IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 1, January 2018

NMF and its Interdisciplinary Applications: A Review

Sumisha Kalithodi[1], Nandakumar Paramparambath[2]

Department of ECE, NSS College of Engineering, Palakkad, India [1,2]

Abstract: Non-negative matrix factorisation (NMF or NNMF) is a group of algorithms for matrix decomposition with the non-negative constraint, that is the matrix to be factorised and the factors of the matrix should be non-negative. NMF is a very popular technique for data decomposition. Some of the data's such as image pixels, power spectra, stock indices are non- negative. NMF has found wide variety of applications mostly in pre-processing of signals and classification. So in this paper we review about the concepts of NMF and its applications in various fields.

Keywords: NMF, signal processing, power spectra, classification.

I. INTRODUCTION

Non-negative Matrix Factorisation (NMF) is a widely known decomposition technique in data analysis. The basic idea behind data analysis is rank reduction. There are many techniques for rank reduction such as Singular Value Decomposition (SVD), Semi Discrete Decompositions (SDD) and Principal Component Analysis (PCA). The main difference between NMF and the above mentioned techniques is the non-negative constraint. NMF was introduced by Paatero and Tapper in 1990's and initially it is named as Positive Matrix Factorisation which is given in [1]. After that Lee and Seung done research on the properties of the algorithm and published some simple algorithms for two types of factorisations described in [2-3].

What NMF does is that, according to our choice of factorisation rank ($\boldsymbol{r}$), it decomposes a non-negative matrix $\boldsymbol{V}$ of order ($\boldsymbol{m}$ by $\boldsymbol{n}$) to two smaller matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ of order ($\boldsymbol{m}$ by $\boldsymbol{r}$) and ($\boldsymbol{r}$ by $\boldsymbol{n}$) respectively. NMF is applied in various applications due to its ability to extract sparse and easily interpretable factors. The main issues related to NMF are NMF is NP-hard and ill posed. And the choice of factorisation rank is also important.

Paper is organized as following sections. Section II gives the basic idea of NMF and brief introduction of its algorithms. Section III contains the various types of NMF. Section IV gives various applications of NMF.

II. ALGORITHMS OF NMF

There are various algorithms for NMF. Each algorithm differs on the cost function which minimises the difference between input data matrix $\boldsymbol{V}$ and the product of factor matrices $\boldsymbol{W}$ and $\boldsymbol{H}$.

A. Lee and Seung Multiplicative Update (MU)

(1) Kullback-Leibler (KL) Divergence Based Cost Function:

KL Divergence between input data matrix $\boldsymbol{V}$ and estimate of input data matrix $\boldsymbol{Z}$ is minimised. It is done through updating $\boldsymbol{W}$ and $\boldsymbol{H}$ iteratively until the cost function converges. $\boldsymbol{W}$ and $\boldsymbol{H}$ should be updated simultaneously. That is after updating one row of $\boldsymbol{W}$ we need to update corresponding column of $\boldsymbol{H}$.

$$\boldsymbol{Z} = \boldsymbol{WH} \qquad \text{..... (1)}$$

$$\min_{\boldsymbol{W},\boldsymbol{H}\geq 0} \boldsymbol{D}(\boldsymbol{V}|\boldsymbol{WH}) \qquad \text{..... (2)}$$

For any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ divergence is given by $\boldsymbol{D}(\boldsymbol{A}|\boldsymbol{B}) = \sum_{i,j}(\boldsymbol{A_{ij}}\, \boldsymbol{log}\left(\frac{\boldsymbol{A_{ij}}}{\boldsymbol{B_{ij}}}\right) - \boldsymbol{A_{ij}} + \boldsymbol{B_{ij}})$ ..... (3)

Multiplicative updating rules are,

$$\boldsymbol{W_{i,a}^{k+1}} \leftarrow \boldsymbol{W_{i,a}^{k}} \frac{\sum_m \frac{\boldsymbol{H_{a,m}V_{i,m}}}{\boldsymbol{(WH)_{i,m}}}}{\sum_n \boldsymbol{H_{a,n}}} \qquad \text{..... (4)}$$

$$\boldsymbol{H}^{k+1}_{a,m} \leftarrow \boldsymbol{H}^{k}_{a,m} \frac{\sum_i \frac{W_{i,a} V_{i,m}}{(WH)_{i,m}}}{\sum_k W_{k,a}} \qquad ..... (5)$$

(2) Frobenius Norm Based Cost Function:

Here the Frobenius Norm between $\boldsymbol{V}$ and $\boldsymbol{Z}$ is calculated and minimised by updating W and H. Frobenius norm is same as Euclidean distance.

