Indian J.Sci.Res. 17(2): 344-352, 2018 ISSN: 0976-2876 (Print)
ISSN: 2250-0138(Online)

DATAANALYSIS PROCEDURE FOR DISCOVERING RELATIONSHIPS IN A GRAPH WITH RESPECT TO NETWORKING

[1]Chandrasekaran K Arun, [2]Dr. Vilas M Ghodki, [3]Dr.S.B.Kishor
[1]Research Scholar, Gondwana University, Mahaharashtra, India
[2]Associate Professor, Dept. of Computer Science , J.B. College, Wardha, Maharashtra, India
[3]Head, Dept. of Computer Studies and Research, S.P College, Chandrapur, Maharsashtra, India
E-Mail :[1]bmk_17@yahoo.com, [2]vilasghodki@gmail.com,[3]s.b.kishor.spc@gmail.com

Abstract - This paper precisely proposes a link-analysis based technique allowing to discover relationships existing between nodes in a computer network or, more generally, a graph. More specifically, this work is based on a random-walk through the database defining a Markov chain having as many states as nodes in the computer network. Suppose, for instance, we are interested in analyzing the relationships between nodes in a computer network, a two-step procedure is developed in analyzing the relationships. First, a much smaller, reduced, Markov chain, only containing the nodes but preserves the main characteristics of the initial chain, is extracted by stochastic complementation. For extracting the reduced Markov by stochastic complementation, an efficient algorithm is proposed. Secondly, the reduced chain is analyzed by, for instance, projecting the states in the subspace spanned by the right eigenvectors of the transition matrix called the basic diffusion map, or by computing a kernel principal-component analysis on a diffusion-map kernel computed from the reduced graph and visualizing the results. Indeed, a valid graph kernel based on the diffusion-map distance, extending the basic diffusion map to directed graphs, is introduced.

Keywords - *Diffusion Map , Stochastic complementation, Feature Redundancy*

I. INTRODUCTION

Wireless sensor networks (WSNs) are being used for diverse applications such as low cost area monitoring, environment monitoring, industrial and machine health monitoring, structural monitoring and military surveillance [1], [2]. In these applications, WSNs generate a large amount of data in the form of streams. In recent times, data mining techniques have been used to extract useful knowledge from WSN data [3], through discovering relationships among the sensor nodes which are known as behavioral patterns [4]. More recently, research has been focused to mine different types of behavioral patterns, e.g., sensor association rules [5], [6], [9] from stored (static) sensor data, context association rules [10] from sensor data stream, associated sensor patterns [7] and regularly frequent sensor patterns [8] from static as well as stream data. Traditional statistical, machine learning, pattern recognition, and data mining approaches [28] usually assume a random sample of independent objects from a single relation. Many of these techniques have gone through the extraction of knowledge from data, almost always leading, in the end, to the classical double-entry tabular format, containing features for a sample of the population. These features are therefore used in order to learn from the sample, provided that it is representative of the population as a whole. However, real-world data coming from many fields such as World Wide Web, marketing, social networks, or biology [16] are often multi relational and interrelated. The work recently performed in statistical relational learning [22], aiming at working with such data sets, incorporates research topics, such as link analysis [36] web mining [1],[9], social network analysis [8], or graph mining[11]. All these research fields intend to find and exploit links between objects which could be of various types and involved in different kinds of relationships. On the other hand, when dealing with a starschema database , this two-step procedure reduces to multiple correspondence analysis. The proposed methodology therefore extends correspondence analysis to the analysis of a relational database. In short, this paper has three main contributions:A two-step procedure for analyzing weighted graphs or relational databases is proposed. .

- It is shown that the suggested procedure extends correspondence analysis.

[1]Corresponding Author

alizing the results. Indeed, a valid graph kernel based on the diffusion-map distance, extending the basic diffusion map to directed graphs, is introduced.

