Journal of Banking & Finance xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Risk evaluations with robust approximate factor models ☆

Ray Yeutien Chou [a], Tso-Jung Yen [b], Yu-Min Yen [c,*]

[a] Institute of Economics, Academa Sinica, 128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan
[b] Institute of Statistical Science, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 11529, Taiwan
[c] Department of International Business, National Chengchi University, 64, Section 2, Zhi-nan Road, Wenshan, Taipei 116, Taiwan

ARTICLE INFO

Article history:
Received 19 February 2015
Accepted 25 May 2016
Available online xxxx

JEL classification:
C58
G12
G31

Keywords:
Approximate factor model
PCA
Norm penalty
Common factor
Idiosyncratic risk
VaR

ABSTRACT

Approximate factor models and their extensions are widely used in economic analysis and forecasting due to their ability to extracting useful information from a large number of relevant variables. In these models, candidate predictors are typically subject to some common components. In this paper we propose a new method for robustly estimating the approximate factor models and use it in risk assessments. We consider a class of approximate factor models in which the candidate predictors are additionally subject to idiosyncratic large uncommon components such as jumps or outliers. By assuming that occurrences of the uncommon components are rare, we develop an estimation procedure to simultaneously disentangle and estimate the common and uncommon components. We then use the proposed method to investigate whether risks from the latent factors are priced for expected returns of Fama and French 100 size and book-to-market ratio portfolios. We find that while the risk from the common factor is priced for the 100 portfolios, the risks from the idiosyncratic factors are not. However, we find that model uncertainty risks of the idiosyncratic factors are priced, suggesting that with effective diversifications, only the predictable idiosyncratic risks can be reduced, but the unpredictable ones may still exist. We also illustrate how the proposed method can be adopted on evaluating value at risk (VaR) and find it can delivery comparable results as the conventional methods on VaR evaluations.

1. Introduction

Approximate factor models and their extensions are widely used in economic analysis and forecasting due to their ability to extracting useful information from a large number of relevant variables (Stock and Watson, 2002; Bernanke et al., 2005; Ludvigson and Ng, 2009). In these models, data generating processes are often specified as a linear combination of relevant common factors and error terms. Estimating these models can pose some difficulties when the relevant common factors are unobservable. An important goal for estimating such models is therefore to identify the latent common factors and their factor loadings. Popular methods for carrying out this estimation task include the maximum likelihood method (MLE), Markov Chain Monte Carlo (MCMC) and Principal Component Analysis (PCA) method. Nevertheless researchers in econometrics often estimate these models with high dimensional data. Therefore PCA method, which is less computational intensive than MLE and MCMC, is often more preferred in practice.

Although PCA method has a computational advantage, it is widely known that the method may fail to yield accurate estimations of the latent factors and factor loadings when large idiosyncratic uncommon components are present in the data (Jolliffe, 2002). To overcome this difficulty, we propose a simple and efficient method for estimating latent factors and factor loadings. This estimation method aims to reduce estimation biases in the latent common factors and their factor loadings by simultaneously disentangling and estimating the common and uncommon components. To develop this method, we first formulate the estimation problem as a penalized least squares problem in which a norm penalty function is imposed on the uncommon components. We then solve the estimation problem by building an algorithm to iteratively solve a principal component analysis (PCA) problem and a one dimensional shrinkage estimation problem. The algorithm can flexibly incorporate with the methods for selecting the number of common components. We call the proposed estimation method P-PCA method (*Penalized least squares plus PCA method*).

☆ We thank Carol Alexander (Editor) and an anonymous referee whose suggestions help to improve the initial version of the paper. For helpful comments, we thank Haiqiang Chen, Ren-Raw Chen, Iftekhar Hasan, Ho-Mou Wu and seminar participants in 2014 Financial Innovations and Bank Regulation Conference in Xiamen University, Academia Sinica, National Chengchi University, National Taiwan University and National Yunlin University of Science and Technology. Yu-Min Yen's work was supported by the Ministry of Science and Technology, Taiwan, Research Grant 103-2410-H-004-213.

* Corresponding author.
E-mail addresses: rchou@econ.sinica.edu.tw (R.Y. Chou), tjyen@stat.sinica.edu.tw (T.-J. Yen), yyu_min@nccu.edu.tw (Y.-M. Yen).

http://dx.doi.org/10.1016/j.jbankfin.2016.05.008
0378-4266/© 2016 Elsevier B.V. All rights reserved.

