Factorizable Sparse Tail Event Curves with Expectiles

Shih-Kang Chao
Wolfgang K. Härdle
Chen Huang

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Department of Statistics, Purdue University
http://lvb.wiwi.hu-berlin.de
http://www.case.hu-berlin.de
http://irtg1792.hu-berlin.de
http://www.stat.purdue.edu
Holding on the two ends...
FASTEC: FActorisable Sparse Tail Event Curves

- **Common structure**
  - High-dimensional time series with factors
  - Sparse penalization

- **Individual variety**
  - Tail behaviour
  - Spread analysis on factor loadings

Chao et al. (2015)
fMRI Application

- High-dimensional, high frequency & large data set
  - 19 volunteers, 256 investment decisions tasks
  - Around $100^3$ voxels’ data points, Blood Oxygenation Level Dependent (BOLD) effect every 2 sec

- Investment decisions and brain reactions
- Economics, Psychology and Statistics
- Spectral clustering identifies active zones

Majer et al. (2015)
Chinese Temperature

- Daily data from 1957 to 2013
- 159 Chinese weather stations
- Temperature distribution and extreme weather forecasting
- Weather derivatives in financial industry
DWD Climate Data

- Daily wind speed data for German stations from 1964 to 2014

207 stations in 2014

FASTEC with Expectiles
Aging and Growing over the World

- About 40 countries or areas, 1921-2011
- Mortality trend over ages
- Extremes and expectiles, tail events

Figure 1: Log death rates over ages
Challenges

- Dimension reduction
- Multivariate tail event regression
- Fast Iterative Shrinkage Thresholding Algorithm?
- Oracle inequalities for the estimator
- How do Tail Event Curves vary in time?
Outline

1. Motivation ✓
2. FASTEC with Expectiles
3. fMRI data & risk perception
4. Empirical Results
5. Conclusions
Tail Event Curve

- **Quantile**
  - Ratio of areas
  - Local influence

\[
\frac{\tau}{1 - \tau} = \frac{\int_{-\infty}^{\tau} dF(y)}{\int_{\tau}^{\infty} dF(y)}
\]

- **Expectile**
  - Ratio of weighted averaged distances
  - Capture the tail moments, not robust

\[
\frac{\tau}{1 - \tau} = \frac{\int_{-\infty}^{e_{\tau}} |y - e_{\tau}| dF(y)}{\int_{e_{\tau}}^{\infty} |y - e_{\tau}| dF(y)}
\]

**Example**

When \( \tau = 0.5 \), quantile = median, expectile = mean.
Quantile and Expectile

- Loss function
  \[ \rho_{\tau,\alpha}(u) = |\tau - I\{u < 0}| |u|^\alpha, \text{ with } \alpha = 1, 2, \tau \in (0, 1) \]

  - Quantile
    \[ q_\tau = \arg\min_\theta E \rho_{\tau,1}(Y - \theta) \]

  - Expectile
    \[ e_\tau = \arg\min_\theta E \rho_{\tau,2}(Y - \theta) \]

Note: the MLE of the location parameter of an ALD/AND correspond to the quantile(expectile) regression estimator
Quantile and Expectile

- Loss function

Figure 2: Expectile and quantile loss functions at $\tau = 0.5$ (dashed), $\tau = 0.9$ (solid).
Model Specification

- \( \{ Y_i \}_{i=1}^{n} \in \mathbb{R}^m \): multivariate curves to be jointly modelled
- \( \{ X_i \}_{i=1}^{n} \in \mathbb{R}^p \): \( p \) increases with \( n \), B-spline basis or other regression variables

Example
Chinese temperature: for each year, \( m = 159 \) (stations), \( n = 365 \) (days), \( p = n^{0.6} \approx 34 \).
fMRI: \( m = 19 \) (individuals) \( \times 256 \) (questions) = 4864, \( n = 50 \) (data points), \( p = n^{0.8} \approx 23 \).
Demographic data: for each year, \( m = 38 \) (countries), \( n = 111 \) (ages), \( p = n^{0.8} \approx 43 \).
Model Specification - ctd

- Conditional expectile function $e(\tau|X_i)$ is approximated by linear factor model:

$$e(\tau|X_i) = \sum_{k=1}^{r} \psi_k(\tau) f_k^T(X_i),$$

where $f_k^T(X_i)$ is the $k$th factor, $r$ is the number of factors, $\psi_k(\tau)$ are the factor loadings.