Keywords - *Diffusion Map , Stochastic complementation, Feature Redundancy*

I. INTRODUCTION

Wireless sensor networks (WSNs) are being used for diverse applications such as low cost area monitoring, environment monitoring, industrial and machine health monitoring, structural monitoring and military surveillance [1], [2]. In these applications, WSNs generate a large amount of data in the form of streams. In recent times, data mining techniques have been used to extract useful knowledge from WSN data [3], through discovering relationships among the sensor nodes which are known as behavioral patterns [4]. More recently, research has been focused to mine different types of behavioral patterns, e.g., sensor association rules [5], [6], [9] from stored (static) sensor data, context association rules [10] from sensor data stream, associated sensor patterns [7] and regularly frequent sensor patterns [8] from static as well as stream data. Traditional statistical, machine learning, pattern recognition, and data mining approaches [28] usually assume a random sample of independent objects from a single relation. Many of these techniques have gone through the extraction of knowledge from data, almost always leading, in the end, to the classical double-entry tabular format, containing features for a sample of the population. These features are therefore used in order to learn from the sample, provided that it is representative of the population as a whole. However, real-world data coming from many fields such as World Wide Web, marketing, social networks, or biology [16] are often multi relational and interrelated. The work recently performed in statistical relational learning [22], aiming at working with such data sets, incorporates research topics, such as link analysis [36] web mining [1],[9], social network analysis [8], or graph mining[11]. All these research fields intend to find and exploit links between objects which could be of various types and involved in different kinds of relationships. On the other hand, when dealing with a starschemadatabase , this two-step procedure reduces to multiple correspondence analysis. The proposed methodology therefore extends correspondence analysis to the analysis of a relational database. In short, this paper has three main contributions:A two-step procedure for analyzing weighted graphs or relational databases is proposed. .

- It is shown that the suggested procedure extends correspondence analysis.
- A kernel version of the diffusion map distance, applicable to directed graphs, is introduced.

The paper is organized as follows: the basic diffusion map distance and its natural kernel on a graph in Section II. In Section IV we some experimental resultsinvolving several data sets.

II.THE DIFFUSION MAP DISTANCE AND ITSNATURAL KERNEL MATRIX

In this section, the basic diffusion map distance [24] is briefly reviewed and some of its theoreticaljustifications are detailed. Then, a natural kernel matrix isderived from the diffusion map distance, providing ameaningful similarity measure between nodes.

A. The Diffusion Map Distance

In our two-step procedure, a diffusion map projection,based on the so-called diffusion map distance, will beperformed after stochastic complementation. Now, sincethe original definition of the diffusion map distance dealsonly with undirected, aperiodic, Markov chains, it will firstbe assumed in Section 2 that the reduced Markov chain,obtained after stochastic complementation, is indeed undirected,aperiodic, and connected—in which case thecorresponding random walk defines an irreducible reversibleMarkov chain. Notice, that it is not required that the

original adjacency matrix is irreducible and reversible; theseassumptions are only required for the reduced adjacencymatrix obtained after stochastic complementation.The original derivation of the diffusion map, introducedindependently by Nadler et al., and Pons and Latapy [22],[13], is detailed in but other interpretationsof this mapping appeared in the literature . For an application of the basicdiffusion map to dimensionality reduction, see [35].Since P is aperiodic, irreducible, and reversible, it is wellknown that all the eigenvalues of P are real and theeigenvectors are also real [7]. Moreover, allitseigenvalues $\in [-1, +1]$, and the eigenvalue 1 has multiplicityone [7]. With these