Here the cost function is,

$$\min_{W,H\geq 0} \| \boldsymbol{V} - \boldsymbol{WH} \|^2 \qquad ..... (6)$$

and the updating rules are ,

$$\boldsymbol{W}^{k+1}_{i,a} \leftarrow \boldsymbol{W}^{k}_{i,a} \frac{(VH^T)_{i,a}}{(WHH^T)_{i,a}} \qquad ..... (7)$$

$$\boldsymbol{H}^{k+1}_{a,m} \leftarrow \boldsymbol{H}^{k}_{a,m} \frac{(W^T V)_{a,m}}{(W^T WH)_{a,m}} \qquad ..... (8)$$

B. Alternating Least Squares (ALS)

ALS is the first proposed algorithm for solving NMF. Fixing either W or H, the problem becomes a least squares problem with non-negativity constraint. Fixing either W or H, the problem becomes a least squares problem with non-negativity constraint.

C. Alternating Non-Negative Least Squares (ANLS)

Each iteration of ANLS computes an optimal solution of the NNLS sub problem and it decreases the error.

D. Hierarchical Alternating Least Squares **(***HALS)*

Solves the NNLS sub problem using an exact coordinate descent method, updating one column of W at a time. Advantage of HALS is it converges much faster than the MU with almost same computational cost.

III. TYPES OF NMF

There are different types of NMF which are differing based on the constraint imposed and structure Also there comes the extension of NMF to higher dimensional data. The various NMF types are discussed in [5].

A. Basic NMF

In basic NMF in order to quantify the difference between the input data matrix V and the product of the factor matrices W and H , simple multiplicative update rule is used. And the only constraint on this type of NMF is the non-negative constraint.

B. Constrained NMF

In this type of NMF some additional constraints are imposed as regularization.

(1) Sparse NMF (SNMF) : Sparseness Constraint is introduced in basic NMF through adding an additional term to the objective function. For improving uniqueness of decomposition sparseness is required. Cost function is given by ,

$$\min_{W,H} \frac{1}{2} \{ \| \boldsymbol{V} - \boldsymbol{WH} \|^2 + \boldsymbol{\alpha} \| \boldsymbol{W} \|^2 + \boldsymbol{\beta} \| \boldsymbol{H} \|^2 \} \qquad ..... (9)$$

(2) Orthogonal NMF: Orthogonal constraint is imposed on basic NMF that is either on factor $\boldsymbol{W}$ or $\boldsymbol{H}$. The product of matrix W (or H) and its transpose gives identity matrix $\boldsymbol{I}$. Orthogonality results in sparseness and the constraint is given by,

$$\boldsymbol{W^T W} = \boldsymbol{I} \text{ or } \boldsymbol{H^T H} = \boldsymbol{I} \qquad ..... (10)$$

(3) Discriminant NMF: It contains the information of discrimination and classification. Basic NMF is an unsupervised learning technique. Through coupling discriminant information with the decomposition, basic NMF can be considered as a supervised learninig technique.

(4) NMF on Manifold: This type of NMF preserves the topological properties. Basic NMF fits the data into a Euclidean space while NMF on manifold is to explicitly incorporate NMF with the proper local invariant properties.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 1, January 2018

With corresponding manifold learning methods, highly improved performance in tasks like document and image clustering can be achieved.

C. Structured NMF

(1) Weighted NMF: According to relative importance of elements some weights are attached to them. Weighted NMF can be solved by introducing the weight matrix in the standard multiplicative update rules.