- Dimension reduced from $p$ to $r$
Model Specification - ctd

- Factors are constructed by linear combination of $X_i$:

$$f_k^T(X_i) = \varphi_k^T(\tau)X_i$$  \hspace{1cm} (2)

- Substituting (2) into (1):

$$e(\tau|X_i) = \gamma^\top(\tau)X_i$$  \hspace{1cm} (3)

with $\gamma(\tau) = (\sum_{k=1}^{r} \psi_k(\tau)\varphi_{k,1}(\tau), \ldots, \sum_{k=1}^{r} \psi_k(\tau)\varphi_{k,p(\tau)})^\top$, which is one column of the coefficient matrix $\Gamma$. 
Estimation

Coefficient matrix $\Gamma$:

$$\hat{\Gamma}_\lambda(\tau) = \arg \min_{\Gamma \in \mathbb{R}^{p \times m}} \left\{ (mn)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m} \rho_\tau \left( Y_{ij} - X_i^\top \Gamma_j \right) + \lambda \| \Gamma \|_* \right\}$$

(4)

- $\Gamma_j$ is the $j$th column of $\Gamma \in \mathbb{R}^{p \times m}$
- nuclear norm $\| \Gamma \|_* = \sum_{j=1}^{\min(p,m)} \sigma_j(\Gamma)$, given the eigenvalues of $\Gamma$: $\sigma_1(\Gamma) \geq \sigma_2(\Gamma) \geq \ldots \geq \sigma_{\min(p,m)}(\Gamma)$
- # of factors is # of nonzero eigenvalues of $\Gamma$
- Solved by fast iterative shrinkage thresholding algorithm
- Identify factors and loadings
Fast Iterative Shrinkage Thresholding Algorithm

Beck and Teboulle (2009)

- Objective: \( \min_{\Gamma} \{ F(\Gamma) \overset{\text{def}}{=} g(\Gamma) + h(\Gamma) \} \)

- \( g \): smooth convex function with Lipschitz continuous gradient

\[
\| \nabla g(\Gamma_1) - \nabla g(\Gamma_2) \|_F \leq L_{\nabla g} \| \Gamma_1 - \Gamma_2 \|_F, \forall \Gamma_1, \Gamma_2
\]

where \( L_{\nabla g} \) is the Lipschitz constant of \( \nabla g \)

- \( h \): continuous convex function, possibly nonsmooth

\[
| F(\Gamma_t) - F(\Gamma^*) | \leq \frac{2L_{\nabla g} \| \Gamma_0 - \Gamma^* \|_F^2}{(t+1)^2}
\]
Loss Error Bound and Convergence Analysis

Theorem 1

- Lipschitz continuity of expectile loss gradient:
  \[ L_{\nabla g} = 2(mn)^{-1} \max(\tau, 1 - \tau) \| X \|_F^2 \]

- In the \( t \)-th step of the iteration

\[
| F(\Gamma_t) - F(\Gamma^*) | \leq \frac{4(mn)^{-1} \max(\tau, 1 - \tau) \| X \|_F^2 \| \Gamma_0 - \Gamma^* \|_F^2}{(t + 1)^2} 
\]

(5)

- To achieve \( | F(\Gamma_t) - F(\Gamma^*) | \leq \varepsilon, \ \forall \varepsilon > 0 \), we need

\[
t \geq \frac{2 \sqrt{\max(\tau, 1 - \tau) \| X \|_F \| \Gamma_0 - \Gamma^* \|_F}}{\sqrt{mn\varepsilon}} - 1
\]

(6)

