

# SOLMU

MATEMATIIKKALEHTI 3/2023

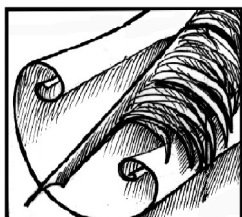
---

[matematiikkalehtisolmu.fi](http://matematiikkalehtisolmu.fi)



## Sisällys

Pääkirjoitus: Mitä matematiikan kirjoissa on? (Anne-Maria Ernvall-Hytönen).....	3
Neuroverkot ja koneoppiminen (Joonas von Lerber) .....	5
Matematiikkadiplomit vuonna 2023 (Marjatta Näätänen) .....	14
Kiehtovaa matematiikkaa, Cauchyn funktionaaliyhtälö (Lasse Pantsar) .....	16
Tarinoita polynomeista (osa 4) (Jukka Tuomela) .....	20
Rekonstruktiokonjektuuri verkkoteoriassa (Milo Orlich) .....	29
Solmun tehtäviä .....	33



## Mitä matematiikan kirjoissa on?

### *Pääkirjoitus*

PISA-tulosten julkistamisen jälkeen on kaikilla ollut mielipiteitä siitä, miten matematiikkaa pitäisi opettaa, mutta jopa siitä, miten tärkeää matematiikka oikeastaan on. Tässä lehdessä en koe tarpeelliseksi argumentoida sen puolesta, että matematiikka todellakin on tärkeää. Itse en myöskään usko, että on yhtä oikeaa totuutta tai helppoja ratkaisuja siihen, mitä ja miten pitäisi opettaa. Tästä huolimatta haluan kirjoittaa muuttaman sanan alakoulun matematiikasta.

Olen kuullut hyvin ristiriitaisia kommentteja siitä, mitä alakoulussa tehdään. Toinen äärlaita on väite siitä, että tehtävät ovat vain tylsiä rutiinitehtäviä, jotka eivät kehitä ajattelua tai joustavuutta, ja joiden kanssa on todella hankala saada ketään kiinnostumaan matematiikasta. Toinen äärlaita taas on väite siitä, että tehtävät ovat niin avoimia, että edes aikuinen ei tajua, mitä pitäisi tehdä.

Olen viime vuosina tuijotellut lasteni matikan kirjoja. En kommentoi sitä, mitä luokassa tapahtuu, vaan ainoastaan kirjojen sisältöjä, sillä vaikka tietäisin mitä lasteni luokissa tapahtuu, ei se tieto yleisty muualle. Kirjojen sisällöt sen sijaan ovat samat kaikissa kouluissa, joissa näitä kirjoja käytetään, jolloin tarkastelen mielelläni niitä. Kirjat edustavat kahta eri kirjasarjaa. Kirjoista löytyy sekä niitä tylsiä rutiinitehtäviä, että niitä tehtäviä, joissa joutuu miettimään, mitä pitäisi tehdä. Minusta tämä ei ole ongelma, sillä tasapaino on vaikuttanut fiksulta. Lisäksi näkisin, että kumpaakin tehtävätyyppiä tarvitaan. On tietysti makuasia, miten paljon kummankinlaisia tehtäviä pitäisi olla. Päätelytehtäviä ja ongelmatehtäviä on hankala ratkaista, jos

peruslaskutoimitukset eivät suju ja laskurutiinia ei ole lainkaan. Toisaalta pelkillä rutiinitehtävillä tylsistyy, eikä myöskään opi soveltamaan taitoja.

Muutama vuosi sitten kuuntelin esitelmää, jossa ruodittiin sitä, miten tietynlaiset arviointitehtävät puuttuvat täysin: tehtävät on tosi usein rakennettu niin, että vain yksi vastaus on oikein. Pointti tässä oli siis se, että kirjasta voisi löytyä esimerkiksi tehtävä, joka olisi vaikka jotain tällaista: *Omenapiirakkaan tarvitaan 3 isoa omenaa. Yksi iso omena on yhtä iso kuin 2 pientä omenaa. Miten monta pientä omenaa tarvitaan omenapiirakkaan?* Sen sijaan tällaista tehtävää ei löytyisi: *Omenapiirakkaan tarvitaan 3 isoa omenaa. Yksi iso omena on yhtä iso kuin 2 pientä omenaa. Anna haluaa tehdä omenapiirakan ja lisäksi ostaa jonkin verran omenoita kotiin syötäväksi. Anna esimerkki siitä, miten monta pientä omenaa Anna voisi ostaa kaupasta.*

Ymmärrän kritiikin liian vaikeista tai avoimista tehtävistä silloin, jos ihmisellä ei ole minkäänlaista kokemusta päätelytehtävistä tai ongelmatehtävien ratkaisusta. Jos aina on sovellettu matematiikkaa vain tuttuihin ja turvallisiin tilanteisiin, niin jopa alaluokkien tehtävät voivat yllättää. Alaluokkien tehtävissä näkisin itse, että oleellista on uskaltaa kokeilla ja testata sitä, mikä toimii. Jotkut tehtävät ovat myös sellaisia, jotka voisivat olla rutiinitehtäviä jossain muualla, mutta alaluokilla ne ovat ongelmatehtäviä. Tällaisia ovat minusta esimerkiksi alakouluun sopiviksi tehdyt yhtälöryhmät. Näissä minusta jujuna on se, että pitää nähdä se kohta, josta lähtemällä tehtävä ratkeaa paljon helpommin kuin muualta lähtemällä. Jos yhtälöryhmä on vaikkapa



tällainen:

$$\begin{cases} \star + \triangle + \circ = 6 \\ \triangle + \triangle + \circ = 7 \\ \star + \star + \star = 3, \end{cases}$$

niin järkevintä on lähteä liikkeelle viimeisestä yhtälöstä, eikä jostain muualta. Minusta on hyödyllistä oppia tarkastelemaan tehtäviä niin, että etsii helppoja aloituskohtia, eikä vain etene ylhäältä alas, vaikka mitä tulisi vastaan. Tämä luultavasti kehittää vähitellen koululaisten joustavuutta.

Osa tehtävistä taas on sellaisia, joissa tällaista helppoa aloituskohtaa ei ole. Tällöin testaaminen, mahdollisesti jotkut vaihtoehdot poissulkien, on hyvä strategia. Testaamiseen oma suhtautumiseni on hiukan ristiriitainen: toisaalta se on joskus erinomainen työkalu. On kuitenkin huolestuttavaa, jos esimerkiksi lukiolainen ei osaa ratkaista yksinkertaista yhtälöä muuten kuin testaamalla. Oleellista on nähdä, milloin testaaminen on hyvä työkalu tai hahmottaa se, että usein se on hyvä ensimmäinen askel. Jos tehtävä vaikuttaa hankalalta, niin on fiksua testata, mitä tapahtuu pienillä luvuilla. Samaan aikaan täytyy kuitenkin ymmärtää, että testaaminen vailla systemaattisuutta ja logiikkaa harvoin antaa varmuuden siitä, että kaikki ratkaisut on löydetty. Tätä pohdintaa soisin olevan paljon enemmänkin: jos löydän yhtälön toteuttavat luvut, niin voinko luot-

taa, että olen löytänyt kaikki yhtälön toteuttavat luvut? Onko tällä väliä tehtävänannon kannalta?

Äsken katsellessani toisen luokan kirjaa, löysin sieltä myös johdattelua kombinaatioihin. Eri asioita piti yhdistää eri tavoin ja listata kaikki vaihtoehdot. Pitää myöntää, että itseäni luultavasti olisi koululaisena jossain vaiheessa alkanut puuduttaa vaihtoehtojen listaaminen. Toisaalta tämä on hyvin konkreettinen tapa johdatella kombinaatioihin ja mahdollisesti huomata itse sääntö kombinaatioiden lukumäärissä.

Ehkä oma johtopäätökseni onkin se, että kirjoissa on yllättävän paljon erilaisia ja monipuolisia tehtäviä. Täysin toinen asia on tietysti se, ehtivätkö kaikki koululaiset tekemään kaikki mielenkiintoisimmatkin tehtävät, samoin se, miten näihin tehtäviin johdatellaan, tai autetaanko koululaisia näkemään säännönmukaisuuksia. Ne säännönmukaisuudet eivät nimittäin välttämättä löydy samoista kirjoista. Peruskoulun opetussuunnitelma 1. ja 2. luokalle sanoo näin: ”Oppilaille tarjotaan mahdollisuuksia löytää yhtäläisyyksiä, eroja ja säännönmukaisuuksia.” Tämä on minusta hyvä tavoite. Menisin kuitenkin itse vielä hiukan kunnianhimoisempaan suuntaan, enkä ainoastaan tarjoaisi mahdollisuuksia löytää näitä, vaan toivoisin, että tarpeen tullen näihin myös johdateltaisiin.

Anne-Maria Ernvall-Hytönen



## Neuroverkot ja koneoppiminen

*Joonas von Lerber*  
Otaniemen lukio

### Johdanto

Tekoäly mullistaa maailmaa monilla eri tavoilla. Tekoälyä voi pyytää kirjoittamaan romaaneja, koodia, runoja tai luomaan maalauksia. Lähes lukematon määrä erilaisia malleja on luotu eri tehtäviä varten. Neuroverkot ovat yksi varhaisimmista tekoälyn malleista, joilla voidaan ratkaista hankalia regressio- ja luokitteluongelmia. Neuroverkot koostuvat pienemmistä tekoneuroneista, jotka ovat yksinkertaisia malleja biologisista hermosoluista.

### Tekoneuronit ja perseptronit (Perceptron)

Tekoneuroni on neuroverkon pienin osanen, sen yksinkertaisin solu. Tekoneuroni on algoritmi, joka pyrkii arvioimaan biologisen neuronin käyttäytymistä. Perseptroni on tietynlainen tekoneuroni, joka pystyy luokittelemaan datan kahteen eri kategoriaan, eli se on binäärinen luokittelija. Nämä kategoriat voivat olla vaikkariistit ja nollat tai 1 ja 0, tai vaikka hymynaama ja surunaama [9].

Perseptronimallia voidaan käyttää luokitteluun, jos hermoverkon data eli sen käsittelemä tieto on lineaarisesti separoituvaa eli kategorioiden raja voidaan kuvaila suorana. Jos ehto ei täyty, perseptroni ei voi oppia dataa kunnollisesti [9].

Tekoneuronin idean alkuperä on biologisessa hermosolussa ja sen käyttäytymisessä, joskin myöhempi tutkimus [5] on osoittanut hermosolujen olevan huomattavasti monimutkaisempia kuin tekoneuronit.

Tekoneuroni-algoritmi ottaa syötteenä yhden data-alkion, laskee sen painotetun summan ja syöttää sen aktivointifunktioon (perseptronin tilanteessa askelfunktioon). Perseptronilla askelfunktion ulostulo on sitten klassifikaatio kahden eri luokan välillä. Kaava 1 on kaksiulotteiselle perseptronille, joka ottaa syötteenä pisteen  $(x_1, x_2)$  ja antaa ulos luokan

$$H(w_1x_1 + w_2x_2 + b) = 0 \text{ tai } 1,$$

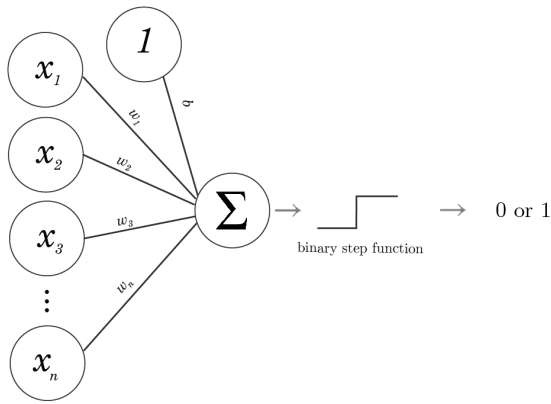
$$H(x) = \begin{cases} 1, & \text{jos } x > 0, \\ 0, & \text{jos } x \leq 0. \end{cases} \quad (1)$$

Tämä lauseke voidaan yleistää  $n$ -ulotteiseen perseptroniin

$$H(b + \sum_{i=1}^n w_i x_i). \quad (2)$$

Yleisen tekoneuronin kaava saadaan muuntamalla kaavan 2 askelfunktio  $H$  yleiseksi aktivointifunktioksi  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ ,

$$\varphi \left( b + \sum_{i=1}^n w_i x_i \right). \quad (3)$$



Kuva 1: Perseptronin rakenne, lähde: [1].

Huomaamme, että voimme esittää pisteen ja perseptronin parametrit vektoreina, ja voimme tehdä siitä laskennallisesti helpomman hyödyntämällä vektorien pistetuloa

$$b + \sum_{i=1}^n w_i x_i = \begin{bmatrix} b \\ w_1 \\ \vdots \\ w_n \end{bmatrix} \cdot \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Täten saamme  $n$ -ulotteisen perseptronin yleisen kaavan

$$H(\mathbf{w} \cdot \mathbf{x} + b), \quad \mathbf{x}, \mathbf{w} \in \mathbb{R}^n. \quad (4)$$

Voimme siis järjestää perseptronin parametrit vektoriin, jonka avulla saamme vastauksen käyttämällä pistetuloa.

Geometrisesti perseptronin voidaan ajatella piirtävän rajaviivan. Rajaviivan alapuoliset pisteet ovat ensimmäistä luokkaa ja yläpuoliset pisteet toista luokkaa. Huomaamme, että perseptroni on lineaarinen avaamalla sen kaksiulotteista kaavaa

$$w_1 x_1 + w_2 x_2 + b = 0.$$

Kuten huomaamme, yhtälö on sama kuin kaksiulotteisella viivalla suhteessa muuttujiin  $x_1, x_2$ . Sama geometrinen intuitio pätee myös korkeampiin ulottuvuuksiin. Yksi  $n$ -ulotteisen perseptronin raja on  $n$ -ulotteinen hypertaso.

Perseptronin toimintaa voi ymmärtää tarkastelemalla muutamaa esimerkkiä.

Kuvassa 2 näemme datajoukon, jolla on kaksi eri kategoriala, valkoiset ja mustat. Huomaamme, että datajoukko on lineaarisesti separoituva, koska voimme piirtää viivaimella suoran erottamaan luokat toisistaan.

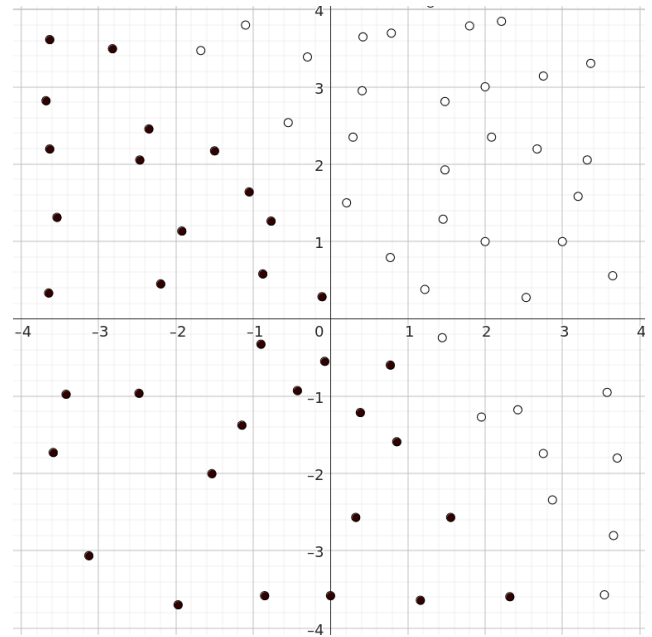
Mutta hipsuliveisulat sentään! Tämähän on juuri perseptronin käyttötarkoitus. Jos säädämme perseptronin, jolla kyseiset parametrit ovat  $w_1 = 4, w_2 = 3, b = -2$ , voimme erottaa pisteet toisistaan. Asettamalla mustien

pisteiden ulostuloksi 0 ja valkoisten pisteiden ulostuloksi 1 voimme sijoittaa kaavaan 1 pisteitä. Esimerkiksi sijoittamalla pisteen  $(2, 1)$  saamme ulostuloksi:

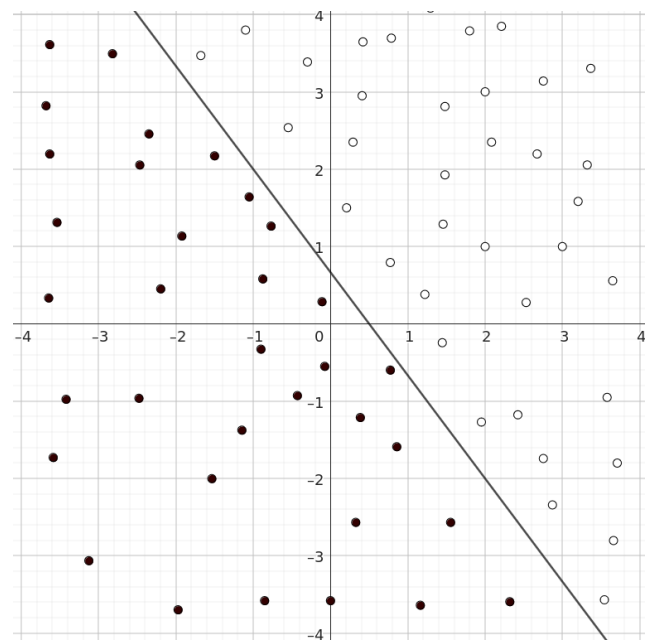
$$H(4 \cdot 2 + 3 \cdot 1 - 2) = H(8 + 3 - 2) = H(9) = 1.$$

Tämän voimme tarkistaa katsomalla kuvaa, ja huomaamme, että perseptroni osasi kertoa oikean luokan. Kokeillaanpa seuraavaksi joukkomme ulkopuolista pistettä  $(-2, 3)$ :

$$H(4 \cdot (-2) + 3 \cdot 3 - 2) = H(-8 + 9 - 2) = H(-1) = 0.$$



Kuva 2: Eräs lineaarisesti separoituva datajoukko  $A$  koordinaatistossa.



Kuva 3: Datajoukon  $A$  rajaviiva piirrettyinä.

Tästä näemme perseptronin kyvyn luokitella eri pisteitä eri luokkiin. Yhdistämällä useita luokittelijoita voimme piirtää useita eri viivoja ja muodostaa mielivaltaisen määrän luokkia.

## Perseptronin opettaminen

Seuraavaksi voimme kysyä, miten saamme sopivat luokittelun parametrit perseptronille?

Käytämme tähän perseptronin opetuskaavaa [9]

$$\begin{aligned} y_j &= H(\mathbf{w}(t) \cdot \mathbf{x}_j), \\ \mathbf{w}(t+1) &= \mathbf{w}(t) + r \cdot (d_j - y_j) \cdot \mathbf{x}_j, \end{aligned} \quad (5)$$

missä

- $\mathbf{w}(t)$  on painovektori ajankohdassa  $t$ ,
- $\mathbf{x}_j$  on datajoukon alkio  $j$ ,
- $d_j$  on odotettu ulostulo datajoukon alkion  $j$  osalta,
- $y_j$  on laskettu ulostulo datajoukon alkion  $j$  osalta,
- $r$  on askelkoko.

Kaavassa laskemme ensimmäiseksi, mikä on kyseisen pisteen luokka, ja jos se on väärin, parametreja muutetaan askelkoon suuruisesti. Askelkoko on meidän itse määrittelemämme hyperparametri, joka on yleensä välillä  $[0, 1]$ . Askelkoko voi olla myös suurempi kuin 1.

Huomaamme vakiotermin  $b$  kadonneen, mutta todellisuudessa se on kiinnitettyä parametrivektorin indeksiin 0. Olen samalla kiinnittänyt datapisteen indeksiin 0 numeron 1, jotta se ei vaikuta vakiotermiin, sillä

$$\begin{bmatrix} b \\ w_1 \\ \vdots \\ w_n \end{bmatrix} \cdot \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + b.$$

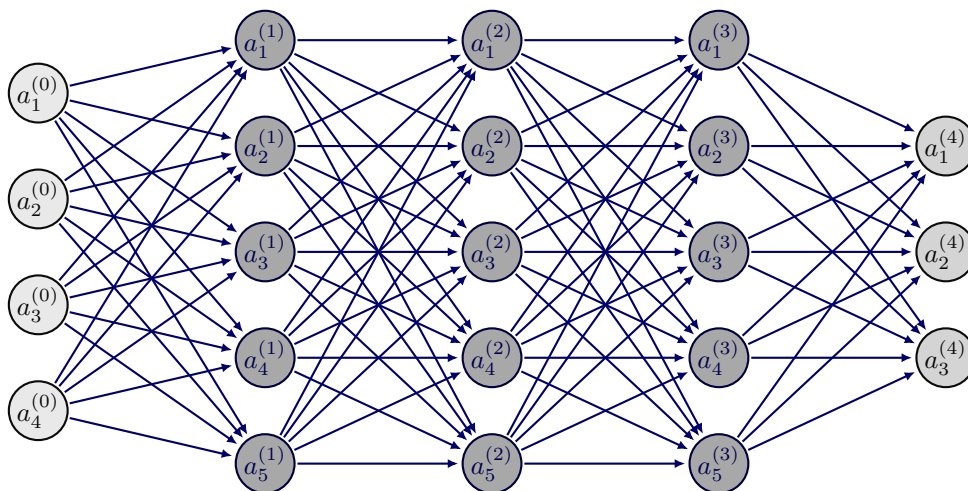
Kaavan 5 voi rakentaa vastavirta-algoritmin avulla, jota käsittelem hieman luvussa Neuroverkkojen opettaminen.

