ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

Latent variable analysis in hospital electric power demand using non-negative matrix factorization

Diego García[1], Ignacio Díaz[1], Daniel Pérez[1], Abel A. Cuadrado [1], Manuel Domínguez [2] *

1- Electrical Engineering Dept. University of Oviedo
Edif. Dept., Campus de Viesques s/n 33204, Gijón - SPAIN
2- SUPPRESS Research Group, University of Leon,
Escuela de Ingenierías, Campus de Vegazana, 24007 León, Spain

Abstract. Energy disaggregation techniques have recently attracted much interest, since they allow to obtain latent patterns from power demand data in buildings, revealing useful information to the user. Unsupervised methods are specially attractive, since they do not require labeled datasets. Particularly, *non-negative matrix factorization* (NMF) methods allow to decompose a single power demand measurement over a certain time period into a set of components or "parts" that are sparse, non-negative and sum up the original measured quantity. Such components reveal hidden temporal patterns and events along this period, related to scheduling events and/or demand patterns from subsystems in the network, that are very useful within an energy efficiency context. In this paper we use this approach on demand data from a hospital during a one-year period, using a calendar visualization of the components, revealing relevant facts about the energy expenditure.

1 Introduction

One interesting approach for power demand analysis [1] is *energy disaggregation*, which decomposes an aggregated electrical measurement into several components, revealing features of downstream demand. An example of this can be found in a household, where we are interested in obtaining different appliance consumptions from the overall energy measurement of the house. In recent years, the interest in energy disaggregation or also named *non-intrusive load monitoring* (NILM) has grown, since the perception and knowledge about consumption is improved through parts-based representation of total energy. This kind of representation can also suggest changes in schedules or detect faults in the network. NILM techniques can be classified into supervised and unsupervised methods. In supervised methods, pattern recognition and optimization algorithms [2] are applied, and its principal disadvantage is that a labeled dataset is required. In many cases, obtaining labeled data can increase set-up costs of NILM systems, making unsupervised methods more suitable. Within unsupervised NILM techniques, *blind source separation* (BSS) and *factorial hidden Markov models*

*The authors would like to thank financial support from the Spanish Ministry of Economy (MINECO)and FEDER funds from the EU under grant DPI2015-69891-C2-1/2-R.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

(FHMM) are the most typical approaches as it is explained in [3]. In this paper we focus in BSS methods, specifically in *non-negative matrix factorization* (NMF). NMF, as PCA or ICA, is a matrix factorization with certain constraints. Its principal constraint is the non-negativity in the matrix values, while the PCA and ICA constraints are orthogonality and statistical independence, respectively. Electric consumption measurements are always positive, hence NMF is a suitable decomposition in order to obtain parts-based representation of data as it is shown in [4].

In a real case of electrical consumption in buildings it is well-known that *thermal comfort systems* are almost half of the overall consumption [5]. Thus, if temporal patterns associated with this kind of systems can be obtained, visual tools can be very useful in order to optimize timetables or make changes in certain behaviours to reduce the total consumption. The visualization of NMF components of electrical time series is suggested in our study in order to highlight significant events in the network which can be associated with *thermal comfort systems*.

The remainder of this paper is organized as follows: first, in section 2, we explain basis and limitations of NMF applied to electric measurements. Then, in section 3, results of the application of NMF to hospital power consumption data are presented. Finally, in section 4 conclusions are drawn and some limitations and future works are discussed.

2 Methods

NMF is formulated [6, 4] as an approximate matrix factorization of non-negative input expressed as $\mathbf{V} \approx \mathbf{WH}$. Let us define a M-dimensional vector $\mathbf{v}_i$ whose elements are non-negative and the matrix $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_N] \in \mathbb{R}^{M \times N}$ as N observations of $\mathbf{v}_j, j = 1...N$. Considering $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_L] \in \mathbb{R}_+^{M \times L}$ and $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_N] \in \mathbb{R}_+^{L \times N}$, NMF can be reformulated as:

