Exam questions
2020

Elements of Information Theory

Random vectors and data models

1. Explain the transformation of random vectors for the general case. Explain the properties of linear transformations.
2. The central limit theorem. Explain the moments of sum of i.i.d. random variables with bounded variance. Explain Chebychev inequality and its link to the weak law of large numbers.
3. Explain the difference between the orthogonal, uncorrelated and independent random vectors. Explain the structure of covariance matrix. Exemplify the covariance matrix of Gaussian random vectors.
4. Explain the eigenvalue decomposition of covariance matrices of random vectors. Main properties.
5. Explain the linear transformation of multivariate Gaussian random vectors: mean and covariance matrix after the linear transformation.
6. Explain the linear transformation of multivariate Gaussian random vectors: decorrelation and whitening.
7. Explain the concentration properties of Euclidean norm applied to the random vectors: mean and variance. Why is it useful in practice?
8. Explain the concentration properties of Hamming weight applied to the random vectors: mean and variance. Why is it useful in practice?
9. Explain the difference in application of eigenvalue decomposition and singular value decomposition. To which matrices we can apply these decompositions? Explain the main properties of transformations.

Information theoretic measures

1. Entropy of random variable: definition and properties. Explain the entropy of binary random variables and its main properties.
2. Joint entropy and conditional entropy: definitions and properties. Explain the role of conditioning on entropy. Use the Venn diagram for your justification or any other properties such as chain rule to argument your answer.

3. Joint entropy for the case of multiple random variables. Explain the chain rule application to the joint entropy decomposition. Show and explain the upper bound on the joint entropy. Explain why it is useful in practice?

4. Explain which distribution has the maximum entropy in the class of distributions for all discrete random variables. Justify your answer by proof and examples.

5. Relative entropy: definition and properties. Explain the usage of relative entropy in practice.

6. Cross entropy: definition and properties. Explain the usage of relative entropy in practice. Explain the link to KLD.

7. Mutual information: definition and properties. Explain the link to KLD. The Venn diagram interpretation.

8. Mutual information for multiple random variables. Explain the chain rule decomposition of mutual information. Explain the notion of conditional mutual information. The Venn diagram interpretation.

9. Data processing inequality. The main consequences for practice and its applicability.

10. Differential entropy: definition and properties. The main difference between the entropy and differential entropy.

11. Differential entropy of Gaussian and uniform pdfs. Explain which distribution has maximum differential entropy and under which conditions (intuitive explanation with examples).

12. Explain why Gaussian pdf has maximum entropy among all distributions with the bounded variance (possible ways how to prove it).

13. Differential entropy of multivariate Gaussian vectors. Properties.

14. KLD between Gaussian random variables and random vectors.

15. Approximation of pdfs based on KLD: forward and reverse KLDs. Explain mean and mode seeking modes.

16. KLD and hypothesis testing. Explain the formulation of density ratio test by difference of KLDs. Why is it interesting for practice?

17. Mutual information for univariate Gaussians.

18. Mutual information for multivariate Gaussians.

* we will confirm these questions before May 29, 2020.