Kurssimateriaalia K3/P3-kursille syksyllä 2003.
8.10.2003 Heikki Apiola
Sisältää otteita *Timo Eirolan* L3-kurssin lineaarialgebramonisteesta,
jonka lähdekoodin Timo on ystävällisesti antanut käyttööni.

1. Normi ja sisätulo

Vektoriavaruuden määritelmässä riitti olettaa, että joukon alkioille on määritelty aksioomat toteuttavat yhteenlasku ja skalaarilla kertominen. Kuitenkin monissa vektoriavaruuksissa voidaan tunnetusti tehdä muitakin laskutoimituksia. Esim. $\mathbb{R}^2$:ssa tai $\mathbb{R}^3$:ssa voidaan laskea vektoreiden pituuksia, välisiä kulmia ja pistetuloja. Jatkuvia funktioita voidaan kertoa keskenään, integroida, niiden maksimeja voi etsiä jne.

Normiavaruus on sellainen vektoriavaruus, jossa vektoreille on määritelty pituusfunktio, jota kutsutaan normiksi. Sisätuloavaruus on puolestaan normiavaruus, jossa lisäksi kulmien mittaaminen on mahdollista ja erityisesti kohtisuoruus eli ortogonaalisuus on määritelty. Seuraavassa tarkastellaan lähemmin, miten tällaisia pituus- ja kulmafunktioita voidaan määritellä.

Määritelmä 1.1. Olkoon V $\mathbb{K}$-kertoiminen vektoriavaruus. Kuvaus $\|\cdot\| : V \mapsto \mathbb{R}$ on *normi*, jos se toteuttaa

(1) $\|\boldsymbol{v}\| \geq 0 \quad \forall\, \boldsymbol{v} \in V$.

(2) $\|\boldsymbol{v}\| = 0 \implies \boldsymbol{v} = 0$.

(3) $\|\boldsymbol{u}+\boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\| \quad \forall\, \boldsymbol{u},\boldsymbol{v} \in V$.

(4) $\|\alpha\,\boldsymbol{v}\| = |\alpha|\;\|\boldsymbol{v}\| \quad \forall\, \alpha \in \mathbb{K},\ \boldsymbol{v} \in V$.

Vektoriavaruutta, jossa on määritelty jokin normi kutsutaan normiavaruudeksi.

Esimerkki 1.1. Vektoriavaruudessa $\mathbb{R}^n$ tavallisin normi on nk. euklidinen normi

$$\|\boldsymbol{x}\|_2 = \Big(\sum_{i=1}^n |x_i|^2\Big)^{\frac{1}{2}}.$$

Selvästi tämä toteuttaa ehdot (1), (2) ja (4). Ominaisuuden (3) eli *kolmioepäyhtälön* näytämme hieman myöhemmin. Muita usein käytettyjä normeja $\mathbb{R}^n$:ssä ovat[1]

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^n |x_i| \qquad \text{ja} \qquad \|\boldsymbol{x}\|_\infty = \max_{1\leq i\leq n} |x_i|\ .$$

Näistä on helppo näyttää ominaisuudet (1)-(4). Ellei toisin mainita, käytetään $\mathbb{R}^n$:ssä normia $\|\cdot\| = \|\cdot\|_2$.

Avaruudessa $\mathbb{C}^n$ käytetään myös aivan samalla tavalla määriteltyjä normeja.

Määritelmä 1.2. Olkoon V $\mathbb{K}$-kertoiminen vektoriavaruus. Kuvaus $\langle\,\cdot\,,\cdot\,\rangle : V\times V \mapsto \mathbb{K}$ on *sisätulo*, jos se toteuttaa ehdot

(1) $\langle \boldsymbol{v},\boldsymbol{v}\rangle \geq 0$ kaikilla $\boldsymbol{v} \in V$.

(2) $\langle \boldsymbol{v},\boldsymbol{v}\rangle = 0 \implies \boldsymbol{v} = 0$.

(3) $\langle \boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}\rangle = \langle \boldsymbol{u},\boldsymbol{w}\rangle + \langle \boldsymbol{v},\boldsymbol{w}\rangle$ kaikilla $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in V$.

(4) $\langle \alpha\boldsymbol{u},\boldsymbol{v}\rangle = \alpha\,\langle \boldsymbol{u},\boldsymbol{v}\rangle$ kaikilla $\alpha \in \mathbb{K},\ \boldsymbol{u},\boldsymbol{v} \in V$.

(5) $\langle \boldsymbol{v},\boldsymbol{u}\rangle = \overline{\langle \boldsymbol{u},\boldsymbol{v}\rangle}$ kaikilla $\boldsymbol{u},\boldsymbol{v} \in V$.

