

(M A T R I I S I) (L A S K E N T A)

Aleksis Koski

16. lokakuuta 2026

Johdanto

Tämä luentomoniste käsittää matriisilaskennan perusteoriaa, alunperin tuotettuna Aalto-yliopiston kurssin *MS-A0001 Matriisilaskenta* sisältöä vastaten. Monisteen ei ole tarkoitus kattaa kaikkia matriisilaskennan aiheita yksityiskohtineen, vaan se on suunniteltu mahdollisimman lähestyttäväksi lukumateriaaliksi aiheeseen ensimmäistä kertaa tutustuvalle opiskelijalle.

Tästä syystä joidenkin todistuksien yksityiskohtia on kevennetty tai ne on sivuutettu. Teoreettisesta tarkkuudesta kiinnostunutta lukijaa kehoitetaan täydentämään mahdolliset puuttuvat yksityiskohdat itse, ja osa näistä on tarkoitettu tehtäviksi kurssin laskuharjoitustilaisuuksiin. Matematiikan opiskelussa laskuharjoitusten tekemisen merkitys on hyvin tärkeä, ja tätä monistetta ei ole tarkoitettu opiskeltavaksi ilman, että ohessa osallistuu myös vastaaviin harjoitustilaisuuksiin.

Opiskelijoita kannustetaan etsimään myös vaihtoehtoja lukumateriaalia tämän monisteen lisäksi, etenkin niille, jotka haluavat syventyä aiheeseen jatkossa enemmänkin. Mikäli lukijalla on toiveita lisäysten suhteen tähän monisteeseen, tai tässä materiaalissa ilmenee virheitä, näistä otetaan yhteydenottoja lämpimästi vastaan.

Monisteessa esiintyvät suomenkieliset termit on valittu kirjoittajan mieltymyksen mukaan vastaamaan mahdollisimman paljon puhekielessä käytettäviä ilmaisuja. Joillain käsitteillä on useita eri nimiä, ja näitä on pyritty mainitsemaan myös melko kattavasti.

Luentomonisteen lopusta löytyy sanasto, jossa esitellään tässä materiaalissa käytetyt termit. Sanaston on laatinut Etna Lindy.

Sisällysluettelo

1	Vektorit ja Matriisit	4
	Vektorit $\mathbb{R}^n$:ssä	4
	Matriisit	10
	Matriisit lineaarikuvauksina	13
	Transpoosi	18
	Bonus: Kompleksiluvut matriiseissa	20
2	Aliavaruudet	22
	Kanta	23
	Matriisin ydin ja kuva	24
3	Lineaariset yhtälöryhmät	26
	Yhtälöryhmän ratkaisujoukko	26
	Gaussin eliminaatio	27
	Pivotit	30
4	Käänteismatriisi	32
	Käänteismatriisi Gauss-Jordanin eliminaatiolla	32
	Kääntävyys ehdot	34
5	Ortogonaalisuus	36
	Ortogonaaliset jonot	36
	Ortogonaalinen projektio	36
	Ortogonalisointi	37
	Ortogonaaliset matriisit	38
	Pienimmän neliösumman ratkaisu yhtälöryhmälle	38
6	Determinantti	40
7	Ominaisarvot ja ominaisvektorit	43
	Ominaisavaruudet	44
	Diagonalisointi	45
8	Neliömuodot	47
9	Matriisihajotelmia	49
	LU-hajotelma	49
	SVD-hajotelma	52
10	Sanasto	53

1. Vektorit ja Matriisit

Tässä luvussa esittelemme tämän materiaalin perusobjektit, vektorit ja matriisit, joiden käsittelyyn myöhemmät aiheet perustuvat. Tarkalleen ottaen käsittelemme tässä materiaalissa lähes yksinomaan **reaalisia** vektoreita ja matriiseja. Tämä tarkoittaa sitä, että vektorimme vastaavat n -ulotteisen reaalisen avaruuden $\mathbb{R}^n$ pisteitä, ja matriisimme näiden avaruuksien lineaarikuvauksia.

Vektorit $\mathbb{R}^n$:ssä

Vektori on matemaattinen käsite, jolla voidaan näppärästi pakata useampia lukuja yhteen objektiin. Esimerkiksi luvut 3, -5, ja 1.5 voidaan asettaa kolmiulotteiseen vektoriin

$$\mathbf{v} = \begin{pmatrix} 3 \\ -5 \\ 1.5 \end{pmatrix}.$$

Vektorissa esiintyvillä luvuilla eli sen **komponenteilla** on järjestys, eli esim vektorit

$$\mathbf{v} = \begin{pmatrix} 3 \\ -5 \\ 1.5 \end{pmatrix} \quad \text{ja} \quad \mathbf{w} = \begin{pmatrix} -5 \\ 1.5 \\ 3 \end{pmatrix}$$

esittävät eri vektoria. Yleisessä n -ulotteisessa vektorissa on n komponenttia, eli vektorin

$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

komponentit ovat $x_1, \dots, x_n$, ja ne ovat tässä luentomonisteessa aina reaalilukuja $\mathbb{R}$:ssä, ellei toisin sanota. Tämä vektori $\mathbf{v}$ on tarkalleen ottaen **pystyvektori** eli **sarakevektori**. Vastaava **vaakavektori** eli **rivivektori** olisi

$$\mathbf{v}^T = (x_1 \quad x_2 \quad \cdots \quad x_n).$$

Tällä hetkellä pysty- ja vaakavektorit eroavat toisistaan ainoastaan kirjoitustavassa, mutta palaamme matriisien yhteydessä asiaan uudestaan. Tarkoitamme vektorilla tyypillisesti aina pystyvektoria.

Vektoreiden tulkintatavat

Yksi vektorin käsitteen vahvuuksista on se, että vektoria voi ajatella monella eri tavalla:

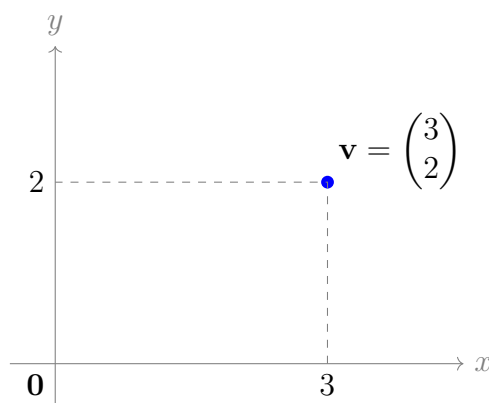
Vektori on lista: Vektoria voidaan ajatella listana informaatiota. Esimerkiksi jos tietyssä lähikaupassa on 2400 tuotetta myynnissä, niiden hinnat voisivat muodostaa 2400-ulotteisen vektorin. Tämä vektori voi olla tallennettuna kaupan tietokantaan, ja hintojen käsittely onnistuu tämän vektorin kautta.

Vektori on piste avaruudessa: Vektoria voidaan ajatella pisteenä n -ulotteisessa avaruudessa. Esimerkiksi kolmiulotteisen avaruuden jokainen piste määräytyy sen x-, y-, ja z-

koordinaatin avulla, joten jokainen kolmiulotteisen avaruuden piste (x, y, z) vastaa yhtä vektoria

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Tällä tavalla vektoreille voidaan antaa geometrinen merkitys, ja samalla tavalla myös yleisen n -ulotteisen avaruuden $\mathbb{R}^n$ pisteitä voidaan ajatella n -ulotteisina vektoreina (vaikka geometrinen hahmottaminen menee n :ssä ulottuvuudessa vaikeammaksi).



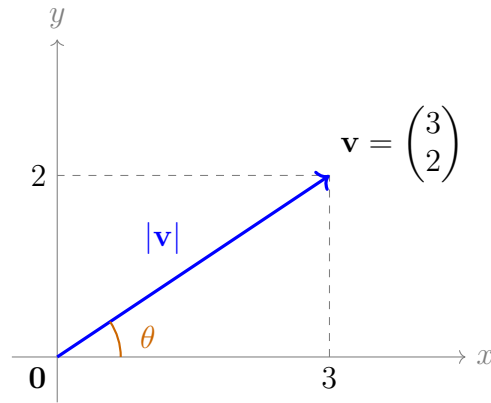
Eräs tärkeä piste $\mathbb{R}^n$:ssä on **origo**, eli piste $(0, 0, \dots, 0)$, jonka kaikki n koordinaattia ovat nollia. Tätä vastaava pystyvektori, jossa on n nollaa, tunnetaan myös **nollavektorina**

$$\mathbf{0}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

yleensä nollavektoria merkitään kuitenkin vain symbolilla **0**.

Vektori on nuoli, jolla on pituus ja suunta: Etenkin fysiikassa vektorit tulkitaan usein nuolina, jotka määräytyvät niiden pituuden ja suunnan avulla. Tämä tulkintatapa voidaan samaistaa edelliseen tapaan seuraavasti. Annettuun pisteeseen $\mathbb{R}^n$:ssä voidaan piirtää nuoli origosta, ja tämän nuolen pituus ja suunta määräytyy pisteestä yksikäsitteisesti. Tämä antaa erityisesti jokaiselle $\mathbb{R}^n$:n vektorille **pituuden** eli **normin**, joka voidaan myös laskea kaavasta

$$|\mathbf{v}| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}, \quad \text{missä } \mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$



Tämä tulkintatapa antaa vektoreille lisää geometrista merkitystä, ja se auttaa monesti hahmoittamaan vektoreiden ominaisuuksia ja käsittelyä.

Vektori on matemaattinen objekti: Abstraktein tulkintatapa on ajatella vektoria $\mathbf{v}$ yksinkertaisesti matemaattisena oliona, jolla on tiettyjä ominaisuuksia ja laskutoimituksia. Tässä kontekstissa vektori voi esittää paljon yleisempää asiaa kuin äärellistä listaa lukuja. Vektorit voidaan myös esitellä abstraktimmalla tavalla kuin tässä monisteessa matemaattisten **vektoriavaruuksien** teorian kautta. Emme käsittele vektoriavaruuksia tässä materiaalissa yleisesti, mutta suosittelemme niihin tutustumista niille, jotka haluavat syventää matemaattista ymmärrystään aiheesta.

Tutustutaan seuraavaksi vektorien peruslaskutoimituksiin.

Vektoreiden laskutoimitukset

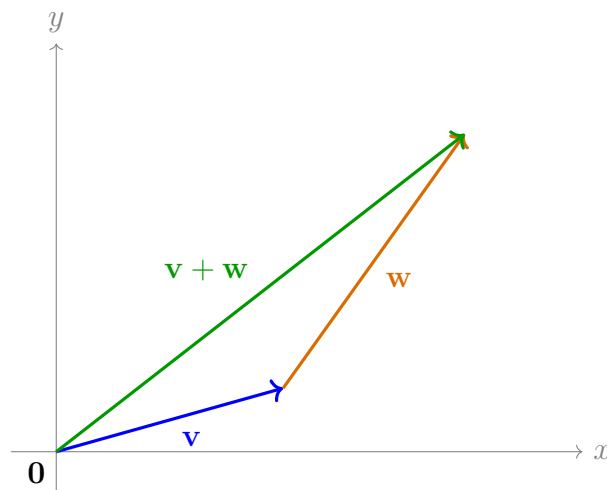
Vektoreiden hyöty matemaattisena käsitteenä ilmenee vielä paremmin erilaisten vektoreiden laskutoimitusten ja toimenpiteiden kautta.

Vektoreiden peruslaskutoimitukset ovat:

Yhteenlasku. Kaksi samankokoista vektoria $\mathbf{v}, \mathbf{w}$ voidaan laskea yhteen komponenteittain:

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \text{missä } \mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Tämä voidaan tulkita geometrisesti myös seuraavan kuvan osoittamalla tavalla.



Skalaarikertolasku. Mikäli a on jokin reaaliluku (eli *skalaari*), ja $\mathbf{v}$ on vektori, voidaan vektoria $\mathbf{v}$ kertoa luvulla a seuraavasti:

$$a\mathbf{v} = a \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{pmatrix}, \quad \text{missä } \mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Positiivisella skalaarilla $a > 0$ kertominen pitää vektorin suunnan samana, mutta kertoo sen pituutta luvulla a (mieti miksi). Negatiivisella skalaarilla $a < 0$ kertominen tekee melkein saman, mutta vaihtaa vektorin suunnan käänteiseksi. Vektorin kertominen nolllalla $a = 0$ antaa aina nollavektorin tulokseksi.

Vektoreiden laskutoimituksilla on samanlaiset laskusäännöt kuin reaaliluvuilla, kuten

- **Symmetrisyys:** $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.
- **Liitännäisyys:** $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.
- **Osittelulaki:** $a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w}$.

Nämä peruslaskutoimitukset määrittävät avaruuden $\mathbb{R}^n$ vektoreiden rakenteen. Tähän liittyvä tärkeä käsite on **lineaarikombinaatio** (tai **lineaariyhdistely**).

Määritelmä 1.1. Kahden vektorin $\mathbf{v}, \mathbf{w}$ **lineaarikombinaatio** on mikä tahansa vektori muotoa

$$a\mathbf{v} + b\mathbf{w}, \quad \text{missä } a, b \in \mathbb{R}.$$

Yleisemmin N vektorin $\mathbf{v}_1, \dots, \mathbf{v}_N$ **lineaarikombinaatio** on mikä tahansa vektori muotoa

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_N\mathbf{v}_N, \quad \text{missä } a_1, \dots, a_N \in \mathbb{R}.$$

Palaamme vektoreiden lineaarikombinaatioiden tarkempaan käsittelyyn osiossa Aliavaruudet.

Esimerkki 1.1. (Standardikanta). Avaruudessa $\mathbb{R}^n$ on hyödyllistä erotella joukosta seuraavat ns. **standardit kantavektorit**:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Nämä n vektoria muodostavat $\mathbb{R}^n$:n vektoreille ns. **kannan**, eli jokainen $\mathbb{R}^n$:n vektori voidaan esittää yksikäsitteisellä tavalla näiden vektoreiden lineaarikombinaationa. Tämä on helppo tarkistaa, sillä jokaiselle vektorille $\mathbf{v}$ voidaan laskea:

$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n.$$

Näiden kahden peruslaskutoimituksen lisäksi vektoreilla on monia muita hyödyllisiä operaatioita. Yksi tärkeimmistä on vektoreiden **pistetulo** eli **sisätulo**:

Määritelmä 1.2. Kahden vektorin **pistetulo** lasketaan kaavasta

$$\mathbf{v} \cdot \mathbf{w} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Pistetuloa merkitään myös $\langle \mathbf{v}, \mathbf{w} \rangle$, ja se antaa siis aina reaaliluvun tulokseksi.

Esimerkki 1.2. Vektorin pituus voidaan laskea pistetulon avulla kaavasta

$$|\mathbf{v}| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

Esimerkki 1.3. Eräässä kaupassa on 2400 tuotetta myynnissä. Tuotteiden hinnat ja kokonaismäärät on tallennettu vektoreihin

$$\mathbf{v}_{\text{hinnat}} = \begin{pmatrix} 1,99 \text{ e} \\ 5,99 \text{ e} \\ \vdots \\ 12,49 \text{ e} \end{pmatrix}, \quad \mathbf{w}_{\text{määrät}} = \begin{pmatrix} 12 \text{ kpl} \\ 0 \text{ kpl} \\ \vdots \\ 24 \text{ kpl} \end{pmatrix}.$$

Kaupan kaikkien tuotteiden yhteenlaskettu arvo voidaan nyt laskea suoraan pistetulosta

$$\text{Kokonaisarvo} = \mathbf{v}_{\text{hinnat}} \cdot \mathbf{w}_{\text{määrät}}.$$

Toinen geometrisesti merkittävä käsite on **vektoreiden välinen kulma**. Kaksi vektoria $\mathbf{v}$ ja $\mathbf{w}$ voidaan tulkita kahtena origosta lähtevänä nuolena. Näiden nuolien välillä on jokin kulma α , joka kulkee radiaaneissa välillä $[0, \pi)$, ja jonka voi aina selvittää kaavasta

$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}||\mathbf{w}| \cos \alpha.$$

Tarkista seuraavat ominaisuudet:

- Pistetulo on **symmetrinen**, eli pätee $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$.
- Pistetulo on **bilineaarinen**, eli jos $\mathbf{u}, \mathbf{v}, \mathbf{w}$ ovat vektoreita ja a skalaari, niin

$$(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w} \quad \text{ja} \quad (a\mathbf{v}) \cdot \mathbf{w} = a(\mathbf{v} \cdot \mathbf{w}),$$

sekä vastaavasti

$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \quad \text{ja} \quad \mathbf{v} \cdot (a\mathbf{w}) = a(\mathbf{v} \cdot \mathbf{w}).$$

- Vektorit $\mathbf{v}, \mathbf{w}$ ovat **kohtisuorassa** eli **ortogonaaliset**, jos niiden välinen kulma on $\pi/2$. Tämä tapahtuu tasan silloin, kun $\mathbf{v} \cdot \mathbf{w} = 0$.
- Vektorit $\mathbf{v}, \mathbf{w}$ ovat samansuuntaisia tasan silloin, kun $\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}||\mathbf{w}|$.
- Kahden vektorin välinen kulma ei riipu niiden pituuksista.

Matriisit

Siinä missä vektori tarjoavaa tavan asettaa useampi reaalityyppi yhteen listaan, matriisi koostuu kokonaisesta taulukosta lukuja. Esimerkiksi matriisista voisi olla vaikka

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 5 & 2 & -5 \end{pmatrix}. \quad (1.1)$$

Tässä matriisissa on kaksi **riviä** ja kolme **saraketta**, joten sitä sanotaan 2×3 -matriisiksi. Tätä merkitään $A \in \mathbb{R}^{2 \times 3}$, ja lukuja $2, 3$ kutsutaan myös matriisin A **dimensioiksi** tai sen **kooksi**. Matriisissa esiintyviä lukuja sanotaan sen **alkioiksi**. Yleistä $m \times n$ -matriisia voidaan merkitä

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Tällöin matriisin A kohdassa (i, j) , eli rivillä i ja sarakkeella j , löytyy alkio a_{ij} . Tätä voidaan lyhyesti merkitä myös $A = (a_{ij})$. Alkiota kohdassa (i, j) merkitään myös A_{ij}

Avaruuden $\mathbb{R}^n$ pystyvektorit ovat itse asiassa vain $n \times 1$ -matriiseja, ja vaakavektorit $1 \times n$ -matriiseja. Matriiseja voidaan ajatella myös listana vaaka- tai pystyvektoreita. Tutustutaan tähän esimerkin kautta:

Esimerkki 1.4. Kaavan (1.1) matriisi A voidaan tulkita myös seuraavalla kahdella eri tavalla:

Rivivektorit: Matriisi A koostuu kahdesta rivivektorista

$$\mathbf{R}_1 = (2 \quad 0 \quad -1) \quad \text{ja} \quad \mathbf{R}_2 = (5 \quad 2 \quad -5).$$

Tätä voidaan merkitä myös $A = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix}$.

Sarakevektorit: Matriisi A koostuu kolmesta sarakevektorista

$$\mathbf{S}_1 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \quad \mathbf{S}_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \text{ja} \quad \mathbf{S}_3 = \begin{pmatrix} -1 \\ -5 \end{pmatrix}.$$

Tätä voidaan merkitä myös $A = (\mathbf{S}_1 \quad \mathbf{S}_2 \quad \mathbf{S}_3)$.

Matriiseja voidaan myös laskea yhteen ja kertoa skalaareilla samalla tavalla kuin vektoreita, kunhan yhteenlaskettavien matriisien dimensiot ovat samat. Esimerkiksi:

$$\begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad -2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 2 \\ 6 & -2 \end{pmatrix}.$$

Matriisien yhteenlaskun suhteen tärkeä perusmatriisi on **nollamatriisi**, jonka alkiot ovat kaikki nollia. Esimerkiksi 2×3 -nollamatriisi on siis

$$0_{2 \times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Nollamatriisi on matriisien yhteenlaskun suhteen neutraalialkio, eli sen lisääminen ei muuta lisättävää matriisia.

Matriisien olennaisin peruslaskutoimitus on kuitenkin niiden **kertolasku**.

Matriisien kertolasku

Matriisien kertolasku on toiminto, jossa otetaan kaksi matriisia A ja B , ja muodostetaan kolmas matriisi AB . Kertolasku tapahtuu yhdistämällä jokainen **ensimmäisen matriisin** A **rivi** ja **toisen matriisin** B **sarake** seuraavasti.

Olkoon A $m \times n$ -matriisi ja B $n \times k$ -matriisi. Tällöin matriisitulo AB on hyvin määritelty, ja se tulee olemaan $m \times k$ -matriisi. Muistisääntö tälle on seuraava:

$$(m \times n)(n \times k) \rightarrow (m \times k).$$

Olennaista on, että matriiseissa A on yhtä monta saraketta, kuin B :ssä rivejä. Muuten matriisikertolaskua AB ei ole määritelty.

