1 Kannat ja kannanvaihto

1.1 Koordinaattivektori

Oletetaan, että V on K-vektoriavaruus, jolla on kanta $S = (v_1, v_2, \ldots, v_n)$. Avaruuden V vektori v voidaan kirjoittaa kannan vektorien lineaarikombinaationa: $v = \sum_{i=1}^{n} a_i v_i$ joillakin $a_i \in K$. Vektorin v ilmaisemiseksi riittää siis tietää kertoimien a_i arvot, ja tieto voidaan tiivistää sarakevektoriin

$$[v]_S = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Huomaa, että kertoimet a_i määräytyvät aina valitun kannan mukaan. Näin saatua sarakevektoria kutsutaan v:n *koordinaattivektoriksi* kannan S suhteen.

Koordinaattivektorit voidaan kirjoittaa sarakevektorien sijasta myös rivivektoreina. Tällöin kaikissa tuloksissa pitää kuitenkin muuttaa sarakkeet riveiksi sekä tehdä muita pieniä muutoksia.

Koordinaattivektori ei kerro yhtään mitään, ellei tiedetä, minkä kannan suhteen se on kirjoitettu. Ilmiötä voi verrata eri kiellillä kirjoitettuihin sanoihin. Jos kirjoitetaan “helmet”, on sanalla eri merkitys sen mukaan, oletetaanko kielen olevan suomi vai englanti. Samalla tavalla vaikkapa koordinaattivektori $[1, 2, 3]^T$ voi tarkoittaa esimerkiksi vektoria $1 \cdot (1, 0, 0) + 2 \cdot (1, 0, 0) + 3 \cdot (1, 0, 0) = (1, 2, 3)$ tai vektoria $1 \cdot (1, -1, 0) + 2 \cdot (0, 1, -1) + 3 \cdot (0, 0, 1) = (1, 1, 2)$ sen mukaan, mikä kanta on käytössä. Jos taas ajatellaan vektoriavaruutta

$$P_2 = \{p \mid p \text{ on reaalikertoiminen polynomi, } \deg p \leq 2\}$$

saadaan aivan erilaisia vektoreita. Jos kannaksi valitaan polynomijono $(1, t, t^2)$, vastaa koordinaattivektoria $[1, 2, 3]^T$ vektori $1 + 2t + 3t^2$.

Voidaan ajatella, että kannan S valinnan jälkeen saadaan kuvaus $[\ ]_S$ avaruudesta V sarakeavaruuteen K^n:

$$[\ ]_S \colon V \to K^n, \quad v \mapsto [v]_S$$

Tämä kuvaus on isomorfismi. Jos taas valitaan jokin toinen kanta, saadaan jokin toinen isomorfismi.

Lineaarialgebran käännöskoneina toimivat *kannanvaihtomatriisit*. Jos kielten välillä toimiva käännöskone muuttaa vaikkapa sanan “kukka” sanaksi “flower”, muuttaa kannanvaihtomatriisi yhden kannan suhteen kirjoitetun koordinaattivektorin toisen kannan suhteen kirjoitetuksi koordinaattivektoriksi.

Olkoot S ja T avaruuden V kantoja. Voidaan osoittaa, että tällöin on olemassa kääntyvä matriisi $P_{S,T}$, jolle pätee

$$P_{S,T}[v]_S = [v]_T \quad \text{kaikilla } v \in V$$

Matriisia $P_{S,T}$ kutsutaan kannanvaihtomatriisiksi ja se on aina yksikäsitteinen.

1.2 Kannanvaihto avaruudessa K^n

Tutkitaan vektoriavaruutta K^n ja sen kahta kantaa

$$S = (v_1, v_2, \ldots, v_n) \qquad \text{ja} \qquad T = (w_1, w_2, \ldots, w_n).$$

Etsitään kannanvaihtomatriisi näiden kahden kannan välille käyttäen apuna luonnollista kantaa E.

Määritetään ensin kannanvaihtomatriisi $P_{S,E}$ kannasta S kantaan E. Matriisin $P_{S,E}$ täytyy siis toteuttaa ehto

$$P_{S,E}[v]_S = [v]_E \quad \text{kaikilla } v \in K^n. \tag{1.1}$$

Matriisi $P_{S,E}$ voidaan määrittää kantavektoreita tutkimalla. Jos ehto nimittäin toteutuu kantavektoreilla, se toteutuu kaikilla muillakin vektoreilla. Muut vektorit ovat näet kantavektoreiden lineaarikombinaatioita.