- Convergence rate \( \mathcal{O}(1/\sqrt{\varepsilon}) \)
Oracle Inequalities

- Upper bounds for $\|\hat{\Gamma}_\lambda - \Gamma^*\|^2_F$ in finite sample
- $\Gamma^*$ can be exactly sparse or not
- High-dimensional framework: rank $(\Gamma^*)$ and $p + m$ are both allowed to tend to infinity (but no quicker than $n$)
- $\{(X_i, Y_i)\}_{i=1}^n \in \mathbb{R}^{p+m}$ are identically distributed observations
Error Bounds for the Estimator

- Unified framework for high-dimensional $M$–Estimators with decomposable regularizers by Negahban et al. (2012)
- Conditions need to be verified:
  - Restricted strong convexity holds for expectile loss function
  - Nuclear norm is decomposable with respect to appropriately chosen subspaces
Error Bounds for the Estimator

Theorem 2
Suppose \( \{X_i\}_{i=1}^n \in \mathbb{R}^p \) are i.i.d. samples from \( N(0, \Sigma) \), for \( n \geq 2 \min(m, p) \), any optimal solution \( \hat{\Gamma}_\lambda \) with a strictly positive tuning parameter \( \lambda \geq 2 \| \nabla g(\Gamma^*) \| \) satisfies the bound

\[
\| \hat{\Gamma}_\lambda - \Gamma^* \|_F^2 \leq \frac{9^3 m^2 \lambda^2}{\{ \min(\tau, 1 - \tau) \sigma_{\min}(\Sigma) \}^2} \psi^2 (\mathcal{M}) + \frac{36 m \lambda}{\min(\tau, 1 - \tau) \sigma_{\min}(\Sigma)} \| \Gamma^*_{\mathcal{M}^\perp} \|_*,
\]

with probability greater than \( 1 - 4 \exp(-n/2) \).
Error Bounds for the Estimator

- $\Psi(M) \overset{\text{def}}{=} \sup_{Z \in M \setminus \{0\}} \frac{\|Z\|_*}{\|Z\|_F}$ is the subspace compatibility constant, which is $\sqrt{\text{rank}(\Gamma^*)}$ when $\Gamma^*$ is exactly sparse.

- $\Gamma_{M^\perp}^* \overset{\text{def}}{=} \arg \min_{Z \in M^\perp} \|Z - \Gamma^*\|_F$

- Best choice of $\lambda$ (Tuning Parameter)
Error Bounds for the Estimator

Corollary

Under the assumptions on sample setting, selecting
\( \lambda = 2m^{-1} S \max (\tau, 1 - \tau) \sqrt{K_u^2 \| \Sigma \|} \sqrt{\frac{p + m}{n}} \), for \( n \geq 2 \min (m, p) \), any optimal solution \( \hat{\Gamma}_\lambda \) satisfies the bound

\[
\| \hat{\Gamma}_\lambda - \Gamma^* \|_F^2 \leq \frac{9^3 \cdot \{2S \max (\tau, 1 - \tau) K_u\}^2 \| \Sigma \| (p + m) \psi^2 (\mathcal{M})}{n \left\{ \min (\tau, 1 - \tau) \sigma_{\min} (\Sigma) \right\}^2} + \frac{72S \max (\tau, 1 - \tau) \sqrt{K_u^2 \| \Sigma \|} \sqrt{p + m}}{\sqrt{n} \min (\tau, 1 - \tau) \sigma_{\min} (\Sigma)} \| \Gamma^*_{\mathcal{M}^\perp} \|_*,
\]

with probability greater than \( 1 - 3 \times 8^{-(p + m)} - 4 \exp (-n/2) \).
Investment Decisions and Brain Reactions

- How does an individual perceive risk?
- Is risk attitude reflected in brain activity?
**Investment Decision Experiment**

- Survey by Department of Education and Psychology, FU Berlin
- **19** healthy volunteers

- Investment Decision (ID) task ($\times 256$)
  - safe vs. random ($\mu, \sigma$)

- fMRI images: 2 sec $\times 1400 \approx 48$ min
Investment Decision

Choose between:

A) **Safe**, fixed return 5%
B) **Random**, investment return (3 types)
   - Single Investment
   - Portfolio of 2 (perfectly) [correlated investments]
   - Portfolio of 2 [uncorrelated investments]

- Each type of portfolio $\times 64$, single $\times 128$
- Display and decision time: 7 sec
ID Experiment

Figure 3: Decide between A) 5% return and displayed B) portfolio/investment FASTEC with Expectiles
fMRI Dynamics

Hemodynamic response (1 voxel) ▶ HRF

Figure 4: Hemodynamic response of a stimulus signal
Risk Attitude Parameter

Figure 5: Estimated risk attitude for 19 subjects

Risk attitude parameter
FASTEC with Expectiles
Importance of Tails

Figure 6: Boxplot of maximum responses over questions for 19 subjects (ordered by risk attitude parameters)
FASTEC with Expectiles
Empirical Results

fMRI Data

- Three ID-related active clusters: aINS_Left and aINS_Right, DMPFC. Majer et al. (2015)

- At each $t$, take different quantile levels (0.1, 0.5, 0.9) among all voxels in each cluster

FASTEC with Expectiles
Data Smoothing

- 19 individuals, 256 questions, $19 \times 256 = 4864$ curves
- Use 4 scans (6 seconds) after each stimulus
- Linear interpolation, take 50 points from the fitted curve

Figure 7: Two examples in aINS_Left cluster at 50% quantile level, with Hemodynamic response. FASTEC with Expectiles
Factor Analysis

- $n = 50$ observations, $m = 4864$ curves
- $X_i$: $B$-spline basis (cubic splines) with $p = n^{0.8} \approx 23$, $t = i/n$, $i = 1, \ldots, n$

<table>
<thead>
<tr>
<th></th>
<th>aINS_L</th>
<th>aINS_R</th>
<th>DMPFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st factor</td>
<td>0.624</td>
<td>0.631</td>
<td>0.612</td>
</tr>
<tr>
<td>2nd factor</td>
<td>0.791</td>
<td>0.793</td>
<td>0.779</td>
</tr>
<tr>
<td>3rd factor</td>
<td>0.907</td>
<td>0.913</td>
<td>0.898</td>
</tr>
<tr>
<td>4th factor</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 1: Proportion of variance explained by the first four factors under $\tau = 50\%$
Empirical Results

Factor Curves

Figure 8: The first 2 factors under $\tau = 99\%$ and $\tau = 1\%$ respectively (1% quantile level in aINS_Left cluster)

FASTEC with Expectiles
Empirical Results

Risk attitude - Stimulus Response

- Standard deviation of the factor loadings

\[ \beta_i = \alpha_0 + \alpha_1 \cdot \text{sd} (\psi_1)^T_{i, \text{ainsL}} + \alpha_2 \cdot \text{sd} (\psi_1)^T_{i, \text{ainsR}} + \alpha_3 \cdot \text{sd} (\psi_1)^T_{i, \text{DMPFC}} + \varepsilon_i \]  

(9)

Figure 9: \( R^2 \) in the regressions under different \( \tau \) levels

FASTEC with Expectiles
Risk attitude - Stimulus Response

\[
\beta_i = \alpha_0 + \alpha_1 \cdot \text{sd} (\psi_1)_i, ainsL + \alpha_2 \cdot \text{sd} (\psi_1)_i, ainsR + \alpha_3 \cdot \text{sd} (\psi_1)_i, DMPFC + \varepsilon_i
\]

(10)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_0)</td>
<td>0.961</td>
<td>0.201</td>
<td>4.777</td>
<td>0.245 (\cdot 10^{-3})</td>
</tr>
<tr>
<td>(\alpha_1)</td>
<td>(-34.414)</td>
<td>12.250</td>
<td>(-2.809)</td>
<td>0.013</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>(-37.571)</td>
<td>15.623</td>
<td>(-2.405)</td>
<td>0.029</td>
</tr>
<tr>
<td>(\alpha_3)</td>
<td>30.984</td>
<td>14.152</td>
<td>2.189</td>
<td>0.044</td>
</tr>
</tbody>
</table>

Table 2: Coefficients estimation results, \(R^2 = 0.422\), \(\text{adj.} R^2 = 0.307\).
Risk attitude - Stimulus Response

Figure 10: Fitted risk attitude by model given in (10) with $\tau = 0.1$. 