Recently many different approximate factor models and their corresponding estimation procedures were developed. Moench et al. (2013) proposed a multilevel factor model for large panel data with between-block variations and idiosyncratic noise, and developed an estimation procedure that can both separate block-level shocks and genuinely common factors and achieve dimension reduction. Ando and Bai (2014) proposed a multifactor model for data with a large number of observable factors and unobservable common and group-specific pervasive factors. Their estimation procedure for the model can simultaneously select relevant observable factors and determine the number of common and group-specific unobservable factors. Cheng et al. (2016) proposed a factor model in which both factor loadings and number of factors can have a behavior of structure break. They adopted a shrinkage estimator that can simultaneously and consistently estimate the number of common factors before and after the structure break. Their estimation procedure is carried out by solving a convex optimization with principal components of the data matrix as its inputs.

A main difference between the methods proposed by aforementioned research in previous paragraph and P-PCA method is that we develop the proposed method by considering a data generating process in which certain large idiosyncratic uncommon components are present. This means that in the data generating process observations are occasionally blurred by extremely large noises such as asset price jumps. They are not broken by a permanent change of common factors or factor loadings. Indeed, under suitable assumptions on the idiosyncratic uncommon components, one can still estimate the factors and factor loadings consistently by using PCA method (Bai and Ng, 2002; Stock and Watson, 2002). However, through intensive simulations we show that the proposed P-PCA estimation procedure can outperform PCA method in term of finite sample efficiency when estimating model parameters under a wide range of model settings. In addition, we discuss how the proposed method can be used to deal with a more general data structure, such as panel data.[1] Throughout these works, we believe the proposed method can serve as a complementary tool for robust estimations rather than a competitive approach to those established approximate factor models.

We apply the proposed method on two empirical analyses with real data. We first investigate whether risks from the latent factors are priced for expected returns of Fama and French 100 size and book-to-market ratio portfolios. Recently Ando and Bai (2014) analyzed possible latent common and group-specific pervasive factors of the expected stock returns in China stock market. They assumed that expected returns of stocks traded in the same exchanges are governed by the same latent within-group common factors. They found the expected returns of stocks traded in different stock exchanges are affected by different observable and latent factors. Unlike their research, we focus on how both the latent common and idiosyncratic, uncommon factors affect cross sectional expected returns. Our analysis relies on decomposing the noise term in the Fama and French three factor model into latent common and idiosyncratic factors. We find that for the 100 portfolios, risk from the common factor is priced but risks from the idiosyncratic factors are not. The latter result is consistent with Arbitrage Pricing Theory: The idiosyncratic risk of a well diversified portfolio should be negligible. It should not have an effect on expected return of the well diversified portfolio. However, we also find that model uncertainty risks of the idiosyncratic factors are priced. The result implies that with effective diversifications, only the predictable idiosyncratic risks can be reduced, but the unpredictable ones may still exist. We next use the proposed method to evaluate monthly value at risk (VaR) of individual stocks and a portfolio of these stocks. We find that the proposed method can deliver comparable results as the conventional methods on VaR evaluations.

The rest of paper is organized as follows. In Section 2 we review PCA method and then introduce P-PCA estimation method. We next discuss how to select number of the latent common factors in the estimation procedure. In Section 3 we discuss some possible extensions of our proposed method. In Section 4 we perform empirical applications. Section 5 is the conclusion.

2. Methodology

In this Section we describe a method for estimating an approximate factor model in which the candidate predictors are subject to idiosyncratic uncommon components. Specifically we assume the N dimensional time series of candidate predictors $\mathbf{X}_t$ and the variable to be forecast Y_t are subject to the following data generating process:

$$\mathbf{X}_t = \mathbf{\Lambda}\mathbf{F}_t + \mathbf{J}_t + \mathbf{e}_t, \quad (1)$$
$$Y_{t+h} = \boldsymbol{\beta}_F^{\mathbf{T}}\mathbf{F}_t + \boldsymbol{\beta}_W^{\mathbf{T}}\mathbf{W}_t + \varepsilon_{t+h}, \quad (2)$$

for $t = 1, \ldots, T$, where $\mathbf{\Lambda}$ is an $N \times r$ factor loading matrix, $\mathbf{F}_t$ is an $r \times 1$ vector of latent factors, $\mathbf{e}_t$ is an $N \times 1$ vector of measurement errors, and $\mathbf{J}_t$ is an $N \times 1$ vector of the idiosyncratic uncommon components. In addition, $\boldsymbol{\beta}_F$ is an $r \times 1$ vector of regression coefficients corresponding to latent factor $\mathbf{F}_t$, $\mathbf{W}_t$ is an $m \times 1$ vectors of observable exogenous variables, and $\boldsymbol{\beta}_W$ is an $m \times 1$ vector of regression coefficients corresponding to $\mathbf{W}_t$. The index h is the forecast horizon, and Y_{t+h} and ε_{t+h} are the variable to be forecast and error term h periods ahead, respectively. The setting is similar to the dynamic factor model considered in Stock and Watson (2002) except $\mathbf{X}_t$ has an additional idiosyncratic uncommon component $\mathbf{J}_t$. By assuming that occurrences of the uncommon components are rare, $\mathbf{J}_t$ is generically a sparse vector, i.e. some of its elements are zero, and practically it can be viewed as a jump or outlier in $\mathbf{X}_t$.