## Neuroverkot

Kuten mainitsin artikkelin alkupuolella, tekoneuronien voidaan ajatella olevan neuroverkkojen rakennuspalikoita. Neuroverkoissa käytetään aktivointifunktioina ”pehmeitä” funktioita, kuten Sigmoid-perheen- tai lineaarisen suoristimen mukaisia funktioita. Yksinkertaisin variaatio tällaisesta funktiosta on nimeltään ReLU (Rectified Linear Unit), joka määritellään  $\text{ReLU}(x) = \max\{0, x\}$  [16].

Neuroverkot rakentuvat siis monista tekoneuroneista, jotka voivat olla järjestettyinä riveihin tai olla järjestelemättömässä asetelmassa. Eteenpäin syötettävissä neuroverkoissa neuronit ovat järjestyneet riveihin, joissa yhteydet kulkevat rivien välissä, mutta eivät rivien sisällä. Jokainen neuroni edelliseltä riviltä on yhdistetty jokaiseen neuroniin seuraavalla rivillä. Kuvassa 4 havainnollistetaan yksinkertaista eteenpäin kytkettyä neuroverkkoa. On myös olemassa neuroverkkomalleja, joissa neuronit eivät ole riveissä, vaan ne voivat olla erilaisissa järjestyksissä. Esimerkki tällaisesta on Hopfieldin verkko [10].

Neuroverkon ensimmäistä kerrosta kutsutaan syötteeksi, seuraavaksi on yksi tai enemmän piilokerroksia, ja viimeinen kerros on nimeltään ulostulo. Kuvassa 4 syöttö on väriltään vaaleampi, piilokerrokset tummempia ja ulostulo taas vaaleampi.



Kuva 4: Yksinkertainen eteenpäin syöttävä neuroverkko.

Neuroverkot ovat erityisen hyviä regressio- ja luokitteluongelmissa, mutta niiden teoreettinen hyöty tulee niiden kyvystä toimia yleisenä funktion approksimoijana. Kaikki laskettavissa olevat funktiot voidaan approksimoida neuroverkolla [11]. Tämä kattaa esimerkiksi oikean biologisen neuronin toiminnan arvioinnin, kuvien luokittelun, talouden arvioinnin ja muut vastaavat sovellukset. Tämä mahdollistaa neuroverkkojen käytön moneen eri sovellukseen [6]. Neuroverkon ominaisuudesta approksimoida jokaista funktiota  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  kerrotaan tarkemmin esitelmässä [12].

Eteenpäin syöttäviä neuroverkkoja käytetään yleisesti. Monet modernit tekoälyt, kuten ChatGPT, perustuvat GPT-4-malliin, jossa eteenpäin syöttävä neuroverkko on yksi osa monien muiden joukossa [17]. Eteenpäin syöttäviä neuroverkkoja käytetään monissa muissa tekoälymalleissa, ja ne ovat yksi yksinkertaisimmista ja parhaiten tutkituista malleista.

Puhun paljon arvoista, ulostuloista, painoista ja vakio-termeistä. Arvo tarkoittaa tekoneuronin arvoa ennen sen syöttämistä aktivointifunktioon, kuten sigmoidiin, ReLUun tai askelfunktioon. Sitä merkitään tässä symbolilla  $x$ . Ulostulo on aktivointifunktion läpi syötetty arvo ja sen symboli on  $o$ . Paino on kahden tekoneuronin välisen yhteyden eli synapsin kerroin, joka kertoo yhteyden voimakkuudesta. Vakio-termi, englanniksi bias, on tekoneuronin ominainen parametri, joka kertoo, kuinka aktiivinen tekoneuroni on. Painon symboli on  $w$  ja vakio-termin  $b$ . Painoa edellisen kerroksen ulostulon  $i$  ja seuraavan kerroksen arvon  $j$  välillä merkitään  $w_{i,j}$ .

Neuroverkko koostuu kerroksista. Jokainen kerros sisältää rivin tekoneuroneita, joista jokainen liittyy jokaiseen seuraavan kerroksen tekoneuroniin kuvan 4 mukaisesti. Jokaisella neuronilla on omat painot ja vakio-termit. Seuraavan kerroksen arvot lasketaan käyttäen tekoneuronin kaavaa 3. Yläindeksit viittaavat kerroksen indeksiin eli monesko kerros neuroverkossa on kyseessä. Alaindeksit taas viittaavat tekoneuronin indeksiin, eli monesko neuroni kyseisessä kerroksessa on kyseessä:

$$\begin{aligned} b_j^L + \sum_{i=1}^n w_{i,j}^L o_i^L &= x_j^{L+1}, \\ \varphi(x_j^{L+1}) &= o_j^{L+1}, \end{aligned} \quad (6)$$

missä

- $j, i$  on neuronin indeksi,
- $L$  on kerroksen indeksi,
- $x$  on neuronin arvo,
- $o$  on neuronin ulostulo,
- $w$  on kyseisen synapsin paino,
- $b$  on kyseisen neuronin vakio-termi,
- $\varphi(x)$  on aktivointifunktio.

Siirrymme siis kerroksesta toiseen tietäen edellisen kerroksen ulostulon, painot ja vakio-termit. Algoritmi jatkuu, kunnes olemme saavuttaneet viimeisen kerroksen eli neuroverkon ulostulon.

Kaava 6 on vektorimuodossaan

$$\begin{aligned} \begin{bmatrix} \sum_{i=1}^n w_{i,1}^L o_i^L \\ \vdots \\ \sum_{i=1}^n w_{i,m}^L o_i^L \end{bmatrix} + \begin{bmatrix} b_1^L \\ \vdots \\ b_m^L \end{bmatrix} &= \begin{bmatrix} x_1^{L+1} \\ \vdots \\ x_m^{L+1} \end{bmatrix}, \\ \varphi \begin{pmatrix} x_1^{L+1} \\ \vdots \\ x_m^{L+1} \end{pmatrix} &= \begin{bmatrix} o_1^{L+1} \\ \vdots \\ o_m^{L+1} \end{bmatrix}. \end{aligned}$$

Yhden neuronin tapauksessa voimme järjestää sen painot vektoriin ja laskea neuronin arvon pistetulona datapisteen kanssa. Monen neuronin systeemissä toimimme samoin, joskin painot järjestetään matriisiin:

$$\begin{bmatrix} w_{1,1}^L & \dots & w_{1,n}^L \\ \vdots & \ddots & \vdots \\ w_{m,1}^L & \dots & w_{m,n}^L \end{bmatrix} \begin{bmatrix} o_1^L \\ \vdots \\ o_n^L \end{bmatrix} + \begin{bmatrix} b_1^L \\ \vdots \\ b_m^L \end{bmatrix} = \begin{bmatrix} x_1^{L+1} \\ \vdots \\ x_m^{L+1} \end{bmatrix}.$$

Huomaamme tässä esitetyn kaavan olevan saman kuin aikaisemmin esitetty summaa käyttävä kaava ja voimme kirjoittaa sen ytimekkäämmin

$$\mathbf{W}^L \mathbf{o}^L + \mathbf{b}^L = \mathbf{x}^{L+1}, \quad (7)$$

$$\varphi(\mathbf{x}^{L+1}) = \mathbf{o}^{L+1}, \quad (8)$$

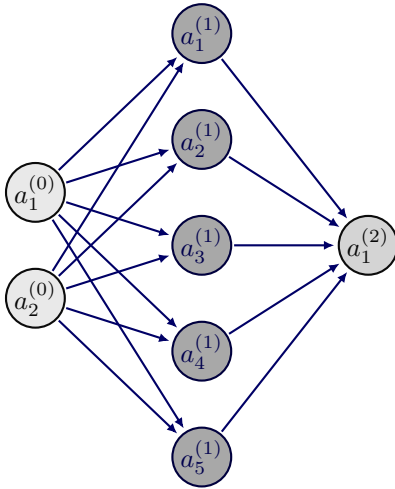
$$\mathbf{W} \in \mathbb{R}^{m \times n}, \mathbf{o} \in \mathbb{R}^n, \mathbf{b}, \mathbf{x} \in \mathbb{R}^m.$$

Kuten huomaamme, mystiset ja ihmeelliset tekoälyt ovat oikeasti vain kasa matriiseja ja vektoreita ryydittynä epälineaarisilla aktivointifunktioilla.

Tiivistykseenä: Eteenpäin syöttävä neuroverkko muodostuu riveistä tekoneuroneita, joista jokainen edellisen kerroksen neuroneista on kytköksissä kaikkiin seuraavan kerroksen neuroneihin. Jokaisella yhteydellä on paino, jotka voidaan esittää painomatriisina. Tekoneuroneilla on myös vakio-termi. Voimme syöttää neuroverkolle dataa, ja neuroverkko syöttää dataa eteenpäin kerrosten välillä käyttäen annettuja kaavoja antaen lopulta ulos haluamme vastauksen.

Esimerkin avulla voimme saada paremman ymmärryksen eteenpäin syöttävien neuroverkkojen toiminnasta. Neuroverkkomme voisi olla vaikka tämän muotoinen:





Neuroverkkomme koostuu siis sisääntulosta, yhdestä piilotetusta kerroksesta ja ulostulosta. Kerrosten välillä on painomatriisi ja vakiotermivektori. Piilotetun kerroksen ja ulostulon aktivointifunktiona toimii yksi sigmoidifunktioista nimeltään logistinen funktio, joka määritellään

$$\sigma(x) = \frac{1}{1 + e^{-x}}.$$

Sen muoto koordinaatistossa muistuttaa S-kirjainta. Se rajoittaa neuroverkkomme ulostulot välille ]0, 1[.

Syötämme sisään vektorin  $\begin{bmatrix} 1.2 \\ 0.4 \end{bmatrix}$ . Ensimmäisen kerroksen arvot:

$$\mathbf{o}^1 = \begin{bmatrix} 1.2 \\ 0.4 \end{bmatrix}, \mathbf{W}^1 = \begin{bmatrix} -1.2 & -1.1 \\ -0.6 & 1.0 \\ 0.8 & -0.5 \\ 0.3 & 1.1 \\ 0.4 & 0.5 \end{bmatrix}, \mathbf{b}^1 = \begin{bmatrix} 1.4 \\ 0.1 \\ -0.5 \\ -0.3 \\ -0.5 \end{bmatrix}.$$

Piilotetun kerroksen arvot  $\mathbf{x}^2$  lasketaan edellisen kerroksen painomatriisin matriisitulosta ulostulojen kanssa käyttäen kaavaa 7:

$$\begin{aligned} \mathbf{x}^2 &= \begin{bmatrix} -1.2 & -1.1 \\ -0.6 & 1.0 \\ 0.8 & -0.5 \\ 0.3 & 1.1 \\ 0.4 & 0.5 \end{bmatrix} \begin{bmatrix} 1.2 \\ 0.4 \end{bmatrix} + \begin{bmatrix} 1.4 \\ 0.1 \\ -0.5 \\ -0.3 \\ -0.5 \end{bmatrix} \\ &= \begin{bmatrix} -1.88 \\ -0.32 \\ 0.76 \\ 0.80 \\ 0.68 \end{bmatrix} + \begin{bmatrix} 1.4 \\ 0.1 \\ -0.5 \\ -0.3 \\ -0.5 \end{bmatrix} = \begin{bmatrix} -0.48 \\ -0.22 \\ 0.26 \\ 0.5 \\ 1.18 \end{bmatrix}. \end{aligned}$$

Nämä ovat siis seuraavan kerroksen arvot. Käyttämällä kaavaa 8 ja sijoittamalla funktion  $\varphi$  paikalle  $\sigma$ :

$$\mathbf{o}^2 = \sigma \begin{pmatrix} -0.48 \\ -0.22 \\ 0.26 \\ 0.5 \\ 1.18 \end{pmatrix} \approx \begin{bmatrix} 0.38 \\ 0.45 \\ 0.56 \\ 0.62 \\ 0.76 \end{bmatrix}.$$

Voimme tämän nyt syöttää seuraavaan kerrokseen käyttämällä kaavaa 7:

$$\mathbf{W}^2 = [-0.3 \quad 0.8 \quad -1.4 \quad -1.4 \quad 1.5], \mathbf{b}^2 = [-0.3],$$

jolloin

$$\begin{aligned} \mathbf{x}^3 &= [-0.3 \quad 0.8 \quad -1.4 \quad -1.4 \quad 1.5] \begin{bmatrix} 0.38 \\ 0.45 \\ 0.56 \\ 0.62 \\ 0.76 \end{bmatrix} + [-0.3] \\ &= -0.04 - 0.3 = -0.34. \end{aligned}$$

Pyöritämme tämän vielä sigmoidin kautta ja saamme neuroverkon ulostuloksi

$$\mathbf{o}^3 = \sigma(-0.34) \approx 0.41.$$


Tämä ulostulo ei kerro mitään järkevää ja on lähes satunnainen, sillä neuroverkkoa ei oltu opetettu suorittamaan mitään tehtävää, vaan painomatriisit ja vakioarvot luotiin satunnaislukugeneraattorilla.

## Neuroverkot luokittelijoina

Eteenpäin syötäviä neuroverkkoja käytetään luokitteluun perustuvissa ongelmissa kuten objektien tunnistaminen kuvista tai videosta. Yksi tunnetuimmista koneoppimisen ongelmista oli saada kone tunnistamaan kokoa  $28 \times 28$  olevista kuvista käsin piirrettyjä numeroita käyttäen MNIST-datajoukkoa. Tätä ongelmaa ja datajoukkoa on jopa kutsuttu koneoppimisen ”Hello World!” -ongelmaksi. Nykytekniikoita käyttäen eteenpäin syötävä neuroverkko pystyy luokitteluun 10 000 uudesta kuvasta noin 99,65 %, ja konvoluutionaalilla neuroverkolla on saatu jopa 99,77 % oikein [18]. Nykyään monia uusia malleja ja tekniikoita kokeillaan ja mitataan tällä datajoukolla.

Neuroverkon ulostulo luokitteluongelmissa on yleensä One-Hot -muodossa, jossa luokat on järjestelty vektoriin. Neuroverkolle annetaan jokin alkio datajoukosta ja neuroverkon tulee antaa ulostulona vektoriin nollia, mutta toivotun luokan kohdalla tulisi olla arvo 1. Kuvassa 5 on esimerkki One-Hot -muodosta väreillä.

Kuvien tunnistamisessa käytetään usein paremmin soveltuvia konvoluutionaalisia neuroverkkoja, joissa yhdistelmä konvoluutio-operaatioita ja eteenpäin syötäviä neuroverkkoja kykenevät tunnistamaan ympyrämäisiä muotoja, suoria viivoja ja monia muita kuvien muodollisia ominaisuuksia [7].

id	color		id	color_red	color_blue	color_green
1	red		1	1	0	0
2	blue		2	0	1	0
3	green		3	0	0	1
4	blue		4	0	1	0

Kuva 5: One-Hot -koodaus väreillä, lähde: [2].

## Neuroverkot regressio-ongelmissa

Eteenpäin syöttävät neuroverkot soveltuvat regressio-ongelmiin mainiosti ja niitä sovelletaan monilla eri aloilla arvioimaan tuotteiden hintoja, taloutta tai vaikka biologisia neuroneita. Neuroverkkojen etu perinteiseen lineaariseen regressioon on neuroverkon kyky approksimoida epälineaarista funktiota saaden tarkemman arvion suureesta [6].

## Neuroverkkojen opettaminen

Koneoppiminen on laaja käsite, joka käsittelee eri malleja ja tapoja, joilla tietokoneet voivat ”oppia”. Neuroverkot oppivat tietyn datajoukon käyttämällä gradienttimenetelmää, joka on optimointimenetelmä käyttäen gradienttioperaattorin ominaisuutta minimoimaan häviöfunktioita.

## Gradienttimenetelmä (Stochastic Gradient Decent)

Gradienttimenetelmä [4] on optimointialgoritmi, joka voi löytää funktion minimin käyttäen funktion gradienttia. Gradientti on pinnalle piirretyn tangenttita-son jyrkin muutossuunta. Kuvittele olevasi lasketu-  
massa vuorta alas ja tehtäväsi on päästä laakson matalimpaan kohtaan. Tehtävään on monia ratkaisuja, mutta yksi yksinkertaisimmista tavoista on ensin löytää jyrkin suunta alaspäin ja ottaa askel tähän suuntaan. Tätä jatketaan, kunnes jyrkin suunta alaspäin on vaakatasossa. Tämä algoritmi varmistaa, että löydät aina jonkinlaisen laakson. Algoritmi voidaan esittää matemaattisesti näin:

$$\begin{aligned} \text{toista } \mathbf{x}_{n+1} &= \mathbf{x}_n - \alpha \nabla f(\mathbf{x}_n), \\ \text{kunnes } \|\nabla f(\mathbf{x}_n) - \vec{0}\| &< \delta, \end{aligned} \quad (9)$$

missä  $\delta > 0$ . Algoritmi pyörii  $N$  kertaa. Algoritmin lopussa  $\mathbf{x}_N$  on funktion lokaali minimi tarkkuudella  $\delta$  eli

$$\exists \epsilon > 0 \text{ siten, että } \|\mathbf{x} - \mathbf{x}_N\| < \epsilon \implies f(\mathbf{x}_N) \leq f(\mathbf{x}).$$

Edellä  $\alpha$  on hyperparametri, joka kuvaa askeleen suuruutta. Se kuvailee, kuinka suuren harppauksen otamme. Jos  $\alpha$  on liian suuri, saatat hypätä koko minimin yli.

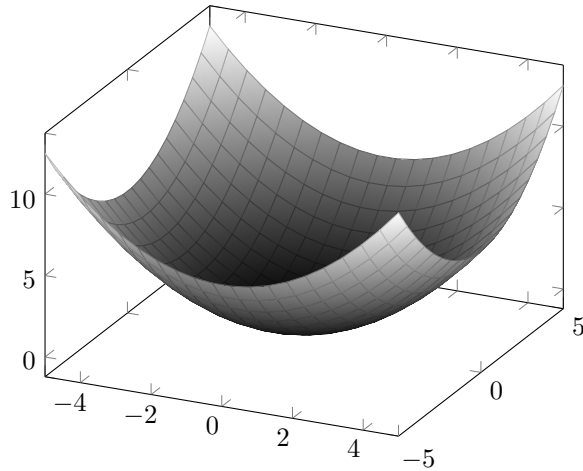
Konvekseilla ja pseudokonvekseilla funktioilla algoritmi löytää lähes aina globaalin minimin [13], mutta ei-konvekseilla funktioilla algoritmi yleensä tökkähtää lokaaliin minimiin. Lokaali minimi ei ole aina hyvä minimi, sillä vuoren harjanteessakin on pieniä kuoppia. Huonojen lokaalien minimien lisäksi moniulotteisissa ei-konvekseissa optimointiongelmissa ovat myös satulapisteet merkittävä ongelma [8].

Neuroverkolle voidaan määritellä häviöfunktio, joka määrittelee suureen kuvaamaan neuroverkon vääryyttä. Neuroverkon treenaamisen alussa se ei vielä tunne datajoukkoa, joten sen antamat ulostulot ovat lähes sattumia. Neuroverkon häviö on siis korkea, koska sen ulostulot ovat roimasti väärässä haluamastamme ulostulosta. Häviöfunktioita on monia ja yksi käytetyimmistä on MSE (Mean Squared Error):

$$\begin{aligned} MSE(\mathbf{o}, \mathbf{t}) &= \frac{1}{N} \sum_{i=1}^N (o_i - t_i)^2 \\ &= \frac{1}{N} \mathbf{e}^\top \mathbf{e}, \text{ missä } \mathbf{e} = \mathbf{o} - \mathbf{t}. \end{aligned} \quad (10)$$

MSE on siis saadun ulostulon ja haluamamme ulostulon erotuksien neliöiden keskiarvo. Jos neuroverkkomme ulostulo olisi  $\begin{bmatrix} 2.3 \\ 0.5 \end{bmatrix}$ , mutta haluaisimme ulostulon  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ , olisi häviö

$$\frac{(2.3 - 1)^2 + (0.5 - 0)^2}{2} = \frac{1.69 + 0.25}{2} = 0.97.$$



Kuva 6: Häviö kasvaa mitä kauempana odotetusta vektorista ulostulomme on,  $z = \text{MSE} \left( \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right)$ .

Neuroverkon häviö datajoukolle  $A$  määritellään laske-  
malla keskiarvo jokaisen datajoukon alkion häviöstä.  
Valitaan symboli  $E(\theta)_A$  kuvaamaan tätä häviötä,  $\theta$  ku-  
vaa neuroverkon parametreja.