(2) Convolutive NMF: It consider the time-frequency domain factorization. Conventional Basic NMF can be regarded as a kind of instantaneous decomposition, where each object is described by the basis matrix W and the coefficient matrix H . To incorporate the time domain information it is necessary to take into account the time-varying characteristic of the spectrum. And typically the temporal relationship between multiple observations over nearby intervals of time is described using a convolutive generative model.

(3) Non-negative matrix tri factorization (NMTF): NMTF decomposes the data matrix into three factor matrices

$$V \approx WSH \qquad \text{..... (11)}$$

Constrained NMTF provides additional degrees of freedom.

D. Generalized NMF

(1) Semi-NMF: Here the non-negativity constraint is relaxed on specific factor matrix. Conventional NMF restricts every element in data matrix $\boldsymbol{X}$ to be non-negative. When V is unconstrained , which may have mixed signs, the factor matrix W is still restricted to be non-negative while placing no restriction on the signs of H.

(2) Non-negative Tensor Factorization (NTF): NTF generalizes the matrix data to higher dimensional tensors. NMF is a particular case of nonnegative n-dimensional tensor factorization (n-NTF) when n = 2.

(3) Non-negative Matrix-set Factorization (NMSF): In NMSF the data is extended from matrices to matrix set. NMSF is implemented directly on the matrix set. Each sample matrix is decomposed into the product of K factor matrices, where the public $K - 1$ factor matrices represent the learnt features which generalize the feature matrix in NMF to a feature matrix set, and the remaining factor matrix varying from individual sample matrix describes the activation patterns which generalizes the coefficient vector in NMF to a coefficient matrix.

(4) Kernel NMF (KNMF): KNMF is a non-linear model of NMF. NMF and its variants mentioned above are linear models, which are unable to extract nonlinear structures and relationships hidden in the data. This restricts the scope of application of NMF. To overcome these limitations, a nature extension is to apply kernel based methods by mapping input data into an implicit feature space using nonlinear functions, just as kernel PCA, kernel LDA, and kernel ICA.

IV. APPLICATIONS OF NMF

NMF is used in various applications due to the non-negative nature of real life data. The various fields in which NMF is applied are feature extraction in biomedical signal processing, audio signal processing, clustering, and image unmixing and so on. The below section briefly describes various applications of NMF.

A. Biomedical Signal Processing

(1) NMF on Electroencephalographic (EEG) recordings:

NMF is widely used in biomedical applications. To extract non-negative components that have meaningful physical or physiological interpretations a new NMF algorithm with smoothness constraint is proposed. Two additional constraints named as temporal smoothness constraint and spatial decorrelation constraint are introduced and new multiplicative update rules are introduced.

To evaluate human brain degeneration processing it is very important to detect of Alzheimer disease (AD). But the early detection of AD is a challenging problem. For early detection of Alzheimer disease using clinical EEG recordings NMF combined with advanced time-frequency analysis and machine learning techniques is proposed in [6]. NMF is used to extract features from the artifacts removed EEG recording. Using Welch method power spectra is calculated. Then on the power spectra constraint NMF is applied.

NMF was employed as feature extraction tool, which leads to more localized and sparse features. The basis vectors, which are obtained by applying NMF on the EEG power spectrum, form a feature space. NMF together with artificial neural network (ANN) for EEG classification is proposed in [7]. Spatial-temporal characteristics of EEG signal are preserved by NMF-ANN structure. As the data size increases NMF suffers from the out of memory problem. To avoid that, the data matrix is down sized using the CUR method before applying NMF and it is discussed in [8].

Artifact removal is one of the big challenges in applications of EEG recording. EEG recordings give a mixture of endogenous brain activities and extraneous environmental and physiological artifacts such as eye movements, power grid noise, muscle activities or heart beat. And these artifacts may overlap with lower energy cerebral signals. So EEG signal separation from this mixture of artifacts is needed for a good analysis of EEG. Non-negative Matrix Factorization (NMF) in a Gaussian source separation frame work can be used for removal of arifacts from the single channel EEG recordings. Ocular artifacts can be removed from the single channel EEG through the analysis of electrooculographic recordings, which gives prior information about the ocular artefact, is proposed in [9].