[1]Corresponding Author

ISSN: 0976-2876 (Print)
ISSN: 2250-0138(Online)

assumptions, Nadler et al. andPons and Latapy [42], [43], [46], [47] proposed to use asdistance between states i and j

$$d_{ij}^2(t) = \sum_{k=1}^{n} \frac{(x_{ik}(t) - x_{jk}(t))^2}{\pi_k} \quad (1)$$

$$\propto (\mathbf{x}_i(t) - \mathbf{x}_j(t))^{\mathrm{T}} \mathbf{D}^{-1} (\mathbf{x}_i(t) - \mathbf{x}_j(t)), \quad (2)$$

since, for a simple random walk on an undirected graph,the entries of the steady-state vector $\boldsymbol{\pi}$ are proportional (the $\propto$ sign) to the generalized degree of each node (the total ofthe elements of the corresponding row of the adjacencymatrix [28]). This distance, called the diffusion mapdistance, corresponds to the sum of the squared differencesbetween the probability distribution of being in any stateafter t transitions when starting (i.e., at time t = 0) from twodifferent states, state i and state j. In other words, twonodes are similar when they diffuse through the network—and thus influence the network—in a similar way. This is anatural definition which quantifies the similarity betweentwo states based on the evolution of the states' probabilitydistribution. Of course, when $i = j, d_{ij}(t) = 0$ .Nadler et al. [22] showed that this distance measurehas a simple expression in terms of the right eigenvectorsof P:

$$d_{ij}^2(t) = \sum_{k=1}^{n} \lambda_k^{2t} (u_{ki} - u_{kj})^2, \quad (3)$$

where $u_{ki} = [\mathbf{u}_k]_i$ is component i of the kth right eigenvector, $\mathbf{u}_k$, of P and λ_k is its corresponding eigenvalue. Asusual, the λ_k are ordered by decreasing modulus, so that thecontributions to the sum in (3) are decreasing with k. Onthe other hand, $\mathbf{x}_i(t)$ can easily be expressed in thespace spanned by the left eigenvectors of $\mathbf{P}$, the $\mathbf{v}_k$,

$$\mathbf{x}_i(t) = (\mathbf{P}^{\mathrm{T}})^t \mathbf{e}_i = \sum_{k=1}^{n} \lambda_k^t \mathbf{v}_k \mathbf{u}_k^{\mathrm{T}} \mathbf{e}_i = \sum_{k=1}^{n} (\lambda_k^t u_{ki}) \mathbf{v}_k, \quad (4)$$

where $\mathbf{e}_i$ is the ith column of $\mathbf{I}$,

$\mathbf{e}_i = [0, \ldots, 0, 1, 0, \ldots, 0]^{\mathrm{T}}$ with the single 1 in position i. The resulting mapping aims to represent each state i in an-dimensional euclidean space with coordinates $(|\lambda_2^t|u_{2i}, |\lambda_3^t|u_{3i}, \ldots, |\lambda_n^t|u_{ni})$, as in (4). Dimensions areordered by decreasing modulus, $|\lambda_k^t|$. This original mappingintroduced by Nadler and coauthors will be referred to asthe basic diffusion map in this paper, in contrast with thediffusion map kernel (KDM) that was introduced inSectionII.The weighting factor, $\mathbf{D}^{-1}$, in (2) is necessary to obtain(3), since the $\mathbf{v}_k$ are not orthogonal. Instead, it can easily beshown that we have $\mathbf{v}_i^{\mathrm{T}} \mathbf{D}^{-1} \mathbf{v}_j = \delta_{ij}$, which aims to redefinethe inner product as $\langle x, y \rangle = \mathbf{x}^{\mathrm{T}} \mathbf{D}^{-1} \mathbf{y}$, where the metric ofthe space is $\mathbf{D}^{-1}$ [7].Notice also that there is a close relationship betweenspectral clustering (the mapping provided by the normalizedLaplacian matrix; see, for instance, [15], [35]) andthe basic diffusion map. Indeed, a common embedding ofthe nodes consists of representing each node by thecoordinates of the smallest nontrivial eigenvectors (correspondingto the smallest eigenvalues) of the normalizedLaplacian matrix, $\tilde{\mathbf{L}} = \mathbf{D}^{-1/2} \tilde{\mathbf{L}} \mathbf{D}^{-1/2}$. More precisely, if $\mathbf{u}_k$isthe$k$th largest right eigenvector of the transition matrix Pand $\tilde{\mathrm{l}}_k$ is the kth smallest nontrivial eigenvector of thenormalizedLaplacian matrix