$$\mathbf{v}_j = \sum_{\alpha=1}^{L} \mathbf{w}_\alpha h_{\alpha,j}$$

where $\mathbf{v}_j$ is the j-th column of $\mathbf{V}$ and $h_{\alpha,j}$ is the α-th element in j-th column of $\mathbf{H}$. Therefore $\mathbf{v}_j$ is a linear combination of columns in $\mathbf{W}$. The coefficients of the combination are imposed by the elements of columns in $\mathbf{H}$, thus, all the input observations are weighted sums of non-negative basis. This non-negativity suggests that the columns of $\mathbf{W}$ are "the parts of a whole" as it is shown in [7], and they will be referred as *basis consumptions* or simply *components*. Therefore, the elements of the rows in $\mathbf{H}$ contain information about the influence of the basis consumptions along the N observations. The application of NMF to electrical consumptions aims to extract a certain number of basis consumptions that are a parts-based representation of the whole measurement. Basis consumptions show latent patterns, which may contain useful information in order to improve the knowledge about the system.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

One difficult issue is to determine the number L of basis components that achieve the exact decomposition $\mathbf{V} = \mathbf{WH}$. Computational NP-hardness of the estimation of L is introduced in [8]. Due to the difficulty of finding L, we dismiss the exact model [8] and we use the basic NMF under only the non-negativity constraints. Typically, $\mathbf{W}$ and $\mathbf{H}$ are estimated using a constrained optimization problem whose objective function is a similarity measurement between $\mathbf{V}$ and $\mathbf{WH}$ [7] and it is solved by means of gradient descent methods [9]. This optimization problem is not convex with respect to $\mathbf{W}$ and $\mathbf{H}$ at the same time; however, it is separately convex in either $\mathbf{W}$ or $\mathbf{H}$. As a nonconvex gradient-based problem, results of basic NMF rely on the initialization of $\mathbf{W}$ and $\mathbf{H}$ matrices. We have considered *nonnegative double singular value decomposition* (*nndsv*) [10].

One variant of the basic method, introduced in [11], is *sparse* NMF, that results in components that are easier interpret, since only a few elements of $\mathbf{W}$ and $\mathbf{H}$ are significant and the rest are near zero (sparseness). Although basic NMF in some cases reach sparse solutions, sparse NMF variants include new constraints in the objective function, so that sparsity and parts-based representation are enforced. The new optimization problem is formulated as:

$$\arg\min_{\mathbf{W},\mathbf{H}} \frac{1}{2}\|\mathbf{V}-\mathbf{WH}\|_F^2+\alpha\rho\|\mathbf{W}\|_1+\alpha\rho\|\mathbf{H}\|_1+\frac{\alpha(1-\rho)}{2}\|\mathbf{W}\|_F^2+\frac{\alpha(1-\rho)}{2}\|\mathbf{H}\|_F^2$$

where $\|\mathbf{V}-\mathbf{WH}\|_F^2$ is a similarity measurement between $\mathbf{W}$ and $\mathbf{H}$. The rest of the terms in the loss function are elements of regularization. $\|\mathbf{W}\|_1, \|\mathbf{H}\|_1$ are L_1-norm of $\mathbf{W}$ and $\mathbf{H}$ respectively and $\|\mathbf{W}\|_F^2, \|\mathbf{H}\|_F^2$ are L_2-norm. Such regularization terms are weighted by two parameters: α, that affects the whole regularization; and ρ, that controls L_1 or L_2-norm priority. When ρ is 1, only L_1-norm members are significant in the loss function and more sparse results are obtained. On the contrary, if ρ is 0, the loss function is only affected by L_2-norm members and therefore dense basic components are obtained.

