Määritelmä

Neliömatriisit A ja B ovat **similaareja** *toistensa suhteen, $A \sim B$, jos on olemassa kääntyvä matriisi P, jolle pätee $A = PBP^{-1}$.*

Lause

Jos matriisit ovat similaarit, on niillä sama karakteristinen polynomi ja siten samat ominaisarvot monikertoineen.

Tod. Jos $B = P^{-1}AP$, niin

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(A - \lambda I)P.$$

Siten

$$\begin{aligned} det(B - \lambda I) =& det\left(P^{-1}(A - \lambda I)P\right) \\ =& det\left(P^{-1}\right) det(A - \lambda I) det(P) = det(A - \lambda I) \end{aligned}$$

eli molemmat karakteristiset polynomit ovat samat. QED

Similaarisuus 2

Esimerkkejä, 1.: Similaareilla matriiseilla on sama determinantti:

$$\det A = \det\left(PBP^{-1}\right) = \det P \det B \det\left(P^{-1}\right)$$
$$= \det P \det B (\det P)^{-1} = \det B \ .$$

2.: $n \times n$ matriisin A **jälki** (engl. trace) on päädiagonaalisumma:

$$Tr(A) = a_{11} + \cdots + a_{nn} \ .$$

Voidaan osoittaa, että

$$Tr(A) = \lambda_1 + \cdots + \lambda_n \ .$$

Koska spektri säilyy similariteetissä, tarkoittaa tämä sitä, että myös jälki on muuttumaton similariteetissä.

3.: Lause ei yleisty ekvivalenssiksi. Esimerkiksi matriiseilla

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{ja} \quad \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

on sama karakteristinen polynomi, $(2-\lambda)^2$, ja siten sama spektri, mutta ne eivät ole similaareja. Miksi?

Diagonalisointi 1 (Kr. 8.4, Lay 5.3)

Matriisin A diagonalisointi tarkoittaa sellaisen diagonaalimatriisin määräämistä, joka on similaari A:n kanssa: $A = PDP^{-1}$. Jos tämä onnistuu, on siitä paljon etua; esimerkiksi $A^k = PD^kP^{-1}$, $k \geq 0$.

Lause

$n \times n$ matriisi A on diagonalisoituva joss A:lla on n lineaarisesti riippumatonta ominaisvektoria. Nämä ominaisvektorit muodostavat matriisin P sarakevektorit ja A:n ominaisarvot antavat D:n alkiot vastaavassa järjestyksessä.

Diagonalisointi 1 (Kr. 8.4, Lay 5.3)

Matriisin A diagonalisointi tarkoittaa sellaisen diagonaalimatriisin määräämistä, joka on similaari A:n kanssa: $A = PDP^{-1}$. Jos tämä onnistuu, on siitä paljon etua; esimerkiksi $A^k = PD^kP^{-1}$, $k \geq 0$.

Lause

$n \times n$ matriisi A on diagonalisoituva joss A:lla on n lineaarisesti riippumatonta ominaisvektoria. Nämä ominaisvektorit muodostavat matriisin P sarakevektorit ja A:n ominaisarvot antavat D:n alkiot vastaavassa järjestyksessä.

Esimerkki 1:

Matriisin $A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$ karakteristinen polynomi on $-(\lambda - 1)(\lambda + 2)^2$, joten ominaisarvot ovat 1 ja -2 (kaksinkertainen). Edelliselle voidaan laskea ominaisvektori $(1, -1, 1)^T$, jälkimmäiselle $(-1, 1, 0)^T$ ja $(-1, 0, 1)^T$, jotka ovat lineaarisesti riippumattomia.

Diagonalisointi 2

Näistä muodostetaan kannanvaihto- ja diagonaalimatriisit

$$P = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

P on ei-singulaari, koska se on täyttä rangia. Näille siis pätee $A = PDP^{-1}$.

Diagonalisointi 2

Näistä muodostetaan kannanvaihto- ja diagonaalimatriisit

$$P = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

P on ei-singulaari, koska se on täyttä rangia. Näille siis pätee $A = PDP^{-1}$.

Esimerkki 2:
Matriisin $B = \begin{pmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix}$ karakteristinen yhtälö on sama kuin yllä. Mutta vaikka ominaisarvot ovat samat (monikertoineen), niin ominaisarvoon -2 liittyviä lineaarisesti riippumattomia ominaisvektoreita on vain yksi. Siten B ei ole diagonalisoituva.

