

Matriisilaskenta/MS-A0004

Luentopruju

Syksy 2021

Tuomas Sahlsten

Perustuu Riikka Kangaslammen ja Nuutti Hyvösen materiaaleihin

Versio 1.0 (15. lokakuuta 2021)

Matematiikan ja Systeemianalyysin laitos, Aalto Yliopisto

Sisällys

1	Vektorit ja kompleksiluvut	3
1.1	Vektorit	3
1.2	Vektorien koordinaatisto	4
1.3	Vektorien skalaaritulo ja ortogonaalisuus	4
1.4	Vektoritulo	5
1.5	Napakoordinaatit	8
1.6	Kompleksiluvut	9
2	Lineaarinen yhtälöryhmä matriiseilla	14
2.1	Matriisit ja yhtälöryhmä matriisimuodossa	14
2.2	Gaussin eliminaatio	17
3	Matriisialgebra	24
3.1	Lineaarikuvaukset	24
3.2	Matriisien laskutoimitukset	28
3.3	Transpoosi ja käänteismatriisi	31
3.4	Käänteismatriisin laskeminen	33
3.5	Determinantti	35
4	Ominaisarvot ja -vektorit	39
4.1	Motivaatio ja määritelmät	39
4.2	Laskeminen	43
4.3	Sovelluksia dynaamisiin systeemeihin ja Markovin ketjuihin	46
4.4	Kertaluvut	49
5	Matriisihajotelmat	51
5.1	Diagonalisointi	51
5.2	Unitaarinen diagonalisointi	54
5.3	Singulaariarvohajotelma (SVD)	57
5.4	SVD:n laskeminen	59
5.5	Matriisin approksimointi SVD:llä ja kohinan poistaminen datasta	61

Tervetuloa!

Tervetuloa *Matriisilaskenta* / MS-A0004 kurssille Aalto yliopistossa! Kurssin tavoitteena on antaa johdanto *matriiseihin*, jotka ovat matemaattisia objekteja mitkä esiintyvät keskeisessä roolissa matematiikassa ja monissa moderneissa käytännön sovelluksissa kuten fysiikassa, tekoälyssä ja koneoppimisessa, suurten tietokantojen käsittelyssä, ja monissa muissa aiheissa. Kurssin suoritettuaan opiskelija osaa esittää lineaariset yhtälöryhmät matriisimuotoisina, ratkaista matriisimuotoiset yhtälöryhmät Gaussin eliminaatiolla, suorittaa matriisien peruslaskutoimitukset, laskea neliömatriisin ominaisarvot ja ymmärtää matriisihajotelmien merkityksen. Käsittelemme lisäksi näistä käsitteistä eri sovelluksia jotta kurssin jälkeen jää käteen toivotavasti ymmärrys miksi matriisit ovat niin keskeisessä roolissa monilla aloilla.

Kurssilla ei ole kirjaa mutta muutamissa kohdissa viitataan kirjaan

Gilbert Strang: *Introduction to Linear Algebra*, 5th Edition - MIT Mathematics, ISBN : 978-09802327-7-6, <http://math.mit.edu/~gs/linearalgebra/>.

Onnea kurssille!

- Tuomas

Luku 1

Vektorit ja kompleksiluvut

1.1 Vektorit

Reaalinen n -ulotteinen avaruus on joukko $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in \mathbb{R}\}$. Sen pisteet voidaan mieltää myös (paikka)vektoreiksi, merkitään niitä pystyvektoreina:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

Määritellään kaksi laskutoimitusta avaruudessa $\mathbb{R}^n$:

- i) yhteenlasku, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$; $\mathbf{x} + \mathbf{y} \in \mathbb{R}^n$
- ii) skalaarilla kertominen, $\alpha \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$; $\alpha \mathbf{x} \in \mathbb{R}^n$

Alkioittain, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \alpha \mathbf{x} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}$$

Vektori $\mathbf{0} \in \mathbb{R}^n$, jonka kaikki alkiot ovat nollia, on *nollavektori*.

Kaikille $\mathbf{x} \in \mathbb{R}^n$ pätee $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

1.2 Vektorien koordinaatisto

Koordinaatisto avaruuteen $\mathbb{R}^n$ muodostetaan kiinnittämällä origo, ns. *kantavektorit* ja suunnistuvuus.

Tasossa kantavektorit ovat esimerkiksi yksikkövektorit $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ja $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, jolloin yleinen vektori $\mathbf{r} \in \mathbb{R}^2$ voidaan kirjoittaa ns. lineaarikombinaationa:

$$\mathbf{r} = \overrightarrow{OP} = x_1\mathbf{i} + x_2\mathbf{j} \hat{=} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Avaruudessa kantavektorit ovat esimerkiksi yksikkövektorit $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ ja $\mathbf{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, jolloin $\mathbf{r} \in \mathbb{R}^3$ voidaan kirjoittaa lineaarikombinaationa:

$$\mathbf{r} = \overrightarrow{OP} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \hat{=} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

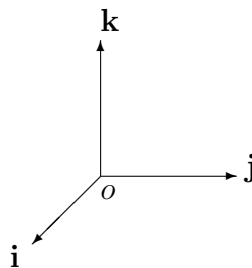
Huom! Kantavektoreita voi olla useampia kuin yllämainitut yksikkövektorit, kannassa tarkoituksena on vain pystyä löytämään joku kokoelma vektoreita niin, että jokainen $\mathbb{R}^n$ alkio voidaan kirjoittaa niiden lineaarikombinaationa. Palaamme tähän myöhemmin.

Suunnistuvuus:

Tasossa: Jos $\mathbf{a}$ kiertyy $\mathbf{b}$:n päälle vastapäivään lyhintä reittiä, on pari $\{\mathbf{a}, \mathbf{b}\}$ *suunnistettu positiivisesti*.

Avaruudessa: Jos kolmikko $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ toimii oikean käden säännön $\{P, E, K\}$ mukaisesti, se on *positiivisesti suunnistettu*.

Kolmiulotteisessa avaruudessa siis esim. karteesinen koordinaatisto $\{o, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$, missä yksikkövektorit $\mathbf{i}$, $\mathbf{j}$ ja $\mathbf{k}$ muodostavat positiivisesti suunnistetun kannan.



1.3 Vektorien skalaaritulo ja ortogonaalisuus

Yksikkövektorit yllä toteuttavat myös geometrisesti enemmän kuin ominaisuuden, että jokainen vektori voidaan kirjoittaa niiden lineaarikombinaationa. Nimittäin nämä vektorit ovat myös kuvan perusteella kohtisuorassa toisiinsa nähden ja niiden pituus on yksi. Kirjallisuudessa kutsumme tällöin, että vektorit ovat ortonormeerattuja keskenään. Tämän voi formalisoida ns. skalaaritulon avulla.

Määritelmä 1.1

- Vektoreiden $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ *skalaaritulo* (eli sisätulo eli pistetulo) on

$$\mathbf{a} \cdot \mathbf{b} = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \alpha_3 \beta_3.$$

Vektorit ovat kohtisuorassa, jos $\mathbf{a} \cdot \mathbf{b} = 0$, merkitään $\mathbf{a} \perp \mathbf{b}$. Sanomme myös, että tällöin ne ovat *ortogonaalisia*.

- Vektorin $\mathbf{a}$ (Euklidinen) *pituus* on luku

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}$$

Ortogonaaliset vektorit $\mathbf{a}$ ja $\mathbf{b}$ ovat tällöin *ortonormeerattuja*, jos niiden pituudet ovat molemmilla 1

Esimerkkinä $\mathbf{i} \cdot \mathbf{j} = 0$ yksikkökantavektoreilla.

Lause 1.2

$$\mathbf{a} \cdot \mathbf{b} = \begin{cases} |\mathbf{a}||\mathbf{b}| \cos \angle(\mathbf{a}, \mathbf{b}), & \text{jos } \mathbf{a} \neq 0, \mathbf{b} \neq 0 \\ 0, & \text{jos } \mathbf{a} = 0 \text{ tai } \mathbf{b} = 0 \end{cases}$$

Esimerkki 1.3 (lasketaan luennolla)

Olkoon $\mathbf{a} = 2\mathbf{i} + 3\mathbf{j}$, $\mathbf{b} = -\mathbf{i} + \mathbf{j}$, $\mathbf{c} = \mathbf{i} + 2\mathbf{j}$. Laske

- $\mathbf{a} \cdot \mathbf{b}$
- $\mathbf{c} \cdot (\mathbf{a} + \mathbf{b})$
- $3\mathbf{a} \cdot (4\mathbf{c} - 3\mathbf{b})$

1.4 Vektoritulo

Vektoritulo voidaan muodostaa vain (kolmiulotteisessa) avaruudessa.

Määritelmä 1.4

Olkoot $\mathbf{a}$, $\mathbf{b}$ kolmiulotteisia vektoreita. *Vektori- eli ristitulo* on vektori $\mathbf{a} \times \mathbf{b}$, jolle pätee:

1. $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin \angle(\mathbf{a}, \mathbf{b})$
2. $\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$, $\mathbf{a} \times \mathbf{b} \perp \mathbf{b}$
3. vektorit $\mathbf{a}$, $\mathbf{b}$, $\mathbf{a} \times \mathbf{b}$ muodostavat oikeakätisen systeemin

Jos $\mathbf{a} = \mathbf{0}$ tai $\mathbf{b} = \mathbf{0}$, on $\mathbf{a} \times \mathbf{b} = \mathbf{0}$.

Esimerkki 1.5

Vektorien $\mathbf{i}$ ja $\mathbf{j}$ vektoritulo $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, sillä

$$|\mathbf{i} \times \mathbf{j}| = |\mathbf{i}||\mathbf{j}| \sin \angle(\mathbf{i}, \mathbf{j}) = 1 \cdot 1 \cdot \sin(\pi/2) = 1,$$

$\mathbf{k}$ on kohtisuorassa kumpaakin vektoria vastaan, ja $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ muodostavat oikeakätisen systeemin.

Vektoritulo $\mathbf{ei}$ ole

- vaihdannainen: $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
- liitännäinen: $(\mathbf{i} \times \mathbf{j}) \times \mathbf{j} = \mathbf{k} \times \mathbf{j} = -\mathbf{i} \neq \mathbf{i} \times (\mathbf{j} \times \mathbf{j}) = \mathbf{i} \times \mathbf{0} = \mathbf{0}$.

Sille kuitenkin pätee

- $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$
- $(\alpha \mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\alpha \mathbf{b}) = \alpha(\mathbf{a} \times \mathbf{b})$.

Kannassa $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ vektoritulo saa nasevan muodon:

Olkoon

$$\begin{aligned}\mathbf{a} &= \alpha_1 \mathbf{i} + \alpha_2 \mathbf{j} + \alpha_3 \mathbf{k}, \\ \mathbf{b} &= \beta_1 \mathbf{i} + \beta_2 \mathbf{j} + \beta_3 \mathbf{k}.\end{aligned}$$

Tällöin voidaan suoralla laskulla osoittaa, että

$$\mathbf{a} \times \mathbf{b} = (\alpha_2 \beta_3 - \alpha_3 \beta_2) \mathbf{i} + (\alpha_3 \beta_1 - \alpha_1 \beta_3) \mathbf{j} + (\alpha_1 \beta_2 - \alpha_2 \beta_1) \mathbf{k}.$$

Myöhemmin tällä kurssilla opimme determinantin käsitteen ja näemme, että yo. voidaan kirjoittaa lyhyesti muodossa

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{vmatrix}.$$

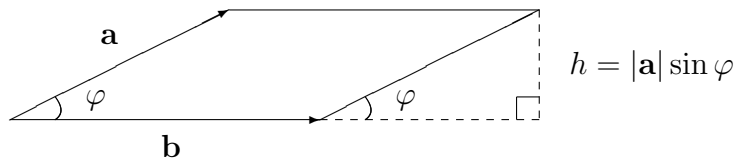
Esimerkki 1.6 (lasketaan luennolla)

Olkoon $\mathbf{a} = \mathbf{i} + \mathbf{j}$, $\mathbf{b} = -\mathbf{i} + 2\mathbf{j}$, $\mathbf{c} = 2\mathbf{j} + \mathbf{k}$. Laske

1. $\mathbf{a} \times \mathbf{b}$
2. $\mathbf{b} \times \mathbf{c}$
3. $\mathbf{a} \times \mathbf{c}$

Lause 1.7

Vektoreiden $\mathbf{a}$ ja $\mathbf{b}$ virittämän suunnikkaan ala on $|\mathbf{a} \times \mathbf{b}|$.

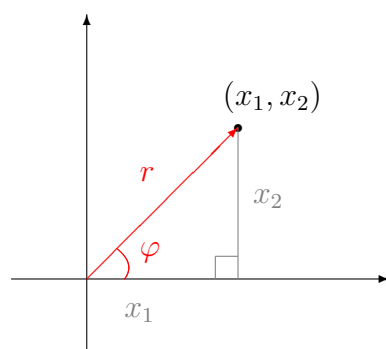
Todistus

[Todistus]

$$\text{Ala} = \text{kanta} \cdot \text{korkeus} = |\mathbf{a}||\mathbf{b}| \sin \varphi = |\mathbf{a} \times \mathbf{b}|.$$

□

1.5 Napakoordinaatit



$$r = \sqrt{x_1^2 + x_2^2}$$

$$\tan \varphi = \frac{x_2}{x_1}$$

$$x_1 = r \cos \varphi$$

$$x_2 = r \sin \varphi$$

Jos $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, niin

$$\mathbf{x} = (r \cos(\varphi), r \sin(\varphi)) = r (\cos(\varphi), \sin(\varphi)),$$

missä $r = |\mathbf{x}| = \sqrt{x_1^2 + x_2^2}$ on pisteen etäisyys origosta

ja $\varphi = \arg(\mathbf{x}) \in \mathbb{R}$ on $\mathbf{x}$:n *argumentti* eli kulma $\mathbf{x}$:n paikkavektorin ja vaaka-akselin välillä.

Luvut

$$\begin{cases} r = |\mathbf{x}| & \in [0, \infty[, \\ \varphi = \arg(\mathbf{x}) & \in \mathbb{R} \end{cases}$$

ovat $\mathbf{x}$:n *napakoordinaatit*.

Huom.

- $\cos(\varphi) = \cos(\varphi + 2\pi)$ ja $\sin(\varphi) = \sin(\varphi + 2\pi)$, joten argumentti on monikäsitteinen (kulmaan voi lisätä 2π -monikertoja)!
Usein halutaan, että $\arg(\mathbf{x}) \in [0, 2\pi[$ tai $\arg(\mathbf{x}) \in]-\pi, \pi]$.
- Jos $x_1 \neq 0$, niin

$$\tan(\arg(\mathbf{x})) = \frac{x_2}{x_1} \quad \text{eli} \quad \arg(\mathbf{x}) = \arctan\left(\frac{x_2}{x_1}\right).$$

Huomaa, että funktion $\arctan$ “päähaara” saa arvoja vain välillä $] -\pi/2, \pi/2[$, joten sen antamaan kulmaan joutuu joskus lisäämään/vähentämään luvun π (piirrä kuva!).

- Origin vaihekulma $\arg(0)$ on “epämääräinen”.

Esimerkki 1.8 (lasketaan luennolla)

Mitkä ovat karteesisissa koordinaateissa ilmoitetun pisteen $\mathbf{x} = (\sqrt{3}, 1)$ napakoordinaatit?

1.6 Kompleksiluvut

Luonnolliset luvut $\mathbb{N} = \{1, 2, 3, 4, 5, \dots\}$ ovat luonnollinen joukko aloittaa laskeminen.

Jos halutaan ratkaista muotoa $x + n = m$, $n, m \in \mathbb{N}$, olevat yhtälöt, kaipaa lukujen joukko kuitenkin laajennusta kokonaislukujen joukoksi: $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$.

Tämäkään ei riitä muotoa $nx = m$, $m, n \in \mathbb{Z}$, $n \neq 0$, olevien yhtälöiden ratkaisemiseen, vaan tarvitaan laajennus rationaalilukuihin:

$$\mathbb{Q} = \left\{ \frac{m}{n} \mid m, n \in \mathbb{Z}, n \neq 0 \right\}.$$

Rationaalilukujenkaan joukosta ei löydy ratkaisua yhtälölle $x^2 = 2$. Tätä varten tehdään vielä laajennus, otetaan mukaan myös irrationaaliluvut, jolloin saadaan reaalilukujen joukko $\mathbb{R}$.

Onko kaikki nyt hyvin?

Mikä on yhtälön $x^2 = -1$ ratkaisu?

Määritellään *imaginääriyksikkö* i olemaan luku, jolle pätee

$$i^2 = -1.$$

Imaginääriyksikköä i voitaisiin siis tavallaan kutsua myös -1 :n neliöjuureksi.