A subtle, still important, difference between this mappingand the one provided by the basic diffusion map concerns theorder in which the dimensions are sorted. Indeed, for the basicdiffusion map, the eigenvalues of the transition matrix P areordered by decreasing modulus value. For this spectralclustering model, the eigenvalues are sorted by decreasingvalue (and not modulus), which can result in a differentrepresentation if P has large negative eigenvalues. Thisshows that the mappings provided by spectral clustering andby the basic diffusion map are closely related.Notice that at least three other justifications of thiseigenvector-based mapping appeared before in the literature,and are briefly reviewed here. It has been shownthat the entries of the subdominant right eigenvector of thetransition matrix P of an aperiodic, irreducible, reversible,Markov chain can be interpreted as a relative distance to its"stationary distribution" . This distance may be regarded as an indicatorof the number of iterations required to reach this equilibriumposition, if the system starts in the state from whichthe distance is being measured. These quantities are onlyrelative, but they serve as a means of comparison among thestates [30]. The same embedding can be obtained byminimizing the criterion

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} (z_i - z_j)^2 = \mathbf{z}^{\mathrm{T}} \mathbf{L} \mathbf{z}$$

[1]Corresponding Author

ISSN: 0976-2876 (Print)
ISSN: 2250-0138(Online)

Here, z_i is the coordinate ofnode i on the axis and the vector *z* contains the z_i. Theproblem sums up in finding the smallest nontrivialeigenvectorof $(\mathbf{I} - \hat{\mathbf{P}})$, which is the same as the secondlargest eigenvector of **P**, and this is once more similar to thebasic diffusion map. Notice that this mapping has beenrediscovered and reinterpreted by Belkin and Niroyi [2], [3]in the context of nonlinear dimensionality reduction. Thelast justification of the basic diffusion map, introduced in[15], is based on the concept of two-way partitioning of agraph. Minimizing a normalized cut criterion whileimposing that the membership vector is centeredwithrespect to the metric D leads to exactly the same embeddingas in the previous interpretation. Moreover, some authors

showed that applying a specific cut criteria to bipartitegraphs leads to simple correspondence analysis. More generally, these mappings are, of course, alsorelated to graph embedding and nonlinear dimensionalityreduction, which have been highly studied topics in recentyears, especially in the manifold learning community (see,i.e., [21], [30], [37] for recent surveys or developments).Experimental comparisons with popular nonlinear dimensionalityreduction techniques are presented in the following section.

