Nonlinear Discriminative Data Visualization

Kerstin Bunte, Barbara Hammer and Michael Biehl

Intelligent Systems Group
Mathematics and Computing Science
University of Groningen
The Netherlands

April 20, 2009
Outline

Introduction

Limited Rank Matrix Learning Vector Quantization

Charting

Experiments

Summary/Outlook
Introduction

- Reduce dimension by eliminating redundancies
- Visualize data to cooperate with human capabilities
- Use label information to bypass typical unsupervised preservation problems
- Keep most relevant information for the classification task during dimension reduction
Limited Rank Matrix LVQ

Training data: $\{\bar{x}_i, y_i\}_{i=1}^S \in \mathbb{R}^N \times \{1, \ldots, C\}$

Limited Rank Matrix LVQ (LiRaM LVQ):
- Prototype based classifier, extension of LVQ
Limited Rank Matrix LVQ (LiRaM LVQ):

- Prototype based classifier, extension of LVQ
- Modified Euclidean distance, can take correlations of features into account

\[ d(\vec{w}_j, \vec{x}) = (\vec{x} - \vec{w}_j)^T \Omega_j^T \Omega_j (\vec{x} - \vec{w}_j) \]

- Adapt local matrices \( \Omega_j \in \mathbb{R}^{M \times N} \) during training (minimize a cost function by gradient descent)
Limited Rank Matrix LVQ (LiRaM LVQ):

- Prototype based classifier, extension of LVQ
- Modified Euclidean distance, can take correlations of features into account

\[ d(\bar{w}_j, \bar{x}) = (\bar{x} - \bar{w}_j)^T \Omega_j^T \Omega_j (\bar{x} - \bar{w}_j) \]

- Adapt local matrices \( \Omega_j \in \mathbb{R}^{M \times N} \) during training (minimize a cost function by gradient descent)
- Provide \( k \) local linear transformations of the data

\[ P_j(\bar{x}_i) = \Omega_j (\bar{x}_i - \bar{w}_j) \]
Charting [Brand 2003, Teh 2003]

Charting:

- Nonlinear combination of local linear projections
- Uses responsibilities $r_{ji}$ of projection $P_j$ for data point $\vec{x}_i$ with $\sum_j r_{ji} = 1$, for example:

$$r_{ji} \propto \exp((\vec{x}_i - \vec{w}_j)^\top \Omega_j^\top \Omega_j (\vec{x}_i - \vec{w}_j) / \sigma_j)$$

with an appropriate bandwith $\sigma_j$

- Find affine transformations of local coordinates with matching overlapping regions
- Analytical solution based on a generalized eigenvalue problem, see [Brand 2003, Teh 2003]
Experiment 1: artificial dataset

Original data

\[ \in \mathbb{R}^{10} \]
Experiment 1: artificial dataset

Original data \( \in \mathbb{R}^{10} \)

LiRaM LVQ/Charting
Experiment 1: artificial dataset

Original data

LiRaM LVQ/Charting

\[ \in \mathbb{R}^{10} \]

C1
C2
w

Local projection 2

Local projection 5
Experiment 1: artificial dataset

Visualization of the artificial data set with four different unsupervised methods.
Experiment 2: Letter recognition

Supervised visualization of the UCI Letter recognition data (26 class problem, \( N = 16 \))
Experiment 2: Letter recognition

Unsupervised visualization of the UCI Letter data
Experiment 3: Segmentation

Supervised visualization of the UCI segmentation data (7 class problem, $N = 16$)
Experiment 3: Segmentation

Unsupervised visualization of the UCI segmentation data (7 class problem, $N = 16$)
Summary/Outlook

The combination of LiRaM LVQ and Charting:

- enables the dimension reduction with incorporating class label information
- provides a nonlinear embedding
- has linear complexity in the number of examples
- shows promising results on artificial and real world data sets
Summary/Outlook

The combination of LiRaM LVQ and Charting:
- enables the dimension reduction with incorporating class label information
- provides a nonlinear embedding
- has linear complexity in the number of examples
- shows promising results on artificial and real world data sets

Outlook:
- Prototypes could be used to compress visual information
- Merge the charting step into the training process
Thank you for your attention! 😊