Diagonalisointi 2

Näistä muodostetaan kannanvaihto- ja diagonaalimatriisit

$$P = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

P on ei-singulaari, koska se on täyttä rangia. Näille siis pätee $A = PDP^{-1}$.

Esimerkki 2:

Matriisin $B = \begin{pmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix}$ karakteristinen yhtälö on sama kuin yllä. Mutta vaikka ominaisarvot ovat samat (monikertoineen), niin ominaisarvoon -2 liittyviä lineaarisesti riippumattomia ominaisvektoreita on vain yksi. Siten B ei ole diagonalisoituva.

Lause

Jos $n \times n$ matriisilla on n erillistä ominaisarvoa, on se diagonalisoituva.

Jordanin muodot 1

Yleisesti voidaan osoittaa, että jokainen matriisi "melkein diagonalisoituu". Jos $n \times n$ matriisin A spektri on $\{\lambda_1, \ldots, \lambda_p\}$ monikerroin $\{m_1, \ldots, m_p\}$, $(p \le n)$, niin $A = PJP^{-1}$, jossa J on ns. **Jordanin matriisi**. Se on muotoa

$$\begin{pmatrix} J_1 & 0 & \ldots \\ 0 & J_2 & \ldots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

jossa **Jordanin lohko** J_i on skalaari λ_i, jos $m_i = 1$. Jos taas $m_i > 1$, on J_i $m_i \times m_i$ yläkolmiomatriisi, jonka diagonaali on identtisesti λ_i ja sen viereinen sivudiagonaali on identtisesti 1. J_i on siis itse asiassa yksinkertainen **nauhamatriisi**, jolla on varsin helppo operoida. Erityisesti J:n potenssit voidaan laskea vaivattomasti ja siten PJP^{-1}-esityksen avulla myös A^k. Näiden avulla taas päästään (myöhemmin kurssilla) käsiksi matriisieksponenttiin e^A.

Jordanin muodot 2

Esimerkki 2, jatkoa: Ominaisarvoon 1 liittyvä ominaisvektori on $u_3 = (1, -1, 1)^T$ ja arvoon -2 $u_1 = (-1, 1, 0)^T$. Puuttuva kolmas, lineaarisesti riippumaton, vektori saadaan ratkaistua **yleistettynä ominaisvektorina** yhtälöstä

$$(B - \lambda I)\, u_2 = u_1,$$

jossa siis nyt $\lambda = -2$. Ratkaisu on $u_2 = (-1, 0, 1)^T$. Edelleen *Mathematicalla*

```
In[42]:= a = {{2, 4, 3}, {-4, -6, -3}, {3, 3, 1}}
Out[42]= {{2, 4, 3}, {-4, -6, -3}, {3, 3, 1}}
In[43]:= Eigensystem[a]
Out[43]= {{-2, -2, 1}, {{-1, 1, 0}, {0, 0, 0}, {1, -1, 1}}}
In[44]:= p = Transpose[{{-1, 1, 0}, {-1, 0, 1}, {1, -1, 1}}]
Out[44]= {{-1, -1, 1}, {1, 0, -1}, {0, 1, 1}}
In[45]:= MatrixForm[ Inverse[p] . a . p ]
Out[45]//MatrixForm=
```

$$\begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

jossa $\begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}$ on siis ominaisarvoon -2 liittyvä Jordanin lohko.

Gerschgorinin lause (Kr. 20.7)

Lause

Olkoon λ $n \times n$ matriisin $A = [a_{ij}]$ ominaisarvo. Silloin jollekin indeksille $j \in \{1, 2, \ldots, n\}$ pätee $|\lambda - a_{jj}| \leq \sum_{k=1,\ k \neq j}^{n} |a_{jk}|$.