Määritelmä 1.9

Kompleksiluku on muotoa

$$a + ib$$

oleva esitys, jossa a ja b ovat reaalilukuja ja i imaginääriyksikkö.

$\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$ on kompleksilukujen joukko.

Samaistetaan vektori $(x, y) \in \mathbb{R}^2$ ja kompleksiluku $x + iy \in \mathbb{C}$, eli

$$\begin{aligned} \mathbb{R}^2 &= \{(x, y) : x, y \in \mathbb{R}\} \\ \cong \mathbb{C} &= \{x + iy : x, y \in \mathbb{R}\}. \end{aligned}$$

Esimerkki 1.10 (lasketaan luennolla)

Piirrä kompleksiluvut $1 + 2i$, $2 - i$, i , -2 ja $-3 - 2i$ koordinaatistoon.

Tulkitsemalla kompleksiluvut tason vektoreiksi nähdään heti, miten niiden yhteenlasku pitää suorittaa.

Olkoot $z = x + iy$ ja $w = a + ib$ kompleksilukuja.

Tällöin lukujen w ja z summa on

$$\begin{aligned} w + z &= (a + ib) + (x + iy) \\ &= (a + x) + i(b + y), \end{aligned}$$

erotus

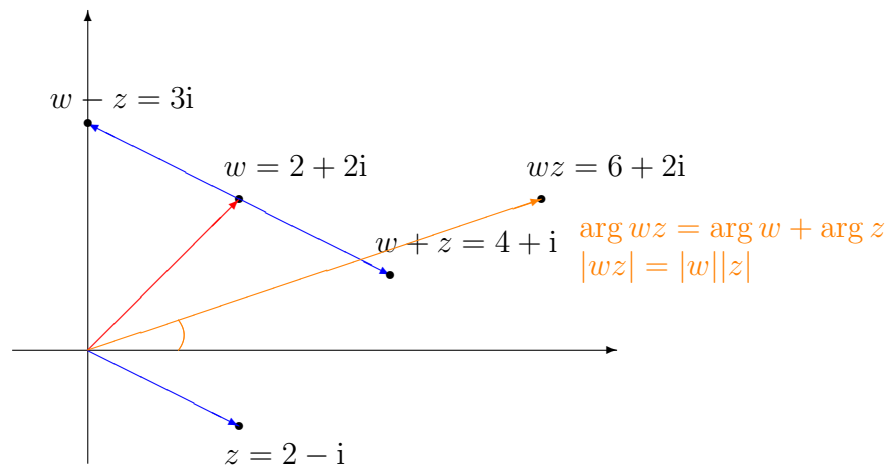
$$\begin{aligned} w - z &= (a + ib) - (x + iy) \\ &= (a - x) + i(b - y), \end{aligned}$$

ja tulo

$$\begin{aligned} wz &= (a + ib)(x + iy) \\ &= ax + aiy + ibx + ibiy \quad (\text{Muista: } i^2 = -1!) \\ &= (ax - by) + i(ay + bx). \end{aligned}$$

Esimerkki 1.11 (lasketaan luennolla)

Olkoot $w = 2 + 2i$ ja $z = 2 - i$. Laske $w + z$, $w - z$ ja wz . Tarkista piirtämällä kuvat.



Määritelmä 1.12

Luvun $z = z_1 + i z_2 \in \mathbb{C}$ *reaaliosa* on $\operatorname{Re}(z) := z_1 \in \mathbb{R}$ ja *imaginaariosa* on $\operatorname{Im}(z) := z_2 \in \mathbb{R}$.

Luku $\bar{z} = z_1 - i z_2 = \operatorname{Re}(z) - i \operatorname{Im}(z)$ on luvun z *kompleksikonjugaatti*.

Määritelmä 1.13 (Itseisarvo ja argumentti)

Jos $z = z_1 + iz_2 \in \mathbb{C}$, olkoon

$$|z| := \sqrt{z_1^2 + z_2^2} \quad \text{ja} \quad \arg(z) := \arg((z_1, z_2)).$$

Esimerkki 1.14

Nyt $|\bar{z}| = |z|$ ja $\arg(\bar{z}) = -\arg(z)$, joten $\bar{z}z = |z|^2$.

Tarkistus:

$$\bar{z}z = (z_1 - iz_2)(z_1 + iz_2) = (z_1^2 + z_2^2) + i(z_1z_2 - z_2z_1) = |z|^2.$$

Huom. Jos $0 \neq z \in \mathbb{C}$, niin

$$\begin{aligned} z &= |z| \left(\frac{z_1}{|z|} + i \frac{z_2}{|z|} \right) \\ &= |z| (\cos(\varphi) + i \sin(\varphi)) \end{aligned}$$

missä $\varphi = \arg(z)$.

Käänteisluku:

Kompleksiluvun $z = x + iy$ käänteisluvulle saadaan laskusääntöjen avulla muoto

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{x - iy}{x^2 + y^2} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2},$$

jos $z \neq 0$. Yleisemmin kompleksiluku w/z voidaan sieventää muotoon $a + ib$ laventamalla se nimittäjän liittoluvulla $\bar{z}$:

$$\frac{w}{z} := w \frac{1}{z} = w\bar{z}/|z|^2.$$

Siis $|w/z| = |w|/|z|$ ja $\arg(w/z) = \arg(w) - \arg(z)$.

Esimerkki 1.15

$$\frac{2 + 3i}{4 - 5i} = \frac{(2 + 3i)(4 + 5i)}{(4 - 5i)(4 + 5i)} = \frac{-7 + 22i}{41} = -\frac{7}{41} + \frac{22}{41}i.$$

Esimerkki 1.16

Jos $n \in \mathbb{Z}$, niin $|z^n| = |z|^n$ ja $\arg(z^n) = n \arg(z)$.

Tapauksessa $z = \cos(\varphi) + i \sin(\varphi)$ saadaan **de Moivren kaava**:

$$(\cos(\varphi) + i \sin(\varphi))^n = \cos(n\varphi) + i \sin(n\varphi).$$

Tekniikassa kompleksiluvut kirjoitetaan usein ns. polaarimuodossa

$$re^{i\varphi}.$$

Tämä tarkoittaa tason pistettä, jonka napakoordinaatit ovat (r, φ) . $e^{i\varphi}$ on yksinkertaisesti lyhyempi merkitätapa kompleksiluvulle $\cos \varphi + i \sin \varphi$. Siis

$$\begin{aligned} re^{i\varphi} &= r(\cos \varphi + i \sin \varphi) \\ &= r \cos \varphi + ir \sin \varphi \\ &= x + iy. \end{aligned}$$

Fakta $e^{i\varphi} = \cos \varphi + i \sin \varphi$ voidaan osoittaa esim. sarjakehitelmien avulla, tämä tehdään kurssilla Differentiaali- ja integraalilaskenta 1.

Esimerkki 1.17

Mitkä kompleksiluvut toteuttavat yhtälön $z^3 = 1$?

Vastaus: Ratkaise kirjoittamalla luvut polaarimuodossa $z = re^{i\varphi}$, $1 = 1e^{in2\pi}$, $n \in \mathbb{Z}$. Tällöin yhtälö saa muodon

$$(re^{i\varphi})^3 = r^3 e^{i3\varphi} = 1e^{in2\pi},$$

josta $r^3 = 1$ ja $3\varphi = n2\pi$. Yhtälön toteuttavat siis pisteet, joille $r = 1$ ja $\varphi = n\frac{2}{3}\pi$, $n \in \mathbb{Z}$, eli $z = 1$, $z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ ja $z = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$. Nämä kolme pistettä sijaitsevat tasavälein yksikköympyrällä.

Seuraava tulos on seurauksiltaan järjestyttävä (vaikkei ehkä heti uskoisi). Sen mukaan ei-vakiolla polynomilla on nollakohta:

Lause 1.18 (Algebran peruslause)

Olkoon $p : \mathbb{C} \rightarrow \mathbb{C}$ polynomi eli

$$p(z) = \sum_{k=0}^n a_k z^k,$$

missä $a_k \in \mathbb{C}$ ja $a_n \neq 0$.

Jos $n \geq 1$, niin $p(w) = 0$ jollakin $w \in \mathbb{C}$.

Seuraus: Jos polynomi p on kuten edellä, niin

$$p(z) = a_n(z - r_1)(z - r_2) \cdots (z - r_n),$$

missä $r_1, r_2, \dots, r_n \in \mathbb{C}$.

Esimerkki 1.19

$$p(z) = z^2 + 1 = (z - i)(z + i)$$

eli polynomin p juuret ovat $-i, i \in \mathbb{C}$ (eivät reaalisia!).

Esimerkki 1.20

Etsi polynomi, jolla on sekä reaalisia että imaginäärisiä juuria.

Ratkaisu: $p(z) = z^3 - 1$. Aikaisemmin huomasimme, että $z^3 = 1$ jos ja vain jos $z = 1$, $z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ tai $z = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$ eli $p(z)$ on sekä reaalisia että kompleksisia juuria. Jos haluaa vielä puhtaasti imaginäärisiä juuria kompleksisten sijaan ja lisäksi reaalisia, niin $p(z) = z^4 - 1$ juuret ovat $1, i, -1, -i$.

Esimerkki 1.21 (lasketaan luennolla)

Etsi polynomin $p(z) = z^2 + (1 - 5i)z - (4 + 4i)$ juuret.

Ratkaisu: Toisen asteen yhtälön ratkaisukaavalla saadaan

$$\begin{aligned} z &= \frac{-(1 - 5i) \pm \sqrt{(1 - 5i)^2 - 4(-4(1 + i))}}{2} \\ &= \frac{5i - 1 \pm \sqrt{1 - 10i - 25 + 16i + 16}}{2} \\ &= \frac{5i - 1 \pm \sqrt{-8 + 6i}}{2} \end{aligned}$$

Neliöjuuri voidaan laskea seuraavasti: $\sqrt{-8 + 6i} = x + iy$ korotetaan puolittain neliöön, jolloin saadaan $-8 + 6i = x^2 + 2ixy - y^2$. Ratkaistaan yhtälöryhmä

$$\begin{cases} x^2 - y^2 = -8 \\ 2xy = 6 \end{cases}$$

(taululla) ja saadaan ratkaisuiksi $x = 1, y = 3$ ja $x = -1, y = -3$, eli $\sqrt{-8 + 6i} = \pm(1 + 3i)$.

Näin ollen polynomin nollakohdat ovat $z = \frac{5i-1-(1+3i)}{2} = -1 + i$ ja $z = \frac{5i-1+(1+3i)}{2} = 4i$.

Huom. Neliöjuuri voitaisiin laskea myös samalla idealla kuin kompleksiluvun potenssi aiemmin, eli muuntamalla $-8+6i$ polaarimuotoon, ottamalla sitten neliöjuuri ($\sqrt{re^{i\varphi}} = r^{1/2}e^{i\varphi/2}$) ja muuntamalla lopuksi takaisin kompleksimuotoon.

Luku 2

Lineaarinen yhtälöryhmä matriiseilla

2.1 Matriisit ja yhtälöryhmä matriisimuodossa

Matriisi on taulukko reaalilukuja muodossa, esimerkiksi,

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & -1 & 5 \\ 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 6 & 2 \\ 5 & 1 \\ 7 & 6 \end{pmatrix}$$

Monia suuria tietokantoja ja käytännön ongelmia kuten piirilaskuja ja fysiikan ongelmia voi esittää matriisien avulla. Erityisesti piirilaskuissa usein käytetään lineaarisia yhtälöryhmiä ja niiden ratkaiseminen voidaan tulkita matriisien avulla.

Tarkastellaan esimerkkinä lineaarista yhtälöparia

$$\begin{cases} 2x_1 - x_2 = 1 \\ x_1 + x_2 = 5. \end{cases}$$

Matriisimuodossa tämä kirjoitetaan $A\mathbf{x} = \mathbf{b}$, missä

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 5 \end{pmatrix},$$

eli

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}.$$

Tulkinta 1: kumpikin yhtälöparin yhtälö kuvaa suoraa tasossa $\mathbb{R}^2$, ja mahdollinen ratkaisu $\mathbf{x} = (x_1, x_2)$ on suorien leikkauspiste.

Tulkinta 2: matriisi A määrittelee *kuvauksen* (= funktion) $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\mathbf{x} \mapsto A\mathbf{x}$. Halutaan löytää piste $\mathbf{x} \in \mathbb{R}^2$, jolle $A\mathbf{x} = (1, 5)$.

Yleisesti: Lineaarinen yhtälöryhmä $A\mathbf{x} = \mathbf{b}$, missä annettuina ovat $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ ja

$\mathbf{b} = (b_1, \dots, b_m) \in \mathbb{R}^m$, ja halutaan ratkaista $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}$$

Tulkinta 1: kukin rivi $a_{k1}x_1 + \dots + a_{kn}x_n = b_k$ ($1 \leq k \leq m$) on yhtälö *hypertasolle* avaruudessa $\mathbb{R}^n$. (Suora, kun $n = 2$; taso, kun $n = 3$.) Mahdollinen ratkaisu $\mathbf{x} \in \mathbb{R}^n$ on kaikille hypertasoille (m kpl) yhteinen piste.

Tulkinta 2: Matriisi A määrittelee kuvauksen $\mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \mapsto A\mathbf{x}$. Etsitään pistettä $\mathbf{x} \in \mathbb{R}^n$, joka kuvautuu pisteeksi $\mathbf{b} \in \mathbb{R}^m$.

Esimerkki 2.1 (lasketaan luennolla)

Olkoon $A = \begin{pmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{pmatrix}$. Määritellään kuvaus $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ säännöllä $T(\mathbf{x}) = A\mathbf{x}$, ts.

$$T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{pmatrix}.$$

- a) Etsi pisteen $\mathbf{u} = (2, -1)$ kuva $T(\mathbf{u})$.
- b) Etsi $\mathbf{x} \in \mathbb{R}^2$ siten, että $T(\mathbf{x}) = (3, 2, -5)$.
- c) Löytyykö b-kohdassa useampia ratkaisuja?
- d) Löytyykö pistettä $\mathbf{x} \in \mathbb{R}^2$ siten, että $T(\mathbf{x}) = (3, 2, 5)$?

Tarkastellaan hieman laajempaa esimerkkiä. On annettu kolme $\mathbb{R}^3$:n vektoria:

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 3 \\ 8 \\ 11 \end{pmatrix}.$$

Etsitään vektorin $\mathbf{b} = \begin{pmatrix} 1 \\ 6 \\ 7 \end{pmatrix}$ esitys vektoreiden $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ avulla.

Tarkasteltavana on siis yhtälö

$$\mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \mathbf{a}_3x_3 = \mathbf{b},$$

missä $x_1, x_2, x_3 \in \mathbb{R}$ ovat tuntemattomia, jotka on tarkoitus ratkaista, jos mahdollista.

Kirjoitetaan tämä yhtälöryhmäksi:

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ -2x_1 + 4x_2 + 8x_3 = 6 \\ 3x_1 + 6x_2 + 11x_3 = 7 \end{cases}$$

Etsitään siis toisaalta kolmen tason leikkauspisteitä.

Sama matriisimuodossa $A\mathbf{x} = \mathbf{b}$:

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 \\ -2 & 4 & 8 \\ 3 & 6 & 11 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} 1 \\ 6 \\ 7 \end{pmatrix}}_{\mathbf{b}}$$

Siis:

$$\begin{aligned} & \begin{pmatrix} 1 & 2 & 3 \\ -2 & 4 & 8 \\ 3 & 6 & 11 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \\ 7 \end{pmatrix} \\ \Leftrightarrow & \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} x_1 + \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} x_2 + \begin{pmatrix} 3 \\ 8 \\ 11 \end{pmatrix} x_3 = \begin{pmatrix} 1 \\ 6 \\ 7 \end{pmatrix} \\ \Leftrightarrow & \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ -2x_1 + 4x_2 + 8x_3 = 6 \\ 3x_1 + 6x_2 + 11x_3 = 7 \end{cases} \end{aligned}$$

Mitä tiedämme ratkaisujen lukumäärästä?

Ratkaisuja voi olla

1. 0 kpl: Tasot eivät leikkaa, eli A ei kuvaa mitään vektoria b :lle.
2. 1 kpl: Tasot leikkaavat yhdessä pisteessä, eli $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ on kanta.
3. ∞ kpl: Tasot leikkaavat pitkin suoraa/tasoa, eli ” A :n ydin on ei-triviaali”.

Hyödyllinen käsite on A :n kuva-avaruus $\mathcal{R}(A)$, jonka alkiot ovat kaikki vektoreiden $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ lineaarikombinaatiot. Jos ratkaisua ei ole olemassa, tulkitaan, että $b \notin \mathcal{R}(A)$.