Uuden matriisin AB alkioit voidaan laskea yksitellen seuraavasti. Jos halutaan selvittää matriisin AB alkio kohdassa (i, j) , otetaan matriisin A rivi i , matriisin B sarake j , kerrotaan vastaavat luvut keskenään, ja summataan ne yhteen. Siis esimerkiksi

$$\begin{pmatrix} 2 & 0 & 2 \\ 1 & 4 & 2 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 0 & 2 \\ 6 & 3 \end{pmatrix} = \begin{pmatrix} 18 & 16 \\ 15 & 19 \end{pmatrix}, \quad 19 = 1 \cdot 5 + 4 \cdot 2 + 2 \cdot 3.$$

Yleisemmin, jos $A = (a_{ij})$ ja $B = (b_{ij})$, niin matriisin AB alkio kohdassa (i, j) saadaan kaavasta

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Matriisien kertolaskun laskeminen käsin on lähinnä työlästä kirjanpitoa, mutta käsitteellisesti matriisien kertolasku avaa niiden todelliset merkitykset ja sovellukset.

Kertolaskun suhteen tärkeä perusmatriisi on **identiteettimatriisi** eli **yksikkömatriisi**. Identiteettimatriisissa on aina sama määrä rivejä ja sarakkeita, eli se on ns. **neliömatriisi**. Yleinen $n \times n$ -identiteettimatriisi on muotoa

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Siis se on $n \times n$ -matriisi, jonka **diagonaalilla** eli (i, i) -kohdissa olevat alkioit ovat kaikki ykkösiä, ja muut alkioit nolliä. Identiteettimatriisia merkitään yleensä pelkästään symbolilla I , kun sen koko on selvä kontekstista.

Identiteettimatriisilla on se tärkeä ominaisuus, että sillä kertominen ei muuta kerrottavaa matriisia. Siis jos A on $m \times n$ -matriisi, niin

$$AI_n = A \quad \text{ja myös} \quad I_m A = A.$$

Huom!! Matriisien kertolaskun määritelmästä on melko suoraviivaista nähdä, että **matriiseja kerrottaessa järjestyksellä on väliä**. Esimerkiksi $m \times n$ -matriisin A ja $n \times k$ -matriisin B tulo AB on määritelty, mutta tulo BA ei ole määritelty, ellei $m = k$. Vaikka molemmat kertolaskut olisivatkin määriteltyjä, on tyypillistä, että tulot AB ja BA antavat täysin eri tuloksen.

Palaamme matriisin kertolaskun eri laskusääntöihin ja ominaisuuksiin pian uudestaan. Käydään kuitenkin heti pari uutta sovellusta, joita matriisikertolasku avaa.

Esimerkki 1.5. Yliopistokurssilla opiskelijan arvostelu perustuu harjoitustehtäviin, sekä väli- ja loppukokeen tuloksiin. Opiskelijoiden arvosanat on tallennettu matriisiin

$$M_{\text{arvosanat}} = \begin{pmatrix} 0.9 & 0.85 & 0.89 \\ 0.84 & 0.77 & 0.81 \\ \vdots & \vdots & \vdots \\ 0.64 & 0.82 & 0.93 \end{pmatrix},$$

missä jokainen rivi vastaa yhtä opiskelijaa, eli sarakkeiden luvut ilmaisevat opiskelijan saaman osuuden kokonaispisteistä harjoituksissa/välikokeessa/loppukokeessa. Jos kurssiarvosanan painotukset ovat

$$\mathbf{v}_{\text{painotus}} = \begin{pmatrix} 0.3 \\ 0.2 \\ 0.5 \end{pmatrix},$$

niin jokaisen opiskelijan lopullinen suoritus saadaan laskemalla tulo

$$M_{\text{arvosanat}} \mathbf{v}_{\text{painotus}} = \begin{pmatrix} 0.9 & 0.85 & 0.89 \\ 0.84 & 0.77 & 0.81 \\ \vdots & \vdots & \vdots \\ 0.64 & 0.82 & 0.93 \end{pmatrix} \begin{pmatrix} 0.3 \\ 0.2 \\ 0.5 \end{pmatrix} = \begin{pmatrix} 0.885 \\ 0.808 \\ \vdots \\ 0.821 \end{pmatrix}.$$

Esimerkki 1.6. Mustavalkoista kuvaa, jonka resoluutio on 1024×1024 pikseliä, vastaa samankokoinen neliömatriisi, esim.

$$A = \begin{pmatrix} 0.012 & 0.013 & \cdots & 0.434 \\ 0.014 & 0.016 & \cdots & 0.442 \\ \vdots & \vdots & \ddots & \vdots \\ 0.671 & 0.677 & \cdots & 0.96 \end{pmatrix},$$

missä jokainen alkio kuvaa vastaavan pikselin kirkkautta. Tutkitaan samankokoista matriisia

$$B = \frac{1}{4} \begin{pmatrix} 3 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 2 & 1 & \cdots & 0 \\ 0 & 0 & 1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 3 \end{pmatrix}.$$

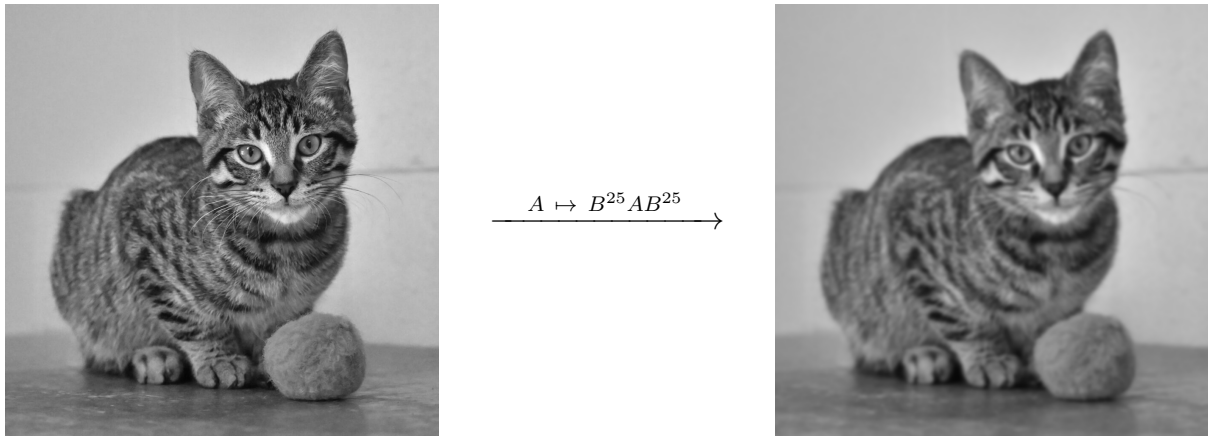
Tämän matriisin B avulla kuvaa A voidaan *sumentaa* seuraavasti. Matriisitulon $A \mapsto BA$ muodostaminen vastaa toimenpidettä, jossa jokaisen pikselin kirkkaus a_{ij} korvataan painotetulla keskiarvolla sen viereisten pikseleiden kirkkauksien kanssa:

$$a_{ij} \mapsto \frac{1}{4}(a_{(i-1)j} + 2a_{ij} + a_{(i+1)j}),$$

eli painotuksella $1 : 2 : 1$. Reunapisteitä painotetaan suhteessa $3 : 1$ viereiseen pikseliin.

Vastaavasti pystysuuntaan saman toimenpiteen voi toistaa kertomalla A :ta B :llä oikealta, eli $A \mapsto AB$.

Molempiin suuntiin sumentaminen onnistuu tekemällä nämä molemmat toimenpiteet, eli soveltamalla muunnosta $A \mapsto BAB$. Tätä muunnosta voidaan myös toistaa, ja sumentaa kuvaa A niin paljon kuin halutaan. Kuvassa esimerkki muunnoksen soveltamisesta 25 kertaa.



Kuva: Matriisimuunnoksen $A \mapsto BAB$ toistuva soveltaminen sumentaa kuvaa enemmän ja enemmän.

Myös monia muita kuvankäsittelytoimenpiteitä voidaan tuottaa matriisien avulla.

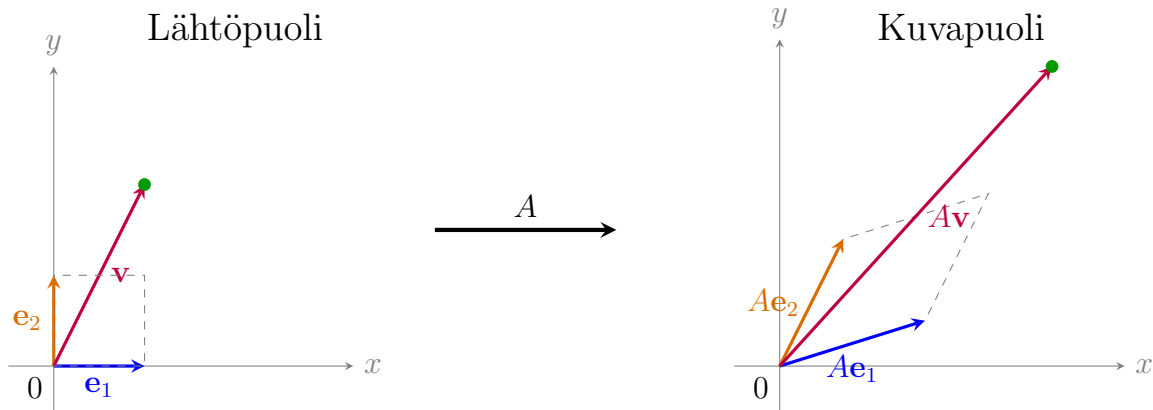
Matriisit lineaarikuvauksina

Myös matriiseja voidaan tulkita geometrisesti, ja tämä tulkinta selvenee, kun matriisilla kerrotaan vektoreita.

Olkoon A $m \times n$ -matriisi, ja $\mathbf{v}$ pystyvektori $\mathbb{R}^n$:ssä, eli matriisina tulkittuna $\mathbf{v}$:n dimensiot ovat $n \times 1$. Tällöin matriisikertolasku $A\mathbf{v}$ on määritelty, ja sen dimensiot ovat $m \times 1$, eli kyseessä on m -ulotteinen pystyvektori.

Siispä $m \times n$ -matriisilla A kertomista voidaan ajatella **operaationa**, joka ottaa syötteen n -ulotteisen vektorin, eli pisteen avaruudessa $\mathbb{R}^n$, ja tulostaa m -ulotteisen vektorin, eli pisteen $\mathbb{R}^m$:ssä. Tällöin matriisia A voidaan ajatella myös kuvauksena/funktiona

$$A : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \mathbf{v} \mapsto A\mathbf{v}.$$



Käytämme tässä luentomateriaalissa samaa symbolia esittämään sekä matriisia A , että sen määräämää lineaarikuvausta $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Täsmällisesti ottaen nämä ovat kaksi eri matemaattista objektia, jotka vastaavat toisiaan luonnollisella tavalla.

Esimerkki 1.7. Lineaarisenä kuvauksena tulkittuna identiteettimatriisi $I : \mathbb{R}^n \rightarrow \mathbb{R}^n$ on yksinkertaisesti avaruuden $\mathbb{R}^n$ **identiteettikuvaus**, eli kuvaus

$$I\mathbf{v} = \mathbf{v},$$

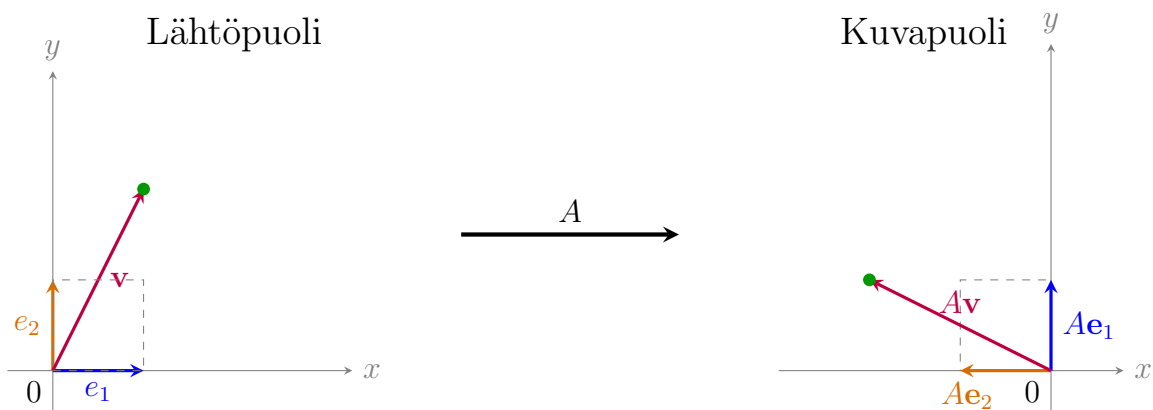
joka kuvaa jokaisen pisteen itselleen.

Esimerkki 1.8. Matriisi

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

vastaa kuvausta tasossa, eli $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Tarkalleen ottaen matriisi kuvaa tason pisteen (x, y) uudelle pisteelle kaavalla

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



Mitä tämä tarkoittaa geometrisesti? Kuvasta voidaan havaita, että matriisi A kiertää jokaista pistettä origon ympäri vastapäivään (eli positiiviseen suuntaan) tasan 90 astetta. Tällaista matriisia sanotaan **kiertomatriisiksi** tai **rotaatiomatriisiksi**, ja jokainen

origokeskinen kierto tasossa annetulla kulmalla θ voidaan esittää yksikäsitteisellä tavalla 2×2 -matriisina:

$$A_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Esimerkki 1.9. Matriisi

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

Vastaa kuvausta $A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$. Matriisi kuvaa kolmiulotteisen avaruuden pisteen (x, y, z) kaksiulotteisen tason pisteelle kaavalla

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ z \end{pmatrix}.$$

Geometrisesti tätä matriisia voi ajatella ns. **projektiona**. Se lähettää jokaisen avaruuden pisteen sen (y, z) -koordinaatteihin tasossa, eli litistää kolmiulotteisen avaruuden kahteen ulottuvuuteen.

Matriisikertolaskun laksusäännöt

Matriisikertolaskulla on myös monta tärkeää laksusääntöä, jotka avaavat matriisien geometrista tulkintaa vielä enemmän. Tutustutaan seuraavaksi näihin laksusääntöihin ja niiden geometrisiin tulkintoihin. Kaikki seuraavaksi esiteltävät kaavat on mahdollista todistaa oikeiksi suoraan määritelmistä pienellä työllä, mutta pidämme esityksen tiiviinä ja sivuutamme todistukset.

Kertolaskun liitännäislaki. Jos A on $m \times n$ -matriisi, B on $n \times k$ -matriisi, ja C on $k \times l$ matriisi, niin pätee

$$(AB)C = A(BC).$$

Koska sulkujen paikalla ei ole väliä, voidaan yllä olevaa tuloa merkitä yksinkertaisesti ABC . Erityisesti jos C :n paikalle laitetaan on $k \times 1$ -vektori v , niin huomataan, että

$$(AB)v = A(Bv).$$

Tämä antaa myös kahden matriisin kertolaskulle **geometrisen tulkinnan**: Tulo AB vastaa yhdistettyä kuvausta $A \circ B : \mathbb{R}^k \rightarrow \mathbb{R}^m$, joka saadaan käyttämällä syötteen v ensin kuvausta $B : \mathbb{R}^k \rightarrow \mathbb{R}^n$, ja sitten tulokseen Bv kuvausta $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Kertolaskun lineaarisuus. Mikäli A ja B ovat molemmat $m \times n$ -matriiseja, ja C ja D $n \times k$ -matriiseja, pätevät molemmat osittelulait

$$(A + B)C = AC + BC, \quad \text{ja} \quad A(C + D) = AC + AD.$$

Lisäksi jos $t \in \mathbb{R}$ on skalaari, pätee

$$A(tC) = (tA)C = t(AC).$$

Erityisesti seuraa, että matriisilla kertominen pelaa hyvin yhteen vektoreiden lineaarikombinaatioiden kanssa. Mikäli $v_1, \dots, v_N$ ovat $n \times 1$ -vektoreita ja $q_1, \dots, q_N$ ovat skalaareja, niin

$$A(q_1 v_1 + \dots + q_N v_N) = q_1 A v_1 + \dots + q_N A v_N.$$

Kun matriisia A ajatellaan kuvauksena $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$, niin ylläoleva ominaisuus tunnetaan **linearisuutena**. Jokainen matriisi määrittelee siis lineaarisen geometrisen kuvauksen, eli **lineaarikuvauksen**.

Käänteinen väite pätee myös: Jokainen kuvaus $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, joka on lineaarinen eli toteuttaa ominaisuuden

$$F(q_1 \mathbf{v}_1 + \cdots + q_N \mathbf{v}_N) = q_1 F(\mathbf{v}_1) + \cdots + q_N F(\mathbf{v}_N)$$

kaikilla mahdollisilla vektoreiden v_i ja skalaarien q_i valinnoilla, määrittyy eräästä $m \times n$ -matriisista $A = (a_{ij})$. (Todistus harjoituksissa)

Tarinan opetus: Matriisit voidaan tulkita geometrisesti täsmälleen lineaarisina kuvauksina $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Tällaisiin lineaarikuvauksiin kuuluvat mm. erilaiset venytykset, kierrot, peilaukset, ja niiden yhdistelmät.

Matriisien geometrisen tulkinnan ansiosta niillä on laajasti sovelluksia tietokonegrafiikassa, robotiikassa, navigoinnissa (ilmailu, avaruus, yms.), ja monilla muilla aloilla, joissa käsitellään koordinaatteja ja niiden välisiä muunnoksia.

Lineaarikuvaukset ovat kuitenkin vain yksi tapa hyödyntää matriiseja, ja esittelemme tulevaisuudessa kappaleissa lisää matriisien ominaisuuksia sekä sovelluksia. Tutkitaan vielä esimerkkien kautta pari uutta tulkintaa matriisien kertolaskulle.

Esimerkki 1.10. (Markovin prosessit). Eräässä kylässä sää on päivittäin joko aurinkoista tai sateista. Seuraavan päivän sää riippuu aina edellisen päivän säästä seuraavalla tavalla:

- Jos tänään paistoi aurinko, huomenna auringonpaisteen todennäköisyys on 0.8 ja sateen 0.2.
- Jos tänään satoi, huomenna auringonpaisteen todennäköisyys on 0.4 ja sateen 0.6.

Ilmaistaan nyt päivittäinen säätilanne vektorina

$$\mathbf{p} = \begin{pmatrix} \text{auringonpaisteen todennäköisyys} \\ \text{sateen todennäköisyys} \end{pmatrix}.$$

Mikäli tänään paistaa aurinko, tiedetään, että tämän päivän todennäköisyysjakauma on

$$\mathbf{p} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Huomisen todennäköisyysjakauma saadaan aina kertomalla muunnosmatriisilla

$$M = \begin{pmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{pmatrix}.$$

Eli huomisen todennäköisyysjakauma olisi

$$M\mathbf{p} = \begin{pmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}.$$

Viikon päästä jakauman voi laskea kertomalla matriisilla M seitsemän kertaa:

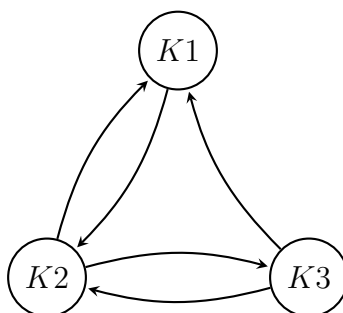
$$M(M(M(M(M(M(M\mathbf{p})))))) = M^7\mathbf{p} = \begin{pmatrix} 0.6672128 \\ 0.3327872 \end{pmatrix}.$$

Tämän toimenpiteen toistaminen suppenee (varsin nopeasti) vektoriin $\mathbf{p}_\infty = \begin{pmatrix} 0.66\dots \\ 0.33\dots \end{pmatrix}$, joka kuvaa sitä, että kylässä satunnaisesti valittuna päivänä vuodesta todennäköisyys auringonpaisteelle on 66.6... % ja sateelle 33.3... %.

Esimerkki 1.11. (Vierusmatriisi). Erään yhtiön lentokoneet liikennöivät kolmen kaupungin välillä. Lentoratoja kuvaa ns. **vierusmatriisi**, 3×3 -matriisi

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

missä matriisin alkio kohdassa (i, j) on yksi, jos kaupungista i on kaupunkiin j kulkeva lentoyhteys, ja 0 muulloin.



Kuva: Kenttien väliset lentoradat.

Tutkitaan, mitä matriisi M^2 kuvaa:

$$M^2 = MM = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Hieman tätä laskua tutkimalla huomataan, että matriisin M^2 alkioista voidaan lukea, montako yhden vaihdon reittiä löytyy kaupungista i kaupunkiin j :

Esimerkiksi kaupungista 1 pääsee kaupunkiin 3 yhdellä vaihdolla vain yhtä eri lentoreittiä pitkin – kaupungin 2 kautta. Kertolaskussa tämä ilmeni siinä, että matriisin M alkiot kohdissa $(1, 2)$ ja $(2, 3)$ ovat molemmat 1, ja muita vastaavia pareja rivillä 1 ja sarakkeella 3 ei ollut. Korostaen:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \text{sillä } 1 = 0 \cdot 0 + 1 \cdot 1 + 0 \cdot 0$$

Vastaavasti kahden vaihdon lentoreittejä voi laskea matriisilla M^3 , ja niin edespäin. Yleisesti ottaen matemaattisten **verkkojen** teoriaa voidaan tutkia myös matriisien avulla, ja matriisit yksinkertaistavat monen verkoissa esiintyvien suureiden laskemista.

Edellisen kahden esimerkin opetus on, että matriiseilla voidaan myös kuvata erilaisia muuttuvia systeemejä, kuten satunnaisesti muuttuvan järjestelmän kehitystä ja polkujen määriä matemaattisessa verkossa.

