On recent non-negative matrix factorization applications in finance

Hellinton Hatsuo Takada 1,2 and Julio Michael Stern 2

1 Quantitative Research, Itaú Asset Management, Av. Brig. Faria Lima 3500, 04538-132, São Paulo, SP, Brazil
2 Institute of Mathematics and Statistics, University of São Paulo, R. do Matão 1010, 05508-090, São Paulo, SP, Brazil

Abstract: Non-negative matrix factorization (NMF) is a multivariate data analysis technique for dimensionality reduction aimed to estimate non-negative factors and factor loadings from non-negative data. In the literature, NMF has a wide variety of applications from text mining to identification of concentrations in chemistry. Early, NMF was presented by Paatero and Tapper in 1994 under the name positive matrix factorization. The name NNMF was established by Lee and Seung in 1999. In this paper, we review some applications of NNMF in finance with a historical organization of the related literature. There are several financial applications from estimation of credit default risk, recognition of financial trends, development of asset allocation strategies, computation of correlation matrix decomposition, modeling of the term structure of interest rates, detection of bankruptcy to identification of market micro-structure patterns. Since factors are not observable, the financial decision-making process based on them is easier when they are interpretable. In general, NNMF provides more interpretable factors due to its non-negativeness when compared to other techniques such as principal component analysis. Ultimately, we assemble such ideas to associate NNMF with investment/risk factor concepts.

Keywords: factorization, finance, optimization

INTRODUCTION

Undoubtedly, a common idea behind several solutions for noise removal, model reduction, blind source separation, feasibility reconstruction, and so on, is to approach the original data by a lower dimension representation obtained via subspace or low rank approximation. In terms of factorization techniques, the most popular approach is the principal component analysis (PCA) which was introduced by Pearson (1901) and developed by Hotelling (1933). In the literature, there are other well-known matrix factorization techniques such as independent component analysis (ICA) (Comon, 1994), sparse linear models (Tibshirani, 1996), or even vector quantization (Gray, 1984).

Basically, the objective of low rank approximations is to reduce the dimensionality of the original data matrix \( D = [d_{ij}] \in \mathbb{R}^{m \times p}, m \land p \in \mathbb{N}_+ \), using an approximation \( \Delta = [\delta_{ij}] \in \mathbb{R}^{m \times p} \) such that

\[
D \cong \Delta = \Phi \Lambda,
\]

where \( \Phi = [\phi_{ij}] \in \mathbb{R}^{m \times k}; \lambda = [\lambda_{ij}] \in \mathbb{R}^{k \times p}; k \in \mathbb{N}_+ \) such that \( k \leq \min(m, p) \). For instance, when each column of \( D \) represents a time series, \( \Phi \) is the matrix containing the called factors or unobserved (latent) variables, \( \Lambda \) is the matrix containing the called factor loadings or weights; \( k \) represents the number of factors.

Conceptually, non-negative matrix factorization (NMF) is a multivariate data analysis technique for dimensionality reduction (Paatero and Tapper, 1994; Lee and Seung, 1999). Basically, it is a subspace or low rank approximation aimed to estimate non-negative factors and non-negative factor loadings from non-negative data. In other words, NMF attempts to find a correspondence for a high dimensional non-negative matrix as the product of two low dimensional matrices under the non-negativity constraint. Clearly, the dimensionality reduction is crucial to obtain parsimonious data representations and to improve the predictive capability of models through projections.

Obviously, factorization techniques allow the identification of hidden patterns and/or trends behind the observed data. Consequently, NMF has been applied in multiple fields as different as air pollution research (Hopke, 1991; Kim et al., 2003), image processing (Lee and Seung, 1999; Piper et al., 2004; Plemmons et al., 2004), sound processing (Smaragdis and Brown, 2003), text mining (Berry, 2001; Hastie et al., 2001; Xu et al., 2003; Pauca et al., 2004; Shahnaz et al., 2006), chemometrics (Paatero and Tapper, 1994; Hopke, 1985), bioinformatics (Brunet et al., 2004), complex networks (Long et al., 2007), and so on. Fortunately, the application fields are very diverse showing the spread of NMF over several areas and, unfortunately, it is impossible to cover here all of them. Additionally, the relationships between NMF and other unsupervised learning models in some application fields have also been addressed in the literature (Ding et al., 2005; Ding et al., 2008).

