

## Monimuuttujamenetelmät

2. harjoitukset, 23.1.2003

- 2.1. (a) Pisteiden  $A$  ja  $B$  koordinaatit ortogonaalisten akselien  $x_1$  ja  $x_2$  suhteen ovat  $A : (3, -2)'$ ;  $B : (5, 1)'$ . Jos akseleita kierretään  $20^\circ$  vastapäivään, niin mitkä ovat  $A$ :n ja  $B$ :n koordinaatit uusien akselien  $x_1^*$  ja  $x_2^*$  suhteen.
- (b) Pisteiden  $A$  koordinaatit ortogonaalisten akselien  $x_1$  ja  $x_2$  suhteen  $(5, 2)'$ . Kierretään akseleita myötäpäivään kulman  $\theta$  verran. Pisteiden  $A$  koordinaatit uusien kierrettyjen akselien suhteen ovat  $(3.69, 3.93)$ . Laske  $\theta$ .

2.2. Laske pääkomponentit korrelaatiomatriisista

$$\mathbf{R} = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}.$$

2.3. Tarkastellaan korrelaatiomatriisia

$$\boldsymbol{\rho} = \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}.$$

- (a) Anna  $\rho$ :n arvoksi 0.7. Laske pääkomponentit ja niiden varianssit.
- (b) Anna  $\rho$ :lle jokin toinen sopiva arvo (Tarkista, että  $\boldsymbol{\rho}$  on korrelaatiomatriisi). Laske pääkomponentit ja vertaa edellisen kohdan tulokseen.
- (c) Tulkitse 1. pääkomponentti.
- 2.4. Eräessä geneettisessä kokeessa mitattiin 150 naarashiiren paino (grammoina) heti synnytyksen jälkeen neljässä ensimmäisessä peräkkäisessä synnytyksessä. Painojen keskiarvovektori on

$$\mathbf{x} = (39.88, 45.08, 48.11, 49.95)'$$

ja otoskorrelaatiomatriisi

$$\mathbf{R} = \begin{pmatrix} 1.000 & 0.7501 & 0.6329 & 0.6363 \\ 0.7501 & 1.000 & 0.6925 & 0.7386 \\ 0.6329 & 0.6925 & 1.000 & 0.6625 \\ 0.6363 & 0.7386 & 0.6625 & 1.000 \end{pmatrix}.$$

Tee aineistosta pääkomponenttianalyysi ja tulkitse tulos.

- 2.5. Tee FIRMAT-aineistosta pääkomponenttianalyysi.
- Laske muuttujien ja pääkomponenttien väliset korrelaatiot. Piirrä korrelaatiodiagrammit.
  - Pääkomponenttien lausekkeet sekä pääkomponenttipistemäärät.
  - Löytyykö tulkinta?
- 2.6. Tee Pojat-aineistosta 2. poikien pään pituuden ja leveyden pääkomponenttianalyysi (Luentomonisteessa on tehty 1. poikien analyysi). Esitä
- kovarianssimatriisin ominaisarvot ja ominaisvektorit,
  - pääkomponenttien lausekkeet,
  - 10 ensimmäistä pääkomponenttipistemäärää ja
  - piirrä koordinaatistoon 10 ensimmäistä pääkomponenttipistettä ja vastaavat keskistetyt havaintopisteet.
- 2.7. Mikä olisi sopiva pääkomponenttien lukumäärä
- Asanat-aineistossa?
  - Osakkeet-aineistossa (Vertaile eri kriteereitä)?
- 2.8. Määritä pääkomponentit, kun kovarianssimatriisi on

$$\begin{pmatrix} \sigma^2 & \sigma^2 \rho & 0 \\ \sigma^2 \rho & \sigma^2 & \sigma^2 \rho \\ 0 & \sigma^2 \rho & \sigma^2 \end{pmatrix}, \quad -\frac{1}{\sqrt{2}} < \rho < \frac{1}{\sqrt{2}}.$$

Kuinka suuren osan eri komponentit selittävät kokonaisvaihtelusta?