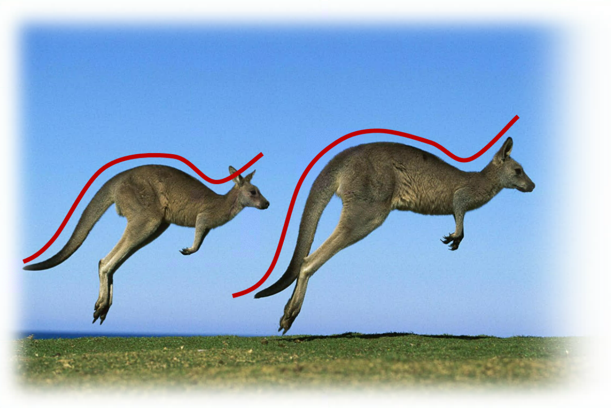
DYTEC - Dynamic Tail Event Curves

Petra Burdejová
Wolfgang Karl Härdle

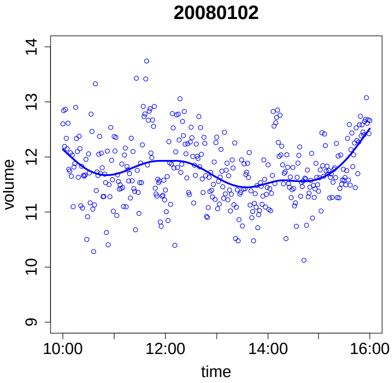
Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://case.hu-berlin.de

Importance of tail event (curves)



Intra-day trading volume

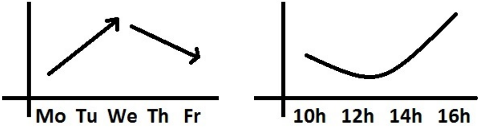
- NASDAQ Market
  - CISCO stocks
- daily 01-12/2008
- 250 trading days
  10:00 - 16:00
- cumulated 1-min volume



Intra-day trading volume

Jain and Joh (1988)

- the day of the week and the hour have effect on trading volume



Darrat et al. (2003) and **Spierdijk et al. (2003)**

- lagged values of volatility and trading volume simultaneously

Bialkowski et al. (2008) and **Brownlees et al. (2011)**

- dynamic volume approach for VWAP

VWAP trading strategy

- Buying/Selling fixed amount of shares at average price $p_j \, \mathbf{I}_{\{p_j > VWAP\}}$ that tracks the VWAP benchmark

$$VWAP = \frac{\sum_{j=1}^{J} v_j \cdot p_j}{\sum_{j=1}^{J} v_j}$$

with price p_j and volume v_j of the j-th transaction

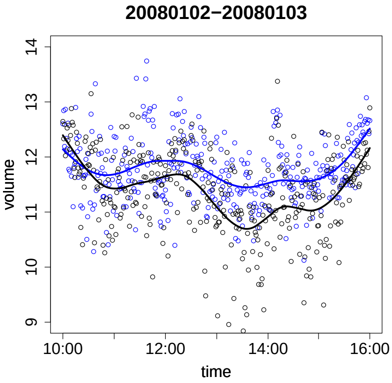
- 50 % of trades are VWAP orders

- Implementation requires **model for intraday evolution of volume**

Intra-day trading volume

High frequency data

- Model curves
- High dimensions
- Focus on dynamics
- Comprising tail events
- Make forecasts



Intra-day trading volume

Figure: Expectiles for $\tau \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.5, , \ldots\}$

Intra-day trading volume

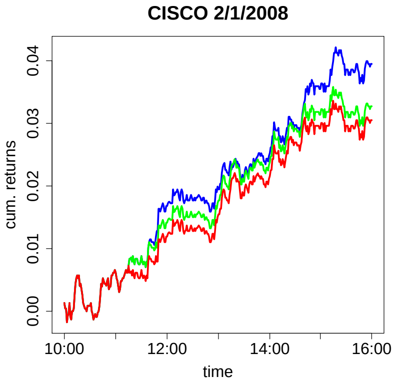


Figure: Cumulative returns of VWAP strategy with weights based on τ-expectiles of volume, $\tau = 0.05, 0.5, 0.95$.

Temperature data

Are we getting stronger extremes?

- Daily average temperature
- Berlin, 1948 - 2013
- Model residuals
  Model
- Usage: Pricing weather derivatives

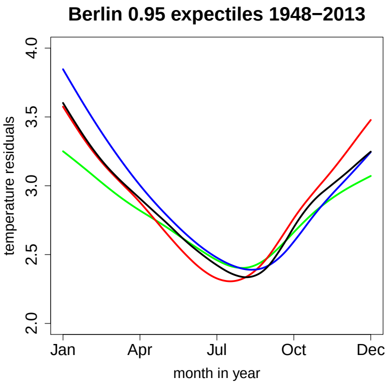
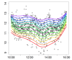


Figure: 0.95-expectile of Berlin temperature residuals in 1948-1969 1970-1991 1992-2013 1948-2013



Hurricane predictions


NATIONAL GEOGRAPHIC Daily News

Home Animals Ancient Energy Environment Travel/Cultures Space/Tech Water Weird News Photos News Video News Blogs

Will U.S. Hurricane Forecasting Models Catch Up to Europe's?

A year after Hurricane Sandy, Europe's forecasting technology is still tops.