As previously described, there is an extensive number of works with NNMF applications. Inevitably, there is the necessity of organizing such works and there are some publications addressing the matter. For instance, Zhang and Zhang (2010) present a survey of some variations of the NNMF methodology together with some applications of NNMF in some fields related to clustering and data mining. Additionally, Smaragdis et al. (2014) try to summarize the
On recent non-negative matrix factorization applications in finance

existing NNMF methodologies and extensions with some applications for blind source separation. Finally, Cichocki et al. (2009) present a complete book about NNMF techniques and applications for data analysis and blind source separation.

In the finance related publications, there are several applications of NNMF in distinct areas. The objective of this paper is to review and organize such references to assist the interested researchers to advance in such studies. Concisely, there are applications of NNMF in different areas of finance like estimation of credit default risk (Vandendorpe et al., 2008), recognition of financial trends (Drakakis et al., 2008), development of asset allocation strategies (Frein et al., 2008), computation of correlation matrix decomposition (Sonneveld et al., 2009), detection of bankruptcy (Chen et al., 2011), modeling of the term structure of interest rates (Takada and Stern, 2015a) and identification of market microstructure patterns (Takada and Stern, 2015b). Ultimately, our idea is to associate NNMF with investment and risk factor concepts.

Briefly, the paper is organized as follows: firstly, the concept of NNMF is presented together with some earliest publications on the subject and a condensed description of the main variations of the factorization techniques available in the literature. Subsequently, we organize and report some recent and diverse publications with applications of NNMF in finance with a succinct description of the associated contexts to give an idea of the miscellaneous areas. Over and above that, we introduce a discussion relating investment and risk factors to NNMF. Finally, we provide some conclusion in conjunction with more comments about the theme at the end.

Non-negative matrix factorization

Historically, the demand for non-negative constraints in decompositions started a long time ago in fields such as chemometrics, spectrophotometry, chromatography, and so on. By the way of illustration, there is the self modeling curve resolution technique, a method for determining the shapes of two overlapping functions from an observed set of additive mixtures of them, that was developed by Lawton and Sylvestre (1971). Aligned to the current idea of NNMF, a technique called positive matrix factorization was invented by Paatero and Tapper (1994). Actually, the name NNMF was established by Lee and Seung (1999).

Conceptually, considering a data matrix \( D \) that is non-negative \( D = [d_{ij}] \in \mathbb{R}^{m \times p}, m \wedge p \in \mathbb{N}_+ \) and given the number of factors \( k \), the NNMF approach aims to find the following approximation

\[
D \cong \bar{D} = \bar{F}\bar{L},
\]

where \( \bar{D} = [\bar{d}_{ij}] \in \mathbb{R}^{m \times p} \) is the low rank approximation; \( \bar{F} = [\bar{f}_{ij}] \in \mathbb{R}^{m \times k} \) is the non-negative matrix containing the factors; \( \bar{L} = [\bar{l}_{ij}] \in \mathbb{R}^{k \times p} \) is the non-negative matrix containing the factor loadings. For instance, when each column of \( D \) represents a time series, the columns of \( \bar{F} \) are the factors and the rows of \( \bar{L} \) are the factor loadings.

Typically, the NNMF is implemented through an optimization problem with an objective function and a set of constraints. Specifically, different objective functions have been proposed leading to a number of variants of NNMF algorithms. The commonly used objective functions are the squared error or Euclidean distance (Paatero and Tapper, 1994) and different types of divergence measures (e.g. Kullback-Leibler divergence (Lee and Seung, 2001), Itakura-Saito divergence (Fevotte et al., 2009), alpha-beta divergence (Cichocki et al., 2011), Bregman divergence (Dhillon and Sra, 2005)). See Cichocki et al. (2009) for a comprehensive review of divergences used. Additionally, the objective functions can be modified to reflect the application needs. For instance, penalty function terms can be added in order to enforce sparsity or to enhance smoothness in the solution.

