

Review Questions for Course
“Machine Learning I - Probabilistic Unsupervised Learning”
Winter Semester 2015/2016

Basics:

- What types of learning do you know (name three)?
- Give examples.

Elementary Probability Theory:

Concepts:

- probability mass function, probability density function and their properties
- joint probability, conditional probability, independence, conditional independence
- expected value for discrete/continuous random variables
- density estimation (parametric, non-parametric)
- Bayes' theorem, likelihood, prior and posterior probability

Example Questions:

- How can probability functions be manipulated (product rule, sum rule, ...)?
- How can one compute the optimal decision boundary between two labelled sets of input?

Generative Models / Mixture of Gaussians:

Concepts:

- optimal decisions
- latent (hidden) variables
- mixture of Gaussians model
- EM algorithm
- k-means clustering

Example Questions:

- What is a generative model and what are generative models used for?
- What generative models do you know?
- What is the likelihood and what is it good for?
- How does one derive the formula for the data likelihood under generative model X (e.g., mixture of Gaussians or probabilistic PCA)?
- What is the EM algorithm and what does it do?
- Give the form of the EM algorithm which was given in the lecture!
- For what type of data is a mixture of Gaussians model used for?
- Given this type of data (if a picture is shown), how would you or how can you proceed in analysing it?
- What is the E-step of the EM algorithm and what is the M-step?
- How can we understand k-means in the context of a mixture of Gaussians?
- If you want to implement (i.e., program) an algorithm that finds structure in this type of data (if a picture is shown), how would you proceed?

Binary Sparse Coding (BSC)

Concepts:

- multiple causes models
- linear projection from hidden to observed space

Example Questions:

- What is a multiple-cause model and what is a single-cause model?
- What is the generative model of BSC and for what data can it be applied?
- How do you find the optimal parameters for BSC?
- What is the problem with the BSC algorithm?

Principal Component Analysis

Concepts:

- data covariance matrix
- objective function of classical PCA
- PCA as dimensionality reduction
- probabilistic PCA
- EM for probabilistic PCA

Example Questions:

- What is PCA and what is it used for (what can it be used for)?
- What is the generative model of PCA and how does it relate to classical PCA?
- How do you find the optimal parameters for probabilistic PCA?
- What are the differences to BSC or Sparse Coding?
- What is Factor Analysis?

Sparse Coding

Concepts:

- blind source separation problem
- non-Gaussian priors
- generative model of Sparse Coding
- EM for Sparse Coding
- intractable EM steps

Example Questions:

- For what data can you apply Sparse Coding?
- What is the Sparse Coding generative model?
- What is a heavy-tail distribution?
- What is the difference to probabilistic PCA?
- How are Sparse Coding and Independent Component Analysis (ICA) related?
- Why is it difficult to find the maximum likelihood parameters of a generative model with non-Gaussian prior distributions? How do you address this problem?

Hidden Markov Models

Concepts:

- time series analysis problem
- Markov property as low order approximation
- hidden Markov process, transition matrix, emission probabilities

Example Questions:

- What type of inference can be done in HMMs?
- What parameters are there for an HMMs and what are their roles?
- How can we do inference (what are the challenges)?
- What are filtered and smoothed marginals?
- How do you compute the marginals?
- How do you find/learn the parameters of an HMM?
- What is the relation between HMMs and Gaussian Mixture Models?
- How can we do learning (what are the challenges)?

Approximation Schemes / Approximate EM

Concepts:

- approximate inference
- intractable EM steps
- approximations to exact EM
- deterministic approximation schemes
- stochastic approximation schemes / sampling

Example Questions:

- Why are approximation schemes needed for some generative models?
- What generative models do you know that do not require an approximation scheme?
- What generative models do you know that do require an approximation scheme?
- What deterministic approximation schemes do you know?
- What is a stochastic approximation scheme?
- What are the advantages/disadvantages for the different schemes?