29 Neural Networks for Nonlinear Principal Component Analysis, Independent Component Analysis, and Blind Signal Separation

Erkki Oja, Juha Karhunen, Jyrki Joutsensalo, Aapo Hyvärinen, and Petteri Pajunen

During the past ten years or so, many neural network architectures and corresponding unsupervised learning rules have been introduced for performing standard Principal Component Analysis (PCA); for a review, see for example [4]. These world-wide developments are largely based on the pioneering work of the first two authors, initiated by E. Oja, in the beginning of the 1980's. We showed that relatively simple, neurobiologically justified Hebbian-type learning rules can provide PCA [1] and made a mathematical analysis of the related learning rules [2]. This early work was collected in the book [3]. Our work on neural PCA is now covered in many of the present-day textbooks on artificial neural networks. In fact, neural PCA is often seen as the other major paradigm in unsupervised neural learning, the other one being competitive learning and especially the Self-Organizing Map.

PCA networks have many applications in optimal linear representation of data in pattern recognition, data compression, and signal processing. They are especially suitable for on-line learning in situations, where it is not expedient to collect a data set and compute the PCA in batch mode. However, they have some inherent limitations, too, that have led researchers to study various forms of unsupervised neural learning beyond PCA. Such techniques are often collectively called nonlinear PCA methods. The main advantages of nonlinear PCA networks over standard PCA networks are:

1. The input-output mapping may be nonlinear while standard PCA is able to realize only linear mappings.

2. Higher than second-order statistics are taken into account in processing the input data via nonlinearities at least implicitly. This property is especially important in blind signal processing. Standard PCA is based on the use of covariances. These second-order statistics are sufficient for complete characterization of Gaussian data only, and for standard linear signal processing.

3. Neural realizations become more competitive compared to conventional numerical methods in nonlinear cases, because closed form solutions do not exist. Standard PCA can be determined efficiently using standard eigenvector computation routines.

Nonlinear PCA methods have applications in at least the following areas:

1. Robust PCA. Using suitable nonlinearities which grow less than linearly makes the analysis results more robust against outliers and non-Gaussian noise in the data. See Section "Robust fitting by nonlinear neural units".

2. Blind signal separation and Independent Component Analysis (ICA). See the corresponding section. These methods have applications in telecommunications, sensor array processing, medical signal processing, speech processing, financial time series analysis, and image feature extraction, to mention just a few of the most important application areas. We have applied methods developed in our laboratory to some of these problems with very interesting results; this will be discussed in more detail in later Sections.

3. Clustering of data and neural projection pursuit.

It is noteworthy that Nonlinear PCA methods, especially Independent Component Analysis, which is closely related to the blind signal separation problem, provide often a very meaningful representation of the input data. Furthermore, this representation is data dependent, and emerges in a completely unsupervised manner from the input data. Recently, Independent Component Analysis has been shown to be closely related to certain fundamental information-theoretic principles, such as maximization of output entropies of a neural network, minimization of mutual information, and information maximization. Currently, many leading neural network researchers share the opinion that these principles are fundamental in designing efficient neural network based information processing methods. Together with interesting applications, these facts have over the past few years prompted a great worldwide interest in neural realizations of Independent Component Analysis and related approaches.
We started a research project on ICA in 1994, based on our earlier theories of nonlinear PCA. Our research group is presently one of the leading ones in the world in this area, which is demonstrated for example by the many invited talks by Prof. Oja and Dr. Karhunen, invitations to international co-operation, visits etc.
In the next few sections of this report, sub-projects of the research effort in Nonlinear PCA and ICA neural networks will be covered in more detail by the members of the research group.

References

[1] E. Oja. A Simplified Neuron Model as a Principal Component Analyzer. *J. of Mathematical Biology*, vol. 16, 1982, pp. 267-273.

[2] E. Oja and J. Karhunen. On Stochastic Approximation of the Eigenvectors and Eigenvalues of the Expectation of a Random Matrix. *J. of Math. Analysis and Applications*, vol. 106, 1985, pp. 69-84.

[3] E. Oja. Subspace Methods of Pattern Recognition. *Research Studies Press and J. Wiley*, 1983.

[4] K. I. Diamantaras and S. Y. Kung. Principal Component Neural Networks: Theory and Applications. *J. Wiley*, 1996.