In particular, there are a lot of interest in enforcing sparsity in the solutions. Some researchers have also shown that the standard NNMF model does not necessarily give the correct part-of-whole representations (Li et al., 2001; Hoyer, 2004), hence many efforts have been done to improve the sparseness of NNMF in order to identify more localized features that are building parts for the whole representation. The relationship between semi-supervised clustering, where some background information concerning the pairwise relations of some samples are known and used into the clustering model in order to guide the clustering process, and NNMF was previously focused in the literature (Kulis et al., 2005; Li and Ding, 2006; Chen et al., 2007).

Computationally, optimization schemes were derived to minimize the underlying objective functions by iterative update rules for solving the NNMF. The algorithms were proposed more or less following the principles of alternating direction iterations, the projected Newton, the reduced quadratic approximation, and the descent search. In general, specific implementations can be categorized into alternating least squares algorithms (Paatero and Tapper, 1994), multiplicative update algorithms (Lee and Seung, 1999; Lee and Seung, 2001; Hoyer, 2002), gradient descent algorithm, and hybrid algorithm (Piper et al., 2004; Paush et al., 2004). Some general assessments of these methods can be found in Liu and Yi (2003), Tropp (2003) and Chu et al. (2004). Although schemes and approaches are different, the major part of the numerical methods is essentially centered around satisfying the first order optimality conditions derived from the Kuhn-Tucker theory.

There are many other matrix factorizations with different constraints. It is difficult to give an exhaustive list of variations of matrix factorization techniques. Consequently, we list some of them for illustration purposes: the binary
matrix factorization (BMF) was presented for solving the bi-clustering problem (i.e. clusterization of the rows and the columns of a matrix simultaneously) (Cheng and Church, 2000; Prelić et al., 2006); the local non-negative matrix factorization (LNMF) was presented for learning a spatially localized, parts-based subspace representation of visual patterns and the objective function is defined to impose the localization constraint (Li et al., 2001; Feng et al., 2002); the sparse non-negative matrix factorization (SNNMF) was presented using several different approaches (Hoyer, 2002; Hoyer, 2004; Liu et al., 2003; Gao and Church, 2005; Pauca et al., 2006; Kim and Park, 2007; Pascual-Montano et al., 2006); the non-negative tri-matrix factorization (Tri-NMF) decomposes the original data into three instead of two matrices (Ding et al., 2006); the symmetric nonnegative matrix factorization (Symmetric-NMF) that imposes some symmetry constraints (Kuang et al., 2012); the CUR that stands for the low-rank approximation of an original matrix A into three matrices C, U, and R such that C is made from columns of A, R is made from rows of A, and the product CUR closely approximates A (Mahoney and Drineas, 2009); the convex non-negative matrix factorization (Convex-NMF) that imposes some convexity constraints (Ding et al., 2010); the semi-non-negative matrix factorization (Semi-NMF) that allows the non-negativity to be relaxed for some elements (Ding et al., 2010); the probabilistic latent semantic indexing (PLSI) (Ding and Peng, 2008; Gaussier and Goutte, 2005), specifically developed for text indexing, and its extension called probabilistic latent component analysis (PLCA) (Smaragdis and Shashanka, 2006); and so on.

APPLICATIONS IN FINANCE

Concisely, finance is the study of money management and it has several branches like financial economics, financial mathematics, behavioral finance, and so forth. Evidently, there are different data exploration needs and models in such areas. Therefore, the objective of the present section is to provide a historical review of applications of NNMF in distinct fields of finance. Particularly, we describe the corresponding context with related details and challenges. Regrettably, the number of papers being published every day is substantial and, consequently, we hope our presented list of NNMF applications in finance is as complete as possible at the time of publication of the paper.