2.2 Gaussin eliminaatio

Perinteisesti yhtälöryhmät ratkaistaan lisäämällä ja vähentämällä yhtälöitä toisistaan, jollakin kertoimilla painotettuina. Matriisimuodossa vastaavat operaatiot voidaan tehdä yksinkertaisemmin merkinnöin. Gaussin eliminaatiomenetelmässä lineaarinen matriisiyhtälö $A\mathbf{x} = \mathbf{b}$ kirjoitetaan liittomatriisina

$$[A \mid b] \quad \text{eli} \quad \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right],$$

jota muokataan *rivioperaatioin*:

1. lisäämällä (painotettu) rivi toiseen riviin
— vastaa (painotetun) yhtälön lisäämistä toiseen
2. vaihtamalla kahden rivin paikkaa keskenään
— vastaa yhtälöiden paikan vaihtoa
3. kertomalla yksittäinen rivi vakiolla $c \neq 0$
— vastaa yhden yhtälön kertomista vakiolla $c \neq 0$

Jos lineaarisesta yhtälöstä $A\mathbf{x} = \mathbf{b}$ saadaan rivioperaatioin $C\mathbf{x} = \mathbf{d}$, merkitään

$$[A \mid b] \quad \sim \quad [C \mid d].$$

Tässä suoraan esimerkki miten algoritmi toimii:

$$\begin{aligned} \begin{cases} x_1 + 2x_2 + 3x_3 & = & 1 \\ -2x_1 + 4x_2 + 8x_3 & = & 6 \\ 3x_1 + 6x_2 + 11x_3 & = & 7 \end{cases} & \quad \text{eli} \quad \begin{bmatrix} 1 & 2 & 3 & | & 1 \\ -2 & 4 & 8 & | & 6 \\ 3 & 6 & 11 & | & 7 \end{bmatrix} \begin{array}{l} \left[\begin{array}{c} \xrightarrow{+} \\ \xleftarrow{+} \end{array} \right]_{-3}^2 \\ \left[\begin{array}{c} \xleftarrow{+} \\ \xrightarrow{+} \end{array} \right]_{-3}^+ \end{array} \\ \\ \begin{cases} x_1 + 2x_2 + 3x_3 & = & 1 \\ 8x_2 + 14x_3 & = & 8 \\ 2x_3 & = & 4 \end{cases} & \quad \text{eli} \quad \begin{bmatrix} 1 & 2 & 3 & | & 1 \\ 0 & 8 & 14 & | & 8 \\ 0 & 0 & 2 & | & 4 \end{bmatrix} \begin{array}{l} \left[\begin{array}{c} \xleftarrow{+} \\ \xrightarrow{+} \end{array} \right]_{-14}^+ \\ \left[\begin{array}{c} \xleftarrow{+} \\ \xrightarrow{+} \end{array} \right]_{-3}^+ \end{array} \\ \\ \begin{cases} x_1 + 2x_2 & = & -5 \\ 8x_2 & = & -20 \\ x_3 & = & 2 \end{cases} & \quad \text{eli} \quad \begin{bmatrix} 1 & 2 & 0 & | & -5 \\ 0 & 8 & 0 & | & -20 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \begin{array}{l} \left[\begin{array}{c} \xleftarrow{+} \\ \xrightarrow{+} \end{array} \right]_{-2}^+ \\ \left[\begin{array}{c} \xleftarrow{+} \\ \xrightarrow{+} \end{array} \right]_{-2}^+ \end{array} \\ \\ \begin{cases} x_1 & = & 0 \\ x_2 & = & -5/2 \\ x_3 & = & 2 \end{cases} & \quad \text{eli} \quad \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & -5/2 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \end{aligned}$$

Ratkaisu on siis $\mathbf{x} = (x_1, x_2, x_3) = (0, -5/2, 2)$.
(Tarkista sijoittamalla!)

Tämä piste on alkuperäisten tasojen ainoa leikkauspiste. Se on myös piste/vektori, jonka matriisi A kuvaa pisteeksi/vektoriksi $\mathbf{b}$. Toisaalta, nämä kertoimet ovat vektorin $\mathbf{b}$ koordinaatit, kun se ilmoitetaan kannassa $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$, eli $0 \cdot \mathbf{a}_1 - \frac{5}{2}\mathbf{a}_2 + 2\mathbf{a}_3 = \mathbf{b}$.

Esimerkki 2.2 (lasketaan luennolla)

Etsi Gaussin eliminaatiomenetelmällä yhtälöryhmän

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases}$$

ratkaisu.

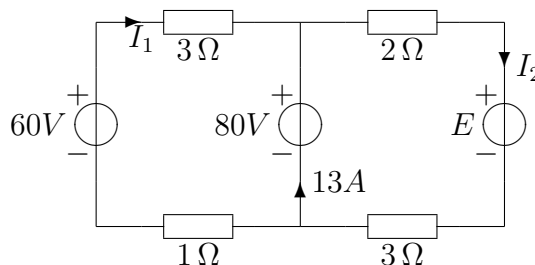
Vastaus: Gaussin eliminaatiomenetelmällä löydetään

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 29 \\ 16 \\ 3 \end{pmatrix}.$$

Katsotaan sitten käytännön esimerkkiä Gaussin eliminaatiosta, joka tulee virtapiirilaskuista:

Esimerkki 2.3 (Virtapiiri)

Virtapiiri on diagrammi, jossa vastukset ja virtojen suunnat on piirretty kuten kuvassa:



Virtapiireihin liittyvät sähködynamiikan Kirchhoffin virta- ja jännitelait:

- Virtapiirissä tiettyyn pisteeseen tulevien ja siitä lähtevien virtojen summa on sama
- Potentiaalierojen summan virtapiirin ympäri täytyy olla nolla

Etsi näiden lakien avulla virrat I_1 , I_2 ja jännite E . Vastuksen aiheuttama potentiaalin muutos on $U = RI$.

Ratkaisu: Kirchhoffin virtalain mukaan virtapiirissä tiettyyn pisteeseen tulevien ja siitä lähtevien virtojen summa on sama, joten piirin yläreunan keskellä olevassa risteyksessä täytyy

pätee $I_1 + 13 = I_2$ (A). Kirchhoffin jännitelain mukaan potentiaalierojen summan virtapiirin ympäri täytyy olla nolla, joten vasemman puoleisesta piiristä saadaan $60 = 3I_1 + 80 + I_1$ (V) ja oikeasta $E = -2I_2 - 3I_2 + 80$ (V), kun muistetaan, että vastuksen aiheuttama potentiaalimuutos on $U = RI$. Saadaan siis yhtälöryhmä

$$\begin{cases} I_1 - I_2 = -13 \\ 4I_1 = -20 \\ 5I_2 + E = 80 \end{cases}.$$

Matriisimuodossa

$$\begin{pmatrix} 1 & -1 & 0 \\ 4 & 0 & 0 \\ 0 & 5 & 1 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \\ E \end{pmatrix} = \begin{pmatrix} -13 \\ -20 \\ 80 \end{pmatrix}.$$

Gaussin eliminaatioaskeleilla (tarkista!) tämä saadaan muotoon

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \\ E \end{pmatrix} = \begin{pmatrix} -5 \\ 8 \\ 40 \end{pmatrix}$$

(Huomaa, että kahden ylimmän rivin järjestystä on vaihdettu!). Vastaus on siis $I_1 = -5A$, $I_2 = 8A$ ja $E = 40V$. Olisikin näemmä kannattanut valita virran I_1 suunta toisin päin.

Siirrytään nyt vaikeampaan esimerkkiin, jossa on ääretön määrä ratkaisuja. Tämä on tärkeä ja usein vaikea esimerkki, joka kannattaa opiskella huolella koska siinä tulee vapaita muuttujia. Olemme jakaneet esimerkin kolmeen osaan ratkaisussa, jotta on helpompi oppia pääaskeleet.

Esimerkki 2.4 (Tärkeä!)

Etsi yhtälöryhmän kaikki ratkaisut, kun

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 1 \\ 2x_1 + 4x_2 + 8x_3 + 10x_4 = 6 \\ 3x_1 + 6x_2 + 11x_3 + 14x_4 = 7 \end{cases} \iff A\mathbf{x} = \mathbf{b}$$

Ratkaisu: Tämä on

(1) Kirjoitetaan yhtälö matriisimuotoon $A\mathbf{x} = \mathbf{b}$, eli

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 8 & 10 \\ 3 & 6 & 11 & 14 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \\ 7 \end{pmatrix}.$$

(2) Ennen kuin sijoitamme liittomatriisiin oikealle puolelle vektorin $\mathbf{b} = \begin{pmatrix} 1 \\ 6 \\ 7 \end{pmatrix}$, suoritetaan

eliminaatioaskeleet yleisellä $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$.

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & b_1 \\ 2 & 4 & 8 & 10 & b_2 \\ 3 & 6 & 11 & 14 & b_3 \end{array} \right] \begin{array}{l} \left[\begin{array}{l} \leftarrow -2 \\ \leftarrow + \end{array} \right]^{-3} \\ \leftarrow + \end{array}$$

$$\sim \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 2 & 2 & b_3 - 3b_1 \end{array} \right] \begin{array}{l} \left[\begin{array}{l} \leftarrow -1 \\ \leftarrow + \end{array} \right] \end{array}$$

$$\sim \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - b_2 - b_1 \end{array} \right]$$

Jotta viimeiselle riville ei syntyisi ristiriitaa, on pädetävä $b_3 - b_2 - b_1 = 0$. Tämä on *konsistenssiehto*. Annetulla vektorilla $7 - 6 - 1 = 0$, joten ristiriitaa ei synny.

(2) Palataan sitten annettuun vektoriin b , jolloin saadaan

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 1 \\ 0 & 0 & 2 & 2 & 6 - 2 \\ 0 & 0 & 0 & 0 & 7 - 6 - 1 \end{array} \right] \begin{array}{l} \left[\begin{array}{l} \leftarrow + \\ | :2 \end{array} \right]^{-3} \end{array}$$

$$\sim \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & -5 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Matriisi A on nyt saatettu *redusoituun porrasmuotoon*. Tämä tarkoittaa muotoa, jossa jokaisen rivin ensimmäinen nollasta poikkeava alkio on 1 ja alemmalla rivillä on alussa nollia aina useampi kuin ylemmällä.

(3) Jaetaan muuttujat

a) kiinnitetyiksi (x_1, x_3)

b) vapaiksi (x_2, x_4)

Miksi nämä nimet?

Vapaat voi korvata parametreilla ja ratkaista kiinnitetyt niiden avulla.

Olkoon $x_2 = \sigma$, $x_4 = \tau$, $\sigma, \tau \in \mathbb{R}$. Ratkaistavana on siis

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \sigma \\ x_3 \\ \tau \end{pmatrix} = \begin{pmatrix} -5 \\ 2 \\ 0 \end{pmatrix}.$$

Helpoiten loppu onnistuu kirjoittamalla ongelma takaisin yhtälöryhmäksi

$$\begin{cases} x_1 + 2\sigma + \tau = -5 \\ x_3 + \tau = 2 \end{cases}.$$

Tästä saadaan ratkaistua kiinnitetyt muuttujat x_1 ja x_3 vapaiden avulla:

$$\begin{cases} x_1 = -5 - 2\sigma - \tau \\ x_3 = 2 - \tau \end{cases}.$$

Kun lisäksi muistetaan, että $x_2 = \sigma$ ja $x_4 = \tau$ ovat mielivaltaisia reaalilukuja, nähdään, että yhtälöt ratkeavat millä tahansa lukunelikolla x_1, x_2, x_3, x_4 , joka on muotoa

$$\begin{cases} x_1 = -5 - 2\sigma - \tau \\ x_2 = \sigma \\ x_3 = 2 - \tau \\ x_4 = \tau \end{cases},$$

missä $\sigma, \tau \in \mathbb{R}$. Toisin sanoen, kaikki muotoa

$$\mathbf{x} = \begin{pmatrix} -5 \\ 0 \\ 2 \\ 0 \end{pmatrix} + \sigma \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \tau \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \quad \sigma, \tau \in \mathbb{R},$$

olevat vektorit toteuttavat siis alkuperäisen yhtälön $A\mathbf{x} = \mathbf{b}$, eli ratkaisuita on ääretön määrä.

Huom. Vapaiden muuttujien kerroinvektorit $\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ ja $\begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$ ratkaisevat yhtälön $A\mathbf{x} = 0$,

eli samaa matriisia vastaavan homogeenisen yhtälön. Myös kaikki niiden lineaarikombinaatiot ratkaisevat homogeenisen yhtälön.

Sanotaankin, että matriisin A *ydin* on yhtälöryhmän $A\mathbf{x} = 0$ ratkaisuiden joukko, eli tässä tapauksessa

$$\mathcal{N}(A) = \text{span} \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\}; \quad \dim \mathcal{N}(A) = 2.$$

Tässä span tarkoittaa annettujen vektorien lineaarikombinaatioiden muodostamaa joukkoa.

Esimerkki 2.5 (lasketaan luennolla)

Etsi yhtälöryhmän

$$\begin{cases} -x_1 + x_2 - x_3 + 3x_4 = 0 \\ 3x_1 + x_2 - x_3 - x_4 = 0 \\ 2x_1 - x_2 - 2x_3 - x_4 = 0 \end{cases}$$

kaikki ratkaisut.

Vastaus: $\mathbf{x} = \alpha \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \alpha \in \mathbb{R}.$

Huomasimme aikaisemmissa esimerkeissä, että matriisit voidaan saattaa Gaussin eliminaatiolla tiettyyn muotoon, josta voi löytää ratkaisuavaruudet. Tämä ajatus voidaan tiivistää seuraavassa teoreettisessa lauseessa:

Lause 2.6

Lineaarinen yhtälöryhmä $A\mathbf{x} = \mathbf{b}$, missä A on $m \times n$ -matriisi, voidaan aina saattaa muotoon

$$\left(\begin{array}{cc|c} I & F & c_1 \\ 0 & 0 & c_2 \end{array} \right),$$

missä I on $r \times r$ -identiteettimatriisi, F on $r \times (n - r)$ -matriisi, c_1 on r -vektori ja c_2 on $(m - r)$ -vektori.

Ratkaisuiden lukumäärälle saadaan ehdot: Jos

$$\begin{cases} r < m \text{ ja } c_2 \neq 0 & \text{lukumäärä on } 0 \\ (r = m \text{ tai } c_2 = 0) \text{ ja } r = n & \text{lukumäärä on } 1 \\ (r = m \text{ tai } c_2 = 0) \text{ ja } r < n & \text{lukumäärä on } \infty \end{cases}$$

Jatketaan vielä muutamalla käytännön esimerkillä.

Esimerkki 2.7

Eräs yksinkertainen talous koostuu hiili-, sähkö- ja teräsektoreista. Sähkösektorin tuotannosta myydään 40% hiilisektorin käyttöön, 50% teräsektorin käyttöön ja loput jää omaan käyttöön. Hiilisektorin tuotannosta sähköteollisuus ostaa 60% ja terästeollisuus 40%. Teräsektorin tuotannosta puolestaan 60% myydään hiilisektorin käyttöön, 20% sähkösektorille ja loput omaan käyttöön.

Merkitään sähkösektorin vuosituotannon arvoa p_s , hiilisektorin p_h ja teräsektorin p_t . Etsi tasapainotila, jossa kunkin sektorin tulot ja menot vastaavat toisiaan.

Ratkaisu: Tasapainotilassa hiilisektorin vuosituoannon arvo p_h on yhtä suuri kuin sen menot. Menot koostuvat siitä, että ostetaan 40% sähkösektorin tuotannosta ja 60% terässektorin tuotannosta. Siis:

$$p_h = 0.40p_s + 0.60p_t.$$

Vastaavasti sähkö- ja terässektoreille: $p_s = 0.60p_h + 0.20p_t + 0.10p_s$ ja $p_t = 0.50p_s + 0.40p_h + 0.20p_t$. (Huomaa, että näillä sektoreilla osa tuotannosta menee omaan käyttöön!) Saadaan siis yhtälöryhmä:

$$\begin{cases} p_h - 0.40p_s - 0.60p_t = 0 \\ -0.60p_h + 0.90p_s - 0.20p_t = 0 \\ -0.40p_h - 0.50p_s + 0.80p_t = 0 \end{cases}$$

Kirjoitetaan tämä matriisimuodossa:

$$\begin{pmatrix} 1 & -0.40 & -0.60 \\ -0.60 & 0.90 & -0.20 \\ -0.40 & -0.50 & 0.80 \end{pmatrix} \begin{pmatrix} p_h \\ p_s \\ p_t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Gaussin eliminaatiolla saadaan pyöristettynä kahden luvun tarkkuudelle (tarkista!):

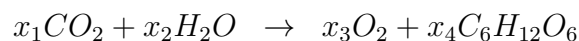
$$\left[\begin{array}{ccc|c} 1 & -0.40 & -0.60 & 0 \\ -0.60 & 0.90 & -0.20 & 0 \\ -0.40 & -0.50 & 0.80 & 0 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -0.94 & 0 \\ 0 & 1 & -0.85 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right],$$

joten yleinen ratkaisu on

$$\begin{pmatrix} p_h \\ p_s \\ p_t \end{pmatrix} \approx p_t \begin{pmatrix} 0.94 \\ 0.85 \\ 1 \end{pmatrix}, \quad p_t \in \mathbb{R}.$$

Esimerkki 2.8 (lasketaan luennolla)

Fotosynteesissä kasvi muuttaa auringonvalosta saamallaan energialla hiilidioksidia CO_2 ja vettä H_2O hapeksi O_2 ja glukoosiksi $C_6H_{12}O_6$. Reaktion kemiallinen yhtälö on siis



Etsi kertoimet x_1, x_2, x_3, x_4 .