IV. EXPERIMENT AND ANALAYSIS

A. Graph Reduction Influence and Embedding Comparison

The objective of this experiment is twofold. The first aim isto study the influence of stochastic complementation ongraph mapping. The second one is to compare five populardimensionality reduction methods, namely, the diffusionmap kernel PCA (KDM PCA or simply KDM), the LaplacianEigenmap (LE) [3], the Curvilinear Component Analysis(CCA) [14], Sammon's nonlinear Mapping (SM) [25], andthe classical Multidimensional Scaling [6], [12], based ongeodesic distances (MDS). For CCA, SM, and MDS, thedistance matrix is given by the shortest path distancecomputed on the reduced graph whose weights are set tothe inverse of the entries of the adjacency matrix obtainedby stochastic complementation. Notice that the MDSmethod computed from the geodesic distance on a graphis also known as the ISOMAP method after [6]. Providedthat the resulting reduced Markov chain is usually dense,the time complexity of each algorithm is as follows: ForKDM PCA, LE, and MDS, the problem is to compute the dominant eigenvectors of a square matrix since the graphis mapped on a d-dimensional space, which is $O(d\,\tau\,n_1^2)$,where$n1$ is the number of nodes of interest being displayedand τ is the number of iterations of the power method. ForSM and CCA, the complexity is about $O(\tau\,n_1^2)$, where τ isthe number of iterations (these algorithms are iterative byrecorded. On the other hand, computing the shortest pathdistances matrix takes $O(n_1^2\log(n_1))$ Thus, each algorithmhas a time complexity between $O(n_1^2)$ and $O(n_1^3)$.In this experiment, we address the task of classificationofunlabeled nodes in partially labeled graphs, that is,semisupervised classification on a graph . Notice thatthe goal of this experiment is not to design a state-of-theartsemisupervised classifier; rather it is to study theperformance of the proposed method, in comparison withother embedding methods.Three graphs are investigated. The first graph isconstructed from the well-known Iris data set [4]. Theweight (affinity) between nodes representing samples isprovided by $w_{ij} = exp[-d_{ij}^2/\sigma^2]$, where d_{ij} is the euclidean distance in the feature space and σ^2 is simply the samplevariance. The classes are the three iris species. The secondgraph is extracted from the IMDb movie database [37]. The last graph,extracted from the CORA data set , is composed ofscientific papers from three topics. A citation graph is builtupon the data set, where two papers are linked if the firstpaper cites the second one. The tested graph contains1,410 nodes divided into three classes representing machinelearning research topics.For each of these three graphs, extra nodes are added torepresent the class labels (called the class nodes). Each classnode is connected to the graph nodes of the correspondingclass. Moreover, in order to define cross-validation folds,these graph nodes are randomly split into training sets andtest sets (called the training nodes and the test nodes,respectively), the edges between the test nodes and the classnodes being removed. The graph is then reduced to the testnodes and to the class nodes by stochastic complementation(the training nodes are rejected in the S2 subset, and thus,censored), and projected into a 2D space by applying one ofthe projection algorithms described before. Terms and topic nodes are displayed jointly.

[1]Corresponding Author

between the test nodes and the class nodes is accuratelyreconstructed in the reduced graph, these nodes from the testset should be projected close to the class node of theircorresponding class. Wereport the classification accuracy forseverallabeling rates, i.e., portions of unlabeled nodes whichconstitute the test set. The proportion of the test nodes variesbetween 50 percent of the graph nodes (twofold crossvalidation)to 10 percent (10-fold cross validation). Thismeans that the proportion of training nodes left apart(censored) by stochastic complementation increases withthe number of folds. The whole cross-validation procedure isrepeated 10 times (10 runs) and the classification accuracyaveraged on these 10 runs is reported, as well as the95 percent confidence interval.For classification, the assigned label of each test node issimply the label provided by the nearest class node, in termsofeuclidean distance in the 2D embedding space. This willpermit to assess if the class information is correctly preservedduring stochastic complementation and 2D dimensionalityreduction. The parameter t of theKDM PCA is set to 5, in viewof our preliminary experiments.Figs. 1a, 1b, and 1c show the classification accuracy, aswell as the 95 percent confidence interval, obtained on thethree investigated graphs for different training/test setpartitioning (folds). The x-axis represents the number offolds, and thus, an increasing number of nodes left apart(censored) by stochastic complementation (from 0, 50, . . . , upto 90 percent). As a baseline, the whole original graph(corresponding to one single fold and referred to as 1-fold) isalso projected without removing any class link and withoutperforming a stochastic complementation; this situationrepresents the ideal case, since all the class information iskept. All the methods should obtain a good accuracy score inthis setting—this is indeed what is observed.First, we observe that, although obtaining very goodperformance when projecting the original graph (1-fold),CCA and SM perform poorly when the number of folds,and thus, the amount of censored nodes, increases. On theother hand, LE is quite unstable, performing poorly on theCORA data set. This means that stochastic complementationcombined with CCA, SM, or LE does not workproperly. On the contrary, the performance of KDM PCAand MDS remains fairly stable; for instance, the averagedecrease of performance of KDM PCA is around 10 percent,in comparison with the mapping of the original graph(from 1-fold to 2-fold—50 percent of the nodes arecensored), which remains reasonable. MDS offers a goodalternative to KDM PCA, showing competitive performance;however, it involves the computation of the all-pairsshortest path distance.These results are confirmed when displaying the mappings.Figs. 1a, 1b, and 1c show a mapping example ofthe test nodes, as well as the class nodes (the white markers)of the CORA graph, for the 10-fold cross-validation setting.Thus, only 10 percent of the graph nodes are unlabeledandprojected after stochastic complementation of the 90 percentremaining nodes. It can be observed that the LaplacianEigenmap managed toseparate the different classes, but mostly in terms of angularsimilarity. On the KDM PCA mapping (Fig. 8d), the classnodes are well-located, at the center of the set of nodesbelonging to the class. On the other hand, the mappingsprovided by CCA and SM after stochastic complementationdo not accurately preserve the class information.