In the credit risk area, the objective is the assessment of the risk of default on a debt that usually arises from a borrower failing to make required payments. Therefore, the credit risk is that of the lender and its measurement is crucial because it is related to the costs the borrower pays for the money. In December 1996, Credit Suisse Group introduced a credit risk management framework called CreditRisk+. In a few words, CreditRisk+ is an industry standard model to estimate the credit default risk of a portfolio of credit loans. It considers default rates as continuous random variables and incorporates the volatility of default rates in order to capture the uncertainty in the level of default rates. Additionally, the effects of background factors are incorporated into the model. Usually, it is necessary that the default probabilities of CreditRisk+ are apportioned using a number of non-negative factor loadings. Nonetheless, in practice the factor loadings are not available. On the other hand, default correlations are often available. Vandendorpe et al. (2008) deduced factor loadings from a given set of default correlations using NNMF presenting the corresponding optimization algorithm.

In the stock market, it has been reported in the related literature that the stock price movements do not behave independently of each other. Actually, they are mainly dominated by several underlying and unobserved factors. Hence, the identification problem of the underlying trends from the stock market data is an interesting problem, which can be solved by NNMF. Initially, Drakakis et al. (2008) focused the identification of common trends using NNMF. Their most interesting result is that the stocks of the same sector is not necessarily assigned into the same cluster and vice versa, which is of potential use to guide diversified portfolio, in other words, investors should diversify their money into not only different sectors, but also different clusters. Takada and Stern (2015c) empirically concluded that NNMF seems to be a better approach to capture such trends when compared to PCA.

In terms of financial crisis, there is a lot of effort to develop models for distress prediction. Obviously, an early detection of crisis is important for policy makers and financial agents. Consequently, building appropriate financial distress prediction model based on the extracted discriminative features is more and more important under the background of financial crisis. In particular, Ribeiro et al. (2009) presents a prediction model which is indeed a combination of k-means, NNMF and support vector machine (SVM). The basic idea is to train a SVM classifier in the reduced dimensional space which is spanned by the discriminative features extracted by NMF, the algorithm of which is initialized by k-means.

In the asset allocation area, the common idea is the implementation of an investment strategy that seeks to balance risk and reward according to some future expectations and an investment policy statement through the determination of the proportion of each asset or asset class in the portfolio. Clearly, the assets or asset classes could be clustered using a latent trend based approach assisting in the allocation process. Frén et al. (2008) applied modified versions of NNMF called semi-NMF and sparse-semi-NMF in a portfolio of simulated assets. The semi-NMF decomposes a data matrix into two other matrices and it has a non-negative constraint only in one of the resulting matrices. The sparse-semi-NMF also decomposes the data matrix into matrices as the semi-NMF but an sparsity constraint is imposed. Frén et al. (2008) clustered the assets into latent trend based groupings and concluded that the approach should be considered in the investment process to reduce the risk in portfolio selection.

In the financial econometrics, several statistical techniques are applied in finance. Remarkably, statistical time series techniques are extensively used in financial modeling of asset returns. In particular, asset returns presents
heteroskedasticity, i.e. the standard deviations of the asset returns over-time is non-constant. Usually, these periods of large and small variances are clustered together which gives rise to the term volatility clustering and the generalized autoregressive conditional heteroskedasticity (GARCH) model is a common approach. Fréin et al. (2009) investigate the relationship between periods of low and high volatility across stock log-returns using NNMF. They formulated a latent autoregressive conditional heteroskedasticity (ARCH) factor model and presented the robust NNMF, which is robust in the presence of noise and suited to their low rank latent ARCH model. They modeled the forces that control the rise and fall of stock prices using variable volatility as features and reveals the dependencies between different return-series. Actually, they performed a clustering of stocks based on volatility trend.

The correlation matrices are very important in finance because it describes how several securities move in relation to each other. Additionally, correlation matrices are fundamental to portfolio management and diversification of investment strategies. Sonneveld et al. (2009) proposes a NNMF of a correlation matrix. The problem of interest is essentially finding the nearest low-rank correlation matrix, but with an additional non-negativity constraint. The nearest low-rank correlation matrix question has drawn broad attention in the financial community. For the purpose of decomposing a correlation matrix, a variety of methods have been proposed. To name a few, the called geometric programming approach, the Lagrange multiplier method, and majorization have been introduced for this purpose. However, none of these approaches accommodate an additional non-negativity constraint. Consequently, NNMF seems to be a natural choice.