News | Videos | Blogs | Personalities

Funding to Allow Better US Hurricane Prediction Following Sandy

By **Vickie Frantz**, AccuWeather.com Staff Writer
May 31, 2013; 4:36 AM

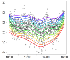
Share | g+1

Congress approved supplemental funding of $23.7 million to the National Weather Service (NWS) for improvements to their forecasting computers in Reston, Va., and Orlando, Fla.

The funding is provided under the Disaster Relief Appropriations Act of 2013. Chris Vaccaro Spokesman for the NWS said while there were already plans to upgrade the computers, the additional funding will allow more improvements than were planned.

This satellite image of Hurricane Sandy is courtesy of NOAA/NASA.

100%



Hurricane predictions

Are we getting stronger extremes?

- Strength of wind in knots
- West Atlantic
- Years 1946 - 2011
- observe different trend pattern

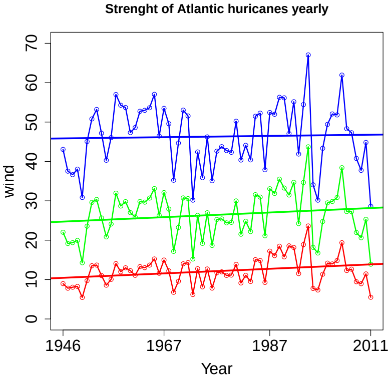
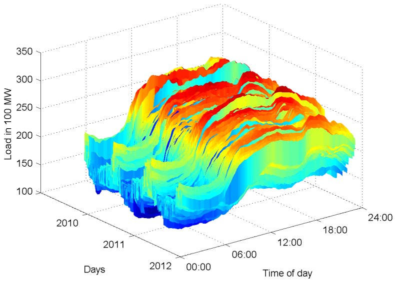
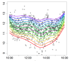


Figure: Yearly expectiles for $\tau = 0.25, 0.5, 0.75$ and trend

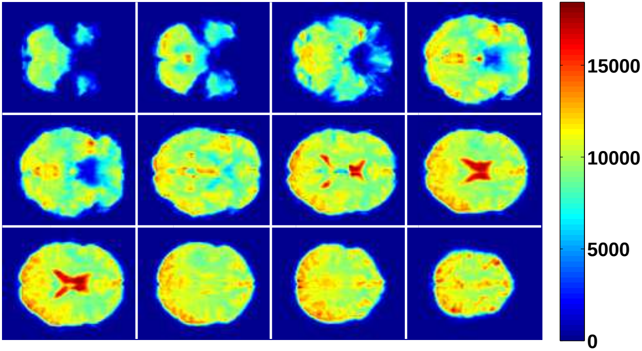
Dynamic demand models

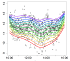
- **Electricity demand**
  - Quarter-hourly
  - Jan.2010 - Dec.2012
  - Amprion company in west of Germany
- **Water demand**
- **Gas demand**





Risk attitude via brain activity

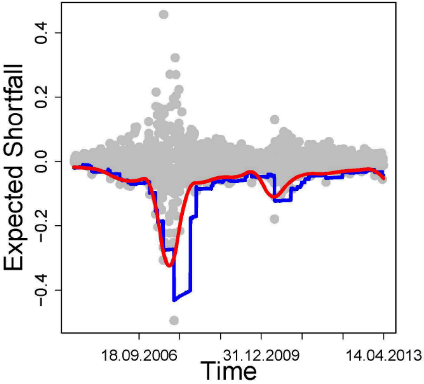


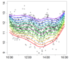


Expected shortfall

- FTSE 100
- 01/09/1997 to 02/05/2005
- ES related to expectiles

▸ Details





Objectives & Challenges

- High frequency data
- Time-varying with intraday pattern
- Dependent

- Model curves
- Focus on dynamics and dependence
- Comprising tail events
- Make forecasts

Outline

1. Motivation ✓
2. Quantiles and Expectiles
3. Modeling Time-varying Curves
4. Outlook
5. Empirical Study
6. References

Quantiles and Expectiles

For r.v. Y obtain τ-quantile

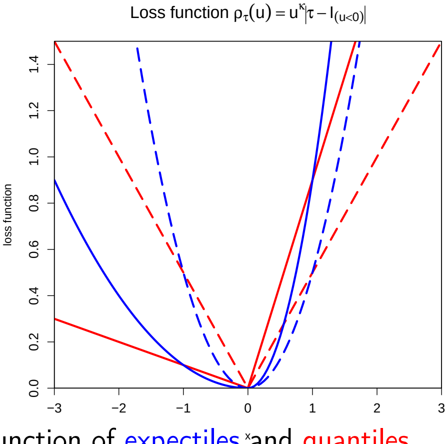
$$q_\tau = \arg\min_\theta \mathsf{E}\left\{\rho_\tau\left(Y - \theta\right)\right\}$$

with asymmetric loss function

$$\rho_\tau\left(u\right) = \left|u\right|^\alpha \left|\tau - \mathbf{I}_{\{u<0\}}\right|$$

where $\alpha = 1$ for quantiles and $\alpha = 2$ for expectiles

Quantiles and Expectiles



LQRcheck

Figure: Loss function of expectiles and quantiles
for $\tau = 0.5$ (dashed) and $\tau = 0.9$ (solid)