FASTEC with Expectiles
Empirical Results

Risk attitude - Stimulus Response

Figure 11: Fitted risk attitude (hollow points) by model given in (10) with $\tau = 0.1$. FASTEC with Expectiles
Empirical Results

**Factor Curves**

Figure 12: The first 2 factors under $\tau = 99\%$ and $\tau = 1\%$ respectively (99% quantile level in aINS_Left cluster)

FASTEC with Expectiles
Risk attitude - Stimulus Response

- Mean of the factor loadings

$$\beta_i = \alpha_0 + \alpha_1 \cdot (\bar{\psi}_1)_{i,ainsL} + \alpha_2 \cdot (\bar{\psi}_1)_{i,ainsR} + \alpha_3 \cdot (\bar{\psi}_1)_{i,DMPFC} + \varepsilon_i$$

(11)

Figure 13: $R^2$ in the regressions under different $\tau$ levels

FASTEC with Expectiles
Risk attitude - Stimulus Response

Figure 14: Fitted risk attitude by model given in (11) with $\tau = 0.7$. 

FASTEC with Expectiles
Empirical Results

Risk attitude - Stimulus Response

Figure 15: Fitted risk attitude (hollow points) by model given in (11) with $\tau = 0.7$. FASTEC with Expectiles.
Figure 16: The first 2 factors under $\tau = 99\%$ and $\tau = 1\%$ respectively (50% quantile level in aINS_Left cluster)
Empirical Results

Factor Loadings

Figure 17: The first factor loadings for #1 and #19 individuals with FASTEC and Expectiles
Factor Loadings

Figure 18: Iso-contour lines of $\psi_1$ for #1 and #19 individuals
FASTEC with Expectiles
Risk attitude - Stimulus Response

Dispersion of the factor loadings under two \( \tau \) levels

\[
\text{dis} (\psi_1)^{\tau_1, \tau_2}_i = \frac{1}{256} \sum_{q=1}^{256} \sqrt{\left\{ \psi_1 (\tau_1)_i,q - \bar{\psi}_1 (\tau_1)_i \right\}^2 + \left\{ \psi_1 (\tau_2)_i,q - \bar{\psi}_1 (\tau_2)_i \right\}^2}
\]

\[
\beta_i = \alpha_0 + \alpha_1 \cdot \log \text{dis} (\psi_1)^{\tau_1, \tau_2}_i,\text{ainsL} + \alpha_2 \cdot \log \text{dis} (\psi_1)^{\tau_1, \tau_2}_i,\text{ainsR} + \alpha_3 \cdot \log \text{dis} (\psi_1)^{\tau_1, \tau_2}_i,\text{DMPFC} + \varepsilon_i
\]

Figure 19: \( R^2 \) in the regressions under different \( \tau \) levels (by pairs)
Empirical Results

Risk attitude - Stimulus Response

Figure 20: Fitted risk attitude by model given in (12) with $\tau_1 = 0.1$, $\tau_2 = 0.9$. FASTEC with Expectiles
Empirical Results

Risk attitude - Stimulus Response

Figure 21: Fitted risk attitude (hollow points) by model given in (12) with $\tau_1 = 0.1, \tau_2 = 0.9$. FASTEC with Expectiles
Empirical Results

Factor Loadings

Figure 22: The first factor loadings for all curves $j = 1, \ldots, 4864$, where "3-171" denotes #3 individual’s #171 question and so on.

FASTEC with Expectiles
Temperature Data

Figure 23: Top figure: detrended temperature series; bottom figure: trend

FASTEC with Expectiles
Temperature Data - Factors

Figure 24: The first factor under 1% and 99% tail levels.