Vastaus: Hiili-, vety- ja happiatomien lukumäärien täytyy pysyä vakioina, joten yhtälön kummallakin puolella niitä kutakin on sama määrä. Tästä saamme yhtälöryhmän, joka ratkaistaan esim. Gaussin eliminaatiomenetelmällä. Vastaukseksi saadaan $x_1 = x_2 = x_3 = 6\tau$, $x_4 = \tau$, $\tau \in \mathbb{R}$.

Luku 3

Matriisialgebra

3.1 Lineaarikuvaukset

Lineaariset yhtälöt ovat “vektoreille luonnollisia” yhtälöitä, joita ratkotaan mm. sähkömagnetiikassa, mekaniikassa, tietotekniikassa, taloustieteissä, ekologiassa jne.

Määritelmä 3.1

Funktio $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ on *lineaarinen* (eli lineaarikuvaus), jos

- (i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ kaikille $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$
- (ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ kaikille $c \in \mathbb{R}$ ja kaikille $\mathbf{u} \in \mathbb{R}^n$.

Huomioita

- Ehdot (i) ja (ii) voidaan lausua myös yhdessä: $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$ kaikille $c, d \in \mathbb{R}$ ja kaikille $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.
Edelleen (induktiolla) $T(c_1\mathbf{u}_1 + \dots + c_k\mathbf{u}_k) = c_1T(\mathbf{u}_1) + \dots + c_kT(\mathbf{u}_k)$ kaikille $k \in \mathbb{N}$, kaikille $c_1, \dots, c_k \in \mathbb{R}$ ja kaikille $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{R}^n$. *Lineaarikuvaus säilyttää lineaarikombinaatiot.*
- Mielivaltaiselle $\mathbf{u} \in \mathbb{R}^n$ pätee $T(\mathbf{0}) = T(0\mathbf{u}) = 0T(\mathbf{u}) = \mathbf{0}$. *Lineaarikuvaus pitää origon paikallaan.*
- Matriisin $A \in \mathbb{R}^{m \times n}$ määräämä kuvaus $T(\mathbf{u}) = A\mathbf{u}$ on lineaarinen (määritelmä toteutuu).
- Lineaarikuvaus $\mathbb{R}^n \rightarrow \mathbb{R}^m$ voidaan aina esittää matriisikuvauksena (nähdään esimerkin jälkeen).

Esimerkki 3.2 (lasketaan luennolla)

Määritellään lineaarikuvaus $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $T(\mathbf{x}) = A\mathbf{x}$, missä

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

Määritä pisteiden $\mathbf{u} = (4, 1)$, $\mathbf{v} = (2, 3)$ ja $\mathbf{u} + \mathbf{v} = (6, 4)$ kuvapistet. Totea, että $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$.

Standardikantavektorit avaruudessa $\mathbb{R}^n$ ovat

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, \dots, 0, 0), \\ \mathbf{e}_2 &= (0, 1, \dots, 0, 0), \\ &\vdots \\ \mathbf{e}_n &= (0, 0, \dots, 0, 1). \end{aligned}$$

Identtinen matriisi I_n on matriisi, jonka j :s sarake (yhtä lailla rivi) on j :s standardikantavektori:

$$I_n = (\mathbf{e}_1 \cdots \mathbf{e}_n) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Lause 3.3

Jokainen lineaarikuvaus $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ voidaan esittää matriisikuvauksena $T(\mathbf{x}) = A\mathbf{x}$, missä matriisin A sarakkeet ovat kantavektorien $\mathbf{e}_j$ kuvat:

$$A = (T(\mathbf{e}_1) \cdots T(\mathbf{e}_n)).$$

Esimerkki 3.4

Olkoon $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ lineaarikuvaus, jolle $T(\mathbf{e}_1) = (5, -7, 2)$ ja $T(\mathbf{e}_2) = (-3, 8, 0)$. Etsitään matriisi $A \in \mathbb{R}^{3 \times 2}$ siten, että $T(\mathbf{x}) = A\mathbf{x}$ kaikille $\mathbf{x} \in \mathbb{R}^2$.

Esimerkki 3.5 (jatkuu)

Koska mielivaltainen $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ voidaan esittää muodossa $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ ja koska T on lineaarinen, niin

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) \\ &= x_1(5, -7, 2) + x_2(-3, 8, 0) = (5x_1 - 3x_2, -7x_1 + 8x_2, 2x_1 + 0x_2) \end{aligned}$$

eli sarakemuodossa

$$T(\mathbf{x}) = \begin{pmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0x_2 \end{pmatrix} = \begin{pmatrix} 5 & -3 \\ -7 & 8 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A\mathbf{x}.$$

Matriisi A on lineaarikuvauksen T esitys standardikannassa.

Avaruuden $\mathbb{R}^n$ suora on joukko

$$\{\mathbf{u} + t\mathbf{v} \mid t \in \mathbb{R}\} \subset \mathbb{R}^n,$$

missä $\mathbf{u} \in \mathbb{R}^n$ on eräs suoran piste ja $\mathbf{v} \in \mathbb{R}^n$ on suoran suuntavektori.

Lause 3.6

Lineaarikuvaus kuvaa suoran suoraksi.

Todistus

[Todistus] Luentoharjoitus. □

Kaikki tason lineaarikuvaukset (eli kuvaukset $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ ovat)

- venytyksiä
- peilauksia
- kiertoja
- projektioita

tai näiden yhdisteitä.

Projektiomatriiseja

Projektio x -akselille: $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Projektio y -akselille: $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

Venytysmatriiseja

Venytys x -suunnassa kertoimella k : $\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$

Venytys y -suunnassa kertoimella k : $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

Peilausten matriiseja

Peilaus x -akselin suhteen: $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Peilaus y -akselin suhteen: $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

Peilaus suoran $y = x$ suhteen: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Peilaus suoran $y = -x$ suhteen: $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

Peilaus origon suhteen: $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

Kierto/Pyöritysmatriisi

Matriisi

$$A_\varphi = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix}$$

pyörittää xy -tason pistettä kulman φ verran origon ympäri vastapäivään. Suorittamalla samaistus $(x, y) \in \mathbb{R}^2$ ja kompleksiluvun $x + iy \in \mathbb{C}$ kanssa, kiertomatriisilla on syväallinen yhteys ekponenttifunktion kanssa:

Esimerkki 3.7 (Lasketaan luennolla)

Käytä kompleksilukujen kertolaskua ja kaavaa $e^{i\varphi} = \cos(\varphi) + i\sin(\varphi)$ todistamaan, että kaikilla $z = x + iy \in \mathbb{C}$ vektorina kompleksiluku $e^{i\varphi}z = A_\varphi(x, y)$ jos $(x, y) \in \mathbb{R}^2$ samaistetaan kompleksiluvun $z \in \mathbb{C}$ kanssa.

3.2 Matriisien laskutoimitukset

Idea: Lineaarikuvausten laskutoimitusten avulla määritellään vastaavat matriisien laskutoimitukset.

Vakiolla kertominen ja summa. Olkoon $t \in \mathbb{R}$ ja $A, B \in \mathbb{R}^{n \times m}$. Silloin $tA, A+B \in \mathbb{R}^{n \times m}$ ja määritellään

$$tA = \begin{bmatrix} tA_{11} & tA_{12} & \dots & tA_{1m} \\ tA_{21} & tA_{22} & \dots & tA_{2m} \\ \vdots & \vdots & \dots & \vdots \\ tA_{n1} & tA_{n2} & \dots & tA_{nm} \end{bmatrix},$$

$$A + B = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} & \dots & A_{1m} + B_{1m} \\ A_{21} + B_{21} & A_{22} + B_{22} & \dots & A_{2m} + B_{2m} \\ \vdots & \vdots & \dots & \vdots \\ A_{n1} + B_{n1} & A_{n2} + B_{n2} & \dots & A_{nm} + B_{nm} \end{bmatrix}$$

Esimerkki 3.8

$$(-2) \begin{bmatrix} +1 & -2 & +3 \\ -4 & +5 & -6 \end{bmatrix} = \begin{bmatrix} -2 & +4 & -6 \\ +8 & -10 & +12 \end{bmatrix},$$

$$\begin{bmatrix} 1 & -2 \\ 3 & -4 \\ 5 & -6 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 4 & 8 \\ 16 & 32 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 7 & 4 \\ 21 & 26 \end{bmatrix}.$$

Huom. Jotta $A + B$ olisi määritelty, on matriisien A ja B oltava samankokoiset!

Olkoon $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ ja $G : \mathbb{R}^n \rightarrow \mathbb{R}^p$. Yhdistetty kuvaus on mielekäs vain muodossa $G \circ F : \mathbb{R}^m \rightarrow \mathbb{R}^p$.

Jos F ja G ovat lineaarikuvauksia, niin $G \circ F$:kin on.

Olkoon B kuvauksen F matriisi ja A kuvauksen G matriisi. Tällöin $G \circ F$ on

$$(G \circ F)(x) = G(F(x)) = A(B(x)) = ABx.$$

AB on *matriisitulo*.

Määritelmä 3.9 (Matriisitulo)

Olkoot $\underbrace{A}_{p \times n} = (\alpha_{ij})$ ja $\underbrace{B}_{n \times q} = (\beta_{ij})$. Tällöin $\underbrace{C}_{p \times q} = AB = (\gamma_{ij})$, missä

$$\gamma_{ij} = \sum_{k=1}^n \alpha_{ik} \beta_{kj}$$

Muistisääntö: $\underbrace{C}_{p \times q} = \underbrace{A}_{p \times n} \underbrace{B}_{n \times q}$.

Huom! Matriisitulo ei ole vaihdannainen eli se ei kommutoi: yleisesti $AB \neq BA$!

Huom! Tulo voi olla nolla, vaikka kumpikaan matriisi ei olisi nollamatriisi!

Olkoot

$$B = \begin{bmatrix} 9 & 8 \\ 7 & 6 \\ 5 & 4 \end{bmatrix} \quad \text{ja} \quad A = \begin{bmatrix} -1 & -2 \\ -3 & -4 \end{bmatrix}.$$

Nyt $A+B$, AB ja BA eivät ole mielekkäitä (vastaavilla lineaarikuvauksilla menisivät dimensiot solmuun tällaisista yhdistelmistä). Kuitenkin voidaan laskea

$$BA = \begin{bmatrix} 9(-1) + 8(-3) & 9(-2) + 8(-4) \\ 7(-1) + 6(-3) & 7(-2) + 6(-4) \\ 5(-1) + 4(-3) & 5(-2) + 4(-4) \end{bmatrix} = \begin{bmatrix} -33 & -50 \\ -25 & -38 \\ -17 & -26 \end{bmatrix}$$

ja

$$\begin{aligned} A^2 := AA &= \begin{bmatrix} -1 & -2 \\ -3 & -4 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ -3 & -4 \end{bmatrix} \\ &= \begin{bmatrix} (-1)(-1) + (-2)(-3) & (-1)(-2) + (-2)(-4) \\ (-3)(-1) + (-4)(-3) & (-3)(-2) + (-4)(-4) \end{bmatrix} \\ &= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}. \end{aligned}$$

Esimerkki 3.10 (lasketaan luennolla)

Olkoot

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \quad \text{ja} \quad C = \begin{bmatrix} -1 & 3 \\ 2 & -2 \end{bmatrix}$$

Laske AB , AC , BC ja CB .

Vastaus: $AB = B$, $AC = C$ (vrt. $1 \cdot x = x$),

$$BC = \begin{bmatrix} 0 & 4 \\ 5 & -3 \end{bmatrix} \quad \text{ja} \quad CB = \begin{bmatrix} 1 & 8 \\ 2 & -4 \end{bmatrix}.$$

Esimerkki 3.11

Olkoon $\mathbf{x} = (-1, -2, -3) \in \mathbb{R}^3$ ja lineaarikuvauksen $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ matriisi

$$A_F = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix}.$$

Silloin

$$\begin{aligned} A_F \mathbf{x} &= \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix} \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix} \\ &= \begin{bmatrix} 2(-1) + 3(-2) + 4(-3) \\ 5(-1) + 6(-2) + 7(-3) \end{bmatrix} = \begin{bmatrix} -20 \\ -38 \end{bmatrix}, \end{aligned}$$

joten $F(-1, -2, -3) = A_F \mathbf{x} = (-20, -38)$.

Katsotaan nyt miten lineaarikuvauksen voi hajoittaa yksinkertaisempien matriisien tuloksi.

Esimerkki 3.12

Tason suorakulmion sivut ovat koordinaattiakselien suuntaisia ja sen kulmat ovat vastapäivään lueteltuina $a = (0, 0)$, $b = (2, 0)$, $c = (2, 4)$ ja $d = (0, 4)$. Suorakulmiolle suoritetaan affiini muunnos f , joka kuvaa sen peilatuksi neliöksi toiseen paikkaan tasossa. Suorakulmion kulmat kuvautuvat muunnoksessa siten, että $f(a) = (0, 0)$ ja $f(c) = (1, 1)$, sivut koordinaattiakselien suuntaiset ja kulmat ovat vastapäivään lueteltaessa $f(a)$, $f(d)$, $f(c)$ ja $f(b)$.

Etsi matriisi A jolla voit esittää lineaarikuvauksen f

$$f(\mathbf{x}) = A\mathbf{x}$$

ja selvitä mikä on pisteen $(3, 1)$ kuva muunnoksessa f .

Ratkaisu: Tästä piirrämme kuvan luennolla, joka helpottaa visualisoimaan tilannetta. Muunnoksen pienentämällä kuvio puolittamalla x -koordinaattiarvot ja jakamalla y -koordinaattiarvot luvulla 4 (matriisi M_1) ja peilaamalla suoran $x = y$ suhteen (koordinaattiarvojen vaihto keskenään, matriisi M_2). Tällöin saamme kuvapisteen järjestykseen $f(a)$, $f(d)$, $f(c)$ ja $f(b)$.

Matriisi, joka kutistaa puoleen x -suunnassa ja neljännekseen y -suunnassa, eli tekee operaation $(x, y) \mapsto (x/2, y/4)$ on

$$\underbrace{\begin{bmatrix} 1/2 & 0 \\ 0 & 1/4 \end{bmatrix}}_{M_1} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x/2 \\ y/4 \end{bmatrix}.$$

Ppeilaus $(x, y) \mapsto (y, x)$:

$$\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{M_1} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}.$$

Koko operaation suorittava matriisi saadaan tekemällä nämä peräjäälkeen:

$$M = M_2 M_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1/4 \end{bmatrix} = \begin{bmatrix} 0 & 1/4 \\ 1/2 & 0 \end{bmatrix}.$$

Piste (x, y) kuvautuu siis pisteeksi $(y/4, x/2)$ ja erityisesti pisteen $(3, 1)$ kuva on $(1/4, 3/2)$.

3.3 Transpoosi ja käänteismatriisi

Katsotaan nyt muutamaa tärkeää operaatiota matriiseille, transpoosi ja käänteismatriisi.

Transpoosi on operaatio matriiseille, joka tekee $n \times m$ -matriisista $m \times n$ -matriisin vaihtamalla rivit ja sarakkeet.

Määritelmä 3.13 (Transpoosi)

Olkoon $\underbrace{A}_{n \times m} = (\alpha_{ij})$. Tällöin matriisin A transpoosi on matriisi

$$A^T = (\gamma_{ij}), \quad \text{missä } \gamma_{ij} = \alpha_{ji}$$

Esimerkki 3.14 (lasketaan luennolla)

Mikä on matriisin

$$A = \begin{pmatrix} 3 & -4 & 2 \\ 1 & 3 & 7 \\ 5 & -1 & 1 \end{pmatrix}$$

transpoosi?

Huom. Tulomatriisille pätee $C = AB$ pätee $C^T = B^T A^T$.