Häviöpinta määritellään laskemalla häviö kaikille sa-  
man muotoisille neuroverkoille ja merkitsemällä jokai-  
nen häviö koordinaatistoon, jossa muuttujina toimivat  
neuroverkon parametrit. Saman muotoiset neuroverkot  
ovat neuroverkkoja, joilla on sama määrä kerroksia ja  
jokaisessa kerroksessa on sama määrä neuroneita. Neu-  
ronien painot ja vakioarvot voivat vaihdella.

Neuroverkon häviöpinnan esittäminen ihmiselle ym-  
märrettävässä muodossa on vaikeaa, sillä neuroverkon  
parametrien määrää voi olla miljoonia. Kuva 7b on vii-  
pale neuroverkon häviöpinnasta [14].

Kuvan 7a musta piste on neuroverkkomme tässä koor-  
dinaatistossa. Musta nuoli voisi kuvata treenattavan  
neuroverkon vaellusta vuoriston pohjalle. Kun häviö-  
funktio on mahdollisimman alhainen, on neuroverkko

tarkimmissa parametreissa. Tällöin musta piste on al-  
haisimmassa kohdassa ja neuroverkko on oppinut teh-  
tävän. Käytännössä emme laske häviöpintaa, sillä gra-  
dienttimenetelmän ei tarvitse nähdä häviöpintaa toi-  
miakseen.

### Vastavirta-algoritmi (Backpropagation)

Neuroverkon gradientin tietylle pisteelle voi laskea eri  
tavoilla. Numeerisesti laskettu gradientti voi olla jos-  
kus hyödyllinen, mutta usein käytetään vastavirta-  
algoritmia, jossa lasketaan häviöfunktion osittaisderi-  
vaatat jokaisen parametrin kanssa käyttäen apuna de-  
riivaatan ketjusääntöä ja matriisilaskennan työkaluja  
[4].

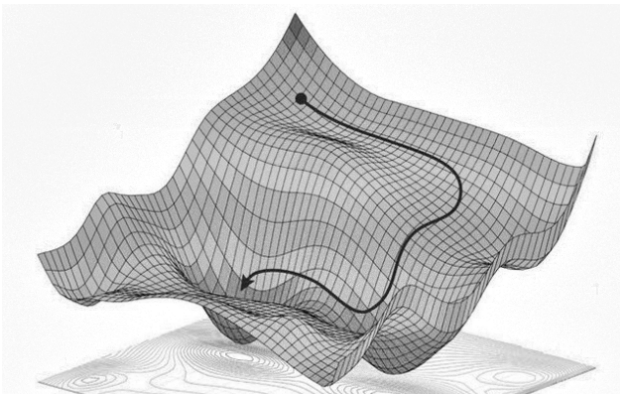
Jos neuroverkkosten opettaminen on herättänyt kiin-  
nostuksesi, voit tutustua tarkemmin vastavirta-algorit-  
miin. YouTube-kanava 3Blue1Brown teki intuitiivisen  
ja laskennallisen selityksen vastavirta-algoritmistä sar-  
jassaan Deep learning. Ehdotan näiden videoiden kat-  
somista ymmärtääksesi algoritmin toiminnan.

### Gradienttimenetelmän optimisaatiota

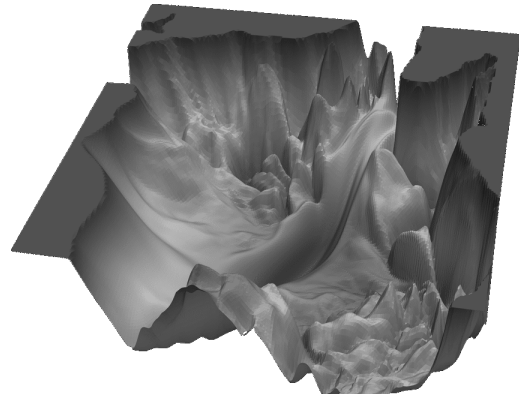
Tutustumalla tarkemmin gradienttimenetelmään ja  
koneoppimiseen saattaa törmätä gradienttimenetel-  
män optimisaatioalgoritmeihin. Optimisaatioalgorit-  
missa muunnellaan gradienttimenetelmää paremman  
minimin saavuttamiseksi nopeammin. Monet näistä al-  
goritmeista hyödyntävät vanhojen gradienttien muis-  
tamista. Yksi intuitiivisimmista ja yksinkertaisimmis-  
ta on gradienttimenetelmä liikemäärällä (Momentum),  
jossa käytetään edellisten gradienttien eksponentaali-  
sesti vähenevää summaa [15]:

$$v_t = \beta v_{t-1} + (1 - \beta) \nabla \theta_t(x),$$

$$\theta_{t+1}(x) = \theta_t(x) - v_t,$$



(a) Möykkyinen häviöpinta, lähde: [3].



(b) Oikean neuroverkon häviöpinta, lähde: [14], kuva 3.

Kuva 7: Esimerkkejä neuroverkkojen häviöpinnoista.

missä  $\beta$  on liikemäärän termi, joka on puoliaukinaisella välillä  $[0, 1]$ . Mitä suurempi beta sitä paremmin algoritmi muistaa vanhat gradientit ja kasvattaa gradientin kokoa.

Algoritmi toimii paremmin kuin perinteinen gradienttimenetelmä, ja se suppenee nopeammin minimiin kuin perinteinen gradienttimenetelmä. Se saavuttaa minimin nopeammin, koska siihen on lisätty nopeuden tyyppinen termi. Tämä termi kasvattaa gradientin suuruutta, jos gradientti on pitkään samaan suuntaan.

Intuitiivisesti Momentum-algoritmin voi ajatella olevan pallo, jolla on kitkaa ja massaa ja joka pyörii häviöpin-taa alas. Kuten pallo, voi se pyöriä minimin ympärillä, kunnes hiljalleen osuu minimin kohdalle.

On myös monia muita tapoja kuten Adam (Adaptive Moment Estimation), jotka lisäävät lisää erilaisia nopeus- ja arviointitermejä, minkä avulla voidaan poistaa askelkoon säätäminen.

## Tiivistelmä

Tekoneuroni on algoritmi, jonka idea syntyi tutkimalla biologisia neuroneita. Perseptroni on eräänlainen tekoneuroni, jonka aktivointifunktio on askelfunktio. Perseptroni on binäärinen luokittelija, joka voi jakaa datajoukon eri kategorioihin. Perseptronit saavuttavat tämän laskemalla painotetun summan datapisteen komponenteista. Tämän jälkeen tulos syötetään binääriiseen askelfunktioon, joka antaa ulos luokan. Voimme organisoida perseptronin painot vektoriin ja nopeuttaa painotetun summan laskemista pistetulon kautta.

Neuroverkko on kasa tekoneuroneita järjestettynä keroksiin. Neuroverkkojen läpi syötetään dataa laskemalla edellisestä kerroksesta seuraava kerros ja suorittamalla tätä, kunnes päästään viimeiseen kerrokseen. Neuroverkkoja voi opettaa tunnistamaan ja klassifikoidaan kuvia, suorittamaan regressio-ongelmia ja tekemään monenlaisia operaatioita, sillä ne ovat pohjimmiltaan universaaleja funktion approksimoijia.

Neuroverkkoja opetetaan gradienttimenetelmällä, jossa hypimme häviötasoa alaspäin gradientin suuntaan eli siis minimoimme häviöfunktion. Häviöfunktio määrittelee, kuinka väärässä neuroverkko oli näillä parametreilla. Esimerkki tällaisesta funktiosta on vaikka MSE (Mean Squared Error). Gradientti lasketaan vastavirta-algoritmillä.

Gradienttimenetelmää voi optimoida lisäämällä nopeus- ja arviointitermejä, joiden avulla neuroverkko suppenee nopeammin minimiin.

## Loppuajatuks

Toivon, että tämän artikkelin avulla sain kiinnostuksen herätettyä koneoppimisen ja erityisesti neuroverkkojen monimutkaiseen maailmaan. Tein itse Rust-ohjelmointikielellä neuroverkkokirjaston, jossa pystyy luomaan neuroverkkoja ja opettamaan niitä käyttämällä yksinkertaisia funktioita.

Lisämateriaaleja (linkkejä) aiheista:

- Perseptronit: Samuli Siltasen sarja Tekoälyn Taikaa
- Neuroverkot: 3Blue1Brownin sarja neuroverkoista
- Neuroverkot: StatQuestin with Josh Starmerin sarja neuroverkoista
- Neuroverkot: Brilliantin kurssi neuroverkoista

Omia neuroverkkoja voi joko testaila omalla kirjastolani tai käyttää Googlen omaa Tensorflow- tai Metan Pytorch-kirjastoa. Nämä kirjastot on optimoitu käytännöllisyyteen ja nopeuteen, ja uskon kirjastojen ominaisuuksien olevan sinulle uusia oppimiskokemuksia.

## Kuka olen?

Hei, olen Joonas. Olen toisen vuosikurssin oppilas Otoniemen lukiosta ja minulla heräsi kiinnostus neuroverkkoihin vuoden 2022 lokakuussa, kun tein kaverini Pyyri Jäppisen kanssa Rust-ohjelmointikurssin projektiksi pienen neuroverkkoesittelyn Rust-ohjelmointikielellä. Jatkoin tämän tekemistä ja projekti hiljalleen kehittyi omaksi harrastukseksi. Olen ollut aina kiinnostunut matematiikasta ja tietokoneista, ja aihe tuntui täydelliseltä yhdistelmältä. Pyyriin jatkamaan tiedonjanoni tyydyttämistä ja oppimaan lisää neuroverkoista ja erilaisista koneoppimisen malleista.

Jos herää kysymyksiä artikkelista ja sen aiheista, niitä voi lähettää sähköpostiin

joonas.vonlerber@gmail.com

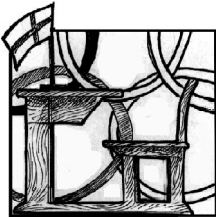
Minut löytää githubista, jossa voit seurata tulevaisuuden projektejani tai nykyistä neuroverkkoprojektiani. Teen tätä omalla vapaa-ajallani ja kirjoittamani kirjasto voi muuttua ajan myötä huomattavasti.

## Viitteet

- [1] URL: [https://miro.medium.com/v2/resize:fit:720/format:webp/1\\*IgF01c6tq4Tz1iwJraw.png](https://miro.medium.com/v2/resize:fit:720/format:webp/1*IgF01c6tq4Tz1iwJraw.png).
- [2] URL: [https://miro.medium.com/v2/resize:fit:720/format:webp/1\\*ggtp4a5YaRx6109KQaY0nw.png](https://miro.medium.com/v2/resize:fit:720/format:webp/1*ggtp4a5YaRx6109KQaY0nw.png).



- [3] URL: <https://www.chegg.com/homework-help/questions-and-answers/10-consider-path-gradient-descent-takes-image-3d-loss-surface-label-parts-path-momentum-w-q85140081>.
- [4] Shun-ichi Amari. “Backpropagation and stochastic gradient descent method”. *Neurocomputing* 5.4-5 (1993), s. 185–196.
- [5] David Beniaguev, Idan Segev ja michael London. “Single Cortical Neurons as Deep Artificial Neural Networks”. *SSRN Electronic Journal* (tammi-kuu 2020). DOI: 10.2139/ssrn.3717773.
- [6] Chris M Bishop. “Neural networks and their applications”. *Review of scientific instruments* 65.6 (1994), s. 1803–1832.
- [7] Dan Cireşan, Ueli Meier ja Juergen Schmidhuber. *Multi-column Deep Neural Networks for Image Classification*. 2012. arXiv: 1202.2745 [cs.CV].
- [8] Yann Dauphin et al. “Identifying and attacking the saddle point problem in high-dimensional non-convex optimization” (2014). arXiv: 1406.2572 [cs.LG].
- [9] Yoav Freund ja Robert Schapire. “Large Margin Classification Using the Perceptron Algorithm”. *Machine Learning* 37 (helmikuu 1999). DOI: 10.1023/A:1007662407062.
- [10] J J Hopfield. “Neural networks and physical systems with emergent collective computational abilities”. en. *Proc. Natl. Acad. Sci. U. S. A.* 79.8 (huhtikuu 1982), s. 2554–2558.
- [11] Kurt Hornik, Maxwell Stinchcombe ja Halbert White. “Multilayer feedforward networks are universal approximators”. *Neural Networks* 2.5 (1989), s. 359–366. ISSN: 0893-6080. DOI: [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8). URL: <https://www.sciencedirect.com/science/article/pii/0893608089900208>.
- [12] Kurt Hornik, Maxwell Stinchcombe ja Halbert White. “Multilayer feedforward networks are universal approximators”. *Neural networks* 2.5 (1989), s. 359–366. URL: [https://www.cs.cmu.edu/~bhiksha/courses/deeplearning/Fall.2016/notes/Sonia\\_Hornik.pdf](https://www.cs.cmu.edu/~bhiksha/courses/deeplearning/Fall.2016/notes/Sonia_Hornik.pdf).
- [13] Yaohua Hu, Jiawen Li ja Carisa Kwok Wai Yu. “Convergence Rates of Subgradient Methods for Quasi-convex Optimization Problems” (2019). arXiv: 1910.10879 [math.OA].
- [14] Hao Li et al. “Visualizing the Loss Landscape of Neural Nets” (2018). arXiv: 1712.09913 [cs.LG].
- [15] Sebastian Ruder. “An overview of gradient descent optimization algorithms” (2017). arXiv: 1609.04747 [cs.LG].
- [16] Tomasz Szandala. “Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks”. *CoRR* abs/2010.09458 (2020). arXiv: 2010.09458. URL: <https://arxiv.org/abs/2010.09458>.
- [17] Ashish Vaswani et al. *Attention Is All You Need*. 2017. arXiv: 1706.03762 [cs.CL].
- [18] Christopher J.C. Burges Yann LeCun Corinna Cortes. *THE MNIST DATABASE of handwritten digits*. Accessed: 2023. URL: <http://yann.lecun.com/exdb/mnist/>.



## Matematiikkadiplomit vuonna 2023

*Marjatta Näätänen*

Uusia vastauspyyntöjä on vuonna 2023 12.12.23 mennessä tullut 381 näistä 106 kunnasta:

Alavus, Asikkala, Espoo, Eurajoki, Forssa, Haapavesi, Hamina, Helsinki, Hollola, Hyvinkää, Hämeenkyrö, Hämeenlinna, Ii, Iisalmi, Ikaalinen, Imatra, Janakkala, Joensuu, Joutsa, Jyväskylä, Jämsä, Järvenpää, Kaarina, Kajaani, Kangasala, Kankaanpää, Karkkila, Karsula, Kauhajoki, Kauniainen, Kaustinen, Keminmaa, Kempele, Kerava, Kihniö, Kirkkonummi, Kokemäki, Kokkola, Kolari, Kontiolahti, Koski tl, Kotka, Kouvolaa, Kuopio, Kurikka, Kuusamo, Kärkölä, Lahti, Laitila, Lapinlahti, Lappeenranta, Laukaa, Lempäälä, Lieto, Lohja, Loimaa, Loppi, Masku, Mikkeli, Muonio, Mustasaari, Muurame, Mäntsälä, Naantali, Nakkila, Nivala, Nokia, Nurmijärvi, Oripää, Oulainen, Oulu, Parainen, Petäjävesi, Pirkkala, Polvijärvi, Pori, Porvoo, Pudasjärvi, Punkalaidun, Raahe, Raasepori, Raisio, Rauma, Ristijärvi, Salo, Sauvo, Savitaipale, Savonlinna, Seinäjoki, Siilinjärvi, Sipoo, Tampere, Turku, Tuusula, Urjala, Utsjoki, Uusikaupunki, Vaala, Vaasa, Valkeakoski, Vantaa, Varkaus, Virrat, Ylivieska, Ylöjärvi, äänekoski.

Laaja-alaisten tehtäväpakettien vastauspyyntöjä tuli 7, näistä 5 kunnasta:

Haapavesi, Helsinki, Lempäälä, Oulu, Rauma.

Solmun Matematiikkadiplomit ovat osoittaneet tarpeellisuutensa, opettajien palaute on hyvin palkitsevaa: ”Huippua, kun löysin matikkaankin diplomitehtävät.” Lukudiplomitehtävillä on jo vuosia kannustettu

kirjastojenkin kautta oppilaiden lukuharrastusta, mutta opettajan palautteen mukaan ”matikka on jossakin pimennossa ollut luonnonvara 😊”. Media ei todellakaan paljoa matematiikka-aihetta suosi, joten tietoa Solmusta levittävät opettajat toisilleen.

Osoitteessa

[matematiikkalehtisolmu.fi/diplomi.html](https://matematiikkalehtisolmu.fi/diplomi.html)

on ohjeet ja kymmenen tulostettavaa diplomia tehtävienneen. Vastaukset koulun sähköpostiin voi opettaja pyytää osoitteesta [juha.ruokolainen \(at\) yahoo.com](mailto:juha.ruokolainen(at)yahoo.com). Julkiseen jakeluun vastauksia ei sijoiteta, eikä niitä ole tarkoitus jakaa koulun ulkopuolelle.

Tehtävien monipuolisuudesta ja siitä, että mukana on myös haastavia tehtäviä, on tullut paljon kiitosta; samalla luokallahan saattaa olla hyvinkin eri tasoisia oppilaita. ”Ope nauttii helposta ylöspäin eriyttämisen mahdollisuudesta.” ”Kiitos ihan valtavan paljon matikkadiplomista!” ”Täältä tulee viidennen luokan oppilaalta positiivista palautetta siitä, että tehtävät ovat kivoja kokonaisuuksia ja ihan erilaisia kuin nää meidän omat matikan tehtävät.” ”Tehtävät kirvoittivat toivottua ryhmätyötä, ääneen pohtimista ja vastausten vertailua. Tämä on kyllä arvokas paketti, joten kiitos siitä.”

Matematiikan tehtäviä tekemällä saa hyvää harjoitusta, paitsi matematiikassa, myös loogisessa ajattelussa, ongelmanratkaisussa ja monenlaisessa kehittävässä puuhailussa. Aivoja pitäisi harjoittaa samoin

kuin lihaksia. Hienomotoriikankin merkitys näyttää peruskoulussa unohtuneen. Erityisesti kasvuiässä aivot muokkautuvat nopeasti, joten silloin on tärkeää harjoittaa monipuolista toimintaa, koska luodaan pohja koko elämää varten.

Diplomit eivät ole tiukasti sidottuja vuosiluokkiin, vaikkakin numerointi kertoo varsinkin alimmilla tasoilla etenemisestä suunnilleen vuosiluokkien mukaisesti. Viimeisessä diplomissa päästään myös hiukan kurkistamaan sellaisiin matematiikan aloihin, joihin ei ehkä koulussa törmätä. Ylimmät diplomit voi suorittaa myös osissa. IX diplomissa tarjotaan joitain lukiollekin sopivia yhteiskunnallisia aiheita ja X:ssä käytetään myös GeoGebraa.

Matematiikan perustiedot ja asenteet syntyvät koulussa. Matematiikkadiplomi antaa mahdollisuuden hyödylliselle ja hauskalle harrastukselle, tarjoaa haasteita ja monipuolista toimintaa jo ensimmäisestä luokasta lähtien. Diplomitehtävät sopivat myös kerhotoimintaan ja kertaukseen. Tämä toiminta ei ole oppilaiden välistä kilpailua, vaan oppilas voi ottaa mittaa itsestään, tarvittaessa kysyä neuvoa ja tehdä yhteistyötä. Harrastus palkitaan ponnistelun jälkeisen onnistumisen ilolla, oppilaat huomaavat oppivansa ja ymmärtävänsä matematiikkaa. Diplomit toimivat matematiikan suppean sanaston takia myös kielikylpynä. Myös suuruusluokkien arviointi on huomattu tarpeelliseksi. Tehtävillä tarjotaan lapsille kokemuksia matematiikan käsitteistä, jotka tarkentuvat myöhemmin noustaessa portaita konkreettisesta abstraktiin. Tehtävät ovat matemaatikkojen itsensä laatimia ja valitsemia, joten matematiikan omaa rakennetta käytetään pohjana ja tuloksena rakentuu kestävä ja mielekäs pohja jatkolle.

## Vuonna 2023 havahduttiin Suomessa PISA-tulosten selvään alamäkeen

Matemaatikoille tämä ei ollut yllätys; hälytystä heräämiseen on yritetty vuosikymmeniä, huonolla menestyksellä. Vallitseva mielipide päättäjien taholta on vakaasti ja horjumatta ollut, että koneet ja pelit ovat nykyäikää, perusmatematiikka voidaan vanhanaikaisena unohtaa ihmisten aivoja kiusaamasta. On aika tehdä ”matematiikkaa”, joka on hauskaa eikä vaadi ponnisteluja, muistia eikä keskittymistä. PISAn tehtävät ovat tä-

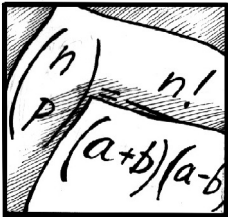
hän suuntaan kehitettyjä, ”matematiikan lukutaitoa”, ts. jokseenkin sanomalehtien uutisten ja taulukoiden ymmärtämiseen tähtäviä – mutta edes tällä tasolla menestymisessä ei Suomen peruskoulun matematiikan yleinen osaaminen pysynyt.