Expectile Curves

Y associated with covariates X

Define generalized regression τ-expectile

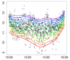
$$e^{\tau}(x) = \arg\min_{\theta} \mathsf{E}\left\{\rho_{\tau}\left(Y-\theta\right) \mid X = x\right\}$$

Use method of penalized splines to estimate vector of coefficients α

$$\widehat{\alpha} = \arg\min_{\alpha} \sum_{i=1}^{n} \rho_{\tau}\left\{y_i - \alpha^{\top} b(x_i)\right\} + \lambda \alpha^{\top}\Omega\alpha$$

with b basis vector (e.g. B-splines) Definition of B-splines

penalization matrix Ω and shrinkage parameter λ Penalization matrix



Estimation of Expectile Curves

Schnabel and Eilers (2009): iterative LAWS algorithm ▸ LAWS

Schnabel (2011): expectile sheets for joint estimation of curves

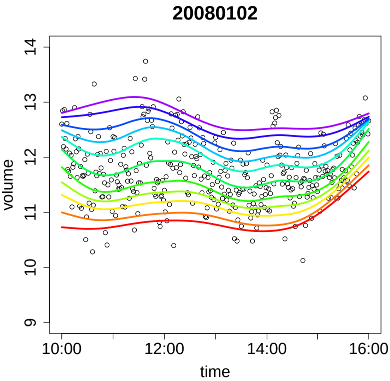
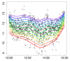


Figure: CISCO trading volume expectiles ($\lambda = 5$)



Dynamic tail event curves

Fix τ :

$$e^{\tau}(t) = \sum_{k=1}^{K} \alpha_k \phi_k(t)$$

with basis $\Phi = (\phi_1, \ldots, \phi_K)^\top$ and $t = 1, \ldots, T$.

Variation in time, $s = 1, \ldots, S$:

$$e_s^{\tau}(t) = \sum_{k=1}^{K} \alpha_{sk} \phi_k(t)$$

Independent curves

Guo et al. (2013)

- $e_s(t)$ **independent** realizations of stationary process
- Following Karhunen-Loève expansion: Functional princ. components

$$e_s(t) = \mu(t) + \sum_{k=1}^{K} \alpha_{sk}\phi_k(t) = \mu(t) + \alpha_s^\top \Phi(t)$$

- Method of penalized splines for mean function and FPC with empirical loss function

$$\sum_{s=1}^{S}\sum_{t=1}^{T} \rho_\tau \left\{ Y_{st} - \theta_\mu^\top b(t) - \alpha_s^\top \Theta_\Phi b(t) \right\} + \text{pen.mtx}$$

Details

Temporal (weak) dependent curves

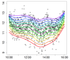
Sequence $\{X_n\}$ is *m*-**dependent** if for any k the σ-algebras $\mathcal{F}_k^- = \sigma\,(\ldots, X_{k-1}, X_k)$ and $\mathcal{F}_{k+m}^+ = \sigma\,(X_{k+m}, X_{k+m+1}, \ldots)$ are independent.

Hörmann and Kokoszka (2010)

- Karhunen-Loève expansion applicable for *m*-dependent
- asymptotic properties of FPC estimates remain the same

- most of time series ARE NOT *m*-dependent
- fail if i.i.d. curves are too noisy
- fail if curves are sufficiently regular but dependency is too strong

How to model stronger dependency?



Cramér-Karhunen-Loève representation

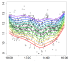
Panaretos and Tavakoli (2013)

- spectral decomposition of stationary functional time series

$$e_s(t) \approx \sum_{j=1}^{J} \exp(\mathrm{i}\omega_j s) \sum_{k=1}^{K} \alpha_{jk}\phi_{j,k}(t)$$

$-\pi = \omega_1 < \ldots < \omega_{J+1} = \pi$
$\{\phi_{j,k}\}_{k\geq 1}$ eigenfunctions
$\{\alpha_{j,k}\}_{k\geq 1}$ corresponding coefficients

Theoretical details



Cramér-Karhunen-Loève representation

Simplification: Assume $\phi_{j,k}(t) = \phi_k(t)$ for each j

Then:

$$
\begin{aligned}
e_s(t) &\approx \sum_{j=1}^{J} \exp(\mathrm{i}\omega_j s) \sum_{k=1}^{K} \alpha_{jk}\phi_k(t) \\
&\approx U(s)^\top A\Phi(t)
\end{aligned}
$$

with

$U(s) = \left(\exp\left(\mathrm{i}\omega_1 s\right), \ldots, \exp\left(\mathrm{i}\omega_J s\right)\right)^\top$
A_{JxK} matrix of coefficients
$\Phi(t) = (\phi_1(t), \ldots, \phi_K(t))^\top$

Empirical loss function

Method of penalized splines for $\phi_k(t) = \sum_{l=1}^{L} \beta_{kl} b_l(t)$
Then

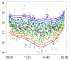
$$\Phi(t) = B_{KxL} b(t)$$

Minimize loss function

$$\sum_{s=1}^{S} \sum_{t=1}^{T} \rho_\tau \left\{ Y_{s,t} - U(s)^\top C b(t) \right\} + \lambda \| C_{JxL} \|_{\mathcal{G}_J}$$

with $C_{JxL} = A_{JxK} B_{KxL}$ matrix of coefficients
and Group lasso penalization

$$\| C_{JxL} \|_{\mathcal{G}_J} = \sum_{j=1}^{J} \| c_{\mathcal{G}_j} \|_2 = \sum_{j=1}^{J} \sqrt{\sum_{l=1}^{L} c_{jl}^2}$$