Jos $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ on lineaarikuvaus samanulotteisten avaruuksien välillä, jolloin sitä vastaava matriisi A on $n \times n$ matriisi. Joillakin lineaarikuvauksilla on olemassa niin sanottu *käänteiskuvaus* $T^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ joka toteuttaa ehdon

$$T \circ T^{-1}(\mathbf{x}) = T^{-1} \circ T(\mathbf{x}) = \mathbf{x}$$

kaikilla $\mathbf{x} \in \mathbb{R}^n$. Tällöin T^{-1} on myös lineaarikuvaus, jonka vastaava matriisiä kutsutaan *käänteismatriisiksi*. Tällöin ehto $T \circ T^{-1}(\mathbf{x}) = T^{-1} \circ T(\mathbf{x}) = \mathbf{x}$ voidaan esittää tulomatriisien avulla seuraavaksi:

Määritelmä 3.15

Matriisi B on matriisin A *käänteismatriisi*, jos

$$AB = I \quad \text{ja} \quad BA = I,$$

missä I on sopivan kokoinen identiteettimatriisi.

Käänteismatriisia merkitään A^{-1} .

Käänteismatriisi ei ole aina olemassa, esim.

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

Heuristisesti syynä on se, että A lineaarikuvauksena ”projisoi” tai rutistaa koko avaruuden $\mathbb{R}^2$ x -akseliin koska se kuvaa kaikkien vektorien y koordinaatin nolllaksi! Miten siis sillä voisi olla käänteiskuvaus? Käymme tämän esimerkin tarkemmin luennolla formaalisti.

Lause 3.16

Jos käänteismatriisi on olemassa, niin se on yksikäsitteinen.

Huom.

- Jos matriisi A on kääntyvä, niin myös matriisi A^{-1} on kääntyvä, ja

$$(A^{-1})^{-1} = A.$$

- Jos A ja B ovat kääntyviä matriiseja, niin myös AB on kääntyvä, ja

$$(AB)^{-1} = B^{-1}A^{-1}.$$

- Jos A on kääntyvä, niin myös A^T on kääntyvä, ja

$$(A^T)^{-1} = (A^{-1})^T.$$

Terminologiaa:

- $\underbrace{A}_{n \times n}$ on neliömatriisi; symmetrinen, jos $A = A^T$.
- A on säännöllinen (sanotaan myös kääntyvä), jos käänteismatriisi on olemassa, muutoin singulaarinen.

- $A = \text{diag}(\alpha_{11}, \alpha_{22}, \dots, \alpha_{nn})$

$$= \begin{pmatrix} \alpha_{11} & 0 & 0 & \dots & 0 \\ 0 & \alpha_{22} & 0 & \dots & 0 \\ & & \ddots & & \\ 0 & \dots & 0 & \alpha_{(n-1)(n-1)} & 0 \\ 0 & 0 & \dots & 0 & \alpha_{nn} \end{pmatrix}$$

on lävistäjä- eli diagonaalimatriisi.

- Yläkolmiomatriisi: $\alpha_{jk} = 0$, kun $j > k$
- Alakolmiomatriisi: $\alpha_{jk} = 0$, kun $j < k$
- A on ortogonaalinen, jos $A^{-1} = A^T$.

Esimerkki 3.17 (lasketaan luennolla)

Laadi esimerkkimatriisit näistä kaikista.

3.4 Käänteismatriisin laskeminen

Olko $A \in \mathbb{R}^{n \times n}$ kääntyvä. Miten lasketaan $A^{-1} \in \mathbb{R}^{n \times n}$? Nyt $AA^{-1} = I_n$ eli jos $\mathbf{x}_j$ on matriisin A^{-1} j :s sarake, niin saadaan n kpl lineaarisia yhtälöitä

$$A\mathbf{x}_j = \mathbf{e}_j \quad (j = 1, \dots, n).$$

Ratkotaan nämä n yhtälöä yhtä aikaa Gauss-eliminaatiolla: liittomatriisi

$$[A \mid I_n] = \left[\begin{array}{cccc|cccc} A_{11} & A_{12} & \dots & A_{1n} & 1 & 0 & \dots & 0 \\ A_{21} & A_{22} & \dots & A_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} & 0 & 0 & \dots & 1 \end{array} \right]$$

muunnetaan liittomatriisiksi $[I_n \mid A^{-1}]$. Siis

$$[A \mid I_n] \sim [I_n \mid A^{-1}]$$

Esimerkki 3.18

Lasketaan matriisin $A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$ käänteismatriisi:

$$\begin{aligned} [A \mid I_2] &= \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array} \right] \sim \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & \frac{1}{2} & -\frac{5}{2} & 1 \end{array} \right] \\ &\sim \left[\begin{array}{cc|cc} 2 & 0 & 6 & -2 \\ 0 & \frac{1}{2} & -\frac{5}{2} & 1 \end{array} \right] \\ &\sim \left[\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -5 & 2 \end{array} \right] = [I_2 \mid A^{-1}] \end{aligned}$$

Siten $A^{-1} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$. (Tarkista kertolaskulla!)

Esimerkki 3.19 (Laskaritehtävä)

Laske matriisin

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{4} & 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix},$$

käänteismatriisi.

Ylemmässä esimerkissä huomaa, että jos diagonaalilta muutetaan yksi termi nolaksi, esimerkiksi siis matriisi

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix},$$

niin saamme ei kääntyvän matriisin. Tämä on työläs tarkistaa Gaussin eliminaatiolla, mutta ns. *determinantti* on paljon helpompi tapa todistaa ettei B ole kääntyvä koska sen determinantti on helposti tarkistettavissa, että se on nolla. Seuraavassa osiossa määrittelemme determinantin ja miksi näin tosiaan on.

3.5 Determinantti

Idea: Matriisin $A \in \mathbb{R}^{n \times n}$ determinantti $\det(A) \in \mathbb{R}$ ilmaisee miten paljon matriisia vastaava lineaarikuvaus “skaalaa&peilaa” avaruutta $\mathbb{R}^n$: Kuution

$$[0, 1]^n := \{x \in \mathbb{R}^n \mid \forall j : x_j \in [0, 1]\}$$

“ n -tilavuus” on 1 (1-til.=pituus, 2-til.=pinta-ala, 3-til.=tavallinen tilavuus, ...) ja A :n sarakkeiden virittämän “särmion”

$$A[0, 1]^n = \{Ax \in \mathbb{R}^n \mid x \in [0, 1]^n\}$$

n -tilavuus on $|\det(A)|$ (determinantin etumerkki kertoo peilautumisista).

Determinantin laskemiseen tarvitaan neljä lakia:

Laki 1:

$$\det(I_n) = 1.$$

Tulkinta 1: Identiteettikuvaus $x \mapsto I_n(x) = x$ ei muuta n -tilavuutta tai suuntia.

Laki 2:

$$\text{Matriisin sarake } t \text{ - kertaistuu} \Rightarrow \text{determinantti } t \text{ - kertaistuu.}$$

Tulkinta 2: “Särmion n -tilavuus” t -kertaistuu yhden särmän t -kertaistuessa.

Esimerkki 3.20

$$\det \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix} \stackrel{2}{=} (-2)(3) \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \stackrel{1}{=} -6.$$

$$\det \begin{bmatrix} 6 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix} \stackrel{2}{=} (6)(5)(4) \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \stackrel{1}{=} 120.$$

$$\det \begin{bmatrix} 0 & 1 & 2 \\ 0 & 3 & 4 \\ 0 & 5 & 6 \end{bmatrix} \stackrel{2}{=} (0) \det \begin{bmatrix} 0 & 1 & 2 \\ 0 & 3 & 4 \\ 0 & 5 & 6 \end{bmatrix} = 0.$$

Laki 3: Jos matriisissa $A \in \mathbb{R}^{n \times n}$ on (ainakin) kaksi samaa saraketta, niin

$$\det(A) = 0.$$

Tulkinta 3: Tällöin särmio $A[0, 1]^n \subset \mathbb{R}^n$ on “litistynyt” ja sen n -tilavuus on 0.

Esimerkki 3.21

$$\det \begin{bmatrix} 1 & 1 & 5 \\ 2 & 2 & 7 \\ 3 & 3 & 9 \end{bmatrix} \stackrel{3}{=} 0, \quad \det \begin{bmatrix} 1 & 5 & 1 \\ 2 & 7 & 2 \\ 3 & 9 & 3 \end{bmatrix} \stackrel{3}{=} 0, \quad \det \begin{bmatrix} 5 & 1 & 1 \\ 7 & 2 & 2 \\ 9 & 3 & 3 \end{bmatrix} \stackrel{3}{=} 0.$$

Laki 4: Olkoon A_j matriisin $A \in \mathbb{R}^{n \times n}$ j :s sarake. Kun $A_k = B_k + C_k$ jollakin $k \in \{1, \dots, n\}$, niin

$$\begin{aligned} & \det \begin{bmatrix} A_1 & \dots & A_{k-1} & B_k + C_k & A_{k+1} & \dots & A_n \end{bmatrix} \\ &= \det \begin{bmatrix} A_1 & \dots & A_{k-1} & B_k & A_{k+1} & \dots & A_n \end{bmatrix} \\ &+ \det \begin{bmatrix} A_1 & \dots & A_{k-1} & C_k & A_{k+1} & \dots & A_n \end{bmatrix}. \end{aligned}$$

Tulkinta 4: Särmiön yhden särmän summaus näkyy n -tilavuudessa summana.

Esimerkki 3.22

$$\begin{aligned} 0 & \stackrel{3}{=} \det \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \stackrel{4}{=} \det \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} + \det \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \\ & \stackrel{4}{=} \det \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} + \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \det \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \\ & \stackrel{3,1}{=} \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 1, \end{aligned}$$

$$\text{joten } \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -1.$$

Esimerkki 3.23 (lasketaan luennolla)

Laske matriisin $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ determinantti lakien §1-§4 avulla.

Ratkaisu:

$$\begin{aligned} \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} & \stackrel{4}{=} \det \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \\ & \stackrel{4}{=} \det \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} + \det \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} \\ & \stackrel{2}{=} a \det(I) + ab \det \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + bc \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + cd \det \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \\ & \stackrel{1\&3}{=} ad + 0 - bc + 0, \end{aligned}$$

$$\text{sillä äsken laskettiin } \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -1.$$

Lait 1, 2, 3, 4 ovat ristiriidattomat ja riittävät determinantin laskemiseen. Käytännössä determinantit kuitenkin lasketaan käsin

- muistamalla, että

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

- ja purkamalla matriisi 2×2 -osiin *alideterminanttisäännöllä*

$$\det(A) = (-1)^k \sum_i (-1)^i A_{ki} \det(A^{*ki}),$$

missä k on jokin rivinumero ja A^{*ki} matriisi A , josta on poistettu rivi k ja sarake i .

Determinanttia merkitään myös lyhyesti pystyviivoilla:

$$|M| := \det(M), \text{ kun } M \text{ on matriisi.}$$

Esimerkki 3.24

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= a(ei - fh) + b(fg - di) + c(dh - eg). \end{aligned}$$

Huom. Myös 3×3 -matriisin determinantille on muistisääntö:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - afh - bdi - ceg.$$

“alamäkeen positiivisina, ylämäkeen negatiivisina”.

Esimerkki 3.25 (lasketaan luennolla)

Laske matriisin $A = \begin{bmatrix} 2 & -1 & 3 \\ 4 & 1 & 2 \\ 6 & 3 & 8 \end{bmatrix}$ determinantti.

Vastaus: $\det(A) = 42$

Lause 3.26

Olkoon $A \in \mathbb{R}^{n \times n}$. Silloin

$$\exists A^{-1} \iff \det(A) \neq 0 \quad \text{ja} \quad \det(A^T) = \det(A).$$

Lause 3.27

$$\det(AB) = \det(A) \det(B)$$

Todistukset kirjassa.

Esimerkki 3.28 (lasketaan luennolla)

Millä kertoimilla $a, b, c, d \in \mathbb{R}$ koskevalla ehdolla matriisilla $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ on käänteismatriisi?
Mikä silloin on käänteismatriisi A^{-1} ?
Tarkista, että $A^{-1}A = I = AA^{-1}$.

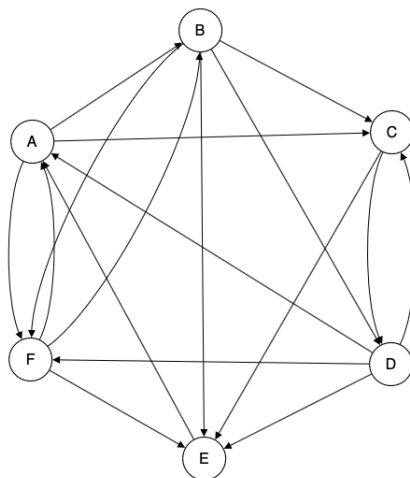
Luku 4

Ominaisarvot ja -vektorit

4.1 Motivaatio ja määritelmät

Kerromme nyt ominaisarvoteoriasta, jonka aloitamme motivoivasta esimerkistä ns. *Googlen PageRank* algoritmista. Tämä on yksinkertainen mutta tehokas Googlen perustajien Larry Pagen ja Sergey Brinin kehittämä menetelmä, jolla voidaan nopeasti luokitella internetin sivuja niiden merkittävyyden perusteella.

Alla meillä on pieni internet verkko nettisivuja, annetaan niille nimet A, B, C, D, E, F. Jokainen nuoli tarkoittaa, että sivulta on toiseen sivuun hyperlinkki, esimerkiksi sivulta A on sivuille B, C ja F linkit.



Nyt verkossa sivu on *merkittävä*, jos moni linkkaa sinne ja/tai merkittävältä sivulta on linkki sivulle. Tämän heuristiikan avulla voimme nyt formaalisti määritellä internet sivun merkittävyys reaalilukuna $x \in [0, 1]$ seuraavasti linkkien määrän ja siihen linkkaavien sivujen vaikuttavuuden avulla:

- (1) Jokaisen internet sivun merkittävyys vaikuttaa tasaisesti jokaiselle sivulle, jolle se linkittää
- (2) Internet sivun merkittävyys x on vaikuttavuuksilla painotettu summa kaikkien niiden sivujen merkittävyyksistä, jotka linkkaavat siihen.

Tällöin sääntö (1) sanoo esimerkiksi koska sivulta B on linkit sivuille C , D , E ja F , sivun *vaikutus* sivujen C , D , E ja F merkittävyyteen on jokaiseen tasan 25% eli $\frac{1}{4}$. Toisaalta sivulta C on linkki vain kahteen sivuun D ja E , joten C :n vaikutus sivujen D ja E merkittävyyteen on jokaiseen tasan 50% eli $\frac{1}{2}$. Lisäksi esimerkiksi sivun C ja F välillä ei ole lainkaan linkkiä, joten kumpikaan ei vaikuta suoraan toisensa merkittävyyteen eli vaikutus toisiinsa on 0% eli 0. Voimme siis säännön (1) avulla muodostaa 6×6 matriisin, jolla voimme mallintaa miten internet sivut *vaikuttavat* toistensa merkittävyyteen (indeksöimällä $A = 1$, $B = 2$, $C = 3$, ..., $F = 6$):

$$P = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{4} & 1 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 \end{pmatrix}$$

Miten sitten saamme tämän matriisin avulla löydettyä sivujen A , B , C , D , E , F merkittävyydet ja siten luokitella ne lopulta vaikka kun Googlaamme vastausta? Jos $x_1, x_2, x_3, x_4, x_5, x_6$ on sivujen A , B , C , D , E , F merkittävyydet, niin koska jokaisen sivun merkittävyys saadaan säännön (2) nojalla vaikuttavuuksilla painotettuna summana muiden sivujen merkittävyydestä, saamme yhtälöryhmän:

$$\begin{cases} \frac{1}{4}x_4 + x_5 + \frac{1}{3}x_6 = x_1 \\ \frac{1}{3}x_1 + \frac{1}{3}x_6 = x_2 \\ \frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{4}x_4 = x_3 \\ \frac{1}{4}x_2 + \frac{1}{2}x_3 = x_4 \\ \frac{1}{4}x_2 + \frac{1}{2}x_3 + \frac{1}{4}x_4 + \frac{1}{3}x_6 = x_5 \\ \frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{4}x_4 = x_6 \end{cases}$$

Toisin sanoen jos $\mathbf{x} = [x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6]^T$, niin

$$P\mathbf{x} = \mathbf{x}.$$

Tällöin sanomme, että $\mathbf{x}$ on ns. *ominaisvektori* matriisille P ominaisarvolla 1. Sitä kutsutaan verkkomme *PageRank vektoriksi*.

Ratkaisemalla nyt $\mathbf{x}$, esimerkiksi rivioperaatioilla yhtälöstä $(P - I)\mathbf{x} = 0$, missä I on 6×6 identiteettimatriisi, saamme kolmen desimaalin tarkkuudella

$$\mathbf{x} \approx \begin{pmatrix} 0.265 \\ 0.138 \\ 0.150 \\ 0.110 \\ 0.187 \\ 0.150 \end{pmatrix}$$

Siis verkkosivu A on merkittävin, sitten E ja niin edelleen.