Sisätulolla varustettua vektoriavaruutta sanotaan sisätuloavaruudeksi.

[1] Normia $\|\cdot\|_1$ kutsutaan taksikuskin normiksi. Miksiköhän?

Reaalisessa tapauksessa (5) saa muodon $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle$ eli reaalinen sisätulo on symmetrinen. Ominaisuudet (3) ja (4) sanovat, että sisätulo on lineaarinen ensimmäisen argumentin suhteen. Toisen argumentin suhteen saadaan:

$$
\begin{aligned}
(1.1) \qquad \langle \boldsymbol{u} \,, \alpha \, \boldsymbol{v} + \beta \, \boldsymbol{w} \rangle &\overset{(5)}{=} \overline{\langle \alpha \, \boldsymbol{v} + \beta \, \boldsymbol{w} \,, \boldsymbol{u} \rangle} \overset{(3),(4)}{=} \overline{\alpha \, \langle \boldsymbol{v} \,, \boldsymbol{u} \rangle + \beta \, \langle \boldsymbol{w} \,, \boldsymbol{u} \rangle} \\
&= \overline{\alpha} \, \overline{\langle \boldsymbol{v} \,, \boldsymbol{u} \rangle} + \overline{\beta} \, \overline{\langle \boldsymbol{w} \,, \boldsymbol{u} \rangle} \overset{(5)}{=} \overline{\alpha} \, \langle \boldsymbol{u} \,, \boldsymbol{v} \rangle + \overline{\beta} \, \langle \boldsymbol{u} \,, \boldsymbol{w} \rangle \ .
\end{aligned}
$$

Täten sisätulo on *konjugoidusti lineaarinen* toisen argumentin suhteen: skalaarit saadaan ulos kompleksikonjugaatteina. Reaalisessa tapauksessa sisätulo on siten lineaarinen myös toisen argumentin suhteen.

Vektoriavaruudesta $\mathbb{R}^n$ tuttu vektoreiden välinen pistetulo[2]: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \boldsymbol{y} = \sum_{i=1}^n x_i \, y_i$ toteuttaa sisätulon ehdot.
Vastaavasti $\mathbb{C}^n$:n vektoreille määritellään $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \, \overline{\boldsymbol{y}} = \sum_{i=1}^n x_i \, \overline{y_i}$.

Esimerkki 1.2. Avaruudessa $C[a, b]$ voidaan määritellä

$$\langle f, g \rangle = \int_a^b f(x) g(x) \, dx \ .$$

Ehdot (1)-(5) seuraavat suoraan integraalin ominaisuuksista.

Esimerkiksi $C[-\pi, \pi]$:ssä funktioiden $f(x) = \sin x$ ja $g(x) = \cos x$ väliset sisätulot ovat

$$
\begin{aligned}
\langle f, g \rangle &= \int_{-\pi}^{\pi} \sin x \cos x \, dx = \int_{-\pi}^{\pi} \tfrac{1}{2} \sin 2x \, dx = 0 \\
\langle f, f \rangle &= \int_{-\pi}^{\pi} \sin^2 x \, dx = \int_{-\pi}^{\pi} \tfrac{1}{2} (1 - \cos 2x) \, dx = \pi \ .
\end{aligned}
$$

Samoin $\langle g, g \rangle = \pi$.

Sisätulon tärkeä ominaisuus on, että se määrittelee heti myös normin: jos V on sisätuloavaruus, asetetaan

$$(1.2) \qquad \| \boldsymbol{v} \| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \ .$$

Sisätulon ehdoista saadaan normin ehdot (1),(2) ja (4) helposti. (3) eli kolmioepäyhtälö vaatii hieman laskemista.

Esitellään ensin *Schwarzin epäyhtälö*[3]: sisätulo ja sen avulla kaavalla (1.2) määritelty $\| \cdot \|$ (jota vielä ei tiedetä normiksi) toteuttavat:

$$(1.3) \qquad | \langle \boldsymbol{u}, \boldsymbol{v} \rangle | \leq \| \boldsymbol{u} \| \, \| \boldsymbol{v} \| \ .$$

Tod. Viittaamme L3-prujuun [TE] tai moninisiin oppikirjoihin. Todistus on tyylipuhdas minimointitehtävä, jossa tarkastellaan toisen asteen polynomia, sopiva vaikka koulukurssiin. Emme kuitenkaan tässä paneudu siihen. $\square$

Näytetään nyt, että kaavalla (1.2) määritelty $\| \cdot \|$ toteuttaa normin ehdon (3) eli kolmioepäyhtälön

$$\| \boldsymbol{u} + \boldsymbol{v} \| \leq \| \boldsymbol{u} \| + \| \boldsymbol{v} \| \ .$$