2.1 Structure of matrix V

If we consider electrical consumption data as a time series $\{x(t)\}$, a restructuring is needed in order to construct $\mathbf{V}$ as an input matrix to NMF. Thus we can restructure the sequence by breaking it up into N contiguous windows of length M:

$$\{\underbrace{x(0), x(1), \ldots, x(M-1)}_{\mathbf{v}_1}, \ldots, \underbrace{x((N-1)M), x((N-1)M+1), \ldots, x(NM-1)}_{\mathbf{v}_N}\}$$

The new windows conform the matrix $\mathbf{V}$ of observations as follows:

$$\mathbf{V} = \begin{pmatrix} x(0) & x(M) & \ldots & x((N-1)M) \\ x(1) & x(M+1) & \ldots & x((N-1)M+1) \\ \vdots & \vdots & \ddots & \vdots \\ x(M-1) & x(2M-1) & \ldots & x(NM-1) \end{pmatrix}$$

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

The results of NMF are closely associated to the length of the windows, so a careful choice of this parameter should be done to obtain more interpretable basis components. Power demand time series in buildings typically present a strong cyclic structure for periods of days, weeks, months or years. If the size of the window is chosen according to these periods of time, more interpretable results will be obtained, since the structuring of the results allows a better association between basis components of consumption and events in the network. In other words, different patterns in the basis components are obtained by NMF if a column of **V** is chosen as a one day, week or month. Accordingly, NMF will return components with patterns in daily, weekly or monthly scale.

3 Experimental Results

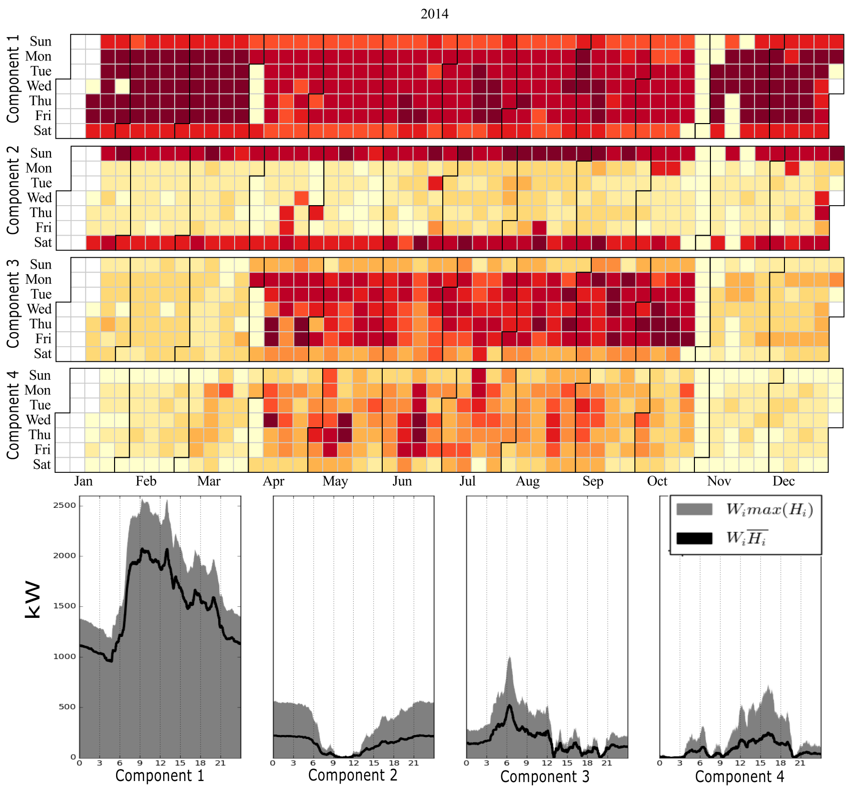

Fig. 1: Four components and their relevance during one year. Above, the calendar visualizations show the component weights (**H** rows) along one year; below, their corresponding maximum and average hourly power demand profiles.

The concepts explained above were tested with real electric power demand data. We have decided to apply sparse NMF to an electrical dataset obtained

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

from a hospital complex, which includes different measurements of current, voltage, power and quality in several measuring points over the network, collected with a sampling period of one minute. In our analysis, we focus on the total power demand of the hospital, including all the heating, ventilation and air conditioning systems (HVAC), and two cold machines, which feed with water the cold ring that keeps diagnostic devices cool. These two machines have a large peak of consumption in their start. Power data were obtained along 2014 but some days were lost due to problems in the meter caused by a temporary network malfunction.