[1]Corresponding Author

ISSN: 0976-2876 (Print)
ISSN: 2250-0138(Online)

Figure 1(a): Classification accuracy obtained by the five compared projection methods for the Iris ((a), three classes), IMDb

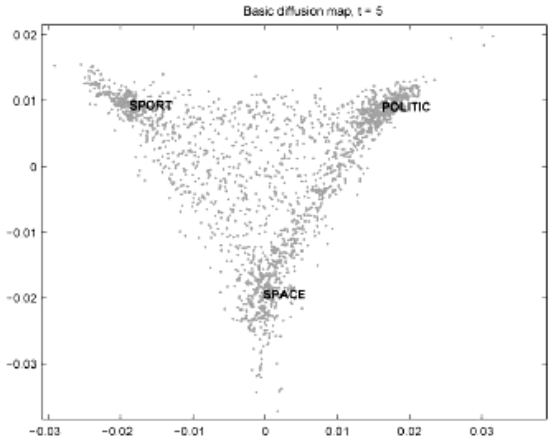


Figure 1 (b) : KDM PCA, or KDM), the LaplacianEigenmap ((e), LE), the Curvilinear Component Analysis

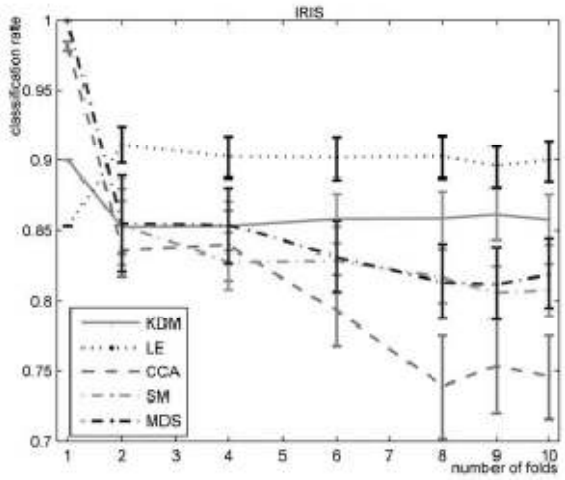
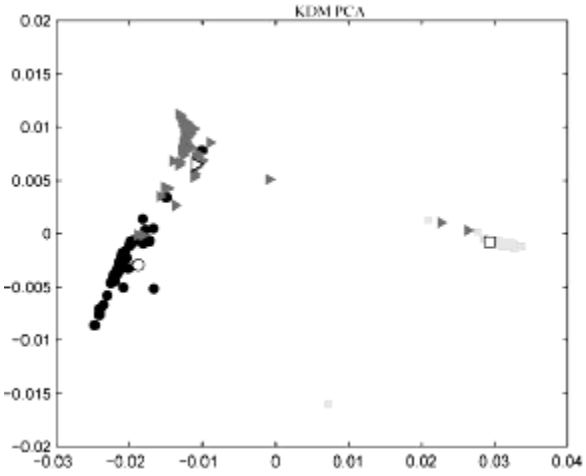


Figure 1(c) : The mapping of 10 percent of the Cora graph (10-folds setting) obtained by the five projection methods



