

Dimensionality reduction by unsupervised regression

Prof. Miguel A. Carreira-Perpinan
School of Engineering
University of California, Merced

ABSTRACT: Dimensionality reduction, also known as manifold learning, is the problem of representing a high-dimensional dataset using a small number of dimensions. It is often used as a feature extraction stage for further processing, such as classification or tracking, or for data visualisation. I will present a recent method that, given a high-dimensional dataset, allows to estimate two mappings: from high to low dimension (dimensionality reduction) and from low to high dimension (reconstruction). The method follows an unsupervised regression point of view, where we introduce the unknown low-dimensional coordinates of the data as parameters, and formulate a regularised objective functional of the mappings and low-dimensional coordinates. Alternating minimisation of this functional is straightforward: for fixed low-dimensional coordinates, the mappings have a unique solution; and for fixed mappings, the coordinates can be obtained by finite-dimensional nonlinear minimisation.

Besides, the coordinates can be initialised to the output of a spectral method such as Laplacian eigenmaps or Isomap. The model generalises PCA and several recent methods that learn one of the two mappings but not both; and, unlike spectral methods, our model provides out-of-sample mappings by construction. Experiments with toy and real-world problems show that the model is able to learn mappings for convoluted manifolds, avoiding bad local optima that plague other methods.

BIOGRAPHY: Miguel Carreira-Perpinan is an assistant professor in Electrical Engineering and Computer Science at the University of California, Merced. He received a PhD in Computer Science from the University of Sheffield, UK in 2001, and did postdoctoral work at Georgetown University and the University of Toronto. He is the recipient of an NSF CAREER award for machine learning approaches to articulatory inversion. His research interests lie in machine learning, with applications to speech processing, computer vision and computational neuroscience.