TECHNISCHE UNIVERSITÄT MÜNCHEN
LEHRSTUHL FÜR STEUERUNGS- UND REGELUNGSTECHNIK
ORDINARIUS: UNIV.-PROF. DR.-ING./UNIV. TOKIO MARTIN BUSS
EXTRAORDINARIA: UNIV.-PROF. DR.-ING. SANDRA HIRCHE

Name:	Matriculation number: (right-aligned, please)
	...
First name:	Faculty:
	Field of study:

1	2	3				Sum:
						Grade:

Test Exam

Machine Learning in Robotics

Prof. Dongheui Lee

Date:

Lecture room:

Please pay attention to the following details:

- The solutions must be written on different pages. Use, if given, the respective patterns and diagrams.
- Write clearly and structure your solutions with the description's indexing.
- Do not give only the final result. Justifications and approaches for the solutions must also be contained in your answer.
- Cross out clearly the solutions that should not be evaluated.
- This problem booklet must be returned.

1. <u>Problem:</u> EM algorithm and mixture models (Points)

Given the following set of data points in a two-dimensional case, we want to build a probability distribution that represents them. For this aim, we have chosen a mixture model with two Gaussians, $\sum_{k=1}^{2} \pi_k \mathcal{N}(x|\mu_k, \sigma_k^2 I)$, and to optimize its parameters we will use the EM algorithm. The initial random means for the Gaussians are given by μ_1 and μ_2 and the covariance matrices are initially equal, but constrained to the form $\sigma_k^2 I$.

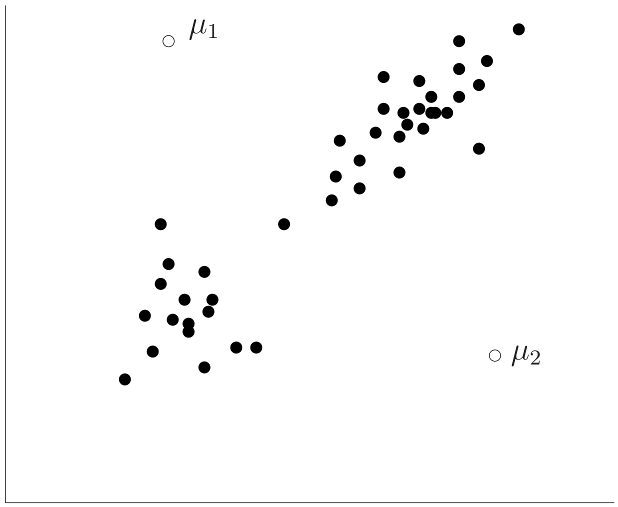

Figure 1

a) Explain the EM algorithm steps and draw on the figure approximately the component parameters (the mean and the shape of the covariance) after one interation.

(Points)

b) Derive the maximum likelihood re-estimation of the component parameters σ_k given a set of data points.

$$\mathcal{N}(x|\mu, \Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} \exp\{-1/2(x-\mu)^T \Sigma^{-1}(x-\mu)\}$$

(Points)

c) Explain the difference between the re-estimation of the mean of the clusters in the K-means algorithm and the EM algorithm. (Points)

2. Problem: Density estimation (Points)

Given an m-dimensional dataset, we desire to estimate the density of the set at a point x. By using non-parametric density estimation, the density at x is given by

$$p(x) = \frac{K}{NV} \tag{1}$$

a) Please explain the variables involved in (1) (K, N, V) and all the assumptions under which this relationship is valid. (Points)

b) Based on (1), explain two different approaches to estimate the density of a dataset at a point and discuss on the convergence of the approaches. (Points)

c) A dataset X is given $X = \{3, 8, 10, 12, 15, 9\}$. Estimate the density of the dataset at $x = 5$ by using Parzen windows with width $h = 4$. How will the probability density be modified if the width of the Parzen window increases? (Points)

Tip: The density estimate by Parzen windows is given by $p(x) = \frac{1}{Nh^m} \sum_{i=1}^{N} \kappa(\frac{x - x^{(i)}}{h})$ where

$$k(u) = \begin{cases} 1, & |u_j| \leq 0.5, \forall j = 1, ..., m \\ 0, & otherwise \end{cases}$$

3. <u>Problem:</u> PCA (Points)

Use proof by induction to show that the linear projection onto an M-dimensional subspace that maximizes the variance of the projected data is defined by the M eigenvectors of the data covariance matrix S, corresponding to the M largest eigenvalues.
It is given that the covariance matrix is given by

$$\mathbf{S} = \frac{1}{N}\sum_{n=1}^{N}(x_n - \tilde{x})(x_n - \tilde{x})^T$$

where $\tilde{x}$ is the sample set mean.