Tod.: Olkoon x ominaisarvoa λ vastaava ominaisvektori ja x_j sen itseisarvoltaan suurin komponentti. Yhtälöryhmän $(A - \lambda I)x = 0$ j:s rivi on $\sum_{k=1,\ k \neq j}^{n} a_{jk} x_k + (a_{jj} - \lambda) x_j = 0$. Jakamalla tämä x_j:lla, soveltamalla kolmioepäyhtälöä, sekä tulosta $|x_i / x_j| \leq 1$, $\forall i$ kaava seuraa. QED

Gerschgorinin lause (Kr. 20.7)

Lause

Olkoon λ $n \times n$ matriisin $A = [a_{ij}]$ ominaisarvo. Silloin jollekin indeksille $j \in \{1, 2, \ldots, n\}$ pätee $|\lambda - a_{jj}| \leq \sum_{k=1,\ k \neq j}^{n} |a_{jk}|$.

Tod.: Olkoon x ominaisarvoa λ vastaava ominaisvektori ja x_j sen itseisarvoltaan suurin komponentti. Yhtälöryhmän $(A - \lambda I)x = 0$ j:s rivi on $\sum_{k=1,\ k \neq j}^{n} a_{jk}x_k + (a_{jj} - \lambda)x_j = 0$. Jakamalla tämä x_j:lla, soveltamalla kolmioepäyhtälöä, sekä tulosta $|x_i/x_j| \leq 1, \ \forall i$ kaava seuraa. QED

Esimerkki: Olkoon

$$A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 5 & 1 \\ \frac{1}{2} & 1 & 1 \end{pmatrix} .$$

Tällöin Lauseen määräämät ns. **Gerschgorinin kiekot** (kompleksitasossa) ovat $|z| \leq 1$, $|z - 5| \leq \frac{3}{2}$ ja $|z - 1| \leq \frac{3}{2}$, joissa ominaisarvot siis sijaitsevat.
Lause on siis sitä hyödyllisempi, mitä pienemmät ovat matriisin ei-diagonaalialkiot (pienemmät kiekkojen säteet).

Dominoiva ominaisarvo 1

Esimerkki: Tarkastellaan kaksitilaista systeemiä, joka päivittyy satunnaisesti diskreetein välein. Siirtymätodennäköisyys $p_{ij} = Tn(j|i),\ i,j \in \{1,2\}$ antaa ehdollisen todennäköisyyden siirtymälle tilasta i tilaan j yhdessä aikayksikössä. Jos t_i ja t_i' ovat tilan i (absoluuttiset) todennäköisyydet perättäisinä ajanhetkinä, pätee päivitykselle **kokonaistodennäköisyyden laki:**

$$\begin{cases} t_1' = t_1 p_{11} + t_2 p_{21} \\ t_2' = t_1 p_{12} + t_2 p_{22} \end{cases} \quad \text{eli} \quad t' = tP, \quad P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}.$$

Yllä t, t' ovat vaakavektoreita. Kun tälläisen **Markovin ketjun**, annetaan kehittyä, asymptoottisesti $tP^n \to p$. Vektori p on systeemin **tasapainotila**. Siten täytyy päteä $p = pP$ eli p on matriisin P **vasen ominaisvektori**, joka liittyy ominaisarvoon 1. Aiemmin on määritelty vain matriisin *oikea* ominaisvektori. Ekvivalentisti pystyvektoreilla: $P^T p^T = p^T$.

Olkoon nyt $P = \begin{pmatrix} 0.6 & 0.4 \\ 0.9 & 0.1 \end{pmatrix}$ ja P^T:n ominaisarvoon 1 liittyvä oikea ominaisvektori $p^T = (0.692308, 0.307692)^T$. Toisaalta voidaan laskea vaikkapa

```
In[127]:=  MatrixForm[ MatrixPower[Transpose[p], 100] ]
Out[127]//MatrixForm=
 0.692308  0.692308
 0.307692  0.307692
```

Näyttääkin siltä, että $(P^T)^n \to \begin{pmatrix} 0.692308 & 0.692308 \\ 0.307692 & 0.307692 \end{pmatrix}$, miksi?

