MUSTERERKENNUNG
5. SEMINAR – EM, DISKRIMINATIVES LERNEN

Aufgabe 1. Die Variable $x \in \mathbb{R}^n$ ist entsprechend der Mischung der Gaussiane verteilt:

$$p(x) = \sum_{i=1}^{m} w_i \cdot \mathcal{N}(x; \mu^i, \sigma^i) = \sum_{i=1}^{m} w_i \cdot \frac{1}{(\sqrt{2\pi}\sigma^i)^n} \exp\Big[-\frac{\|x-\mu^i\|^2}{2\sigma^{i2}}\Big],$$

mit den Zentren μ^i, Streuungen σ^i und Gewichten w_i, $\sum_i w_i = 1$. Schätzen Sie die unbekannten Parameter der Mischung (d.h. μ^i, σ^i und w_i, $i = 1 \dots m$) anhand einer gegebenen Lernstichprobe $L = (x^1, x^2, \dots, x^{|L|})$ nach dem Maximum-Likelihood Prinzip mittels EM-Algorithmus.
Hinweise:

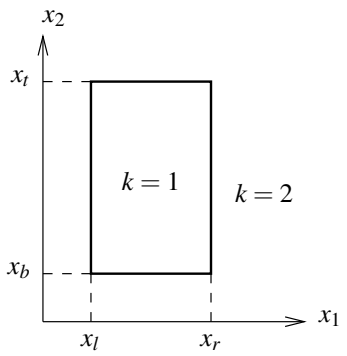
a) Definieren Sie auf eine geeignete Art und Weise die zusätzlichen Variablen α.
b) Wie sind diese bei den bekannten Parametern des Modells zu wählen (E-Schritt)?
c) Wie ist das Likelihood bei fixierten α zu maximieren (M-Schritt)?

Aufgabe 2. Die Merkmale eines Objektes, welches sich in zwei Zuständen $k = 1, 2$ befinden kann, sind Vektoren $x = (x_1, x_2) \in \mathbb{R}^2$. Die Wahrscheinlichkeitsverteilung ist

$$\begin{aligned} p(k=1) &= p(k=2), \\ p(x|k=1) &= C \cdot \exp\left[-\frac{(x_1-\mu_1)^2}{\sigma^2}\right], \\ p(x|k=2) &= C \cdot \exp\left[-\frac{(x_2-\mu_2)^2}{\sigma^2}\right], \end{aligned}$$

mit den Parametern $\mu_1, \mu_2, \sigma \in \mathbb{R}$.

a) Wie sieht die Klasse der Entscheidungsregeln für dieses Wahrscheinlichkeitsmodell aus?
b) Seien die Parameter μ_1 und μ_2 unbekannt. Gegeben sei eine Lernstichprobe $\big((x^1, k^1), \dots, (x^m, k^m)\big)$. Finden Sie die unbekannten Parameter nach dem Maximum Likelihood Prinzip.
c) Finden Sie die Entscheidungsregel, die die Anzahl der Fehlklassifikationen (Empirisches Risiko mit $C(k,k') = \mathbb{1}(k \neq k')$) auf der Lernstichprobe minimiert. Formulieren Sie dafür einen möglichst effizienten Algorithmus.

Aufgabe 3. Bestimmen Sie die Vapnik-Chervonenkis Dimension für die Entscheidungsregel, die im zweidimensionalen Raum ein achsenparalleles Rechteck bildet.

$$e(x) = \begin{cases} 1 & \text{wenn } x_l \leq x_1 \leq x_r \text{ und } x_b \leq x_2 \leq x_t \\ 2 & \text{sonst} \end{cases}$$