Transpoosi

Vielä yksi hyödyllinen matriisioperaatio on **transpoosi**, joka määritellään seuraavasti. Mikäli $A = (a_{ij})$ on $m \times n$ -matriisi, niin sen transpoosi on $n \times m$ -matriisi $A^T = (a_{ji})$. Käytännössä tämä näyttää tältä:

$$A = \begin{pmatrix} 3 & 2 & 0 \\ -1 & -3 & 5 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 3 & -1 \\ 2 & -3 \\ 5 & 0 \end{pmatrix}.$$

Transpoosi siis vaihtaa rivit sarakkeiksi, ja päinvastoin. Pystyvektorin transpoosi on aina vastaava vaakavektori, ja tästä saadaan pystyvektoreiden pistetulolle myös kaava

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w} = \mathbf{w}^T \mathbf{v},$$

missä kaksi jälkimmäisintä lauseketta tulkitaan matriisitulona.

Neliömatriisia A sanotaan **symmetriseksi**, mikäli $A^T = A$. Siis esim. matriisi

$$A = \begin{pmatrix} 1 & 0 & 4 \\ 0 & -2 & 6 \\ 4 & 6 & 0 \end{pmatrix}$$

on symmetrinen. Transpoosin kaksi peruslaskukaavaa ovat

$$\boxed{(A + B)^T = A^T + B^T \quad \text{ja} \quad (AB)^T = B^T A^T.} \quad (1.2)$$

Lisäksi transpoosilla on tärkeä yhteys vektorien pistetuloon

$$\boxed{\mathbf{v} \cdot (A\mathbf{w}) = (A^T \mathbf{v}) \cdot \mathbf{w} \quad \text{kaikilla vektoreilla } \mathbf{v} \in \mathbb{R}^m, \mathbf{w} \in \mathbb{R}^n.} \quad (1.3)$$

Matriisitulon rivi- ja sarakekehitemä

Matriisituloa voidaan toisinaan yksinkertaistaa jakamalla kerrottava(t) matriisi(t) pienempiin palasiin. Yksi tärkeä esimerkki tästä on matriisin ja vektorin tulo, joka voidaan esittää näppärästi rivi- ja sarakevektoreiden avulla. Mikäli A on $m \times n$ -matriisi ilmaistuna muodoissa

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \vdots \\ \mathbf{R}_m \end{pmatrix} = (\mathbf{S}_1 \quad \mathbf{S}_2 \quad \cdots \quad \mathbf{S}_n),$$

niin sen tulo mielivaltaisen $n \times 1$ -vektorin $\mathbf{v} = (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n)^T$ kanssa voidaan laskea myös kaavoista

$$\boxed{A\mathbf{v} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \vdots \\ \mathbf{R}_m \end{pmatrix} \mathbf{v} = \begin{pmatrix} \mathbf{R}_1 \mathbf{v} \\ \mathbf{R}_2 \mathbf{v} \\ \vdots \\ \mathbf{R}_m \mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_1^T \cdot \mathbf{v} \\ \mathbf{R}_2^T \cdot \mathbf{v} \\ \vdots \\ \mathbf{R}_m^T \cdot \mathbf{v} \end{pmatrix}}$$

ja

$$A\mathbf{v} = (\mathbf{S}_1 \quad \mathbf{S}_2 \quad \cdots \quad \mathbf{S}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{S}_1 + x_2\mathbf{S}_2 + \cdots + x_n\mathbf{S}_n$$

Vielä yleisemmin, on usein hyödyllistä jakaa matriisi pienempiin palasiin eli **lohkoihin**. Esimerkiksi 3×4 -matriisi voitaisiin jakaa seuraavasti

$$A = \left(\begin{array}{c|cc} 1 & 0 & 3 & 1 \\ -2 & -1 & 6 & 5 \\ 0 & -1 & 2 & -1 \end{array} \right).$$

Lohkoihin jaettaessa matriisi pilkotaan siis pienempiin osiin, joista jokainen osa itsessään esittää matriisia. Jos kaksi matriisia A ja B on jaettu sopivan kokoisiin lohkoihin, voidaan matriisitulo laskea myös kaavalla

$$AB = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} = \begin{pmatrix} A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4 \\ A_3B_1 + A_4B_3 & A_3B_2 + A_4B_4 \end{pmatrix},$$

mikäli kaikki tässä esiintyvät lohkojen matriisitulot ovat määriteltyjä, eli lohkot on valittu sopivan kokoisiksi, että matriisitulot onnistuvat.

Esimerkki 1.12. (Yksinkertainen hakualgoritmi). Tietokannassa on 95000 eri tietuetta. Jokaista tietuetta vastaa yksi 500×1 -vektori $\mathbf{v}_i$, ja niistä muodostetaan normalisoidut vektorit $\mathbf{u}_i = \mathbf{v}_i/|\mathbf{v}_i|$. Vektori $\mathbf{u}_i$ on siis se ykkösen pituinen vektori, joka on samansuuntainen kuin vektori $\mathbf{v}_i$. Vektoreita $\mathbf{u}_i$ vastaavat vaakavektorit tallennetaan 95000×500 -matriisiin

$$M = \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_{95000}^T \end{pmatrix}.$$

Olkoon $\mathbf{h}$ hakusana, joka on myös normalisoitu ykkösen pituiseksi 500×1 -vektoriksi. Tällöin matriisitulon rivikehitelmä antaa

$$M\mathbf{h} = \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_{95000}^T \end{pmatrix} \mathbf{h} = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{h} \\ \vdots \\ \mathbf{u}_{95000} \cdot \mathbf{h} \end{pmatrix}.$$

Tuloksena saadun vektorin komponentit $\mathbf{u}_i \cdot \mathbf{h}$ vastaavat vektoreiden $\mathbf{u}_i$ ja $\mathbf{h}$ välisen kulman θ_i kosinia, sillä normalisoinnin takia

$$\mathbf{u}_i \cdot \mathbf{h} = |\mathbf{u}_i||\mathbf{h}| \cos \theta_i = \cos \theta_i.$$

Mitä lähempänä tämä kosini on ykköstä, sitä lähempänä vektorit $\mathbf{u}_i$ ja $\mathbf{h}$ ovat toisiaan. Nyt jos vektorin $M\mathbf{h}$ komponentit järjestetään suuruusjärjestyksessä, niin voidaan nähdä, mitkä tietueet parhaiten vastaavat hakusanaa $\mathbf{h}$.

Tämä menetelmä on varsin alkeellinen, mutta myös hienostuneemmat hakualgoritmit käyttävät paljon matriisilaskentaa. Esimerkiksi Googlen asema suosituimpana hakukoneena pohjautuu sen 2000-luvun alussa kehittämään algoritmiin *PageRank*, joka soveltaa matriiseja vahvasti. Lisää tästä myöhemmin.

Bonus: Kompleksiluvut matriiseissa

Tässä materiaalissa keskitymme matriiseihin, joiden alkiot ovat kaikki reaalilukuja. Sekä sovelluksia, että matriisien täyttää teoriaa varten on kuitenkin tarve esitellä myös **kompleksiluvut** ja tutkia myös matriiseja, joiden alkiot ovat näitä kompleksilukuja.

Kompleksiluvut ovat jälleen yksi uusi tapa ajatella tason pisteitä. Tason pistettä (x, y) vastaa kompleksiluku

$$x + iy,$$

missä i on ns. **imaginääriyksikkö**, joka vastaa y -akselin pistettä $(0, 1)$. Kompleksiluvun $z = x + iy$ koordinaattia x kutsutaan sen **reaaliosaksi** $\operatorname{Re}(z)$ ja koordinaattia y sen **imaginääriosaksi** $\operatorname{Im}(z)$.

Kompleksiluvut elävät siis tasossa $\mathbb{R}^2$, jota kutsutaan tässä kontekstissa myös **kompleksitasoksi**, ja sitä merkitään symbolilla $\mathbb{C}$. Jokaista reaalilukua x voidaan ajatella kompleksilukuna $x + i0$, ja kompleksitasoa usein ajatellaan reaaliakselin laajennuksena.

Kompleksilukuja voi laskea yhteen samalla tavalla kuin vektoreita:

$$(x + iy) + (a + ib) = (x + a) + i(y + b).$$

Kompleksilukuja voidaan kuitenkin kertoa keskenään samoilla säännöillä kuin reaalilukuja, kun asetetaan lisäehto $i^2 = i * i = -1$. Tällöin yleinen kertolasku menee kaavalla

$$\begin{aligned}(x + iy)(a + ib) &= xa + x(ib) + (iy)a + (iy)(ib) \\ &= xa + i(xb + ya) + i^2yb \\ &= xa - yb + i(xb + ya).\end{aligned}$$

Kuten vektoreita, myös kompleksilukuja voi ajatella niiden **pituuden** eli **modulin**

$$|z| = |x + iy| = \sqrt{x^2 + y^2}$$

sekä suunnan avulla. Tasossa suunta määräytyy täysin kulman suhteen, joka kompleksilukuun z piirretty jana origosta muodostaa reaaliakselin suhteen. Tätä kulmaa kutsutaan kompleksiluvun **argumentiksi** ja merkitään $\arg z$.

Kompleksilukujen kertolaskulla on yksinkertainen geometrinen tulkinta: Kahden kompleksiluvun z_1 ja z_2 kertolaskussa lukujen z_1 ja z_2 modudit kerrotaan keskenään, ja argumentit summataan. Siis

$$\boxed{|z_1 z_2| = |z_1| |z_2|, \quad \arg(z_1 z_2) = \arg z_1 + \arg z_2.}$$

Kompleksiluvun $z = x + iy$ **liittoluku** eli **kongugaatti** on

$$\bar{z} = x - iy.$$

Liittoluvun ottaminen vastaa kompleksitasossa peilausta x -akselin suhteen. Tällöin pätee myös, että $|\bar{z}| = |z|$ sekä $\arg \bar{z} = -\arg z$. Liittoluvun hyödyllinen ominaisuus on, että kompleksiluvun moduli saadaan myös kaavasta

$$|z|^2 = x^2 + y^2 = (x + iy)(x - iy) = z\bar{z}.$$

Kompleksiset matriisit

Myös kompleksilukuja voi asettaa matriiseihin samaan tapaan kuin reaalityyppisiä lukuja. Jos A on $m \times n$ -matriisi jonka alkiot ovat kompleksilukuja, merkitään $A \in \mathbb{C}^{m \times n}$. Esimerkiksi

$$A = \begin{pmatrix} 1+i & -1 & i \\ 1-2i & -i & 3i \end{pmatrix}$$

on kompleksinen 2×3 -matriisi.

Kompleksimatriisien kertolasku lasketaan aivan samalla tavalla kuin reaalisten matriisien, mutta toki alkioiden välisiä kertolaskuja laskettaessa käytetään kompleksilukujen kertolaskua.

Ainoa oleellinen muutos on transpoosia laskiessa. Kompleksisten matriisien yhteydessä sovelletaan transpoosin sijaan lähes aina **konjugaattitranspoosia**. Konjugaattitranspoosissa annetun matriisin $A = (a_{ij})$ rivit ja sarakkeet vaihdetaan jälleen, mutta lisäksi jokainen alkio korvataan sen liittoluvulla. Sitä merkitään $A^* = (\overline{a_{ji}})$. Esim. yllä määritellyn matriisin A konjugaattitranspoosi on

$$A^* = \begin{pmatrix} 1-i & 1+2i \\ -1 & i \\ -i & -3i \end{pmatrix}.$$

Syy, miksi tavallinen transpoosi ei kelpaa, piilee kompleksisten vektorien pistetulossa. Kahden kompleksilukuja sisältävän vektorin pistetulo määritelläänkin kaavalla

$$\langle \mathbf{v}, \mathbf{w} \rangle = v_1 \overline{w_1} + \cdots + v_n \overline{w_n}, \quad \text{missä } \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}.$$

Jälkimmäisen vektorin komponenteista on otettu liittoluku, sillä tämä takaa, että pistetulosta saadaan edelleen vektorin pituus kaavalla

$$\langle \mathbf{v}, \mathbf{v} \rangle = |\mathbf{v}|^2.$$

Nyt konjugaattitranspoosille pätee aina kaava

$$\langle A\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, A^*\mathbf{w} \rangle.$$

2. Aliavaruudet

Tässä luvussa vahvistamme ymmärrystä siitä, miten matriisit käyttäytyvät lineaarikuvauksina, ja mitä avaruuden osajoukkoja niiden käsittelyyn liittyy.

Aliavaruuden virittäminen

Aloitetaan esimerkkitilanteesta, jossa meillä on käytössä kaksi vektoria

$$\mathbf{v} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \quad \text{ja} \quad \mathbf{w} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

jotka elävät siis kolmiulotteisessa avaruudessa $\mathbb{R}^3$. Näiden vektoreiden **virittämä** joukko eli **span** koostuu kaikista niiden lineaarikombinaatioista, ja sitä merkitään

$$\text{span}(\mathbf{v}, \mathbf{w}) = \{a\mathbf{v} + b\mathbf{w} : a, b \in \mathbb{R}\}.$$

Koska tämä joukko varmasti sisältää myös kaikki jäsentensä lineaarikombinaatiot, on se esimerkki ns. **aliavaruudesta** eli **vektorialiavaruudesta**.

Määritelmä 2.1. Avaruuden $\mathbb{R}^n$ epätühjä osajoukko V on **aliavaruus**, mikäli kaikki sen vektoreiden $\mathbf{v}_1, \dots, \mathbf{v}_N \in V$ lineaarikombinaatiot

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_N\mathbf{v}_N, \quad a_1, \dots, a_N \in \mathbb{R}$$

kuuluvat edelleen osajoukkoon V .

Geometrisesti aliavaruus on aina joko suora, taso, kolmiulotteinen avaruus, tai niin edespäin. Myös triviaali aliavaruus $V = \{\mathbf{0}\}$, joka sisältää pelkän origon/nollavektorin, on tärkeä erikoistapaus. Jokaisen aliavaruuden täytyy aina sisältää nollavektori (koska yllä voidaan valita $a_1, \dots, a_N = 0$).

Edellä annetut vektorit $\mathbf{v}, \mathbf{w}$ virittävät täsmälleen xz -tason, sillä jokainen xz -tason piste voidaan lausua muodossa

$$\begin{pmatrix} x \\ 0 \\ z \end{pmatrix} = \begin{pmatrix} -(z-x) + z \\ 0 \\ z \end{pmatrix} = (z-x) \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = (z-x)\mathbf{v} + z\mathbf{w}.$$

Taso on kaksiulotteinen, ja käy järkeen, että kaksi vektoria $\mathbf{v}, \mathbf{w}$ virittäisivät kaksiulotteisen objektin. Tämä sääntö ei kuitenkaan päde ihan aina, sillä esimerkiksi vektorit

$$\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \quad \text{ja} \quad \mathbf{p} = \begin{pmatrix} -2 \\ 2 \\ -2 \end{pmatrix}$$

virittävät ainoastaan suoran, sillä molemmat vektorit ovat tällä suoralla, ja toinen saadaan aina toisesta kertomalla sopivalla skalaarilla. Ongelma on, että vektorit $\mathbf{u}$ ja $\mathbf{p}$ ovat ns. **lineaarisesti riippuvia** toisistaan.

Määritelmä 2.2. Vektorit $\mathbf{v}_1, \dots, \mathbf{v}_N$ ovat **lineaarisesti riippuvia**, mikäli löytyy sellaiset reaalityyppiset $a_1, \dots, a_N$, jotka eivät kaikki ole nollia, mutta

$$a_1\mathbf{v}_1 + \dots + a_N\mathbf{v}_N = \mathbf{0}.$$

Muulloin sanotaan, että vektorit $\mathbf{v}_1, \dots, \mathbf{v}_N$ ovat **lineaarisesti riippumattomia**.

Lineaarinen riippuvuus ja -riippumattomuus ovat hyvin keskeisiä käsitteitä vektoreiden ja matriisien teoriassa, ja tämä käsite tulee myöhemmin olemaan tärkeässä roolissa monissa matriisien sovelluksissa.

Yleisesti: Aina jos vektorijoukkoon lisää jonkin niiden kanssa lineaarisesti riippuvan vektorin, niin tämän vektorijoukon virittämä aliavaruus ei muutu.

Sen sijaan N kappaletta **lineaarisesti riippumattomia** vektoreita $\mathbf{v}_1, \dots, \mathbf{v}_N$ määrittää aina aidosti N -ulotteisen aliavaruuden $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_N)$. Eli yksi tällainen vektori määrittää aina suoran, kaksi tason, ja niin edelleen.

Kanta

Sanotaan, että joukko vektoreita $\mathbf{v}_1, \dots, \mathbf{v}_N$ on avaruuden $\mathbb{R}^n$ (tai jonkin sen aliavaruuden V) **kanta**, mikäli nämä vektorit ovat lineaarisesti riippumattomia ja niiden virittämä joukko $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_N)$ on koko avaruus $\mathbb{R}^n$ (tai vastaavasti annettu aliavaruus V). Mikäli kyseessä on koko avaruus $\mathbb{R}^n$, täytyy pakosti olla $N = n$.

Aiemmin totesimme jo, että standardikantavektorit $\mathbf{e}_1, \dots, \mathbf{e}_n$ (kts. Esimerkki 1.1) määrittävät yhden kannan avaruudelle $\mathbb{R}^n$. Avaruudella (tai aliavaruudella) voi kuitenkin olla useampia kantoja. Esimerkiksi mitkä tahansa kaksi vektoria $\mathbf{v}, \mathbf{w}$ tasossa $\mathbb{R}^2$ määrittävät tason kannan, kunhan $\mathbf{v}$ ja $\mathbf{w}$ eivät ole saman- tai vastakkaissuuntaisia.

Kanta on loistava käsite aliavaruuksien rakenteen ymmärtämiseen. Emme tässä käy yksityiskohtaisia todistuksia seuraaville väitteille, mutta kehotamme lukijaa pohtimaan niitä omassa rauhassa:

- Jokaiselle aliavaruudelle voidaan löytää kanta.
- Jos aliavaruudella V on kaksi eri kantaa, on niissä kuitenkin sama määrä vektoreita. Tätä lukumäärää sanotaan aliavaruuden **dimensioksi** tai **ulottuvuudeksi** (suora = 1, taso = 2, jne.), ja sitä merkitään $\dim V$. Triviaalin aliavaruuden $\{\mathbf{0}\}$ dimensio on nolla.
- Jokaisen k -ulotteisen aliavaruuden V kantaan voidaan lisätä $n - k$ vektoria niin, että ne muodostavat yhdessä avaruuden $\mathbb{R}^n$ kannan.
- Jos vektorit $\mathbf{v}_1, \dots, \mathbf{v}_N$ virittävät aliavaruuden V , niin $m \geq \dim V$. Lisäksi näistä vektoreista voidaan poistaa $N - \dim V$ kappaletta niin, että jäljelle jää jokin V :n kanta.

Matriisin ydin ja kuva

Jokaiselle $m \times n$ -matriisille A on mahdollista määritellä kaksi tärkeää joukkoa. Nämä ovat matriisin A **ydin** eli **nolla-avaruus** (eng. *kernel*)

$$\text{Ker } A = \{\mathbf{v} \in \mathbb{R}^n : A\mathbf{v} = \mathbf{0}\}.$$

Ydin $\text{Ker } A$ koostuu siis niistä avaruuden $\mathbb{R}^n$ vektoreista, jotka matriisi A kuvaa nollavektoriksi $\mathbb{R}^m$:ssä. Toinen tärkeä joukko on matriisin A **kuva** eli **kuva-avaruus** (eng. *image*)

$$\text{Im } A = \{A\mathbf{v} \in \mathbb{R}^m : \mathbf{v} \in \mathbb{R}^n\}.$$

Kuva $\text{Im } A$ siis koostuu kaikista niistä vektoreista avaruudessa $\mathbb{R}^m$, jotka saadaan jonkun vektorin $\mathbf{v} \in \mathbb{R}^n$ kuvavektorina $A\mathbf{v}$ kuvauksessa $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Kuva-avaruus voidaan myös rakentaa seuraavasti:

Matriisin kuva-avaruus on sen sarakkeiden $A = (\mathbf{S}_1 \ \mathbf{S}_2 \ \cdots \ \mathbf{S}_n)$ virittämä aliavaruus $\text{span}(\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_n)$.

Tämä siksi, että matriisitulon sarakkekehityksen avulla

$$A\mathbf{v} = (\mathbf{S}_1 \ \mathbf{S}_2 \ \cdots \ \mathbf{S}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{S}_1 + x_2\mathbf{S}_2 + \cdots + x_n\mathbf{S}_n,$$

eli matriisin kertominen vektorilla $\mathbf{v} = (x_1 \ \cdots \ x_n)^T$ antaa saman tuloksen, kuin jos sarakkeista ottaa lineaarikombinaation kertoimilla x_i . Kuva-avaruus on siis erityisesti $\mathbb{R}^m$:n aliavaruus, ja sen ulottuvuuden määrä tunnetaan matriisin A **asteena** (myös **säännöllisyysaste** tai **rangi**)

$$\text{rank } A = \dim \text{Im } A.$$

On myös varsin suoraviivaista todistaa, että matriisin ydin muodostaa aliavaruuden $\mathbb{R}^n$:ssä.

Ytimen ja kuvan rooli selviää pian, kun tutkimme, miten lineaariset yhtälöryhmät ratkaistaan matriisien avulla. Käsitteellisesti ne auttavat myös suuresti ymmärtämään matriisien yleistä käyttäytymistä, kuten toivottavasti seuraava esimerkki hieman havainnollistaa.

