Insinöörimatematiikka: Lineaarialgebra

Mika Hirvensalo
mikhirve@utu.fi

Matematiikan ja tilastotieteen laitos
Turun yliopisto

2024

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 $\mathbf{1} = (1, 1, \ldots, 1)$
- $\mu_{\boldsymbol{x}} = \frac{1}{n} \sum_{i=1}^{n} x_i$ ja $\mu_{\boldsymbol{y}} = \frac{1}{n} \sum_{i=1}^{n} y_i$
- 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)
- $$\mathrm{Corr}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\boldsymbol{x}' \cdot \boldsymbol{y}'}{||\boldsymbol{x}'||||\boldsymbol{y}'||}$$

Korrelaatiokerroin

Huomautus

$$
\begin{aligned}
\mathbf{x}' \cdot \mathbf{y}' &= (\mathbf{x} - \mu_{\mathbf{x}}\mathbf{1}) \cdot (\mathbf{y} - \mu_{\mathbf{y}}\mathbf{1}) \\
&= \mathbf{x} \cdot \mathbf{y} - \mu_{\mathbf{y}}\mathbf{x} \cdot \mathbf{1} - \mu_{\mathbf{x}}\mathbf{1} \cdot \mathbf{y} + \mu_{\mathbf{x}}\mu_{\mathbf{y}}\mathbf{1} \cdot \mathbf{1} \\
&= \mathbf{x} \cdot \mathbf{y} - n\mu_{\mathbf{y}}\mu_{\mathbf{x}} - n\mu_{\mathbf{x}}\mu_{\mathbf{y}} + n\mu_{\mathbf{x}}\mu_{\mathbf{y}} \\
&= \mathbf{x} \cdot \mathbf{y} - n\mu_{\mathbf{x}}\mu_{\mathbf{y}} \\
&= \sum_{i=1}^{n} x_i y_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i .
\end{aligned}
$$

Pääkomponenttianalyysin perusteita

Käsitteitä

- Lähtökohta: Satunnaismuuttujat $X_1, \ldots, X_n$.
- Tavoite: Satunnaismuuttujat $Y_1, \ldots, Y_n$,

$$Y_i = \sum_{k=1}^{n} B_{ik} X_k,$$

  jotka selittäisivät satunnaisvaihtelua "paremmin" kuin alkuperäiset havaintoihin perustuvat muuttujat eivätkä korreloi.
- Lisätavoite: Vain muutamalla muuttujalla $Y_1, \ldots, Y_l$ ($l < n$) voisi selittää suurimman osan datan vaihtelusta.

Käsitteitä

- $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{im})$ datavektori, jossa m havaintoa satunnaismuuttujasta X_i
- Määritelmä: $\mu_i = \mathbb{E}(X_i) = \frac{1}{m}\sum_{k=1}^{m} x_{ik}$.
- Määritelmä: $\boldsymbol{x}_i' = \boldsymbol{x}_i - \mu_i \mathbf{1}$.
- Määritelmä: $\mathrm{Cov}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \mathbb{E}((\boldsymbol{x}_i - \mu_i)(\boldsymbol{x}_j - \mu_j)) = \frac{1}{m}\boldsymbol{x}_i' \cdot \boldsymbol{x}_j'$
- Huom: $\mathrm{Cov}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \frac{||\boldsymbol{x}_i'|| \cdot ||\boldsymbol{x}_j'||}{m} \mathrm{Corr}(\boldsymbol{x}_i, \boldsymbol{x}_j)$
- Määritelmä: Kovarianssimatriisi

$$\Sigma = \frac{1}{m} \begin{pmatrix} \boldsymbol{x}_1' \cdot \boldsymbol{x}_1' & \boldsymbol{x}_1' \cdot \boldsymbol{x}_2' & \ldots & \boldsymbol{x}_1' \cdot \boldsymbol{x}_n' \\ \boldsymbol{x}_2' \cdot \boldsymbol{x}_1' & \boldsymbol{x}_2' \cdot \boldsymbol{x}_2' & \ldots & \boldsymbol{x}_2' \cdot \boldsymbol{x}_n' \\ \ldots & \ldots & \ddots & \ldots \\ \boldsymbol{x}_n' \cdot \boldsymbol{x}_1' & \boldsymbol{x}_n' \cdot \boldsymbol{x}_2' & \ldots & \boldsymbol{x}_n' \cdot \boldsymbol{x}_n' \end{pmatrix}$$

Pääkomponenttianalyysin perusteita

Havainto:

$$
\begin{aligned}
& (c_1 c_2 \ldots c_n) \begin{pmatrix} \boldsymbol{x}_1 \cdot \boldsymbol{x}_1 & \boldsymbol{x}_1 \cdot \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_1 \cdot \boldsymbol{x}_n \\ \boldsymbol{x}_2 \cdot \boldsymbol{x}_1 & \boldsymbol{x}_2 \cdot \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_2 \cdot \boldsymbol{x}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x}_n \cdot \boldsymbol{x}_1 & \boldsymbol{x}_n \cdot \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_n \cdot \boldsymbol{x}_n \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \ddots \\ c_n \end{pmatrix} \\
= \quad & (c_1 \boldsymbol{x}_1 + \ldots + c_n \boldsymbol{x}_n) \cdot (c_1 \boldsymbol{x}_1 + \ldots + c_n \boldsymbol{x}_n) \geq 0.
\end{aligned}
$$

Matriisi on siis symmetrinen ja positiivisemidefiniitti.