Myöhemmin kun puhumme dynaamisista systeemeistä ja Markovin ketjuista, huomaamme, että vektorin $\mathbf{x}$ voi myös löytää ns. *internet surffaajan satunnaisklikkailuna*. Tällöin lähdettyä

esimerkiksi tasaisesta alkutilasta $\mathbf{y} = [\frac{1}{6} \frac{1}{6} \frac{1}{6} \frac{1}{6} \frac{1}{6} \frac{1}{6}]^T$, jossa surffaaja ei vielä tiedä mitkä ovat sivujen merkittävyydet, niin iteraatiot matriisin P tuloja itsensä kanssa vektoriin

$$\mathbf{y} = \begin{pmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \end{pmatrix}, P\mathbf{y} \approx \begin{pmatrix} 0.264 \\ 0.111 \\ 0.139 \\ 0.125 \\ 0.222 \\ 0.139 \end{pmatrix}, P^2\mathbf{y} \approx \begin{pmatrix} 0.300 \\ 0.134 \\ 0.147 \\ 0.097 \\ 0.175 \\ 0.147 \end{pmatrix}, \dots, P^{12}\mathbf{y} \approx \begin{pmatrix} 0.265 \\ 0.138 \\ 0.150 \\ 0.110 \\ 0.187 \\ 0.150 \end{pmatrix}, \dots$$

ovat nopeasti hyvin lähellä ominaisvektoria $\mathbf{x}$. Tämä on usein nopea tapa löytää $\mathbf{x}$ kuin algebrallisesti ratkoa valtava matriisiyhtälö $P\mathbf{x} = \mathbf{x}$. Tämä ilmiö johtuu siitä, että P voidaan tulkita kertomaan todennäköisyyksiä mille sivulle satunnainen klikkaaja menisi miltäkin sivulta oltuaan ja muutaman iteraation jälkeen tulee jo hyvä kuva mitkä sivut ovat merkittävimpiä koska prosessi lähestyy nopeasti ns. *tasapainotilaa* (vektori $\mathbf{x}$). Riskinä tässä satunnaisklikkailussa on se, että surffaaja voi juuttua nettisivuille, joissa ei ole linkkejä muille sivuille. Silloin PageRank algoritmia voi muokata niin, että surffaaja vain sulkee sivun ja valitsee satunnaisesti jonkin toisen toisen nettisivun, siirtyy sinne ja jatkaa satunnaisklikkailuaan. Puhumme tästä uudestaan hieman enemmän muilla esimerkeillä myöhemmin kurssilla.

Mennään nyt yleiseen ominaisarvoteoriaan.

Kun matriisilla kerrotaan vektoria, vektorin suunta ja pituus yleensä muuttuvat. Jotkin vektorit kuitenkin säilyttävät suuntansa. Näitä sanotaan matriisin ominaisvektoreiksi:

Määritelmä 4.1 (Ominaisarvot ja vektorit)

Jos $n \times n$ -matriisille A pätee

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

jollakin vektorilla $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ ja skalaarilla $\lambda \in \mathbb{C}$, niin λ on matriisin A *ominaisarvo* ja $\mathbf{x}$ sitä vastaava *ominaisvektori*.

Esimerkki 4.2 (lasketaan luennolla)

Olkoon $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$ ja $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Ovatko $\mathbf{u}$ ja $\mathbf{v}$ matriisin A ominaisvektoreita?

Vastaus: $\mathbf{u}$ on, sillä $A\mathbf{u} = -4\mathbf{u}$. $\mathbf{v}$ ei ole, sillä $A\mathbf{v} \neq \lambda\mathbf{v} \ \forall \lambda \in \mathbb{C}$.

Esimerkki 4.3 (lasketaan luennolla)

Osoita, että 7 on matriisin $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$ ominaisarvo.

Ratkaisu: $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Lause 4.4

Erisuuria ominaisarvoja vastaavat ominaisvektorit ovat lineaarisesti riippumattomia.

Muistetaan, että vektorit $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ ovat lineaarisesti riippumattomat, jos mitään niistä ei voida lausua lineaarikombinaationa toisista, eli yhtälön

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n = 0$$

ainoa ratkaisu on $c_1 = c_2 = \dots = c_n = 0$.

Todistus

[Todistus] Luennolla, vastaoleuksesta johdetaan ristiriita. □

Huomioita

- reaalisia ominaisvektoreita ei aina ole olemassa
- ominaisvektori on määritelmän mukaan nolasta eroava
- ominaisarvo voi olla nolla
- $A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow A(t\mathbf{x}) = \lambda(t\mathbf{x})$ kaikilla $t \in \mathbb{C}$, joten ominaisvektorin $\mathbf{x}$ sijaan voidaan puhua $\mathbf{x}$:n suuntaisesta *ominaissuorasta* $\{t\mathbf{x} \mid t \in \mathbb{C}\} \subset \mathbb{C}^n$. (Kulkee origon kautta.)
- Jos matriisin $A \in \mathbb{R}^{n \times n}$ ominaisarvo $\lambda \neq 0$, niin vastaava ominaissuora kuvautuu itselleen ja ominaisarvo λ ilmoittaa ominaissuoran suuntaisen (mahdollisesti kompleksisen!) venytyksen.
- Jos $\mathbb{R} \ni \lambda < 0$, niin suunnistus ominaissuoralla kääntyy, ts. venytyksen lisäksi lineaarikuvaus peilaa ominaissuoran normaalin suhteen.
- Jos $\lambda = 0$, niin kuvaus litistää ominaissuoran origoksi.

Ominaisarvoille pätevät seuraavat tulokset:

- Matriisi A on kääntyvä, jos ja vain jos 0 ei ole sen ominaisarvo.
- Jos $\lambda \neq 0$ on kääntyvän matriisin ominaisarvo, niin $1/\lambda$ on käänteismatriisin A^{-1} ominaisarvo.
- Kolmiomatriisin ominaisarvot ovat sen diagonaalialkiot.
- Matriisin determinantti on yhtä kuin sen ominaisarvojen tulo:

$$\det(A) = \lambda_1 \lambda_2 \dots \lambda_n$$

- Matriisin $A = (a_{ij})$ diagonaalialkioiden summa eli *jälki* (engl. trace) on yhtä kuin sen ominaisarvojen summa:

$$a_{11} + a_{22} + \dots + a_{nn} = \lambda_1 + \lambda_2 + \dots + \lambda_n$$

4.2 Laskeminen

Ominaisyyhtälö $A\mathbf{x} = \lambda\mathbf{x}$ on yhtäpitävästi $(A - \lambda I)\mathbf{x} = \mathbf{0}$, missä I on identtinen matriisi.

Tälle löytyy nollasta eroava ratkaisu $\mathbf{x}$ täsmälleen silloin, kun $\det(A - \lambda I) = 0$ jollekin $\lambda \in \mathbb{R}$.

Tarkastellaan esimerkkinä lineaarikuvausta

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

Muodostetaan lineaarikuvauksen *karakteristinen polynomi* $p(\lambda) = \det(A - \lambda I)$:

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) = \det\left(\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ &= \det\left(\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) = \det\begin{pmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} \\ &= (1-\lambda)^2 - 4. \end{aligned}$$

Haluttiin $\det(A - \lambda I) = 0$ eli $p(\lambda) = 0$:

$$\begin{aligned} (1-\lambda)^2 - 4 &= 0 \\ \Leftrightarrow (1-\lambda)^2 &= 4 \\ \Leftrightarrow (1-\lambda) &= \pm 2 \\ \Leftrightarrow \lambda &= -1 \text{ tai } \lambda = 3. \end{aligned}$$

Nämä ovat A :n ominaisarvot. Etsitään niitä vastaavat ominaisvektorit ratkaisemalla $\mathbf{x}$ yhtälöstä $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

Kun $\lambda = -1$, niin yhtälö on $(A + I)\mathbf{x} = \mathbf{0}$, ts.

$$\left(\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ts.

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

eli saadaan yhtälöpari

$$\begin{cases} 2x_1 + 2x_2 = 0 \\ 2x_1 + 2x_2 = 0. \end{cases}$$

(Lineaarisesti riippuvat yhtälöt juuri kuten pitääkin, sillä halutaan ratkaisuksi ominaissuora.)

Yhtälöparin ratkaisujoukko on $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = -x_1\}$. (Suora.)

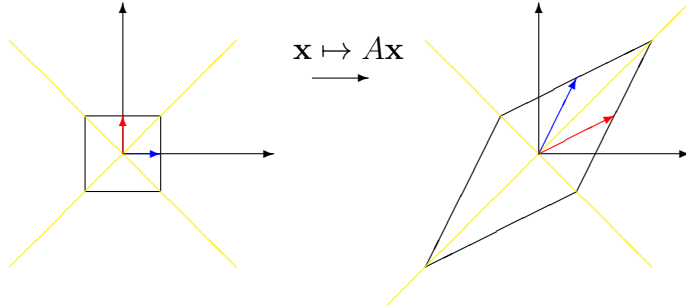
Vastaavasti kun $\lambda = 3$, niin ominaisyyhtälö on $(A - 3I)\mathbf{x} = \mathbf{0}$ ja saadaan yhtälöpari

$$\begin{cases} -2x_1 + 2x_2 = 0 \\ 2x_1 - 2x_2 = 0. \end{cases}$$

Jälleen lineaarisesti riippuvat yhtälöt; ratkaisujoukko on suora $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = x_1\}$.

Yhteenveto: ominaisarvoa -1 vastaava ominaissuora on $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = -x_1\}$ ja ominaisarvoa 3 vastaava ominaissuora on $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = x_1\}$. Ominaisvektoreita ovat näillä suorilla olevat vektorit, esim. $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ja $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Nämä tiedot kertovat lineaarikuvauksesta kaiken!



Kuvassa skaalataan ominaissuorien suuntaisesti kertoimilla 3 ja -1 .

Ominaisarvot ja $-$ -vektorit lasketaan siis seuraavasti:

- Muodosta karakteristinen polynomi $p(\lambda) = \det(A - \lambda I)$.
- Etsi karakteristisen polynomin nollakohdat $p(\lambda) = 0$, nämä ovat ominaisarvot.
- Ratkaise kullakin ominaisarvolla λ_i sitä vastaava ominaisvektori/suora yhtälöstä $(A - \lambda_i I)\mathbf{x} = \mathbf{0}$.

Esimerkki 4.5 (lasketaan luennolla)

Määritä matriisin

$$A = \begin{pmatrix} 9 & -5 & 2 \\ 14 & -8 & 4 \\ 10 & -7 & 5 \end{pmatrix}$$

ominaisarvot ja vastaavat ominaissuorat. Piirrä kuva.

Vastaus: ominaisarvot ovat 1 , 2 , ja 3 , ja vastaavat ominaissuorat ovat $\{t \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} : t \in \mathbb{R}\}$,

$\{t \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} : t \in \mathbb{R}\}$ ja $\{t \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} : t \in \mathbb{R}\}$.

Esimerkki 4.6

Matriisin

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

ominaisarvot ovat 1 ja -1 , sillä sen määräämä lineaarikuvaus on peilaus suoran suhteen.

Esimerkki 4.7

Matriisilla

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

ei ole reaalisia ominaisvektoreita, sillä sen määräämä lineaarikuvaus on 90 asteen pyöritys. Ominaisarvot ovat i ja $-i$.

Kuten edellisessä esimerkissäkin nähtiin, reaaliselle matriisille kompleksiset ominaisarvot esiintyvät konjugaattiparina.

Lause 4.8

Jos reaalisella matriisilla A on kompleksinen ominaisarvo $\lambda = x + yi$, jota vastaa ominaisvektori $\mathbf{v}$, niin myös $\bar{\lambda} = x - yi$ on ominaisarvo ja $\bar{\mathbf{v}}$ on sitä vastaava ominaisvektori.

Todistus

[Todistus] $A\mathbf{v} = \lambda\mathbf{v} \Rightarrow A\bar{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}.$

□

4.3 Sovelluksia dynaamisiin systeemeihin ja Markovin ketjuihin

Markovin ketjuja käytetään matemaattisina malleina useissa erilaisissa biologian, talouden, kemian, fysiikan jne. tilanteissa. Malli sopii tilanteeseen, jossa samaa koetta tai mittausta toistetaan, tulos on yksi äärellisestä joukosta vaihtoehtoja ja kunkin toiston tulos riippuu vain edellisen toiston tuloksesta:

Systeemin tilaa hetkellä k kuvaa vektori $\mathbf{x}_k$. Tiedetään, että tila seuraavalla hetkellä saadaan laskettua matriisin A avulla, eli

$$\mathbf{x}_{k+1} = A\mathbf{x}_k.$$

Tyypillisesti matriisin A alkiot kuvaavat siirtymätodennäköisyyksiä, joten kunkin sen sarakkeen alkioden summa on 1.

Esimerkki 4.9

Systeemin tilaa hetkellä k kuvaa vektori $\mathbf{x}_k$. Tiedetään, että tila seuraavalla hetkellä saadaan laskettua matriisin $A = \begin{pmatrix} .95 & .03 \\ .05 & .97 \end{pmatrix}$ avulla: $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Jos alussa $\mathbf{x}_0 = [.6, .4]^T$, niin mihin tilaan systeemi päättyy lopulta?

Ratkaisu: Lasketaan A :n ominaisarvot ja ominaisvektorit: $\lambda_1 = 1$, $\lambda_2 = 0.92$, $\mathbf{v}_1 = [3, 5]^T$, $\mathbf{v}_2 = [1, -1]^T$.

Kirjoitetaan $\mathbf{x}_0$ ominaisvektoreiden kombinaationa: $\mathbf{x}_0 = 0.125\mathbf{v}_1 + 0.225\mathbf{v}_2$.

Tällöin $\mathbf{x}_k = A^k\mathbf{x}_0 = 0.125 \cdot 1^k\mathbf{v}_1 + 0.225 \cdot 0.92^k\mathbf{v}_2 \rightarrow 0.125\mathbf{v}_1$, kun $k \rightarrow \infty$. Systeemi päättyy siis tilaan $0.125\mathbf{v}_1 = [0.375, 0.625]^T$.

Esimerkki 4.10 (lasketaan luennolla)

Erään kaupungin säätilaa voidaan kuvata yksinkertaistetusti siten, että sadepäivän jälkeen seuraavanakin päivänä sataa todennäköisyydellä 0.5 ja poutapäivän jälkeen on seuraavanakin päivänä poutaa todennäköisyydellä 0.9. Olkoon vektori

$$\mathbf{x}_k = \begin{bmatrix} \text{sateettoman sään todennäköisyys päivänä } k \\ \text{sateisen sään todennäköisyys päivänä } k \end{bmatrix}.$$

Muodosta matriisi A , jonka avulla saat laskettua säävektorin seuraavalle päivälle, eli $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

Miten suurella todennäköisyydellä satunnaisena päivänä sataa?

Ratkaisu:

Sateen ja aurinkoisen sään todennäköisyys päivänä $k+1$ saadaan siis päivän k todennäköisyyksistä seuraavasti:

$$\mathbf{x}_{k+1} = \begin{bmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{bmatrix} \mathbf{x}_k$$

Rajakäyttäytymisen $k \rightarrow \infty$ selvittämiseksi lasketaan A :n ominaisarvot ja -vektorit:

$$\begin{aligned} \lambda_1 + \lambda_2 &= \text{Tr}(A) = 1.4 \\ \lambda_1 \lambda_2 &= \det(A) = 0.4 \\ \Rightarrow \lambda_1 &= 1, \lambda_2 = 0.4 \end{aligned}$$

$\lambda_1 = 1$:

$$\begin{bmatrix} -0.1 & 0.5 & | 0 \\ 0.1 & -0.5 & | 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 5 & | 0 \\ 0 & 0 & | 0 \end{bmatrix} \Rightarrow \mathbf{v}_1 = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

$\lambda_2 = 0.4$:

$$\begin{bmatrix} 0.5 & 0.5 & | 0 \\ 0.1 & 0.1 & | 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & | 0 \\ 0 & 0 & | 0 \end{bmatrix} \Rightarrow \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Ominaisvektorit ovat riippumattomat, joten kirjoitetaan $\mathbf{x}_0$ muodossa $\mathbf{x}_0 = w_1 \mathbf{v}_1 + w_2 \mathbf{v}_2$, jolloin saadaan

$$\mathbf{x}_k = A^k \mathbf{x}_0 = w_1 \cdot 1^k \cdot \begin{pmatrix} 5 \\ 1 \end{pmatrix} + w_2 \cdot 0.4^k \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rightarrow w_1 \begin{pmatrix} 5 \\ 1 \end{pmatrix}, \text{ kun } k \rightarrow \infty.$$

Vielä pitää selvittää w_1 . Koska alkutilan $\mathbf{x}_0$ alkioiden täytyy summautua ykköseen (tn:llä 1 joka sataa tai paistaa), alkutilasta riippumatta saadaan

$$\begin{aligned} w_1 \mathbf{v}_1 + w_2 \mathbf{v}_2 = \mathbf{x}_0 &= [x_{0,1}, x_{0,2}]^T \Leftrightarrow \begin{pmatrix} 5w_1 + w_2 \\ w_1 - w_2 \end{pmatrix} = \begin{pmatrix} x_{0,1} \\ x_{0,2} \end{pmatrix} \\ \Rightarrow 6w_1 = x_{0,1} + x_{0,2} = 1 &\Rightarrow w_1 = \frac{1}{6}. \end{aligned}$$

Näin ollen

$$\mathbf{x}_k \rightarrow \begin{pmatrix} \frac{5}{6} \\ \frac{1}{6} \end{pmatrix}, \text{ kun } k \rightarrow \infty,$$

joten satunnaisena päivänä paistaa todennäköisyydellä $\frac{5}{6}$ ja sadetta on luvassa todennäköisyydellä $\frac{1}{6}$.

