Insinöörimatematiikka: Lineaarialgebra

Mika Hirvensalo
mikhirve@utu.fi

Matematiikan ja tilastotieteen laitos
Turun yliopisto

2024

Koordinaattivektorin etsiminen

Jos $B_1 = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\} \subseteq \mathbb{R}^n$ on jokin kanta, on jokaiselle $\boldsymbol{x} \in \mathbb{R}^n$ olemassa kantaesitys:

$$\boldsymbol{x} = x_1\boldsymbol{b}_1 + \ldots + x_n\boldsymbol{b}_n,$$

Koordinaattivektori $\boldsymbol{x}_B = (x_1, \ldots, x_n)$ voidaan löytää Gaussin-Jordanin prosessilla (Yleensä työläs prosessi)

Koordinaattivektorin etsiminen, Ortonormaali kanta

Jos $B_1 = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\} \subseteq \mathbb{R}^n$ on ortonormaali kanta, voidaan kantaesitys

$$\boldsymbol{x} = x_1\boldsymbol{b}_1 + \ldots + x_n\boldsymbol{b}_n,$$

löytää helpommin. Koordinaattivektori $(x_1, \ldots, x_n)$ voidaan löytää laskemalla sisätulo $(\boldsymbol{x}, \boldsymbol{b}_i)$, jolloin

$$\begin{aligned}(\boldsymbol{x}, \boldsymbol{b}_i) &= x_1(\boldsymbol{b}_1, \boldsymbol{b}_i) + \ldots + x_i(\boldsymbol{b}_i, \boldsymbol{b}_i) + \ldots + x_n(\boldsymbol{b}_n, \boldsymbol{b}_i) \\ &= x_i(\boldsymbol{b}_i, \boldsymbol{b}_i) = x_i||\boldsymbol{b}_i||^2,\end{aligned}$$

siis

$$x_i = \frac{(\boldsymbol{x}, \boldsymbol{b}_i)}{||\boldsymbol{b}_i||^2}$$

Määritelmä

Olkoot $\boldsymbol{x}$ ja $\boldsymbol{y} \neq \boldsymbol{0}$ avaruuden $\mathbb{R}^n$ vektoreita. Tällöin *vektorin* $\boldsymbol{x}$ *projektio vektorilla* $\boldsymbol{y}$ tarkoittaa vektoria x', joka on vektorin $\boldsymbol{y}$ kanssa yhdensuuntainen ja jolle $\boldsymbol{x} - \boldsymbol{x}'$ on kohtisuorassa vektoria $\boldsymbol{y}$ vastaan. Merkitään $\boldsymbol{x}' = \mathrm{Proj}_{\boldsymbol{y}}(\boldsymbol{x})$

Gramin-Schmidtin ortogonalisointiprosessi

Lause: Jos $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ on lineaarisesti riippumaton joukko ja $(\cdot, \cdot)$ jokin sisätulo, on olemassa ortogonaali joukko $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k\}$, jolle $L(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) = L(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$

Etäisyyden säilyttävät kuvaukset

Määritelmä

Lineaarikuvaus $f : V \to V$ säilyttää etäisyyden, jos $d(f(\boldsymbol{x}), f(\boldsymbol{y})) = d(\boldsymbol{x}, \boldsymbol{y})$ aina, kun $\boldsymbol{x}$, $\boldsymbol{y} \in V$. Kuvaus f säilyttää normin, jos $||f(\boldsymbol{x})|| = ||\boldsymbol{x}||$. Kuvaus f säilyttää sisätulon, jos $(f(\boldsymbol{x}), f(\boldsymbol{y})) = (\boldsymbol{x}, \boldsymbol{y})$.

Lause

Olkoon etäisyys normin ja tämä puolestaan sisätulon indusoima. Jos lineaarikuvaus $f : V \to V$ säilyttää sisätulon, säilyttää se myös normin ja etäisyyden.

Etäisyyden säilyttävät kuvaukset

Määritelmä

Matriisi A on ortogonaali, jos $A^T = A^{-1}$, siis $A^T A = AA^T = I$.

