Non–Linear Dimensionality Reduction Techniques

Dilip Kumar Pratihar

Indian Institute of Technology, Kharagpur, India

INTRODUCTION

Most of the complex real-world systems involve more than three dimensions and it may be difficult to model these higher dimensional data related to their input-output relationships, mathematically. Moreover, the mathematical modeling may become computationally expensive for the said systems. A human being can visualize only up to three dimensions (3-D). So, any system involving more than 3-D cannot be visualized. To overcome this difficulty, higher dimensional data are generally mapped into either 2-D or 3-D, for visualization and ease of modeling. Dimensionality reduction techniques are nothing but the mapping methods, with the help of which the higher dimensional data can be mapped into the lower dimension after ensuring a reasonable accuracy. It is to be noted that the precision of modeling depends on the said accuracy in mapping. Thus, it is worthy to study the dimensionality reduction techniques.

BACKGROUND

A number of dimensionality reduction techniques are available in the literature; those are classified into two groups, namely linear and non-linear methods (Siedlecki, Seidlecka & Sklansky, 1988; Konig, 2000; Pratihar, Hayashida & Takagi, 2001). In linear methods, each of the lower dimensional components is considered as a linear combination of the higher dimensional components. These methods include principal component analysis (Jolliffe, 1986; Jackson, 1991), projection pursuit mapping (Crawford & Fall, 1990), factor analysis (Mardia, Kent & Bibby, 1995), independent component analysis (Cardoso, 1999), and others. On the other hand, there are some non-linear mapping methods in use, in which the relationships among the lower dimensional and higher dimensional components are non-linear in nature. The non-linear methods are further classified into two sub-groups, namely distance preserving and topology preserving techniques. Distance preserving techniques include Sammon's non-linear mapping (NLM) (Sammon, 1969), VISOR algorithm (Konig, Bulmahn & Glessner, 1994), triangulation method (Konig, 2000), and others, whereas the techniques like self-organizing map (SOM) (Kohonen, 1995), topology preserving mapping of sample sets (Konig, 2000) are known as the topology preserving tools. Two other dimensionality reduction techniques, namely locally linear embedding (LLE) (Roweis & Saul, 2000) and Isomap (Tenenbaum, de Silva & Langford, 2000) proposed in 2000, have also gained the popularity. Dimensionality reduction problem has also been solved using an optimization tool like genetic algorithm (GA) (Raymer, Punch, Goodman, Kuhn & Jain, 2000). In this connection, it is important to mention that more recently, the author of this chapter along with one of his students have proposed a GA-like approach for dimensionality reduction (Dutta & Pratihar, 2006). Another dimensionality reduction technique, namely HyperMap has been proposed by An, Yu, Ratanamahatana & Phoebe Chen (2007), in which the limitation of Euclidean space has been overcome by representing an axis as a line, a plane or a hyperplane. Moreover, an interactive technique has been developed to vary the weights associated with each hyperaxis for ease of visualization. The present chapter deals with some of the non-linear dimensionality reduction techniques (also known as mapping methods). It is important to mention that a particular mapping method may differ from others in terms of accuracy in mapping, visualization capability and computational complexity. An ideal dimensionality reduction technique is one, which can carry out the mapping from a higher dimensional space to a lower dimensional space accurately at a low computational cost and at the same time, can also provide with the widely distributed mapped data suitable for visualization.

MAIN FOCUS

In this section, the principles of some of the non-linear dimensionality reduction techniques have been explained.

Sammon's Nonlinear Mapping (NLM)

It is a distance preserving technique of mapping. Here, the error in mapping from a higher dimensional space to a lower dimensional space/plane is minimized using a gradient descent method (Sammon, 1969).

Let us consider N points in an L-dimensional (L>3) space represented by X_i, where $i = 1,2,3,....,N$. The aim is to map these N-points from L-dimensional space to 2-D plane or 3-D space. Let us also suppose that the mapped data in 2-D plane or 3-D space are represented by Y_i, where $i = 1,2,....,N$. The scheme is shown in Figure 1.

