kaikki lävistäjän alapuolella olevat alkiot:

$$\mathbf{A}_1 = \mathbf{P}_1\mathbf{A} = \begin{pmatrix} 3.7416574 & 2.4053512 & 2.9398737 \\ 0. & 0.4534522 & -0.6155927 \\ 0. & -0.0930955 & -2.2311854 \end{pmatrix}$$

Sitten poimitaan toiselta sarakkeelta vektori

$$\mathbf{x}_2 = \mathtt{a}(2:3,1) = \begin{pmatrix} 0.4534522 \\ -0.0930955 \end{pmatrix},$$

josta

$$\mathbf{u}_2 = \mathbf{x}_2 - \parallel \mathbf{x}_2 \parallel \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -0.0094578 \\ -0.0930955 \end{pmatrix}.$$

Tästä saadaan toinen muunnosmatriisi

$$\mathbf{P}_2 = \mathbf{I} - 2\frac{\mathbf{u}_2\mathbf{u}_2^{\mathrm{T}}}{\parallel \mathbf{u}_2 \parallel^2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.9795688 & -0.2011093 \\ 0 & -0.2011093 & -0.9795688 \end{pmatrix}$$

Kerrotaan $\mathbf{A}_1$ muunnoksella, jolloin toisen sarakkeen lävistäjän alapuoliset alkiot nollautuvat.

$$\mathbf{A}_2 = \mathbf{P}_2\mathbf{A}_1 = \begin{pmatrix} 3.7416574 & 2.4053512 & 2.9398737 \\ 0 & 0.4629100 & -0.1543033 \\ 0 & 0 & 2.3094011 \end{pmatrix}$$

Näin matriisi onkin muunnetta yläkolmiomatriisiksi. Jos matriisi on isompi, toistetaan samaa menettelyä kullekin pystyriville, kunnes kaikki lävistäjän alapuolella olevat alkiot on nollattu.

Hajotelman matriisit saadaan nyt muunnosmatriisien avulla:

$$\mathbf{Q} = \mathbf{P}_1\mathbf{P}_2 = \begin{pmatrix} 0.8017837 & 0.1543033 & -0.5773503 \\ 0.2672612 & 0.7715167 & 0.5773503 \\ 0.5345225 & -0.6172134 & 0.5773503 \end{pmatrix}$$

$$\mathbf{R} = \mathbf{A}_2 = \mathbf{P}_2\mathbf{P}_1\mathbf{A} = \begin{pmatrix} 3.7416574 & 2.4053512 & 2.9398737 \\ 0 & 0.4629100 & -0.1543033 \\ 0 & 0 & 2.3094011 \end{pmatrix}.$$

Matriisi $\mathbf{R}$ on siis itse asiassa viimeisessä muunnoksessa laskettu matriisi $\mathbf{A}_k$, joten alkuperäistä matriisia $\mathbf{A}$ ei tarvita. Jos siis tilaa on säästettävä, kukin matriiseisti $\mathbf{A}_i$ voidaan aina kirjoittaa edellisen päälle. Myöskään aikaisempia muunnosmatriiseja $\mathbf{P}_i$ ei tarvitse tallettaa, mutta matriisin $\mathbf{Q}$ alkuarvoksi asetetaan $\mathbf{P}_1$, ja kussakin vaiheessa $\mathbf{Q}$ kerrotaan aina uudella muunnosmatriisilla $\mathbf{P}_i$.

Tarkistuksen vuoksi voidaan laskea, että hajotelman tekijöiden tulo on todellakin sama kuin alkuperäinen matriisi:

$$\mathbf{QR} = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{pmatrix}.$$

Matriisin $\mathbf{Q}$ ortogonaalisuus nähdään esimerkiksi laskemalla tulo $\mathbf{QQ}^{\mathrm{T}}$:

$$\mathbf{QQ}^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Yleisessä tapauksessa hajotelman matriisit ovat

$$\begin{aligned}\mathbf{Q} &= \mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_n,\\ \mathbf{R} &= \mathbf{P}_n\cdots\mathbf{P}_2\mathbf{P}_1\mathbf{A}.\end{aligned}$$

Givensin rotaatiot

Toinen yleisesti käyttety menetelmä perustuu Givensin kiertomatriiseihin.

$$P_{kl}(\theta) = \begin{pmatrix} 1 & & & & \\ & \cos\theta & \cdots & \sin\theta & \\ & \vdots & 1 & vdots & \\ & -\sin\theta & \cdots & \cos\theta & \\ & & & & 1 \end{pmatrix}.$$

Tämä on ortogonaalinen matriisi.

