Oulun yliopisto
Matemaattisten tieteiden yksikkö
Tilastollinen hahmontunnistus
Harjoitus 14

1. Tarkastellaan c:n luokan tilannetta, satunnaisvektorin X kovarianssimatriisia Σ ja luennoilla määriteltyjä matriiseja S_b ja S_w. Osoita, että $\Sigma = S_b + S_w$.

2. Tarkastellaan kahden luokan tilannetta. Eräs mahdollinen luokkien etäisyyttä mittaava kriteeri on *Kolmogorovin etäisyys*

$$J_\varphi = \int_{\mathbb{R}^d} |P_1 f_1 - P_2 f_2|.$$

Osoita, että Bayesin virhetodennäköisyydelle pätee

$$e(g^*) = \frac{1}{2}(1 - J_\varphi).$$

3. Osoita, että satunnaisvektorin Y pääkomponentit eivät korreloi keskenään.

4. Olkoon Y D-ulotteinen satunnaisvektori. Haetaan yksikkövektoreita $a_1, \ldots, a_D \in \mathbb{R}^D$, joille

1. $a = a_1$ maksimoi satunnaismuuttujan $a^T Y$ varianssin.
2. $a = a_{i+1}$ maksimoi satunnaismuuttujan $a^T Y$ varianssin ehdolla, että $a^T Y$ ei korreloi satunnaismuuttujien $a_1^T Y, \ldots, a_i^T Y$ kanssa.

Osoita, että tehtävä voidaan ratkaista Y:n kovarianssimatriisin Σ ortonormaalien ominaisvektoreiden avulla ottamalla $a_i = u_i$, missä $\Sigma u_i = \lambda_i u_i$, $i = 1, \ldots, D$ ja $\lambda_1 \geq \cdots \geq \lambda_D$.

5. Karhunen-Loève-muunnoksen johdossa käytetty suure $E\|Y'' - \mu''\|^2$ ei ole ainoa mahdollinen menetetyn informaation mitta. Esimerkiksi eräs luonteva lähestymistapa olisi vaatia, että ortonormaalin joukon $\{v_1, \ldots, v_d\}$ tulee olla sellainen, että rekonstruktiovirhe $E\|Y - Y'\|^2$ minimoituu. Esitä ortonormaali joukko, joka on tämän kriteerin mielessä paras mahdollinen. (Vihje: tämä optimointitehtävä palautuu lauseen 9.6 todistukseen, jossa kovarianssimatriisin Σ voi korvata jollain toisellakin positiivisesti semidefiniitillä matriisilla.)