SFB 649 Discussion Paper 2016-018

Factorisable Sparse Tail Event Curves with Expectiles

Wolfgang K. Härdle*
Chen Huang*
Shih-Kang Chao*²

* Humboldt-Universität zu Berlin, Germany
*² Purdue University, United States of America

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Factorisable Sparse Tail Event Curves with Expectiles*

Oberwolfach Report: New Developments in Functional and Highly Multivariate Statistical Methodology

Wolfgang K. Härdle†, Chen Huang‡, Shih-Kang Chao§

March 15, 2016

JEL classification: C38, C55, C61, C91, D87

Keywords: multivariate functional data, high-dimensional M-estimators, nuclear norm regularizer, factor analysis, expectile regression, fMRI, risk perception

Data are observed more and more in form of curves, thus prompting a joint modelling to find out common patterns and also individual variations. Real data curve modelling occurs e.g. in neuroeconomics, wind speed analysis, demographics among many other disciplines.

Functional data analysis studies variation of random objects in a high dimensional contact and provides insight into main factors, typically extracted as principal components via a Karhunen-Loève decomposition. However, in a variety of applications one is more interested in the tail behavior rather than the variations around the mean. Thus the analysis of curve variation is around a tail event curve (TEC) rather than around a mean curve as in functional PCA. TECs may be identified through tail probabilities or more general through functions based on conditional tail events. Modeling such Tail Event Curves (TEC) requires to deviate from Hubert ℓ_2 geometry and to introduce asymmetric norms or check functions. For example, quantile regression is a widely used method can be exploited to grasp the whole information on the conditional distribution and especially the tail structure, which plays crucial roles in risk management. Concerning multivariate

*Financial support from the Deutsche Forschungsgemeinschaft via CRC 649 ”Economic Risk” and IRTG 1792 ”High Dimensional Non Stationary Time Series”, Humboldt-Universität zu Berlin, is gratefully acknowledged.

†Humboldt-Universität zu Berlin, C.A.S.E. - Center for Applied Statistics and Economics, Spandauer Str. 1, 10178 Berlin, Germany. Sim Kee Boon Institute for Financial Economics, Singapore Management University, 50 Stamford Road, Singapore 178899, Singapore.

‡Humboldt-Universität zu Berlin, C.A.S.E. - Center for Applied Statistics and Economics, Spandauer Str. 1, 10178 Berlin, Germany. Email: chen.huang@hu-berlin.de.

§Department of Statistics, Purdue University, 250 N University St., West Lafayette IN 47907-2066, U.S.A.

quantile regression, many previous works study in this direction under different frameworks. But none of them worked in high-dimensional case. [2] introduced factorisable sparse tail event curves (FASTEC) method to implement high-dimensional multivariate quantile regression.

fMRI risk perception analysis requires to study the shape (e.g., amplitude, delay, and duration) of the estimated hemodynamic response function (HRF) to particular tasks answered by every individual. More noteworthily, extreme behaviors of the response function may reveal unobserved neuronal activation information. Therefore, we need a global measure which can capture the tail moments and be more sensitive to the outliers. Expectiles can be a better choice than quantiles in consideration of extremes although it is not robust. This fact motivates us to build an expectile based FASTEC model.

Denote $\mathbf{Y} = (Y_{ij}) \in \mathbb{R}^{n\times m}$ as the multivariate curves we want to jointly model, where n is the length of observations and m is the number of curves. $\{\boldsymbol{X}_i\}_{i=1}^n \in \mathbb{R}^p$ are the covariates with dimension p, e.g., B-spline basis. Both p and m are allowed to tend to infinity with the sample size n (but no quicker than n).

Let $e_j(\tau|\boldsymbol{X}_i)$ be the conditional expectile function of Y_{ij} given $\boldsymbol{X}_i$, for $i = 1,\dots,n$ and $j = 1,\dots,m$ with $\tau \in (0,1]$, and approximate it by a linear factor model,

$$Y_{ij} = e_j(\tau|\boldsymbol{X}_i) + u_{ij}$$

$$e_j(\tau|\boldsymbol{X}_i) = \sum_{k=1}^{r} \psi_{j,k}(\tau) f_k^\tau(\boldsymbol{X}_i), \tag{1}$$

where $f_k^\tau(\boldsymbol{X}_i)$ is the kth factor, r is the number of factors (much less than p), $\psi_{j,k}(\tau)$ are the factor loadings. Furthermore, factors are constructed by linear combination of covariates $\boldsymbol{X}_i$

$$f_k^\tau(\boldsymbol{X}_i) = \boldsymbol{X}_i^\top \boldsymbol{\varphi}_k(\tau). \tag{2}$$

