מכון ויצמן למדע

THE WEIZMANN INSTITUTE OF SCIENCE

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Machine Learning and Statistics Seminar

Room 1 ,Ziskind Building
on Wednesday, Mar 16, 2016
at 11:15

Amichai Painsky
Tel Aviv University

Generalized Independent Component Analysis Over Finite Alphabets

Abstract:

Independent component analysis (ICA) is a statistical method for transforming an observable multidimensional random vector into components that are as statistically independent as possible from each other. Usually the ICA framework assumes a model according to which the observations are generated (such as a linear transformation with additive noise). ICA over finite fields is a special case of ICA in which both the observations and the independent components are over a finite alphabet. In this work we consider a generalization of this framework in which an observation vector is decomposed to its independent components (as much as possible) with no prior assumption on the way it was generated. This generalization is also known as Barlow's minimal redundancy representation problem [Barlow, '89] and is considered an open problem. We propose several theorems and show that this hard problem can be accurately solved with a branch and bound search tree algorithm, or tightly approximated with a series of linear problems. Moreover, we show that there exists a simple transformation (namely, order permutation) which provides a greedy yet very effective approximation of the optimal solution. We further show that while not every random vector can be efficiently decomposed into independent components, the vast majority of vectors do decompose very well (that is, with a small constant cost), as the dimension increases. The minimal redundancy representation (also known as factorial coding) has many applications, mainly in the fields of neural networks and deep learning. In this work we show that this formulation further applies to large alphabet source coding. Joint work with Prof. Saharon Rosset from the Statistics Department and Prof. Meir Feder from the EE department, Tel Aviv University