V Conclusion

Let us now come back to our research questions. As a firstobservation, we can say that the two-step procedure(stochastic complementation followed by a diffusion mapprojection) provides an embedding in a low-dimensionalsubspace from which useful information can be extracted.Indeed, the experiments show that highly related elementsare displayed close together while poorly related elementstend to be drawn far apart. This is quite similar tocorrespondence analysis to which the procedure is closelyrelated. Second, it seems that stochastic complementationreasonably preserves proximity information, when combinedwitha diffusion map (KDM PCA) or an ISOMAPprojection (MDS). For the diffusion map, this is normal,since both stochastic complementation and the diffusionmap distance are based on a Markov chain model—stochasticcomplementation is the natural technique allowingto censor states of a Markov chain. On the contrary,stochastic complementation should not be combined withaLaplacianEigenmap, a curvilinear component analysis, oraSammon nonlinear mapping—the resulting mapping isnot accurate. Finally, the KDM PCA provides exactly thesame results as the basic diffusion map when t is large.However, when the parameter t is low, the resultingprojection tends to highlight the outlier nodes and tomagnify the relative differences between nodes. It istherefore recommended to display a whole range ofmappings for several different values of t.

[1]Corresponding Author

References

[1] P. Baldi, P. Frasconi, and P. Smyth, Modeling the Internet and theWeb: Probabilistic Methods and Algorithms. John Wiley & Sons, 2003.

[2] M. Belkin and P. Niyogi, "LaplacianEigenmaps and SpectralTechniques for Embedding and Clustering," Advances in NeuralInformationProcessing Systems, vol. 14, pp. 585-591, MIT Press, 2001.

[3] M. Belkin and P. Niyogi, "LaplacianEigenmaps for DimensionalityReduction and Data Representation," Neural Computation,vol. 15, pp. 1373-1396, 2003.

[4] C. Blake, E. Keogh, and C. Merz, "UCI Repository of MachineLearning Databases," Univ. California, Dept. of Information andComputer Science, http://www.ics.uci.edu/~mlearn/MLRepository.html, 1998.

[5] J. Blasius, M. Greenacre, P. Groenen, and M. van de Velden,"Special Issue on Correspondence Analysis and Related Methods,"Computational Statistics and Data Analysis, vol. 53, no. 8,pp. 3103-3106, 2009.

[6] I. Borg and P. Groenen, Modern Multidimensional Scaling: Theoryand Applications. Springer, 1997.

[7] P. Bremaud, Markov Chains: Gibbs Fields, Monte Carlo Simulation,and Queues. Springer-Verlag, 1999.

[8] P. Carrington, J. Scott, and S. Wasserman, Models and Methods inSocial Network Analysis.Cambridge Univ. Press, 2006.

[9] S. Chakrabarti, Mining the Web: Discovering Knowledge fromHypertext Data. Elsevier Science, 2003.

[10] F.R. Chung, Spectral Graph Theory. Am. Math. Soc., 1997.

[11] D.J. Cook and L.B. Holder, Mining Graph Data. Wiley and Sons,2006.

[12] T. Cox and M. Cox, Multidimensional Scaling, second ed. Chapmanand Hall, 2001.

[13] N. Cressie, Statistics for Spatial Data. Wiley, 1991.

[14] P. Demartines and J. Herault, "Curvilinear Component Analysis:A Self-Organizing Neural Network for Nonlinear Mapping ofData Sets," IEEE Trans. Neural Networks, vol. 8, no. 1, pp. 148-154,Jan. 1997.

[15] C. Ding, "Spectral Clustering," Tutorial presented at the 16thEuropean Conf. Machine Learning (ECML '05), 2005.

[16] P. Domingos, "Prospects and Challenges for Multi-Relational DataMining," ACM SIGKDD Explorations Newsletter, vol. 5, no. 1,pp. 80-83, 2003.

[17] F. Fouss, A. Pirotte, J.-M. Renders, and M. Saerens, "Random-Walk Computation of Similarities between Nodes of a Graph,with Application to Collaborative Recommendation," IEEE Trans.Knowledge and Data Eng., vol. 19, no. 3, pp. 355-369, Mar. 2007.