Esimerkki 4.11

Kokeile laskea seuraavien dynaamisia systeemejä kuvaavien matriisien A ja A^∞ ominaisarvot ja ominaisvektorit

$$A = \begin{pmatrix} .6 & .2 \\ .4 & .8 \end{pmatrix}, \quad A^\infty = \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{pmatrix}$$

Jos nyt laskisi lisäksi matriisin A^{100} , niin voidaan huomata, että matriisin A^{100} esittämä lineaarikuvaus on jo hyvin lähellä matriisin A^∞ lineaarikuvausta. Voimme tehdä tämän formaalisti kun ns. *diagonalisoimme* matriiseja, joka samalla idealla kuin edellisessä esimerkissä antaa meille tavan laskea miten potenssi A^k käyttäytyy kun k kasvaa.

4.4 Kertaluvut

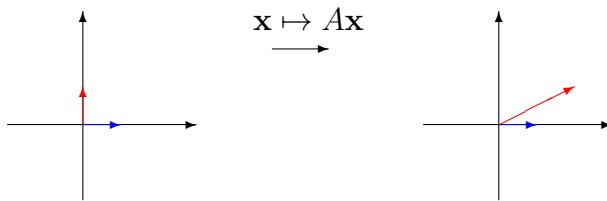
Määritelmä 4.12

Polynomin $\det(A - \lambda I)$ juuren kertaluku on kyseisen ominaisarvon λ *algebrallinen kertaluku* $m_a(\lambda)$. Ominaisarvon λ *geometrinen kertaluku* $m_g(\lambda)$ on sitä vastaavan ominaisavaruuden dimensio, eli lineaarisesti riippumattomien ominaisvektorien lukumäärä.

Esimerkki 4.13

Matriisin $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ karakteristinen polynomi on $(1 - \lambda)^2$, joten ominaisarvon $\lambda = 1$ algebrallinen kertaluku on $m_a(1) = 2$. Sitä vastaavat ominaisvektorit toteuttavat $x_2 = 0$, joten on vain yksi ominaissuora $\{(x, 0) \mid x \in \mathbb{R}\}$, ja geometrinen kertaluku on siis $m_g(1) = 1$.

Piirretään kuva esimerkin tilanteesta:



Vain x_1 -akseli pysyy siis paikallaan, toista muuttumatonta suoraa ei ole. Tässä tapauksessa ominaisarvot ja -vektorit **eivät** kerro kaikkea kuvauksesta!

Esimerkki 4.14 (lasketaan luennolla)

Etsi matriisiin

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

ominaisarvojen algebralliset ja geometriset kertaluvut.

Ratkaisu: Lasketaan ominaisarvot karakteristisen polynomin nollakohtina: $\det(A - \lambda I) = \dots = (3 - \lambda)^2(\lambda + 1) = 0$, joten ominaisarvon $\lambda = 3$ algebrallinen kertaluku on 2 ja ominaisarvon $\lambda = -1$ on 1.

Lasketaan sitten ominaisvektorit:

Kun $\lambda = -1$, yhtälö on $(A + 1I)\mathbf{x} = 0$, ja saadaan $x_2 = 0$ ja $x_3 = -x_1$, eli ominaissuora $\{t(1, 0, -1) \mid t \in \mathbb{R}\}$. Näin ollen geometrinen kertaluku on $m_g(-1) = m_a(-1) = 1$.

Arvolle $\lambda = 3$ saadaan yhtälöstä $(A - 3I)\mathbf{x} = 0$ ehdot $x_2 \in \mathbb{R}$ ja $x_3 = x_1$, joten tätä ominaisarvoa vastaten saadaanakin ominaistaso

$$\{s(1, 0, 1) + t(0, 1, 0) \mid s, t \in \mathbb{R}\}.$$

(Lineaarisesti riippumattomat ominaisvektorit esim. $(1, 0, 1)$ ja $(0, 1, 0)$). Geometrinen kertaluku on siis $m_g(3) = 2$.

Huom 1. Geometrinen kertaluku ei koskaan voi olla suurempi kuin algebrallinen kertaluku, $m_g(\lambda) \leq m_a(\lambda)$.

Huom 2. Matriisit A ja B ovat *similaariset*, jos on olemassa sellainen kääntyvä matriisi S , että $S^{-1}AS = B$. Similaaristen matriisien A ja B karakteristiset polynomit ovat samat, ja niillä on siis samat ominaisarvot.

Luku 5

Matriisihajotelmat

Laskentaongelmissa käsiteltävät matriisit ovat tyypillisesti valtavia. Jotta laskut menisivät läpi edes jossain määrin järjellisessä ajassa, käytetään hyväksi mm. matriisihajotelmia. Idea-
na niissä on purkaa matriisi useamman, yksinkertaisemman matriisin tuloksi. Matriisihajotel-
mia on koko joukko, ne soveltuvat eri käyttötarkoituksiin ja erityyppisille matriiseille. Tällä
kurssilla tarkastelemme varsinaisesti näistä kahta: diagonalisointia ja singulaariarvohajotelmaa
(SVD). Lisäksi harjoituksessa 4 esiintyi LU-hajotelma.

5.1 Diagonalisointi

$\mathbb{R}^{n \times n}$ neliömatriisi A , joilla on n riippumatonta ominaisvektoria, voidaan diagonalisoida. Tämä
tarkoittaa, että matriisi $A \in \mathbb{R}^{n \times n}$ voidaan kirjoittaa muodossa

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

missä matriisin S sarakkeet ovat matriisin A ominaisvektorit, ja matriisi Λ on diagonaalimat-
riisi, jossa kussakin sarakkeessa on matriisissa S samassa sarakkeessa olevaan ominaisvektoriin
liittyvä ominaisarvo.

Esimerkki 5.1 (Luentoharjoitus)

Aiemmin laskimme matriisille

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

ominaisarvot ja -vektorit, ja saimme ominaisarvoa -1 vastaavaksi ominaisvektoriksi
 $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ja ominaisarvoa 3 vastaavaksi $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Diagonalisoi A .

Vastaus:

$$A = SAS^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

Huom $n \times n$ -matriisi diagonalisoituu, jos sillä on n lin. riippumatonta ominaisvektoria.
Erityisesti tämä tapahtuu silloin, kun matriisilla on n erisuurta ominaisarvoa, mutta voi siis
tapahtua muulloinkin!

Esimerkki 5.2

Matriisi

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

ei diagonalisoidu. Sen ainoa ominaisarvo on nimittäin 0, eli $m_a(0) = 2$. Ominaisvektoreita laskettaessa kuitenkin havaitaan, että löytyy vain yksi ominaisvektori, eli $m_g(0) = 1$, eikä A diagonalisoidu.

Esimerkki 5.3

Matriisi

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

diagonalisoituu: Sen ominaisarvot ovat 3 ja -1 , $m_a(3) = 2$ ja $m_a(-1) = 1$, mutta onneksi (kuten aiemmin laskettiin) ominaisarvoa 3 vastaten löytyy ominaistaso, eli saadaan kaksi lin. riippumatonta ominaisvektoria om. arvolle 3. Saadaan hajotelma

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

Huom. Jos matriisi on diagonalisoituvaa, sen potensseja on hyvin näppärä laskea:

$$\begin{aligned} A^2 &= S \underbrace{\Lambda S^{-1} S}_{=I} \Lambda S^{-1} = S \Lambda^2 S^{-1} \\ A^3 &= S \underbrace{\Lambda S^{-1} S}_{=I} \underbrace{\Lambda S^{-1} S}_{=I} \Lambda S^{-1} = S \Lambda^3 S^{-1} \\ A^k &= S \Lambda^k S^{-1} \end{aligned}$$

Diagonaalimatriisin potenssithan ovat helppoja:

$$\Lambda^k = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n^k \end{pmatrix}.$$

Esimerkki 5.4 (Luentoharjoitus)

Laske A^{2014} , kun

$$A = \begin{pmatrix} 5 & 1 \\ 0 & 4 \end{pmatrix}$$

Vastaus:

$$A = S\Lambda S^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix},$$

$$\begin{aligned} A^{2014} &= \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5^{2014} & 0 \\ 0 & 4^{2014} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 5^{2014} & 5^{2014} - 4^{2014} \\ 0 & 4^{2014} \end{pmatrix} \end{aligned}$$

Huom. Käänteismatriisin löytäminenkin on diagonalisoidulle matriisille helppoa:

$$A^{-1} = (S\Lambda S^{-1})^{-1} = S\Lambda^{-1}S^{-1}.$$

Diagonaalimatriisillehan

$$\Lambda^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n^{-1} \end{pmatrix}.$$

5.2 Unitaarinen diagonalisointi

Joissakin tapauksissa diagonalisointi onnistuu erityisen mukavasti: tästä on kyse normaalin matriisin unitaarisessa diagonalisoinnissa.

Ensin hieman terminologiaa:

- Matriisi $A^* \in \mathbb{C}^{n \times m}$ on matriisin $A \in \mathbb{C}^{m \times n}$ *konjugaattitranspoosi*, jos $(A^*)_{kj} = \overline{A_{jk}}$.
- Matriisi $A \in \mathbb{C}^{m \times n}$ on *symmetrinen*, jos $A^* = A$.
- Matriisi $A \in \mathbb{C}^{m \times n}$ on *normaali*, jos $A^*A = AA^*$.
- Matriisi $U \in \mathbb{C}^{n \times n}$ on *unitaarinen*, jos $U^* = U^{-1}$.

Huomaa, että ehdoista $A^* = A$ ja $A^*A = AA^*$ seuraa heti, että matriisi voi itseasiassa olla symmetrinen tai normaali vain, jos se on neliömatriisi, eli $n = m$. (Miksi?)

Lause 5.5 (Normaalin matriisin unitaaridiagonalisointi)

Matriisille $A \in \mathbb{C}^{n \times n}$ seuraavat ehdot ovat yhtäpitäviä:

1. A on normaali.
2. $A = U\Lambda U^*$, missä $U \in \mathbb{C}^{n \times n}$ on unitaarinen ja $\Lambda \in \mathbb{C}^{n \times n}$ diagonaalinen.

Esimerkki 5.6

Jos A on symmetrinen, eli $A^* = A$, niin A on normaali:

$$A^*A = A^2 = AA^*.$$

Symmetrinen matriisi on siis aina unitaarisesti diagonalisoituva.

Esimerkki 5.7 (lasketaan luennolla)

Osoita, että matriisi $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ei ole normaali.

Lause 5.8 (Erikoistapaus unitaaridiagonalisoinnista)

Jos $A \in \mathbb{C}^{n \times n}$ on symmetrinen ja sen ominaisarvot ovat erisuuret, niin $A = U\Lambda U^*$, missä $U \in \mathbb{C}^{n \times n}$ on unitaarinen ja $\Lambda \in \mathbb{R}^{n \times n}$ diagonaalinen. Lisäksi matriisi U saadaan sarakematriisina $U = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_n]$ ominaisvektoreista joiden pituus on 1 ja Λ ominaisarvoista.

Todistus

[Ei kysytä tentissä/kokeessa] Kompleksisille vektoreille $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ pistetulo on

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n \overline{u_i} v_i = \mathbf{u}^* \mathbf{v},$$

missä jälkimmäinen yhtäsuuruus seuraa tulkitsemalla $\mathbf{u}$ ja $\mathbf{v}$ pystyvektoreiksi: $(1 \times n)$ -matriisi kertaa $(n \times 1)$ -matriisi on (1×1) -matriisi eli skalaari. Lisäksi tällöin saamme konjugaattitranspoosille hyödyllisen ominaisuuden:

$$\mathbf{u}^*(A^* \mathbf{v}) = \mathbf{u} \cdot (A^* \mathbf{v}) = A \mathbf{u} \cdot \mathbf{v} = (A \mathbf{u})^* \mathbf{v}, \quad \mathbf{u}, \mathbf{v} \in \mathbb{C}^n,$$

missä $\mathbf{u} \cdot (A^* \mathbf{v}) = A \mathbf{u} \cdot \mathbf{v}$ seuraa konjugaattitranspoosin määritelmästä ja sisätulosta.

Tämän havainnon avulla voimme todistaa, että A :n ominaisarvot ovat reaalisia. Jos λ on A :n jokin ominaisarvo ominaisvektorilla $\mathbf{v} \neq 0$, niin kompleksi konjugaatin laskusäännöllä saamme tarkistettua

$$\lambda |\mathbf{v}|^2 = \lambda \mathbf{v} \cdot \mathbf{v} = \lambda \mathbf{v}^* \mathbf{v} = \mathbf{v}^*(\lambda \mathbf{v}) = \mathbf{v}^*(A \mathbf{v}) = \mathbf{v}^* A^* \mathbf{v} = (A \mathbf{v})^* \mathbf{v} = (\lambda \mathbf{v})^* \mathbf{v} = \overline{\lambda} \mathbf{v}^* \mathbf{v} = \overline{\lambda} |\mathbf{v}|^2$$

Koska $\mathbf{v} \neq 0$, on $|\mathbf{v}|^2 \neq 0$, joten voimme jakaa pois ja saamme $\lambda = \overline{\lambda}$. Siis $\lambda \in \mathbb{R}$. Tästä saamme diagonaalimatriisin $\Lambda \in \mathbb{R}^{n \times n}$.

Entä sitten unitaarimatriisi U ? Normaalin matriisin eri ominaisarvoja vastaavat normeeratut ominaisvektorit ovat ortonormaaleja. Samaa ominaisarvoa vastaavat ominaisvektorit voidaan valita siten, että ne ovat ortonormaaleja. Matriisi U saadaan siis normeeratuista, ortonormaaleista ominaisvektoreista. Toisin sanoen unitaarimatriisi U saadaan nyt sarakematriisista

$$U = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_n]$$

muodostettuna ominaisvektoreista $\mathbf{v}_j$, joiden pituus on $|\mathbf{v}_j| = 1$. Huomaa, että jotta matriisi U voisi olla unitaarinen, sen sarakevektorien pitää olla ortonormaaleja. Kompleksiset vektorit ovat ortogonaaliset, jos niiden pistetulo on 0. Siis tarkistetaan vielä, että $\mathbf{v}_j \cdot \mathbf{v}_k = 0$ kaikilla $j \neq k$. Koska matriisi A on symmetrinen (ja $\overline{\lambda_j} = \lambda_j$), saamme

$$\lambda_j (\mathbf{v}_j \cdot \mathbf{v}_k) = (\lambda_j \mathbf{v}_j) \cdot \mathbf{v}_k = (A \mathbf{v}_j) \cdot \mathbf{v}_k = (A \mathbf{v}_j)^* \mathbf{v}_k = \mathbf{v}_j^* (A^* \mathbf{v}_k) = \mathbf{v}_j^* (A \mathbf{v}_k) = \mathbf{v}_j^* (\lambda_k \mathbf{v}_k) = \lambda_k (\mathbf{v}_j \cdot \mathbf{v}_k).$$

Siispä

$$(\lambda_j - \lambda_k)(\mathbf{v}_j \cdot \mathbf{v}_k) = 0.$$

Koska $\lambda_j \neq \lambda_k$ (tämä oli oletus), niin tämä on mahdollista vain jos $\mathbf{v}_j \cdot \mathbf{v}_k = 0$.

□

Esimerkki 5.9 (lasketaan luennolla)

Kvanttilaskennan malleissa ja spin ketjuissa käytetään usein ns. *Pauli matriiseja* (lisää näistä esim. https://en.wikipedia.org/wiki/Pauli_matrices). Yksi näistä on

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Diagonalisoi se unitaarisesti.