Dominoiva ominaisarvo 2

P^T:n spektri on $\{1, -0.3\}$ ja vastaavat normalisoidut lineaarisesti riippumattomat oikeat ominaisvektorit ovat $p^T = (0.692308, 0.307692)^T$ ja $(-0.5, 0.5)^T$. Määritellään näistä normaaliin tapaan

$$M = \begin{pmatrix} 0.692308 & -0.5 \\ 0.307692 & 0.5 \end{pmatrix} \quad \text{ja} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -0.3 \end{pmatrix}.$$

Tällöin $P^T = MDM^{-1}$ josta edelleen $\left(P^T\right)^n = MD^nM^{-1}$, $\forall n \geq 1$. Mutta koska tietenkin $(-0.3)^n \to 0$, tämä implikoi välittömästi että

$$\left(P^T\right)^n \longrightarrow M \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} M^{-1} = \begin{pmatrix} 0.692308 & 0.692308 \\ 0.307692 & 0.307692 \end{pmatrix}.$$

Jos q on arvoa $\lambda = 1$ vastaava vasen ominaisvektori eli toteuttaa $qP^T = q$, ja se normalisoidaan siten, että $qp^T = 1$, saadaan $q = (1, 1)$ ja nähdään, että yllä viimeisenä oleva matriisi on p^Tq. Se on rangia 1 oleva ns. **projektiomatriisi**.

Ylläoleva yleistyy huomattavasti: Markov-matriisi ja $\lambda = 1$ eivät ole välttämättömiä oletuksia. Sen sijaan on oleellista, että $m \times m$ matriisilla A on **dominoiva ominaisarvo** $\lambda_1 > |\lambda_i|$, $\forall i \neq 1$. Silloin voidaan osoittaa, että $\frac{A^n}{\lambda_1^n}$ suppenee kohti λ_1:tä vastaavien normeerattujen vasemman ja oikean ominaisvektorin ulkoista tuloa.

Ortogonaalisuus vektoreille 1 (Kr. 7.9, 9.2, Lay 6.1)

$n \times 1$ matriisit $u, v \in \mathbf{R}^n$ ovat pystyvektoreja.

Määritelmä

Skalaari $u^T v = u \cdot v$ on vektoreiden u ja v **sisä- eli pistetulo**.

Pistetulo on liitännäinen, vaihdannainen ja distributiivinen skalaarilla kertomisen suhteen. Selvästi $u^T u \geq 0$, josta seuraa

Määritelmä

Vektorin **normi eli pituus** *on $\|u\| = \sqrt{u^T u}$.*

Tämä määritelmä on itse asiassa täsmälleen Pythagoraan teoreema.

Ortogonaalisuus vektoreille 1 (Kr. 7.9, 9.2, Lay 6.1)

$n \times 1$ matriisit $u, v \in \mathbf{R}^n$ ovat pystyvektoreja.

Määritelmä

Skalaari $u^T v = u \cdot v$ on vektoreiden u ja v **sisä- eli pistetulo**.

Pistetulo on liitännäinen, vaihdannainen ja distributiivinen skalaarilla kertomisen suhteen. Selvästi $u^T u \geq 0$, josta seuraa

Määritelmä

Vektorin **normi eli pituus** *on* $\|u\| = \sqrt{u^T u}$.

Tämä määritelmä on itse asiassa täsmälleen Pythagoraan teoreema.

Määritelmä

Vektorit u ja v ovat **kohtisuorassa eli ortogonaaliset** *jos* $u^T v = 0$.

Ortogonaalisuus vektoreille 2

Määritelmä

Vektorin $u \in \mathbf{R}^n$ **ortogonaalikomplementti**, *merkitään* $u^\perp$, *on joukko* $\{v \in \mathbf{R}^n | \; v \cdot u = 0\}$. *Aliavaruuden* $W \subset \mathbf{R}^n$ *ortogonaalikomplementti* $W^\perp$ *on kaikkien niiden vektoreiden joukko, jotka ovat ortogonaalisia jokaista* W*:n nollasta eroavaa alkiota vastaan.*

Esimerkki: Vektorin $u \in \mathbf{R}^3, \; u \neq 0$ ortogonaalikomplementti on vektoria vastaan kohtisuora, origon kautta kulkeva taso. Kyseisen tason ortogonaalikomplementti taas on tason normaalivektoreiden joukko (joista u on eräs).

Voidaan helposti osoittaa, että vektoreille $u, v \in \mathbf{R}^n, \; u, v \neq 0$ pätee aina $u \cdot v = \|u\|\|v\| \cos \alpha$, jossa α on vektoreiden välinen kulma. Huomaa, että dimensio $n \geq 2$ on mielivaltainen; kaksi vektoria u ja v määräävät korkeintaan kaksiulotteisen aliavaruuden, joten kulma siinä on aina määriteltävissä.