Substituting (2) into (1) yields

$$e_j(\tau|\boldsymbol{X}_i) = \boldsymbol{X}_i^\top \boldsymbol{\gamma}_j(\tau), \tag{3}$$

with $\boldsymbol{\gamma}_j(\tau) = (\sum_{k=1}^r \psi_{j,k}(\tau)\varphi_{k,1}(\tau),\dots,\sum_{k=1}^r \psi_{j,k}(\tau)\varphi_{k,p}(\tau))^\top$ as the unknown coefficient vector. In multivariate case, what needs to be estimated becomes a $p\times m$ coefficient matrix $\mathbf{\Gamma}$, where $\boldsymbol{\gamma}_j(\tau)$ in (3) is the jth column of $\mathbf{\Gamma}$.

With increasing dimension of both explanatory and response variables one faces the difficulty of estimating a very high dimensional coefficient matrix. A natural way to reduce the burden of this estimation task is to introduce a penalty term. [4] proposed a penalization approach with nuclear norm, the sum of the singular values of the coefficient matrix as the penalty. Numerically the estimator can be easily obtained since it involves a convex optimization. Moreover, compare with previous traditional works such as reduced rank approach, the number of factors does not need to be predetermined. Dimension reduction and coefficient estimation can be done simultaneously.

To be more precise it is proposed to estimate the coefficient matrix $\mathbf{\Gamma}$ by solving:

$$\widehat{\mathbf{\Gamma}}_\lambda(\tau) = \arg \min_{\mathbf{\Gamma}\in\mathbb{R}^{p\times m}} F(\mathbf{\Gamma}), \tag{4}$$

$$F(\mathbf{\Gamma}) = (mn)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m} \rho_\tau \left(Y_{ij} - \mathbf{X}_i^\top \mathbf{\Gamma}_{\cdot j}\right) + \lambda \|\mathbf{\Gamma}\|_*, \tag{5}$$

$$\rho_\tau(u) = |\tau - \mathbf{1}\{u < 0\}| \, |u|^2 . \tag{6}$$

Nuclear norm $\|\mathbf{\Gamma}\|_*$ is defined by $\sum_{l=1}^{\min(p,m)} \sigma_l(\mathbf{\Gamma})$ given the singular values of $\mathbf{\Gamma}$: $\sigma_1(\mathbf{\Gamma}) \geq \sigma_2(\mathbf{\Gamma}) \geq \ldots \geq \sigma_{\min(p,m)}(\mathbf{\Gamma})$. The convexity of the nuclear norm results in a convex optimization problem that can be solved via various of efficient methods. The number of nonzero singular values of $\mathbf{\Gamma}$ is identified as r. A high dimension $p \times m$ is reduced to $r \times \max(p,m)$ by regularization, when $\mathbf{\Gamma}$ is sparse. After obtaining the $\widehat{\mathbf{\Gamma}}_\lambda(\tau)$ from (4), singular value decomposition (SVD) can be employed to estimate the factors and normalized factor loadings respectively.

Moreover, the loss function for expectile regression has a smooth convex function form. Combining with the nuclear norm penalty, we can use Fast Iterative Shrinkage-Thresholding Algorithm proposed by [1] to solve the optimization directly. Without smoothing the asymmetric absolute check function, the convergence rate in the iterative procedure is quicker than in quantile regression case. Based on the unified framework for high-dimensional M-estimators with decomposable regularizers provided by [3], the finite sample oracle properties of the estimator associated expectile loss and nuclear norm regularizer are studied formally in this paper.

As an empirical illustration, our model is applied on fMRI data to see if individual's risk perception can be recovered by brain activities. Results show that main factors can reflect the common patterns of curves. Factor loadings over different tail levels can help to find out the most risk-seeking and averse behaviours. Taking tail risks into consideration, individual's risk attitudes can be predicted more precisely, especially the extremes.