In terms of detection of bankruptcy, it is a main concern when analyzing markets and companies. Specifically, bankruptcy is a legal status of a company that cannot repay the debts it owes to creditors. The detection of bankruptcy aims to predict the probability that the company may become bankrupt in the following years given a set of financial ratios that describe the situation of a company over a given period. Chen et al. (2011) have shown empirically using data from France that a modified version of NNMF called graph regularized non-negative matrix factorization (GNMF) seems to be suitable to detect bankruptcy. In other words, Chen et al. (2011) results show GNMF is applicable to explore the intrinsic structure of the high dimensional bankrupt data and enhance the predictive ability of rating classification models.

In the fixed income area, there are many studies involving interest rate models. Particularly, the modeling of the term structure of interest rates or the yield curve is of special interest. The yield curve is important in economy and finance because it reflects current expectations of market participants about future changes in the interest rates. There are several factor models for the yield curve: Litterman and Scheinkman (1991) proposed a three factor model based on PCA and suggested names for these factors: level, steepness (or slope) and curvature. Since then, these factors became attributes of the yield curve. Independently, Nelson and Siegel (1987) published a parametric model for the yield curve which was rewritten by Diebold and Li (2006) in terms of the yield curve attributes. Takada and Stern (2015a) applied NNMF to obtain factors from the term structure of interest rates and the procedure is compared with other very popular techniques: PCA and Nelson-Siegel model. The NNMF approximation for the term structure of interest rates is better in terms of fitting.

In the following subsection, we review the details of an application in finance related to market microstructure to illustrate the advantages of NNMF when compared with other techniques.

**Sample application in finance: market microstructure**

In the market microstructure, the objective is the detailed understanding of how exchanges occur in markets. The intraday trading volume of a security is the total amount of traded contracts distributed over the day. Consequently, the intraday trading volume captures part of the intraday trading activity and represents a proxy for the intraday liquidity of a market. Clearly, the intraday trading volume is important when developing execution strategies. In the literature, the intraday trading volume for equities has been reported to possess an intraday U-shaped pattern, i.e. heavy trading volume at the beginning and at the end of the trading day and the relatively light trading volume at the middle of the trading day (for example, see Jain and Joh (1988) for an early work on the subject). As a consequence, several approaches were developed to model the intraday trading volume (e.g. Panas (2005) uses a beta density function to fit the U-shaped pattern).

In this section, we review the application of NNMF to capture the intraday trading volume pattern presented by Takada and Stern (2015b). Here, the data matrix $D$ represents the intraday traded volume in number of contracts, $p$ represents the number of time bins during a trading day (e.g. a time bin is from 10:00 a.m. until 11:00 a.m.) and $m$ is the amount of different days and/or equity names. The securities selected are from the Brazilian stock exchange (BM&F Bovespa) for the period from April 2013 until September 2013. Additionally, we present the results obtained with time bins equal to 1 hour with a total of $p = 8$ time bins a day. The idea is to investigate factors and factor loadings from intraday traded volume for only one individual equity name and for a set of different equities. Consequently, we focus not only individual estimations but also joint estimations to obtain intraday trading volume patterns.

In particular, the ticker names of the equities chosen are: PETR3, PETR4, VALE3, VALE5, BBDC3 and BBDC4. PETR3 is the ordinary stock of an oil and gas company, PETR4 is the preferential stock of the same company and one of the most liquid shares traded in Brazil; VALE3 is the ordinary stock of an mining company and VALE5 is the...
preferential stock of the same company; BBDC3 is the ordinary stock of a financial services company and BBDC4 is the preferential stock of the same company.