Lause

Jos lineaarikuvauksen $f : V \to V$ matriisi A_f on ortogonaali, niin kuvaus f säilyttää pistetulon (ja siis myös tämän indusoiman normin ja etäisyyden)

Määritelmä

Matriisin A *adjugaatti* määritellään $A^* = \overline{A}^T$. Matriisi A on *unitaarinen*, jos $A^* = A^{-1}$, siis $A^* A = AA^* = I$.

Lause

Jos lineaarikuvauksen $f : V \to V$ matriisi A_f on unitaarinen, niin f säilyttää Hermiten pistetulon (ja siis myös sen indusoiman normin ja etäisyyden)

Kulma avaruudessa $\mathbb{R}^n$

Vektoreiden $\boldsymbol{x}$ ja $\boldsymbol{y} \in \mathbb{R}^n$ välinen *kulma* tarkoittaa kulmaa $\theta \in [0, \pi]$, joka muodostuu origosta (pisteestä $\mathbf{0}$) pisteeseen $\boldsymbol{x}$ ja origosta pisteeseen $\boldsymbol{y}$ kulkevien janojen välille.
Huomaa, että kolme pistettä $\boldsymbol{x}$, $\boldsymbol{y}$ ja $\mathbf{0}$ kuuluvat aina samalle tasolle, joten vektoreiden välinen kulma on aina *tason* ominaisuus.

Määritelmä

Jos θ on vektoreiden $\boldsymbol{x}$ ja $\boldsymbol{y}$ välinen kulma, niin

$$\boldsymbol{x} \cdot \boldsymbol{y} = ||\boldsymbol{x}|| \cdot ||\boldsymbol{y}|| \cos \theta.$$

Huomautus

Jos kumpikaan $\boldsymbol{x}$ tai $\boldsymbol{y}$ ei ole nollavektori, on

$$\cos \theta = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{||\boldsymbol{x}||||\boldsymbol{y}||}$$

Yleistys

Nollasta poikkeavien vektoreiden $\boldsymbol{x}$ ja $\boldsymbol{y}$ välinen kulma θ määritellään

$$\cos\theta = \frac{(\boldsymbol{x}, \boldsymbol{y})}{||\boldsymbol{x}||||\boldsymbol{y}||}.$$

Nollavektorin välinen kulma toiseen vektoriin määritellään aina suoraksi kulmaksi.

Huomautus

Cauchyn-Schwarzin epäyhtälön perusteella

$$|(\boldsymbol{x}, \boldsymbol{y})|^2 \leq (\boldsymbol{x}, \boldsymbol{x})(\boldsymbol{y}, \boldsymbol{y}) = ||\boldsymbol{x}||^2 \cdot ||\boldsymbol{y}||^2,$$

joten

$$|(\boldsymbol{x}, \boldsymbol{y})| \leq ||\boldsymbol{x}||||\boldsymbol{y}||$$

Huomioita

Vektorin normi ortogonaalissa kannassa

Olkoon $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ ortogonaali kanta $\mathbb{R}^n$:ssä. Jos

$$\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \ldots + \alpha_n \boldsymbol{b}_n,$$

on

$$||\boldsymbol{x}||^2 = (\sum_{i=1}^{n} \alpha_i \boldsymbol{b}_i, \sum_{j=1}^{n} \alpha_j \boldsymbol{b}_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j (\boldsymbol{b}_i, \boldsymbol{b}_j) = \sum_{i=1}^{n} \alpha_i^2 ||\boldsymbol{b}_i||^2.$$

Erikoistapaus ($n = 2$): $\boldsymbol{x} = \boldsymbol{b}_1 + \boldsymbol{b}_2$. Tällöin

$$||\boldsymbol{x}||^2 = ||\boldsymbol{b}_1||^2 + ||\boldsymbol{b}_2||^2$$

Erikoistapaus: Ortonormaali kanta.