Esimerkki 2.1. Tutkitaan lineaarikuvausta $A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, jonka matriisi on

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

Tällöin A kuvaa kolmiulotteisia vektoreita kaavalla

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + z \\ y + z \end{pmatrix}.$$

Matriisin ydin $\text{Ker } A$ muodostuu niistä vektoreista, joilla $x + z = 0 = y + z$. Siis yleinen ytimen vektori on muotoa

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -z \\ -z \\ z \end{pmatrix} = z \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

Täten ydin $\text{Ker } A$ on yksiulotteinen. Vastaavasti matriisin kuva $\text{Im } A$ on kaksiulotteinen, sillä jokainen tason $\mathbb{R}^2$ piste saadaan kuvapisteenä esimerkiksi vektorista

$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Siis matriisin aste on kaksi. Matriisi A tiputtaa siis kolmiulotteisesta avaruudesta yhden ulottuvuuden pois (tämä ulottuvuus jää näkyviin sen ytimeen), ja lopputuloksena on kaksiulotteinen kuvajoukko $\text{Im } A$.

Seuraavaksi vielä yksi perusesimerkki matriisien sovelluksista, tällä kertaa differentiaaliyhtälöiden numeerisessa ratkaisemisessa.

Esimerkki 2.2. (Differentiaaliyhtälön ratkaisuaapproksimaatio). Harmonisen värähtelijän liikettä kuvaa lineaarinen toisen asteen differentiaaliyhtälö, kuten

$$x'' - 2x' + 3x = 0,$$

jossa ratkaistava funktio $x(t)$ kuvaa värähtelijän sijaintia. Tämän yhtälön voi ratkaista numeerisesti seuraavalla tavalla. Määritellään ensin $y = x'$, jolloin yhtälön voi esittää kahden yhtälön systeeminä

$$\begin{cases} x' = y \\ y' = -3x + 2y \end{cases}$$

Ratkaistaan yhtälö numeerisesti alkuarvolla $x(0) = x_0$ ja $x'(0) = y(0) = y_0$. Olkoon aika-askel pieni luku Δt , ja määritellään $t_n = n\Delta t$, $x_n = x(t_n)$ ja $y_n = y(t_n)$. Tällöin

$$\begin{cases} x_{n+1} \approx x_n + x'(t_n)\Delta t = x_n + y_n\Delta t \\ y_{n+1} \approx y_n + y'(t_n)\Delta t = y_n + (-3x_n + 2y_n)\Delta t \end{cases}$$

Jos $\mathbf{v}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$, tämä voidaan ilmaista yksinkertaisessa vektorimuodossa

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \Delta t \begin{pmatrix} 0 & 1 \\ -3 & 2 \end{pmatrix} \mathbf{v}_n,$$

josta ratkaisuaapproksimaatio saadaan laskettua. Samaa ideaa voidaan soveltaa myös monimutkaisempiin tapauksiin, kuten useamman muuttujan differentiaaliyhtälösystemeihin.

3. Lineaariset yhtälöryhmät

Tutkitaan seuraavaa käytännön ongelmaa:

Tehtävä: Kaupassa omena maksaa **kaksi euroa**, ja banaani **yhden euron**. Haluat ostaa yhteensä **viisi hedelmää**, ja käyttää niihin **seitsemän euroa**.

Tämän tehtävän voi ratkaista kirjoittamalla se yhtälöryhmänä

$$\begin{cases} x + y = 5 \\ 2x + y = 7 \end{cases},$$

missä x ja y vastaavat omenoiden ja banaanien määrää. Tästä ratkaisu on varsin helppo löytää, mutta kiinnitetään ennemminkin huomiota siihen, että se voidaan ilmaista myös matriisitulon avulla:

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix},$$

eli muodossa

$$A\mathbf{v} = \mathbf{b}, \quad \text{missä } A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}.$$

Yleisesti ottaen, jos A on $m \times n$ -matriisi, $\mathbf{b}$ on $m \times 1$ -vektori, ja $\mathbf{v}$ on tuntematon $n \times 1$ -vektori, niin yhtälö

$$A\mathbf{v} = \mathbf{b}$$

kuvaa yleistä lineaarista yhtälöryhmää, jossa on n tuntematonta, ja m yhtälöä.

Tämä avaa matriiseille jälleen uuden tärkeän sovelluksen. Esittelemme ensin, mitä uutta teoreettista ymmärrystä matriisit avaavat lineaarisiin yhtälöryhmiin, ja palataan sitten yhtälöryhmien ratkaisemiseen käytännössä.

Yhtälöryhmän ratkaisujoukko

Tavoitteemme on ymmärtää, milloin lineaarisella yhtälöryhmällä

$$A\mathbf{v} = \mathbf{b}$$

on ratkaisu, ja voiko ratkaisuja olla useampia. Tässä edelleen $A \in \mathbb{R}^{m \times n}$ ja $\mathbf{b} \in \mathbb{R}^m$ ovat kiinnitettyjä, ja $\mathbf{v} \in \mathbb{R}^n$ on tuntematon.

Tehdään ensin kriittinen huomio:

$\text{Yhtälöryhmällä } A\mathbf{v} = \mathbf{b} \text{ on ratkaisu } \mathbf{v} \text{ tasan silloin, kun } \mathbf{b} \in \text{Im } A.$

Tämä huomio seuraa suoraan kuvan $\text{Im } A$ määritelmästä, sillä vektori $\mathbf{b} \in \mathbb{R}^m$ kuuluu kuva-avaruuteen tasan silloin, kun on olemassa $\mathbf{v} \in \mathbb{R}^n$, jolle $A\mathbf{v} = \mathbf{b}$. Huomio ei välttämättä kerro vielä mitään siitä, miten yhtälöryhmä ratkaistaan käytännössä, mutta se kertoo meille esimerkiksi, että ne vektorit $\mathbf{b}$, joilla yhtälö ratkeaa, muodostavat aliavaruuden $\mathbb{R}^m$:ssä.

Esimerkiksi jos $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ on projektio xy -tasoon, niin yhtälöryhmällä $A\mathbf{v} = \mathbf{b}$ on ratkaisu vain, jos vektori $\mathbf{b}$ on xy -tasossa.

Ratkaisujoukon rakenne

Mikäli $\mathbf{b} \in \text{Im } A$, täytyy yhtälöryhmällä $A\mathbf{v} = \mathbf{b}$ olla siis ainakin yksi ratkaisu $\mathbf{v}$. Mutta voiko näitä ratkaisuja olla useampi, ja millaisen joukon kaikki mahdolliset ratkaisut muodostavat?

Tehdään ensin seuraava päättely: Olkoon $\mathbf{v}_1$ ja $\mathbf{v}_2$ kaksi ratkaisua. Tällöin

$$A(\mathbf{v}_1 - \mathbf{v}_2) = A\mathbf{v}_1 - A\mathbf{v}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

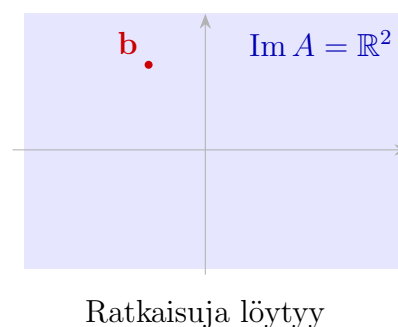
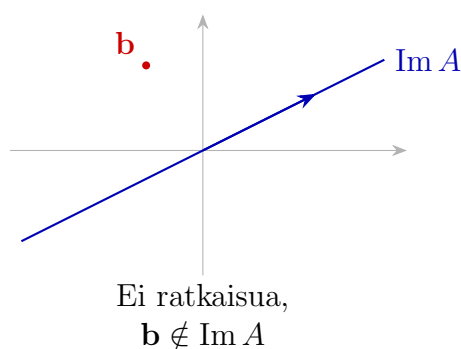
Siispä erotus $\mathbf{v}_1 - \mathbf{v}_2$ kuuluu määritelmän mukaan matriisin A ytimeen $\text{Ker } A$. Vastaavasti jos $\mathbf{v}$ on yhtälön $A\mathbf{v} = \mathbf{b}$ ratkaisu ja $\mathbf{w} \in \text{Ker } A$, niin $\mathbf{v} + \mathbf{w}$ on myös ratkaisu:

$$A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Siispä matriisin ydin paljastaa ratkaisujoukon rakenteen: Mikäli yhtälöryhmällä $A\mathbf{v} = \mathbf{b}$ on yksi ratkaisu $\mathbf{v}$, niin kaikkien ratkaisujen joukko saadaan siirtämällä aliavaruutta $\text{Ker } A$ vektorilla $\mathbf{v}$.

Tämän pohdinnan lopputulos on, että yhtälöryhmän $A\mathbf{v} = \mathbf{b}$ ratkeavuudelle on oleellisesti kolme eri tapausta:

- **Tapaus 1:** Yhtälöllä ei ole ratkaisua, koska $\mathbf{b}$ ei kuulu kuvaan $\text{Im } A$.
- **Tapaus 2a:** Yhtälöllä on ratkaisu $\mathbf{v}$, ja $\mathbf{v}$ on yksikäsitteinen, sillä $\text{Ker } A = \{\mathbf{0}\}$ on triviaali aliavaruus.
- **Tapaus 2b:** Yhtälöllä on ratkaisu $\mathbf{v}$, ja $\text{Ker } A \neq \{\mathbf{0}\}$. Tällöin kaikki ratkaisut muodostavan äärettömän joukon (suora, taso, jne.), joka on samanulotteinen kuin $\text{Ker } A$.



Gaussin eliminaatio

Haluaisimme seuraavaksi löytää menetelmän ratkaista lineaarisia yhtälöryhmiä, joka teepii kaikkiin kolmeen yllä johdettuun tapaukseen. Tätä varten on olemassa kätevä, mutta varsin mekaaninen algoritmi nimeltä **Gaussin eliminaatio**.

Gaussin eliminaatio perustuu seuraavaan kahteen huomioon:

- Yhtälöryhmän ratkaisu ei muutu, jos sen kaksi riviä vaihtaa keskenään paikkaa. Esim. yhtälöryhmillä

$$\begin{cases} x + y = 1 \\ 3x - 2y = 8 \end{cases} \quad \text{ja} \quad \begin{cases} 3x - 2y = 8 \\ x + y = 1 \end{cases} \quad \text{on samat ratkaisut.}$$

- Yhtälöryhmän ratkaisu ei muutu, jos sen riviin i lisää jonkin toisen rivin j kerrottuna jollain reaaliluvulla. Esim. yhtälöryhmillä

$$\begin{cases} x + y = 1 \\ 3x - 2y = 8 \end{cases} \quad \downarrow \times 2 \quad \text{ja} \quad \begin{cases} x + y = 1 \\ 5x = 10 \end{cases} \quad \text{on samat ratkaisut.}$$

Jälkimmäinen huomio voidaan perustella sillä, että toisesta yhtälöryhmästä päästään aina takaisin edelliseen vähentämällä rivistä i rivi j samalla kertoimella. Näitä kahta toimenpidettä sanotaan **rivioperaatioiksi**.

Huom! Nämä huomiot voi myös perustella täsmällisesti matriiseilla. Kahdessa äskeisessä esimerkissä tehdyt rivioperaatiot voi myös toteuttaa kertomalla vasen yhtälöryhmä puolittain matriiseilla

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

Tarkista, että näin todellakin on.

Gaussin eliminaatio perustuu edellä esitettyjen rivioperaatioiden toistamiseen, kunnes yhtälöryhmä on helposti ratkaistavassa muodossa. Havainnollistetaan algoritmia esimerkillä. Ratkaistaan yhtälöryhmä

$$\begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 10 \end{pmatrix}.$$

Kirjoitetaan yhtälöryhmä hieman tiiviimmässä muodossa:

$$\begin{array}{ccc|c} 2 & 4 & -2 & 2 \\ 4 & 9 & -3 & 8 \\ -2 & -3 & 7 & 10 \end{array}$$

Nyt sovelletaan rivioperaatioita, ja eliminoidaan ensimmäisestä sarakkeesta kaikki luvut ensimmäisen luvun alapuolelta seuraavasti:

$$\begin{array}{ccc|c} 2 & 4 & -2 & 2 \\ 4 & 9 & -3 & 8 \\ -2 & -3 & 7 & 10 \end{array} \begin{array}{l} \leftarrow -2 \cdot (\text{Rivi } 1) \\ \leftarrow 1 \cdot (\text{Rivi } 1) \end{array} \Rightarrow \begin{array}{ccc|c} 2 & 4 & -2 & 2 \\ 0 & 1 & 1 & 4 \\ 0 & 1 & 5 & 12 \end{array}$$

Sitten eliminoidaan toisesta sarakkeesta kaikki toisen luvun alapuolelta käyttämällä toista riviä:

$$\begin{array}{ccc|c} 2 & 4 & -2 & 2 \\ 0 & 1 & 1 & 4 \\ 0 & 1 & 5 & 12 \end{array} \begin{array}{l} \leftarrow -1 \cdot (\text{Rivi } 2) \end{array} \Rightarrow \begin{array}{ccc|c} 2 & 4 & -2 & 2 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 4 & 8 \end{array}$$

Yhtälöryhmä on nyt saatu haluttuun muotoon:

$$\begin{array}{ccc|c} 2 & 4 & -2 & 2 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 4 & 8 \end{array}$$

Tätä sanotaan **porrasmuodoksi**. Yhtälöryhmä on porrasmuodossa, kun:

- Nollarivit (jos niitä on) ovat alimpina.
- Muiden rivien ensimmäinen nollasta poikkeava alkio (ns. **pivot**) on edellisen rivin pivotin oikealla puolella.
- Jokaisen pivotin alapuolella on pelkkiä nollia.

Porrasmuodosta ratkaisu on helppo löytää aloittaen viimeisestä yhtälöstä

$$4z = 8 \quad \Rightarrow \quad z = 2.$$

Tämä voidaan sijoittaa toiseen yhtälöön, ja saadaan

$$y + z = 4 \quad \Rightarrow \quad y = 2.$$

Lopulta ensimmäisestä yhtälöstä saadaan

$$2x + 4y - 2z = 2 \quad \Rightarrow \quad x = 1.$$

Tämä algoritmi tepsii myös tapauksiin, joissa ratkaisuja ei löydy, tai löytyy äärettömästi. Käydään lisää esimerkkitapauksia vielä läpi.

Esimerkki 3.1. Ratkaistaan yhtälöryhmä

$$\begin{array}{ccc|c} 1 & 2 & 1 & 10 \\ 1 & -1 & 2 & 8 \\ 2 & 1 & 3 & 16 \end{array}$$

Ensimmäinen eliminaatio antaa

$$\begin{array}{ccc|c} 1 & 2 & 1 & 10 \\ 1 & -1 & 2 & 8 \\ 2 & 1 & 3 & 16 \end{array} \begin{array}{l} \leftarrow -1 \cdot (\text{Rivi } 1) \\ \leftarrow -2 \cdot (\text{Rivi } 1) \end{array} \Rightarrow \begin{array}{ccc|c} 1 & 2 & 1 & 10 \\ 0 & -3 & 1 & -2 \\ 0 & -3 & 1 & -4 \end{array}$$

Toinen antaa

$$\begin{array}{ccc|c} 1 & 2 & 1 & 10 \\ 0 & -3 & 1 & -2 \\ 0 & -3 & -1 & -4 \end{array} \begin{array}{l} \leftarrow -1 \cdot (\text{Rivi } 2) \end{array} \Rightarrow \begin{array}{ccc|c} 1 & 2 & 1 & 10 \\ 0 & -3 & 1 & -2 \\ 0 & 0 & 0 & -2 \end{array}$$

Yhtälöryhmä on nyt porrasmuodossa, mutta viimeisessä yhtälössä lukee $0 = -2$. Yhtälöllä ei siis ole ratkaisua.

Esimerkki 3.2. Ratkaistaan yhtälöryhmä

$$\begin{array}{cc|c} 1 & -1 & 2 \\ -2 & 2 & -4 \end{array}$$

Eliminaatio antaa ensimmäisellä askeleella suoraan porrasmuodon

$$\begin{array}{cc|c} 1 & -1 & 2 \\ 0 & 0 & 0 \end{array}$$

Jälkimmäinen yhtälö ei anna mitään ehtoa muuttujalle y , joten y on vapaa parametri. Sen sijaan ensimmäinen yhtälö antaa muuttujalle x ehdon $x = 2 + y$, joten yleinen ratkaisu on muotoa

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 + y \\ y \end{pmatrix}.$$

Esimerkki 3.3. Edellisissä tapauksissa rivien vaihtaminen keskenään ei ollut tarpeellista. Tämä kuitenkin tulee välillä tarpeeseen, esimerkiksi yhtälöryhmässä

$$\begin{array}{ccc|c} 0 & -1 & 1 & 2 \\ 2 & 1 & -1 & 2 \end{array}$$

Tässä edellinen tapa aloittaa ei toimi, koska ensimmäisen rivin ensimmäinen alkio on 0, eli ensimmäisellä rivillä ei voi eliminoida alemman rivin ensimmäistä alkioita. Siis vaihdetaan rivit keskenään, ja saadaan porrasmuodossa oleva yhtälöryhmä:

$$\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & -1 & 1 & 2 \end{array}$$

Tämä on kolmen muuttujan (esim. x, y, z) yhtälöryhmä, jossa kolmatta yhtälöä ei ole. Siis viimeiselle muuttujalle z ei tule ehtoa, joten se on vapaa parametri. Lisäksi saadaan

$$y = z - 2 \quad \text{ja} \quad 2x + y - z = 2 \Rightarrow x = 2.$$

Siis ratkaisu on

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ z - 2 \\ z \end{pmatrix}.$$

Pivotit

Gaussin eliminaation päättyessä yhtälöryhmä $A\mathbf{v} = \mathbf{b}$ on palautunut muotoon kuten

$$\begin{array}{cccc|c} 3 & -2 & 0 & 6 & 2 \\ 0 & 0 & -1 & 3 & -7 \\ 0 & 0 & 0 & 4 & 0 \end{array}$$

Yllä eroteltuja johtavia alkioita sanotaan siis **pivoteiksi**. Pivotien sijainneista voidaan päätellä suoraan, miltä yhtälöryhmän ratkaisu näyttää.

Tapaus 1: Yhtälöryhmällä ei ole ratkaisua tasan silloin, kun pivotien alapuolelle jää jokin rivi, jossa vasemmalla on pelkkiä nollia ja oikealla on nollasta poikkeava luku. Siis kuten Esimerkissä 3.1.

Mikäli näin ei tapahdu, yhtälö ratkeaa, ja tapauksia on taas kaksi:

Tapaus 2a: Yhtälöryhmällä on yksikäsitteinen ratkaisu tasan silloin, kun jokaisella sarakkeella on yksi pivot. Tällöin pivotit asettuvat kaikki diagonaalille. Siis esimerkiksi:

$$\begin{array}{ccc|c} 1 & 0 & 4 & 2 \\ 0 & -5 & 1 & -3 \\ 0 & 0 & 6 & 12 \end{array}$$

Tapaus 2b: Yhtälöryhmällä on äärettömän monta ratkaisua silloin, kun jollain sarakkeella ei ole pivotia. Tällöin näitä sarakkeita vastaavat muuttujat ovat **vapaita muuttujia**, ja pivotien sarakkeita vastaavat muuttujat ovat ratkaistavia **pivotmuuttujia**. Esim.

$$\begin{array}{ccccc|c} 1 & 4 & -1 & 0 & 2 & 0 \\ 0 & 0 & 2 & 5 & -2 & 1 \\ 0 & 0 & 0 & 4 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Tässä yhtälöryhmässä on 5 saraketta, eli 5 muuttujaa (esim. x, y, z, a, b). Toinen ja viimeinen (eli y, b) ovat vapaita, ja muut (eli x, z, a) ovat pivotmuuttujia.

Vapaiden muuttujien lukumäärä kertoo ratkaisuavaruuden ulottuvuuden määrän, mikä on sama kuin yhtälöryhmän matriisin A ytimellä $\text{Ker } A$.

Pivotmuuttujien lukumäärä taas kertoo, kuinka monta vapausastetta kuva-avaruudessa on, eli matriisin A asteen $\text{rank } A$. Koska pivotmuuttujien ja vapaiden muuttujien määrä on yhteensä sarakkeiden lkm, voimme päätellä seuraavan tuloksen.

Lause 3.1. *Olko A $m \times n$ -matriisi. Tällöin*

$$\dim \text{Ker } A + \text{rank } A = n.$$

Tämä tulos tunnetaan **Rank-Nullity** lauseena.

4. Käänteismatriisi

Olkoon A neliömatriisi, eli $n \times n$ -matriisi jollain n . Tällöin on mahdollista, että A :lla on olemassa **käänteismatriisi** A^{-1} , mikäli löytyy sellainen $n \times n$ -matriisi A^{-1} , jolle pätee

$$A^{-1}A = I = AA^{-1}.$$

Käänteismatriisilla kertominen molemmilta puolilta siis antaa samankokoisen identiteettimatriisin tulokseksi. Tällöin sanotaan, että matriisi A on **kääntövä** tai **säännöllinen**.