Empirical loss function

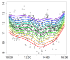
$$\underbrace{\sum_{s=1}^{S}\sum_{t=1}^{T}\rho_\tau\left\{Y_{s,t}-U(s)^\top Cb(t)\right\}}_{\text{l(C)}}+\lambda\sum_{j=1}^{J}\|c_{\mathcal{G}_j}\|_2$$

$l(C)$ continuously differentiable

K-K-T conditions for $\widehat{C}$ to be a solution:

$$\begin{aligned}\nabla l(\widehat{C})_{\mathcal{G}_j}+\lambda\frac{C_{\mathcal{G}_j}}{\|C_{\mathcal{G}_j}\|_2} &= 0 \quad \text{if } C_{\mathcal{G}_j}\neq 0\\ \|\nabla l(\widehat{C})_{\mathcal{G}_j}\|_2 &\leq \lambda \quad \text{if } C_{\mathcal{G}_j}=0\end{aligned}$$

If $\tau=0.5$, closed form solution available



Block-Coordinate Gradient Descent Algorithm (BCGD)

Tseng & Yun (2009)

Solve nonconvex nonsmooth optimization problem

$$\min_x f(x) + \lambda P(x)$$

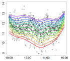
where $\lambda > 0$

$P : \mathbb{R}^n \to (-\infty, \infty]$ block-separable convex function

f smooth on an open subset containing $\operatorname{dom} P$

- combination of quadratic approximation and coordinate descent algorithm
- global convergence

▸ Theoretical details



B-C-G-D Algorithm

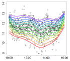
Stepwise and blockwise minimize

$$S_\lambda(\widehat{C}^{(t)}+d) = l(\widehat{C}^{(t)})+d^\top \nabla l(\widehat{C}^{(t)})+\frac{1}{2}d^\top H^{(t)} d+\lambda \sum_{j=1}^{J} \| C_{\mathcal{G}_j}^{(t)}+d\|_2$$

Details

Notation: $\min\limits_{d_{\mathcal{G}_j}} S_\lambda(\widehat{C}^{(t)} + d)$

minimization, where $d = (d_{\mathcal{G}_1}, \dots, d_{\mathcal{G}_J})$ with $d_{\mathcal{G}_k} = 0$ for $k \neq j$



B-C-G-D Algorithm

repeat

  t=t+1

  for $j = 1$ to J **do**

    if $\|\nabla l(\widehat{C}^{(t)})_{\mathcal{G}_j} - h^{(t)}_{\mathcal{G}_j}\widehat{C}^{(t)}_{\mathcal{G}_j}\|_2 \leq \lambda$ **then**

      $d^{(t)}_{\mathcal{G}_j} = -\widehat{C}^{(t)}_{\mathcal{G}_j}$

    else

      $d^{(t)}_{\mathcal{G}_j} = \min_{d_{\mathcal{G}_j}} S_\lambda(\widehat{C}^{(t)} + d)$

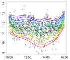
    end if

  end for

until convergence criterion met

update $\widehat{C}^{(t+1)} - \widehat{C}^{(t)} - \alpha^{(t)} d^{(t)} = 0$

Details



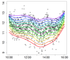
DYTEC Algorithm

Idea:
start with initial weights $w_{s,t}$ (obtained separately for each $s = 1, \ldots, S$) and iterate between following steps:

- compute $\widehat{C}$ using BCGD algorithm
- update weights

$$w_{s,t} = \begin{cases} \tau & \text{if } Y_{s,t} > U(s)^\top \widehat{C} b(t), \\ 1 - \tau & \text{otherwise.} \end{cases}$$

- stop if there is no change in weights $w_{s,t}$.



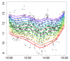
Dynamic functional factor model

- generalization (capture nonstationarity)
- Hays et al.(2012) & Kokoszka et al.(2014)
- extend with the idea of two spaces of basis function

Model

$$e_s(t) = \sum_{k=1}^{K} Z_{sk} m_k(t) = Z_s^\top m(t)$$

with time-varying factor loadings Z_k
and functional factors



Dynamic functional factor model

Time basis:

$$Z_{sk} = \sum_{j=1}^{J} \alpha_{kj} u_j(s)$$

$$Z_s = A_{KxJ} \cdot U(s)$$

Space basis:

$$m_k(t) = \sum_{l=1}^{L} \beta_{kl} b_l(t)$$

$$m(t) = B_{KxL} b(t)$$

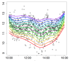
Dynamic functional factor model

$$e_s(t) = Z_s^\top m(t) = U(s)^\top Cb(t)$$

with $C_{JxL} = A_{JxK}^\top B_{KxL}$ matrix of coefficients,
space basis vector $b(t) = \{b_1(t), \dots, b_L(t)\}^\top$
and time basis vector $U(s) = \{u_1(s), \dots, u_J(s)\}^\top$

Same loss function:

$$\sum_{s=1}^{S} \sum_{t=1}^{T} \rho_\tau \left\{ Y_{s,t} - U(s)^\top Cb(t) \right\} + \lambda \| C_{JxL} \|_{\mathcal{G}_J}$$