Tod. Käyttäen sisätulon ominaisuuksia ja Schwarzin epäyhtälöä saadaan

$$
\begin{aligned}
\| \boldsymbol{u} + \boldsymbol{v} \|^2 &= \langle \boldsymbol{u} + \boldsymbol{v} \,, \boldsymbol{u} + \boldsymbol{v} \rangle = \langle \boldsymbol{u}, \boldsymbol{u} \rangle + \langle \boldsymbol{u}, \boldsymbol{v} \rangle + \langle \boldsymbol{v}, \boldsymbol{u} \rangle + \langle \boldsymbol{v}, \boldsymbol{v} \rangle \\
&\leq \| \boldsymbol{u} \|^2 + 2 \ | \langle \boldsymbol{u}, \boldsymbol{v} \rangle | + \| \boldsymbol{v} \|^2 \\
&\leq \| \boldsymbol{u} \|^2 + 2 \, \| \boldsymbol{u} \| \, \| \boldsymbol{v} \| + \| \boldsymbol{v} \|^2 = (\| \boldsymbol{u} \| + \| \boldsymbol{v} \|)^2 \ ,
\end{aligned}
$$

josta väite seuraa. $\square$

---

[2] Lausekkeessa $\boldsymbol{x}^T \boldsymbol{y}$ vektorit on ajateltu $n \times 1$ -matriiseiksi, jolloin $\boldsymbol{x}^T$ on $1 \times n$ -matriisi ja $\boldsymbol{x}^T \boldsymbol{y}$ on 1×1 -matriisi eli skalaari.

[3] Täydellisemmin: Cauchy-Schwarz-Bunjakovskin epäyhtälö.

Kysmys: Onko jokaisen normin taustalla aina sisätulo ?
Vastaus: Ei. Esimerkiksi edellä esiintyneet $\|.\|_1$ (taksikuski) ja $\|.\|_\infty$ eivät ole peräisin mistään sisätulosta.

1.1. **Ortogonaalisuus.**

Vektorit $\boldsymbol{u}$ ja $\boldsymbol{v}$ ovat *ortogonaaliset*, kun $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$. Ortogonaalisuus määritellään samoin kompleksikertoimisissa vektoriavaruuksissa. Täten $\left[\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right]$ ja $\left[\begin{smallmatrix} i \\ 1 \end{smallmatrix}\right]$ ovat ortogonaaliset $\mathbb{C}^2$:ssa.

Sisätuloavaruuden vektorijoukkoa $S = \{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k\}$ sanotaan ortogonaaliseksi, jos kaikki sen vektorit ovat keskenään ortogonaaliset: $\langle \boldsymbol{v}^i, \boldsymbol{v}^j \rangle = 0$, kun $i \neq j$.

Ortogonaalinen vektorijoukko $\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n\}$ on myös lineaarisesti riippumaton edellyttäen, että se ei sisällä nollavektoria. Tämä nähdään seuraavasti. Jos $c_1 \boldsymbol{v}^1 + \cdots + c_n \boldsymbol{v}^n = 0$, otetaan tämän sisätulo $\boldsymbol{v}^k$:n kanssa, jolloin

$$0 = \langle c_1 \boldsymbol{v}^1 + \cdots + c_n \boldsymbol{v}^n \,,\, \boldsymbol{v}^k \rangle = c_1 \langle \boldsymbol{v}^1, \boldsymbol{v}^k \rangle + \cdots + c_k \langle \boldsymbol{v}^k, \boldsymbol{v}^k \rangle + \cdots + c_n \langle \boldsymbol{v}^n, \boldsymbol{v}^k \rangle = c_k \|\boldsymbol{v}^k\|^2$$

ja koska $\boldsymbol{v}^k \neq 0$, saadaan $c_k = 0$. Näin kaikki kertoimet saadaan yksitellen nolliksi, joten $\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n\}$ on lineaarisesti riippumaton.

Jos ortogonaalisen joukon vektorit ovat lisäksi pituudeltaan ykkösiä kutsutaan joukkoa *ortonormaaliksi.*

Samoin, jos matriisin $\boldsymbol{Q} \in \mathbb{R}^{m \times n}$ sarakkeet ovat ortonormaalit (jolloin välttämättä $m \geq n$), saadaan $\boldsymbol{Q}^T \boldsymbol{Q} = \boldsymbol{I}$. Jos $m > n$, niin $\boldsymbol{Q}$ ei kuitenkaan ole invertoituva; sillä on vain *vasemmanpuoleinen inverssi.*

