LIITE 1: Matriisilaskentaa

Tämän liitten tarkoituksena on antaa kurssilla tarvittavat perustiedot matriisilaskennasta. Aloitetaan eräästä perustyökalusta, *Householderin projektiosta.* Olkoon $\nu \in \mathbb{R}^n$, $\nu \neq 0$. Vektoriin ν liittyvä Householderin projektio määritellään kaavalla

$$P_\nu = \mathbf{1}_n - \frac{2}{\|\nu\|^2}\nu\nu^T.$$

Tällä matriisilla on seuraavat ominaisuudet:

1. P_ν määrittelee heijastuksen tason $\mathrm{sp}\{\nu\}^\perp = \{p \in \mathbb{R}^n \mid p^T\nu = 0\}$ suhteen,
2. P_ν on symmetrinen ja ortogonaalinen.

Näiden ominaisuuksien toteaminen jätetään harjoitustehtäväksi.

Matriisin hyödyllisyys perustuu siihen, että sen avulla voidaan nollata annetun vektorin komponentit yhtä lukuunottamatta. Olkoon $x \in \mathbb{R}^n$. Haluamme valita vektorin ν siten, että $P_\nu x \in \mathrm{sp}\{e_1\}$, missä $e_1 = (1, 0, \ldots, 0)^T$. Koska P_ν:n määritelmästä seuraa, että $P_\nu x \in \mathrm{sp}\{x, \nu\}$, on ν valittava avaruudesta $\mathrm{sp}\{x, e_1\}$. Kirjoitetaan siis $\nu = x + \alpha e_1$, ja etsitään sopiva $\alpha \in \mathbb{R}$. Todetaan, että

$$\nu^T x = \|x\|^2 + \alpha x_1,$$

ja

$$\|\nu\|^2 = \|x + \alpha e_1\|^2 = \|x\|^2 + 2\alpha x_1 + \alpha^2.$$

Näin ollen saamme

$$\begin{aligned} P_\nu x &= x - 2\frac{\nu^T x}{\|\nu\|^2}\nu = \left(1 - 2\frac{\nu^T x}{\|\nu\|^2}\right)x - 2\frac{\nu^T x}{\|\nu\|^2}\alpha e_1 \\ &= \left(1 - 2\frac{\|x\|^2 + \alpha x_1}{\|x\|^2 + 2\alpha x_1 + \alpha^2}\right)x - 2\frac{\nu^T x}{\|\nu\|^2}\alpha e_1 \\ &= \frac{\alpha^2 - \|x\|^2}{\|x\|^2 + 2\alpha x_1 + \alpha^2}x - 2\frac{\nu^T x}{\|\nu\|^2}\alpha e_1. \end{aligned}$$

Nähdään, että α tulee valita siten, että vektorin x kerroin häviää. Tämä ehto täyttyy, jos $\alpha = \pm\|x\|$. Periaatteessa kumpikin merkki käy. Numeerisen stabiilisuuden takaamiseksi on kuitenkin merkin valinnassa oltava tarkka. Osoittautuu, että laskennallisesti on viisasta valita

$$\nu = x + \mathrm{sign}(x_1)\|x\|e_1.$$

Jos nimittäin x on jo alun perin hyvin lähellä vektoria e_1, eli $x = x_1 e_1 + x'$, missä $e_1^T x' = 0$ ja $\|x'\|$ on pieni verrattuna $|x_1|$:een, niin tällöin

$$x - \mathrm{sign}(x_1)\|x\|e_1 = x_1 e_1 + x' - \mathrm{sign}(x_1)\sqrt{x_1^2 + \|x'\|^2}e_1$$

$$= x_1 e_1 + x' - \mathrm{sign}(x_1)|x_1|\sqrt{1 + \frac{\|x'\|^2}{x_1^2}}e_1 \approx \left(x_1 - x_1\left(1 - \frac{\|x'\|^2}{2x_1^2}\right)\right) + x'$$

$$= \frac{\|x'\|^2}{2x_1}e_1 + x',$$

mistä nähdään, että kyseisen vektorin normi on pieni, jolloin matriisin P_ν lasku olisi numeerisesti epästabiilia.