NMF is effective in blink artifact rejection. However, the above method was not formulated the accuracy. Therefore an artifactitious rejection method using single-channel EEG recordings and two steps NMF is proposed in [10]. Extraction of discriminative spectral features from the power spectrum of EEG can also done using NMF. When KNMF with linear kernel is used spectral features can be easily computed and is shown in [11]**.** And by increasing the difference of spectral EEG energies the classification accuracy can be improved. It can be achieved by imposing some constraints on KNMF as proposed in [12].

(2) NMF on ECG signals:

It is always possible to capture effect of EMG signal, when ECG is recorded and it act as noise. Since the TF patterns of EMG and ECG differ widely they can be separated using NMF. NMF factorizes the mixed signal and those factors give ECG and EMG estimates and it is discussed in [13].

(3) NMF on EMG signals:

Biomedical signals are non-stationary in nature. Due to time varying nature of such signals the processing will be difficult and the small variations in the processed data may create inefficient classification. In order to overcome this problem a time-frequency NMF (TF-NMF) is used. Sleep signal can be classified efficiently from the EMG signal using this method as proposed in [14].

Surface Electromyography (sEMG) is mainly used for evaluating the functioning of hand. NMF is used for feature extraction and selection through exploiting sparse nature of the signal sEMG. EMG finger movements can be identified through combining NMF and ANN and is proposed in [15]. Cross fuzzy entropy employed NMF can be used for muscle synergy extraction which is a best method for illustrating the coordinated motor patterns and it is discussed in [16-17].

B. Audio Signal Processing

(1) Blind Speech Separation:

Every communication scenario suffers from background noise. When speech and noise superpose it results in low intelligibility of speech signal. And speech enhancement technique aims to improve the speech quality and there by speech intelligibility. NMF can be applied on both monaural and binaural speeches and can separate the speech signal from the mixture of speech and noise. Speech separation can either supervised or unsupervised.

To eliminate the effect of background noise Bayesian formulation NMF together with HMM can be used. And Enhanced NMF gives better results. Through this the additive noise can be removed and is shown in [26]. In order to separate speech and music different cost functions are used. For training stage of speech KL divergence cost function is used and that of music is Itakura-Saito (IS) divergences. And in decomposition stage a linear combination of these two cost functions are employed as proposed in [27].

And to extract spectro-temporal structures, NMF can be employed on magnitude spectrum of noisy speech signal. Through factorizing magnitude spectrum using NMF the basis vectors can be extracted and giving these basis vectors to Deep Neural Network (DNN) the original speech samples and noise can be reconstructed as proposed in [28]. Speech separation can be done effectively through combining NMF with Time Delay of Arrival (TDOA) estimation technique such as Generalized Cross Correlation (GCC) as in [29].

C. Spectral Data Analysis in Astronomy

Data analysis is an important process in various fields like engineering, science and business. And in most of the cases the data to be analyzed is non-negative. The spectral data which is obtained using astronomical spectrometers is analyzed and NMF is applied to unmix spectral reflectance data for space object identification and classification purposes as in [30].

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 1, January 2018

D. Image Processing

(1) Hyperspectral Image Unmixing:

Hyperspectral unmixing is used for analyzing hyperspectral images. Through this a mixed pixel is factorized into a collection of constituent materials with appropriate weight. Since NMF can utilize the sparse data also NMF can be applied for hyperspectral unmixing. Manifold regularized sparse NMF gives better results and is given in [18]. Introducing sparse NMF into kernel space, which can analyze non-linear data also, results in a better hyperspectral unmixing which is proposed in [19].

Cluster constraint based sparse NMF which strengthen the constrained relationship between original image and abundance maps is proposed in [20] , and Dual Graph Regularized NMF which captures the latent manifold structure is given in [21]. The KL Divergence Constrained Sparse NMF performs better than Euclidean based NMF through measuring statistical distribution difference is also used for hyperspectral unmixing and is discussed in [22].