FASTEC with Expectiles
Temperature Data - Factor Loadings

Figure 25: The first factor loadings for each station.
Empirical Results

Temperature Data - Chinese Map

Figure 26: Chinese map marked with three selected weather stations.

FASTEC with Expectiles
Wind Data - Factors

Figure 27: The first factor under 1% and 99% tail levels.
Wind Data - Factor Loadings

Figure 28: The first factor loadings for each station.

FASTEC with Expectiles
Wind Data - German Map

1964

Figure 29: German map marked with two most extreme weather stations.
FASTEC with Expectiles
Mortality Data - Curves

In each year, a bundle of $m$ curves over ages $0, 1, \ldots, 110$.

Estimate conditional expectile curves applying functional data analysis.

Figure 30: Log death rate curves from 1921 to 2011.
Mortality Data - Factors

The common trend concerning quinquagenarian group

Use factor loadings to detect the outliers

- Good ones: Japan, Switzerland
- Bad ones: Latvia, Russia

Figure 31: The first factor under 70% tail level.
Mortality Data - Factor Loadings

Figure 32: The first factor loadings for four representative countries in all years under 70% tail level: Switzerland, Japan, Latvia, Russia
FASTEC with Expectiles
Conclusions

- Principal factors capture the common patterns among curves
- TEC study discovers the extreme behaviors
- Consistency and convergence rate of the estimator are demonstrated by theorems
- Risk attitude implied by individual’s choices and his brain reactions can be linked by statistical model
Factorizable Sparse Tail Event Curves with Expectiles

Shih-Kang Chao
Wolfgang K. Härdle
Chen Huang

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Department of Statistics, Purdue University
http://lvb.wiwi.hu-berlin.de
http://www.case.hu-berlin.de
http://irtg1792.hu-berlin.de
http://www.stat.purdue.edu
References

A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems
SIAM Journal on Imaging Sciences, 2(1), 183-202

M-Quantiles
Biometrika, 75(4), 761-771
A Singular Value Thresholding Algorithm for Matrix Completion

Factorisable Sparse Tail Event Curves
References

*Portfolio Decisions and Brain Reactions via the CEAD method* 

*A Unified Framework for High-Dimensional Analysis of M-Estimators with Decomposable Regularizers* 
Statistical Science, 27(4), 538-557
Value at Risk and Expected Shortfall

Figure 33: The relationship between VaR and ES in one plot.
History of Expectiles

- Gravile - A. Goldberger
  - motivated by the interpretation of expectation as a center of gravity
- Projectile - G. Chamberlain
  - motivated by the fact that it solves a least squares problem
- Other alternative terminologies: Heftile, Loadile

Tail Event Curve

FASTEC with Expectiles
Iterative Algorithm

- Initialize: \( \Gamma_0 = 0, \Sigma_1 = 0 \), step size \( \delta_1 = 1 \)
- For \( t = 1, 2, \ldots, T \)
  - \( \Gamma_t = \arg \min_{\Gamma} \left\{ \frac{g(\Gamma)}{L_{\nabla g}} + \frac{1}{2} \left\| \Gamma - \left\{ \Omega_t - \frac{1}{L_{\nabla g}} \nabla g(\Omega_t) \right\} \right\|^2 \right\} \)
  - when penalizing nuclear norm, \( \Gamma_t = P \left( R - \frac{\lambda}{L_{\nabla g}} I_{p \times m} \right) Q_{+} \top \), see Cai et al. (2010), where \( \Omega_t - \frac{1}{L_{\nabla g}} \nabla g(\Omega_t) = PRQ_{+} \top \), by SVD
  - \( \delta_{t+1} = \frac{1+\sqrt{1+4\delta_t^2}}{2} \)
  - \( \Omega_{t+1} = \Gamma_t + \frac{\delta_{t-1}}{\delta_{t+1}} (\Gamma_t - \Gamma_{t-1}) \)
- \( \hat{\Gamma} = \Gamma_T \)
Factorize $\hat{\Gamma}_\lambda(\tau)$