Ominaisarvojen laskeminen

Ominaisarvot voidaan ratkaista iteratiivisesti QR-algoritmilla, jossa käytetään edellä ollutta QR-hajotelmaa. Jos lähdetään liikkeelle suoraan alkuperäisestä matriisista, tehtävä on laskennallisesti hyvin raskas. Siksi matriisi kannattaa ensin muuntaa sopivampaan muotoon.

Tehtävänä on löytää joukko similariteettimuunnoksia, joilla matriisi saadaan johonkin yksinkertaisempaan muotoon.

Neliömatriisi on lohkokolmiomatriisi, jos se on muotoa

$$\begin{pmatrix} T_{11} & T_{12} & T_{13} & \cdots & T_{1n} \\ 0 & T_{22} & T_{23} & \cdots & T_{2n} \\ 0 & 0 & \vdots & & \\ 0 & 0 & 0 & \cdots & T_{nn} \end{pmatrix},$$

missä alimatriisit T_{ij} ovat neliömatriiseja. Voidaan osoittaa, että tällaisen matriisin ominaisarvot ovat lävistäjälohkojen T_{ii} ominaisarvot.

Niinpä jos matriisi on lävistäjä- tai kolmiomatriisi, ominaisarvot ovat lävistäjällä olevat alkiot. Jos tällainen muoto löytyy, tehtävä on ratkaistu. Yleensä äärellisellä määrällä similariteettimuunnoksia ei kuitenkaan päästä aivan niin pitkälle.

Jos alkuperäinen matriisi on symmetrinen, se voidaan muuntaa tridiagonaaliseksi ilman, että ominaisarvot muuttuvat. Yleisen matriisin tapauksessa tuloksena on Hessenbergin matriisi, joka on muotoa

$$\mathbf{H} = \begin{pmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \end{pmatrix}$$

Tarvittavat muunnokset voidaan taaskin tehdä joko Householderin muunnoksilla tai Givensin rotaatioilla. Nyt menetelmää muutetaan sen verran, että välittömästi lävistäjän alapuolella olevan vinorivin alkioita ei nollata.

Matriisin muuntaminen Householderin muunnoksilla

Tarkastellaan ensimmäisenä esimerkkinä symmetristä matriisia

$$\mathbf{A} = \begin{pmatrix} 4 & 3 & 2 & 1 \\ 3 & 4 & -1 & 2 \\ 2 & -1 & 1 & -2 \\ 1 & 2 & -2 & 2 \end{pmatrix}$$

Ruvetaan muuntamaan tätä Householderin muunnoksilla. Nyt ensimmäiseltä pystyriviltä poimitaan vektoriin $\mathbf{x}_1$ vain lävistäjän alapuoliset alkiot:

$$\mathbf{x}_1 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

Näistä muodostetaan vektori

$$\mathbf{u}_1 = \mathbf{x}_1 - \parallel \mathbf{x}_1 \parallel \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} 0 = \begin{pmatrix} -0.7416574 \\ 2 \\ 1 \end{pmatrix}$$

Tämän avulla saadaan matriisi

$$\mathbf{p}_1 = \mathbf{I} - 2\mathbf{u}_1\mathbf{u}_1^{\mathrm{T}} / \parallel \mathbf{u}_1 \parallel = \begin{pmatrix} 0.8017837 & 0.5345225 & 0.2672612 \\ 0.5345225 & -0.4414270 & -0.7207135 \\ 0.2672612 & -0.7207135 & 0.6396433 \end{pmatrix}$$

ja tästä edelleen Householderin muunnosmatriisi

$$\mathbf{P}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.8017837 & 0.5345225 & 0.2672612 \\ 0 & 0.5345225 & -0.4414270 & -0.7207135 \\ 0 & 0.2672612 & -0.7207135 & 0.6396433 \end{pmatrix}$$

Nyt voidaan suorittaa matriisin **A** similariteettimuunnos. Muunnosmatriisi on symmetrinen, joten sitä ei tarvitse transponoida.

$$\mathbf{A}_1 = \mathbf{P}_1\mathbf{A}\mathbf{P}_1 = \begin{pmatrix} 4 & 3.7416574 & 0 & 0 \\ 3.7416574 & 2.4285714 & 1.2977396 & 2.1188066 \\ 0 & 1.2977396 & 0.0349563 & 0.2952113 \\ 0 & 2.1188066 & 0.2952113 & 4.5364723 \end{pmatrix}$$