[18] F. Fouss, J.-M.Renders, and M. Saerens, "Links betweenKleinberg's Hubs and Authorities, Correspondence Analysis andMarkov Chains," Proc. Third IEEE Int'l Conf. Data Mining (ICDM),pp. 521-524, 2003.

[19] F. Fouss, L. Yen, A. Pirotte, and M. Saerens, "An ExperimentalInvestigation of Graph Kernels on a Collaborative RecommendationTask," Proc. Sixth Int'l Conf. Data Mining (ICDM '06), pp. 863-868, 2006.

[20] F. Geerts, H. Mannila, and E. Terzi, "Relational Link-BasedRanking," Proc. 30th Very Large Data Bases Conf. (VLDB), pp. 552-563, 2004.

[21] X. Geng, D.-C.Zhan, and Z.-H. Zhou, "Supervised NonlinearDimensionality Reduction for Visualization and Classification,"IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics,vol. 35, no. 6, pp. 1098-1107, Dec. 2005.

[22] Introduction to Statistical Relational Learning, L. GetoorandB. Taskar, eds. MIT Press, 2007.

[23] J. Gower and D. Hand, Biplots. Chapman & Hall, 1995.

[24] M.J. Greenacre, Theory and Applications of Correspondence Analysis.Academic Press, 1984.

[1]Corresponding Author

[25] A. Greenbaum, Iterative Methods for Solving Linear Systems. Soc.ForIndustrial and Applied Math., 1997.

[26] K.M. Hall, "An R-Dimensional Quadratic Placement Algorithm,"Management Science, vol. 17, no. 8, pp. 219-229, 1970.

[27] D.A. Harville, Matrix Algebra from a Statistician's Perspective.Springer-Verlag, 1997.

[28] T. Hastie, R. Tibshirani, and J. Friedman, The Elements ofStatistical Learning: Data Mining, Inference, and Prediction,second ed. Springer-Verlag, 2009.

[29] H. Hwang, W. Dhillon, and Y. Takane, "An Extension of MultipleCorrespondence Analysis for Identifying Heterogeneous Subgroupsof Respondents," Psychometrika, vol. 71, no. 1, pp. 161-171,2006.

[30] A. Izenman, Modern Multivariate Statistical Techniques: Regression,Classification, and Manifold Learning. Springer, 2008.

[31] R. Johnson and D. Wichern, Applied Multivariate StatisticalAnalysis, sixth ed. Prentice Hall, 2007.

[32] R. Kimball and M. Ross, The Data Warehouse Toolkit: The CompleteGuide to Dimensional Modeling. John Wiley & Sons, 2002.

[33] J.M. Kleinberg, "Authoritative Sources in a Hyperlinked Environment,"J. ACM, vol. 46, no. 5, pp. 604-632, 1999.

[34] P. Kroonenberg and M. Greenacre, "Correspondence Analysis,"Encyclopedia of Statistical Sciences, S. Kotz, ed., second ed., pp. 1394-1403, John Wiley & Sons, 2006.

[35] S. Lafon and A.B. Lee, "Diffusion Maps and Coarse-Graining: AUnified Framework for Dimensionality Reduction, Graph Partitioning,and Data Set Parameterization," IEEE Trans. PatternAnalysis and Machine Intelligence, vol. 28, no. 9, pp. 1393-1403, Sept.2006.

[36] A.N. Langville and C.D. Meyer, Google's PageRank and Beyond: TheScience of Search Engine Rankings. Princeton Univ. Press, 2006.

[37] J. Lee and M. Verleysen, Nonlinear Dimensionality Reduction.Springer, 2007.

[38] D. Lusseau, K. Schneider, O. Boisseau, P. Haase, E. Slooten, andS. Dawson, "The Bottlenose Dolphin Community of DoubtfulSound Features a Large Proportion of Long-Lasting Associations.Can Geographic Isolation Explain This Unique Trait?"Behavioral Ecology and Sociobiology, vol. 54, no. 4, pp. 396-405,2003.

[1]Corresponding Author