Householderin projektion eräänä tärkeänä sovelluksena todistetaan matriisien *QR–hajotelma.*

Lause L 1 *Jokainen matriisi $A \in \mathbb{R}^{n\times m}$ voidaan kirjoittaa tulomuotoon*

$$A = QR,$$

missä $Q \in \mathbb{R}^{n\times n}$ on ortogonaalinen ja matriisi $R \in \mathbb{R}^{n\times m}$ on yläkolmiomatriisi, eli $R_{i,j} = 0$, mikäli $i > j$.

Todistus: Käytetään induktiota rivien lukumäärän suhteen. Kun $A \in \mathbb{R}^{1\times m}$, on lause triviaali, onhan jokainen vaakavektori yläkolmiomatriisi.

Oletetaan siis, että lause on voimassa kaikilla m, kun $n = k$. Olkoon $A \in \mathbb{R}^{(k+1)\times m}$. Olkoon $P \in \mathbb{R}^{(k+1)\times(k+1)}$ Householderin projektio, joka kiertää matriisin A ensimmäisen sarakkeen a vektorin e_1 suuntaiseksi. Jos merkitään $A = (a^{(1)}, a^{(2)}, \ldots, a^{(m)})$, niin

$$Pa^{(1)} = P^T a^{(1)} = \begin{pmatrix} a \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ \mathbf{0}_{k\times 1} \end{pmatrix}.$$

Tällöin PA on lohkomuodossa kirjoitettuna

$$PA = \begin{pmatrix} a & w^T \\ \mathbf{0}_{k\times 1} & C \end{pmatrix},$$

missä $w \in \mathbb{R}^{m-1}$, $C \in \mathbb{R}^{k\times(m-1)}$. Induktio–oletuksen nojalla löydetään nyt ortogonaalinen matriisi $Q \in \mathbb{R}^{k\times k}$ siten, että $Q^T C = R \in \mathbb{R}^{k\times(m-1)}$, missä R on yläkolmiomatriisi. Määritellään

$$\tilde{Q} = \begin{pmatrix} 1 & \mathbf{0}_{1\times k} \\ \mathbf{0}_{k\times 1} & Q \end{pmatrix} = \mathrm{diag}(1, Q).$$

Tällöin pätee

$$\tilde{Q}^T PA = \mathrm{diag}(1, Q) \begin{pmatrix} a & w^T \\ \mathbf{0}_{k\times 1} & C \end{pmatrix} = \begin{pmatrix} a & w^T \\ \mathbf{0}_{k\times 1} & R \end{pmatrix},$$

missä siis oikea puoli on yläkolmiomuotoa. Koska matriisi $\tilde{Q}^T P$ on ortogonaalinen, on lause todistettu. $\square$

Edellä todistettu lause antaa samalla algoritmin QR–hajotelman laskemiseksi Householderin projektioilla.

Toinen kanoninen matriisihajotelma, jota näissä luennoissa tarvitaan, on matriisin *singulaariarvohajotema* (lyhennetään yleensä SVD). Tämän hajotelman johtaminen käy ehkä helpoimmin käyttäen matriisin normia. Matriiseille voidaan määritellä useita eri normeja, joista tässä tarkastellaan vain euklidista operaatorinormia. Olkoon $A \in \mathbb{R}^{n\times m}$. Määritellään

$$\|A\| = \sup\{\|Ax\| \mid x \in \mathbb{R}^m,\ \|x\| = 1\}.$$

Koska $\mathbb{R}^m$:n yksikköpallo on kompakti, voidaan supremum luonnollisesti korvata maksimilla.

Lause L 2 *Jokainen matriisi $A \in \mathbb{R}^{m\times n}$ voidaan kirjoittaa muotoon*

$$A = UDV^T,$$

missä $U \in \mathbb{R}^{m\times m}$ ja $V \in \mathbb{R}^{n\times n}$ ovat ortogonaalisia ja $D \in \mathbb{R}^{m\times n}$ on diagonaalimatriisi, jonka diagonaalielementit ovat ei–negatiivisia.