Using PCA, we show in Figures 1 and 2 the first four factor loadings and the percentage of total explained variance of the data according to the number of factors for PETR4 and for the set of equities, respectively. It is possible to notice that the first factor loading (with higher percentage of total explained variance) for both cases has the U-shaped pattern. However, the second, third and fourth factor loadings do not possess a direct interpretation. Additionally, the percentage of total explained variance for the first factors are very low indicating the need of more factors (not only one) to be used in the PCA based factor model for intraday trading volume (for example, in both cases illustrated in Figures 1 and 2, it is necessary to include at least three factors to explain more than 70% of the total data variance).

Using NNMF, the first objective is to explore the existence of one main factor ($k = 1$) in the intraday traded volume. In Figure 3, we present the obtained factor loadings for our data set. As expected, it is possible to identify the U-shaped pattern observing the factor loadings of PETR4 and the set of equities. For the estimation using the set of equities, we present in Figure 4 the first factors obtained from PCA and from NNMF ($k = 1$) for each equity. The NNMF factors can be easily interpreted as the volume level for each equity. Consequently, it is clear that PETR4 has the highest traded volume over the time while BBDC3 has the lowest traded volume over the time. Concerning the PCA factors, since they become negative such an interpretation is not possible.

Using NNMF, the first objective is to explore the existence of one main factor ($k = 1$) in the intraday traded volume. In Figure 3, we present the obtained factor loadings for our data set. As expected, it is possible to identify the U-shaped pattern observing the factor loadings of PETR4 and the set of equities. For the estimation using the set of equities, we present in Figure 4 the first factors obtained from PCA and from NNMF ($k = 1$) for each equity. The NNMF factors can be easily interpreted as the volume level for each equity. Consequently, it is clear that PETR4 has the highest traded volume over the time while BBDC3 has the lowest traded volume over the time. Concerning the PCA factors, since they become negative such an interpretation is not possible.

Using NNMF, the first objective is to explore the existence of one main factor ($k = 1$) in the intraday traded volume. In Figure 3, we present the obtained factor loadings for our data set. As expected, it is possible to identify the U-shaped pattern observing the factor loadings of PETR4 and the set of equities. For the estimation using the set of equities, we present in Figure 4 the first factors obtained from PCA and from NNMF ($k = 1$) for each equity. The NNMF factors can be easily interpreted as the volume level for each equity. Consequently, it is clear that PETR4 has the highest traded volume over the time while BBDC3 has the lowest traded volume over the time. Concerning the PCA factors, since they become negative such an interpretation is not possible.
On recent non-negative matrix factorization applications in finance

Figure 4 – The PCA first factors (left figure) and the NNMF (k = 1) factors (right figure) for each equity in the set of equities.

Using NNMF with two factors (k = 2), we present the obtained factor loadings for our data set in Figure 5. As expected, they are easily interpretable. We identify an increasing end of day pattern (first factor loading) and a decreasing start of day pattern (second factor loading) observing both the factor loadings for PETR4 and the set of equities. In Figure 6, we present the first and second factors obtained for each selected equity. The two factors can be easily interpreted as the volume level at the beginning and at the end of the day for each equity name. Consequently, using the two NNMF factors a financial analysis could study individually the volume levels at the beginning and at the end of the day. Again, concerning the PCA factors, since they become negative such an interpretation is not possible.

Figure 5 – The NNMF (k = 2) factor loading for PETR4 (left figure) and the NNMF (k = 2) factor loading for the set of equities (right figure).

Figure 6 – The NNMF (k = 2) first factors (left figure) and the NNMF (k = 2) second factors (right figure) for each selected equity.

DESIGN OF INVESTMENT FACTORS

For modern investments, the idea of investment factors is widespread. According to Bender et al. (2013), a factor can be thought of as any characteristic relating a group of securities that is important in explaining their return and risk. Additionally, Bender et al. (2013) also state that since factors cannot be directly observed, there is a vigorous debate about how to define and estimate them. The use of factors in finance is not new, for instance, the arbitrage pricing theory (APT) proposed by Ross (1976) states that the expected return of a financial asset can be modeled as a function of various factors such as macroeconomic or theoretical market indices.