The N points in L-dimensional space are indicated as follows:

$$X_1 = \begin{bmatrix} x_{11} \\ x_{12} \\ . \\ . \\ . \\ x_{1L} \end{bmatrix}, X_2 = \begin{bmatrix} x_{21} \\ x_{22} \\ . \\ . \\ . \\ x_{2L} \end{bmatrix}, ..., X_N = \begin{bmatrix} x_{N1} \\ x_{N2} \\ . \\ . \\ . \\ x_{NL} \end{bmatrix}.$$

Similarly, the N-points in 2-D plane or 3-D space can be expressed like the following.

$$Y_1 = \begin{bmatrix} y_{11} \\ . \\ . \\ . \\ . \\ y_{1D} \end{bmatrix}, Y_2 = \begin{bmatrix} y_{21} \\ . \\ . \\ . \\ . \\ y_{2D} \end{bmatrix}, ..., Y_N = \begin{bmatrix} y_{N1} \\ . \\ . \\ . \\ . \\ y_{ND} \end{bmatrix},$$

where D indicates the dimension of the lower dimensional space.

This technique consists of the following steps:

- Initially, generate N points in 2-D plane at random, corresponding to N points in L-D space.
- Determine the mapping error as follows: Let d_{ij}^* be the Euclidean distance between two points X_i and X_j in L-dimensional space and d_{ij} represents the Euclidean distance between the corresponding two mapped points Y_i and Y_j, in 2-D plane. For an error-free mapping, the following condition has to be satisfied:

$$d_{ij}^* = d_{ij}.$$

However, the mapping may not be error-free and the mapping error in m^{th} iteration E(m) can be determined mathematically as follows:

$$E(m) = \frac{1}{C} \sum_{i=1}^{N} \sum_{j=1(i<j)}^{N} \frac{[d_{ij}^* - d_{ij}(m)]^2}{d_{ij}^*}$$

where $C = \sum_{i=1}^{N} \sum_{j=1(i<j)}^{N} d_{ij}^*$ and

$$d_{ij}(m) = \sqrt{\sum_{k=1}^{D} [y_{ik}(m) - y_{jk}(m)]^2}.$$

- The aforementioned mapping error can be minimized using the steepest descent method, in which the search follows a rule given as:

$$y_{pq}(m+1) = y_{pq}(m) - (MF)\Delta_{pq}(m),$$

where $y_{pq}(m)$ and $y_{pq}(m+1)$ represent the $q-th$ dimension of point p in 2-D at $m-th$ and $(m+1)-th$ iterations, respectively, MF is a magic factor representing the step length and it varies in the range of 0.0 to 1.0, and:

Figure 1. Mapping from L-dimensional space to 2-D plane or 3-D space using NLM (Dutta & Pratihar, 2006).
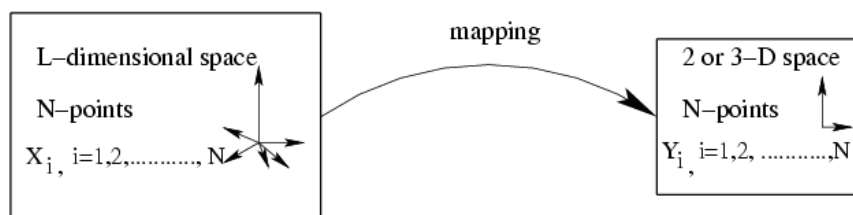
Related Content

Multi-Instance Learning with MultiObjective Genetic Programming

Amelia Zafra (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1372-1379).*

www.irma-international.org/chapter/multi-instance-learning-multiobjective-genetic/11000

Database Security and Statistical Database Security

Edgar R. Weippl (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 610-616).*

www.irma-international.org/chapter/database-security-statistical-database-security/10884

Quantization of Continuous Data for Pattern Based Rule Extraction

Andrew Hamilton-Wrightand Daniel W. Stashuk (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1646-1652).*

www.irma-international.org/chapter/quantization-continuous-data-pattern-based/11039

Deep Web Mining through Web Services

Monica Maceliand Min Song (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 631-637).*

www.irma-international.org/chapter/deep-web-mining-through-web/10887

Data-Driven Revision of Decision Models

Martin Žnidaršic, Marko Bohanecand Blaž Zupan (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 617-623).*

www.irma-international.org/chapter/data-driven-revision-decision-models/10885