(2) Image Quality Assessment: The communication and digital imaging technologies are growing rapidly. And as a result it is very important to measure the quality of images for many applications like compression, restoration and watermarking. The existing quality measures like luminance, gradient and contrast are inconsistent and a new image quality metric which is based on the part based representation of NMF is proposed. NMF measures the image degradation of distorted image using the reference image as proposed in [23].

(3) Image Clustering:

The fundamental information of an image is contained in its low rank parts. To extract those low rank features NMF can be applied which performs clustering of images and is discussed detail in [24].

E. Seismic signal analysis for Footstep separation

Detecting human motion based on footstep signature using seismic/acoustic sensors is an emerging research topic. For home land security applications like border security, Seismic footstep detection based systems can be used. Seismic footstep signal separation for a single channel recording can be done using Sparsity-based NMF (SNMF).

To separate the human footstep signatures from the horse footstep signatures, supervised NMF technique is employed. First the spectrogram of human footstep signals and horse footstep signals is computed then the algorithm is applied on those spectrograms. The method has the ability to decompose a complex signal automatically into objects that have a meaningful interpretation, which is proposed in [25].

V. CONCLUSION

NMF is a very popular technique which comprises many algorithms for decomposing a non-negative matrix into two low rank non-negative matrices. NMF has been applied in wide variety of applications and have good results. NMF is very popular due to the non-negative nature of real life data's we are dealing with. In this paper the main algorithms and various types of NMF are described. Also the various applications of NMF are also explained.

REFERENCES

[1] P. Paatero and U. Tapper, "Positive Matrix Factorization: a Non-negative Factor Model with Optimal Utilization of Error Estimates of Data Values," Environmetrics, vol. 5, no. 2, pp. 111–126, 1994.
[2] DD.Lee, H.Seung. "Learning the parts of objects by non- negative matrix factorization "[J]. Nature, 1999, 401: 788-791.
[3] DD.Lee, H.Seung. "Algorithms for non-negative matrix factorization" [C]. In Adv. Neural Info. Proc. NIPS2000, 2000.
[4] "Algorithms for Non-negative Matrix Factorization," in Proc. Advance.
[5] Yu-Xiong Wang, Yu-Jin Zhang, " Non-negative Matrix Factorization: a Comprehensive Review" , IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, TKDE-2011-07-0429, 2013.
[6] Zhe Chen, Andrzej Cichocki and Tomasz M. Rutkowski , "CONSTRAINED NON-NEGATIVE MATRIX FACTORIZATION METHOD FOR EEG ANALYSIS IN EARLY DETECTION OF ALZHEIMER DISEASE" ,IEEE ICASSP 2006.
[7] LIU Mingyu, WANG Jue, ZHENG Chongxun, "Non-Negative Matrix Factorizations of Spontaneous Electroencephalographic Signals for Classification ",Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005.
[8] Hyekyoung Lee and Seungjin Choi , "CUR+NMF for Learning Spectral Features from Large Data Matrix", IEEE, 2008.
[9] Cécilia Damon, Antoine Liutkus, Alexandre Gramfort and Slim Essid , " NON-NEGATIVE MATRIX FACTORIZATION FOR SINGLE-CHANNEL EEG ARTIFACT REJECTION " , ICASSP 2013.
[10] Suguru Kanoga and Yasue Mitsukura , " Proposed Blink Artifacts Rejection Method Mixing with Electroencephalographic Recordings Using Single-Channel Electroencephalographic Recording and Non-Negative Matrix Factorization", IEEE, 2014.
[11] HyekyoungLee , AndrzejCichocki and SeungjinChoi, " Kernel non-negative matrix factorization for spectral EEG feature extraction " , ELSEVIER, eurocomputing72(2009)3182–3190.
[12] Motoki Sakai , "Kernel Nonnegative Matrix Factorization with Constraint Increasing the Discriminability of Two Classes for the EEG Feature Extraction ",International Conference on Signal-Image Technology & Internet-Based Systems , IEEE, 2013.
[13] Maciej Niegowski and Miroslav Zivanovic , "ECG-EMG Separation by Using Enhanced Non-Negative Matrix Factorization " , IEEE, 2014.
[14] Mehrnaz Shokrollahi, and Sridhar Krishnan , " Non-negative Matrix Factorization and Sparse Representation for Sleep Signal Classification " , 35th Annual International Conference of the IEEE EMBS Osaka, Japan, 3 - 7 July, 2013.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 1, January 2018