Categorically, there are several ways of classifying factors. In the context of financial investments, Connor (1995) presents three main categories of factors: macroeconomic, statistical and fundamental. The first category, macroeconomic factors, makes use of economic time series to capture the permeating behavior of financial asset returns.
According to Connor (1995), the macroeconomic factors are the simplest and most intuitive type of factors. Consequently, they are extensively used in practice by investment professionals in the decision making process. Usually, the observable economic time series used as basis for building macroeconomic factors are inflation, activity, yield curves, and so on. For illustration purposes, see Chen et al. (1986) for one of the earliest most well-known models in the literature.

The second category, statistical factors, is generated by models trying to identify latent variables using statistical techniques such as PCA or ICA, where the factors are not pre-specified in advance. The statistical factors have been used in several models in finance and, according to Connor (1995), they outperform the macroeconomic factor models in terms of explanatory power. Unfortunately, the major part of the statistical factors does not possess direct interpretations by the financial analysts and economists. In other words, they should possess an intuitive meaning related to financial or economic concepts or the possibility of direct identification to observable financial and economic variables.

Finally, the third category, fundamental factors, are the most widely used factors in finance. Usually, they capture stock features such as industry sector, some country membership, valuation ratios, prices and negotiation volumes based indicators, and so on. In the last decades, the most popular fundamental factors in the related literature are related to the following concepts: value (it is a reference to the value stocks - underpriced stock according to some fundamental analysis), growth (it is a reference to the growth stocks - stocks expected to present returns at an above-average rate relative to its peers), size (it captures the size of the companies) and momentum (it captures some trends the stock prices). Fama and French (1992, 1993), seminal researches related to fundamental factors, introduced a model to explain the USA equity market returns using three fundamental factors: the market factor (it is based on the expected market return and comes from the traditional capital asset pricing model), the size factor (it is based on the small versus large capitalization stocks) and the value factor (it represents the stocks with low versus the stocks with high book to market).

According to Bender et al. (2013), the question of which type of factor to use and the best way to construct a model and specify factors continue to be debated. Obviously, the categorization of factors in macroeconomic, statistical and fundamental is not perfect. It is possible to use statistical techniques such as PCA to help to build macroeconomic or fundamental factors. Practitioners discern the macroeconomic factors as easy to interpret economically, the fundamental factors as also easy to interpret using financial and accounting concepts, and statistical factors as not necessarily interpretable. Pragmatically, decision-making based on intuitive and interpretable factors are taken more easily.

As it is evident from the literature, NNMF techniques are surprisingly more effective in extracting perceptually meaningful factors from complex mixtures than other statistical factor models such as PCA or ICA. Specifically, several existent applications of NNMF in finance points to the intuitiveness and interpretability of the obtained factors. Obviously, such an interpretation often does not apply to decompositions that allow negative factors and factor loadings. The elements of the decompositions can cancel each other out, obscuring the latent components perceptual meaningfulness. It is also important to state that not only the factors but the joint behavior of factors and factor loadings should be considered intuitive and interpretable to easy the decision-making process.

**FINAL REMARKS**

The applications of NNMF in finance are numerous and there are many other possible applications. The authors hope the paper will help the interested readers to understand the potentials of the applications of NNMF in finance. The intuitiveness and interpretability of the resulting factors, and the possibility of imposing additional restrictions are essential to financial investment application success. The investors and decision-makers need to understand the rationality behind the investment strategies and, consequently, factors to feel comfortable to take decisions based on them.

**ACKNOWLEDGMENTS**

This research was supported by projeto FAPESP-BG CEPID 2014/50279-4.

**REFERENCES**


On recent non-negative matrix factorization applications in finance


Pauca, V. P.; Shahmaz, F.; Berry, M. W.; Plemmons, R. J., 2004, “Text mining using non-negative matrix factorizations”, Proceedings of the 4th SIAM International Conference on Data Mining, Orlando, FL, USA.


On recent non-negative matrix factorization applications in finance


RESPONSIBILITY NOTICE

The authors are the only responsible for the material included in this paper.