Toinen sarake käsitellään samalla tavalla. Ensin muodostetaan vektori $\mathbf{x}_2$

$$\mathbf{x}_2 = \begin{pmatrix} 1.2977396 \\ 2.1188066 \end{pmatrix}$$

ja tästä edelleen

$$\mathbf{u}_2 = \mathbf{x}_2 - \parallel \mathbf{x}_2 \parallel (1,0)^{\mathrm{T}} = \begin{pmatrix} -1.1869072 \\ 2.1188066 \end{pmatrix}$$

ja

$$\mathbf{p}_2 = \begin{pmatrix} 0.5223034 & 0.8527597 \\ 0.8527597 & -0.5223034 \end{pmatrix}$$

ja varsinainen muunnosmatriisi

$$\mathbf{P}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 00 & \\ 0 & 0 & 0.5223034 & 0.8527597 \\ 0 & 0 & 0.8527597 & -0.5223034 \end{pmatrix}$$

Muunnos tuottaa matriisin

$$\mathbf{A}_2 = \mathbf{P}_2\mathbf{A}_1\mathbf{P}_2 = \begin{pmatrix} 4 & 3.7416574 & 0 & 0 \\ 3.7416574 & 2.4285714 & 2.4846467 & 0 \\ 0 & 2.4846467 & 3.5714286 & -1.8708287 \\ 0 & 0 & -1.8708287 & 1 \end{pmatrix}$$

Saimme siis tridiagonaalisen matriisin, kuten symmetrisen matriisin tapauksessa pitääkin.

Muutetaan sitten alkuperäinen matriisi epäsymmetriseksi:

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 3 & 1 \\ 3 & 4 & -2 & 1 \\ 2 & -1 & 1 & -2 \\ 1 & 2 & -2 & 2 \end{pmatrix}$$

Muunnos etenee samalla tavalla kuin edelläkin, ja tulokseksi saadaan

$$\mathbf{A}_2 = \begin{pmatrix} 4 & 3.4743961 & -1.3039935 & -0.4776738 \\ 3.7416574 & 1.7857143 & 2.9123481 & 1.1216168 \\ 0 & 2.0934787 & 4.2422252 & -1.0307759 \\ 0 & 0. & -1.2980371 & 0.9720605 \end{pmatrix},$$

joka on Hessenbergin muotoa.

QR-algoritmi

Nyt matriisi on muunnettu alkuperäistä yksinkertaisempaan tridiagonaaliseen tai Hessenbergin muotoon, jolla on kuitenkin samat ominaisarvot kuin alkuperäisellä matriisilla. Olkoon tämä muunnettu matriisi $\mathbf{H}$. Seuraavaksi voidaan ruveta etsimään matriisin ominaisarvoja. Tämä tapahtuu iteroimalla.

Valitaan aluksi $\mathbf{A}_1 = \mathbf{H}$. Sitten toistetaan seuraavia askelia:

– Lasketaan QR-hajotelma $\mathbf{Q}_i\mathbf{R}_i = \mathbf{A}_i$.

– Lasketaan uusi matriisi $\mathbf{A}_{i+1} = \mathbf{R}_i\mathbf{Q}_i$.

QR-hajotelman matriisi $\mathbf{Q}$ on ortogonaalinen ja $\mathbf{A} = \mathbf{QR}$, joten $\mathbf{R} = \mathbf{Q}^{\mathrm{T}}\mathbf{A}$. Niinpä

$$\mathbf{RQ} = \mathbf{Q}^{\mathrm{T}}\mathbf{AQ}$$

on similariteettimuunnos ja säilyttää ominaisarvot.

Matriisien $\mathbf{A}_i$ jono suppenee kohti yläkolmio- tai lohkomatriisia, josta ominaisarvot voidaan poimia.

Viimeisen esimerkin tapauksessa raja-arvo on 50 iteraatiokierroksen jälkeen

$$\begin{pmatrix} 7.0363389 & -0.7523758 & -0.7356716 & -0.3802631 \\ 4.265E-08 & 4.9650342 & -0.8892339 & -0.4538061 \\ 0 & 0 & -1.9732687 & -1.3234202 \\ 0 & 0 & 0 & 0.9718955 \end{pmatrix}$$

Matriisin ominaisarvot ovat nyt lävistäjällä olevat luvut. Jos ominaisarvot ovat kompleksiarvoisia, matriisi on lohkomatriisi, jonka ominaisarvot ovat lävistäjällä olevien 2×2-alimatriisien ominaisarvoja.