We choose $M = 1440$ so that one column of $\mathbf{V}$ corresponds to one day; therefore matrix $\mathbf{V}$ will be composed of $N = 365$ observations of daily total power consumption. Regarding the NMF regularization parameter, ρ is set to 1, because we want results as sparse as possible, and α will be greater than 1 in order to increase the influence of L_1-norm in the loss function. Finally, we empirically choose $L = 4$ basic disaggregated components.

The results are shown in Figure 1. Each gray chart shows the product between a basis consumption profile and the highest value in the corresponding row in $\mathbf{H}$, and each black line shows the product between a basis consumption profile and the average value in the corresponding row in $\mathbf{H}$. The calendars show the values of the rows in $\mathbf{H}$ associated with each basis along the year, in other words, the influence of the basis consumption profile. Note that the bottom four charts are in kW units, because the total power demand is a sum of terms, each one being the product of a component and its associated element in $\mathbf{H}$.

The first one has the form of a normal daily profile in the hospital and it is so high that can be interpreted as a canvas to which the rest of the components will be aggregated. The second component shows an important influence in weekends and holidays, where the first one takes less significance. The third component shows a significant influence from April to November and reveals the activation of cooling elements in HVAC systems. If we focus on its profile, we will find a large peak early in the morning at 6:30 when the HVAC system starts. Between November and April, the cooling subsystem in the HVAC system is disabled, and for this reason, the $\mathbf{H}$ elements in the third component are smaller in this period. Finally, the fourth component reveals large values in the afternoon and at noon that are associated with the start of one of the two cold machines that refrigerate the cold water ring. This fact is known because on April 9th the technical staff carried out a maintenance test of these machines.

4 Conclusions

By means of *NMF*, it is possible to find latent patterns in electric power demand time series of large consumers, as shown here with a hospital, which would be more difficult to obtain by other methods. Due to the non-negativity constraint the user can form a daily profile as one weighted addition of positive and characteristic disaggregated profiles where no cancellation between terms is possible. Such lack of cancellations implies that *NMF* can be interpreted as parts-based

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

representation and these parts reveal connections with relevant events in the network, as, for instance, the start of $HVAC$ systems. Moreover, the elements of $\mathbf{H}$ associated with one of the components contain information about its influence along the N observations. In the hospital results, these elements represent how much relevance the disaggregated profile has along the days in one year. Temporal information and parts-based representation are interpretable and highly useful knowledge, that suggest to create visual analytics applications with the results in *NMF*. We have displayed the weights in $\mathbf{H}$ matrix using calendar layouts that present a simple representation of the influence of the components in overall daily consumption along a year and allow comparisons between them. In future works *NMF* can be the base of more intuitive and interactive applications. Despite its advantages, some drawbacks are presented by *NMF* techniques applied to electrical consumptions such as the choice of the number of components or the fact that the first component may not be sparse even when sparseness constraints are included. These open issues, as well as exploring other electrical measurements downstream in the network of the hospital, are the guidelines of future works.

References

[1] J Zico Kolter and Matthew J Johnson. Redd: A public data set for energy disaggregation research. In *Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA*, volume 25, pages 59–62. Citeseer, 2011.

[2] George William Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.

[3] A. Zoha, A. Gluhak, M.A. Imran, and S. Rajasegarar. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors (Switzerland)*, 12(12):16838–16866, 2012. cited By 127.

[4] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[5] Luis Pérez-Lombard, José Ortiz, and Christine Pout. A review on buildings energy consumption information. *Energy and buildings*, 40(3):394–398, 2008.

[6] P. Paatero and U. Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):111–126, 1994. cited By 1714.

[7] Daniel D Lee and H Sebastian Seung. Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems*, pages 556–562, 2001.

[8] Stephen A Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization*, 20(3):1364–1377, 2009.

[9] Chih-Jen Lin. Projected gradient methods for nonnegative matrix factorization. *Neural computation*, 19(10):2756–2779, 2007.

[10] Christos Boutsidis and Efstratios Gallopoulos. Svd based initialization: A head start for nonnegative matrix factorization. *Pattern Recognition*, 41(4):1350–1362, 2008.

[11] Patrik O Hoyer. Non-negative matrix factorization with sparseness constraints. *Journal of machine learning research*, 5(Nov):1457–1469, 2004.