Nonmetric Multidimensional Scaling with Neural Networks

Michiel C. van Wezel, Walter A. Kosters, Peter van der Putten, and Joost N. Kok

Leiden Institute of Advanced Computer Science (LIACS)
Universiteit Leiden
P.O. Box 9512, 2300 RA Leiden, The Netherlands.
{michiel,kosters,putten,joost}@liacs.nl

Abstract In this paper we present a neural network for nonmetric multidimensional scaling. In our approach, the monotone transformation that is a part of every nonmetric scaling algorithm is performed by a special feedforward neural network with a modified backpropagation algorithm. Contrary to traditional methods, we thus explicitly model the monotone transformation by a special purpose neural network. The architecture of the new network and the derivation of the learning rule are given, as well as some experimental results. The experimental results are positive.

1 Introduction

This paper concerns visualization of multidimensional data using neural nonmetric multidimensional scaling. Generally stated, multidimensional scaling (abbreviated MDS) is a collection of techniques for embedding dissimilarity data in a space with a chosen dimensionality. The embedding is often used for the purpose of data visualization and exploratory data analysis. In this sense, MDS is a competitor for other data visualization techniques, such as a Kohonen neural network and principal component analysis. Traditional MDS techniques are subdivided into metric MDS, where the dissimilarities between objects are assumed to be proportional to Euclidean distances, and nonmetric MDS, where the dissimilarities are only assumed to be monotonically related to Euclidean distances. In the case where the dissimilarities represent, e.g., distances between the capitals of the European countries, the Euclidean distance assumption is realistic. However, in the case of dissimilarities between soda brands that have been reported by a panel of test persons, nonmetric MDS seems more appropriate.

Although traditional multidimensional scaling is a well established pattern recognition technique (see, e.g., [12][10][7]), to our knowledge this subject has not received a lot of attention from researchers in the field of neural networks. In particular, all publications concerning neural networks for multidimensional scaling that are known to us concern metric multidimensional scaling. In [9] a simple neural network is given for metric multidimensional scaling. This neural
network merely performs a gradient descent on the cost function, which carries
the risk of getting stuck in local minima in the error function. To prevent this, in
Klöck and Buhmann apply annealing methods from statistical mechanics to
the metric MDS problem. No neural algorithms have been applied to nonmetric
MDS so far.

This paper is organized as follows. First, an exact problem statement is pre-

tered for multidimensional scaling in Section 2. Next, the proposed feedforward
neural network for nonmetric multidimensional scaling and the corresponding
learning rule are given in Section 3. After that, the results of some experiments
are given in Section 4 and finally Section 5 gives conclusions and a summary.
This paper does not give an in-depth treatment of multidimensional scaling in

general. The interested reader is referred to [2,3,10] for more information on the
subject.

2 Multidimensional Scaling: Problem Statement

It is impossible to give a definition of the term multidimensional scaling that
covers the general use of this term and is yet formal enough to be precise.
Some people use the term for a very broad class of data analysis techniques and
problems including cluster analysis and factor analysis (see, e.g., [7]), while others
use a more narrow definition, and refer to MDS as a class of techniques used
to develop a spatial representation of proximity data. The distinction between
a broad and a narrow definition is also made in a useful taxonomy introduced
in a paper by Carroll and Arabie [1]. This paper is concerned with MDS in the
narrow sense: finding a spatial representation of objects based on proximity data.

Using this narrow definition, there are two “kinds” of MDS: metric MDS
and nonmetric MDS. Before we are able to describe these variants we need some
terminology, which is often used in MDS literature and is also adopted in this
paper.

**Dissimilarities:** In both metric and nonmetric MDS the starting point is a
matrix $\delta$ of dissimilarities, of which an element $\delta_{ij}$ denotes the dissimilarity
between the objects $i$ and $j$. The number of objects is denoted by $n$.

**Embedding:** An embedding of the objects in Euclidean space. The coordinates
of an object $i$ in this embedding are denoted by $x_i$. The dimensionality of the
embedding space is denoted by $m$, so $x_i = (x_{i1}, \ldots, x_{im})^T$, where $^T$ denotes
transpose.

**Distances:** The (real) Euclidean distance between objects $i$ and $j$ is denoted
by $\hat{d}_{ij}$. The distance between the estimates for the spatial representations $x_i$
and $x_j$ of objects $i$ and $j$ is denoted by $d_{ij} = \| x_i - x_j \|$, where $\| \cdot \|$ denotes
the Euclidean norm. Collectively the distances are denoted by matrix $d$.

**Disparities:** These quantities are used in nonmetric scaling. Disparities $\hat{\delta}$ are as
close as possible to distances between the corresponding coordinate estimates
$d$ but with the restriction that they are monotonically related to the original
dissimilarity data $\delta$. 