Olkoon $\boldsymbol{U}$ reaalinen tai kompleksinen matriisi, jonka sarakkeet ovat ortonormaalit Tällöin[4] $\boldsymbol{U}^* \boldsymbol{U} = \boldsymbol{I}$ ja

$$\langle \boldsymbol{U}\boldsymbol{x} \,,\, \boldsymbol{U}\boldsymbol{y} \rangle = (\boldsymbol{U}\boldsymbol{y})^* \boldsymbol{U}\boldsymbol{x} = \boldsymbol{y}^* \boldsymbol{U}^* \boldsymbol{U} \boldsymbol{x} = \boldsymbol{y}^* \boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \; .$$

Erityisesti: unitaarisella (reaalisessa tapauksessa ortogonaalisella) matriisilla kerrottaessa vektoreiden pituudet ja niiden väliset sisätulot säilyvät.

Annetun vektorin koordinaatit ortonormaalin kannan suhteen on helppo laskea:

Olkoon $B = \{\boldsymbol{b}^1, \ldots, \boldsymbol{b}^n\}$ sisätuloavaruuden V ortonormaali kanta. Jos $\boldsymbol{v} = c_1 \boldsymbol{b}^1 + \cdots + c_n \boldsymbol{b}^n$, otetaan tämän sisätulo $\boldsymbol{b}^k$:n kanssa, jolloin $\left\langle \boldsymbol{v}, \boldsymbol{b}^k \right\rangle = c_k \left\langle \boldsymbol{b}^k, \boldsymbol{b}^k \right\rangle = c_k$. Näin saadaan kaikki kertoimet. Siis esitys ortonormaalissa kannassa saadaan:

$$\boldsymbol{v} = \sum_{k=1}^{n} \left\langle \boldsymbol{v}, \boldsymbol{b}^k \right\rangle \boldsymbol{b}^k \;, \qquad \text{kaikilla } \boldsymbol{v} \in V \,.$$

Ortonormaaleja kantoja voidaan muodostaa nk. Gram–Schmidtin prosessilla. Olkoon $(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots)$ (äärellinen tai ääretön) jono lineaarisesti riippumattomia sisätuloavaruuden vektoreita. Muodostetaan yhtä pitkä jono $(\boldsymbol{q}^1, \boldsymbol{q}^2, \ldots)$ ortonormaaleja vektoreita seuraavasti:

$$\begin{aligned} \boldsymbol{q}^1 &= \boldsymbol{v}^1 / \|\boldsymbol{v}^1\| \,, \\ \left.\begin{aligned} \boldsymbol{w}^k &= \boldsymbol{v}^k - \textstyle\sum_{j=1}^{k-1} \langle \boldsymbol{v}^k, \boldsymbol{q}^j \rangle \, \boldsymbol{q}^j \,, \\ \boldsymbol{q}^k &= \boldsymbol{w}^k / \|\boldsymbol{w}^k\| \,. \end{aligned}\right\} & \quad k = 2, 3, \ldots \end{aligned} \tag{1.4}$$

Tässä keskimmäisellä rivillä $\boldsymbol{v}^k$:sta poistetaan sen komponentit jo muodostetuilla suunnilla $\boldsymbol{q}^1, \ldots, \boldsymbol{q}^{k-1}$. Viimeisellä rivillä jäljelle jäävä osa normeerataan ykkösen pituiseksi.

Lause 1.1. *Edellä esitetylle Gram-Schmidtin prosessille pätee:*

[4]Kompleksiselle matriisille $\boldsymbol{M}^* = \overline{\boldsymbol{M}}^T$ ja reaaliselle $\boldsymbol{M}^* = \boldsymbol{M}^T$.

a) $(\boldsymbol{q}^1, \boldsymbol{q}^2, \dots)$ *on ortonormaali.*

b) $\mathrm{sp}(\boldsymbol{q}^1, \dots, \boldsymbol{q}^k) = \mathrm{sp}(\boldsymbol{v}^1, \dots, \boldsymbol{v}^k)$ *kaikilla* $k \geq 1$.

Erityisesti, jos V *on äärellisdimensioinen ja* $\{\boldsymbol{v}^1, \dots, \boldsymbol{v}^n\}$ *on sen kanta, niin* $\{\boldsymbol{q}^1, \dots, \boldsymbol{q}^n\}$ *on* V*:n ortonormaali kanta.*