Time basis

- capture periodic variation
- capture trend

- Proposal by Song et al. (2013):
  - Legendre polynomial basis
  - Fourier series

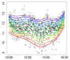
See Appendix

Space basis

- capture daily patterns
- capture specific structure

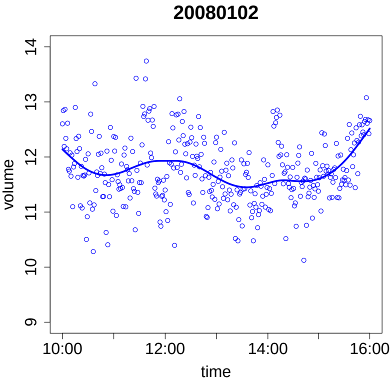
- Proposal by Song et al (2013):
  - Data driven
  - Based on combination of smoothing techniques and FPCA
- B-splines

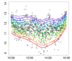
Definition of B-splines



Empirical Study

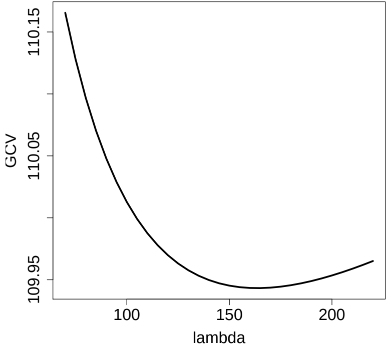
- NASDAQ Market
  - CISCO stocks
- daily 01-12/2008
- 250 trading days
  10:00 - 16:00
- cumulated 1-min volume

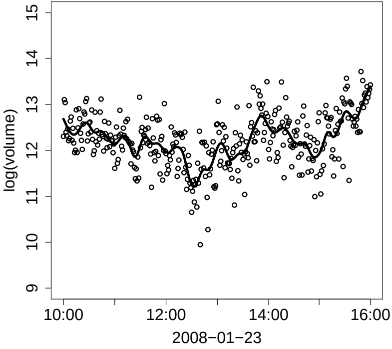




Trading volume - smoothing

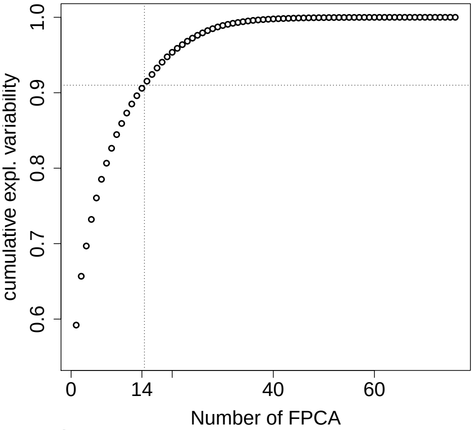
B-splines with knots every 5 minutes (i.e. 76 splines)
Optimal $\lambda = 164$



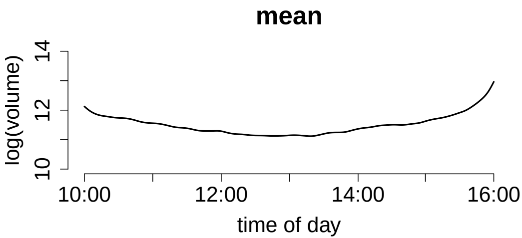


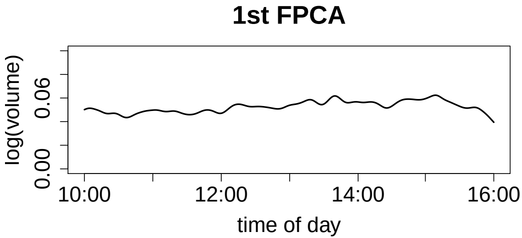
Trading volume - FPCA

Use 14 FPCs to explain 90% of variation



Trading volume - FPCA





Trading volume - FPCA

Fisher's G-test: p-value=0.000

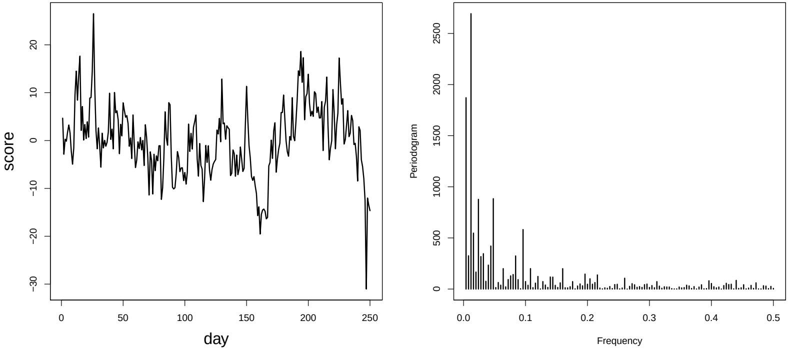


Figure: Scores of 1FPC and periodogram

Outlook

- Algorithm - Code in R
- Simulation and Empirical studies

DYTEC - Dynamic Tail Event Curves

Petra Burdejová
Wolfgang Karl Härdle

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://case.hu-berlin.de

References

Bialkowski, J., Darolles, S., Le Fol, G.
Improving VWAP strategies: A dynamic volume approach
Journal of Banking & Finance **32**(9): 1709-1722, 2008,
DOI: 10.1016/j.jbankfin.2007.09.023

