MACHINE LEARNING, 2. SEMINAR – BAYESSCHE ENTSCHEIDUNGEN

In den unten formulierten Aufgabenstellungen bezeichnet der obere Index die Nummer der Klasse und der untere Index die Nummer der Komponente eines Vektors, z.B. μ_1^2 bedeutet „die erste Komponente des Vektors $\mu^2 \in \mathbb{R}^n$.

Aufgabe 1. Ein Objekt kann sich mit den bekannten a-priori Wahrscheinlichkeiten $p(k)$ in den zwei Zuständen $k = 1,2$ befinden. Die bedingten Wahrscheinlichkeiten für die Merkmale $x \in \mathbb{R}^n$ sind Gaussch verteilt:

$$p(x \mid k) = \frac{1}{(\sqrt{2\pi}\sigma^k)^n} \exp\left[-\frac{\|x-\mu^k\|^2}{2\sigma^{k2}}\right].$$

a) Beide Verteilungen haben *dieselbe* Streuung, d.h. $\sigma^1 = \sigma^2 = \sigma$. Die Kosten für Fehlklassifikationen $C(k,k')$ sind als eine (allgemeine) 2×2 Matrix $C_{kk'}$ angegeben. Leiten Sie die zugehörige Bayessche Strategie ab und geben Sie eine geometrische Interpretation.

b) Beide Verteilungen haben *dasselbe* Zentrum, d.h. $\mu^1 = \mu^2 = \mu$ und *unterschiedliche* Streuungen σ^k. Für dieses Wahrscheinlichkeitsmodell soll der Bayessche Klassifikator konstruiert werden. Die Kostenfunktion für Fehlklassifikationen ist die Deltafunktion $\delta(k \neq k')$. Welche geometrische Form hat die Entscheidungsregel? Wie ergeben sich die Parameter der Entscheidungsregel aus den bekannten Parametern des Wahrscheinlichkeitsmodells $p(k{=}1)$, $p(k{=}2)$, σ^1, σ^2, μ?

Aufgabe 2. Ein Objekt kann sich mit den a-priori Wahrscheinlichkeiten $p(k)$ in den Zuständen $k = 1,2$ befinden. Die bedingten Wahrscheinlichkeiten für das skalare Merkmal $x \in \mathbb{R}$ sind

$$p(x|k) = C \cdot \exp\left[-\tau \cdot |x - \mu^k|\right]$$

(τ und die μ^k, $k = 1,2$ sind reellwertige Parameter).

a) Bestimmen Sie den Normierungskoeffizient C.

b) Wie ergibt sich die Bayessche Entscheidung für den Objektzustand k bei bekannten τ, μ^k und $p(k)$?

c) Geben Sie die Parameter an, bei welchen für eine der Klassen nie entschieden wird. Kann man eine solche Situation auch bei Gausschen bedingten Wahrscheinlichkeitsverteilungen konstruieren?

Aufgabe 3. Ein Objekt (x,k), dessen beobachtbare Merkmale $x = (x_1, x_2) \in \mathbb{R}^2$ Punkte im zweidimensionalen Raum sind, kann sich in zwei Zuständen $k = 1,2$ befinden. Seine Wahrscheinlichkeitsverteilung $p(x,k) = p(k) \cdot p(x \mid k)$ sei bekannt, wobei die bedingten Wahrscheinlichkeiten $p(x \mid k)$ wie folgt definiert sind:

$$p(x \mid k) = C \cdot \exp\Big(-\frac{\max\big[(x_1-\mu_1^k)^2, (x_2-\mu_2^k)^2\big]}{2\sigma^2}\Big),$$

mit den Zentren $\mu^k = (\mu_1^k, \mu_2^k) \in \mathbb{R}^2$.

a) Bestimmen Sie den Normierungskoeffizient C.

b) Wie sieht die Bayessche Entscheidungsstrategie (mit $C(k,k') = \delta(k \neq k')$) aus?

Aufgabe 4. Sei x und y zwei diskrete Zufallsvariablen. Konstruieren Sie eine Wahrscheinlichkeitsverteilung $p(x,y)$ derart, dass die Maximum Marginal Entscheidung

$$x^* = \arg\max_x \sum_y p(x,y)$$
$$y^* = \arg\max_y \sum_x p(x,y)$$

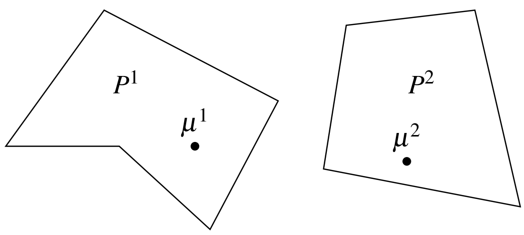
die Wahrscheinlichkeit $p(x^*, y^*) = 0$ hat.

Hinweis: Es genügt solche Variablen x und y zu betrachten, die jeweils nur drei Werte annehmen können. Somit ist die Wahrscheinlichkeitsverteilung $p(x,y)$ eine 3×3 Matrix.

Aufgabe 5. Ein Objekt (x,k), dessen beobachtbare Merkmale $x \in \mathbb{R}^2$ Punkte im zweidimensionalen Raum sind, kann sich a-priori gleichwahrscheinlich in zwei Zuständen $k = 1,2$ befinden. Die bedingten Wahrscheinlichkeiten $p(x \mid k)$ sind Gaussche Verteilungen:

$$p(x \mid k) = \frac{1}{(\sqrt{2\pi}\sigma)^2} \exp\left[-\frac{||x-\mu^k||^2}{2\sigma^2}\right].$$

Von den Zentren μ^k, $k = 1,2$ sei nur bekannt, dass jedes von ihnen in einem Polyeder P^k, $k = 1,2$ liegt (siehe Skizze).



Bei welchen Lagen der Zentren wird das Risiko der Bayesschen Entscheidung maximal?