- the number of nonzero singular values is the number of sparse factors: $r$
- dimension reduced from $p$ to $r$
- singular value decomposition: $\hat{\Gamma}_\lambda(\tau) = \text{SVD}^\top$
- $f_k^\tau(X_i) = \varphi_k^\top(\tau) = \sigma_k S_k^\top X_i$, $\psi = D_j$, which is orthonormal
RSC of Expectile Loss

- RSC holds for \( g(\Gamma) = (mn)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m} \rho_{\tau}(Y_{ij} - X_{i}^\top \Gamma_{j}) \) with curvature \( \kappa > 0 \) and tolerance function \( \xi(\cdot) \) if

\[
g(\Gamma^* + \Delta) - g(\Gamma^*) - \langle \nabla g(\Gamma^*), \Delta \rangle \geq \kappa \| \Delta \|_{F}^{2} - \xi^{2}(\Gamma^*), \forall \Delta \in \mathbb{C}
\]

(13)

- \( \rho_{\tau}(u) = |\tau - 1 \{u < 0\}| u^2 \) satisfies

\[
\rho_{\tau}(u + \delta) - \rho_{\tau}(u) - \rho'_{\tau}(u) \delta \geq \min(\tau, 1 - \tau) \delta^2
\]

(14)
Figure 34: An example when $\tau = 0.9$, $\delta \in [-0.5, 0.5]$, where the two dash lines are the LHS of (14) w.r.t. $\delta$ (for $u = \pm 0.1$ respectively), and the red line is the lower bound FASTEC with Expectiles
RSC of Expectile Loss - ctd

- \( g(\Gamma) \) is RSC with \( \kappa = (m)^{-1} \min(\tau, 1 - \tau) \sigma_{\min} \left( \frac{X^\top X}{n} \right) \) and \( \xi(\cdot) = 0 \)
- If \( \{X_i\}_{i=1}^n \in \mathbb{R}^p \) are i.i.d. samples from \( \mathcal{N}(0, \Sigma) \), for \( n \geq 2 \min(m, p) \),
  \( \kappa = \frac{1}{9} (m)^{-1} \min(\tau, 1 - \tau) \sigma_{\min}(\Sigma) \) with probability greater than \( 1 - 4 \exp(-n/2) \)

Return

FASTEC with Expectiles
Decomposable Regularizers

For $\mathcal{M} \subseteq \overline{\mathcal{M}}$ of $\mathbb{R}^{p \times m}$, a norm-based regularizer $R$ is decomposable with respect to $\left(\mathcal{M}, \overline{\mathcal{M}}\right)$, if

$$R(\Gamma + \Delta) = R(\Gamma) + R(\Delta), \forall \Gamma \in \mathcal{M}, \Delta \in \overline{\mathcal{M}}$$  \hspace{1cm} (15)

Nuclear norm is decomposable with respect to

$$\mathcal{M}(U, V) = \{\Gamma \in \mathbb{R}^{p \times m} | \text{row} (\Gamma) \subseteq U, \text{col} (\Gamma) \subseteq V\}$$

$$\overline{\mathcal{M}}(U, V) = \{\Gamma \in \mathbb{R}^{p \times m} | \text{row} (\Gamma) \subseteq U^\perp, \text{col} (\Gamma) \subseteq V^\perp\}$$
More Assumptions

- \{ (X_i, Y_i) \}_{i=1}^n \in \mathbb{R}^{p+m} are i.i.d., \{ X_i \}_{i=1}^n \in \mathbb{R}^p \sim N(0, \Sigma)
- Conditional on \( X_i \), \( u_{ij} = \{ Y_{ij} - X_i^\top \Gamma \cdot j \}_{j=1}^m \) are cross-sectional independent over \( j \)
- \( u_{i1}, \ldots, u_{im} \) are sub-gaussian: \( \exists C > 0 \) such that \( P(|u_{ij}| > s) \leq \exp \left( 1 - \frac{s^2}{C^2} \right), \quad j \in \{1, \ldots, m\} \)
- \( K_u \triangleq \max_{1 \leq j \leq m} \| u_{ij} \|_{\psi_2} = \max_{1 \leq j \leq m} \sup_{p \geq 1} p^{-1/2} (E |u_{ij}|^p)^{1/p} \)
Best choice of $\lambda$