Brownlees, Ch. T., Cipolliniy, F., Gallo, G.M.
Intra-daily Volume Modeling and Prediction for Algorithmic Trading
Journal of Financial Econometrics **9**(3): 489-518, 2011,
DOI: 10.2139/ssrn.1393993

Darrat, A. F., Rahman, S., Zhong, M.
Intraday trading volume and return volatility of the DJIA stocks: A note
Journal of Banking & Finance **27**: 2035-2043, 2003,
DOI: 10.2139/ssrn.348820

Engle, R. F., Granger, C.W.J., Rice, J., Weiss, A.,
Semiparametric Estimates of the Relation between Weather and Electricity Sales
Journal of the American Statistical Association **81**(394): 310-320, 1986, DOI: 10.2307/2289218

Guo, M., Härdle, W. K.
Simultaneous confidence bands for expectile functions
Advances in Statistical Analysis, **96**: 517-541 published online, 2012, DOI: 10.1007/s10182-011-0182-1

Guo, M., Zhou, L., Huang, J. Z., Härdle, W. K.
Functional data analysis of generalized regression quantiles
Statistics and Computing, published online, 2013, DOI: 10.1007/s11222-013-9425-1

Härdle, W. K., Hautsch, N. and Mihoci, A.
Local Adaptive Multiplicative Error Models for High-Frequency Forecasts
Journal of Applied Econometrics, accepted for publication, 2014, DOI: 10.1002/jae.2376

Hays, S., Shen ,H., Huang, J.Z,
Functional dynamic factor models with application to yield curve forecasting
The Anals of Applied Statistics **6**(3): 870–894, 2012, DOI: 10.1214/12-AOAS551

Hörmann, S., Kokoszka ,P.
Weakly dependent functional data
The Anals of Statistics **38**(3): 1845-1884, 2010,
DOI: 10.1214/09-AOS768

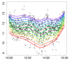
Kokoszka ,P., Miao, H., Zhang, X.
Functional Dynamic Factor Model for Intraday Price Curves
Journal of Financial Econometrics **0**(0): 1-22, 2014,
DOI: 10.1093/jjfinec/nbu004

Panaretos, V. M., Tavakoli, S.
Cramér-Karhunen-Loève representation and harmonic principal component analysis of functional time series
Stochastic Processes and their Applications **123**(7): 2779-2807, 2013, DOI: 10.1016/j.spa.2013.03.015

Jain, P. C., Joh, G. H.
The Dependence between Hourly Prices and Trading Volume
The Journal of Financial and Quantitative Analysis **23**(3): 269-283, 1988, DOI: 10.2307/2331067

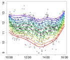
Schnabel, S.K., Eilers, P. H. C.
Optimal expectile smoothing
Computational Statistics and Data Analysis **53**(12): 4168-4177, 2009, DOI: 10.1016/j.csda.2009.05.002

Song, S., Härdle, W. K., Ritov, Y.
Generalized Dynamic Semiparametric Factor Models for High Dimensional Nonstationary Time Series
The Econometrics Journal, accepted for publication, 2013, DOI: 10.1111/ectj.12024



Spierdijk, L., Nijman, T.E., van Soest, A.H.O.
Price Dynamics and Trading Volume: A Semiparametric Approach
Department of Applied Mathematics, University of Twente, 2004, DOI: 10.2139/ssrn.512342

Tseng, P., Yun, S.
A coordinate gradient descent method for nonsmooth separable minimization
Mathematical Programming **117**(1-2): 387-423, 2009, DOI: 10.1007/s10107-007-0170-0



Model of Temperature data

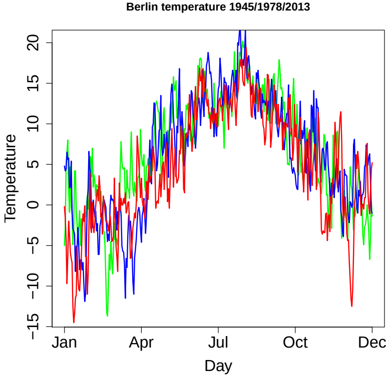
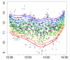


Figure: Daily average temperature in Berlin in 1948, 1980, 2013

Back to Motivation - Temperature Data



Model of Temperature data

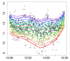
For days $t = 1, \ldots, 24090$ (i.e. 66 years)

$$X_t = T_t - \Lambda_t$$

$$\Lambda_t = a + bt + \sum_{m=1}^{2} \left\{ c_m \cos\left(\frac{m \cdot \pi \cdot t}{365}\right) + d_m \sin\left(\frac{m \cdot \pi \cdot t}{365}\right) \right\}$$

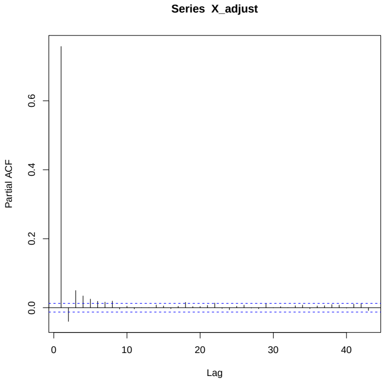
$$X_t = \sum_{l=1}^{L} \beta_l X_{t-l} + \varepsilon_t$$