Ennen lauseen todistusta mainitaan, mitä diagonaalisuus tarkoittaa matriisille, joka ei ole neliömatriisi. Jos $m < n$, niin diagonaalinen D on muotoa

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & d_2 & & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & & & & \vdots \\ 0 & \dots & & d_m & 0 & \dots & 0 \end{pmatrix} = \left(\mathrm{diag}(d_1, \dots, d_m), \mathbf{0}_{m\times(n-m)}\right).$$

Jos taas $m > n$, niin diagonaalinen D on muotoa

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & & \vdots \\ \vdots & & \ddots & \\ & & & d_n \\ 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix} = \begin{pmatrix} \mathrm{diag}(d_1, \dots, d_n) \\ \mathbf{0}_{(m-n)\times n} \end{pmatrix}.$$

Lyhyesti voimme siis merkitä $D = \mathrm{diag}(d_1, \ldots, d_{\min(m,n)})$.

Lauseen L 2 todistus: Merkitään $\|A\| = \sigma$. Voimme olettaa, että $\sigma \neq 0$. Olkoon $x \in \mathbb{R}^n$ yksikkövektori, jolle $\|A\| = \|Ax\|$. Olkoon $y = (1/\sigma)Ax$, jolloin siis myös y on yksikkövektori. Valitaan vektorit $v^{(2)}, \ldots, v^{(n)} \in \mathbb{R}^n$ ja vastaavasti $u^{(2)}, \ldots, u^{(m)} \in \mathbb{R}^m$ siten, että $\{x, v^{(2)}, \ldots, v^{(n)}\}$ muodostaa $\mathbb{R}^n$:n ortogonaalisen normitetun kannan ja $\{y, u^{(2)}, \ldots, u^{(m)}\}$ $\mathbb{R}^m$:n normitetun ortogonaalisen kannan. kirjoittamalla

$$V = (x, v^{(2)}, \ldots, v^{(n)}) \in \mathbb{R}^{n\times n}, \quad U = (y, u^{(2)}, \ldots, u^{(m)}) \in \mathbb{R}^{m\times m},$$

todetaan, että matriisit ovat ortogonaalisia, ja että

$$A_1 = U^T AV = \begin{pmatrix} y^T \\ (u^{(2)})^T \\ \vdots \\ (u^{(m)})^T \end{pmatrix} \left(\sigma y, Av^{(2)}, \ldots, Av^{(n)}\right) = \begin{pmatrix} \sigma & w^T \\ \mathbf{0}_{(m-1)\times 1} & B \end{pmatrix},$$

missä $w \in \mathbb{R}^{(n-1)}$, $B \in \mathbb{R}^{(m-1)\times(n-1)}$. Koska

$$A_1 \begin{pmatrix} \sigma \\ w \end{pmatrix} = \begin{pmatrix} \sigma^2 + \|w\|^2 \\ Bw \end{pmatrix},$$

niin

$$\left\| A_1 \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\| \geq \sigma^2 + \|w\|^2,$$

eli erityisesti

$$\|A_1\| \geq \sqrt{\sigma^2 + \|w\|^2}.$$

Mutta toisaalta ortogonaalimatriiseilla kertominen säilyttää matriisin normin (harjoitustehtävä), joten

$$\|A\| = \|A_1\| \geq \sqrt{\sigma^2 + \|w\|^2}.$$

Tästä päätellään, että $w = 0$. Toistamalla päättely ensin matriisille B ja jatkamalla induktiivisesti saadaan haluttu väite. $\square$

Huomautus: Edellä esitetystä todistuksesta nähdään, että diagonaalielementit tulevat annetulla konstruktiolla samalla järjestetyksi laskevaan järjestykseen:

$$d_1 \geq d_2 \geq \ldots \geq d_{\min(m,n)}$$

(harjoitustehtvä).

