

Metodi statistici *

Prof. Lorenzo Finesso¹

¹ISIB-CNR

Istituto di Ingegneria Biomedica

Email: lorenzo.finesso@isib.cnr.it

Calendario: 24 ore, Lunedì e Mercoledì, dalle 10.30 alle 12.30. Prima lezione Lunedì 10 novembre. Lezioni successive: 12, 17, 19, 24, 26 novembre, 1, 3, 9, 10, 15, 17 dicembre. Aula DEI/G (Piano 3, Dipartimento di Ingegneria dell'Informazione, Via Gradenigo 6/a)

Prerequisiti: Basi di Probabilità e di Algebra Lineare.

Tipologia di esame: Homework assignments and take-home exam.

Obiettivi del corso:

The course will present a survey of statistical techniques which are important in applications. The unifying power of the information theoretic point of view will be stressed.

Programma del corso:

Background material. The noiseless source coding theorem will be quickly reviewed in order to introduce the basic notions of entropy and informational divergence (Kullback-Leibler distance) of probability measures. The analytical and geometrical properties of the divergence will be presented.

Divergence minimization problems. Three basic minimization problems will be posed and, on simple examples, it will be shown that they produce the main methods of statistical inference: hypothesis testing, maximum likelihood, maximum entropy.

Multivariate analysis methods. Study of the probabilistic and statistical aspects of the three main methods: Principal Component Analysis (PCA), Canonical Correlations (CC) and Factor Analysis (FA). In the spirit of the course these methods will be derived also via divergence minimization. Time permitting there will be a short introduction to the Non-negative Matrix Factorization method as an alternative to PCA to deal with problems with positivity constraints.

EM methods. The Expectation-Maximization method was introduced as an algorithm for the computation of Maximum Likelihood (ML) estimator with partial observations (incomplete data). We will present the EM method as an alternating divergence minimization algorithm (à la Csiszár Tusnády) and show its application to the ML estimation of Hidden Markov Models.

The MDL method. The Minimum Description Length method of Rissanen will be presented as a general tool for model complexity estimation.

* corso mutuato dalla Scuola di Dottorato in Ingegneria dell'Informazione