Huomioita

Määritelmä

Matriisi A on *positiividefiniitti*, jos $\boldsymbol{x}^T A\boldsymbol{x} > 0$ kaikille $\boldsymbol{x} \neq 0$.
Matriisi A on *positiivisemidefiniitti*, $\boldsymbol{x}^T A\boldsymbol{x} \geq 0$ kaikille $\boldsymbol{x}$

Lause

Jos matriisi A on positiivisemidefiniitti, ovat sen ominaisarvot $\lambda \geq 0$.

Todistus: Olkoon $\boldsymbol{x} \neq \mathbf{0}$ jokin λ:n kuuluva ominaisvektori. Tällöin

$$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{x}^T \lambda\boldsymbol{x} = \lambda||\boldsymbol{x}||^2,$$

josta

$$\lambda = \frac{\boldsymbol{x}^T A\boldsymbol{x}}{||\boldsymbol{x}||^2} \geq 0.$$

Huomioita

Lause

Symmetrinen $n \times n$ matriisi A on positiivisemidefiniitti jos ja vain jos sen ominaisarvot ovat ≥ 0.

Todistus

Symmetrisen matriisin A ominaisvektorit $\boldsymbol{b}_1$, ..., $\boldsymbol{b}_n$ voidaan valita siten, että ne muodostavat avaruuden $\mathbb{R}^n$ ortonormaalin kannan. Tällöin mikä hyvänsä vektori $\boldsymbol{x} \in \mathbb{R}^n$ voidaan esittää muodossa $\boldsymbol{x} = x_1\boldsymbol{b}_1 + \ldots + x_n\boldsymbol{b}_n$, jolloin

$$\begin{aligned}
\boldsymbol{x}^T A\boldsymbol{x} &= \boldsymbol{x} \cdot A\boldsymbol{x} = \boldsymbol{x} \cdot (x_1 A\boldsymbol{b}_1 + \ldots + x_n A\boldsymbol{b}_n) \\
&= (x_1\boldsymbol{b}_1 + \ldots + x_n\boldsymbol{b}_n) \cdot (x_1\lambda_1\boldsymbol{b}_1 + \ldots + x_n\lambda_n\boldsymbol{b}_n) \\
&= x_1^2\lambda_1 + \ldots + x_n^2\lambda_n.
\end{aligned}$$

Väite seuraa tästä helposti.

Huomioita

Lause

Symmetrisen ($A^T = A$) (reaalisen) $n \times n$-matriisin ominaisarvot ovat reaalisia ja ominaisvektoreista voidaan muodostaa ortonormaali kanta.

Todistus (reaalisuus): Liittoluvut yhtälöstä $A\boldsymbol{x} = \lambda\boldsymbol{x}$ laskemalla saadaan $A\overline{\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}}$.
Nyt

$$\overline{\boldsymbol{x}}^T A\boldsymbol{x} = \overline{\boldsymbol{x}}^T \lambda\boldsymbol{x} = \lambda\overline{\boldsymbol{x}}^T\boldsymbol{x}.$$

Toisaalta

$$\overline{\boldsymbol{x}}^T A\boldsymbol{x} = (\overline{\boldsymbol{x}}^T A^T)\boldsymbol{x} = (A\overline{\boldsymbol{x}})^T\boldsymbol{x} = \overline{\lambda}\overline{\boldsymbol{x}}^T\boldsymbol{x}.$$

Koska

$$\overline{\boldsymbol{x}}^T\boldsymbol{x} = |x_1|^2 + \ldots + |x_n|^2 > 0,$$

yhtälöstä $\lambda\overline{\boldsymbol{x}}^T\boldsymbol{x} = \overline{\lambda}\overline{\boldsymbol{x}}^T\boldsymbol{x}$ seuraa $\overline{\lambda} = \lambda$.