Singulaariarvohajotelman avulla voidaan määritellä eräs hyödyllinen käsite, *matriisin häiriöalttiusluku (eng. condition number).* Häiriöalttiusluku $\kappa(A)$ määritellään suurimman ja pienimmän singulaariarvon suhteena,

$$\kappa(A) = \frac{\max(d_j)}{\min(d_j)}.$$

Mikäli singulaariarvojen minimi on nolla, määritellään $\kappa(A) = \infty$

Käydään läpi vielä yksi kanoninen hajotelma. Olkoon $A \in \mathbb{R}^{n\times n}$ neliömatriisi. Matriisi on *positiividefiniitti*, mikäli kaikilla $x \in \mathbb{R}^n$, $x \neq 0$ pätee

$$x^T Ax > 0.$$

Vastaavasti matriisia A sanotaan *positiivisemidefiniitti*, jos kaikilla $x \in \mathbb{R}^n$ pätee

$$x^T Ax \geq 0.$$

Positiividefiniiteille matriiseille on voimassa seuraava *Choleskyn hajotelma*.

Lause L 3 *Jokainen symmetrinen positiividefiniitti neliömatriisi $A \in \mathbb{R}^{n\times n}$ voidaan esittää muodossa*

$$A = K^T K,$$

missä matriisi K on yläkolmiomatriisi, jonka diagonaalielementit ovat positiivisia.

Todistus: Todistus tapahtuu jälleen induktiolla matrisin koon suhteen. Kun $n = 1$, on tulos triviaali. Oletetaan siis, että tulos pätee, kun $n = k$. Olkoon $A \in \mathbb{R}^{(k+1)\times(k+1)}$. Kirjoitetaan A lohkomuotoon

$$A = \begin{pmatrix} a & b^T \\ b & \tilde{A} \end{pmatrix},$$

missä $a \in \mathbb{R}$, $b \in \mathbb{R}^k$ ja $\tilde{A} \in \mathbb{R}^{k\times k}$. Todetaan aluksi, että $a > 0$, sillä

$$a = e_1^T Ae_1 > 0.$$

Määritellään nyt

$$w = \frac{1}{\sqrt{a}} b \in \mathbb{R}^k.$$

Osoitetaan aluksi, että matriisi $\tilde{A} - ww^T$ on positiividefiniitti. Olkoon sitä varten $x \in \mathbb{R}^k$ mielivaltainen, $x \neq 0$. Tällöin matriisin A:n positiividefiniittisyyden nojalla pätee

$$\begin{pmatrix} -\frac{x^Tb}{a} & x^T \end{pmatrix} A \begin{pmatrix} -\frac{x^Tb}{a} \\ x \end{pmatrix} > 0.$$

Mutta toisaalta

$$A \begin{pmatrix} -\frac{x^Tb}{a} \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{x^Tb}{a} b + \tilde{A}x \end{pmatrix},$$

joten

$$\begin{pmatrix} -\frac{x^Tb}{a} & x^T \end{pmatrix} A \begin{pmatrix} -\frac{x^Tb}{a} \\ x \end{pmatrix} = \frac{(x^Tb)^2}{a} + x^T\tilde{A}x = x^T(\tilde{A} - ww^T)x > 0.$$

Induktio–oletuksen nojalla on siis olemassa yläkolmiomatriisi $\tilde{K}$, jonka diagonaalielementit ovat positiiviset ja jolle

$$\tilde{A} - ww^T = \tilde{K}^T \tilde{K}.$$

Määrittelemällä matriisi K kaavalla

$$K = \begin{pmatrix} \sqrt{a} & w^T \\ 0 & \tilde{K} \end{pmatrix}$$

todetaan suoralla laskulla, että K toteuttaa väitteen vaatimukset. $\square$

Harjoitustehtäviä:

1. Todista Householderin projektion ominaisuudet 1.–2. Totea, että ortogonaalisella matriisilla kertominen säilyttää sisätulon ja normin.
2. Todista, että ortigonaalimatriisien tulo on ortogonaalinen.
3. Lauseessa L2 annetaan induktiivinen konstruktio matriisin singulaariarvohajotelmalle. Näytä, että mainittu rekursiivinen algoritmi antaa singulaariarvot laskevassa järjestyksessä.
4. Pohjautuen lauseen L1 todistukseen kirjoita QR–hajoteman ortogonaalisen matriisin Q eksplisiittinen lauseke Householderin projektioita käyttäen.