Ortogonaalisuus vektoreille 3

Määritelmä

Vektorikokoelma $\{u_1, \ldots, u_k\} \subset \mathbf{R}^n, k \leq n$ *on* **ortogonaalinen joukko**, *jos* $u_i \cdot u_j = 0,\ i \neq j$ *ja* $u_j \cdot u_j > 0\ \forall i, j.$
$\mathbf{R}^n$*:n* **ortogonaalikanta** *on n:n elementin ortogonaalinen joukko.*

Vektori ja mikä tahansa (ei nolla)vektori sen ortogonaalikomplementista muodostavat ortogonaalisen joukon. Jos ortogonaalisessa joukossa on k elementtiä, virittävät ne k-ulotteisen aliavaruuden $\mathbf{R}^n$:än.

Esimerkkejä, 1. $\mathbf{R}^n$:n kanoninen ortogonaalikanta muodostuu **alkeisvektoreista** $e_i = (0, \ldots, 1, \ldots, 0),\ i = 1, \ldots, n$, jossa ainoa ei-häviävä alkio on i:nnessä koordinaatissa.

2. Vektorit $u_1 = (3, 1, 1)^T,\ u_2 = (-1, 2, 1)^T$ ja $u_3 = \left(-\frac{1}{2}, -2, \frac{7}{2}\right)^T$ muodostavat ortogonaalisen kannan $\mathbf{R}^3$:en. Voidaan helposti tarkistaa, että $y = (6, 1, -8)^T = u_1 - 2u_2 - 2u_3$. Miten tehdä tällainen esitys systemaattisesti annetuille vektoreille?

Ortogonaalinen projektio (Lay 6.3)

Lause

Jos $\{u_1, \ldots, u_p\}$ *on ortogonaalinen kanta* $\mathbf{R}^n$*:n aliavaruudelle, niin silloin mielivaltainen aliavaruuden vektori voidaan hajoittaa*

$$y = \sum_{i=1}^{p} c_i u_i, \quad \text{jossa} \quad c_i = \frac{y \cdot u_i}{u_i \cdot u_i}\ .$$

Tod.: $y \cdot u_j = \left(\sum_{i=1}^{p} c_i u_i\right) \cdot u_j = c_j u_j \cdot u_j.$ QED

Suure $\frac{y \cdot u}{u \cdot u} u$ on y:n **ortogonaaliprojektio** vektorin u suuntaan.

Esimerkki (ed. jatkoa): $u_1 \cdot u_1 = 11$, $u_2 \cdot u_2 = 6$ ja $u_3 \cdot u_3 = \frac{33}{2}$. Koska sisätulot ovat $y \cdot u_1 = 11$, $y \cdot u_2 = -12$ ja $y \cdot u_3 = -33$, seuraavat annetut kertoimet 1, -2 ja -2.

Jos ortogonaaliselle joukolle pätee $u_i \cdot u_j = \delta_{ij} = \begin{cases} 0 & , \; i \neq j \\ 1 & , \; i = j \end{cases}$, on kyseessä **ortonormaali** joukko (jolle laskenta on erityisen helppoa!)

Gram-Schmidt ortogonalisointi 1 (Lay 6.4)

Lause

Olkoon $\{u_1, u_2, \ldots, u_p\}$ *kanta* $\mathbf{R}^n$*:n aliavaruudelle* W*. Kun määritellään*

$$\begin{aligned}
v_1 &= u_1 \\
v_2 &= u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\
v_3 &= u_3 - \frac{u_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\
&\vdots \\
v_p &= u_p - \frac{u_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_p \cdot v_2}{v_2 \cdot v_2} v_2 - \cdots - \frac{u_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1} \ ,
\end{aligned}$$

niin $\{v_1, v_2, \ldots, v_p\}$ *muodostaa* W*:n ortogonaalisen kannan.*

Normalisoimalla $\left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \ldots, \frac{v_p}{\|v_p\|} \right\}$ saadaan ortonormaali kanta.