Huomioita

Jos merkitään

$$X = \begin{pmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \dots \\ \boldsymbol{x}_n^T \end{pmatrix}$$

on

$$XX^T = \begin{pmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \dots \\ \boldsymbol{x}_n^T \end{pmatrix} (\boldsymbol{x}_1\, \boldsymbol{x}_2\, \dots\, \boldsymbol{x}_n) = \begin{pmatrix} \boldsymbol{x}_1^T\boldsymbol{x}_1 & \boldsymbol{x}_1^T\boldsymbol{x}_2 & \dots & \boldsymbol{x}_1^T\boldsymbol{x}_n \\ \boldsymbol{x}_2^T\boldsymbol{x}_1 & \boldsymbol{x}_2^T\boldsymbol{x}_2 & \dots & \boldsymbol{x}_2^T\boldsymbol{x}_n \\ \dots & \dots & \ddots & \dots \\ \boldsymbol{x}_n^T\boldsymbol{x}_1 & \boldsymbol{x}_n^T\boldsymbol{x}_2 & \dots & \boldsymbol{x}_n^T\boldsymbol{x}_n \end{pmatrix}$$

Huomioita

Olkoon $Y = PX$. Tällöin

$$YY^T = PX(PX)^T = PXX^TP^T = P\Sigma P^T.$$

Onko mahdollista valita P siten että $P\Sigma P^T$ on diagonaalinen?

Pääkomponenttianalyysin perusteita

Huomioita

- Kovarianssimatriisin Σ ominaisarvot ovat aina reaalisia ja ≥ 0
- Kovarianssimatriisi Σ voidaan diagonalisoida $\Sigma = P\Lambda P^{-1}$.
- Matriisi P voidaan koostaa matriisin Σ ominaisvektoreista ja valita ortogonaaliksi, siis $P^T = P^{-1}$.

Määritelmä

Vektoreiden $\boldsymbol{x} = (x_1, x_2, x_3)$ ja $\boldsymbol{y} = (y_1, y_2, y_3)$ ristitulo määritellään

$$
\begin{aligned}
& \boldsymbol{x} \times \boldsymbol{y} \\
= & \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \boldsymbol{i} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - \boldsymbol{j} \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + \boldsymbol{k} \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \\
= & (x_2 y_3 - x_3 y_2)\boldsymbol{i} - (x_1 y_3 - x_3 y_1)\boldsymbol{j} + (x_1 y_2 - x_2 y_1)\boldsymbol{k} \\
= & (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1)
\end{aligned}
$$

Skalaarikolmitulo

Lause: Olkoot $\boldsymbol{x} = (x_1, x_2, x_3)$, $\boldsymbol{y} = (y_1, y_2, y_3)$ ja $\boldsymbol{z} = (z_1, z_2, z_3)$. Tällöin

- $\boldsymbol{x} \cdot (\boldsymbol{y} \times \boldsymbol{z}) = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$
- $\boldsymbol{x} \cdot (\boldsymbol{y} \times \boldsymbol{z}) = \boldsymbol{y} \cdot (\boldsymbol{z} \times \boldsymbol{x}) = \boldsymbol{z} \cdot (\boldsymbol{x} \times \boldsymbol{y})$
- Tulkinta: Suuntaissärmiön tilavuus

Lause

- $\boldsymbol{x} \times \boldsymbol{y} = -\boldsymbol{y} \times \boldsymbol{x}$
- $(\boldsymbol{x} + \boldsymbol{y}) \times \boldsymbol{z} = \boldsymbol{x} \times \boldsymbol{z} + \boldsymbol{y} \times \boldsymbol{z}$
- $(a\boldsymbol{x}) \times \boldsymbol{y} = \boldsymbol{x} \times (a\boldsymbol{y}) = a(\boldsymbol{x} \times \boldsymbol{y})$
- $\boldsymbol{x} \times \boldsymbol{x} = \boldsymbol{0}$
- $\boldsymbol{x} \cdot (\boldsymbol{x} \times \boldsymbol{y}) = 0$ ja $\boldsymbol{y} \cdot (\boldsymbol{x} \times \boldsymbol{y}) = 0$ (ristitulo on kohtisuorassa alkuperäisiin vektoreihin nähden).
- $||\boldsymbol{x} \times \boldsymbol{y}||^2 + (\boldsymbol{x} \cdot \boldsymbol{y})^2 = ||\boldsymbol{x}||^2||\boldsymbol{y}||^2$.
- $||\boldsymbol{x} \times \boldsymbol{y}|| = ||\boldsymbol{x}||||\boldsymbol{y}|| \sin\theta$, missä θ on vektoreiden $\boldsymbol{x}$ ja $\boldsymbol{y}$ välinen kulma.