Huom! Osoittautuu, että jos A ja B ovat neliömatriiseja ja joko $AB = I$ tai $BA = I$, niin välttämättä $B = A^{-1}$. Siis käänteismatriisin ehto täytyy tarkistaa vain toiseen suuntaan. Perustelun tälle voi päätellä myöhemmin, kun esittelemme eri ehtoja kääntövyydelle.

Lineaarikuvauksena tulkittuna $A^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ on kuvauksen A **käänteiskuvaus**, mikäli sellainen löytyy.

Käänteismatriisi ja yhtälöryhmät

Jos käänteismatriisi A^{-1} sattuu olemaan tiedossa, niin mikä tahansa yhtälöryhmä

$$A\mathbf{v} = \mathbf{b}, \quad \text{missä } \mathbf{v} \text{ ja } \mathbf{b} \text{ ovat molemmat } n \times 1\text{-vektoreita,}$$

on varsin helppo ratkaista. Kerrotaan yhtälö molemmin puolin matriisilla A^{-1} , ja saadaan

$$\begin{aligned} A^{-1}A\mathbf{v} &= A^{-1}\mathbf{b} && \Leftrightarrow \\ I\mathbf{v} &= A^{-1}\mathbf{b} && \Leftrightarrow \\ \mathbf{v} &= A^{-1}\mathbf{b}. \end{aligned}$$

Tämä huomio pätee myös toiseen suuntaan. Mikäli yhtälöryhmällä $A\mathbf{v} = \mathbf{b}$ on jokaisella $\mathbf{b} \in \mathbb{R}^n$ olemassa yksikäsitteinen ratkaisu $\mathbf{v}_{\mathbf{b}} \in \mathbb{R}^n$, on matriisin A pakko olla kääntövä. Tarkka perustelu tälle voidaan löytää osoittamalla, että kuvaus $\mathbf{b} \mapsto \mathbf{v}_{\mathbf{b}}$ määrittelee lineaarikuvauksen, jonka täytyy vastata A :n käänteismatriisia (pohdi miksi). On kuitenkin myös toinen tapa päätellä sama asia, ja tämä antaa myös tavan laskea käänteismatriisi käytännössä.

Käänteismatriisi Gauss-Jordanin eliminaatiolla

Myös käänteismatriisin määrittämiseen voidaan soveltaa Gaussin eliminaatiota, tosin menetelmää täytyy yleistää hieman. Tämän yleistyksen nimi on **Gauss-Jordanin menetelmä**. Otetaan esimerkiksi matriisi:

$$A = \begin{pmatrix} 1 & 3 & -2 \\ 0 & 4 & 1 \\ -2 & -2 & 3 \end{pmatrix}.$$

Käänteismatriisin muodostamista varten kirjoitetaan matriisi A vierekkäin identiteettimatriisin kanssa:

$$\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 & 1 & 0 \\ -2 & -2 & 3 & 0 & 0 & 1 \end{array}$$

Gaussin eliminaation jälkeen tämä saadaan porrasmuotoon

$$\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 & 1 & 0 \\ 0 & 0 & -2 & 2 & -1 & 1 \end{array}$$

Haluamme tästä kuitenkin vielä jatkaa eliminointia, kunnes vasemmalle puolelle jää vain identiteettimatriisi. Tämä saavutetaan ensin jakamalla rivejä sopivilla luvuilla, että pivotien paikalle saadaan ykköset:

$$\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 & 1 & 0 \\ 0 & 0 & -2 & 2 & -1 & 1 \end{array} \begin{array}{l} \\ \leftarrow \times \frac{1}{4} \\ \leftarrow \times \frac{-1}{2} \end{array} \Rightarrow \begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & -\frac{1}{2} \end{array}$$

Sitten alempien rivien pivoteja käytetään eliminoimaan ylemmiltä riveiltä nolasta poikkeavat oikean puolen alkiot:

$$\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & -\frac{1}{2} \end{array} \leftarrow (-3) \times (\text{Rivi } 2) \\ \Rightarrow \begin{array}{ccc|ccc} 1 & 0 & -\frac{11}{4} & 1 & -\frac{3}{4} & 0 \\ 0 & 1 & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & -\frac{1}{2} \end{array} \begin{array}{l} \\ \leftarrow \frac{11}{4} \times (\text{Rivi } 3) \\ \leftarrow \frac{-1}{4} \times (\text{Rivi } 3) \end{array} \\ \Rightarrow \begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{7}{4} & \frac{5}{8} & -\frac{11}{8} \\ 0 & 1 & 0 & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ 0 & 0 & 1 & -1 & \frac{1}{2} & -\frac{1}{2} \end{array}$$

Lopulta vasemmalle jää 3×3 -identiteettimatriisi, ja oikealla puolella oleva matriisi on alkuperäisen matriisin A käänteismatriisi

$$A^{-1} = \begin{pmatrix} -\frac{7}{4} & \frac{5}{8} & -\frac{11}{8} \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ -1 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Perustelu miksi tämä toimii on varsin yksinkertainen: Olkoon $\mathbf{v}$ jokin vektori, ja määritellään $\mathbf{b} = A\mathbf{v}$. Jokainen Gauss-Jordanin rivioperaatio vastaa samaa, kuin yhtälöä $A\mathbf{v} = \mathbf{b}$ kerrotaisiin puolittain jollain sopivalla matriisilla, esimerkiksi

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Rivien 1 ja 2 vaihto}} \quad \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}}_{\text{Riviin 3 lisätään } (-2) \times (\text{Rivi } 1)} \quad \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}}_{\text{Rivi 3 kerrotaan luvulla } \frac{1}{2}}$$

Merkitään eliminaatiossa esiintyviä matriiseja $E_1, E_2, \dots, E_N$. Lopulta yhtälön vasemmalle puolelle jää vain identiteettimatriisi, eli symbolisesti se menee muotoon

$$I\mathbf{v} = E_N \cdots E_1 \mathbf{b}.$$

Tästä voidaan päätellä, että oikealle puolelle jäänyt matriisi $E_N \cdots E_1$ on täsmälleen A :n käänteismatriisi, sillä

$$\mathbf{v} = I\mathbf{v} = E_N \cdots E_1 \mathbf{b} = (E_N \cdots E_1) A\mathbf{v}.$$

Siis täytyy olla $A^{-1} = E_N \cdots E_1$.

Esimerkki 4.1. (Cramerin sääntö). Tutkitaan, milloin 2×2 -matriisilla

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

on käänteismatriisi. Väitämme, että tämä käänteismatriisi saadaan aina kaavasta

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

mikäli lauseke $ad - bc$ ei ole nolla. Tämä tunnetaan **Cramerin sääntönä**. Nopea tarkistus:

$$AA^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} ad - bc & 0 \\ 0 & -bc + ad \end{pmatrix} = I_2.$$

Myös $A^{-1}A = I_2$ samanlaisella laskulla.

Siis kappaleen alussa esiintynyt yhtälöryhmä

$$\begin{cases} x + y = 5 \\ 2x + y = 7 \end{cases}$$

voitaisiin ratkaista muodostamalla matriisin

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \text{ käänteismatriisi } A^{-1} = \frac{1}{1 - 2} \begin{pmatrix} 1 & -1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix},$$

ja kertomalla tällä vektoria $\begin{pmatrix} 5 \\ 7 \end{pmatrix}$, antaen lopputuloksena ratkaisun $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

Kääntyvyys ehdot

Matriisin kääntyvyys voidaan tarkistaa monella eri tavalla. Seuraava lause esittelee osan eri tavoista.

Lause 4.1. *Olko A $n \times n$ -neliomatriisi. Tällöin seuraavat ehdot ovat yhtäpitäviä*

- A on kääntyvä.
- $\text{Ker } A = \{\mathbf{0}\}$.
- $\text{rank } A = n$.
- A :n rivivektorit ovat lineaarisesti riippumattomia.
- A :n sarakevektorit ovat lineaarisesti riippumattomia.

Todistus. Matriisin kääntyvyys on yhtäpitävää sen kanssa, että yhtälöryhmällä $A\mathbf{v} = \mathbf{b}$ on yksikäsitteinen ratkaisu kaikilla vektoreilla $\mathbf{b}$. Gaussin eliminaatiosta huomattiin, että tämä tapahtuu juuri silloin, kun porrasmuodossa on tasan n kappaletta pivoteja (tällöin myöskään nollarivejä ei jää, sillä A on neliömatriisi).

Tämä on ensinnäkin yhtäpitävää ehdon $\text{Ker } A = \{\mathbf{0}\}$ kanssa, sillä ytimen $\text{Ker } A$ dimensio vastaa vapaiden muuttujien lukumäärää. Eli kun vapaita muuttujia ei ole, ydin on nollaulotteinen ja sisältää vain nollavektorin.

Rank-Nullity lauseen perusteella tämä on ekvivalenttia sen kanssa, että matriisi on täysasteinen, $\text{rank } A = n$. Ja koska matriisin kuva-avaruus on sen sarakevektoreiden virittämä, on tämä jälleen sama asia, kuin sarakevektoreiden lineaarinen riippumattomuus.

Lopulta perustellaan, miksi rivivektorit ovat myös lineaarisesti riippumattomia. Yksi helppo perustelu on transpoosin kautta: Matriisi A on kääntyvä tasan silloin, kun transpoosi A^T on myös kääntyvä, sillä

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

Todistus tälle harjoitustehtävissä. Koska transpoosin sarakevektorit ovat A :n rivivektorit, on myös rivivektoreiden lin. riippumattomuus yhtäpitävää kääntyvyyden kanssa.

5. Ortogonaalisuus

Vektorit-osiossa totesimme, että kaksi vektoria $\mathbf{v}, \mathbf{w}$ ovat **kohtisuorassa** tai **ortogonaaliset**, mikäli $\mathbf{v} \cdot \mathbf{w} = 0$. Tyypillisesti geometriassa kohtisuoruus määritellään suorien välisen kulman avulla, mutta osoittautuu, että sen määrittäminen pistetulon avulla on monella tapaa luonnollisempi ja käytännöllisempi vaihtoehto. Tässä kappaleessa tutkimme, mitä uusia mahdollisuuksia tämä lähestymistapa avaa.

Pythagoraan lause: Mikäli vektorit $\mathbf{v}$ ja $\mathbf{w}$ ovat ortogonaaliset, täytyy päteä, että

$$|\mathbf{v} + \mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2.$$

Tämä on varsin suora seuraus pistetulon ominaisuuksista, sillä jos $\mathbf{v} \cdot \mathbf{w} = 0$, niin

$$|\mathbf{v} + \mathbf{w}|^2 = (\mathbf{v} + \mathbf{w}) \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{v} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} = |\mathbf{v}|^2 + 2(\mathbf{v} \cdot \mathbf{w}) + |\mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2.$$

Ortogonaaliset jonot

Olkoon $\mathbf{v}_1, \dots, \mathbf{v}_N$ jono nollasta poikkeavia vektoreita. Sanomme, että vektorit $\mathbf{v}_j$ muodostavat **ortogonaalisen jonon** tai **-kannan**, mikäli ne ovat kaikki toisiaan vastaan kohtisuorassa, eli

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0, \quad \text{kun } i \neq j.$$

Lisäksi sanomme, että jono $(\mathbf{v}_i)$ on **ortonormaali**, mikäli lisäksi jokaisen vektorin $\mathbf{v}_i$ pituus on yksi, eli $\mathbf{v}_i \cdot \mathbf{v}_i = 1$. Esimerkiksi avaruuden $\mathbb{R}^n$ standardikanta $\mathbf{e}_1, \dots, \mathbf{e}_n$ on ortonormaali.

Ortogonaalinen projektio

Olkoon $\mathbf{v}_1, \dots, \mathbf{v}_N$ jokin ortogonaalinen kanta aliavaruudelle

$$V = \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_N).$$

Jos $\mathbf{b} \in \mathbb{R}^n$ on jokin vektori, joka ei kuulu aliavaruuteen V , on luonnollista pohtia, miten lasketaan pisteen $\mathbf{b}$ ja aliavaruuden V etäisyys. Tätä varten täytyy selvittää sellainen aliavaruuden vektori $\mathbf{w} \in V$, jonka etäisyys vektorista $\mathbf{b}$ on pienin mahdollinen.

Geometrian perusteella järkevä arvaus on, että vektori $\mathbf{w}$ pitää valita niin, että vektori $\mathbf{b} - \mathbf{w}$ on kohtisuorassa aliavaruutta V , eli jokaista sen vektoria, vasten. Tällaista vektoria $\mathbf{w}$ sanotaan vektorin $\mathbf{b}$ **ortogonaaliseksi projektioksi** aliavaruudelle V . Sitä merkitään $\text{proj}_V(\mathbf{b})$ (tai $\text{proj}_{\mathbf{v}_1, \dots, \mathbf{v}_N}(\mathbf{b})$) ja sen voi laskea kaavasta

$$\text{proj}_V(\mathbf{b}) = \sum_{i=1}^N \frac{\mathbf{b} \cdot \mathbf{v}_i}{|\mathbf{v}_i|^2} \mathbf{v}_i.$$

Erityisesti ortogonaalinen projektio jonkin vektorin $\mathbf{v}$ määräämälle suoralle saadaan kaavasta

$$\text{proj}_{\mathbf{v}}(\mathbf{b}) = \frac{\mathbf{b} \cdot \mathbf{v}}{|\mathbf{v}|^2} \mathbf{v},$$

ja projektio yleiselle ortogonaalisen kannan määräämälle aliavaruudelle V saadaan täten laskemalla yhteen yksiulotteiset projektiot jokaisen kantavektorin suuntaan.

Ortogonaalinen projektio $\text{proj}_V(\mathbf{b})$ on se yksikäsitteinen vektori, jolle vektori $\mathbf{b} - \text{proj}_V(\mathbf{b})$ on kohtisuorassa jokaista aliavaruuden V vektoria kohti (perustelu hetken päästä). Tarkistetaan nyt, että tämä todellakin antaa $\mathbf{b}$:n lähimmän mahdollisen vektorin aliavaruudessa V . Jos $u \in V$ on mikä tahansa vektori, niin Pythagoraan lauseella

$$|\mathbf{b} - u|^2 = |\mathbf{b} - \text{proj}_V(\mathbf{b}) + (\text{proj}_V(\mathbf{b}) - u)|^2 = |\mathbf{b} - \text{proj}_V(\mathbf{b})|^2 + |\text{proj}_V(\mathbf{b}) - u|^2 \geq |\mathbf{b} - \text{proj}_V(\mathbf{b})|^2.$$

Siis valinta $u = \text{proj}_V(\mathbf{b})$ todellakin minimoi etäisyyden $|\mathbf{b} - u|$.

Tarkistetaan vielä, että $\mathbf{b} - \text{proj}_V(\mathbf{b})$ on varmasti kohtisuorassa aliavaruutta V vasten. Jokaisella $j = 1, \dots, N$ pätee

$$\begin{aligned} (\mathbf{b} - \text{proj}_V(\mathbf{b})) \cdot \mathbf{v}_j &= \sum_{i=1}^N \left(\mathbf{b} \cdot \mathbf{v}_j - \frac{\mathbf{b} \cdot \mathbf{v}_i}{|\mathbf{v}_i|^2} \mathbf{v}_i \cdot \mathbf{v}_j \right) \\ &= \left(\mathbf{b} \cdot \mathbf{v}_j - \frac{\mathbf{b} \cdot \mathbf{v}_j}{|\mathbf{v}_j|^2} \mathbf{v}_j \cdot \mathbf{v}_j \right) \quad (\text{Sillä } i \neq j \Rightarrow \mathbf{v}_i \cdot \mathbf{v}_j = 0) \\ &= (\mathbf{b} \cdot \mathbf{v}_j - \mathbf{b} \cdot \mathbf{v}_j) \\ &= 0. \end{aligned}$$

Ja jos $u = \sum_{j=1}^N a_j \mathbf{v}_j$ on jokin aliavaruuden V vektori esitettynä kannassa $(\mathbf{v}_j)$, niin

$$(\mathbf{b} - \text{proj}_V(\mathbf{b})) \cdot u = \sum_{j=1}^N a_j (\mathbf{b} - \text{proj}_V(\mathbf{b})) \cdot \mathbf{v}_j = \sum_{j=1}^N a_j \cdot 0 = 0.$$

□

Ortogonalisointi

Olkoon $\mathbf{v}_1, \dots, \mathbf{v}_N$ kanta avaruuden $\mathbb{R}^n$ aliavaruudelle. Tällöin on olemassa toinen kanta $\mathbf{u}_1, \dots, \mathbf{u}_N$ samalle aliavaruudelle siten, että vektorit $\mathbf{u}_i$ ovat ortonormaaleja.

Tämän uuden kannan muodostaminen onnistuu **Gram-Schmidtin orthogonalisoinnilla**. Tämän algoritmin idea on edetä vektorit $\mathbf{v}_i$ läpi järjestyksessä, ja poistaa aina jokaisesta vektorista edellisten vektoreiden kanssa samansuuntaiset osuudet. Menetelmä toimii käytännössä näin:

- (1) Olkoon $\mathbf{u}_1 = \mathbf{v}_1 / |\mathbf{v}_1|$.
- (2) Olkoon $\mathbf{u}_2^* = \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2)$. Tällöin $\mathbf{u}_2 = \mathbf{u}_2^* / |\mathbf{u}_2^*|$ on yksikkövektori, joka on kohtisuorassa vektoria $\mathbf{u}_1$ vasten.
- (3) Olkoon yleisesti

$$\mathbf{u}_j^* = \mathbf{v}_j - \text{proj}_{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{j-1}}(\mathbf{v}_j).$$

Tällöin $\mathbf{u}_j = \mathbf{u}_j^* / |\mathbf{u}_j^*|$ on yksikkövektori, joka on kohtisuorassa kaikkia vektoreita $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{j-1}$ vasten.

Tämä tuottaa lopulta ortonormaalin vektorijonon $\mathbf{u}_1, \dots, \mathbf{u}_N$ joka virittää saman aliavaruuden, kuin vektorit $\mathbf{v}_1, \dots, \mathbf{v}_N$.

Ortogonaaliset matriisit

Sanotaan, että neliömatriisi Q on **ortogonaalinen**, mikäli

$$Q^T Q = I.$$

Ortogonaalisuuden voi ilmaista usealla eri tavalla

Lause 5.1. *Olkoon Q $n \times n$ -neliömatriisi. Tällöin seuraavat ehdot ovat yhtäpitäviä*

- Q on ortogonaalinen.
- Q on kääntävä ja $Q^{-1} = Q^T$.
- Matriisin Q sarakkeet muodostavat ortonormaalin kannan avaruudelle $\mathbb{R}^n$.
- Matriisin Q rivit muodostavat ortonormaalin kannan avaruudelle $\mathbb{R}^n$.
- Lineaarikuvaus $Q : \mathbb{R}^n \rightarrow \mathbb{R}^n$ säilyttää vektorien pituudet ja niiden väliset kulmat, eli

$$Q\mathbf{u} \cdot Q\mathbf{v} = \mathbf{u} \cdot \mathbf{v} \quad \text{kaikilla } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$$

Lisäksi näistä voidaan päätellä, että ortogonaaliselle matriisille pätee myös $QQ^T = I$, ja

$$\det(Q)^2 = \det(Q^T) \det(Q) = \det(Q^T Q) = \det(I) = 1 \quad \Rightarrow \quad \det(Q) = \pm 1.$$

Pienimmän neliösumman ratkaisu yhtälöryhmälle

Olemme huomanneet aiemmin, että yleisellä lineaarisella yhtälöryhmällä

$$A\mathbf{v} = \mathbf{b}$$

ei aina ole ratkaisua. Tämä tapahtuu tasan silloin, kun $\mathbf{b}$ ei ole kuva-avaruudessa $\text{Im } A$. Käytännössä näin tapahtuu erityisen usein, kun yhtälöryhmässä on huomattavasti enemmän yhtälöitä kuin muuttujia.

Tällaisessa tapauksessa on kuitenkin järkevä pohtia, onko olemassa jotain vektoria $\mathbf{v}$, joka on jossain mielessä 'mahdollisimman lähellä' yhtälön ratkaisemista. Luonnollinen tapa tehdä tästä täsmällistä, on etsiä vektori $\mathbf{v}$, jolla suure

$$|A\mathbf{v} - \mathbf{b}|^2$$

on pienin mahdollinen. Tämän optimointiongelman ratkaisua $\mathbf{v}$ sanotaan **pienimmän neliösumman ratkaisuksi** yhtälöryhmälle

$$A\mathbf{v} = \mathbf{b}.$$

Suureen $|A\mathbf{v} - \mathbf{b}|^2$ minimointi tarkoittaa erityisesti sitä, että haluamme etsiä matriisin A kuva-avaruuden vektoria $\mathbf{w} = A\mathbf{v}$, jolla etäisyys $|\mathbf{w} - \mathbf{b}|$ on mahdollisimman pieni. Tämä saavutetaan silloin, kun $\mathbf{w}$ on vektorin $\mathbf{b}$ ortogonaalinen projektio kuva-avaruudelle $\text{Im } A$.