Tod. Prosessi pyörii niin kauan, kun $\boldsymbol{w}^k \neq 0$ (tai $\boldsymbol{v}^j$-vektorit loppuvat). Näytetään aluksi, että b) on voimassa tähän asti. Koska

$$\boldsymbol{v}^k = \|\boldsymbol{w}^k\|\, \boldsymbol{q}^k + \sum_{j=1}^{k-1} \left\langle \boldsymbol{v}^k, \boldsymbol{q}^j \right\rangle \boldsymbol{q}^j ,$$

saadaan kaikilla k : $\boldsymbol{v}^k \in \mathrm{sp}(\boldsymbol{q}^1, \dots, \boldsymbol{q}^k)$, josta $\mathrm{sp}(\boldsymbol{v}^1, \dots, \boldsymbol{v}^k) \subset \mathrm{sp}(\boldsymbol{q}^1, \dots, \boldsymbol{q}^k)$. Toisaalta, jokaiselle $\boldsymbol{q}^k$ selvästi pätee $\boldsymbol{q}^k \in \mathrm{sp}(\boldsymbol{q}^1, \dots, \boldsymbol{q}^{k-1}, \boldsymbol{v}^k)$. Täten induktiivisesti

$$\boldsymbol{q}^k \in \mathrm{sp}(\boldsymbol{q}^1, \dots, \boldsymbol{q}^{k-1}, \boldsymbol{v}^k) \subset \mathrm{sp}(\boldsymbol{q}^1, \dots, \boldsymbol{q}^{k-2}, \boldsymbol{v}^{k-1}, \boldsymbol{v}^k) \subset \dots \subset \mathrm{sp}(\boldsymbol{v}^1, \dots, \boldsymbol{v}^k) .$$

Näin kaikilla k, joten $\mathrm{sp}(\boldsymbol{q}^1, \dots, \boldsymbol{q}^k) \subset \mathrm{sp}(\boldsymbol{v}^1, \dots, \boldsymbol{v}^k)$ ja b) on voimassa.

Jos olisi $\boldsymbol{w}^k = 0$ jollakin k, tämä tarkoittaisi, että

$$\boldsymbol{v}^k = \sum_{j=1}^{k-1} \left\langle \boldsymbol{v}^k, \boldsymbol{q}^j \right\rangle \boldsymbol{q}^j \in \mathrm{sp}(\boldsymbol{v}^1, \dots, \boldsymbol{v}^{k-1})$$

(sillä b) on voimassa vielä edellisellä kierroksella). Mutta tämä on mahdotonta, koska $\boldsymbol{v}^1, \dots, \boldsymbol{v}^k$ ovat lineaarisesti riippumattomat. Siispä $\boldsymbol{w}^k$:t eivät koskaan tule nolliksi.

Todistetaan a) induktiolla:

Selvästi $\{\boldsymbol{q}^1\}$ on ortonormaali.

Oletetaan, että $\{\boldsymbol{q}^1, \dots, \boldsymbol{q}^k\}$ on ortonormaali. Tällöin, kun $i \leq k$, saadaan

$$\begin{aligned}
\left\langle \boldsymbol{q}^{k+1}, \boldsymbol{q}^i \right\rangle &= \left\langle \tfrac{1}{\|\boldsymbol{w}^{k+1}\|} \big(\boldsymbol{v}^{k+1} - \textstyle\sum_{j=1}^k \left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^j \right\rangle \boldsymbol{q}^j\big) , \boldsymbol{q}^i \right\rangle \\
&= \tfrac{1}{\|\boldsymbol{w}^{k+1}\|} \big(\left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^i \right\rangle - \textstyle\sum_{j=1}^k \left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^j \right\rangle \left\langle \boldsymbol{q}^j, \boldsymbol{q}^i \right\rangle \big) \\
&= \tfrac{1}{\|\boldsymbol{w}^{k+1}\|} \big(\left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^i \right\rangle - \left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^i \right\rangle \big) = 0 .
\end{aligned}$$

Näin $\boldsymbol{q}^{k+1}$ on kohtisuorassa kaikkia $\boldsymbol{q}^i$, $i \leq k$ vastaan. Selvästi $\|\boldsymbol{q}^{k+1}\| = 1$. Ja kun muutkin ovat keskenään ortonormaalit, $\{\boldsymbol{q}^1, \dots, \boldsymbol{q}^{k+1}\}$ on ortonormaali. $\square$

Huomaa, että saatava ortonormaali joukko riippuu paitsi vektoreista $\boldsymbol{v}^j$ myös niiden järjestyksestä.

Tehtävä 1.1. Näytä, että äärellisdimensioisen sisätuloavaruuden mielivaltainen ortonormaali joukko voidaan täydentää ortonormaaliksi kannaksi.