Huomataan, että kannan S vektorit näyttävät kannan S suhteen kovin siisteiltä: $[v_i]_S = [0\ 0\ \ldots\ 0\ 1\ 0 \ldots\ 0\ 0]^T$, missä luku 1 on kohdassa i. Kun tällaisilla vektoreilla kertoo jotakin matriisia, lähtevät matriisista mukaan sarakkeen i alkiot. Jotta ehto (1.1) toteutuisi, on matriisin $P_{S,E}$ sarakkeessa i siis oltava koordinaattivektorin $[v_i]_E$ alkiot:

$$P_{S,E} = [[v_1]_E\ [v_2]_E\ \ldots [v_n]_E].$$

Koordinaattivektori $[v_i]_E$ on puolestaan helppo määrittää, sillä sen alkiot ovat samat kuin vektorin v_i komponentit.

Kannanvaihtomatriisi voidaan aina muodostaa näin helposti mistä tahansa kannasta luonnolliseen kantaan. Matriisin sarakkeiden alkiot saadaan suoraan kantavektorien komponenteista.

Jos halutaan kannanvaihtomatriisi kannasta E kantaan S, käytetään matriisin $P_{S,E}$ käänteismatriisia. (Kannanvaihtomatriisilla on aina käänteismatriisi. Tämän todistaminen jätetään harjoitustehtäväksi.)

Samalla tavalla kuin edellä löydetään kannanvaihtomatriisi $P_{T,E}$ kannasta T kantaan E:

$$P_{T,E} = [[w_1]_E,\ [w_2]_E\ \dots [w_n]_E].$$

Matriisi $(P_{T,E})^{-1}$ puolestaan toimii kannanvaihtomatriisina kannasta E kantaan T.

Kannanvaihtomatriisi kannasta S kantaan T voidaan muodostaa edellä mainittujen matriisien avulla. Vaihdetaan ensin kannasta S kantaan E ja sitten kannasta E kantaan T. Kannanvaihtomatriisi on siten $(P_{T,E})^{-1}P_{S,E}$. Huomaa, että matriiseja on kerrottava juuri tässä järjestyksessä, sillä kannanvaihtomatriiseilla kerrotaan vektoreita vasemmalta puolelta. Ensin siis operoidaan matriisilla $P_{S,E}$ ja sitten matriisilla $(P_{T,E})^{-1}$.

Kaikkein olennaisinta kannanvaihdossa on kuitenkin se, että kannanvaihtomatriisi on aina olemassa. Esimerkiksi todistuksissa ei yleensä ole kiinnostavaa se, miltä matriisi tarkalleen ottaen näyttää.

1.3 Kannanvaihto yleisessä tapauksessa

Aina ei ole käytettävissä luonnollista kantaa kuten avaruuden K^n tapauksessa. Tällöin voidaan kuitenkin käyttää samanlaisia ideoita kuin yllä, vaikka laskemista tarvitaankin enemmän.

Tutkitaan vektoriavaruutta V ja sen kantoja

$$S = (v_1, v_2, \dots, v_n) \qquad \text{ja} \qquad T = (w_1, w_2, \dots, w_n).$$

Etsitään kannanvaihtomatriisi $P_{S,T}$ kannasta S kantaan T. Nyt täytyy siis päteä

$$P_{S,T}[v]_S = [v]_T \quad \text{kaikilla } v \in V. \tag{1.2}$$

Kuten yllä matriisi $P_{S,T}$ voidaan määrittää pelkkiä kantavektoreita tutkimalla. Kannan S vektorien koordinaattivektorit kannan S suhteen ovat muotoa

$$[v_i]_S = [0\ 0\ \dots\ 0\ 1\ 0 \dots\ 0\ 0]^T.$$

Kannan T suhteen vektorien v_i koordinaattivektorit saattavat näyttää hieman monimutkaisemmilta. Suurin työ kannanvaihdossa onkin selvittää, miltä kyseiset koordinaattivektorit näyttävät.

Kun koordinaattivektorit $[v_i]_T$ on määritetty, on matriisin $P_{S,T}$ muodostaminen helppoa. Huomataan jälleen, että matriisissa P on oltava sarakkeina ne vektorit, jotka matriisikertolaskusta halutaan tulokseksi. Siis

$$P_{S,T} = [[v_1]_T,\ [v_2]_T\ \dots [v_n]_T].$$

Nyt on helppo tarkistaa, että $P_{S,T}[v_i]_S = [v_i]_T$ kaikilla $i \in \{1, 2, \dots, n\}$ ja siten ehto (1.2) pätee.