[15] Ganesh R. Naik and Hung T. Nguyen , "Non Negative Matrix Factorisation for the Identification of EMG finger movements: Evaluation using matrix analysis " , IEEE Journal of Biomedical and Health Informatics 2015.
[16] Beth Jelfs , Ling Li , Chung Tin and Rosa H.M. Chan , "Fuzzy Entropy Based Nonnegative Matrix Factorization For Muscle Synergy Extraction " , IEEE 2016.
[17] Luca Cerina, Pierandrea Cancian, Giuseppe Franco and Marco D. Santambrogio , "A Hardware Acceleration for Surface EMG Non-Negative Matrix Factorization " , IEEE International Parallel and Distributed Processing Symposium Workshops, 2017.
[18] Xiaoqiang Lu, Hao Wu, Yuan Yuan, Pingkun Yan, and Xuelong Li, "Manifold Regularized Sparse NMF for Hyperspectral Unmixing ", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 51, NO. 5, MAY 2013.
[19] Bei Fang, Ying Li, Peng Zhang,and Bendu Bai , "Kernel Sparse NMF for Hyperspectral Unmixing" , IEEE, 2014.
[20] Xinwei Jiang, Lei Ma and Yiping Yang , "CLUSTER CONSTRAINT BASED SPARSE NMF FOR HYPERSPECTRAL IMAGERY UNMIXING ", ICIP 2014.
[21] Lei Tong , Jun Zhou , Xiao Bai and Yongsheng Gao, "Dual Graph Regularized NMF for Hyperspectral Unmixing ", IEEE 2014.
[22] Shaodong Wang, Nan Wang, Dacheng Tao, Lefei Zhang and Bo Du , "A K-L Divergence Constrained Sparse NMF for Hyperspectral Signal Unmixing " , IEEE International Conference on Security, Pattern Analysis , And Cybernetics (SPAC) ,October 18-19, 2014.
[23] Shuigen Wang, Chenwei Deng, Weisi Lin, Guang-Bin Huang and Baojun Zhao , "NMF-Based Image Quality Assessment Using Extreme Learning Machine " , IEEE TRANSACTIONS ON CYBERNETICS 2016.
[24] Xuelong Li, Guosheng Cui and Yongsheng Dong, "Graph Regularized Non-Negative Low-Rank Matrix Factorization for Image Clustering " , IEEE TRANSACTIONS ON CYBERNETICS 2016.
[25] Asif Mehmood, Thyagaraju Damarla and James Sabatier, "Separation of human and animal seismic signatures using non-negative matrix factorization " , ELSEVIER, Pattern Recognition Letters 33 2085–2093 , 2012.
[26] Akarsh K.A and Senthamizh Selvi R , "Speech Enhancement using Non negative Matrix Factorization and Enhanced NMF" , International Conference on Circuit, Power and Computing Technologies [ICCPCT] 2015.
[27] Babak Nasersharif and Sara Abdali ," Speech/Music Separation Using Non-negative Matrix Factorization with Combination of Cost Functions", IEEE , International Symposium on Artificial Intelligence and Signal Processing (AISP), 2015.
[28] Shuai Nie, Shan Liang, Hao Li, XueLiang Zhang, ZhanLei Yang, WenJu Liu and LiKe Dong ," EXPLOITING SPECTRO-TEMPORAL STRUCTURES USING NMF FOR DNN-BASED SUPERVISED SPEECH SEPARATION" , IEEE, ICASSP 2016.
[29] Sean U. N. Wood, Jean Rouat, Stephane Dupont, and Gueorgui Pironkov , "Blind Speech Separation and Enhancement With GCC-NMF" , IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 25, NO. 4, APRIL 2017.
[30] V. Paul Pauca, J. Piper and Robert J. Plemmons, "Nonnegative Matrix Factorization for Spectral Data Analysis", Department of Computer Science, Wake Forest University, Winston-Salem, NC 27109.