Esimerkki 1.3. Lähdetään liikkeelle $\mathbb{R}^3$:n kannasta $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \boldsymbol{v}^3\} = \left\{ \begin{bmatrix} 1\\1\\0 \end{bmatrix}, \begin{bmatrix} 1\\2\\0 \end{bmatrix}, \begin{bmatrix} 1\\1\\2 \end{bmatrix} \right\}$. Saadaan:

$$\begin{aligned}
\boldsymbol{q}^1 &= \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1\\1\\0 \end{bmatrix} \\
\boldsymbol{w}^2 &= \begin{bmatrix} 1\\2\\0 \end{bmatrix} - \tfrac{3}{\sqrt{2}} \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1\\1\\0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\\ \frac{1}{2}\\0 \end{bmatrix} \\
\boldsymbol{q}^2 &= \tfrac{1}{\sqrt{\frac{1}{2}}} \begin{bmatrix} -\frac{1}{2}\\ \frac{1}{2}\\0 \end{bmatrix} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} -1\\1\\0 \end{bmatrix} \\
\boldsymbol{w}^3 &= \begin{bmatrix} 1\\1\\2 \end{bmatrix} - \tfrac{2}{\sqrt{2}} \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1\\1\\0 \end{bmatrix} + 0 \cdot \tfrac{1}{\sqrt{2}} \begin{bmatrix} -1\\1\\0 \end{bmatrix} = \begin{bmatrix} 0\\0\\2 \end{bmatrix} \\
\boldsymbol{q}^3 &= \begin{bmatrix} 0\\0\\1 \end{bmatrix} .
\end{aligned}$$

Näin saatiin ortonormaali kanta $\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} , \ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} , \ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

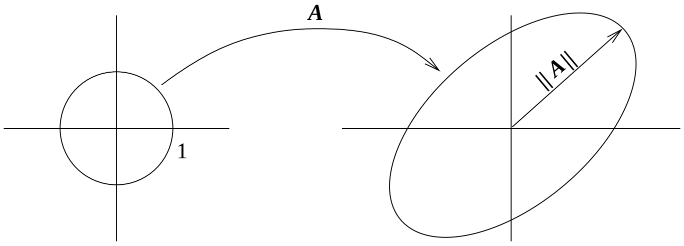
1.2. **Matriisinormi ja häiriöalttius.**

Vektorin normi mittaa vektorin pituutta. Matriiseille ja lineaarikuvauksille voidaan myös määritellä normeja. Erityisen hyödyllisiksi osoittautuvat sellaiset normit, jotka on määritelty vektorinormien avulla. Rajoitumme tässä tarkastelemaan vain matriisien normeja, normiavaruuksien välisten lineaarikuvausten normit määritellään samalla tavalla.

Olkoon $\|\cdot\|$ jokin vektorinormi (esim. $\|\cdot\|_2$ tai $\|\cdot\|_\infty$). Mitataan matriisin kokoa sillä, kuinka pitkiksi vektoreiksi matriisilla kerrottaessa yksikkövektorit saattavat kuvautua. Niinpä matriisille $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ asetetaan

$$\|\boldsymbol{A}\| = \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\| \, . \tag{1.5}$$

Tässä siis oikealla puolella esiintyy vektoreiden $\boldsymbol{x} \in \mathbb{C}^n$ ja $\boldsymbol{A}\boldsymbol{x} \in \mathbb{C}^m$ normeja.



Näin määritelty $\|\boldsymbol{A}\|$ toteuttaa määritelmän 1.1 neljä ehtoa:

(1) (1.5):n oikealla puolella esiintyy vain ei-negatiivisia lukuja, joten $\|\boldsymbol{A}\| \geq 0$.

(2) Jos $\boldsymbol{A} \neq 0$, niin sillä on olemassa ei-nolla elementti a_{ij}. Valitaan $\boldsymbol{x} = \boldsymbol{e}^j$, jolloin $\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \neq 0$ ja $\|\boldsymbol{A}\| \geq \|\boldsymbol{A}\boldsymbol{x}\| > 0$.

(3)
$$\begin{aligned} \|\boldsymbol{A}+\boldsymbol{B}\| &= \max_{\|\boldsymbol{x}\|=1} \|(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}\| \leq \max_{\|\boldsymbol{x}\|=1} (\|\boldsymbol{A}\boldsymbol{x}\| + \|\boldsymbol{B}\boldsymbol{x}\|) \\ &\leq \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\| + \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{B}\boldsymbol{x}\| = \|\boldsymbol{A}\| + \|\boldsymbol{B}\| \, . \end{aligned}$$

Tässä käytettiin aluksi vektorinormin kolmioepäyhtälöä.

(4) $\|\alpha\boldsymbol{A}\| = \max_{\|\boldsymbol{x}\|=1} \|\alpha\boldsymbol{A}\boldsymbol{x}\| = \max_{\|\boldsymbol{x}\|=1} |\alpha|\,\|\boldsymbol{A}\boldsymbol{x}\| = |\alpha|\,\|\boldsymbol{A}\|$
jälleen vektorinormin vastaavan ominaisuuden perusteella.