Vuonna 2005 kirjoitti yli 200 yliopistojen, korkeakoulujen ja ammattikorkeakoulujen matematiikan opettajaa ja tutkijaa huolestuneina matematiikan osaamistason heikkenemisestä – siis aikana, jolloin oltiin ns. PISA-osaamisen huipuilla:

*”PISA-tutkimuksessa jää siten täysin avoimeksi, miten hyvin osataan esimerkiksi laskea murtoluvuilla, ratkaista yksinkertaisia yhtälöitä, tehdä varsinaisia geometrisia päättelyitä, laskea kappaleiden tilavuuksia, käsitellä algebran lausekkeita. Algebra on kuitenkin matematiikassa peruskoulun jälkeisten opintojen kannalta keskeisin yksittäinen osa-alue.*

*Peruskoulussa pitäisi oppia matematiikan perusasiat, joiden varaan voidaan myöhemmin rakentaa lisää. Laskimien käyttökään ei muuta tilannetta: vaikka laskin laskisikin murtoluvuilla, myös käsin laskeminen on osattava, koska se on algebrallisten lausekkeiden käsittelyn pohja. Jatko-opiskelu tulee mahdottomaksi ellei perusta ole kunnossa. Yksi syy lisääntyvään huonoon osaamiseen ylioppilaskokeessa ja korkeakouluopintojen alussa onkin ilmeisesti jo peruskoulussa saadun pohjan heikkous. Uusia vaikeampia asioita ei kyetä omaksumaan, koska huomattava energia menee vielä lukiossa peruskoulutason asioiden pohdiskeluun. Kierre jatkuu jatko-opinnoissa: lukion asioita ei hallita ja eteenpäin meno vaikeutuu.*

*PISA-tutkimus tuo hyödyllistä tietoa arkielämässä tarpeellisesta matemaattisesta lukutaidosta ja yksinkertaisten ongelmien ratkaisukyvyistä. Tällainen taito ei vain riitä yhä voimakkaammin matematiikkaa hyödyntävässä maailmassa. Kunnollista matemaattista pohjaa tarvitaan etenkin teknillisillä ja luonnontieteellisillä aloilla, biologia mukaan luettuna. PISA-tutkimus kertoo hyvin vähän tästä pohjasta, joka tulisi luoda jo peruskoulussa. Sen vuoksi olisi ehdottoman tarpeellista, että jatkossa Suomi osallistuisi myös niihin kansainvälisiin arviointeihin, joissa arvioidaan jatko-opintojen kannalta keskeisten matematiikan taitojen hallintaa.”*



## Kiehtovaa matematiikkaa, Cauchyn funktionaaliyhtälö

Lasse Pantsar

### Additiivisuus ja lineaarisuus

**Määritelmä 1:** Funktio  $f: \mathbb{R} \rightarrow \mathbb{R}$  on *additiivinen*, jos se toteuttaa *Cauchyn funktionaaliyhtälön*

$$f(x + y) = f(x) + f(y) \quad (1)$$

kaikilla  $x \in \mathbb{R}$  ja kaikilla  $y \in \mathbb{R}$ .

Helposti voidaan todistaa

**Lause 1:** Jos funktio  $f: \mathbb{R} \rightarrow \mathbb{R}$  on additiivinen, on

$$f(qx) = qf(x) \quad (2)$$

kaikilla  $q \in \mathbb{Q}$  ja kaikilla  $x \in \mathbb{R}$ .

Jokaisella additiivisella funktiolla  $f$  on siis ominaisuus (2), mutta toteutuuko yhtälö  $f(rx) = rf(x)$  kaikille additiivisille funktioille  $f$  ja kaikille reaaliluvuille  $r$  ja  $x$ ? Tämän selvittämiseksi otetaan käyttöön lineaarisen funktion käsite.

**Määritelmä 2:** Funktio  $f: \mathbb{R} \rightarrow \mathbb{R}$  on *lineaarinen*, jos on olemassa  $c \in \mathbb{R}$  siten, että

$$f(x) = cx \quad (3)$$

kaikilla  $x \in \mathbb{R}$ .

Nyt voidaan todistaa:

**Lause 2:** Funktio  $f: \mathbb{R} \rightarrow \mathbb{R}$  on lineaarinen, jos ja vain jos

$$f(rx) = rf(x) \quad (4)$$

kaikilla  $r \in \mathbb{R}$  ja kaikilla  $x \in \mathbb{R}$ .

Termi lineaarinen tulee siitä, että tällaisen funktion kuvaaja on origon kautta kulkeva suora.

Määritelmästä seuraa välittömästi

**Lause 3:** Jokainen lineaarinen funktio  $f: \mathbb{R} \rightarrow \mathbb{R}$  on additiivinen.

Mutta onko jokainen additiivinen funktio lineaarinen? Jäljempänä osoitetaan, että ei ole. Mutta sitä ennen tutustutaan vähän tällaisen additiivisen ei-lineaarisen funktion ominaisuuksiin.

### Additiivisen ei-lineaarisen funktion ominaisuuksia

**Lause 4:** Jos funktio  $f: \mathbb{R} \rightarrow \mathbb{R}$  on additiivinen ja ei-lineaarinen, sen kuvaaja on tiheä koko tasossa.

Lause ”kuvaaja on tiheä koko tasossa” tarkoittaa, että tason  $\mathbb{R}^2$  jokaiselle pisteelle  $(a, b)$  ja jokaiselle positiiviselle reaaliluvulle  $r$  on olemassa  $x \in \mathbb{R}$  siten, että funktion  $f$  kuvaajan pisteen  $(x, f(x))$  etäisyys pisteestä  $(a, b)$

$$d = \sqrt{(x - a)^2 + (f(x) - b)^2}$$

on pienempi kuin  $r$ .

Lauseen 4 todistus: Valitaan mielivaltainen tason piste  $(a, b)$  ja mielivaltainen positiivinen reaaliluku  $r$  ja oletetaan, että funktio  $f$  on additiivinen eli sillä on omi-



naisuus (1), mutta se ei ole lineaarinen eli sillä ei ole ominaisuutta (3).

Pistettä  $(a, b)$  lähellä olevan funktion  $f$  kuvaajan pisteen  $(x, f(x))$  löytämiseksi todetaan aluksi, että kaikille rationaaliluvuille  $q_1$  ja  $q_2$  ja kaikille reaaliluvuille  $x_1$  ja  $x_2$  on tason piste

$$P_{q_1x_1q_2x_2} = (q_1x_1 + q_2x_2, q_1f(x_1) + q_2f(x_2)) \quad (5)$$

funktion  $f$  kuvaajan piste, sillä lauseen 1 ja määritelmän 1 mukaan

$$\begin{aligned} & (q_1x_1 + q_2x_2, q_1f(x_1) + q_2f(x_2)) \\ &= (q_1x_1 + q_2x_2, f(q_1x_1) + f(q_2x_2)) \\ &= (q_1x_1 + q_2x_2, f(q_1x_1 + q_2x_2)). \end{aligned}$$

Pitää siis löytää reaaliluvut  $x_1$  ja  $x_2$  sekä rationaaliluvut  $q_1$  ja  $q_2$  siten, että pisteen  $P_{q_1x_1q_2x_2}$   $x$ -koordinaatti  $q_1x_1 + q_2x_2$  on lähellä  $a$ :ta ja  $y$ -koordinaatti  $q_1f(x_1) + q_2f(x_2)$  lähellä  $b$ :tä.

Valitaan ensin mielivaltainen  $x_1 \in \mathbb{R} \setminus \{0\}$ . Koska  $f$  ei ole lineaarinen, on määritelmän 2 mukaan jokaiselle  $c \in \mathbb{R}$  olemassa  $x \in \mathbb{R}$  siten, että  $f(x) \neq cx$ . Koska  $x_1 \neq 0$ , voidaan valita  $c = f(x_1)/x_1$ , jolloin on olemassa  $x_2 \in \mathbb{R}$  siten, että

$$f(x_2) \neq \frac{f(x_1)}{x_1}x_2,$$

jolloin

$$x_1f(x_2) - x_2f(x_1) \neq 0. \quad (6)$$

Reaaliluvut  $x_1$  ja  $x_2$  on nyt valittu. Pitää vielä löytää rationaaliluvut  $q_1$  ja  $q_2$ , joille pisteen  $P_{q_1x_1q_2x_2}$   $x$ -koordinaatti  $q_1x_1 + q_2x_2$  on lähellä  $a$ :ta ja  $y$ -koordinaatti  $q_1f(x_1) + q_2f(x_2)$  lähellä  $b$ :tä. Tässä tarkoituksessa etsitään ensin reaaliluvut  $x$  ja  $y$ , joille

$$\begin{cases} xx_1 + yx_2 = a, \\ xf(x_1) + yf(x_2) = b. \end{cases} \quad (7)$$

Koska (6):n mukaan  $x_1f(x_2) - x_2f(x_1) \neq 0$ , on tällä yhtälöparilla ratkaisu

$$\begin{cases} x = \frac{af(x_2) - bx_2}{x_1f(x_2) - x_2f(x_1)}, \\ y = \frac{bx_1 - af(x_1)}{x_1f(x_2) - x_2f(x_1)}. \end{cases}$$

Näin saadut  $x$  ja  $y$  eivät kuitenkaan ole välttämättä rationaalilukuja, joten sijoittamalla yhtälöön (5) pisteen  $P_{q_1x_1q_2x_2}$  koordinaatteihin  $q_1$ :n paikalle  $x$  ja  $q_2$ :n paikalle  $y$  ei aina saada funktion  $f$  kuvaajan pistettä. Mutta nyt riittääkin löytää rationaaliluku  $q_1$  niin läheltä  $x$ :ää ja rationaaliluku  $q_2$  niin läheltä  $y$ :tä, että pisteen  $P_{q_1x_1q_2x_2}$  etäisyys pisteestä  $(a, b)$  on pienempi

kuin alussa valittu mielivaltainen positiivinen reaaliluku  $r$ .

Rationaalilukujen joukko on tiheä reaalilukujen joukossa, eli jokaisen reaaliluvun jokaisessa avoimessa ympäristössä on ainakin yksi rationaaliluku. Tämä tarkoittaa, että olipa  $\delta$  mikä tahansa positiivinen reaaliluku, niin on olemassa rationaaliluku  $q_1$  siten, että  $|q_1 - x| < \delta$ , ja kun edellä oli valittu  $x_1 \neq 0$ , niin

$$|q_1 - x| |x_1| < \delta |x_1|. \quad (8)$$

Tällöin on myös

$$|q_1 - x| |f(x_1)| \leq \delta |f(x_1)|. \quad (9)$$

Samoin on olemassa rationaaliluku  $q_2$  siten, että  $|q_2 - y| < \delta$ , jolloin

$$|q_2 - y| |x_2| \leq \delta |x_2| \quad (10)$$

ja

$$|q_2 - y| |f(x_2)| \leq \delta |f(x_2)|. \quad (11)$$

Kolmioepäyhtälönä tunnetun reaalilukujen ominaisuuden mukaan kaikilla  $a \in \mathbb{R}$  ja kaikilla  $b \in \mathbb{R}$  on  $|a + b| \leq |a| + |b|$ .

Yhtälöparin (7) ensimmäistä yhtälöä, kolmioepäyhtälöä ja epäyhtälöitä (8) ja (10) käyttäen saadaan

$$\begin{aligned} |q_1x_1 + q_2x_2 - a| &= |q_1x_1 + q_2x_2 - xx_1 - yx_2| \\ &\leq |q_1 - x| |x_1| + |q_2 - y| |x_2| \\ &< \delta |x_1| + \delta |x_2| = \delta(|x_1| + |x_2|). \end{aligned}$$

Yhtälöparin (7) toista yhtälöä, kolmioepäyhtälöä ja epäyhtälöitä (9) ja (11) käyttäen saadaan

$$\begin{aligned} |q_1f(x_1) + q_2f(x_2) - b| &= |q_1f(x_1) + q_2f(x_2) - xf(x_1) - yf(x_2)| \\ &\leq |q_1 - x| |f(x_1)| + |q_2 - y| |f(x_2)| \\ &\leq \delta |f(x_1)| + \delta |f(x_2)| = \delta(|f(x_1)| + |f(x_2)|). \end{aligned}$$

Näin valituilla luvuilla  $q_1$ ,  $x_1$ ,  $q_2$  ja  $x_2$  on piste

$$P_{q_1x_1q_2x_2} = (q_1x_1 + q_2x_2, q_1f(x_1) + q_2f(x_2))$$

funktion  $f$  kuvaajan piste. Sen ja pisteen  $(a, b)$  välinen etäisyys

$$\begin{aligned} & \sqrt{(q_1x_1 + q_2x_2 - a)^2 + (q_1f(x_1) + q_2f(x_2) - b)^2} \\ & < \delta \sqrt{(|x_1| + |x_2|)^2 + (|f(x_1)| + |f(x_2)|)^2} \end{aligned}$$

on pienempi kuin edellä valittu pisteen  $(a, b)$  avoimen ympäristön mielivaltainen säde  $r > 0$ , jos

$$\delta \leq \frac{r}{\sqrt{(|x_1| + |x_2|)^2 + (|f(x_1)| + |f(x_2)|)^2}}.$$

Lause 4 on näin todistettu.  $\square$

On siis osoitettu, että jos funktio on additiivinen, sen kuvaaja on joko origon kautta kulkeva suora tai pistejoukko, joka on tiheä koko tasossa.

Mielenkiintoista on erityisesti se, että tason jokaisen pisteen jokaisesta avoimesta ympäristöstä ei löydy ainostaan yhtä, vaan itse asiassa äärettömän monta funktion  $f$  kuvaajan pistettä, mikä voidaan todistaa vastaoletusta käyttäen.

Lauseesta 4 seuraa heti, että jos yhdelläkin tason pisteellä on avoin ympäristö, johon ei kuulu yhtään additiivisen funktion  $f$  kuvaajan pistettä, niin  $f$  on lineaarinen. Näin ollen esim. yhdelläkin avoimella välillä ylhäältä (tai alhaalta) rajoitettu additiivinen funktio on lineaarinen. Tästä taas seuraa, että additiivinen funktio on lineaarinen, jos se on yhdessäkin pisteessä  $x_0 \in \mathbb{R}$  jatkuva.

## Additiivisen ei-lineaarisen funktion olemassaolo

Nyt tiedetään, millaisia additiiviset ei-lineaariset funktiot ovat. Mutta onko sellaisia olemassa?

Yksi ns. *valinta-aksioman* seurauksista on

**Lause 5:** On olemassa reaalilukujen joukon osajoukko  $\mathcal{B}$  siten, että jokainen reaaliluku  $x$  voidaan esittää yksikäsitteisellä tavalla äärellisen monen joukkoon  $\mathcal{B}$  kuuluvan luvun  $x_1, \dots, x_n$  rationaalilukukertoimisena lineaarikombinaationa, siis muodossa

$$x = q_1 x_1 + \dots + q_n x_n,$$

missä  $\{q_1, \dots, q_n\} \subseteq \mathbb{Q}$ ,  $\{x_1, \dots, x_n\} \subseteq \mathcal{B}$ .

Joukkoa  $\mathcal{B}$  kutsutaan reaalilukujen *Hamelin kannaksi* saksalaisen matemaatikon Georg Karl Wilhelm Hamelin (12.9.1877–4.10.1954) mukaan.

Lauseen todistus menee niin syvälle joukko-opin perusteisiin, että se joudutaan tässä yhteydessä sivuuttamaan.

Kiusallista, mutta samalla erittäin viehättävää on se, että tätä joukkoa  $\mathcal{B}$  ei voida konstruoida, joten ei tiedetä, millainen esim. luvun 1 esitys on. Sen verran kuitenkin tiedetään, että luku 0 ei kuulu kantaan, sillä jos se kuuluisi ja esim. luvun 1 esitys kannan  $\mathcal{B}$  lukuja käyttäen olisi

$$1 = q_1 x_1 + \dots + q_n x_n,$$

niin olisi myös

$$1 = q_1 x_1 + \dots + q_n x_n + c \cdot 0$$

jokaiselle reaaliluvulle  $c$  ja eri esityksiä olisi siis äärettömän monta.

Tiedetään myös, että kannassa  $\mathcal{B}$  on vähintään kaksi lukua, sillä jos siinä olisi vain yksi luku  $x_1$ , olisi jokainen reaaliluku muotoa  $q x_1$ , missä  $q$  on rationaaliluku. Olisi siis oltava esim.  $1 = q_1 x_1$  jollekin rationaaliluvulle  $q_1$  ja  $\sqrt{2} = q_2 x_1$  jollekin rationaaliluvulle  $q_2$ , mistä seuraisi, että

$$\sqrt{2} = \frac{\sqrt{2}}{1} = \frac{q_2 x_1}{q_1 x_1} = \frac{q_2}{q_1}$$

olisi rationaaliluku.

Itse asiassa kanta  $\mathcal{B}$  on ylinumeroituvasti ääretön lukujoukko ja sen mahtavuus (kardinaliteetti) on sama kuin reaalilukujen joukon  $\mathbb{R}$ , eli on olemassa bijektio  $g: \mathcal{B} \rightarrow \mathbb{R}$ .

Oletetaan tästä alkaen, että  $x$ :n esityksestä jätetään aina pois kaikki termit, joissa on kertoimena 0. Toisin sanoen, oletetaan, että jos  $x \neq 0$  ja  $x = q_1 x_1 + \dots + q_n x_n$ , niin  $\{q_1, \dots, q_n\} \subseteq \mathbb{Q} \setminus \{0\}$ .

Jos taas  $x = 0$  ja  $x = q_1 x_1 + \dots + q_n x_n$ , niin esityksen yksikäsitteisyys vuoksi  $q_1 x_1 + \dots + q_n x_n = 0$  toteutuu vain, jos  $q_1 = \dots = q_n = 0$ , sillä 0:lla on myös esitys  $0 = 0x_1 + \dots + 0x_n$ .

Hamelin kantaa käyttäen voidaan nyt todistaa

**Lause 6:** On olemassa funktio  $f: \mathbb{R} \rightarrow \mathbb{R}$ , joka on additiivinen mutta ei lineaarinen.

Todistus: Valitaan ensin mielivaltainen funktio  $h: \mathcal{B} \rightarrow \mathbb{R}$  ja sen jälkeen funktio  $f: \mathbb{R} \rightarrow \mathbb{R}$  seuraavasti

$$f(x) = \begin{cases} 0, & \text{jos } x = 0, \\ q_1 h(x_1) + \dots + q_n h(x_n), & \text{jos } x \neq 0, \end{cases} \quad (12)$$

missä  $x = q_1 x_1 + \dots + q_n x_n$ ,  $\{q_1, \dots, q_n\} \subseteq \mathbb{Q} \setminus \{0\}$  ja  $\{x_1, \dots, x_n\} \subseteq \mathcal{B}$ .

Koska  $x$ :n esitys  $x = q_1 x_1 + \dots + q_n x_n$  on yksikäsitteinen, on myös  $f(x)$ :n arvo yksikäsitteinen ja  $f$  on näin ollen funktio.

Näin määritellylle funktiolle  $f$  pätee

$$f(x) = h(x) \quad (13)$$

jokaiselle  $x \in \mathcal{B}$ . Jos nimittäin  $x \in \mathcal{B}$ , niin  $x \neq 0$  ja koska  $x$ :n yksikäsitteinen esitys kannan  $\mathcal{B}$  lukujen lineaarikombinaationa on  $1x$ , niin  $f(x) = f(1x) = 1h(x) = h(x)$ .

Melko suoraviivaisesti edeten voidaan todistaa, että  $f$  on additiivinen.

Todistetaan nyt, että niiden funktioiden joukossa, joilla on ominaisuus (12), on ei-lineaarinen funktio.

Kuten edellä osoitettiin, on kannassa  $\mathcal{B}$  vähintään kaksi lukua, olkoot ne  $x_1$  ja  $x_2$ . Jotta  $f$  olisi lineaarinen, olisi määritelmän 2 mukaan oltava olemassa  $c \in \mathbb{R}$  siten, että  $f(x) = cx$  kaikilla  $x \in \mathbb{R}$ .

Siis  $f(x_1) = cx_1$  ja  $f(x_2) = cx_2$  jolloin

$$c = \frac{f(x_1)}{x_1} = \frac{f(x_2)}{x_2},$$

sillä kannan  $\mathcal{B}$  lukuina  $x_1 \neq 0$  ja  $x_2 \neq 0$ . Tällöin on yhtälön (13) mukaan

$$\frac{h(x_1)}{x_1} = \frac{f(x_1)}{x_1} = \frac{f(x_2)}{x_2} = \frac{h(x_2)}{x_2}.$$

Tämä tarkoittaa, että  $f$  on lineaarinen vain, jos

$$h(x_2) = \frac{h(x_1)}{x_1}x_2,$$

joten jos valitaan

$$h(x_2) \neq \frac{h(x_1)}{x_1}x_2,$$

saadaan ei-lineaarinen funktio  $f$ , jolla on ominaisuus (12) ja joka on näin ollen additiivinen.  $\square$

Tästä voidaan päätellä kaksi mielenkiintoista asiaa.

