Lineaariset mallit, kl 2015, Harjoitus 2, viikko 13

1. Olkoon matriisi $\mathbf{A}$ $(n \times n)$ ortogonaalinen projektio eli symmetrinen ja idempotentti (eli $\mathbf{A}^2 = \mathbf{A} = \mathbf{A}'$). Osoita, että $\mathbf{A}$:n ominaisarvot ovat nollia ja ykkösiä. (*Vihje*: Ota lähtökohdaksi ominaisvektorit määrittävä yhtälö $\mathbf{Ax} = \lambda\mathbf{x}$; ks. monisteen Liite B.6.)

2. Neliömatriisin jälki on sen diagonaalialkioiden summa eli, jos $\mathbf{A} = [a_{ij}]$ on $n \times n$ matriisi, niin sen jälki (engl. trace) on $\mathsf{tr}\,(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$. Osoita seuraavat tulokset.

(i) $\mathsf{tr}\,(\mathbf{A} + \mathbf{B}) = \mathsf{tr}\,(\mathbf{A}) + \mathsf{tr}\,(\mathbf{B})$, kun $\mathbf{B}$ on $n \times n$ matriisi

(ii) $\mathsf{tr}\,(\mathbf{AB}) = \mathsf{tr}\,(\mathbf{BA})$, kun tulot ovat määriteltyjä.

(iii) $\mathsf{tr}\,(\mathbf{A}) = \mathbf{A}$:n ominaisarvojen summa, kun $\mathbf{A}$ on symmetrinen.

(iv) $\mathbf{A}$:n aste $= \mathbf{A}$:n jälki eli $\mathsf{r}(\mathbf{A}) = \mathsf{tr}\,(\mathbf{A})$, kun $\mathbf{A}$ $(n \times n)$ on ortogonaalinen projektio.

Vihje: Kohdassa (iii) voit käyttää symmetrisen matriisin pääakselihajotelmaa; ks. monisteen Liite B.6 ja kohdassa (iv) lisäksi tehtävää 1 ja lineaarialgebrasta tunnetuksi oletettua tulosta, jonka mukaan matriisin aste ei muutu, kun matriisi kerrotaan vasemmalta tai oikealta jollain kääntyvällä matriisilla.

3. Totea, että yhden selittäjän lineaarisen regressiomallin $Y_1, ..., Y_n \perp\!\!\!\perp$, $Y_i \sim \mathsf{N}\left(\beta_1 + \beta_2 x_i, \sigma^2\right)$ normaaliyhtälöt voidaan esittää ilman matriiseja muodossa

$$\beta_1 n + \beta_2 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \tag{1}$$

$$\beta_1 \sum_{i=1}^{n} x_i + \beta_2 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i. \tag{2}$$

4. (i) Osoita, että edellisen tehtävän yhtälöiden (1) ja (2) ratkaisuna saatavat PNS-estimaatit voidaan lausua muodossa

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \quad \text{ja} \quad \hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x},$$

jossa esimerkiksi $\bar{y} = (y_1 + \cdots + y_n)/n$.

(ii) Esitä kohdassa (i) saatu $\hat{\beta}_2$:n lauseke edelleen muodossa

$$\hat{\beta}_2 = r_{xy} \frac{s_y}{s_x},$$

jossa esimerkiksi $s_y = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2}$ on selitettävän muuttujan havainnoista laskettu keskihajonta ja r_{xy} on havainnoista (y_i, x_i), $i = 1, ..., n$, laskettu korrelaatiokerroin eli

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}.$$

5. (Jatkoa tehtävälle 1.5) Johda tehtävän 1.5 varianssianalyysimallissa parametrien $\mu_1, ..., \mu_p$ PNS-estimaattien lausekkeet normaaliyhtälöiden ratkaisukaavaa käyttäen.