Ratkaisu: Huomataan, että σ_1 on symmetrinen:

$$\sigma_1^* = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1$$

joten erityisesti

$$\sigma_1^* \sigma_1 = \sigma_1^2 = \sigma_1^* \sigma_1$$

eli σ_1 on normaali ja se voidaan diagonalisoida unitaarisesti.

Etsitään ominaisarvot:

$$\det(\sigma_1 - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1,$$

joten $\lambda_1 = 1$, $\lambda_2 = -1$.

Etsitään ominaisvektorit: yhtälöstä $(\sigma_1 - I)\mathbf{x} = 0$ saadaan

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

eli ehto $-x_1 + x_2 = 0$. Valitaan ominaisvektoriksi $\mathbf{v}_1 = (1/\sqrt{2}, 1/\sqrt{2})^T$. (Tällöin sen pituus on 1!)

Vastaavasti ominaisarvolle $\lambda_2 = -1$ saadaan ominaisvektori $\mathbf{v}_2 = (1/\sqrt{2}, -1/\sqrt{2})^T$.

Pauli matriisille σ_1 saadaan siis unitaarinen diagonalisointi

$$\begin{aligned} \sigma_1 &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \\ &= U \Lambda U^*. \end{aligned}$$

(Tarkista kertolaskulla!)

Muita kvanttilaskennassa käytettyjä Pauli matriiseja ovat

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

ja

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

jotka ovat myös unitaarisesti diagonalisoitavissa samalla idealla kuin aiemmassa esimerkissä.

Pauli matriisit ovat esimerkkejä ns. *kvanttiporteista*, joilla voi luoda kvanttimekaanisia lo-giikkayksiköitä kvanttilaskennassa. Niiden avulla voidaan suorittaa laskutoimituksia kvant-ti biteille (qubit), jotka ovat vektoreita. Muita esimerkkejä ovat esimerkiksi Hadamard (H) portti ja Phase (S,P) portti. Kiinnostuneelle lukijalle käy katsomassa esimerkkejä https://en.wikipedia.org/wiki/Quantum_logic_gate

5.3 Singulaariarvohajotelma (SVD)

Kuten edellä huomattiin, vain neliömatriisiin voi diagonalisoida — olettaen, että sillä on täysi määrä riippumattomia ominaisvektoreita. Kaikille matriiseille voidaan kuitenkin tehdä *singulaariarvohajotelma* (Singular Value Decomposition, SVD).

SVD:n ideana on purkaa lineaarikuvaus kolmeen osaan: unitaariseen kuvaukseen, venytykseen ja toiseen unitaariseen kuvaukseen. (Unitaariset kuvaukset voi usein tulkita kierroiksi.) SVD:tä käytetään mm. signaalinkäsittelyssä ja statistiikassa. Siihen perustuvat esimerkiksi jotkin kuvanpakkausmenetelmät ja Googlen hakukoneen toiminta. SVD liittyy läheisesti ns. *pääkomponenttianalyysiin* (PCA), jossa tavoitteena on löytää monidimensioisesta datasta ne komponentit, joilla sen keskeisimmät piirteet voidaan esittää menettämättä oleellista informaatiota.

Jotta voimme määritellä singulaariarvohajotelman, määrittelemme ensin ns. *singulaariarvot*.

Määritelmä 5.10

Luku $\sigma > 0$ on matriisin A *singulaariarvo*, jos on olemassa vektorit $\mathbf{u}, \mathbf{v} \neq 0$ siten, että

$$A\mathbf{v} = \sigma\mathbf{u} \quad \text{ja} \quad A^T\mathbf{u} = \sigma\mathbf{v}.$$

Huomaa, että jos σ on matriisin A singulaariarvo, niin σ^2 on matriisin $A^T A$ ominaisarvo:

$$A^T A\mathbf{v} = A^T(\sigma\mathbf{u}) = \sigma A^T\mathbf{u} = \sigma^2\mathbf{v}.$$

Lause 5.11 (Singulaariarvohajotelma, SVD)

Matriisille $A \in \mathbb{R}^{m \times n}$ on olemassa hajotelma

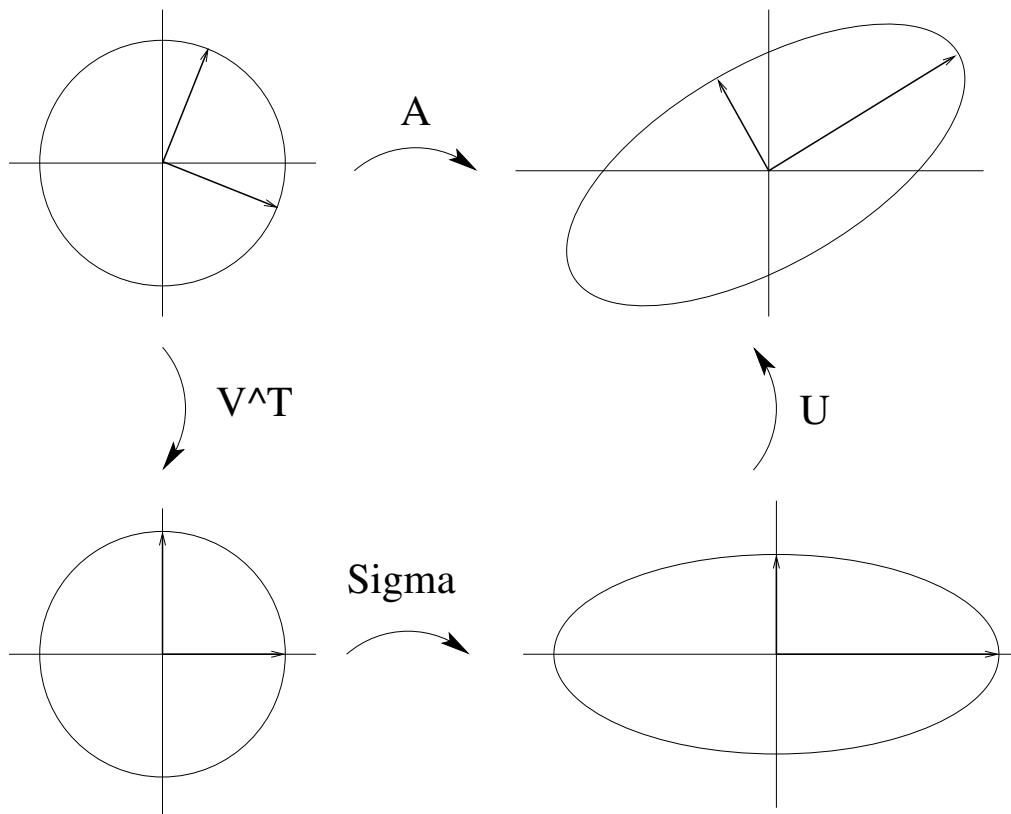
$$A = U\Sigma V^T,$$

missä U on $m \times m$ -ulotteinen ortogonaalinen matriisi, V on $n \times n$ -ulotteinen ortogonaalinen matriisi ja Σ on $m \times n$ -ulotteinen diagonaalimatriisi, joka sisältää matriisin A singulaariarvot.

Muistutus: Matriisi U on ortogonaalinen, jos käänteismatriisi on sen transpoosi: $U^{-1} = U^T$. Tällöin sen sarakevektorit ovat ortonormaaleja. (Huomaa, että reaaliselle matriisille ortogonaalisuus ja unitaarisuus ovat sama asia.)

Huom.

- Vain matriisi Σ on yksikäsitteinen (kun singulaariarvot asetetaan diagonaalille suurimmasta pienimpään), muita voi löytyä useita.
- SVD on olemassa myös kompleksisille matriiseille, silloin transpoosin V^T sijaan tarvitaan adjungaatti V^* (kompleksikonjugaatin transpoosi) ja U ja V ovat unitaarisia, eli $V^{-1} = V^*$, $U^{-1} = U^*$.



Kuva 5.1: SVD geometrisesti: $A = \text{kierto } (U) \text{ kertaa venytys } (\Sigma) \text{ kertaa kierto } V^T$

5.4 SVD:n laskeminen

Matriisit $AA^T \in \mathbb{R}^{m \times m}$ ja $A^T A \in \mathbb{R}^{n \times n}$ ovat aina symmetrisiä, joten niillä on täysi määrä ortogonaalisia ominaisvektoreita. Koko SVD:n idea perustuu tähän. Siten SVD voidaan laskea seuraavien tietojen perusteella:

- (1) Matriisin Σ diagonaalialkiot σ_j ovat matriisin $A^T A$ ominaisarvojen positiiviset neliöjuuret.
- (2) Matriisin V sarakevektorit $\mathbf{v}_j$ (eli V^T :n rivit) ovat matriisin $A^T A$ yksikköpituiset ominaisvektorit.
- (3) Matriisin U sarakevektorit $\mathbf{u}_j$ ovat matriisin AA^T yksikköpituiset ominaisvektorit. Jos singulaariarvo $\sigma_j \neq 0$, niin sarakevektori $\mathbf{u}_j$ voidaan laskea itse asiassa vektorin $\mathbf{v}_j$ avulla kaavalla

$$\mathbf{u}_j = A\mathbf{v}_j/\sigma_j.$$

Jos $\sigma_j = 0$, niin täytyy laskea matriisin AA^T ominaisarvoa 0 vastaava yksikköpituisen ominaisvektori, ja se tulee olemaan $\mathbf{u}_j$.

Käydään tarkkaan läpi yksi esimerkki, jossa tapahtuu myös tilanne $\sigma_j = 0$ jollain j :

Esimerkki 5.12 (Luentoharjoitus)

Etsi matriisin

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$$

singulaariarvohajotelma.

Ratkaisu:

- (1) Matriisin $A^T A = \begin{pmatrix} 10 & 20 \\ 20 & 40 \end{pmatrix}$ ominaisarvot ovat 50 ja 0, joten singulaariarvot ovat $\sigma_1 = \sqrt{50}$, $\sigma_2 = 0$. Siis

$$\Sigma = \begin{pmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{pmatrix}.$$

- (2) Matriisin $A^T A$ yksikköpituiset ominaisvektorit ovat

$$\mathbf{v}_1 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

ja

$$\mathbf{v}_2 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}.$$

Saamme siis

$$V = [\mathbf{v}_1 \quad \mathbf{v}_2]$$

joten

$$V^T = [\mathbf{v}_1 \quad \mathbf{v}_2]^T = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}.$$

(3) Koska $\sigma_1 = \sqrt{50} \neq 0$, voimme laskea vektorin $\mathbf{u}_1$ kaavalla:

$$\mathbf{u}_1 = A\mathbf{v}_1/\sigma_1 = \begin{pmatrix} 5/\sqrt{5} \\ 15/\sqrt{5} \end{pmatrix} \frac{1}{\sqrt{50}} = \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix}.$$

Koska $\sigma_2 = 0$, vektoria $\mathbf{u}_2$ ei saada samalla tavalla, vaan se täytyy laskea matriisin AA^T yksikköpituisena ominaisvektorina:

$$AA^T = \begin{pmatrix} 5 & 15 \\ 15 & 45 \end{pmatrix}$$

ja sen ominaisarvoa 0 vastaava yksikköpituisen ominaisvektori on

$$\mathbf{u}_2 = \begin{pmatrix} 3/\sqrt{10} \\ -1/\sqrt{10} \end{pmatrix}$$

Saamme siis

$$U = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix}$$

Siispä hajotelma on

$$\begin{aligned} A &= U\Sigma V^T \\ &= \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}. \end{aligned}$$

(Tarkista kertolaskulla!)

5.5 Matriisin approksimointi SVD:llä ja kohinan poistaminen datasta

Alkuperäistä $n \times n$ matriisia A voidaan approksimoida sen SVD:n avulla ottamalla k ensimmäistä saraketta matriisista U , $k \times k$ osa diagonaalimatriisin Σ vasemmasta yläkulmasta ja k ensimmäistä riviä matriisista V^T , eli ensimmäistä k :ta singulaariarvoa vastaava osuus. Singulaariarvot oli valittu suuruusjärjestykseen, suurimmasta pienimpään. Näin saadaan alkuperäistä matriisia approksimoiva matriisi, "jonka rangi on k ".

Esimerkki 5.13

Edellisen kappaleen esimerkin matriisille saadaan approksimaatio arvolla $k = 1$:

$$\begin{aligned} & \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix} \\ &\approx \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{50} \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{10} & 2\sqrt{10} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \end{aligned}$$

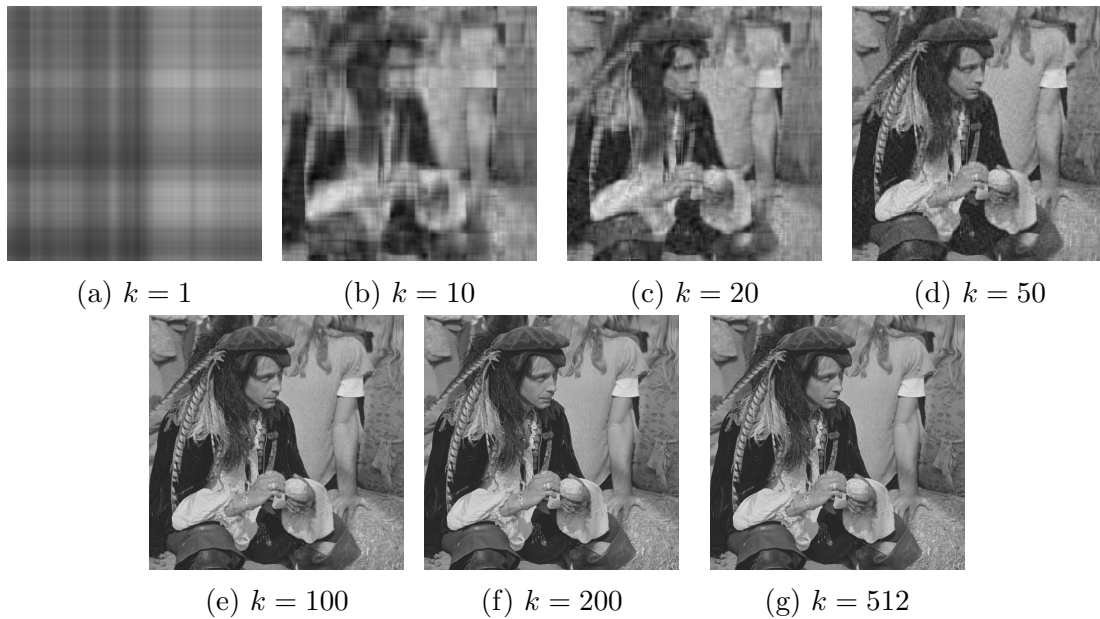
Nyt arvon $k = 1$ approksimaatiomatriisi sattui olemaan sama kuin alkuperäinen.

Tätä ajatusta voi soveltaa melun poistamiseen datasta esimerkiksi kuvan pakkaamisessa.



(a) Alkuperäinen kuva, 210 KB (b) SVD:llä pakattu kuva, 82 KB

Testikuvassa on 512×512 pikseliä, eli se voidaan esittää 512×512 -matriisilla, pikselin väri antaa arvon ko. alkioille matriisissa. Tälle matriisille voidaan tehdä singulaariarvohajotelma, ja approksimoida sitten kuvaa ottamalla hajotelmasta k :ta ensimmäistä singulaariarvoa vastaava osuus. Huomataan, että vähempikin määrä dataa riittää kuvan esittämiseen tunnistattavasti ja jopa silmämääräisesti riittävän tarkasti.



Kuva 5.3: Approksimaatiomatriiseja vastaavat kuvat antavat pakatun version alkuperäisestä kuvasta kun valittujen singulaariarvojen lukumäärä on $k = 1, 10, 20, 50, 100$ ja 200 . Lopussa $k = 512$ on täsmälleen alkuperäinen kuva.

Kun alkuperäisen $m \times n$ -matriisin sijaan käytetään SVD:stä k ensimmäistä singulaariarvoa, tarvittavien lukujen määrä on

$$k + km + kn,$$

eli k singulaariarvoa, niistä vastaavat k ensimmäistä sarakevektoria (pituus m) matriisista U ja k ensimmäistä rivivektoria (pituus n) matriisista V^T .

Neliömatriisille $n \times n$ tämä tarkoittaa, että jos valittujen singulaariarvojen lukumäärä on

$$k < \frac{n^2}{1 + 2n},$$

niin datan määrä pienenee. Edellä testikuvalla $n = 512$, joten $k < 255$ vähentää datamäärää. Huomattiin siis, että SVD:n avulla voidaan poimia datasta olennainen ja päästä turhasta “kohinasta” eroon, pienentäen samalla käsiteltävää datamäärää merkittävästi. Modernin data-analyysin *pääkomponenttianalyysi* (PCA) hyödyntää myös tätä samaa ajatusta (kts. https://en.wikipedia.org/wiki/Principal_component_analysis ja osio yhteydestä SVD).