Erityisesti vektori $\mathbf{b} - \mathbf{w} = \mathbf{b} - A\mathbf{v}$ on tällöin kohtisuorassa kuva-avaruutta vasten, joten se kuuluu kuva-avaruuden ortokomplementtiin, eli transpoosin A^T ytimeen $\text{Ker } A^T$. Siispä

$$A^T(\mathbf{b} - A\mathbf{v}) = \mathbf{0} \quad \Leftrightarrow \quad A^T A\mathbf{v} = A\mathbf{b}.$$

Tällä tavalla

Esimerkki 5.1. (Suoran sovitus). Olkoon annettu N kappaletta pisteitä (x_i, y_i) , $i = 1, \dots, N$ tasossa. Pyrimme etsimään suoraa

$$y = c_1 x + c_2,$$

joka parhaiten approksimoi pistejoukkoa (x_i, y_i) . Tämä on luonnollista muotoilla pienimmän neliösumman ongelmana, jossa

$$A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}, \quad \text{ja} \quad \mathbf{v} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Sillä tällöin suure

$$|A\mathbf{v} - \mathbf{b}|^2 = \sum_{i=1}^N |(c_1 x_i + c_2) - y_i|^2$$

ilmaisee neliösummaa pisteiden (x_i, y_i) poikkeamista suoran $y = c_1 x + c_2$ suhteen.

Samalla idealla voidaan sovittaa pistejoukkoihin myös esim. korkeamman asteen polynomeja, taso kolmessa ulottuvuudessa, jne.

6. Determinantti

Jokaiseen neliömatriisiin voidaan liittää lukuarvo nimeltä **determinantti**. Matriisin A determinanttia merkitään $\det A$.

Geometrisesti determinantti kuvaa sitä, kuinka paljon kuvaus $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ venyttää avaruutta. Otetaan esimerkiksi 2×2 -matriisi

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

Tämä matriisi A tekee geometrisesti kahta asiaa: Se venyttää x -suunnassa kertoimella 2, ja y -suunnassa kertoimella 3. Siis minkä tahansa tasokuvion **pinta-ala** kuusinkertaistuu, kun sitä venytetään kuvauksella A . Tämän takia matriisin A determinantti on $\det A = 6$.

Katsotaan seuraavaksi kahta muuta esimerkkiä

$$B = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \quad \text{ja} \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Huomataan, että matriisin B rivit (ja sarakkeet) ovat lineaarisesti riippuvia. Matriisi B kuvaa mielivaltaisen pisteen seuraavasti:

$$B \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ x - y \end{pmatrix}.$$

Huomataan, että jokainen B :n tuottama kuvapiste on muotoa $\begin{pmatrix} z \\ z \end{pmatrix}$ jollain $z \in \mathbb{R}$, joten B :n kuvajoukko on yksinkertaisesti suora $\{x = y\}$ tasossa. Koska suoran pinta-ala on nolla, kuvaa B jokaisen tasokuvion yksiulotteiseksi kuvioksi, jonka pinta-ala on nolla. Tämän takia B :n determinantti on $\det B = 0$.

Sen sijaan C toimii kuvauksena yksinkertaisesti vaihtamalla koordinaattien paikkaa:

$$C \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}.$$

Tämän voi myös geometrisesti tulkita peilauksena suoran $\{x = y\}$ suhteen. Peilaus ei muuta pinta-alaa, mutta se vaihtaa kuvion **orientaation**, eli oikeakätinen kuvio muuttuu vasenkätiseksi. Tämän takia C :n determinantti on $\det C = -1$.

Determinantti siis kuvaa tasossa pinta-alan muutosta orientaatio mukaanlukien. Yleisen 2×2 -matriisin determinantin voi pienellä vaivalla laskea kaavasta

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

Tämä luku esiintyi myös Cramerin säännössä, ja se oli nolla tasan silloin, kun matriisi ei ollut kääntyvä. Tämä on myös yleinen fakta determinantista.

Lause 6.1. *Neliömatriisi A on kääntyvä täsmälleen silloin, kun $\det A \neq 0$.*

Yleisemmälle $n \times n$ -neliömatriisille determinantti ei kuvaa pinta-alan muutosta, vaan vastaavan n -ulotteisen tilavuuden muutosta, jälleen orientaatio/kätisyys mukaanlukien. Determinantin itseisarvo kertoo tilavuuden varsinaisen muutoksen, ja sen merkki kertoo, vaihtuuko orientaatio vai ei. Ylläolevan lauseen perusteella determinantti paljastaa siis myös, onko lineaarisella yhtälöryhmällä

$$A\mathbf{v} = \mathbf{b}$$

aina yksikäsitteistä ratkaisua olemassa.

Deteminantille pätee myös mainio kaava matriisitulon kanssa:

$$\boxed{\det(AB) = \det A \det B.} \quad (6.1)$$

Tämä kaava on varsin helppo perustella geometrisesti: Jos kuvaus B ensin muuttaa tilavuutta kertoimella $\det B$, ja kuvaus A sitten kertoimella $\det A$, niin yhdistetyn kuvauksen AB tilavuuden muutos on näiden lukujen tulo. Tästä voidaan johtaa, että käänteismatriisin determinantti saadaan kaavasta

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

Determinantin laskeminen

Determinantin laskeminen isommille matriiseille on valitettavan työlästä käsin, mutta siihen on pari vaihtoehtoista menetelmää.

Menetelmä 1. Gaussin eliminaatio.

Gaussin eliminaatiolla voidaan myös laskea determinantteja. Otetaan esimerkiksi matriisi (tiivistetyksi kirjoitettuna)

$$\begin{array}{ccc} 0 & -1 & 1 \\ 2 & -2 & 4 \\ -1 & 3 & 4 \end{array}$$

Idea on, että Gaussin eliminaation jokainen askel ei muuta matriisin determinantin suuruutta, paitsi jokainen kahden rivin vaihto vaihtaa determinantin merkin.

Kun yllä olevalle matriisille tehdään Gaussin eliminaatio, vaihdetaan ensiaskeleella rivit 1 ja 2, ja sitten edetään normaalisti

$$\begin{array}{ccc} 0 & -1 & 1 \\ 2 & -2 & 4 \\ -1 & 3 & 4 \end{array} \Rightarrow \begin{array}{ccc} 2 & -2 & 4 \\ 0 & -1 & 1 \\ -1 & 3 & 4 \end{array} \leftarrow 0.5 \cdot (\text{Rivi } 1) \Rightarrow \begin{array}{ccc} 2 & -2 & 4 \\ 0 & -1 & 1 \\ 0 & 2 & 6 \end{array}$$

Ja vielä lisäämällä kolmanteen riviin $2 \times$ toinen rivi saadaan lopullinen muoto

$$\begin{array}{ccc} 2 & -2 & 4 \\ 0 & -1 & 1 \\ 0 & 0 & 8 \end{array}$$

Tällaisen ns. yläkolmiomatriisin determinantti on sen **diagonaalien alkioden tulo**, eli $2 \cdot (-1) \cdot 8 = -16$. Kun muistetaan vielä, että eliminaatiossa piti tehdä yksi vaihto (eli merkkiä pitää vielä vaihtaa kerran), niin lopulta saadaan

$$\det \begin{pmatrix} 0 & -1 & 1 \\ 2 & -2 & 4 \\ -1 & 3 & 4 \end{pmatrix} = (-1) \cdot (-16) = 16.$$

Menetelmä 2. 3×3 -Matriisit.

Esittelemme näppärän tavan laskea 3×3 -matriisin

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

determinantti. Kirjoitetaan avuksi matriisiin viereen sen kaksi ensimmäistä saraketta:

$$\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}$$

Nyt determinantti saadaan

- Lisäämällä kaikki alas ja oikealle menevien diagonaalien alkioden tulot ja
- Vähentämällä kaikki ylös ja oikealle menevien diagonaalien alkioden tulot

Siis

$$\det A = aei + bfg + cdh - gec - hfa - idb.$$

Vaihtoehtoisesti samat diagonaalit voi hahmottaa suoraan matriisista ilman apukuvaa, kun ajatellaan, että oikean reunan yli mennessä tullaan vasemmasta reunasta takaisin. Esim. diagonaalit bfg ja cdh :

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

Menetelmä 3: Kofaktorikehitelmä.

Olkoon $A = (a_{ij})$ $n \times n$ -matriisi, ja merkitään A_{ij} :llä sitä $(n-1) \times (n-1)$ -matriisia, joka saadaan A :sta poistamalla rivi i ja sarake j . Tällöin mille tahansa valitulle riville i determinantti voidaan kehittää kaavalla

$$\det A = \sum_{j=1}^n a_{ij} (-1)^{i+j} \det A_{ij}.$$

Täten isomman matriisin determinantti voidaan aina laskea pienempien matriisien avulla, joten jokaisen determinantin voi lopulta laskea toistamalla tätä ja pilkkomalla riittävän pieniin osiin. Tämä menetelmä ei kuitenkaan ole kovin käytännöllinen, joten se mainitaan tässä lähinnä yleissivistyksenä.

7. Ominaisarvot ja ominaisvektorit

Olkoon A jälleen $n \times n$ -neliömatriisi, ja ajatellaan sitä nyt kuvauksena $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Kun matriisilla A kuvataan vektoria $\mathbf{v}$, saadaan uusi vektori

$$A\mathbf{v},$$

jonka pituus ja suunta ovat tyypillisesti eri, kuin vektorilla $\mathbf{v}$. Mutta on toisinaan mahdollista löytää myös sellaisia vektoreita, joiden suunta ei muutu, eli joita matriisi A ainoastaan venyttää. Näille vektoreille pätee siis

$$A\mathbf{v} = \lambda\mathbf{v} \quad \text{jollain skalaarilla } \lambda \in \mathbb{R}.$$

Huomaa, että annamme tässä skalaarikertoimen λ olla mahdollisesti myös negatiivinen, missä tapauksessa vektorin $\mathbf{v}$ suunta kääntyy vastakkaiseksi kuvauksessa A .

Mikäli jokin vektori $\mathbf{v} \neq \mathbf{0}$ toteuttaa yhtälön $A\mathbf{v} = \lambda\mathbf{v}$, sanotaan, että $\mathbf{v}$ on matriisin A **ominaisvektori** ja λ on matriisin A **ominaisarvo**.

Ominaisarvot ja -vektorit osoittautuvat erittäin hyödylliseksi käsitteeksi ymmärtämään matriisien ominaisuuksia syvemmin. Tutkitaan ensin, milloin sellaisia voisi löytyä. Yhtälöä

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

voidaan muokata kirjoittamalla $\lambda\mathbf{v} = \lambda I\mathbf{v}$, jolloin

$$A\mathbf{v} = \lambda I\mathbf{v} \quad \Rightarrow \quad A\mathbf{v} - \lambda I\mathbf{v} = \mathbf{0} \quad \Rightarrow \quad (A - \lambda I)\mathbf{v} = \mathbf{0}.$$

Pohditaan, milloin tällä yhtälöllä on nollasta poikkeava ratkaisu $\mathbf{v}$. Koska $A - \lambda I$ on neliömatriisi, voi yhtälöllä olla ratkaisu vain silloin, kun matriisi $A - \lambda I$ ei ole kääntyvä. Siis Lauseen 4.1 nojalla

$$\det(A - \lambda I) = 0.$$

Jos matriisi A on annettu, tämä on muuttujan λ suhteen yhden tuntemattoman muuttujan yhtälö. On mahdollista osoittaa, että lauseke

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

on aina muuttujan λ suhteen polynomi astetta n . Tätä sanotaan matriisin A **karakteristiseksi polynomiksi**, ja sen juuret ovat matriisin A ominaisarvot.

Esimerkki 7.1. Olkoon

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

Selvitetään matriisiin A ominaisarvot ja -vektorit. Karakteristinen polynomi on

$$p(\lambda) = \det(A - \lambda I_2) = \det \begin{pmatrix} 2 - \lambda & 1 \\ 0 & -1 - \lambda \end{pmatrix} = (2 - \lambda)(-1 - \lambda).$$

Siis ominaisarvot ovat $\lambda = 2$ ja $\lambda = -1$. Määritetään sitten ominaisvektorit $\mathbf{v}$.

Tapaus $\lambda = 2$. Saadaan yhtälö

$$(A - 2I_2)\mathbf{v} = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Tästä luetaan suoraan $y = 0$, ja x on vapaa parametri. Siis yleinen ominaisvektori on muotoa

$$\mathbf{v} = \begin{pmatrix} x \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

mielivaltaisella $x \neq 0$.

Tapaus $\lambda = -1$. Saadaan yhtälö

$$(A - (-1)I_2)\mathbf{v} = \begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Tästä luetaan suoraan $3x + y = 0$. Valitaan x vapaaksi parametriksi, jolloin $y = -3x$, ja yleinen ominaisvektori on muotoa

$$\mathbf{v} = \begin{pmatrix} x \\ -3x \end{pmatrix} = x \begin{pmatrix} 1 \\ -3 \end{pmatrix}$$

mielivaltaisella $x \neq 0$.

Lopputulos. Siis omaisvektorit ovat (skaalausta vaille) vektorit $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ja $\begin{pmatrix} 1 \\ -3 \end{pmatrix}$, ja niitä vastaavat ominaisarvot 2 ja -1 .

Esimerkki 7.2. Kaikilla matriiseilla ei ole (reaalisia) ominaisarvoja. Esimerkiksi matriisi

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

vastaa tasossa 90-asteen kiertoa, joten nollavektoria lukuunottamatta se muuttaa jokaisen vektorin suuntaa, eli sillä ei voi olla ominaisvektoreita. Tämä näkyy suoraan myös karakteristisesta polynomista $p_A(\lambda) = \lambda^2 + 1$, jolla ei ole reaalisia juuria.

Huom! Ominaisarvojen teoria muuttuu huomattavasti, kun siirrytään reaalisisista matriiseista kompleksisiin. Edellisellä matriisilla A on nimittäin kompleksitasossa ominaisarvona kompleksiluku i , sillä voimme laskea, että

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

Vastaavasti myös $-i$ on ominaisarvo, ominaisvektorilla $\begin{pmatrix} 1 & i \end{pmatrix}^T$. Lisää kompleksisten matriisien ominaisarvoista myöhemmin tässä luvussa.

Ominaisvaruudet

Joissain tapauksissa samaa ominaisarvoa voi vastata useampi erisuuntainen ominaisvektori. Esimerkiksi identiteettimatriisilla on vain yksi ominaisarvo 1, jolle jokainen vektori

on ominaisvektori. Tällaisten tapausten takia on luonnollista puhua matriisin A ominaisarvon λ **ominaisavaruudesta**, joka on kaikkien ominaisarvoa λ vastaavien ominaisvektoreiden joukko. Tämä joukko on luonnollisesti aliavaruus (pohdi miksi).

Esimerkiksi jos matriisin ydin on epätriviaali aliavaruus (eli sisältää muutakin kuin nollavektorin), niin jokainen ytimen nollasta poikkeava vektori on matriisin ominaisvektori ominaisarvolla 0. Ydin on tällöin luonnollisesti ominaisarvon 0 ominaisavaruus.

Determinantille on myös uusi laskutapa ominaisarvojen avulla.

Lause 7.1. *Olkoon $n \times n$ -matriisilla A n kappaletta ominaisarvoja, jotka ovat $\lambda_1, \dots, \lambda_n$. Determinantille pätee tällöin*

$$\det A = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n.$$

Luonnollisesti determinantti on nolla tasan silloin, kuin joku ominaisarvoista on nolla. Tämän voi päätellä myös siitä, että silloin matriisin A ydin on epätriviaali (kts. Lauseet 4.1 ja 6.1).

Diagonalisointi

Matriisia, jonka diagonaalin ulkopuoliset alkioit ovat kaikki nollia, kutsutaan **diagonaalimatriisiksi** tai **lävistämatriisiksi**. Tutkitaan yleistä diagonaalimatriisia

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Tätä voidaan merkitä lyhyemmin myös $D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Matriisin D diagonaali-alkiot λ_j ovat täsmälleen sen ominaisarvot, ja niitä vastaavat ominaisvektorit ovat standardikannan vektorit e_j .

Sanomme, että neliömatriisi A on **diagonalisoituva**, mikäli löytyy sellainen kääntyvä matriisi S ja diagonaalimatriisi D , jolle

$$A = SDS^{-1}$$

Matriisin diagonalisointi on läheisessä yhteydessä sen ominaisarvoihin.

Lause 7.2. *$n \times n$ -Neliömatriisi A on diagonalisoituva tasan silloin, kun sillä on n kappaletta lineaarisesti riippumattomia ominaisvektoreita $\mathbf{v}_1, \dots, \mathbf{v}_n$. Jos näitä vastaavat ominaisarvot ovat $\lambda_1, \dots, \lambda_n$, niin matriisi A diagonalisoituu kaavalla*

$$A = SDS^{-1}, \quad \text{missä } D = \text{diag}(\lambda_1, \dots, \lambda_n) \quad \text{ja} \quad S = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n).$$

Todistus. Olkoon matriisi A ensin diagonalisoituva. Jos kerrotaan oikealta puolelta matriisilla S , saadaan diagonalisaatiolle myös muoto $AS = SD$. Jos nyt S on sarake-muodossa $S = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n)$, niin standardikantavektoreita $\mathbf{e}_j$ kertomalla saadaan

$$A\mathbf{v}_j = ASe_j = SDe_j = S(\lambda\mathbf{e}_j) = \lambda\mathbf{v}_j.$$

Tästä voidaan päätellä suoraan, että diagonalisointiossa esiintyvän matriisin S sarakkeet paljastavat matriisin A ominaisvektorit, ja niitä vastaavat ominaisarvot löytyvät matriisin D diagonaalilta.

Jos taas matriisilla A on n lineaarisesti riippumatonta ominaisvektoria $\mathbf{v}_1, \dots, \mathbf{v}_n$, määritellään S ja D kuten edellä. Koska matriisin S sarakkeet ovat lineaarisesti riippumattomia, se on kääntövä. Koska kaava $AS\mathbf{e}_j = SD\mathbf{e}_j$ pätee kaikilla standardikannan vektoreilla $\mathbf{e}_j$, se pätee lineaarisuuden perusteella myös jokaisella avaruuden $\mathbb{R}^n$ vektorilla. Siis $AS = SD$, eli $A = SDS^{-1}$.

Symmetrisen matriisin diagonalisointi

Osoittautuu, että jokaisen symmetrisen matriisin voi diagonalisoida, jopa niin, että matriisi S valitaan ortogonaaliseksi.

Lause 7.3. *Olko A symmetrinen matriisi. Tällöin on olemassa ortogonaalinen matriisi Q siten, että*

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

missä $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ on diagonaalimatriisi. Tällöin luvut λ_j ovat matriisin A ominaisarvot, ja matriisin Q sarakkeet ovat niitä vastaavat ominaisvektorit.

Diagonaaliesityksen voi kirjoittaa myös muodossa $A = QDQ^T$, sillä ortogonaalisuudesta seuraa $Q^{-1} = Q^T$.

Diagonalisoidun matriisin potensseja on helppo laskea, esimerkiksi

$$A^2 = QDQ^{-1}QDQ^{-1} = QD^2Q^{-1},$$

Ja yleisesti $A^k = QD^kQ^{-1}$, sekä $D^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$.

Todistus. Lauseen 7.2 perusteella riittää siis osoittaa, että matriisilla A on n kappaletta lineaarisesti riippumattomia ominaisvektoreita.

Vaikka Lause 7.3 puhuu vain reaalisista matriiseista ja ominaisarvoista, tarvitsemme sen todistamiseen kompleksilukuja. Syy tähän piilee karakteristisessa polynomissa: Matriisin A ominaisarvot määräytyvät karakteristisen polynomin juurista, mutta kaikilla polynomeilla ei välttämättä ole reaalisia juuria. Sen sijaan kompleksitasossa pätee seuraava vahva tulos.

Algebran peruslause: Jokaisella polynomilla astetta n on tasan n juurta kompleksitasossa, mukaanlukien juurten kertaluvut.

Esimerkiksi polynomilla $x^2 + 1$ ei ole reaalisia juuria, mutta kompleksitasossa sille löytyy helposti juuret $x = \pm i$.

8. Neliömuodot

Matematiikassa, erityisesti differentiaalilaskennassa ja optimointiongelmassa, tulee usein vastaan lausekkeita kuten

$$q(x, y) = x^2 + 4xy + 3y^2.$$

Tällainen lauseke on esimerkki (kahden muuttujan) **neliömuodosta**, ja sille on olemassa myös matriisiesitys:

$$q(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Matriisiesityksessä kohtiin $(1, 1)$ ja $(2, 2)$ tulee kertoimet termeistä x^2 ja $3y^2$, ja ristitermin $4xy$ kerroin jaettiin kahtia kohtiin $(1, 2)$ ja $(2, 1)$. Näin saatu matriisi $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ on symmetrinen. Yleisesti ottaen jokainen symmetrinen $n \times n$ -matriisi A määrää **neliömuodon**

$$q : \mathbb{R}^n \rightarrow \mathbb{R}, \quad q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x},$$

jossa esiintyy n muuttujaa, eli $n \times 1$ -vektorin $\mathbf{x}$ komponentit.

Huom! Symmetrisyys on tässä lähinnä sopimuskysymys, sillä vaikka lausekkeessa $\mathbf{x}^T B \mathbf{x}$ esiintyvä matriisi B ei olisi symmetrinen, voidaan silti kirjoittaa

$$\mathbf{x}^T B \mathbf{x} = \mathbf{x}^T A \mathbf{x}, \quad \text{missä } A = \frac{B + B^T}{2} \text{ on symmetrinen.}$$

Esimerkki 8.1. Kaksi objektia on yhdistetty toisiinsa jousella, ja molemmat on yhdistetty jousilla myös viereisiin seiniin. Jos muuttujat x ja y kuvaavat objektien poikkeamaa lepotilasta, niin suure

$$E(x, y) = \frac{1}{2}k_1x^2 + \frac{1}{2}k_2y^2 + \frac{1}{2}k_3(x - y)^2$$

kuvaa systeemin potentiaalienergiaa kokonaisuudessa. Tämä lauseke on myös neliömuoto.