1.4 Lineaarikuvauksen matriisi

Jos vektorit voidaan kirjoittaa koordinaattivektoreina, lineaarikuvaukset voidaan kirjoittaa matriiseina. Olkoon $L\colon V \to W$ lineaarikuvaus. Oletetaan, että $S = (v_1, v_2, \dots, v_n)$ on avaruuden V kanta ja $T = (w_1, w_2, \dots, w_m)$ on avaruuden W kanta. Tutkitaan, millaisena matriisina kuvaus L voidaan esittää. Nyt halutaan siis löytää matriisi $[L]_{S,T}$, jolle pätee

$$[L]_{S,T}[v]_S = [L(v)]_T \quad \text{kaikilla } v \in V. \tag{1.3}$$

Toisin sanoen, jos matriisilla $[L]_{S,T}$ kertoo sarakevektoria $[v]_S$, saa saman tuloksen kuin kuvaamalla vektoria v kuvauksella L. Tämä tulos on tietenkin kirjoitettu kannan T suhteen. Voidaan osoittaa, että tällainen matriisi on aina olemassa ja se on yksikäsitteinen. Huomaa, että lineaarikuvauksen matriisiesitys riippuu aina valitusta kannasta.

Kun lineaarikuvaus muutetaan matriisiksi, sitä voidaan ajatella sarakeavaruuksien K^n ja K^m välisenä kuvauksena. Tilannetta voidaan havainnollistaa seuraavalla kaaviolla:

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "{[\ ]_S}"'] & W \arrow[d, "{[\ ]_T}"'] \\ K^n \arrow[r, "{[L]_{S,T}}"'] & K^m \end{tikzcd}$$

Koska lineaarikuvauksen arvot määräytyvät täysin kannan arvojen avulla, saadaan matriisin $[L]_{S,T}$ alkiot kantavektoreiden arvoista kuvauksessa L. Samalla tavalla kuin edellä voidaan päätellä, että matriisin $[L]_{S,T}$ sarakkeina ovat kannan S vektorien arvot kirjoitettuna kannan T suhteen:

$$[L]_{S,T} = [[L(v_1)]_T \ [L(v_2)]_T \ \dots \ [L(v_n)]_T].$$

Tällöin ehto (1.3) toteutuu jokaisella kantavektorilla ja edelleen kaikilla vektoreilla.

Oletetaan, että $L\colon V \to V$ on lineaarikuvaus. Olkoot S ja T avaruuden V kantoja. Kannan S suhteen kirjoitetulle matriisille $[L]_{S,S}$ käytetään lyhyyden vuoksi merkintää $[L]_S$. Samalla tavoin merkitään $[L]_{T,T} = [L]_T$.

Kannanvaihtomatriisin avulla voidaan muuttaa myös lineaarikuvauksen L matriisiesitys kannasta toiseen. Oletetaan, että $P_{S,T}$ on kannanvaihtomatriisi kannasta S kantaan T. Tällöin

$$(P_{S,T})^{-1}[L]_T P_{S,T} = [L]_S.$$

Tulos voidaan selittää seuraavalla tavalla: Oletetaan, että tiedetään matriisi $[L]_T$, eli tiedetään, mitä kuvaus L tekee koordinaattivektoreille, jotka on kirjoitettu kannan T suhteen. Halutaan tietää, miltä näyttää matriisi $[L]_S$, eli miten kuvaus L toimii kannan S suhteen. Ensin siirrytään kannasta S kantaan T kannanvaihtomatriisilla $P_{S,T}$. Nyt ollaan kannassa T ja voidaan käyttää matriisia $[L]_T$. Sen jälkeen siirrytään matriisilla $(P_{S,T})^{-1}$ takaisin kantaan S. Seuraava diagrammi havainnollistaa samaa asiaa:

$$\begin{tikzcd} K^n \arrow[r, "{[L]_S}"] \arrow[d, "P_{S,T}"'] & K^n \arrow[d, "P_{S,T}"'] \\ K^n \arrow[r, "{[L]_T}"'] & K^n \end{tikzcd}$$

Jos tarkastellaan yleistä lineaarikuvausta $L\colon V \to W$, voidaan eri kantojen suhteen kirjoitetut matriisit tällöinkin muuttaa toisikseen kannanvaihtomatriisien avulla. Tällöin on vain käytettävä kahta eri kannanvaihtomatriisia, toista avaruudessa V ja toista avaruudessa W.