Ensiksi se, että pienikin  $h(x_2)$ :n poikkeama luvusta  $x_2 \frac{h(x_1)}{x_1}$  aiheuttaa sen, että  $f$  muuttuu lineaarisesta ei-lineaariseksi. Koska  $f$  on additiivinen, tämä tarkoittaa lauseen 4 mukaan, että funktion  $f$  kuvaaja hajoo origon kautta kulkevasta suorasta pistejoukoksi, joka on tiheä koko tasossa.

Toiseksi se, että kun on vain yksi  $h(x_2)$ :n arvo, jolla  $f$  on lineaarinen, ja äärettömän monta, joilla ei ole, lineaarisuus on harvinaista additiivisten funktioiden joukossa. Jopa niin harvinaista, että umpimähkään valittu additiivinen funktio on lineaarinen todennäköisyydellä 0.

Siitä, millaisia additiiviset ei-lineaariset funktiot ovat, saa jonkinlaisen käsityksen, jos alkaa piirtää sellaisen

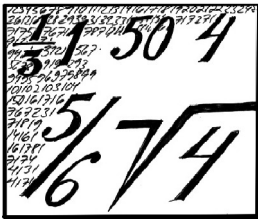
kuvaajaa. Tällöinhän on piirrettävä jokaiselle pystysuoralle suoralle yksi kuvaajan piste ja jokaiseen sen avoimeen ympäristöön äärettömän monta kuvaajan pistettä.

Ei siis ole ihme, jos additiivisia ei-lineaarisia funktioita kutsutaan joskus ”hirviöfunktioiksi” (”monster functions”).

Tästä kirjoituksesta on yksityiskohtaisempi versio Matematiikkalehti Solmun Oppimateriaalit-sivulla: <https://matematiikkalehtisolmu.fi/oppimateriaalit.html>

## Lähteet

1. Augustin-Louis Cauchy:  
[https://en.wikipedia.org/wiki/Augustin-Louis\\_Cauchy](https://en.wikipedia.org/wiki/Augustin-Louis_Cauchy) [25.11.2023]
2. Lauseen 4 todistuksen periaate:  
[https://en.wikipedia.org/wiki/Cauchy%27s\\_functional\\_equation](https://en.wikipedia.org/wiki/Cauchy%27s_functional_equation) [25.11.2023]
3. Georg Karl Wilhelm Hamel:  
[https://en.wikipedia.org/wiki/Georg\\_Hamel](https://en.wikipedia.org/wiki/Georg_Hamel) [25.11.2023]
4. Hamelin kannan olemassaolo ja kardinaliteetti:  
<http://thales.doa.fmph.uniba.sk/sleziak/texty/rozne/pozn/tm/hamel.pdf> [25.11.2023]
5. Additiivisen ei-lineaarisen funktion muodostamisen periaate:  
[https://en.wikipedia.org/wiki/Cauchy%27s\\_functional\\_equation](https://en.wikipedia.org/wiki/Cauchy%27s_functional_equation) [25.11.2023]



## Tarinoita polynomeista (osa 4)

**Jukka Tuomela**

Itä-Suomen yliopisto, Joensuu

jukka.tuomela@uef.fi

### Kertaus

Jatketaan aiempien kirjoitusten polynomiteemaa [6, 7, 8]. Laskuissa viitataan jakolaskuun, jonka algoritmi on esitetty kirjoituksessa [6]. Lukijalta oletetaan, että hän tietää, mitä tarkoittaa polynomien yhteen- ja kertolasku.

Erinomainen kirja polynomeista kiinnostuneille on [3], ja erityisesti tämän kirjoituksen asioista löytyy lisätietoa kirjasta [2]. wxMaxima, jonka avulla esimerkkejä voi helposti laskea, on vapaasti saatavilla oleva symbolisen laskennan ohjelmisto.<sup>1</sup>

Myös Sage on ilmainen ohjelmisto, jolla voi laskea sekä symbolisesti että numeerisesti [1].<sup>2</sup> Jäljempänä ovat tehtävät voi helposti laskea myös Sagen avulla. Samoin Wolfram Alphalla voi kätevästi laskea polynomeilla.<sup>3</sup>

Esimerkkeinä on taas käytetty matematiikkakilpailusivun tehtäviä.<sup>4</sup>

### Epäyhtälöt

Aiemmissa kirjoituksissa on analysoitu polynomiyhtälöitä, ja on havaittu, että muuttujien eliminointi, te-

kijöihinjako ja jakolaskualgoritmi ovat työkaluja, joita jatkuvasti on hyödyllistä käyttää. Katsotaan tässä epäyhtälöitä. Havaitaan, että jakolasku on edelleen jopa yllättävän hyödyllinen, mutta epäyhtälötehtävät johtavat myös aivan uudenslaisiin ideoihin.

Koska tehtävien polynomit ovat usein homogeenisia ja/tai symmetrisiä, niin palautetaan ensin mieleen, mitä näillä tarkoitetaan.

Funktio  $g$  on *homogeeninen* astetta  $k$ , jos pätee

$$g(tx_1, \dots, tx_n) = t^k g(x_1, \dots, x_n).$$

Esimerkiksi jos  $x$ ,  $y$  ja  $z$  ovat kolmion sivut ja  $A$  on pinta-ala, niin selvästi

$$A(tx, ty, tz) = t^2 A(x, y, z),$$

kun  $t > 0$ .  $A$  ei kuitenkaan ole polynomi, mutta kuten kohta nähdään, niin  $A^2$  on polynomi. Sen täytyy siis olla homogeeninen neljännen asteen polynomi.

Geometrian tehtävät usein johtavat homogeenisiin polynomeihin, koska klassisessa geometriassa ei ole mitään valittua skaalaa. Muistetaan myös viime kerralta [8], että klassisen geometrian tehtävät voidaan aina muotoilla polynomitehtävinä.

<sup>1</sup><https://wxmaxima-developers.github.io/wxmaxima/index.html>

<sup>2</sup><https://www.sagemath.org/index.html>

<sup>3</sup><https://www.wolframalpha.com/>

<sup>4</sup><https://matematiikkakilpailut.fi/aiheet/>



Tarkastellaan symmetriaa vain kahden ja kolmen muuttujan tapauksessa. Kahden muuttujan funktio  $f$  on symmetrinen, jos  $f(x, y) = f(y, x)$ . Kolmen muuttujan tapauksessa tulee erilaisia symmetrian asteita. Olkoon  $f$  muuttujien  $x$ ,  $y$  ja  $z$  funktio. Muuttujilla on kuusi permutaatiota:

$$\begin{array}{ccc} (x, y, z), & (z, x, y), & (y, z, x) \\ (y, x, z), & (x, z, y), & (z, y, x). \end{array}$$

Määritellään nyt

(i)  $f$  on symmetrinen, jos

$$\begin{aligned} f(x, y, z) &= f(z, x, y) = f(y, z, x) = \\ f(y, x, z) &= f(x, z, y) = f(z, y, x). \end{aligned}$$

(ii)  $f$  on syklistesti symmetrinen, jos

$$f(x, y, z) = f(z, x, y) = f(y, z, x).$$

Myös symmetriaa esiintyy usein geometrian tehtävissä, koska tyypillisesti geometrisesti mielekkäät asiat eivät voi riippua siitä, miten eri asioita, kuten kolmion sivuja, nimetään.

*Esimerkki 1.* Äsken jo huomattiin, että  $A$  on homogeeninen; sen täytyy olla myös symmetrinen, koska pinta-ala ei voi riippua siitä, miten sivut on nimetty. Edelleen pinta-ala on nolla, jos  $x + y = z$ , joten polynomin

$$g = (x + y - z)(x + z - y)(y + z - x)$$

täytyy olla  $A^2$ :n tekijä. Siis

$$A^2 = c(x + y + z)g$$

jollekin  $c$ . Kutsutaan näin saatua polynomia *Heronin polynomiksi*.<sup>5</sup>

$$\begin{aligned} h_A &= (x + y + z)g \\ &= 2(x^2y^2 + x^2z^2 + y^2z^2) - x^4 - y^4 - z^4 \\ &= (x^2 + y^2 + z^2)^2 - 2(x^4 + y^4 + z^4). \end{aligned}$$

Koska tiedetään, että  $A^2 = 1/4$ , kun  $x = y = 1$  ja  $z = \sqrt{2}$ , niin  $c = 1/16$ , joten  $A^2 = h_A/16$ . Huomaa, että  $h_A$  ei ole positiivinen kaikilla muuttujien arvoilla. ☆

*Esimerkki 2.* Olkoot  $x$ ,  $y$  ja  $z$  kolmion sivut ja  $A$  kolmion pinta-ala. Osoita, että<sup>6</sup>

$$x^2 + y^2 + z^2 \geq 4\sqrt{3}A.$$

Tämä oli IMO:n matematiikkakilpailuissa vuonna 61. Matematiikkavalmennuksen 15. tehtävä kesällä 2022 oli muuten sama, mutta vakion  $4\sqrt{3}$  tilalla oli 6.

<sup>5</sup>[https://www.wikiwand.com/en/Heron%27s\\_formula](https://www.wikiwand.com/en/Heron%27s_formula)

<sup>6</sup>Tämä tunnetaan *Weitzenböckin epäyhtälönä*.

Heronin polynomin avulla nähdään, että

$$\begin{aligned} f &= \frac{1}{2} \left( (x^2 + y^2 + z^2)^2 - 3h_A \right) \\ &= 3(x^4 + y^4 + z^4) - (x^2 + y^2 + z^2)^2 \\ &= (x^2 - y^2)^2 + (x^2 - z^2)^2 + (y^2 - z^2)^2 \geq 0. \end{aligned}$$

Selvästi yhtäsuuruus pätee vain, jos  $x = y = z$ . ☆

Esimerkin ratkaisussa tuli ilmi eräs tärkeä strategia: pyritään kirjoittamaan polynomi summana, jonka jokainen termi on jotenkin selvästi positiivinen. On osoitautunut, että tuo neliöllinen tapaus on niin tärkeä, että sille on annettu oma nimi:

$f$  on *sos-polynomi* (sum of squares), jos on olemassa polynomit  $q_j$  siten, että

$$f = \sum_{j=1}^m q_j^2.$$

Käytännön laskuissa on kätevämpi välttää neliöjuuria ja kirjoittaa  $c_j q_j^2$ , missä  $c_j > 0$ .

Joskus epäyhtälö ikään kuin ratkeaa itsestään, koska annettu polynomi on valmiiksi neliö; esimerkiksi IMO:n 2. tehtävässä vuonna 2008 piti osoittaa, että

$$(x^2y^2 - 3xy + x + y)^2 \geq 0.$$

Samoin matematiikkavalmennuksen 16. tehtävässä syyskuussa 2022 piti osoittaa, että  $h^2 > 0$ , kun  $x$ ,  $y$  ja  $z$  ovat erisuuria, missä

$$h = xyz - (x + y - z)(x + z - y)(y + z - x).$$

Esa Vesalaisen tehtäväkokoelmassa 20. tehtävässä puolestaan piti osoittaa, että

$$(x^2 - 2x - 1)(x^2 - 2)(2x - 1)^2 \geq 0,$$

kun  $0 \leq x \leq 1$ . Tietenkin nämä kaikki tehtävät oli muotoiltu niin, ettei ratkaisua heti nähnyt.

Epäyhtälötehtävissä on usein joitain lisärajoituksia. Esimerkiksi voitaisiin haluta osoittaa, että  $f \geq 0$ , kun lisäksi oletetaan, että  $g_1 = 0$  ja/tai  $g_2 \geq 0$ . Tässä tapauksessa strategiana olisi etsiä esitysmuotoa

$$f = r + q_1g_1 + q_2g_2. \quad (1)$$

Jos  $r \geq 0$  ja  $q_2 \geq 0$ , niin tämä todistaa väitteen. Huomaa, että tavallaan  $r$  on jakojäännös ja  $q_j$  ovat osamääriä, joten varmaankin jakolaskualgoritmi on tässä hyödyllinen.

*Esimerkki 3.* Osoitetaan ideaa (1) käyttäen, että

$$h = xyz - (x + y - z)(x + z - y)(y + z - x) \geq 0,$$

kun  $x \geq 0$ ,  $y \geq 0$  ja  $z \geq 0$ .<sup>7</sup> Koska  $h$  on selvästi symmetrinen, voidaan olettaa, että  $z \geq x \geq y$ . Jaetaan ensin polynomilla  $z - x$ :

$$\begin{aligned} h &= q(z - x) + y(x - y)^2, \\ q &= z(z - y) - (x - y)^2. \end{aligned}$$

Jakojäännös on positiivinen; jaetaan sitten osamäärä polynomilla  $x - y$ :

$$q = (y + z - x)(x - y) + z(z - x).$$

Siis osamääräkin on positiivinen, joten  $h \geq 0$ .

Jos  $x$ ,  $y$  ja  $z$  ovat jonkin kolmion sivujen pituuksia, niin jakamalla polynomilla  $g = (z - x)(z - y)$  saadaan yhdellä jakolaskulla

$$h = zg + (x - y)^2(x + y - z) \geq 0.$$

Toisaalta tämän voi todistaa tavallaan ilman mitään laskuja järjestämällä  $h$ :n termit uudelleen:

$$h = (z - x)(z(z - y) - x(x - y)) + y(z - y)(x - y).$$

Selvästi  $h \geq 0$ , kun  $z \geq x \geq y$ .

Tämän avulla voidaan ratkaista pohjoismaitten matematiikkakilpailuissa vuonna 2004 ollut geometrian tehtävä. Siinä piti osoittaa, että jos  $x$ ,  $y$  ja  $z$  ovat kolmion sivut ja  $r$  on kolmion kärkien kautta kulkevan ympyrän säde, niin

$$f = (x + y + z)r^2 - xyz \geq 0.$$

Muistetaan geometriasta, että<sup>8</sup>

$$4Ar = xyz,$$

missä  $A$  on kolmion pinta-ala. Käytetään taas Heronin polynomia  $h_A$  ja sitä, että  $h_A \geq 0$ , kun  $x$ ,  $y$  ja  $z$  on kolmion sivut. Tästä saadaan

$$\begin{aligned} h_A f &= xyz \left( (x + y + z)xyz - h_A \right) \\ &= xyz(x + y + z)h \geq 0. \end{aligned}$$

Epäyhtälö  $h \geq 0$  ratkaisee myös IMO:n toisen tehtävän vuodelta 2000 ja pohjoismaisen matematiikkakilpailun toisen tehtävän vuodelta 2005. ☆

*Esimerkki 4.* IMO:n tehtävä vuodelta 1983. Oletetaan, että  $x$ ,  $y$  ja  $z$  ovat kolmion sivut; osoita, että

$$f = yx^2(x - y) + zy^2(y - z) + xz^2(z - x) \geq 0. \quad (2)$$

$f$  on syklisesti symmetrinen, joten riittää tarkastella kahta tapausta: (i)  $z \geq x \geq y$  ja (ii)  $y \geq x \geq z$ .

Tapauksessa (i) jaetaan käyttämällä tuttua polynomia  $g = (z - x)(z - y)$ :

$$f = (xz + y(x - y))g + xy(x - y)^2.$$

Tämä on selvästi positiivinen. Tässä ei siis tarvittu sitä tietoa, että kyseessä on kolmion sivut. Tapauksessa (ii) jaetaan polynomilla  $x + z - y$ , jolloin saadaan

$$f = q(x + z - y) + 2x(y - x)^3.$$

Siis jakojäännös on positiivinen. Jaetaan sitten osamäärä polynomilla  $x - z$ , mistä saadaan

$$q = (y(y - x) + x(x - z))(x - z) + x(y - x)^2 \geq 0,$$

joten  $f \geq 0$  tässäkin tapauksessa. ☆

*Esimerkki 5.* Pohjoismaitten matematiikkakilpailuissa vuonna 1987 piti osoittaa, että

$$f = x^2y^4 + y^2z^4 + z^2x^4 - xyz(xy^2 + yz^2 + zx^2) \geq 0,$$

kun muuttujat ovat positiivisia.  $f$  on syklisesti symmetrinen, joten periaatteessa pitäisi tarkastella kahta tapausta: (1)  $z \geq x \geq y$  ja (2)  $z \geq y \geq x$ . Kuitenkin molemmissa tapauksissa ongelma ratkeaa, kun jaetaan polynomilla  $g = (z - x)(z - y)$ :

$$\begin{aligned} f &= qg + r, \\ q &= y^2z(z + y) + y^4 - x^3y + x^4, \\ r &= (x - y)^2(y^2 + xy + x^2)(yz + xz - xy). \end{aligned}$$

Selvästi  $q \geq 0$  ja  $r \geq 0$ .

Myös IMO:n 4. tehtävä vuodelta 1963, Baltian tie kilpailun 5. tehtävä vuodelta 1991 ja Vesalaisen tehtävät 17, 27, 28 ja 49 ratkeavat jakamalla polynomilla  $g = (z - x)(z - y)$ .

$f$  voidaan symmetrisoida sijoituksella  $\hat{x}^3 = xy^2$ ,  $\hat{y}^3 = yz^2$  ja  $\hat{z}^3 = zx^2$ , jolloin saadaan symmetrinen polynomi

$$\hat{f} = \hat{x}^6 + \hat{y}^6 + \hat{z}^6 - \hat{x}\hat{y}\hat{z}(\hat{x}^3 + \hat{y}^3 + \hat{z}^3).$$

Vesalaisen tehtävässä 27 puolestaan piti osoittaa, että

$$f_0 = x^4 + y^4 + z^4 - xyz(x + y + z) \geq 0.$$

Itse asiassa  $\hat{f}$  ja  $f_0$  ovat sos-polynomeja, joten ne ovat positiivisia kaikilla arvoilla. Katsotaan tätä kohta. ☆

Tuntuu siltä, että yleispolynomi  $g = (z - x)(z - y)$  ratkaisee monet ongelmat. Tämä oli minullekin yllätys: ajattelin, että laskemalla kuten kaavassa (1) voisi ratkaista joitain tehtäviä; kuten edellä nähtiin, niin osoitautui kuitenkin, että monet kilpailutehtävät ratkeavat näin. Lisäksi ei ollut mitenkään vaikea keksiä millä polynomilla jakaa; päinvastoin voisi sanoa, että nuo

<sup>7</sup>Tämä on erikoistapaus *Schurin epäyhtälöstä*: [https://www.wikiwand.com/en/Schur%27s\\_inequality](https://www.wikiwand.com/en/Schur%27s_inequality)

<sup>8</sup>[https://www.wikiwand.com/en/Circumscribed\\_circle](https://www.wikiwand.com/en/Circumscribed_circle)

käytetyt polynomit ovat sellaisia, mitkä ensimmäisenä tulevat mieleen.

Päällisin puolin ei näytä olevan mitään selkeää syytä sille, että jakojäännös ja osamäärä osoittautuivat positiiviseksi. Periaatteessahan voisi yhtä hyvin käydä, että osamäärä (jakojäännös) on sopivasti positiivinen, kun jakojäännös (osamäärä) on negatiivinen. Jos siis lähtee laskemaan käyttäen jakolaskua ja yhtälöä (1), niin tietääkseni ei ole mitään lausetta tai edes jotain ideaa, miksi tämä välttämättä johtaisi ratkaisuun. Kuitenkin näin tuntuu usein käyvän, enkä osaa edes muotoilla täsmällisesti, mikä lause tässä pitäisi todistaa.

Luonnollisesti ideaa (1) voidaan systemaattisesti soveltaa; tätä on tarkemmin kuvattu kirjassa [2, section 3.4].

## Epäyhtälöt ja optimointi

Jos halutaan osoittaa, että  $f \geq 0$  joko koko avaruudessa tai jossain sen osajoukossa, niin tähän on oikeastaan optimointitehtävä, koska halutaan osoittaa, että  $f$ :n minimi on einegatiivinen halutussa joukossa. Ottamalla sitten käyttöön tietyt differentiaalilaskennan työkalut tehtävä voidaan joskus näppärästi ratkaista tätä kautta. Tietenkin kilpailutehtävissä tehtävät on muotoiltu niin, että ne ratkeavat muutenkin, mutta yleisesti ottaen kannattaa muistaa tämäkin näkökulma. Jätänkin harjoitustehtäväksi seuraavan:

Olkoon  $h$  esimerkissä 3 ollut polynomi. Osoita differentiaalilaskennan avulla, että  $h \geq 0$ , kun  $x \geq 0$ ,  $y \geq 0$  ja  $z \geq 0$ . Ohje: käytä hyväksi sitä, että  $h$  on homogeeninen.

## SOS-polynomit

Palataan sitten algebrallisiin menetelmiin ja tarkastellaan tässä vain tapausta, jossa haluttaisiin osoittaa, että polynomi  $f \geq 0$  kaikilla muuttujien arvoilla. Eräs idea olisi etsiä  $f$ :lle sos-esitystä. Mutta onko jokainen positiivinen polynomi sos-polynomi? Valitettavasti näin ei ole kuin muutamassa tapauksessa.