Matriisinormilla ja vastaavalla vektorinormilla on lisäksi ominaisuudet (harjoitustehtävä)

$$\|\boldsymbol{A}\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{x}\| \, , \tag{1.6}$$

$$\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{B}\| \, , \tag{1.7}$$

$$\|\boldsymbol{A}^k\| \leq \|\boldsymbol{A}\|^k \, , \ k = 1, 2, \ldots \, . \tag{1.8}$$

Kun halutaan korostaa, minkä vektorinormin avulla matriisinormi on määritelty käytetään vastaavaa merkkiä. Esimerkiksi

$$\|\boldsymbol{A}\|_1 = \max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1 \qquad \text{ja} \qquad \|\boldsymbol{A}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2 \, .$$

Riippuen valitusta vektorinormista matriisin normin laskeminen voi olla hankalaa tai helpompaa. 1- ja ∞-normit ovat laskuissa monesti käteviä:

Lause 1.2. *Olkoon* $A \in \mathbb{C}^{m\times n}$. *Tällöin*

$$\|\boldsymbol{A}\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}| \qquad \textit{ja} \qquad \|\boldsymbol{A}\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|$$

Tod. Jos $\|\boldsymbol{x}\|_1 = \sum_{k=1}^n |x_k| = 1$, niin

$$\begin{aligned}\|\boldsymbol{Ax}\|_1 &= \sum_{i=1}^m |(\boldsymbol{Ax})_i| = \sum_{i=1}^m \left|\sum_{k=1}^n a_{ik}x_k\right| \le \sum_{i=1}^m\sum_{k=1}^n |a_{ik}x_k| \\ &= \sum_{k=1}^n |x_k| \sum_{i=1}^m |a_{ik}| \le \sum_{k=1}^n |x_k| \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}| = \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}| \ .\end{aligned}$$

Siten $\|\boldsymbol{A}\|_1 \le \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}|$. Toisaalta, jos l on siten, että

$$\sum_{i=1}^m |a_{i\,l}| = \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}| \ ,$$

niin

$$\|\boldsymbol{Ae}^l\|_1 = \left\| \begin{bmatrix} a_{1\,l} \\ \vdots \\ a_{m\,l}\end{bmatrix} \right\|_1 = \sum_{i=1}^m |a_{i\,l}| \ ,$$

joten $\|\boldsymbol{A}\|_1 \ge \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}|$.

$\|\cdot\|_\infty$ -normia koskeva väite jätetään harjoitustehtäväksi. □

Tehtävä 1.2. Millaisia yleisesti päteviä epäyhtälöitä saat matriisin $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ normien $\|\boldsymbol{A}\|_1$, $\|\boldsymbol{A}\|_2$ ja $\|\boldsymbol{A}\|_\infty$ välille? Katso vastaavien vektorinormien välisiä epäyhtälöitä (tehtävä **??**).

Seuraava tärkeä tulos tulee käyttöön vielä useasti. Loppupuolella esitämme sille myös toisen todistuksen.

Lause 1.3. *Olkoon* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *siten, että* $\|\boldsymbol{A}\| < 1$. *Tällöin* $\boldsymbol{I} - \boldsymbol{A}$ *on invertoituva ja* $\|(\boldsymbol{I}-\boldsymbol{A})^{-1}\| \le \frac{1}{1-\|\boldsymbol{A}\|}$.

Tod. Jos $\boldsymbol{I} - \boldsymbol{A}$ ei ole invertoituva, niin on olemassa $\boldsymbol{x} \in \mathbb{C}^n$ siten, että $\|\boldsymbol{x}\| = 1$ ja $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x} = 0$. Tällöin $\|\boldsymbol{A}\| \ge \|\boldsymbol{Ax}\| = \|\boldsymbol{x}\| = 1$, mikä on ristiriita.

Jos $\|\boldsymbol{x}\| = 1$ ja $\boldsymbol{v} = (\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{x}$, niin

$$1 = \|(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{v}\| \ge \|\boldsymbol{v}\| - \|\boldsymbol{Av}\| \ge \|\boldsymbol{v}\| - \|\boldsymbol{A}\|\|\boldsymbol{v}\| = (1-\|\boldsymbol{A}\|)\|\boldsymbol{v}\| \ .$$