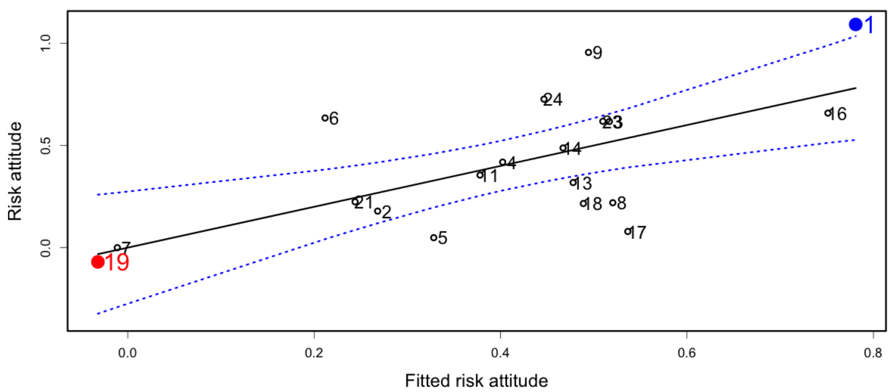

Figure 1: Horizontal axis denotes the fitted risk attitudes by the first factor loadings estimated from the brain data when $\tau = 0.1$, vertical axis denotes the risk attitudes parameters based on their choices. #1 and #19 are the most risk-averse and risk-seeking people respectively.
FASTEC_with_Expectiles

References

[1] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, Topology **2** (2009), 183–202.

[2] S.-K. Chao, W. K. Härdle and M. Yuan, *Factorisable sparse tail event curves*, SFB 649 Discussion Paper 2015-034 (2015).

[3] S. N. Negahban, P. Ravikumar, M. J. Wainwright and B. Yu, *A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers*, Statistical Science **27** (2012), 538–557.

[4] M Yuan, A. Ekici, Z. Lu and R. Monteiro, *Dimension reduction and coefficient estimation in multivariate linear regression*, Journal of the Royal Statistical Society: Series B **69** (2007), 329–346.

SFB 649 Discussion Paper Series 2016

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Downside risk and stock returns: An empirical analysis of the long-run and short-run dynamics from the G-7 Countries" by Cathy Yi-Hsuan Chen, Thomas C. Chiang and Wolfgang Karl Härdle, January 2016.
002 "Uncertainty and Employment Dynamics in the Euro Area and the US" by Aleksei Netsunajev and Katharina Glass, January 2016.
003 "College Admissions with Entrance Exams: Centralized versus Decentralized" by Isa E. Hafalir, Rustamdjan Hakimov, Dorothea Kübler and Morimitsu Kurino, January 2016.
004 "Leveraged ETF options implied volatility paradox: a statistical study" by Wolfgang Karl Härdle, Sergey Nasekin and Zhiwu Hong, February 2016.
005 "The German Labor Market Miracle, 2003 -2015: An Assessment" by Michael C. Burda, February 2016.
006 "What Derives the Bond Portfolio Value-at-Risk: Information Roles of Macroeconomic and Financial Stress Factors" by Anthony H. Tu and Cathy Yi-Hsuan Chen, February 2016.
007 "Budget-neutral fiscal rules targeting inflation differentials" by Maren Brede, February 2016.
008 "Measuring the benefit from reducing income inequality in terms of GDP" by Simon Voigts, February 2016.
009 "Solving DSGE Portfolio Choice Models with Asymmetric Countries" by Grzegorz R. Dlugoszek, February 2016.
010 "No Role for the Hartz Reforms? Demand and Supply Factors in the German Labor Market, 1993-2014" by Michael C. Burda and Stefanie Seele, February 2016.
011 "Cognitive Load Increases Risk Aversion" by Holger Gerhardt, Guido P. Biele, Hauke R. Heekeren, and Harald Uhlig, March 2016.
012 "Neighborhood Effects in Wind Farm Performance: An Econometric Approach" by Matthias Ritter, Simone Pieralli and Martin Odening, March 2016.
013 "The importance of time-varying parameters in new Keynesian models with zero lower bound" by Julien Albertini and Hong Lan, March 2016.
014 "Aggregate Employment, Job Polarization and Inequalities: A Transatlantic Perspective" by Julien Albertini and Jean Olivier Hairault, March 2016.
015 "The Anchoring of Inflation Expectations in the Short and in the Long Run" by Dieter Nautz, Aleksei Netsunajev and Till Strohsal, March 2016.
016 "Irrational Exuberance and Herding in Financial Markets" by Christopher Boortz, March 2016.
017 "Calculating Joint Confidence Bands for Impulse Response Functions using Highest Density Regions" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, March 2016.
018 "Factorisable Sparse Tail Event Curves with Expectiles" by Wolfgang K. Härdle, Chen Huang and Shih-Kang Chao, March 2016.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".