**Lause 1.** *Olkoon  $n$  polynomien muuttujien määrä ja  $d$  polynomien aste. Positiiviset polynomit ovat aina sos-polynomeja vain seuraavissa tapauksissa:*

- (i)  $n = 1$  ja  $d$  on mielivaltainen parillinen luku,
- (ii)  $d = 2$  ja  $n$  mielivaltainen,
- (iii)  $n = 2$  ja  $d = 4$ .

Kohdat (i) ja (ii) ovat varsin selkeitä tapauksia. Katsoaan kohta (i) tässä esimerkin avulla ja palataan kohtaan (ii) myöhemmin.

*Esimerkki 6.* Jos  $f$  on positiivinen yhden muuttujan polynomi, jolla on vain reaalisia nollakohtia, niin se on jo valmiiksi neliö. Katsotaan siis vain kompleksinen tapaus esimerkin avulla. Muistetaan identiteetti

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2. \quad (3)$$

Tämän avulla saadaan

$$\begin{aligned} f &= x^4 + 6x^3 + 39x^2 + 66x + 58 \\ &= (x^2 + 2x + 2)(x^2 + 4x + 29) \\ &= ((x + 1)^2 + 1)((x + 2)^2 + 25) \\ &= (x^2 + 3x + 7)^2 + (4x + 3)^2. \end{aligned}$$

Siis soveltamalla identiteettiä (3) rekursiivisesti nähdään, että mikä tahansa positiivinen yhden muuttujan polynomi voidaan esittää korkeintaan kahden polynomin neliön summana.<sup>9</sup> ☆

Sen sijaan lauseen kohta (iii), ja se, että nämä todella ovat ainoat tapaukset, ei ole mitenkään selvää. Hilbert todisti tämän 1800-luvun lopussa, mutta konkreettinen esimerkki positiivisesta polynomista, joka ei ole sos-polynomi löydettiin vasta paljon myöhemmin. Helpoin esimerkki löytyy tapauksessa  $n = 2$  ja  $d = 6$ ; tämä on kuuluisa Motzkinin polynomi:

$$f_m = x^4 y^2 + x^2 y^4 - 3x^2 y^2 + 1. \quad (4)$$

Katsotaan kohta, miksi tämä ei voi olla sos-polynomi, ja osoitetaan sitten, että  $f_m \geq 0$  kuitenkin pätee.

Vaikka siis kaikki positiiviset polynomit eivät ole sos-polynomeja, niin tilanne on kuitenkin parempi kuin tuon lauseen perusteella voisi kuvitella. Katsotaan kuitenkin ensin, miten sos-esitysmuotoa voisi systemaattisesti etsiä.

*Esimerkki 7.* Äskeisessä esimerkissä polynomien

$$f = x^4 + 6x^3 + 39x^2 + 66x + 58$$

sos-muoto löydettiin, koska nollakohdat tunnettiin. Tyypillisesti nollakohtia ei tunneta, joten miten silloin voisi edetä? Tässä tarvitaan jokin matriisilaskun käsitteitä ja merkintöjä, jotka on esitelty kirjoituksen lopussa.

Jos  $f = \sum_j q_j^2$ , niin  $q_j$ :t ovat toisen asteen polynomeja. Olkoot  $v = (1, x, x^2)$  ja  $A$  symmetrinen  $3 \times 3$  matriisi. Yritetään esittää  $f$  muodossa

$$f = \langle v, Av \rangle.$$

Koska  $A$ :ssa on 6 vapaata parametria ja  $f$ :ssä on 5 kerrointa, niin voisi kuvitella, että yksi parametri voidaan

<sup>9</sup>Tässä on yhdistetty kahta hyvin yleisesti käytettyä todistustekniikkaa: *proof by example* ja *proof by omission*.  
<https://www.math.utah.edu/~cherk/mathjokes.html>

valita vapaasti. Kertomalla auki ja vertaamalla kertoimia nähdäänkin, että

$$A = \begin{pmatrix} 58 & 33 & (39-c)/2 \\ 33 & c & 3 \\ (39-c)/2 & 3 & 1 \end{pmatrix}.$$

Nyt pitäisi valita  $c$  siten, että  $A$  olisi psd (positiivisemidefiniitti). Lauseen 2 ja kriteerin (6) perusteella on järkevää valita parametri niin, että lävistäjälkiot olisivat mahdollisimman isoja ja lävistäjän lukuosuudet alkioit itseisarvoltaan mahdollisimman pieniä. Tuntuu siis luonnolliselta kokeilla arvoa  $c = 39$ .<sup>10</sup> Saadaan siis

$$A = \begin{pmatrix} 58 & 33 & 0 \\ 33 & 39 & 3 \\ 0 & 3 & 1 \end{pmatrix}.$$

Lasketaan sitten  $A$ :n LDL-hajotelma  $A = LDL^T$  (määritelmä 1 ja lause 3). Tämän voi laskea Sagesa seuraavasti. Ensin määritellään

$$A = \text{matrix}(\text{QQ}, [[58,33,0],[33,39,3],[0,3,1]])$$

Tämän jälkeen hajotelma saadaan komennolla

$$[P,L,D] = A.\text{block\_ldlt}()$$

missä  $P$ ,  $L$  ja  $D$  ovat kuten lauseessa 3. Tämä antaa  $P = I$ ,  $D = \text{diag}(58, 1173/58, 217/391)$  ja

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 33/58 & 1 & 0 \\ 0 & 58/391 & 1 \end{pmatrix}.$$

MATLABISSA<sup>11</sup> sama hajotelma saadaan numeerisesti komendoilla

$$A = [58,33,0;33,39,3;0,3,1] \\ [L,D,P] = \text{ldl}(A)$$

Määritellään nyt  $L^T v = q$ , minkä jälkeen

$$f = \langle v, Av \rangle = \langle v, LDL^T v \rangle \\ = \langle L^T v, DL^T v \rangle = 58q_1^2 + \frac{1173}{58}q_2^2 + \frac{217}{391}q_3^2 \\ = 58(1 + 33x/58)^2 + \frac{1173}{58}(x - 58x^2/391)^2 + \frac{217}{391}x^4.$$

Kolmas yhtäsuuruus seuraa siis kaavasta (5). Huomataan, että saatiin aivan erilainen sos-muoto kuin äsken. Tämä onkin tyypillistä: sos-esitys ei ole yksikäsitteinen. ☆

Tässä jo nähtiin, miten sos-esitystä voisi etsiä yleisessä tapauksessa:

1. Laitetaan tarvittavat monomit vektoriin  $v$ .
2. Esitetään annettu polynomi  $f$  muodossa

$$f = \langle v, Av \rangle,$$

missä  $A$  on symmetrinen.

3. Tyypillisesti joitain  $A$ :n alkioita voi valita vapaasti. Pyritään valitsemaan nämä siten, että  $A$  on psd.

4. Lasketaan LDL-hajotelma  $A = LDL^T$  ja asetetaan  $q = L^T v$ . Tästä saadaan

$$f = \sum_i d_{ii} q_i^2.$$

Jos  $A$  on psd, niin  $d_{ii} \geq 0$ .

Tämän avulla voidaankin helposti todistaa äsken lauseen kohta (ii). Jos  $x = (x_1, \dots, x_n)$  ja  $f$  on muuttujien  $x_j$  toisen asteen homogeeninen polynomi, niin on olemassa symmetrinen  $A$  siten, että  $f = \langle x, Ax \rangle$ . Jos  $f \geq 0$ , niin  $A$  on psd ja sos-muoto saadaan kuten yllä kohdassa 4.

## Newton ja Motzkin

Äsken menetelmän kohdassa 1 puhuttiin tarvittavista monomeista. Jos  $f$  on vaikkapa 8. asteen polynomi, niin yleisesti ottaen polynomit  $q_j$  ovat yleisiä 4. asteen polynomeja. Jos muuttujia on  $n$ , niin astetta  $d$  olevassa polynomissa on  $\binom{n+d}{d}$  eri monomia, joten monomien määrä kasvaa varsin nopeasti sekä  $d$ :n että  $n$ :n suhteen. Joskus kuitenkin pärjätään paljon pienemmällä määrällä. Katsotaan tämä vain kahden muuttujan tapauksessa, mutta idea yleistyy sellaisenaan  $n$ :n muuttujan tapaukseen.

Liitetään monomiin  $x^i y^k$  tason piste  $(i, k)$ . Kahden muuttujan polynomiin  $f$  voidaan siis liittää pisteet kaikista sen monomeista, ja *Newtonin monikulmio*  $\mathcal{N}(f)$  on pienin konvekssi monikulmio, joka sisältää kaikki  $f$ :n monomien pisteet. Jos nyt  $g$  on jokin toinen polynomi, niin voidaan osoittaa, että

$$\mathcal{N}(fg) = \mathcal{N}(f) + \mathcal{N}(g).$$

Oikealla puolella olevaa summaa sanotaan *Minkowskin summaksi*. Jos siis  $A, B \subset \mathbb{R}^n$ , niin näitten Minkowskin summa on

$$A + B = \{p \in \mathbb{R}^n \mid p = a + b, a \in A, b \in B\}.$$

Tätä hyväksi käyttäen voidaan edelleen osoittaa, että jos  $f = \sum_j q_j^2$ , niin

$$\mathcal{N}(q_j) \subset \frac{1}{2} \mathcal{N}(f).$$

Osoitetaan tämän avulla, että Motzkinin polynomi (4) ei ole sos-polynomi.

Kuvassa 8 on  $\mathcal{N}(f_m)$  ja sen ”puolikas”. Jos siis pätsi, että  $f_m = \sum_j q_j^2$ , niin  $q_j$  olisi muotoa

$$q_j = c_{j,0} + c_{j,1}xy + c_{j,2}x^2y + c_{j,3}xy^2.$$

<sup>10</sup>Jätän harjoitustehtäväksi sen tutkimisen mitä tapahtuu, kun  $c = 25$  tai  $c = 45$ , ja miksi nämä arvot ovat mielenkiintoisia.

<sup>11</sup><https://se.mathworks.com/help/matlab/ref/ldl.html>

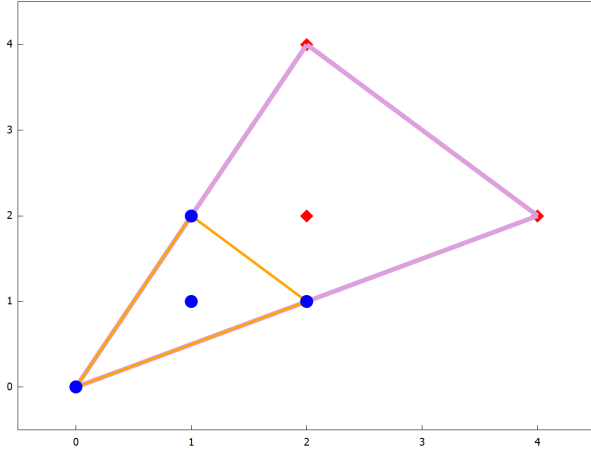


Huomaa, että  $q_j$  on kolmannen asteen polynomi, joten periaatteessa siinä voisi olla 10 monomia ja siis 10 tuntematonta kerrointa. Newtonin monikulmion avulla kertoimien ja monomien määrä voidaan rajoittaa neljään, mikä on huomattava yksinkertaistus.

Mutta jos  $q_j$  on edellä olevaa muotoa, niin

$$q_j^2 = c_{j,1}^2 x^2 y^2 + \dots,$$

mutta toisaalta  $f_m = -3x^2 y^2 + \dots$ , joten yhtäsuuruus on mahdoton.



Kuva 8: Motzkinin polynomien Newtonin monikulmio.

Kun Hilbert oli osoittanut, että yleensä positiiviset polynomit eivät ole sos-polynomeja, hän pohti, voisivatko kaikki positiiviset polynomit olla rationaalifunktioiden neliöitä. Hilbert selvästi piti tätä mielenkiintoisena kysymyksenä, koska tämä oli 17. ongelma, jonka hän esitti kuuluisassa esitelmässään vuonna 1900.<sup>12</sup> Parikymmentä vuotta myöhemmin Artin osoitti, että sopivat rationaalifunktiot aina löytyy. Artinin tulos voidaan muotoilla seuraavasti:

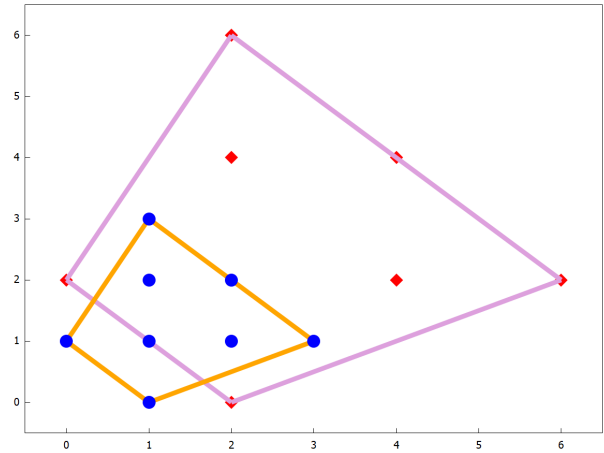
Jos  $f \geq 0$ , niin on olemassa sos-polynomi  $q$  siten, että  $qf$  on sos-polynomi.

Artinin lauseen takia siis sos-tekniikat ovat paljon hyödyllisempiä kuin lauseen 1 perusteella voisi kuvitella. Nyt voitaisiinkin yrittää etsiä jotain sos-polynomia  $q$  siten, että  $qf_m$  olisi sos-polynomi. Jos tällainen  $q$  löytyy, niin tästä tiettävästi seuraa, että  $f_m \geq 0$ .

Nyt osoittautuu, että riittää valita  $q = x^2 + y^2$ . Olkoon siis  $\hat{f}_m = (x^2 + y^2)f_m$ . Kuvassa 9 on esitetty vastaavat Newtonin monikulmiot; tarvitaan siis monomivektori

$$v = (x, y, xy, x^2y, xy^2, x^3y, x^2y^2, xy^3).$$

Huomaa, että ilman Newtonin monikulmiota voisi luulla, että monomivektoriin pitää laittaa kaikki  $\binom{6}{4} = 15$  komponenttia.



Kuva 9: Polynomia  $\hat{f}_m$  vastaava Newtonin monikulmio.

Yritetään nyt etsiä symmetrinen  $A \in \mathbb{R}^{8 \times 8}$  siten, että

$$\hat{f}_m = \langle v, Av \rangle.$$

Periaatteessa  $A$ :ssa on 36 vapaata parametria. Symmetrian avulla voidaan kuitenkin heti pienentää parametrien määrää. Olkoon

$$\tilde{v} = (y, x, xy, y^2x, yx^2, y^3x, x^2y^2, yx^3).$$

Koska  $f_m$  ja siis  $\hat{f}_m$  ovat symmetrisiä, niin täytyy päteä

$$\langle v, Av \rangle = \langle \tilde{v}, A\tilde{v} \rangle.$$

Toisaalta  $\tilde{v} = Pv$ , missä

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

Siispä kaavan (5) perusteella

$$\begin{aligned} \langle \tilde{v}, A\tilde{v} \rangle &= \langle Pv, APv \rangle = \langle v, P^T APv \rangle \\ &= \langle v, PAPv \rangle = \langle v, Av \rangle. \end{aligned}$$

Täytyy siis päteä, että  $A = PAP$ . Vertaamalla matriiseja nähdään, että on enää 21 vapaata parametria:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} \\ a_{12} & a_{11} & a_{13} & a_{15} & a_{14} & a_{18} & a_{17} & a_{16} \\ a_{13} & a_{13} & a_{33} & a_{34} & a_{34} & a_{36} & a_{37} & a_{36} \\ a_{14} & a_{15} & a_{34} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} \\ a_{15} & a_{14} & a_{34} & a_{45} & a_{44} & a_{48} & a_{47} & a_{46} \\ a_{16} & a_{18} & a_{36} & a_{46} & a_{48} & a_{66} & a_{67} & a_{68} \\ a_{17} & a_{17} & a_{37} & a_{47} & a_{47} & a_{67} & a_{77} & a_{67} \\ a_{18} & a_{16} & a_{36} & a_{48} & a_{46} & a_{68} & a_{67} & a_{66} \end{pmatrix}.$$

<sup>12</sup>[https://en.wikipedia.org/wiki/Hilbert%27s\\_problems](https://en.wikipedia.org/wiki/Hilbert%27s_problems)

Yhtälöstä  $\hat{f}_m = \langle v, Av \rangle$  saadaan periaatteessa 24 ehtoa, koska Newtonin monikulmiossa  $\mathcal{N}(\hat{f}_m)$  on 24 monomia vastaavaa pistettä. Symmetrian takia saadaan kuitenkin vain 14 riippumatonta yhtälöä. Kun vain kertoo kaikki auki, niin havaitaan, että yhtälöt ovat lineaarisia ja vieläpä hyvin yksinkertaisia. Kun ratkaistaan nämä yhtälöt, niin jäljelle jää vielä 7 parametria:

$$A = (A_1, A_2),$$

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{-a_{33}}{4} \\ 0 & 0 & a_{33} & -a_{17} - a_{18} \\ 0 & \frac{-a_{33}}{4} & -a_{17} - a_{18} & a_{44} \\ \frac{-a_{33}}{4} & 0 & -a_{17} - a_{18} & -a_{37} \\ 0 & a_{18} & \frac{-a_{44}}{2} - \frac{3}{2} & 0 \\ a_{17} & a_{17} & a_{37} & a_{47} \\ a_{18} & 0 & \frac{-a_{44}}{2} - \frac{3}{2} & -a_{47} \end{pmatrix},$$

$$A_2 = \begin{pmatrix} \frac{-a_{33}}{4} & 0 & a_{17} & a_{18} \\ 0 & a_{18} & a_{17} & 0 \\ -a_{17} - a_{18} & \frac{-a_{44}}{2} - \frac{3}{2} & a_{37} & \frac{-a_{44}}{2} - \frac{3}{2} \\ -a_{37} & 0 & a_{47} & -a_{47} \\ a_{44} & -a_{47} & a_{47} & 0 \\ -a_{47} & 1 & 0 & 1 - \frac{a_{77}}{2} \\ a_{47} & 0 & a_{77} & 0 \\ 0 & 1 - \frac{a_{77}}{2} & 0 & 1 \end{pmatrix}.$$

Nyt muistetaan lause 2 ja pyritään laittamaan lävistäjän ulkopuoliset parametrit nollassi, ja lävistäjälätkiot mahdollisimman isoiksi. Heti nähdään, että voidaan asettaa  $a_{17} = a_{18} = a_{37} = a_{47} = 0$ . Tämän jälkeen 7. rivillä on enää vain  $a_{77}$ , joten valitaan sekin nollassi. Nyt  $A$ :n ensimmäinen rivi toteuttaa ehdon (6), jos  $a_{33} \leq 4$ , joten valitaan  $a_{33} = 4$ . Lopuksi ehto (6) pätee neljännellä rivillä, jos  $a_{44} \geq 1$ , ja kolmannella rivillä, jos  $0 \leq a_{44} \leq 1$  joten valitaan  $a_{44} = 1$ , jolloin

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

Kun nyt lasketaan hajotelma  $A = LDL^T$ , niin huomataan, että  $D$ :ssä on vain kolme nollassa poikkeavaa

alkiota:  $d_{11} = d_{22} = 1$  ja  $d_{33} = 4$ . Kun merkitään  $q = L^T v$ , niin voidaan lopulta kirjoittaa

$$\begin{aligned} \hat{f}_m &= (x^2 + y^2)f_m = q_1^2 + q_2^2 + 4q_3^2, \\ q_1 &= y(1 - x^2), \\ q_2 &= x(1 - y^2), \\ q_3 &= xy(x^2 + y^2 - 2)/2. \end{aligned}$$

*Esimerkki 8.* Esimerkissä 4 mainitussa Vesalaisen tehtävässä 27 piti osoittaa, että

$$f_0 = x^4 + y^4 + z^4 - xyz(x + y + z) \geq 0.$$

Näytetään, että tämä on sos-polynomi. Kyseessä on homogeeninen neljännen asteen polynomi, joten tarvitaan monomivektori

$$v = (x^2, xy, xz, y^2, yz, z^2).$$

Newtonin monikulmion avulla nähdään, että kaikkia näitä monomeja saatetaan tarvita. Nyt yhtälössä  $f_0 = \langle v, A_0 v \rangle$  on periaatteessa 21 parametria, mutta symmetrian perusteella jäljelle jää vain 6 parametria:

$$A_0 = \begin{pmatrix} c_0 & c_1 & c_1 & c_2 & c_3 & c_2 \\ c_1 & c_4 & c_5 & c_1 & c_5 & c_3 \\ c_1 & c_5 & c_4 & c_3 & c_5 & c_1 \\ c_2 & c_1 & c_3 & c_0 & c_1 & c_2 \\ c_3 & c_5 & c_5 & c_1 & c_4 & c_1 \\ c_2 & c_3 & c_1 & c_2 & c_1 & c_0 \end{pmatrix}.$$