Gram-Schmidt ortogonalisointi 2 (Lay 6.4)

Esimerkki: Kokoelma $\{u_i\}_{i=1}^3 = \left\{(1,0,0)^T, (1,1,0)^T, (1,1,1)^T\right\}$ on helppo nähdä lineaarisesti riippumattomaksi. Muodostetaan siitä ortonormaali kanta $\mathbf{R}^3$:lle:

$$
\begin{aligned}
v_1 &= u_1 = (1,0,0)^T \\
v_2 &= (1,1,0)^T - (1,1,0)(1,0,0)^T(1,0,0)^T = (0,1,0)^T \\
v_3 &= (1,1,1)^T - (1,1,1)(1,0,0)^T(1,0,0)^T - (1,1,1)(0,1,0)^T(0,1,0)^T = (0,0,1)^T \ ,
\end{aligned}
$$

jossa vektorit v_i tulivat (sattumalta) valmiiksi normalisoituneina ja saatiin $\mathbf{R}^3$:n kanoninen kanta. Mikäli alkuperäistä joukkoa käsitellään vaikkapa käänteisessä järjestyksessä, saadaan erilainen ortogonaalinen systeemi:

$$
\begin{aligned}
w_1 &= u_3 = (1,1,1)^T \\
w_2 &= (1,1,0)^T - \frac{(1,1,0)(1,1,1)^T}{(1,1,1)(1,1,1)^T}(1,1,1)^T = \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)^T \\
w_3 &= (1,0,0)^T - \frac{(1,0,0)(1,1,1)^T}{(1,1,1)(1,1,1)^T}(1,1,1)^T - \frac{(1,0,0)\left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)^T}{\left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)\left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)^T}\left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)^T \\
&= \left(\frac{1}{2}, -\frac{1}{2}, 0\right)^T \ ,
\end{aligned}
$$

joka on edelleen normalisoitava, jos halutaan ortonormaali joukko.

Ortogonaalinen matriisi 1 (Kr. 8.3, Lay 6.2)

Lause

Jos neliömatriisin U sarakkeet ovat ortonormaali joukko, pätee sille $U^T U = I$.

Tod.: $U = (u_1|u_2|\cdots|u_n)$ (u_i on pystyvektori), joten $U^T U = I$ seuraa välittömästi sisätuloista: $u_i \cdot u_j = \delta_{ij}$. QED

Lause määrittelee **ortogonaalisen matriisin**. Niillä on erinomaisia ominaisuuksia:

- $U^{-1} = U^T$,
- rivit ovat myös ortonormaali joukko,
- $(Ux) \cdot (Uy) = (Ux)^T (Uy) = x^T U^T Uy = x^T y = x \cdot y$, josta erikoistapauksena seuraa, että $\|Ux\| = \|x\|$, kuten myös se, että ortogonaalinen matriisi säilyttää vektorien väliset kulmat.

Ortogonaalinen matriisi 2 (Kr. 8.3, Lay 1.9)

Esimerkkejä, 1.:

$$K(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

kierto tasossa vastapäivään kulman θ verran, on ortogonaalinen transformaatio ja $K(\theta)$ on ortogonaalinen matriisi jokaiselle θ. Kuvauksessa $x \mapsto K(\theta)x$ vektorin pituus ei muutu. Huomaa, että inverssi on harvinaisen helppo: transponointi tai $\theta \to -\theta$! K:n ortogonaalinen lähisukulainen on

$$H(\phi) = \begin{pmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix}.$$

Miten H kuvaa tasovektorit? Mitä merkitsee se, det $K(\theta) = 1$, mutta det $H(\phi) = -1$?

Ortogonaalinen matriisi 2 (Kr. 8.3, Lay 1.9)

Esimerkkejä, 1.:

$$K(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

kierto tasossa vastapäivään kulman θ verran, on ortogonaalinen transformaatio ja $K(\theta)$ on ortogonaalinen matriisi jokaiselle θ. Kuvauksessa $x \mapsto K(\theta)x$ vektorin pituus ei muutu. Huomaa, että inverssi on harvinaisen helppo: transponointi tai $\theta \to -\theta$!
K:n ortogonaalinen lähisukulainen on

$$H(\phi) = \begin{pmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix}.$$

Miten H kuvaa tasovektorit? Mitä merkitsee se, $\det K(\theta) = 1$, mutta $\det H(\phi) = -1$?