Todistus (jatkoa)

Osoitetaan sitten, että symmetrisen matriisin erisuuriin ominaisarvoihin kuuluvat ominaisvektorit ovat ortogonaaleja. Olkoon $A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$ ja $A\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_2$. Koska $(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T A \boldsymbol{y}$, on

$$\begin{aligned} & \lambda_1(\boldsymbol{x}_1, \boldsymbol{x}_2) = (\lambda_1 \boldsymbol{x}_1, \boldsymbol{x}_2) = (A\boldsymbol{x}_1, \boldsymbol{x}_2) = (A\boldsymbol{x}_1)^T \boldsymbol{x}_2 = \boldsymbol{x}_1^T A^T \boldsymbol{x}_2 \\ = \ & \boldsymbol{x}_1^T A \boldsymbol{x}_2 = (\boldsymbol{x}_1, A\boldsymbol{x}_2) = (\boldsymbol{x}_1, \lambda_2 \boldsymbol{x}_2) = \lambda_2(\boldsymbol{x}_1, \boldsymbol{x}_2), \end{aligned}$$

siis $\lambda_1(\boldsymbol{x}_1, \boldsymbol{x}_2) = \lambda_2(\boldsymbol{x}_1, \boldsymbol{x}_2)$.
Jos matriisilla on n erisuurta ominaisarvoa, väite seuraa tästä.
Yleinen tapaus saadaan tarkastelemalla avaruuden ns. ortogonaalihajotelmia.

Seuraus:

Symmetrinen matriisi on diagonalisoituva ja diagonaalisuuden välittävä matriisi on ortogonaali ($P^T P = I$). Jos nimittäin $P = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, niin

$$
\begin{aligned}
P^T P &= \begin{pmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \vdots \\ \boldsymbol{x}_n^T \end{pmatrix} (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \begin{pmatrix} \boldsymbol{x}_1^T \boldsymbol{x}_1 & \boldsymbol{x}_1^T \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_1^T \boldsymbol{x}_n \\ \boldsymbol{x}_2^T \boldsymbol{x}_1 & \boldsymbol{x}_2^T \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_2^T \boldsymbol{x}_n \\ \ldots & \ldots & \ddots & \ldots \\ \boldsymbol{x}_n^T \boldsymbol{x}_1 & \boldsymbol{x}_n^T \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_n^T \boldsymbol{x}_n \end{pmatrix} \\
&= I
\end{aligned}
$$

Korrelaatiokerroin

Ongelmanasettelu

- Kaksi satunnaismuuttujaa X ja Y.
- Näiden havainnot $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ ja $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ tapauksissa 1, 2, …, n.
- Selvitettävä, kuinka yleisesti $y_i \approx c \cdot x_i + b$.

Muuttujien siirto

- Merkitään $\mathbf{1} = (1, 1, \ldots, 1)$
- $\mu_{\boldsymbol{x}} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} x_i$ ja $\mu_{\boldsymbol{y}} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} y_i$
- Määritellään $\boldsymbol{x}' = \boldsymbol{x} - \mu_{\boldsymbol{x}}\mathbf{1}$ ja $\boldsymbol{y}' = \boldsymbol{y} - \mu_{\boldsymbol{y}}\mathbf{1}$.
- Tällöin $\boldsymbol{x} = c\boldsymbol{y} + b\mathbf{1} \Leftrightarrow \boldsymbol{x}' = c\boldsymbol{y}'$.
- Lisäksi vektoreiden $\boldsymbol{x}'$ ja $\boldsymbol{y}'$ koordinaattien summa on 0

Korrelaatiokerroin

Riippuvuus

- Yhdensuuntaisuus $\boldsymbol{x}' = c\boldsymbol{y}'$ merkitsee täydellistä (lineaarista) riippuvuutta.
- Vektoreiden $\boldsymbol{x}'$ ja $\boldsymbol{y}'$ kohtisuoruus merkitsee täydellistä (lineaarista) riippumattomuutta.
- Riippuvuuden määrää on siksi luonteva mitata vektoreiden $\boldsymbol{x}'$ ja $\boldsymbol{y}'$ välisellä kulmalla (tai sen kosinilla)
- $$\mathrm{Corr}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\boldsymbol{x}' \cdot \boldsymbol{y}'}{||\boldsymbol{x}'||||\boldsymbol{y}'||}$$