Tämän jälkeen yhtälö  $f_0 = \langle v, A_0 v \rangle$  kiinnittää 4 parametria, mistä saadaan  $A_0 =$

$$\begin{pmatrix} 1 & 0 & 0 & -\frac{c_4}{2} & c_3 & -\frac{c_4}{2} \\ 0 & c_4 & -\frac{1}{2} - c_3 & 0 & -\frac{1}{2} - c_3 & c_3 \\ 0 & -\frac{1}{2} - c_3 & c_4 & c_3 & -\frac{1}{2} - c_3 & 0 \\ -\frac{c_4}{2} & 0 & c_3 & 1 & 0 & -\frac{c_4}{2} \\ c_3 & -\frac{1}{2} - c_3 & -\frac{1}{2} - c_3 & 0 & c_4 & 0 \\ -\frac{c_4}{2} & c_3 & 0 & -\frac{c_4}{2} & 0 & 1 \end{pmatrix}.$$

Lopuksi valitaan lauseen 2 perusteella  $c_4 = 1$  ja  $c_3 = 0$ , jolloin  $A_0$  on psd:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 1 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 & 1 \end{pmatrix}.$$

Sitten lasketaan  $A_0 = LDL^T$ , ja havaitaan, että  $D$ :ssä on 4 nollassa poikkeavaa alkioita, mistä lopulta saadaan

$$\begin{aligned} f_0 &= q_1^2 + q_2^2 + \frac{3}{4}q_3^2 + \frac{3}{4}q_4^2, \\ q_1 &= x^2 - y^2/2 - z^2/2, \\ q_2 &= xy - xz/2 - yz/2, \\ q_3 &= x(y - z), \\ q_4 &= y^2 - z^2. \end{aligned}$$

Aivan samoin voidaan osoittaa, että esimerkiksi 3 ollut polynomi

$$f_1 = x^6 + y^6 + z^6 - xyz(x^3 + y^3 + z^3)$$

on sos-polynomi. Nyt tarvitaan monomivektori:

$$v = (x^3, x^2y, x^2z, xy^2, xyz, xz^2, y^3, y^2z, yz^2, z^3).$$

Yhtälön  $f_1 = \langle v, A_1 v \rangle$  matriisilla  $A_1$  on periaatteessa 55 vapaata parametria, mutta symmetrian perusteella parametrien määrä onkin vain 13. Yhtäsuuruus  $f_1 = \langle v, A_1 v \rangle$  kiinnittää 7 parametria. Loput 6 parametria valitaan taas lauseen 2 mukaisesti. Kun valitaan lävistäjän ulkopuolella kaikki mahdolliset parametrit nolliksi, niin huomataan, että ehto (6) on voimassa. Kun sitten lasketaan  $A_1 = LDL^T$ , niin nähdään, että  $D$ :n lävistäjällä on 8 nolasta poikkeavaa alkioita, joten  $f_1$  voidaan esittää kahdeksan polynomien summana. ☆

## Lopuksi

Sos-tekniikat eivät selvästikään sovellu käsin laskentaan, mutta yllä olevat laskut olivat helppoja sopivaa ohjelmaa käyttäen. Tarvittiin vain matriisien kertolaskua ja yksinkertaisten lineaaristen yhtälöitten ratkaisua. Lisäksi tietysti piti laskea LDL-hajotelma. Monissa sos-sovelluksissa on kuitenkin välttämätöntä käyttää numeerisia menetelmiä. Tähän on olemassa MATLABIIN tehty paketti `sostools` [5].

## Matriisit

Olkoot  $x = (x_1, \dots, x_n)$  ja  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$  vektoreita ja asetetaan

$$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n.$$

Neliömatriisi  $A$  on

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

Matriisi määrittää lineaarikuvauksen, kun määritellään matriisin ja vektorin kertolasku:

$$A: \mathbb{R}^n \rightarrow \mathbb{R}^n, \\ y = Ax \iff y_i = \sum_{j=1}^n a_{ij} x_j.$$

Matriisin  $A$  transpoosi  $A^T$  saadaan, kun indeksointia vaihdetaan: siis jos  $A = (a_{ij})$ , niin  $A^T = (a_{ji})$ . On helppo tarkistaa, että

$$\langle Ax, y \rangle = \langle x, A^T y \rangle. \quad (5)$$

Matriisien  $A$  ja  $B$  kertolasku määritellään seuraavasti:

$$C = BA \iff c_{ij} = \sum_{\ell=1}^n b_{i\ell} a_{\ell j}.$$

Huomaa, että yleensä  $AB \neq BA$ . Seuraavat matriisiluokat ovat tässä oleellisia:

1.  $A$  on lävistäjämatriisi, jos  $a_{ij} = 0$  kun  $i \neq j$ . Yksikökmatriisi  $I$  on sellainen lävistäjämatriisi, jossa lävistäjäalkiot ovat ykkösiä. Huomaa, että  $AI = IA = A$  kaikille  $A$ . Jos lävistäjäalkiot ovat  $d_1, \dots, d_n$ , niin voidaan merkitä  $D = \text{diag}(d_1, \dots, d_n)$ .
2.  $A$  on alakolmiomatriisi, jos  $a_{ij} = 0$ , kun  $i < j$ .
3.  $A$  on symmetrinen, jos  $A = A^T$ .
4. Symmetrinen matriisi  $A$  on psd (positiivisemidefiniitti), jos

$$\langle x, Ax \rangle \geq 0 \quad \text{kaikilla } x \in \mathbb{R}^n.$$

5.  $P$  on permutaatiomatriisi, jos jokaisella  $P$ :n rivillä ja sarakkeella on yksi alkio ykkönen ja muut alkiot nollia.<sup>13</sup>

Olkoon sitten  $e_j$  vektori, jonka  $j$ :s komponentti on 1 ja muut nollia; tällöin

$$\langle e_j, A e_j \rangle = a_{jj}.$$

Saadaan siis välttämätön ehto:

$$\text{jos } A \text{ on psd, niin } a_{jj} \geq 0 \text{ kaikilla } j.$$

Tämä ei kuitenkaan ole riittävä ehto. Tarkastelemalla  $2 \times 2$  tapausta saadaan kuitenkin lisää käteviä välttämättömiä ehtoja:

$$\text{Jos } A \text{ on psd, niin } a_{jj} a_{kk} - a_{jk}^2 \geq 0 \text{ kaikilla } j \text{ ja } k. \\ \text{Erityisesti jos } a_{kk} = 0, \text{ niin } a_{jk} = 0 \text{ kaikilla } j.$$

Oletetaan sitten, että seuraava ehto pätee kaikilla  $j$ :

$$a_{jj} \geq \sum_{i \neq j} |a_{ij}|. \quad (6)$$

Tällöin on tapana sanoa, että  $A$  on *lävistäjävaltainen* (diagonally dominant). Voidaan osoittaa:

**Lause 2.**  $A$  on psd, jos ehto (6) pätee.

<sup>13</sup>Shakin pelaajat huomaavat, että jokainen  $8 \times 8$  permutaatiomatriisi vastaa yhtä *torniongelman* ratkaisua. Torniongelma on siis  $8! = 40320$  ratkaisua. Harjoitustehtäväksi jätän sen pohtimisen, kuinka monta oleellisesti erilaista ratkaisua on, ja miten oleellinen erilaisuus pitäisi määritellä.

Tämä ehto ei puolestaan ole välttämätön. Eräs tunnettu esimerkki on Hilbertin matriisi  $H$ , jonka alkiot ovat  $h_{ij} = 1/(i+j-1)$ . Baltian tie -matematiikkakilpailuissa vuonna 1990 piti osoittaa, että  $H$  on psd.<sup>14</sup> Olkoot

$$g = c_0 + c_1x + \dots + c_nx^n$$

ja  $c = (c_0, \dots, c_n)$ . Nyt selvästi

$$\langle c, Hc \rangle = \int_0^1 g^2 dx \geq 0.$$

Ehto (6) on silti erittäin hyödyllinen, koska se on niin helppo tarkistaa.

Tarvitaan vielä

**Määritelmä 1.** *Olkoon  $A$  on symmetrinen.  $A$ :n LDL-hajotelma on*

$$A = LDL^T,$$

missä  $D$  on lävistäjämatriisi ja  $L$  on alakolmiomatriisi, ja lisäksi  $l_{ii} = 1$  kaikilla  $i$ .

Kaikilla symmetrisillä matriiseilla ei ihan ole juuri tällaista LDL-hajotelmaa. Yleisesti ottaen voidaan osoittaa:

**Lause 3.** *Jos  $A$  on symmetrinen, niin on olemassa  $P$ ,  $L$  ja  $D$  siten, että*

$$P^TAP = LDL^T,$$

missä  $P$  on permutaatiomatriisi ja  $L$  on alakolmiomatriisi, ja lisäksi  $l_{ii} = 1$  kaikilla  $i$ .  $D$  on lohkolävistäjä-matriisi, jossa lohkot ovat joko  $1 \times 1$  tai  $2 \times 2$ .

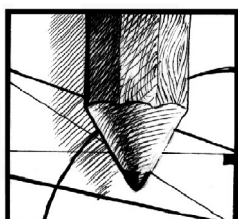
Tarkka muotoilu löytyy Highamin kirjasta [4]. Vaikka kirjan nimessä on *numerical algorithms*, niin LDL-hajotelma on (myös) symbolinen algoritmi siinä mielessä, että jos  $A$ :n alkiot ovat rationaalilukuja, niin myös  $L$ :n ja  $D$ :n alkiot ovat rationaalilukuja.

Tätä yleistystä ei tässä kirjoituksessa tarvita, vaan kaikissa esimerkeissä  $D$  on lävistäjämatriisi ja  $P$  on yksikkömatriisi.

## Viitteet

- [1] G. V. Bard, *Sage for undergraduates*, 2nd ed., American Mathematical Society, 2022.
- [2] G. Blekherman, P. A. Parrilo, and R. R. Thomas (eds.), *Semidefinite optimization and convex algebraic geometry*, MOS-SIAM Series on Optimization, vol. 13, SIAM, 2013.
- [3] D. A. Cox, J. Little, and D. O’Shea, *Ideals, varieties, and algorithms*, 4th ed., Undergraduate Texts in Mathematics, Springer, Cham, 2015, An introduction to computational algebraic geometry and commutative algebra.
- [4] N. Higham, *Accuracy and stability of numerical algorithms*, 2nd ed., SIAM, 2002.
- [5] A. Papachristodoulou, J. Anderson, G. Valmorbidia, S. Prajna, P. Seiler, and P. A. Parrilo, *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*, <http://arxiv.org/abs/1310.4716>, 2013, Available from <http://www.eng.ox.ac.uk/control/sostools>, <http://www.cds.caltech.edu/sostools> and <http://www.mit.edu/~parrilo/sostools>.
- [6] J. Tuomela, *Tarinoita polynomeista (osa 1)*, Solmu (2022), no. 2.
- [7] ———, *Tarinoita polynomeista (osa 2)*, Solmu (2022), no. 3.
- [8] ———, *Tarinoita polynomeista (osa 3)*, Solmu (2023), no. 2.

<sup>14</sup>Hilbertin matriisi on erikoistapaus Gramin matriisista, jotka ovat aina psd. [https://en.wikipedia.org/wiki/Gram\\_matrix](https://en.wikipedia.org/wiki/Gram_matrix)

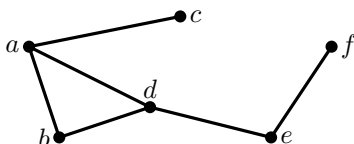


## Rekonstruktiokonjektuuri verkkteoriassa

*Milo Orlich*

Matematiikan ja systeemianalyysin laitos, Aalto-yliopisto  
milo.orlich@aalto.fi

Vapaasti kuvailtuna *verkko* (tai *graafi*) on jotakin, joka näyttää tältä:



Verkko koostuu siis *solmuista* (“paksuista pisteistä”) ja *kaarista* (“viivoista”). Kaksi solmua voi liittyä toisiinsa kaarella tai ei. Joissakin tilanteissa verkossa voi olla “silmukka”, eli jokin solmu voi yhdistyä itseensä, tai kahden solmun välillä voi olla enemmän kuin yksi kaari, tai jokaisella kaarella voi olla suunta, esimerkiksi näin:

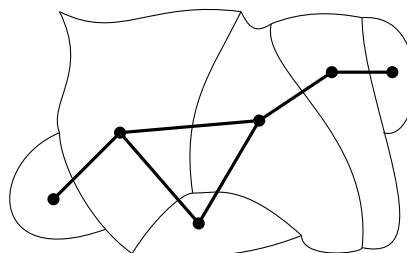


Emme ota huomioon näitä mahdollisuuksia tässä kirjoituksessa.

**Esimerkki 1.** Koska verkon käsite on niin yksinkertainen ja abstrakti, sitä käytetään monella matematiikan alalla ja muissa tieteissä. Verkkojen avulla voi mallintaa monimutkaisilta näyttäviä tosielämän tilanteita:

1. Facebookin käyttäjät ovat ison verkon solmut, ja kaksi käyttäjää yhdistyy kaarella, jos he ovat Facebook-kavereita.

2. Kaikki maailman asukkaat ovat vielä isomman verkon solmut, ja kaksi ihmistä yhdistyy kaarella, jos he ovat jossakin vaiheessa kätelleet.
3. Kartassa olevat valtiot ovat verkon solmut, ja kaksi valtiota yhdistyy kaarella, jos niillä on yhteistä rajaa:



Seuraavaksi esitetään verkon määritelmä.

**Määritelmä 2.** *Verkko*  $V$  on järjestetty pari  $(S_V, K_V)$ , jossa  $S_V$  on äärellinen joukko, jonka alkioita kutsutaan verkon  $V$  *solmuiksi*, ja  $K_V$  on joukko joukkoja, jotka koostuvat tarkalleen kahdesta solmusta ja joita kutsutaan verkon  $V$  *kaariksi*.

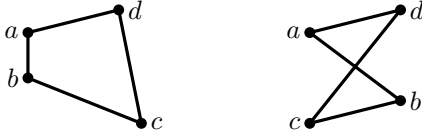
**Esimerkki 3.** Jos  $V$  on ensimmäisessä kuvassa oleva verkko, niin

$$S_V = \{a, b, c, d, e, f\},$$

$$K_V = \{\{a, b\}, \{a, c\}, \{a, d\}, \{b, d\}, \{d, e\}, \{e, f\}\}.$$



Tämä vaivalloinen verkon määritelmä voi vaikuttaa hankalalta ja epäselvältä, mutta sen avulla suurien verkkojen tutkiminen on mahdollista tietokoneella, ja se on täsmällisempi kuin “piirustus” matemaatikon näkökulmasta. On tärkeää huomata, että piirustus on vain visuaalinen apu. Esimerkiksi nämä kaksi piirustusta esittävät samaa verkkoa:



Syy on se, että “kunnon” verkko  $V$  on pari  $(S_V, K_V)$ .

*Viitteet:* Verkoista kirjoitettiin mm. Solmun artikkeleissa [2] ja [3]. Hyvä kirja on esimerkiksi [1], jota käytetään Aalto-yliopiston verkkoteorian kurssilla oppikirjana.

Kun tutustuu uuteen matemaattiseen käsitteeseen, voi olla hyödyllistä ajatella sitä kombinatorisesti, eli kysyä “*Kuinka monta ... ?*” -kysymyksiä.

**Esimerkki 4.** 1. On olemassa vain yksi verkko, jolla on tarkalleen yksi solmu  $a$ :

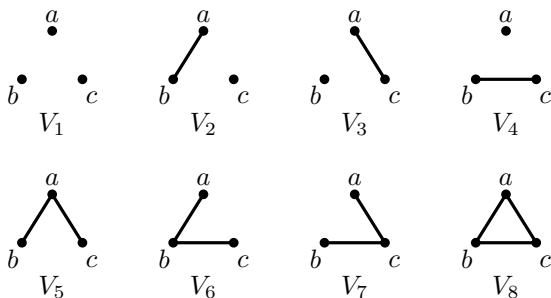


Tällä verkolla ei ole kaarta. (Mutta on olemassa äärettömän monta yksisolmuista verkkoa, koska on olemassa äärettömän monta yksiötä! Solmujoukko voi olla  $\{a\}$ ,  $\{b\}$ ,  $\{c\}$ , ...,  $\{1\}$ ,  $\{2\}$ ,  $\{3\}$ , ...,  $\{\text{Helsinki}\}$ ,  $\{\text{Uranus}\}$ ,  $\{\text{Teräsmies}\}$ , jne.)

2. On olemassa kaksi verkkoa, joiden solmujoukko on  $S = \{a, b\}$ ; toisella ei ole kaarta, ja toisella on tarkalleen yksi kaari:



3. On olemassa kahdeksan verkkoa, joiden solmujoukko on  $S = \{a, b, c\}$ :



4. Entä, jos  $S = \{a, b, c, d\}$ ?

**Tehtävä 5.** Kuinka monta kaarta voi korkeintaan olla verkossa, jossa on  $n$  solmua?

**Tehtävä 6.** Kuinka monta erilaista verkkoa on olemassa, jos solmujoukko on  $S = \{s_1, \dots, s_n\}$ ?

## Isomorfiset verkot

Kaksi verkkoa voi “näyttää samalta”, jos solmujen nimet unohdetaan. Esimerkin 4.3 verkot  $V_2$ ,  $V_3$  ja  $V_4$  ovat samannäköisiä: jokaisella on yksi kaari ja erillinen solmu. Verkot  $V_5$ ,  $V_6$  ja  $V_7$  näyttävät kaikki  $V$ -kirjaimelta. Kuten jo huomattiin, on itse asiassa olemassa *äärettömän monta* verkkoa, jotka näyttävät samalta, kun solmujen nimet unohdetaan!

Tämän koko tarinan tarkka nimi on verkkojen *isomorfismi*. (Isomorfismin täsmällinen määritelmä on alla. Muuten sana “isomorfismi” tulee kreikasta: “isomorfinen” on kirjaimellisesti “samanmuotoinen”.)

**Esimerkki 7.** 1. Kaikki yksisolmuiset verkot ovat keskenään isomorfisia. Toisin sanoen, jos solmun nimi unohdetaan, niin ainoa yksisolmuinen verkko näyttää tältä:



2. Jokainen kaksisolmuinen verkko on isomorfinen joko ensimmäisen tai toisen seuraavassa kuvassa olevan verkon kanssa:



Jos siis solmujen nimet unohdetaan, on olemassa vain kaksi kaksisolmuista verkkoa.

3. Jos solmujen nimet unohdetaan, on olemassa vain neljä kolmisolmuista verkkoa:



**Tehtävä 8.** Kun solmujen nimet unohdetaan, kuinka monta nelisolmuista verkkoa on olemassa? (Voit myös tutkia tapausta, jossa solmujoukko koostuu viidestä solmusta, mutta vastaus on  $> 30$ .)

Kuten verkon käsite itse, verkkojen isomorfismi on melko intuitiivinen asia. Se on valitettavasti vielä liian epämääräinen: tietokone ei voisi tehdä mitään “epämääräisen, mutta intuitiivisen” idean avulla, ja verkkojen piirtämisestä tulee nopeasti liian vaikeaa ihmisille, joten tarvitaan tietokonetta ja täsmällistä määritelmää.

Muista, että *bijektio*  $\varphi: A \rightarrow B$ , missä  $A$  ja  $B$  ovat joukkoja, on sekä injektio että surjektio, eli bijektio on sellainen funktio, että:

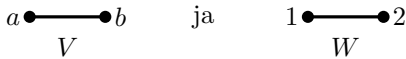
- jos  $a \neq a'$ , niin  $\varphi(a) \neq \varphi(a')$ ,
- jokaisella alkiolla  $b \in B$  on olemassa sellainen  $a \in A$ , että  $\varphi(a) = b$ .

**Määritelmä 9.** Verkkoja  $V$  ja  $W$  kutsutaan *isomorfeisiksi*, jos on olemassa sellainen bijektio  $\varphi: S_V \rightarrow S_W$ , että kaikilla solmuilla  $s, s' \in S_V$  pätee

$$\{s, s'\} \in K_V \Leftrightarrow \{\varphi(s), \varphi(s')\} \in K_W.$$

Tällaista bijektiota kutsutaan verkkojen  $V$  ja  $W$  *isomorfismiksi*. Jos isomorfismi on olemassa, niin kirjoitetaan  $V \cong W$ . Jos ei, niin kirjoitetaan  $V \not\cong W$ . Joukko, joka koostuu kaikista verkon  $V$  kanssa isomorfeisista verkoista, on verkon  $V$  *isomorfialuokka*.

**Esimerkki 10.** Verkoille



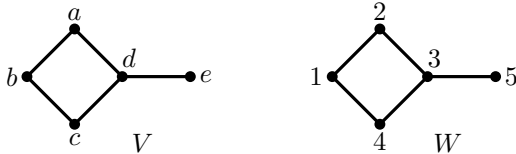
on olemassa kaksi erilaista isomorfismia:

$$\begin{array}{l} a \leftrightarrow 1 \\ b \leftrightarrow 2 \end{array} \quad \text{ja} \quad \begin{array}{l} a \leftrightarrow 2 \\ b \leftrightarrow 1. \end{array}$$

**Tehtävä 11.** Kirjoita eksplisiittisesti edellisen esimerkin tapaan isomorfismi

$$\begin{array}{l} a \leftrightarrow \varphi(a), \quad b \leftrightarrow \varphi(b), \quad c \leftrightarrow \varphi(c), \\ d \leftrightarrow \varphi(d), \quad e \leftrightarrow \varphi(e) \end{array}$$

seuraaville kahdelle verkolle:



Kuinka monta erilaista isomorfismia on olemassa?