Back to Motivation - Temperature Data

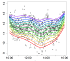


Model of Temperature data

$$\hat{a} = 5.617 \qquad \hat{\beta}_1 = 0.786$$
$$\hat{b} = 2.3 * 10^{-5} \qquad \hat{\beta}_2 = -0.078$$
$$\hat{c}_1 = -4.15 * 10^{-2} \qquad \hat{\beta}_3 = 0.024$$
$$\hat{c}_2 = -7.14 * 10^{-2} \qquad \hat{\beta}_4 = 0.015$$
$$\hat{d}_1 = -7.932 \qquad \hat{\beta}_5 = 0.011$$
$$\hat{d}_2 = -3.109 \qquad \hat{\beta}_6 = 0.007$$
$$\hat{\beta}_7 = 0.001$$
$$\hat{\beta}_8 = 0.019$$



Back to Motivation - Temperature Data



Relating Expectiles and Expected Shortfall

Newey and Powell (1987):

$$e_\tau = \arg\min_e \mathsf{E}\left\{ \left| \tau - \mathbf{I}_{\{Y<e\}} \right| (Y-e)^2 \right\}$$

$$\frac{1-2\tau}{\tau} \mathsf{E}\left\{ (Y - e_\tau)\, \mathbf{I}_{\{Y<e_\tau\}} \right\} = e_\tau - \mathsf{E}(Y)$$

Taylor (2008):

$$\mathsf{E}\left(Y | Y < e_\tau\right) = e_\tau + \frac{\tau \left\{ e_\tau - \mathsf{E}(Y) \right\}}{(1-2\tau) F(e_\tau)}$$

Back

B-splines

Knot vector $t = (t_1, \dots, t_M)$ as nondecreasing sequence in $[0, 1]$

Control points $P_0, \dots, P_N$

Define i-th B-spline basis function $N_{i,j}$ of order j as

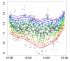
$$N_{i,0}(t) = \begin{cases} 1 & \text{if } t_i < t < t_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

$$N_{i,j}(t) = \frac{t - t_i}{t_{i+j} - t_i} N_{i,j-1}(t) + \frac{t_{i+j+1} - t}{t_{i+j+1} - t_{i+1}} N_{i+1,j-1}(t)$$

$j = 1, \dots, N - M - 1$

▸ Back to Estimation of Expectile Curves

▸ Back to Space Basis



Estimation of Quantile Curves
Definition of Penalty matrix

$$\widehat{\alpha} = \arg\min_{\alpha} \sum_{i=1}^{N} \rho_{\tau} \left\{ y_i - \alpha^\top b(x_i) \right\} + \lambda \alpha^\top \Omega \alpha$$

where $b(x) = (b_1, \ldots, b_q(x))^\top$ is vector of B-spline basis functions

Denote $\tilde{b}(x) = (\tilde{b}_1(x), \ldots, \tilde{b}_q(x))^\top$ the vector of second derivatives of basis functions

and set

$$\Omega = \int \tilde{b}(x)\tilde{b}(x)^\top \, dx$$

Back to Estimation of Expectile Curves

LAWS estimation

Schnabel and Eilers (2009):

$$\min \sum_{i=1}^{n} w_i(\tau)(y_i - \mu_i)^2$$

where

$$w_i(\tau) = \begin{cases} \tau & \text{if } y_i > \mu_i \\ 1-\tau & \text{if } y_i \leq \mu_i, \end{cases}$$

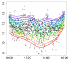
μ_i expected value according to some model.

Iterations:

- fixed weights, closed form solution of weighted regression
- recalculate weights

until convergence criterion met.

▸ Back to Expectiles Curves



LAWS estimation

Example:

Classical linear regression model

$$Y = X\beta + \varepsilon$$

where $\mathsf{E}(\varepsilon \,|X) = 0$ and $\mu = \mathsf{E}(Y\,|X) = X\beta$.

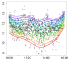
$$\widehat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} w_i (y_i - \mu_i)^2$$

Then:

$$\widehat{\beta} = (X^\top WX)^{-1} XWY$$

with W diagonal matrix of fixed weights w_i.

Back to Expectile Curves



Functional principal components

$X(t)$ stochastic process on compound interval T
with mean function $\mu(t) = \mathsf{E}\{X(t)\}$
and covariance function $K(s,t) = \mathrm{cov}(X(s), X(t))$

There exist orthogonal sequence of eigenfunctions ϕ_j and eigenvalues λ_i such that $K(s,t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s)\phi_j(t)$
We can rewrite process as

$$X(t) = \mu(t) + \sum_{j=1}^{\infty} \sqrt{\lambda_j}\kappa_j \phi_j(t)$$

where $\kappa_j = \frac{1}{\sqrt{\lambda_j}} \int X(t)\phi_j(s) ds$, $\mathsf{E}(\kappa_j) = 0$ and $\mathsf{E}(\kappa_j \kappa_K) = \delta_{jk}$.