Definiittisyys ja ominaisarvot

Neliömuotojen perusteoria keskittyy ymmärtämään sitä, miten matriisin A ominaisuudet vaikuttavat neliömuodon $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ saamiin arvoihin. Osoittautuu, että matriisin A ominaisarvot määräävät neliömuodon $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ yleisen käyttäytymisen täysin. Tämä johtuu siitä, että symmetrinen matriisi A voidaan diagonalisoida

$$A = Q^T D Q, \quad \text{missä } D = \text{diag}(\lambda_1, \dots, \lambda_n), \text{ ja } Q \text{ on ortogonaalinen.}$$

Tällöin voidaan huomata, että koordinaattimuunnos $\mathbf{y} = Q\mathbf{x}$ yksinkertaistaa neliömuotoa huomattavasti muotoon

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{x}^T A \mathbf{x} = \mathbf{x}^T Q^T D Q \mathbf{x} = \mathbf{y}^T D \mathbf{y} \\ &= \lambda_1 y_1^2 + \dots + \lambda_n y_n^2. \end{aligned}$$

Voimme erityisesti päätellä neliömuodon arvojen positiivisuudesta/negatiivisuudesta suoraan ominaisarvojen perusteella. Oleelliset luokitukset ovat:

Luokka	Ominaisarvot	$q(\mathbf{x})$, kun $\mathbf{x} \neq \mathbf{0}$
Positiivisesti definiitti	Kaikki $\lambda_i > 0$	$q(\mathbf{x}) > 0$
Positiivisesti semidefiniitti	Kaikki $\lambda_i \geq 0$	$q(\mathbf{x}) \geq 0$
Negatiivisesti definiitti	Kaikki $\lambda_i < 0$	$q(\mathbf{x}) < 0$
Negatiivisesti semidefiniitti	Kaikki $\lambda_i \leq 0$	$q(\mathbf{x}) \leq 0$
Indefiniitti	Eri merkkisiä ominaisarvoja	$q(\mathbf{x})$ saa molempia merkkejä

9. Matriisihajotelmia

Edellisessä luvussa esitelly diagonalisointi on vain yksi monista tavoista jakaa matriisi tuloksi yksinkertaisemmista matriiseista, ja monet muut eri **matriisihajotelmat** ovat vahvoja työkaluja, joita matriisilaskenta tarjoaa käytännön ongelmien ratkaisemiseen. Tutustumme tässä luvussa kahteen uuteen perushajotelmaan nimeltä **LU-hajotelma** ja **SVD-hajotelma**, mutta monia muitakin hyödyllisiä hajotelmia löytyy näiden lisäksi.

LU-hajotelma

Luvussa 3 tutustuimme lineaarisen yhtälöryhmän $A\mathbf{v} = \mathbf{b}$ ratkaisemiseen Gaussin eliminaatiolla. Osoittautuu, että tätä prosessia voidaan soveltaa myös hajoittamaan (harvoja poikkeuksia lukuunottamatta) jokainen neliömatriisi kahden matriisin tuloksi, joista ensimmäinen on **alakolmiomatriisi** ja toinen **yläkolmiomatriisi**. Neliömatriisia sanotaan (ala/ylä)kolmiomatriisiksi, jos sen diagonaalin (ylä/ala)puolella on pelkkiä nollia. Siis esimerkiksi

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 7 & 1 & 0 \\ -1 & 8 & 3 \end{pmatrix} \text{ on alakolmiomatriisi, ja } U = \begin{pmatrix} -2 & 5 & 9 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \text{ on yläkolmiomatriisi.}$$

Jos neliömatriisin A voi esittää muodossa $A = LU$ sopivilla ala/yläkolmiomatriiseilla L ja U , tätä sanotaan sen **LU-hajotelmaksi**. Esimerkiksi 3×3 -matriisin LU-hajotelma voisi näyttää tältä:

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 3 \\ 8 & 7 & 9 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix}}_L \underbrace{\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}}_U = LU. \quad (9.1)$$

Tulemme näkemään, että LU-hajotelma on oleellisesti matriisimuotoinen Gaussin eliminaatio, ja tästä syystä sen sovelluskohteet ovat myös samat, kuten: Yhtälöryhmien ratkaiseminen, käänteismatriisin muodostaminen, ja determinantin laskeminen. Erityisesti tietokoneet käyttävät pääasiassa LU-hajotelmaa yhtälöryhmien ratkaisemisessa.

LU-hajotelman muodostaminen

Annetun matriisin LU-hajotelman (mikäli sellainen on olemassa) voi löytää soveltamalla Gaussin eliminaatiota matriisiin seuraavasti. Rivioperaatioiden kertoimista täytyy kuitenkin pitää kirjaa, sillä niiden avulla muodostetaan matriisi L . Tutkitaan esimerkkinä matriisia

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 3 \\ 8 & 7 & 9 \end{pmatrix}$$

Gaussin eliminaatiolla saadaan

$$\begin{array}{ccc} \begin{array}{ccc} 2 & 1 & 1 \\ 4 & 3 & 3 \\ 8 & 7 & 9 \end{array} & \begin{array}{l} \leftarrow -2 \cdot (\text{Rivi } 1) \\ \leftarrow -4 \cdot (\text{Rivi } 1) \end{array} & \Rightarrow \begin{array}{ccc} 2 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 3 & 5 \end{array} \begin{array}{l} \\ \leftarrow -3 \cdot (\text{Rivi } 2) \end{array} \Rightarrow \begin{array}{ccc} 2 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{array} \end{array}$$

Lopputuloks on porrasmuodossa, ja porrasmuodossa oleva matriisi on täsmälleen haluttu yläkolmiomatriisi U . Alakolmiomatriisi L löytyy palauttamalla mieleen eliminaatiossa tehdyt rivioperaatiot seuraavasti

- Alakolmiomatriisin L diagonaalille asetetaan pelkkiä ykkösiä (ja sen yläpuolelle nollia).
- Jos Gaussin eliminaatiossa riviin j lisättiin rivi i kerrottuna jollain kertoimella a , niin matriisin L kohtaan (j, i) asetetaan tämän kertoimen vastaluku $-a$.

Edellisessä esimerkissä kertoimien vastaluvut olivat 2, 4, ja 3, joten

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix}.$$

Näin ollaan löydetty hajotelma $A = LU$, joka vastaa yhtälöä (9.1). Sama menetelmä toimii aina, **kun eliminaatiossa rivejä ei tarvitse vaihtaa keskenään**. Mikäli rivien vaihtoja tarvitaan, on olemassa yleisempi PLU -hajotelma, josta lisää hetken päästä. Perustellaan kuitenkin ensin, miksi yllä esitelty menetelmä oikeastaan toimii.

LU-hajotelman todistus: Gaussin eliminaatiossa sovellettuja rivioperaatioita voi tulkita myös matriisimuodossa. Edellisessä esimerkissä matriisille A toteutetut rivioperaatiot

$$\mathbf{R}_2 \Rightarrow \mathbf{R}_2 - 2\mathbf{R}_1, \quad \mathbf{R}_3 \Rightarrow \mathbf{R}_3 - 4\mathbf{R}_1, \quad \text{ja} \quad \mathbf{R}_3 \Rightarrow \mathbf{R}_3 - 3\mathbf{R}_2$$

vastaavat samaa, kuin matriisin A kertoisi vasemmalta puolelta matriiseilla

$$L_{21} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad L_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} \quad \text{ja} \quad L_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{pmatrix}.$$

Siis Gaussin eliminaatiosta saatu lopputuloks voidaan ilmaista tässä tapauksessa yhtälönä

$$L_{32}L_{31}L_{21}A = U.$$

Tästä saadaan ratkaistua

$$A = \underbrace{L_{21}^{-1}L_{31}^{-1}L_{32}^{-1}}_L U,$$

Sievennetään vielä matriisin L määrittävä tulo. Huomataan, että esimerkiksi rivioperaation L_{32} käänteinen toimenpide on operaatio $\mathbf{R}_3 \Rightarrow \mathbf{R}_3 + 3\mathbf{R}_2$, eli sama operaatio, mutta erimerkkisellä kertoimella. Täten

$$L_{32}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}.$$

Jos kirjoitetaan $L = L_{21}^{-1}L_{31}^{-1}L_{32}^{-1}I$, niin voidaan päätellä, että matriisi L voidaan selvittää soveltamalla identiteettimatriisiin I kaikki eliminaatiossa tapahtuneiden operaatioiden käänteisoperaatiot, ja myös käänteisessä järjestyksessä. Tehdään juuri näin:

$$\begin{array}{ccc} \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} & \Rightarrow & \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{array} \leftarrow + \begin{array}{c} 2 \\ 4 \end{array} \cdot (\text{Rivi } 1) \\ \leftarrow + 3 \cdot (\text{Rivi } 2) & & \end{array} \Rightarrow \begin{array}{ccc} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{array}.$$

Siis alakolmiomatriisin L diagonaalien alle todellakin tulee eliminaatiossa käytettyjen kertoimien vastaluvut.

PLU-hajotelma

Tarkalleen ottaen syy, miksi rivioperaatioiden kautta muodostuva matriisi L on juuri alakolmiomatriisi perustuu siihen, että rivioperaatioissa jokaiseen riviin lisätään ainoastaan sitä ylempanä olevia rivejä. Tästä syystä LU-hajotelma ei tepsii esimerkiksi matriisille

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

sillä tämän matriisin rivit täytyy ensin vaihtaa keskenään. Tämä onnistuu matriisimuodossa kertomalla A vasemmalta ns. **permutaatiomatriisilla**

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

jolloin siis

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

Yleinen permutaatiomatriisi on sellainen neliömatriisi, jonka jokaisella rivillä ja sarakeella on tasan yksi ykkönen, ja kaikki muut matriisin alkiot ovat nollia. Esimerkiksi

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

Permutaatiomatriisilla kertominen vaihtaa kerrottavan matriisin rivien paikkaa. Esimerkiksi tämä matriisi P vaihtaa rivejä järjestyksessä $1 \rightarrow 3 \rightarrow 4 \rightarrow 2 \rightarrow 1$.

Jokaisen neliömatriisin rivit voidaan järjestää uudelleen niin, että edellä esitetty menetelmä tuottaa LU-hajotelma jälleen toimii. Siis jokaiselle neliömatriisille A löytyy permutaatiomatriisi P siten, että

$$PA = LU \quad \text{tai} \quad A = P^{-1}LU.$$

Tätä sanotaan PLU-hajotelmaksi.

LU-hajotelman motivaatio

Tässä vaiheessa lukija saattaa vielä pohtia kysymystä:

Miksi tehdä LU-hajotelma? Olemme näyttäneet, että LU/PLU-hajotelman muodostaminen perustuu täysin Gaussin eliminaatioon, joten on syytä pohtia, miksi tehdä tällainen hajotelma, kun voisimme vain käyttää Gaussin eliminaatiota sen sijaan.

Tutkitaan tilannetta, jossa meillä on ratkaistavana tuhat yhtälöryhmää

$$\begin{aligned} A\mathbf{v}_1 &= \mathbf{b}_1 \\ &\vdots \\ A\mathbf{v}_{1000} &= \mathbf{b}_{1000}, \end{aligned}$$

joissa kaikissa esiintyy sama matriisi A . Jokaiseen yhtälöryhmään voidaan toki soveltaa Gaussin eliminaatiota yksitellen, mutta tämä on varsin työlästä. Sen sijaan matriisin

A LU-hajotelma (tai PLU) tarvitsee muodostaa vain kerran, ja sen avulla jokainen yhtälöryhmä $A\mathbf{v}_j = \mathbf{b}_j$ voidaan palauttaa (mahdollista rivien uudelleenjärjestämistä vaille) muotoon $LU\mathbf{v}_j = \mathbf{b}_j$. Tämä voidaan mieltää kahtena eri systeeminä

$$L\mathbf{u}_j = \mathbf{b}_j \quad \text{ja} \quad U\mathbf{v}_j = \mathbf{u}_j.$$

Nämä ovat huomattavasti suoraviivaisempia ratkaista, sillä (ylä/ala)kolmiomatriisin määrittävä yhtälöryhmä voidaan ratkaista yksi muuttuja kerrallaan aloittaen alhaalta/ylhäältä.

Siis siinä missä Gaussin eliminaatio on menetelmä yhtälöryhmän ratkaisemiseen, LU/PLU-hajotelma on esitystapa annetulle matriisille, jota voidaan hyödyntää uudelleen aina tarvittaessa.

SVD-hajotelma

10.

Sanasto

Affinikuvaus

affine transformation

Algebraillinen kertaluku

algebraic multiplicity

Aliavaruus

subspace

(Säännöllisyys)aste, rangi

rank

Cholesky-hajotelma

Cholesky factorization

Cramerin sääntö

Cramer's rule

Determinantin kaava

formula for determinant

Determinantti

determinant

Diagonalisaatio

diagonalization

Lineaarikuvaus sekä vakiotermi, muotoa $T(v) = Av + v_0$.

Ominaisarvon λ_i algebraillinen kertaluku kertoo, kuinka monta kertaa λ_i esiintyy karakteristisen polynomin $\det(A - \lambda I)$ juurena, eli kuinka monta kertaa termi $(\lambda - \lambda_i)$ jakaa karakteristisen polynomin. Toisin sanottuna, polynomin $\det(A - \lambda I)$ voi esittää muodossa $\prod_{i=1}^k (\lambda_i - \lambda)^{n_i}$, missä λ_i ovat A :n ominaisarvot, ja n_i on ominaisarvon λ_i algebraillinen kertaluku.

Joukko V on $\mathbb{R}^n$:n aliavaruus, jos se sisältää kaikki jäsentensä lineaarikombinaatiot. Triviaaleja aliavaruuksia ovat $\emptyset, \{0\}$, ja $\mathbb{R}^n$. Yleisemmin vektoriavaruuden V aliavaruus W on osajoukko, joka on itsessään myös vektoriavaruus.

Matriisin rivien tai sarakkeiden virittämän vektoriavaruuden ulottuvuus. Riippumattomien sarakkeiden tai rivien lukumäärä.

Muotoa $A = CC^T$ tai $A = (L\sqrt{D})(L\sqrt{D})^T$.

Ratkaisukaava yhtälöryhmälle $Ax = b$. Cramerin sääntö antaa $x_j = \det(B_j) / \det(A)$, missä B_j on muuten kuin A , mutta sarake numero j on korvattu vektorilla b .

Matriisin A determinantti on $n!$ termin summa. Jokainen termi vastaa yhtä permutatiomatriisia P niin, että termi on AP :n lävistäjäalkioiden tulo kertaa $\det(P)$. Toisin sanottuna $\det(A) = \sum_P \det(P) \prod_{i=1}^n (AP)_{ii}$.

Määriteltä neliömatriiseille. Merkitään $\det(A)$, joskus $|A|$. Determinantti määritellään seuraavan kolmen ominaisuuden kautta; identiteettimatriisille pätee $\det(I) = 1$; jos A :n kahden rivin paikkaa vaihtaa, determinantti pysyy muuten samana, mutta sen merkki vaihtuu; determinantti on lineaarinen jokaisella rivillä. Tästä seuraa, että $\det(A) = 0$ silloin kun A on epäsäännöllinen.

Hajotelma muotoa $A = SDS^{-1}$, tai vaihtoehtoisesti $AS = SD$, missä D on diagonaalimatriisi joka pitää sisällään A :n ominaisarvot, ja S koostuu A :n ominaisvektoreista. Huom! Diagonalisaatio on olemassa vain jos A :lla on n riippumattonta ominaisvektoria. Diagonalisaatiolla on helppo laskea matriisin potensseja: $A^k = SD^k S^{-1}$. Matriisit A ja D ovat similaarisia.

Diagonalisoituva matriisi*diagonalizable matrix*

Matriisi A on diagonalisoituva, jos sille voidaan muodostaa hajotelma $A = SDS^{-1}$, missä D on diagonaalimatriisi. Tämä on mahdollista silloin, kun A :lla on n riippumatonta ominaisvektoria.

Dimensio, ulottuvuus*dimension*

Vektoriavaruuden V dimensio on sen kantavektorien lukumäärä. Huom. vaikka kanta voidaan valita usealla eri tavalla, sen vektoreiden lukumäärä on aina sama.

Eliminaatio*elimination*

Sarja rivioperaatioita, joilla redusoidaan matriisi A joko yläkolmiomuotoon U tai redusoituun porrasmuotoon R . Tällöin $A = LU$, missä L koostuu *kertoimista* ℓ_{ij} . Vaihtoehtoisesti $PA = LU$, missä P on permutaatiomatriisi joka uudelleenjärjestee A :n rivit, tai $EA = R$, missä E on jokin kääntyvä (säännöllinen) matriisi.

Eliminaatiomatriisi*elimination matrix*

Eliminaatiomatriisi E_{ij} on muuten kuin identiteettimatriisi, mutta sen arvo kohdassa (i, j) on $-\ell_{ij}$. Tämä matriisi vastaa rivioperaatiota, jossa matriisista A vähennetään ℓ_{ij} kertaa rivi j rivistä i . Matriisitulo $E_{ij}A$ on tämän rivioperaation lopputulos.

Ellipsi ja ellipsoidi*ellipse, ellipsoid*

Ellipsin kaava voidaan esittää muodossa $x^T Ax = 1$, missä matriisi A on positiivisesti definiitti. Ellipsin akselit ovat matriisin A ominaisvektoreita. Akselien pituus on $1/\sqrt{\lambda}$, missä λ on A :n ominaisarvo.

Ensimmäisen asteen matriisi*rank-1 matrix*

Matriisi, jonka rivi- ja sarakeavaruudet ovat yksiulotteisia. Kaikki matriisin sarakkeet (ja rivit) voidaan siis muodostaa yhtä saraketta (riviä) skaalaamalla, eli $A = uv^T$, missä u ja v ovat pystyvektoreita. Huom. $A \neq 0$.

Epäsäännöllinen matriisi, singulaarinen matriisi*singular matrix*

Neliömatriisi, jonka aste on pienempi kuin sen koko. Myös, neliömatriisi, jolla ei ole käänteismatriisia tai neliömatriisi jolla on lineaarisesti riippuvia rivejä tai sarakkeita. Neliömatriiseilla $\det(A) = 0$ jos ja vain jos A on epäsäännöllinen.

Erityisratkaisu*special solution*

Ratkaisu yhtälöryhmälle $As = 0$, jossa yksi vapaa muuttuja on $s_i = 1$ ja muut vapaat muuttujat ovat nollaa.

Gauss-Jordanin menetelmä*Gauss-Jordan method*

Menetelmä käänteismatriisin laskemiseksi. Tehdään rivioperaatioita täydennetyille matriisille $[A \ I]$ kunnes saavutetaan muoto $[I \ A^{-1}]$.

Geometrinen kertaluku*geometric multiplicity*

Ominaisarvon λ riippumattomien ominaisvektorien lukumäärä. Vaihtoehtoisesti ominaisvektorien virittämän avaruuden dimensio.

LU-hajotelma*LU factorization*

Hajotelma muotoa $A = LU$, missä U on yläkolmiomatriisi. Jos matriisia A eliminoidessa ei tarvitse käyttää rivipermutaatioita, matriisi L on muotoa $L_{ij} = \ell_{ij}$ (ks. kerroin), ja $L_{ii} = 1$.

Hermiittinen matriisi*Hermitian matrix*

Kompleksisille matriiseille symmetristä matriisia vastaava käsite. Matriisi on hermiittinen, jos $A^* = \overline{A}^T = A$, missä $\overline{A}$ tarkoittaa A :n alkioittaista kompleksikonjugaattia. Tällöin $a_{ij} = \overline{a_{ji}}$.

Identiteettimatriisi*identity matrix*

Matriisi I , jonka lävistäjäalkiot I_{ii} ovat 1 ja muut alkiot nollaa. Matriisille I pätee $AI = IA = A$, eli se toimii matriisitulossa kuin luku 1 skalaarikertolaskussa.

Indefiniitti matriisi*indefinite matrix*

Symmetrinen matriisi, jolla on sekä positiivisia että negatiivisia ominaisarvoja.

Jälki*trace*

Matriisin lävistäjäalkioiden summa.

Karakteristinen polynomi*characteristic polynomial*

Polynomi $\det(A - \lambda I)$, missä λ on muuttuja. Karakteristisen polynomin juuret ovat matriisin ominaisarvoja.

Kerroin ℓ_{ij} *multiplier*

Kerroin, jota käytetään i . rivin pivotsarakkeen eliminomiseen rivin j avulla. Rivi j kerrotaan kertoimella ℓ_{ij} ja vähennetään rivistä i . Jotta haluttu alkio saadaan eliminotua, kerroin on valittava $\ell_{ij} = A_{ik}/A_{jk}$, missä k on pivotsarake.

Kiertomatriisi*rotation matrix*

Kaksiulotteisille vektoreille $R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$. Operaatio Rv pyöryttää vektoria v kulman θ mukaisesti vastapäivään. Matriisin R_θ käänteismatriisi saadaan käyttämällä kulmaa $-\theta$. Huom! $R_\theta^{-1} = R_\theta^T$, eli R_θ on ortogonaalinen matriisi. Matriisin R ominaisarvot ovat $e^{i\theta}$ ja $e^{-i\theta}$, ja sen ominaisvektorit ovat $(1, i)$ ja $(1, -i)$.