Below we review PCA method for estimating the latent factors $\mathbf{F}_t$ and factor loadings $\mathbf{\Lambda}$. Define $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_T)^{\mathbf{T}}$, $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_T)^{\mathbf{T}}$ and $\mathbf{J} = (\mathbf{J}_1, \ldots, \mathbf{J}_T)^{\mathbf{T}}$. Suppose $N > T$ and the number of factors r is known. Without the term $\mathbf{J}_t$, we can estimate the factor matrix $\mathbf{F}$ and factor loading matrix $\mathbf{\Lambda}$ by solving the following optimization problem:

$$\min_{\mathbf{F},\mathbf{\Lambda}} \frac{1}{TN} \|\mathbf{X} - \mathbf{F}\mathbf{\Lambda}^{\mathbf{T}}\|_F^2, \text{ subject to } \frac{\mathbf{F}^{\mathbf{T}}\mathbf{F}}{T} = \mathbf{I}_r. \quad (3)$$

Here $\|.\|_F$ is the Frobenius norm. The optimization problem (3) is closely related to the principal component analysis (PCA). The estimated factor matrix $\hat{\mathbf{F}}$ can be obtained by multiplying $\sqrt{T}$ with a matrix containing the eigenvectors corresponding to the largest r eigenvalues of the $T \times T$ matrix $\mathbf{X}\mathbf{X}^{\mathbf{T}}$. Given $\hat{\mathbf{F}}$, the estimated factor loading matrix $\hat{\mathbf{\Lambda}}$ can be obtained by using the least squares method, i.e. $\hat{\mathbf{\Lambda}} = \left(\left(\hat{\mathbf{F}}^{\mathbf{T}}\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}^{\mathbf{T}}\mathbf{X}\right)^{\mathbf{T}} = \mathbf{X}^{\mathbf{T}}\hat{\mathbf{F}}/T$. On the other hand, when $T \geqslant N$, we can estimate the factor and factor loading matrices by solving the problem (3) with the constraint $\mathbf{F}^{\mathbf{T}}\mathbf{F}/T = \mathbf{I}_r$ replaced by $\mathbf{\Lambda}^{\mathbf{T}}\mathbf{\Lambda}/N = \mathbf{I}_r$. In this situation the estimated factor loading matrix $\bar{\mathbf{\Lambda}}$ can be obtained by multiplying $\sqrt{N}$ with a matrix containing the eigenvectors corresponding to the largest r eigenvalues of the $N \times N$ matrix $\mathbf{X}^{\mathbf{T}}\mathbf{X}$. Given $\bar{\mathbf{\Lambda}}$, the estimated factor matrix $\bar{\mathbf{F}}$ can be obtained by using the least squares method, i.e. $\bar{\mathbf{F}} = \left((\bar{\mathbf{\Lambda}}^{\mathbf{T}}\bar{\mathbf{\Lambda}})^{-1}\bar{\mathbf{\Lambda}}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\right)^{\mathbf{T}} = \mathbf{X}\bar{\mathbf{\Lambda}}/N$. Now define $\mathbf{Z} = \mathbf{F}\mathbf{\Lambda}^{\mathbf{T}}, \hat{\mathbf{Z}} = \hat{\mathbf{F}}\hat{\mathbf{\Lambda}}^{\mathbf{T}}$ and $\bar{\mathbf{Z}} = \bar{\mathbf{F}}\bar{\mathbf{\Lambda}}^{\mathbf{T}}$. Here the matrices $\hat{\mathbf{Z}}$ and $\bar{\mathbf{Z}}$ can be viewed as low rank approximations to the matrix $\mathbf{X}$. It is known that $\hat{\mathbf{Z}} = \bar{\mathbf{Z}}$, and hence the objective function $\|\mathbf{X} - \mathbf{F}\mathbf{\Lambda}^{\mathbf{T}}\|_F^2/(TN)$ has the same value under the two optimal solutions $\left(\hat{\mathbf{F}}, \hat{\mathbf{\Lambda}}\right)$ and $(\bar{\mathbf{F}}, \bar{\mathbf{\Lambda}})$.

[1] Relevant simulation results are shown in Appendices A and B.