▸ Back to Modelling of Time-varying Curves

Guo(2013) - Empirical loss function

$$S^* = S + M_\mu + M_\Phi$$

where

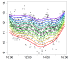
$$S = \sum_{d=1}^{D} \sum_{t=1}^{T} \rho_\tau \left(Y_{dt} - b(t)^\top \theta_\mu - b(t)^\top \Theta_\Phi \alpha_d \right)$$

$$M_\mu = \theta_\mu^\top \int \tilde{b}(x) \tilde{b}(x)^\top \, dx \theta_\mu = \theta_\mu^\top \Omega \theta_\mu$$

$$M_\Phi = \sum_{k=1}^{K} \theta_{\phi,k} \int \tilde{b}(x) \tilde{b}(x)^\top \, dx \, \theta_{\phi,k}$$

and $\tilde{b}(x)$ vector of second derivatives

Back to Guo(2013) - Estimation of Quantile Curves



Cramér-Karhunen-Loève representation

Conditions (Panaretos and Tavakoli (2013))

X_t second order stationary time series in $L^2([0,1],\mathbb{R})$ with zero mean, $\mathsf{E}\,\|X_0\|_2^2 < \infty$ and autocovariance kernel at lag t :

$$r_t(u,v) = \mathsf{E}\left(X_t(u)X_0(v)\right)$$

$u, v \in [0,1],\ t \in \mathbb{Z},\quad$ inducing operator :

$$\mathcal{R}_t : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{R})$$

Assume:

i) $\sum_{t\in\mathbb{Z}} \|\mathcal{R}_t\|_1 < \infty$

ii) $(u,v) \to r_t(u,v)$ continuous $t \in \mathbb{Z}$, and $\sum_{t\in\mathbb{Z}} \|r_t\|_\infty < \infty$

Back to C-K-L representation

Cramér-Karhunen-Loève representation

Theorem (Panaretos and Tavakoli (2013))

X_t admits representation

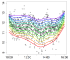
$$X_t = \int_{-\pi}^{\pi} \exp(\mathrm{i}\omega_j t) dZ_\omega \text{ a.s.}$$

where for fixed ω, Z_ω is random element of $\mathrm{L}^2([0,1],\mathbb{C})$
and process $\omega \to Z_\omega$ has orthogonal increments.

Integral can be understood as a Riemann-Stieltjes limit in sense

$$\mathrm{E}\,\|X_t - \sum_{j=1}^{J} \exp(\mathrm{i}\omega_j t)(Z_{\omega_{j+1}} - Z_{\omega_j})\|_2^2 \to 0 \ \text{ as } J \to 0$$

▸ Back to C-K-L representation



Cramér-Karhunen-Loève representation

Remark (Panaretos and Tavakoli (2013))

With spectral density operator $\mathcal{F}_\omega = \frac{1}{2\pi} \sum_{t\in\mathbb{Z}} \exp(-\mathrm{i}\omega t)\mathcal{R}_t$ having eigenfunctions $\{\phi_n^\omega\}_{t\geq 1}$

C-K-L representation can be interpreted as

$$X_t = \int_{-\pi}^{\pi} \exp(\mathrm{i}) \sum_{n=1}^{\infty} \langle \phi_n^\omega, dZ_\omega \rangle \, \phi_n^\omega$$

Back to C-K-L representation

Block-Coordinate Gradient Descent Algorithm

Tseng & Yun (2009)

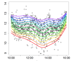
$$\min_x f(x) + \lambda P(x)$$

Solve

$$\min_d \left\{ d^\top \nabla f(x) + \frac{1}{2} d^\top H d + \lambda P(x+d) \right\}$$

- P is block-separable then H is block-diagonal
- solve subproblems
  (for every j take $d = (d_{\mathcal{G}_1}, \dots, d_{\mathcal{G}_J})$ with $d_{\mathcal{G}_k} = 0$ for $k \neq j$)

Back to BCGD algorithm



Block-Coordinate Gradient Descent Algorithm

$$d_{\mathcal{G}_j}^{(t)} = -\frac{1}{h_{\mathcal{G}_j}^{(t)}} \left(\nabla l(\widehat{C}^{(t)})_{\mathcal{G}_j} - \lambda \frac{\nabla l(\widehat{C}^{(t)})_{\mathcal{G}_j} - h_{\mathcal{G}_j}^{(t)} \widehat{C}_{\mathcal{G}_j}^{(t)}}{\| \nabla l(\widehat{C}^{(t)})_{\mathcal{G}_j} - h_{\mathcal{G}_j}^{(t)} \widehat{C}_{\mathcal{G}_j}^{(t)} \|_2} \right)$$

$H^{(t)}$ has submatrices $H_{\mathcal{G}_j}^{(t)} = h_{\mathcal{G}_j}^{(t)} I_{\mathcal{G}_j}$ for scalars $h_{\mathcal{G}_j}^{(t)}$

$\alpha^{(t)}$ set by Amijo rule (See Details: Tseng & Yun (2009))

Back to BCGD algorithm

Song et. al.(2013) - Time basis

Orthogonal Legendre polynomial basis
to capture the global trend in time
$u_1(d) = 1/C_1,\ u_2(d) = d/C_2,\ u_3(d) = (3t^2 - 1)/C_3, \dots$
with generic constant C_i such that $\sum_{d=1}^{D} \mathrm{u}_i(d)/C_i^2 = 1$

Fourier series
to capture periodic variations
$u_4 = \sin(2\pi d/p)/C_4,\ u_5 = \cos(2\pi d/p)/C_5,$
$u_6 = \sin(2\pi d/(p/2))/C_6, \dots$
with given period p

Back to Song(2013) -Time Basis