Kolmioepäyhtälö*triangle inequality*

Vektoreille u ja v pätee $|u + v| \leq |u| + |v|$.

Kolmilävistäjämatrissi, kolmidiagonaalinen matrissi*tridiagonal matrix*

Kuin lävistäjämatrissi, mutta myös alkiot lävistäjän ”lähellä” voivat olla muuta kuin nollaa. Toisin sanottuna matrissi T on kolmilävistäjämatrissi, jos $T_{ij} = 0$ aina kun $|i - j| > 1$.

Kompleksikonjugaatti*complex conjugate*

Kompleksiluvun ”peilaus” x -akselin suhteen. Jos $z = x + iy$, niin z :n kompleksi konjugaatti on $\bar{z} = x - iy$. Huom. $z\bar{z} = x^2 + y^2 = |z|^2$.

Kuva-avaruus $\text{Im}(A)$ *image space, range, column space*

Matriisin A sarakkeiden virittämä vektoriavaruus. A :n sarakkeiden kaikkien mahdollisten lineaariyhdistelyiden joukko.

Käänteismatriisi*inverse*

Neliömatriisin A käänteismatriisi A^{-1} toteuttaa $AA^{-1} = I$. Käänteismatriisi on yksikäsitteinen, ja matriisin A^{-1} käänteismatriisi on A . Matriisilla on käänteismatriisi jos ja vain jos $\det(A) \neq 0$, eli kun se on säännöllinen.

Liitännäisyys*associativity*

Matriisien tulolle pätee $(AB)C = A(BC)$, eli tulot voi laskea missä järjestyksessä tahansa. Näin ollen sulut voidaan jättää pois ja kirjoittaa vain ABC .

Lineaarikuvaus*linear transformation*

Funktio T , joka kuvaa vektorin v joksikin toiseksi vektoriksi $T(v)$. Ollakseen lineaarinen kuvauksen T on täytettävä ehto $T(av + bw) = aT(v) + bT(w)$ kaikille skalaareille a, b ja vektoreille v, w .

Lineaarikombinaatio*linear combination*

Vektorien skaalausta ja summausta. Esim. skalaareille $a_1, \dots, a_n$ ja vektoreille $v_1, \dots, v_n$ summa $\sum_{i=1}^n a_i v_i$ on vektorien $v_1, \dots, v_n$ lineaariyhdistely.

Lineaarinen riippuvuus*linear dependency*

Vektorit $v_1, \dots, v_n$ ovat lineaarisesti riippuvia, jos on olemassa kertoimet $a_1, \dots, a_n$ siten että $\sum_{i=1}^n a_i v_i = 0$ ja $a_j \neq 0$ jollekin $j = 1, \dots, n$.

Lohkomatriisi*block matrix*

Matriisin voi jaotella osiin niin, että matriisi koostuu pienemmistä matriiseista, joita kutsutaan lohkoiksi.

Lohkomatriisien tulo*block product*

Lohkoihin jaettujen matriisien tulon voi laskea lohkojen avulla, mikäli lohkojen koot sopivat yhteen. Esim.

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$
Lävistäjämatrissi, diagonaalimatrissi*diagonal matrix*

Matriisi, jonka alkiot ovat nollaa kaikkialla muualla, paitsi lävistäjällä, eli kohdissa (i, i) .

Markovin matrissi*Markov's matrix*

Kaikki matriisin M arvot ovat epänegatiivisia ja jokaisen sarakkeen alkioden summa on 1. Suurin ominaisarvo on 1. Jos $M_{ij} > 0$, matriisin potenssien M^k sarakkeet lähestyvät tasapainotilan ominaisvektoria $Ms = s > 0$.

Matriisitulo*matrix product*

Matriisit A ja B voidaan kertoa keskenään, jos niiden koot ovat yhteensopivat; matriisilla A on oltava yhtä monta saraketta kun B :llä on rivejä. Matriisitulon AB alkio (i, j) muodostetaan summana $\sum_{k=1}^k A_{ik}B_{kj}$. Toisin sanottuna $(AB)_{ij}$ on rivin A_i ja sarakkeen B_j pistetulo. Jos matriisit A ja B ovat kokoa $m \times k$ ja $k \times n$, matriisi AB on kokoa $m \times n$.

Matriisi-vektoritulo*matrix-vector product*

Matriisin A ja pystyvektorin x tulo on lineaariyhdistely A :n sarakkeista; $Ax = \sum_{i=1}^n A_i x_i$. Lopputulos on siten $m \times 1$ vektori, jos A on kokoa $m \times n$.

Nilpotentti matrissi*nilpotent matrix*

Matriisi A on nilpotentti, jos jollekin $k = 0, 1, \dots$ pätee $A^k = 0$. Esimerkiksi yläkolmiomatriisit, joiden lävistäjäalkiot ovat nollaa, ovat nilpotentteja.

Nolla-avaruus,
ydin $\text{Ker}(A)$
nullspace, kernel

Nolla-avaruusmatriisi
nullspace matrix

Normaali matriisi
normal matrix

Normaaliyhtälö
normal equation

Ortogonaalinen matriisi
orthogonal matrix

Ortogonaaliset aliavaruudet
orthogonal subspaces

Ortonormaalit vektorit
orthonormal vectors

Ominaisarvo,
ominaisvektori
eigenvalue, eigenvector

Ominaisarvolause
spectral theorem

Niiden vektoreiden x joukko, joille $Ax = 0$. Näiden vektorien joukko muodostaa vektoriavaruuden. Jos A on $m \times n$ matriisi jonka säännöllisyysaste on r , nolla-avaruuden dimensio on $n - r$.

Nolla-avaruusmatriisin N sarakkeet ovat yhtälöryhmän $As = 0$ erityisratkaisut. Erityisratkaisuja on aina $n - r$, missä n on A :n sarakkeiden lukumäärä ja r on A :n säännöllisyysaste.

Matriisi A , jolle pätee $AA^T = A^T A$. Tällöin A :n ominaisvektorit ovat ortonormaaleja.

Yhtälöryhmä $A^T A \hat{x} = A^T b$. Jos A on täysasteinen, normaaliyhtälön ratkaisut ovat pienimmän neliösumman ratkaisuja yhtälöryhmälle $Ax = b$. Yhtälö kertoo, että kaikki A :n sarakkeet ovat ortogonaalisia vektorin $b - A\hat{x}$ kanssa.

Matriisi on ortogonaalinen, jos $Q^T Q = I$. Ortogonaalisille matriiseille pätee $Q^T = Q^{-1}$, eli käänteismatriisi on helppo laskea. Ortogonaaliset matriisit säilyttävät vektorien pituudet, eli $|Qx| = |x|$ ja $(Qx)^T(Qy) = x^T Q^T Q y = x^T y$. Kaikki Q :n ominaisarvot ovat kokoa $|\lambda| = 1$, ja ominaisvektorit ovat myös ortogonaalisia. Esimerkiksi kiertomatriisit, peilausmatriisit ja permutaatiomatriisit ovat ortogonaalisia.

Aliavaruudet V ja W ovat ortogonaalisia, jos kaikki vektorit $v \in V$ ja $w \in W$ ovat ortogonaalisia keskenään, eli $v^T w = 0$.

Vektorit $q_1, \dots, q_n$ ovat ortonormaaleja keskenään, jos $q_i^T q_j = 0$ aina kun $i \neq j$ ja $|q_i| = 1$. Jos ortonormaalit vektorit järjestetään matriisin Q sarakkeiksi, niin $Q^T Q = I$, eli Q on ortogonaalinen matriisi. Jos Q on neliömatriisi, sen sarakkeet muodostavat ortonormaalin kannan avaruudelle $\mathbb{R}^n$; jokainen vektori $\mathbb{R}^n$:ssä voidaan ilmaista Q :n sarakkeiden lineaariyhdistelyinä.

Neliömatriisin A ominaisarvo on luku, jolle on olemassa vektori $x \neq 0$ siten että $Ax = \lambda x$. Tällöin x on ominaisarvoa λ vastaava ominaisvektori. Ominaisarvoille λ pätee myös $\det(A - \lambda I) = 0$

Jos A on reaalinen symmetrinen neliömatriisi, niin on olemassa hajotelma $A = QDQ^T$, missä D on reaalinen lävistäjämatriisi, joka pitää sisällään A :n ominaisarvot, ja Q on ortogonaalinen matriisi, jonka sarakkeet ovat A :n ominaisarvoja. Mekaniikassa Q :n sarakkeita kutsutaan pääakseleiksi.

Osittelulaki

distributivity

Peilausmatriisi

reflection matrix

Permutaatiomatriisi

permutation matrix

Pienimmän

neliösumman ratkaisu

least-squares solution

Pistetulo

dot product

Pituus

length

Pivot

pivot

Pivot-sarakkeet

pivot columns

Porrasmuoto

echelon matrix

Positiivisesti

definiitti matriisi

positive-definite matrix

Projektio suoralle

projection onto a line

Matriisitulolle pätee $A(B + C) = AB + AC$.

Matriisi muotoa $Q = I - 2uu^T$, missä $|u| = 1$. Matriisi Q peilaa yksikkövektorin u vektoriksi $Qu = -u$. Vektorit, jotka ovat ortogonaalisia u :n kanssa, pysyvät muuttumattomina, sillä ne ovat "peilausakselin" suuntaisia.

Permutaatiomatriisi on matriisi, jolla kertomalla voi uudelleen- järjestellä matriisin rivit tai sarakkeet. Permutaatiomatriisin sarakkeet ovat samat kuin identiteettimatriisilla I , mutta mahdollisesti toisessa järjestyksessä. Tulo PA järjestee matriisin A rivit ja AP sen sarakkeet. Erilaisia järjestyksiä n alkiolle on olemassa $n!$ kappaletta, ja siten myös permutaatiomatriiseja on näin monta. Permutaatiomatriisille pätee $\det(P) = \pm 1$.

Pienimmän neliösumman ratkaisu $\hat{x}$ on ratkaisu, joka minimoi virheen $|e| = |A\hat{x} - b|$. Se on myös ratkaisu normaaliyhtälölle $A^T A\hat{x} = A^T b$.

Vektorien v ja w pistetulo on $v \cdot w = \sum_{i=1}^n v_i w_i = v^T w = w^T v = v_1 w_1 + \dots + v_n w_n$.

Vektorin v pituus $|v|$ on yleistys Pythagoraan lauseesta, ja sen voi laskea kaavalla $|v| = \sqrt{v_1^2 + \dots + v_n^2} = \sqrt{v^T v}$.

Eliminaatiossa ensimmäinen (vasemmalta laskien) nollasta eroava sarake rivillä, jota sillä hetkellä käsitellään.

Porrasmuodossa olevan matriisin kullekin riville se sarake, jolla on ensimmäinen nollasta poikkeava arvo.

Eliminoitu matriisi on järjestetty niin, että alempien rivien pivot-sarakkeet ovat ylempien pivot-sarakkeiden oikealla puolella ja nollarivit ovat alimpana. Näin ollen kunkin rivin pivot-alkion alapuolella on vain nollaa. Matriisin nollat muodostavat porrasmaisen kuvion.

Symmetrinen neliömatriisi, jonka ominaisarvot ovat positiivisia. Määritelmällisesti kaikille vektoreille $x \neq 0$ pätee $x^T A x > 0$. Porrasmuodossa ollessaan matriisin pivot-alkiot ovat positiivisia.

Olkoon L origon kautta kulkeva suora, joka kulkee pisteen a läpi. Projektion matriisi, joka heijastaa vektorin v suoralle L on $P = aa^T / a^T a$, ja projisoitu vektori on muotoa $Pv = a(a^T v / a^T a)$. Matriisin P säännöllisyysaste on 1.

Ratkeava yhtälöryhmä*solvable system*

Yhtälöryhmä $Ax = b$ on ratkeava, jos sille on olemassa ratkaisu x . Yhtälöryhmälle on olemassa ratkaisu täsmälleen silloin kun b kuuluu A :n kuva-avaruuteen. Jos matriisi A on säännöllinen, yhtälöryhmä on aina ratkeava.

Redusoitu porrasmuoto*reduced row echelon form, $\text{rref}(A)$*

Matriisi A on eliminoitu siten, että pivot-alkioiden ylä- ja alapuolella on vain nollaa, ja rivit on skaalattu siten, että pivot-alkiot saavat arvon 1. Redusoidun porrasmuodon rivit muodostavat kannan matriisiin A kuva-avaruudelle.

Riippumattomat vektorit*independent vectors*

Vektorit $v_1, \dots, v_n$ ovat riippumattomia jos ja vain jos ei ole olemassa lineaariyhdistelyä jossa $c_1v_1 + \dots + c_nv_n = 0$ ja $c_j \neq 0$ jollekin $j = 1, \dots, n$. Toisin sanottuna, yhtäkään vektoreista v_i ei voi muodostaa vektoreiden $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ lineaariyhdistelyä.

Ristitulo*cross product*

Ristitulo kolmiulotteisten vektorien u ja v välillä on niille kohtisuora vektori, jonka pituus on $|u||v|\sin\theta$. Sen koordinaatit saadaan matriisin $\begin{bmatrix} i & j & k; u_1 & u_2 & u_3; v_1 & v_2 & v_3 \end{bmatrix}$ determinantin avulla.

Riviavaruus $\text{Im}(A^T)$ *row space*

Matriisin A rivien lineaariyhdistelyiden virittämä vektoriavaruus.

Semidefiniitti matriisi*semidefinite matrix*

Symmetrinen matriisi A on positiivisesti semidefiniitti, jos $x^T Ax \geq 0$ kaikille $x \in \mathbb{R}^n$. Vastaavasti matriisi on negatiivisesti semidefiniitti, jos $x^T Ax \leq 0$.

Permutaation signum*sign of a permutation*

Jokaista permutaatiota σ vastaa permutaatiomatriisi P_σ , jonka voi rakentaa identiteettimatriisin sarakkeita keskenään vaihtamalla. Permutaation signum, $\text{sgn}(\sigma)$, on 1, jos P_σ voidaan muodostaa I :stä parillisella määrällä sarakevaihtoja, -1 muuten. Pätee myös $\text{sgn}(\sigma) = \det(P_\sigma)$. Voidaan myös laskea kaavalla $\text{sgn}(\sigma) = \prod_{i < j} \frac{\sigma(i) - \sigma(j)}{i - j}$.

Similaariset matriisit*similar matrices*

Matriisit A ja B ovat similaarisia, jos on olemassa jokin säännöllinen matriisi M siten että $B = M^{-1}AM$. Matriisit ovat similaarisia, jos niiden ominaisarvot ovat samat.

Skalaarikolmitulo*volume of a box*

3×3 matriisin A rivit (tai sarakkeet) määrittävät suunnikkaan, jonka tilavuus on $|\det(A)|$.

Spektri*spectrum*

Matriisin ominaisarvojen joukko $\{\lambda_1, \dots, \lambda_n\}$. Arvo $\max |\lambda_i|$ on matriisin spektrisäde.

Symmetrinen hajotelma*symmetric factorization*

Hajotelma muotoa $A = LDL^T$ tai $A = QDQ^T$ (L alakolmiomatriisi ja Q ortogonaalinen). Matriisi D on diagonaalinen ja koostuu matriisin A ominaisarvoista. Matriisin D positiiviset pivot-alkiot ja A :n positiiviset ominaisarvot ovat yhtäsuuret.

Symmetrinen matriisi*symmetric matrix***Standardikanta vektori-
avaruudelle $\mathbb{R}^n$** *standard basis for $\mathbb{R}^n$* **Säännöllinen matriisi,
täysiasteinen matriisi***full-rank matrix***Takaisinsijoitus.***back-substitution***Transpoosi***transpose***Täydellinen ratkaisu.***complete solution***Täydennetty matriisi.***augmented matrix***Täysi sarakeaste, täysi
riviaste***full column rank, full row rank***Ulkotulo***outer product***Unitaarinen matriisi***unitary matrix***Vaihdannaisuus***commutativity***Vandermonde-matriisi***Vandermonde matrix***Vapaa muuttuja***free variable***Vapaat sarakkeet***free columns*

Matriisi on symmetrinen, jos $A = A^T$.

Identiteettimatriisin sarakkeet, eli n -ulotteiset vektorit, joiden ainoa nollasta eroava alkio on 1, muodostavat kan-
nan avaruudelle $\mathbb{R}^n$.

Matriisi, jonka säännöllisyysaste on $\min(m, n)$.
Neliömatriisi, jolla on käänteismatriisi. Matriisi, jonka
sarakkeet (tai rivit) ovat lineaarisesti riippumattomia.

Kun yhtälöryhmä on saatettu yläkolmiomuotoon,
voidaan ratkaisu selvittää muuttuja kerrallaan ”alhaalta
ylös”, muuttujasta x_n muuttuun x_1 .

Matriisin ”peilaaminen” sen lävistäjän suhteen. Päte
 $(A^T)_{ij} = A_{ji}$.

Yhtälöryhmän kaikkien ratkaisujen ilmaiseminen muo-
dossa $x = x_p + x_n$, missä x_p on yksittäisratkaisu ja x_n
toteuttaa yhtälöryhmän $Ax_n = 0$ (eli $x_n \in N(A)$).

Yhtälöryhmä $Ax = b$ ilmaistuna muodossa $[Ab]$, missä
matriisia A on täydennetty sarakkeella b .

Matriisin kaikki sarakkeet (tai rivit) ovat lineaarisesti ri-
ippumattomia. Vastaavasti matriisin nolla-avaruus on
triviaali $N(A) = \emptyset$ ja sen kuva-avaruuden ulottuvuus on
matriisin sarakkeiden lukumäärä.

Vektorioperaatio kahden pystyvektorin välillä, merkitään
 uv^T . Jos $u \in \mathbb{R}^m$ ja $v \in \mathbb{R}^n$, tuloksena on $m \times n$ matriisi
jonka rangi on yksi. Lasketaan $(uv^T)_{ij} = u_i v_j$.

Matriisi U on unitaarinen jos $U^*U = I$. Toisinsan-
ottuna, matriisin U konjugaattitranspoosi on myös sen
käänteismatriisi.

Jos matriiseille A ja B pätee $AB = BA$, sanotaan että
matriisit kommutoivat keskenään. Ei yleensä päde ma-
triisitulolle. Jos matriisit lisäksi diagonalisoituvia, niiden
ominaisvektorit ovat samat.

Matriisi V , jolle $Vc = b$ määrittelee polynomin $p(x) =$
 $c_0 + \dots + c_{n-1}x^{n-1}$ jolle $p(x_i) = b_i$. Määritellään $V_{ij} =$
 $(x_i)^{j-1}$. Matriisin V determinantti on $\prod_{j>i}(x_j - x_i)$.

Muuttuja x_i on vapaa, jos eliminaatiossa sarakkeella i ei
ole pivot-alkiota. Yhtälöryhmä ratkeaa millä tahansa va-
paiden muuttujien x_i arvoilla, ja loput muuttujat voidaan
ratkaista niiden suhteen (jos ratkaisu on olemassa).

Sarakkeet joilla ei ole pivot-alkiota eliminaatiossa.

Vektori avaruudessa $\mathbb{R}^n$
vector in $\mathbb{R}^n$

Vektoriavaruuden kanta
basis of a vector space

**Vektoriavaruuden
virittävä joukko**
spanning set of a vector space

**Vektoriavaruuksien
summa, suora summa**
direct sum

Vektorien yhteenlasku
vector sum

Vinosymmetrinen matriisi
skew-symmetric matrix

Yksittäisratkaisu
particular solution

Vektori avaruudessa $\mathbb{R}^n$ on n reaaliluvun sekvenssi $(x_1, \dots, x_n)$.

Joukko lineaarisesti riippumattomia vektoreita, joiden lineaariyhdistelyinä voidaan esittää mikä tahansa avaruuden vektori. Vektoriavaruudella voi olla monta erilaista kantaa, mutta niiden koko on aina sama (ks. dimensio).

Vektorit $v_1, \dots, v_k$ virittävät vektoriavaruuden V , jos kaikki vektorit V :ssä voidaan muodostaa niiden lineaariyhdistelyinä. Huom! Virittävä joukko ei ole välttämättä kanta; vektorit saavat olla lineaarisesti riippuvia. Siispä avaruuden V dimensiota ei voida laskea virittävän joukon koosta.

Vektoriavaruus $V + W$, joka koostuu vektoreista $v + w$. Jos V :n ja W :n ainoa yhteinen alkio on 0, niin $V + W$:tä kutsutaan myös suoraksi summaksi. Tällöin $\dim(V + W) = \dim V + \dim W$.

Kahden vektorin alkiottainen summa. Vektorien on oltava saman kokoiset. $(v + w)_i = v_i + w_i$.

Matriisi on vinosymmetrinen jos $A^T = -A$. Matriisilla A ei tällöin ole reaalisia ominaisarvoja, vaan kaikki ovat imaginäärisiä. Ominaisvektorit ovat ortogonaalisia, ja e^{Kt} on ortogonaalinen matriisi.

Mikä tahansa yksittäinen ratkaisu yhtälöryhmälle $Ax = b$. Jos yhtälöryhmällä ei ole vapaita muuttujia, yksittäisratkaisu on myös ainoa ratkaisu.