Avaruuden $\mathbb{R}^3$ tasot

Määritelmä

Olkoon $V \leq \mathbb{R}^n$ aliavaruus ja $\boldsymbol{r} \in \mathbb{R}^n$ jokin vektori. Tällöin sanotaan, että joukko

$$\boldsymbol{r} + V = \{\boldsymbol{r} + \boldsymbol{v} \mid \boldsymbol{v} \in V\}$$

on aliavaruuden V *sivuluokka*.

Lause

$$\boldsymbol{r_1} + V = \boldsymbol{r_2} + V \Leftrightarrow \boldsymbol{r}_1 - \boldsymbol{r}_2 \in V.$$

Määritelmä

Avaruuden $\mathbb{R}^3$ taso on kaksiulotteisen aliavaruuden V sivuluokka.

Avaruuden $\mathbb{R}^3$ tasot

Parametriesitys

$$T = \boldsymbol{r} + L(\boldsymbol{s}_1, \boldsymbol{s}_2) = \{\boldsymbol{r} + c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2 \mid c_1, c_2 \in \mathbb{R}\}$$

- Jos vektorit $\boldsymbol{s}_1$ ja $\boldsymbol{s}_2$ *eivät* ole lineaarisesti riippumattomia, ei aliavaruus $L(\boldsymbol{s}_1, \boldsymbol{s}_2)$ ole kaksiulotteinen eikä tällöin kyseessä ole taso.
- Vektoreita $\boldsymbol{s}_1$ ja $\boldsymbol{s}_2$ sanotaan tason T suuntavektoreiksi ja vektoria $\boldsymbol{r}$ tason T paikkavektoriksi.
- Käyttökelpoinen tason T pisteiden generoimiseksi.
- Parametrimuodosta hankala selvittää, onko $\boldsymbol{x} \in T$.
- Hankala selvittää, ovatko kaksi tasoa samat.

Avaruuden $\mathbb{R}^3$ tasot

Normaalimuoto

$$T = \{\boldsymbol{r} + \boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{R}^3, \boldsymbol{n} \cdot \boldsymbol{x} = 0\} = \{\boldsymbol{x} \in \mathbb{R}^3 \mid (\boldsymbol{x} - \boldsymbol{r}) \cdot \boldsymbol{n} = 0\}$$

Vektoria $\boldsymbol{n}$ kutsutaan tason T normaalivektoriksi.

Koordinaattimuoto

Merkitsemällä $\boldsymbol{n} = (a, b, c)$, $\boldsymbol{x} = (x, y, z)$ ja $\boldsymbol{r} = (r_1, r_2, r_3)$ saadaan yhtälö $(\boldsymbol{x} - \boldsymbol{r}) \cdot \boldsymbol{n} = 0$ muotoon
$a(x - r_1) + b(y - r_2) + c(z - r_3) = 0$, ja edelleen

$$ax + by + cz = d,$$

missä $d = ar_1 + br_2 + cr_3$.

- Käyttökelpoinen kysymyksen $\boldsymbol{x} \in T$? ratkaisemiseksi
- Hankala tason pisteiden generoimiseksi.

Kolme pistettä

Jos avaruuden $\mathbb{R}^3$ pisteet $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ ja $\boldsymbol{p}_3$ eivät ole samalla suoralla, ne määrittävät tason T yksikäsitteisesti.

Huomautus

Pisteet $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ ja $\boldsymbol{p}_3$ ovat samalla suoralla tarkalleen silloin kun $\boldsymbol{p}_3 - \boldsymbol{p}_1$ ja $\boldsymbol{p}_2 - \boldsymbol{p}_1$ ovat lineaarisesti riippuvat.

Määritelmä

Tasojen T_1 ja T_2 välisellä kulmalla tarkoitetaan niiden normaalivektorien välistä kulmaa.

Avaruuden $\mathbb{R}^3$ suorat

Määritelmä

Avaruuden $\mathbb{R}^3$ suora on yksiulotteisen avaruuden sivuluokka.

Parametrimuoto

$$L = \boldsymbol{r} + L(\boldsymbol{s}) = \{\boldsymbol{r} + c\boldsymbol{s} \mid c \in \mathbb{R}\}$$

- Vektoria $\boldsymbol{r}$ kutsutaan suoran L paikkavektoriksi ja vektoria $\boldsymbol{s}$ sen suuntavektoriksi.
- Käyttökelpoinen suoran pisteiden generoimiseksi.
- Hankala kysymyksen $\boldsymbol{x} \in L$ ratkaisemiseksi.

Avaruuden $\mathbb{R}^3$ suorat

Merkintä

$\boldsymbol{r} = (x_0, y_0, z_0)$ ja $\boldsymbol{s} = (a, b, c)$, jolloin

$$L = \{(x_0, y_0, z_0) + t(a, b, c) \mid t \in \mathbb{R}\},$$

jolloin suoran pisteet ovat muotoa
$(x, y, z) = (x_0 + ta, y_0 + tb, z_0 + tc)$, josta $t = \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$

Koordinaattimuoto

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$$