Entä, jos otetaan verkot



niin kuinka monta erilaista isomorfismia on olemassa?

## Invariantit

Vaikka verkkojen isomorfismi määritellään huolellisesti, ongelma on se, että on vaikeaa saada selville, onko olemassa kahden annetun verkon isomorfismi vai ei. Joku voisi periaatteessa testata jokaisen mahdollisen bijektion ja tarkistaa, onko se isomorfismi, mutta tämä on todella hidas algoritmi. Siksi verkkoteoreetikot tutkivat mahdollisesti auttavia "invariantteja". Verkon *invariantti* on suure tai ominaisuus  $\varphi$ , joka on sama isomorfeisille verkoille. Eli jos  $V \cong W$ , niin  $\varphi(V) = \varphi(W)$ . Esimerkiksi solmujen tai kaarten määrät ovat invariantteja:

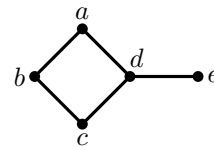
Jos verkot  $V$  ja  $W$  ovat isomorfeisia, niin niillä on sama solmujen määrä ja sama kaarten määrä.

Päinvastainen lause ei pidä paikkaansa:

**Tehtävä 12.** Etsi kaksi sellaista verkkoa  $V$  ja  $W$ , joilla on sama solmujen määrä ja sama kaarten määrä, mutta  $V$  ja  $W$  eivät ole isomorfeisia.

**Määritelmä 13.** Olkoon  $s$  verkon  $V$  solmu. Solmun  $s$  *asteluku*  $al(s)$  on solmuun  $s$  liittyvien solmujen lukumäärä. Verkon *astelukujen jono* koostuu kaikista asteluvuista, suurimmasta pienimpään.

**Esimerkki 14.** Alla olevassa verkossa,  $al(d) = 3$ ,  $al(a) = al(b) = al(c) = 2$  ja  $al(e) = 1$ .



Astelukujen jono on siten  $(3, 2, 2, 2, 1)$ .

Huomaa, että:

- asteluvuista voi laskea kaarten lukumäärän:

$$\frac{1}{2} \sum_{s \in S_V} al(s) = |K_V|,$$

- jos  $V \cong W$ , niin verkkojen  $V$  ja  $W$  astelukujen jonot ovat samat.

**Tehtävä 15.** Etsi kaksi verkkoa, joilla on sama solmujen määrä ja sama kaarten määrä, mutta eri astelukujen jonot.

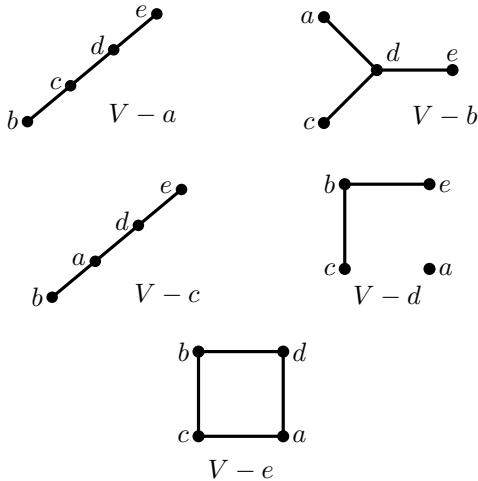
**Tehtävä 16.** Etsi sellaiset verkot  $V$  ja  $W$ , joiden astelukujen jonot ovat samat, mutta  $V \not\cong W$ .

## Verkon pakka ja konjektuuri

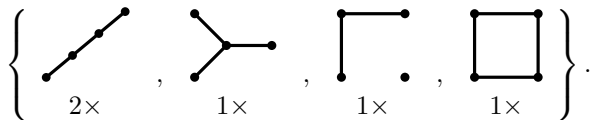
*Multijoukko* on sellainen joukon yleistys, jossa jokaisella alkioilla on oma "toistoluku", joka voi olla suurempi kuin 1. Jokainen alkio voi siis esiintyä multijoukossa enemmän kuin yhden kerran.

**Määritelmä 17.** Jokaiselle solmulle  $s_i \in S_V = \{s_1, \dots, s_n\}$  määritellään verkko  $V - s_i$  poistamalla solmu  $s_i$  ja kaikki solmuun  $s_i$  liittyvät kaaret. Multijoukko, joka koostuu verkkojen  $V - s_1, \dots, V - s_n$  isomorfialuokista, on verkon  $V$  *pakka*, ja sitä merkitään symbolilla  $P(V)$ . Pakan alkiota kutsutaan *korteiksi*.

**Esimerkki 18.** Jos  $V$  on esimerkissä 14 oleva verkko, niin saadaan



Siten verkon  $V$  pakka  $P(V)$  on multijoukko



On tärkeää muistaa, että jokaiselle verkolle  $V - s_i$  laitamme pakkaan vain sen isomorfialuokan, eli emme tiedä, minkä solmun poistamalla saimme tämän kortin.

Verkon pakka muistaa kuitenkin joitakin alkuperäisen verkon tietoja. Esimerkiksi verkon  $V$  solmujen määrä on sama kuin korttien määrä, tai myös jokaisen kortin solmujen määrä plus 1.

**Tehtävä 19.** Todista, että:

1. jos  $k_i$  on verkon  $V - s_i$  kaarten määrä, niin

$$|K_V| = \frac{1}{n-2} \sum_{i=1}^n k_i,$$

2. verkon pakasta voi laskea sen astelukujen jonon,
3.  $V \cong W \Rightarrow P(V) = P(W)$ , eli pakka on verkon invariantti.

Seuraavaksi esitetään vanha konjektuuri, eli avoin ongelma, jota verkkoteoreetikot ovat yrittäneet todistaa jo 70 vuotta!

**Konjektuuri 20** (Kelly, Ulam). Jos  $n \geq 3$ , niin  $P(V) = P(W) \Rightarrow V \cong W$ .

Sanallisesti ilmaistuna konjektuuri sanoo, että pakka määrää verkon isomorfialuokan: ei ole olemassa epäisomorfisia verkkoja, joilla on sama pakka. Tämä on siksi tunnettu myös ”rekonstruktiokonjektuurina”.

**Esimerkki 21.** Osoitetaan eksplisiittisesti, että konjektuurissa oleva implikaatio pätee, jos  $n = 3$ . Kaikki mahdolliset kolmisolmuisten verkkojen pakat ovat

$$P(\cdot \cdot \cdot) = \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \\ 3 \times \end{array} \right\}, \quad P(\cdot \cdot \cdot) = \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \\ 2 \times \end{array}, \begin{array}{c} \cdot \\ \cdot \\ \cdot \\ 1 \times \end{array} \right\},$$

$$P(\Delta) = \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \\ 3 \times \end{array} \right\}, \quad P(\wedge) = \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \\ 1 \times \end{array}, \begin{array}{c} \cdot \\ \cdot \\ \cdot \\ 2 \times \end{array} \right\}.$$

Selvästi pätee: jos  $V \not\cong W$ , niin  $P(V) \neq P(W)$ .

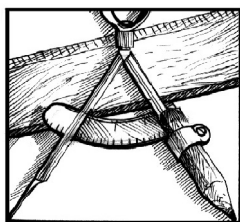
**Tehtävä 22.** Tarkista samalla tavalla, että konjektuurissa oleva implikaatio pätee, jos  $n = 4$ , mutta ei päde, jos  $n = 2$ .

Implikaation pätevyys on tarkistettu tietokoneella, kun  $3 \leq n \leq 13$ . Konjektuuri on todistettu myös muutamassa erikoistapauksessa, eli kun  $V$  ja  $W$  kuuluvat joihinkin tiettyihin verkkoperheisiin. Jos esimerkiksi  $V$  ja  $W$  ovat niin sanottuja ”puita”, tai jos jokaisella solmulla on sama asteluku, niin konjektuurin implikaatio pitää paikkansa. Aiheesta voi lukea lisää verkkosivulta [4].

*Suuri kiitos arvokkaista kommentteista Pekka Alestalolle, Tuomas Kelomäelle ja Teemu Lundströmille!*

## Viitteet

- [1] R. Diestel, *Graph Theory*, Springer-Verlag, Heidelberg, Graduate Texts in Mathematics, Volume 173.
- [2] V. Romanov, *Verkkojen teoriaa Königsbergistä internetiin (Solmu 3/2016)*, [https://matematiikkalehtisolmu.fi/2016/3/verkkojen\\_teoraaa.pdf](https://matematiikkalehtisolmu.fi/2016/3/verkkojen_teoraaa.pdf)
- [3] E. V. Vesalainen, *Tasograafit ja väritykset (Solmu 1/2016)*, <https://matematiikkalehtisolmu.fi/2016/1/varityksia.pdf>
- [4] Wikipedia: Reconstruction conjecture [https://en.wikipedia.org/wiki/Reconstruction\\_conjecture](https://en.wikipedia.org/wiki/Reconstruction_conjecture)



## Solmun tehtäviä

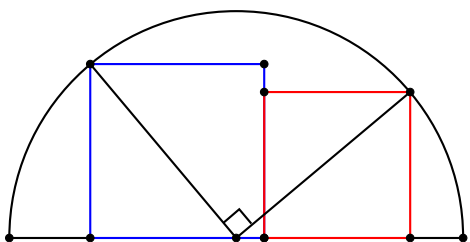
Tehtävät sopivat lukiolaisten lisäksi myös edistyneille yläkoululaisille.

1. Ratkaise yhtälö

$$\sqrt{7 + \frac{3}{\sqrt{x}}} = 7 - \frac{9}{x},$$

kun  $x$  on reaaliluku.

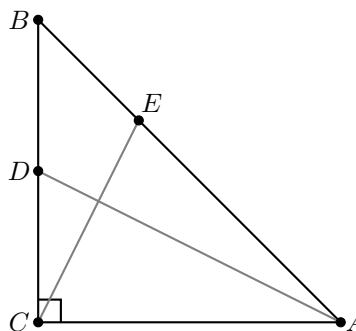
2. Yksikköpuoliympyrän (eli puoliympyrän, jonka säde on 1) sisään piirretään kaksi neliötä siten, että molempien neliöiden yksi sivu on puoliympyrän halkaisijalla ja yksi kärki on puoliympyrällä sekä yksi kummankin neliön sivuista sijaitsee samalla suoralla (katso alla oleva kuva). Ympyrän keskipisteestä neliöiden puoliympyrällä sijaitseviin kärkiin piirrettyjen janojen välinen kulma on suora. Osoita, että näin piirrettyjen neliöiden pinta-alojen summa on aina sama.



3. Ravintolassa tarjottava keitto osoittautui liian suolaiseksi. Kokki päätti laimentaa keittoa edellisenä päivänä yli jääneellä keitolla, jossa ei ollut tarpeeksi suolaa. Suolaisen keiton suolapitoisuus on 5 prosenttia ja edellisen päivän keiton suolapitoisuus on 1,2 prosenttia.

Kuinka paljon kumpaakin keittoa tulee lisätä, jotta sekoitettua keittoa saadaan 72 desilitraa ja sen suolapitoisuus on 3,48 prosenttia?

4. Suorakulmaisessa tasakylkisessä kolmiossa  $ABC$  kateetin  $BC$  keskipiste on  $D$  ja hypotenuusalla  $AB$  lähempänä kärkeä  $B$  oleva piste  $E$  jakaa hypotenuusan suhteessa 1 : 2. Osoita, että janat  $AD$  ja  $CE$  ovat kohtisuorassa toisiaan vastaan.



5. Tarkastellaan funktioita

$$f_0(x) = \frac{1}{1-x} \quad \text{ja} \quad f_n(x) = f_0(f_{n-1}(x)),$$

kun  $n$  on positiivinen kokonaisluku. Laske  $f_{2023}(2023)$ .

6. Merkintä  $[x]$  tarkoittaa *lattiafunktiota* reaaliluvusta  $x$ , jonka arvo on suurin kokonaisluku, joka ei ole suurempi kuin  $x$ , siis

$$[x] = \max\{n \in \mathbb{Z} \mid n \leq x\}.$$

Esimerkiksi  $[2,5] = 2$ ,  $[-0,2] = -1$  ja  $[10] = 10$ .

Tarkastellaan funktioita

$$\begin{aligned} f(x) &= \sqrt{x+5}, \\ g(x) &= \frac{-2x+8}{5}, \\ h(x) &= \lfloor x+3 \rfloor. \end{aligned}$$

Määritä funktioiden kuvaajien yhteiset pisteet.

7. Olkoot  $a > 2$ ,  $b$  ja  $c$  reaalilukuja. Tarkastellaan seuraavia kolmea väittämää.

- (1) Yhtälöllä  $ax^2 + bx + c = 0$  ei ole reaalisia ratkaisuja.
- (2) Yhtälöllä  $(a-1)x^2 + (b-1)x + (c-1) = 0$  on yksi reaalinen ratkaisu.
- (3) Yhtälöllä  $(a-2)x^2 + (b-2)x + (c-2) = 0$  on kaksi reaalista ratkaisua.

(a) Jos väittämät (1) ja (2) ovat totta, niin onko myös väittämä (3) totta?

(b) Jos väittämät (2) ja (3) ovat totta, niin onko myös väittämä (1) totta?

8. Onko totta, että jos kolmion jokaisen kulman sini on rationaalinen, niin myös jokaisen kulman kosini on rationaalinen?

9. Osoita, että jos  $0 < a, b < 1$ , niin

$$(a+b-ab)(a^b + b^a) > a+b.$$

10.  $5 \times 5$ -ruudukko täytetään satunnaisessa järjestyksessä kirjoittamalla jokaiseen ruutuun luku. Ruutuun kirjoitettava luku määräytyy sen mukaan, kuinka moni vieressä olevassa ruutu on jo täytetty. Vierekkäin olevia ruutuja ovat ne, joilla on yhteinen sivu. Esimerkiksi alla oleva ruudukko on täytetty luvuilla seuraavassa järjestyksessä: a5, b5, c5, d5, e5, e4, e3, e2, a4, a3, a2, a1, b1, c1, d1, e1, b4, c4, b3, b2, c2, d2, d3, d4, c3.

5	0	1	1	1	1
4	1	2	2	4	1
3	1	2	4	2	1
2	1	3	2	3	2
1	1	1	1	1	2
	a	b	c	d	e

Täytä ruudukko kahdella muulla tavalla sääntöjä noudattaen. Osoita, että ruudukkoon täytettyjen lukujen summa on aina 40 riippumatta siitä, missä järjestyksessä ruudukko on sääntöjä noudattaen täytetty.

Tehtävien ratkaisut julkaistaan Solmun seuraavassa numerossa.

Lähde: KöMaL

Käännös ja sovitus suomeksi: Mika Koskenoja



## Solmu 3/2023

ISSN-L 1458-8048

ISSN 1459-0395 (Painettu)

ISSN 1458-8048 (Verkkolehti)

Julkaisija:

*Suomen matemaattinen yhdistys ry*  
*PL 68 (Pietari Kalmin katu 5)*  
*00014 Helsingin yliopisto*

Päätoimittaja:

*Anne-Maria Ernvall-Hytönen*, apulaisprofessori, Matematiikan ja tilastotieteen osasto, Helsingin yliopisto

Toimitussihteeri:

*Juha Ruokolainen*, FT

Sähköposti:

*toimitus@matematiikkalehtisolmu.fi*

Verkkosivu:

*matematiikkalehtisolmu.fi*

Toimittajat:

*Pekka Alestalo*, dosentti, Matematiikan ja systeemianalyysin laitos, Aalto-yliopisto  
*Sirkka-Liisa Eriksson*, dosentti, Matematiikan ja tilastotieteen osasto, Helsingin yliopisto  
*Aapo Halko*, FT, Matematiikan ja tilastotieteen osasto, Helsingin yliopisto  
*Olli Järvinen*, jatko-opiskelija, Matematiikan ja tilastotieteen laitos, Turun yliopisto  
*Jyrki Lahtonen*, yliopistonlehtori, Matematiikan ja tilastotieteen laitos, Turun yliopisto  
*Heikki Pokela*, tuntiopettaja, Tapiolan lukio  
*Antti Rasila*, Associate Professor, Guangdong Technion - Israel Institute of Technology  
*Mikko Sillanpää*, professori, Matemaattisten tieteiden laitos ja Biologian laitos, Oulun yliopisto  
*Samuli Siltanen*, professori, Matematiikan ja tilastotieteen osasto, Helsingin yliopisto  
*Kimmo Vehkalahti*, vanhempi yliopistonlehtori, Yhteiskuntatieteiden keskus, Helsingin yliopisto

Tieteelliset asiantuntijat:

*Heikki Apiola*, dosentti, Matematiikan ja systeemianalyysin laitos, Aalto-yliopisto  
*Mika Koskenoja*, dosentti, Matematiikan ja tilastotieteen osasto, Helsingin yliopisto  
*Liisa Näveri*, FT, Opettajankoulutuslaitos, Helsingin yliopisto

Matematiikkadiplomit:

*Juha Ruokolainen*, juha piste ruokolainen 'at' yahoo piste com  
*Marjatta Näätänen*, dosentti, Matematiikan ja tilastotieteen osasto, Helsingin yliopisto

Yliopistojen ja korkeakoulujen yhteyshenkilöt:

*Ari Koistinen*, FM, Metropolia Ammattikorkeakoulu  
*Juha Lehtinen*, yliopistonlehtori, Matematiikan ja tilastotieteen laitos, Jyväskylän yliopisto  
*Jorma Merikoski*, emeritusprofessori, Tietotekniikan yksikkö, Tampereen yliopisto  
*Antti Viholainen*, yliopistonlehtori, Fysiikan ja matematiikan laitos, Itä-Suomen yliopisto

Kansikuva:

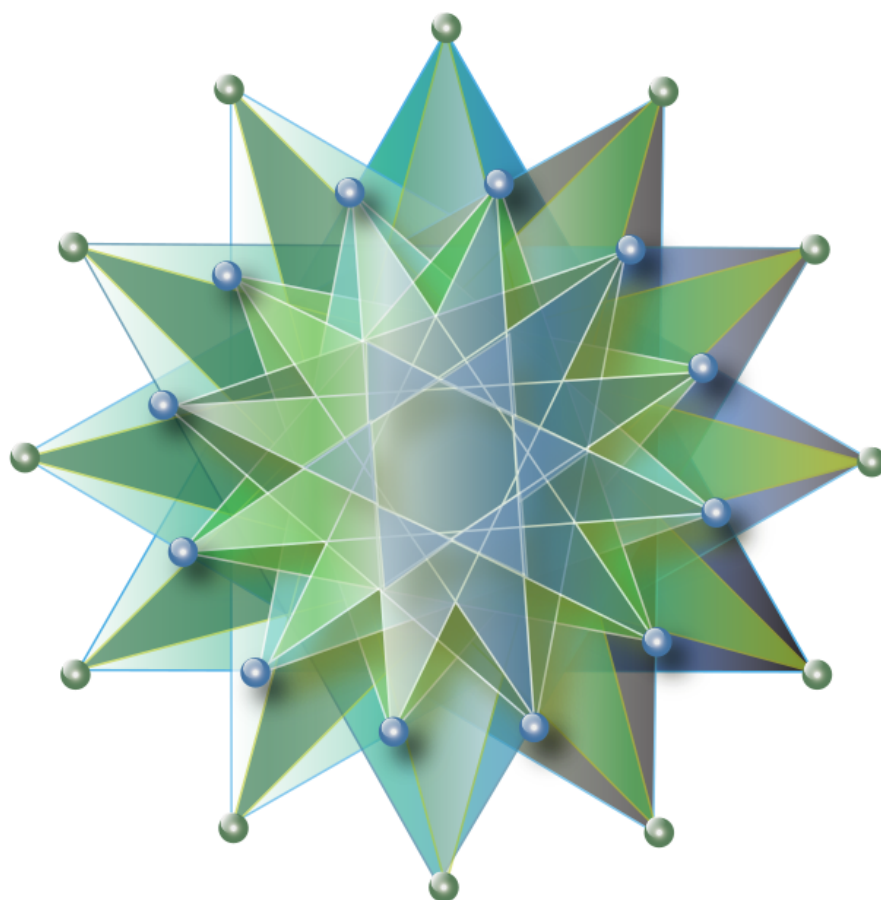
*Noora Isoeskelä*. Kuva on tehty tekoälyä apuna käyttäen ja sen tuottamia kuvia muokaten.

Painopaikka:

*Painosalama Oy*

Numeroon 1/2024 tarkoitetut kirjoitukset pyydämme lähettämään 3.3.2024 mennessä.

Kiitämme taloudellisesta tuesta Jenny ja Antti Wihurin rahastoa.



# MATEMATIIKAN VERKKOSANAKIRJA

MATEMATIIKKALEHTISOLMU.FI