Siten $\|\boldsymbol{v}\| \le \frac{1}{1-\|\boldsymbol{A}\|}$. □

Häiriöalttius. Kun käytännön tehtävissä päädytään lineaariseen malliin $\boldsymbol{Ax} = \boldsymbol{b}$, niin usein yhtälöiden kertoimissa ja datassa eli matriisin $\boldsymbol{A}$ tai vektorin $\boldsymbol{b}$ alkioissa, on epävarmuutta. Kertoimet on voitu saada esimerkiksi mittausten tuloksena. Halutaan tietää, miten suuri virhe tästä voi aiheutua ratkaisuun $\boldsymbol{x}$. Tarkastellaan ensin, miten $\delta\boldsymbol{b}$:n suuruinen häiriövektori oikean puolen vektorissa vaikuttaa ratkaisuun. Merkitään $\delta\boldsymbol{x}$:llä ratkaisuvektorin muutosta. Vähentämällä yhtälöt

$$\boldsymbol{Ax} = \boldsymbol{b} \qquad \text{ja} \qquad \boldsymbol{A}(\boldsymbol{x}+\delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$$

puolittain, saadaan $\delta\boldsymbol{x} = \boldsymbol{A}^{-1}\delta\boldsymbol{b}$. Siten *absoluuttisen virheen* normille saadaan yläraja

$$\|\delta\boldsymbol{x}\| \le \|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{b}\| \ . \tag{1.9}$$

Lineaarisen yhtälöryhmän ratkaisun voi kerroinmatriisia skaalaamalla saada pienemmäksi, jolloin myös absoluuttinen virhe pienenee. Paremmin ratkaisun virhettä kuvaakin *suhteellinen virhe* $\|\delta\boldsymbol{x}\|/\|\boldsymbol{x}\|$. Koska

$$\|\boldsymbol{b}\| \le \|\boldsymbol{A}\|\|\boldsymbol{x}\| \tag{1.10}$$

niin epäyhtälöistä (1.9) ja (1.10) saadaan suhteelliselle virheelle yläraja-arvio

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\| \frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|} .$$

Tämän perusteella asetetaan

Määritelmä 1.3. Matriisin *häiriöalttius* on

$$\kappa(\boldsymbol{A}) = \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\| .$$

Suuri häiriöalttius merkitsee siten, että pienikin suhteellinen virhe $\boldsymbol{b}$:ssä voi aiheuttaa ratkaisuun $\boldsymbol{x}$ suuren epävarmuuden. Aivan vastaavasti voidaan tarkastella matriisin $\boldsymbol{A}$ häiriön $\delta\boldsymbol{A}$ aiheuttamaa virhettä ratkaisuun, ja saadaan

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|} \leq \kappa(\boldsymbol{A}) \frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|} .$$

Häiriöalttius riippuu (hieman) siitä, missä matriisinormissa (ja vastaavassa vektorinormissa) asioita mitataan. Koska $1 = \|\boldsymbol{I}\| = \|\boldsymbol{A}\boldsymbol{A}^{-1}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|$, saadaan $\kappa(\boldsymbol{A}) \geq 1$ jokaiselle (invertoituvalle) matriisille normista riippumatta.

Huomaa, että (toisin kuin determinantti) häiriöalttius ei riipu matriisin skaalauksesta:

$$\kappa(\alpha\boldsymbol{A}) = \|\alpha\boldsymbol{A}\|\|(\alpha\boldsymbol{A})^{-1}\| = |\alpha|\,\|\boldsymbol{A}\|\,\|\tfrac{1}{\alpha}\ \boldsymbol{A}^{-1}\| = \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\| = \kappa(\boldsymbol{A}) .$$

Unitaariselle matriisille $\boldsymbol{U}$ pätee $\|\boldsymbol{U}\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$, joten $\|\boldsymbol{U}\|_2 = 1$ ja samoin $\|\boldsymbol{U}^{-1}\|_2 = \|\boldsymbol{U}^*\|_2 = 1$, joten $\kappa_2(\boldsymbol{U}) = 1$. Siten unitaarisen matriisin häiriöalttius (2-normissa mitattuna) on pienin mahdollinen.

Esimerkki 1.4. Lasketaan $\kappa_1(\boldsymbol{A})$, kun $\boldsymbol{A} = \begin{bmatrix} 1 & \varepsilon \\ 1 & -\varepsilon \end{bmatrix}$. Nyt $\boldsymbol{A}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1/\varepsilon & -1/\varepsilon \end{bmatrix}$ joten häiriöalttiudeksi saadaan lauseella 1.2 (kun $\varepsilon \in (0,1)$)

$$\kappa_1(\boldsymbol{A}) = \|\boldsymbol{A}\|_1\|\boldsymbol{A}^{-1}\|_1 = 2\,\tfrac{1}{2}\,(1 + 1/\varepsilon) = 1 + 1/\varepsilon ,$$

joka on suuri ε:n ollessa pieni.