

Entropic Embedding in High-Dimensional Biological Systems

Capobianco E

Department of Biomedical Engineering, Boston University, Boston

Motivation

The exploratory analysis and computational modeling of complex high-dimensional systems represent important interdisciplinary research areas for many application fields. Challenging inference problems that are often addressed concern model building and variable selection, clustering and feature extraction, non-linear structure detection and relationship between observed and intrinsic dimensionality, signal extraction, de-noising, and network dynamics reconstruction. The attempt of this work is to leverage on such methodological wealth for dealing with problems relevant to systems and computational biology.

Methods

Genomic data are examples of noisy high-dimensional systems whose observed dynamics may be viewed as mixtures of informative sources with unknown statistical distribution and subject to unknown mixing mechanism. Gene expression values have many interesting features that depend on complex network dynamics. This work presents an application of independent component analysis (ICA) interestingly combined with fuzzy rules, embedding principles and entropic measures.

Results

Entropy and embedding turn out to be very useful tools for controlling the robustness and stability of the decomposition of a system with larger than intrinsic dimensionality in the observed variables, and complex dynamics in the hidden ones. We report results from experiments that show convergence to the intrinsic dimensionality from the observed genomic space by the means of its least dependent decomposition in just a minimal number of salient features.

Contact email: enrico_capobianco@yahoo.com

Supplementary informations

The author is currently on leave and re-locating to a new research Institute. This work was conducted completely at Boston University.