2.: Olkoon P $n \times n$-matriisi, jonka jokaisessa sarakkeessa ja jokaisella rivillä on yksi ykkönen muiden alkioiden ollessa nollia. Siten matriisin rivit ja sarakkeet ovat aina alkeisvektoreita $e_i = (0, \ldots, 1, \ldots, 0) \in \mathbf{R}^n$. Selvästi $e_i \cdot e_j = \delta_{ij}$, joten ortogonaalisuus seuraa välittömästi.
Olkoon $w = (1, 2, 3, \ldots, n)^T$. On helppo nähdä, että Pw on pystyvektori, jossa luvut $1, 2, \ldots, n$ ovat uudessa järjestyksessä. P on esimerkki **permutaatiomatriisista**. Sillä ei ole ilmeistä geometrista tulkintaa, mutta se on hyvin tärkeä kombinatoriikassa. Esimerkiksi jos $n = 52$, on kaikkien tällaisten P-matriisien joukko sama kuin korttipakan sekoitusten joukko (joita on $52! \approx 8.0658 \cdot 10^{67}$ kpl...)

QR-hajotelma 1 (Kr. 20.9, Lay 6.4)

Olkoon A $m \times n$-matriisi, jolla on lineaarisesti riippumattomat sarakkeet$\{x_1, x_2, \ldots, x_n\}$ eli nämä siis virittävät n-ulotteisen avaruuden, A:n sarakeavaruuden. Sarakkeet voidaan ortonormalisoida Gram-Schmidtillä ja muodostaa näistä matriisi $Q = (u_1 | \cdots | u_n)$. Siten

$$x_k = \sum_{j=1}^{k} r_{jk} u_j + \sum_{j=k+1}^{n} 0 \cdot u_j$$

Saadaan myös $r_{kk} > 0$ valitsemalla kantaan joko u_k tai $-u_k$. Määrittelemällä $r_k = (r_{1k}, \ldots, r_{kk}, 0 \ldots, 0)^T$ nähdään, että $x_k = Qr_k,\ \ k = 1, \ldots, n$. Asettamalla $R = (r_1 | \cdots | r_n)$ saadaan

Lause

QR-hajotelma *on* $A = QR$, *jossa* Q*:n sarakkeet ovat ortonormaali kanta* A*:n sarakeavaruudelle ja* R *on kääntyvä yläkolmiomatriisi, jonka diagonaali on positiivinen.*

QR-hajotelma 2

Esimerkki: Matriisin

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

sarakkeet voidaan ortonormalisoida Gram-Schmidtin proseduurilla ja saadaan

$$Q = \begin{pmatrix} \frac{1}{2} & -\frac{3}{\sqrt{12}} & 0 \\ \frac{1}{2} & \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{2} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{6}} \\ \frac{1}{2} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{6}} \end{pmatrix}.$$

Koska nyt pätee $Q^T Q = I$, saadaan $Q^T A = Q^T QR = IR = R$, joten

$$R = \begin{pmatrix} 2 & \frac{3}{2} & 1 \\ 0 & \frac{3}{\sqrt{12}} & \frac{2}{\sqrt{12}} \\ 0 & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}.$$

QR-hajotelma on tärkeä numerisen lineaarialgebran algoritmi ja se on implementoitu moniin matemaattisiin ohjelmistoihin.

Erityisesti jos A on $n \times n$ neliömatriisi, ovat hajotelman Q ja R myös samaa kokoa olevia neliömatriiseja. Jos edelleen A on kääntyvä, on se Käänteismatriisilauseen (katso kertaus) nojalla täyttä rangia, $rank(A) = n$. Siten A:n sarakeavaruus (A:n sarakevektorien virittämä avaruus) on $\mathbf{R}^n$. Edellisen konstruktion matriisin Q sarakkeet ovat siten ortonormaali kanta $\mathbf{R}^n$:lle. Siten tässä tapauksessa Q on todellakin ortogonaalimatriisi. Toisaalta R on aina neliömatriisi ja kääntyvä ($det(R) = r_{11} \cdots r_{nn} > 0$). Niinpä tässä tärkeässä erikoistapauksessa pätee

$$A^{-1} = (QR)^{-1} = R^{-1}Q^{-1} = R^{-1}Q^T \ ,$$

jossa myös R^{-1} on helppo laskettava, koska R on kolmiomatriisi.