Under the assumptions on sample setting,

$$\Pr \left( \|\nabla g (\Gamma)\| \leq m^{-1} S \max (\tau, 1 - \tau) \sqrt{K_u^2 \|\Sigma\|} \sqrt{\frac{p + m}{n}} \right) \geq 1 - 3 \times 8^{-(p+m)} - 4 \exp \left(-\frac{n}{2}\right),$$

where $S$ is an absolute constant.

Corollary
Experiment

- **Incentive to be rational**
  - Draw 1 ID task and multiply subject’s choice by 100 EUR
  - \( 9\% \times 100 = 9 \text{ EUR} \)

- **Gaussian returns:**
  - \( \mu = 5\%, 7\%, 9\%, 11\% \)
  - \( \sigma = 2\%, 4\%, 6\%, 8\% \)

ID Experiment

FASTEC with Expectiles
Single Investment fMRI Experiment

Figure 35: An example of return stream from single investment displayed to the subject during the experiment for 7 sec.; returns $r_i \sim N(\mu, \sigma^2)$, here $\mu = 5\%$, $\sigma = 2\%$

FASTEC with Expectiles
Correlated Portfolio

Figure 36: An example of return streams from correlated portfolio displayed to the subject during the experiment for 7 sec.; returns $r_i \sim N(\mu, \sigma^2)$, here $\mu_1 = 5\%, \mu_2 = 9\%$ and $\sigma = 2\%$

FASTEC with Expectiles
Uncorrelated Portfolio

Figure 37: An example of return streams from uncorrelated portfolio displayed to the subject during the experiment for 7 sec.; returns $r_i \sim N(\mu, \sigma^2)$, here $\mu = 7\%$, $\sigma = 2\%$

FASTEC with Expectiles
Hemodynamic response function e.g. Double Gamma function

\[ h(t) = \left( \frac{t}{5.4} \right)^6 \exp\left( - \frac{t - 5.4}{0.9} \right) - 0.35 \left( \frac{t}{10.8} \right)^{12} \exp\left( - \frac{t - 10.8}{0.9} \right), \quad t \geq 0 \text{-time [sec]} \]

Figure 38: Predicted response as a convolution of a stimulus signal and a HRF. Figure modified from FEAT - FMRI.

FASTEC with Expectiles
Risk Attitude Parameter

- Risk-return choice model

\[ V_r^i = \bar{x}_r - \beta_i S_r, \quad 1 \leq i \leq 19, 1 \leq r \leq 256 \]

- \( x_r \) portfolio return stream, \( \bar{x}_r \) average return (\( \mu \))
- \( S_r \) standard deviation of \( x_r \) (risk)
- \( V_r^i \) subjective value (unobserved), 5% risk free return
- \( \beta_i \) risk attitude parameter
Risk Attitude Parameter

- Estimation of individual risk attitude by logistic regression

\[
\begin{align*}
P\{\text{risky choice} \mid x\} &= \frac{1}{1 + \exp\{\bar{x} - \beta S(x) - 5\}} \\
P\{\text{sure choice} \mid x\} &= 1 - \frac{1}{1 + \exp\{\bar{x} - \beta S(x) - 5\}}
\end{align*}
\]

- risky choice - unknown return, sure choice - fixed, 5% return
- $\beta$ estimated by maximum likelihood

FASTEC with Expectiles
Cluster Activation: aINS

Figure 39: Anterior insula (aINS) activated during all type of investment decisions in the group-level analysis.

FASTEC with Expectiles
Cluster Activation: DMPFC

Figure 40: Dorsolateral prefrontal cortex (DMPFC) activated during all type of investment decisions in the group-level analysis. fMRI Data FASTEC with Expectiles
B-Spline Basis for Cubic Splines

Figure 41: B-Spline basis for Cubic splines with 23 basis functions