TEDAS - Tail Event Driven Asset Allocation

Wolfgang Karl Härdle

Sergey Nasekin

David Lee

Phoon Kok Fai

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://www.case.hu-berlin.de

S&P 500 Stocks

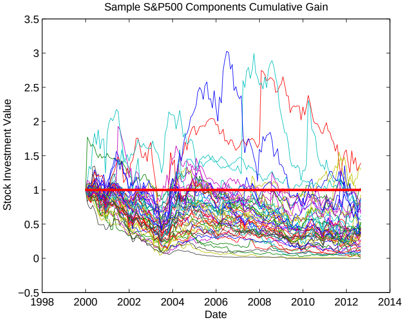
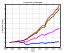


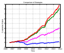
Figure 1: 50 random S&P 500 Sample Components' Cumulative Return: **94%** of stocks lost the value of the initial investment (thick red line)



Hedge Funds

A **hedge fund** is an "aggressively managed portfolio of investments that uses advanced investment strategies such as leveraged, long, short and derivative positions in both domestic and international markets with the goal of generating high returns".

- diversification - reduction of the portfolio risk
- construction - a more diverse universe of assets
- allocation - a higher risk-adjusted return.



Hedge Funds

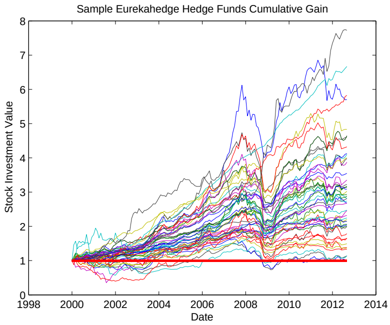
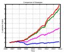


Figure 2: 50 Eurekahedge Hedge Funds Indices' Cumulative Return: **0%** of funds lost the value of the initial investment (thick red line)



Hedge Funds and Diversification

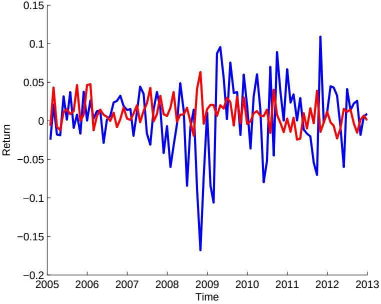
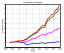


Figure 3: S&P 500 (in blue) and Eurekahedge North America Macro Hedge Fund Index (in red) monthly returns in 31.01.2005-31.12.2012

TEDAS Strategy

- Improving Asset Allocation
  - VaR optimization
  - Higher portfolio moments
- Hedging tail events
  - Lasso quantile regression
  - Applications, Monte-Carlo study



Outline

1. Motivation ✓
2. Hedge Funds and Diversification
3. Asset Allocation with Hedge Funds
4. Portfolio Trading Strategy
5. Simulations
6. TEDAS Strategy
7. Lasso Shrinkage Method

Mean-Variance Asset Allocation

Markowitz diversification rule:

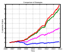
$$\begin{aligned}
\underset{w \in \mathbb{R}^d}{\text{minimize}} \quad & \sigma_p^2(w) \stackrel{\text{def}}{=} w^\top \Sigma w \\
\text{subject to} \quad & w^\top \mu = \mu_p, \\
& \sum_{i=1}^{d} w_i = 1, \\
& w_i \geq 0
\end{aligned}$$

here w_i weights, Σ covariance of asset returns r_i, μ_p is the "target" return

Diversification Concept

Reduce specific risk; remain with market risk. Mean-variance theory implies

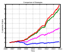
- diversification is beneficial when returns are uncorrelated or negatively correlated;
- diversification increases certainty when returns are uncorrelated and variances are identical;
- avoid investing in securities with high covariances among themselves.



Correlation Examples - Traditional Assets

MSCI Indices	US	UK	SW	GER	JAP
US (US)	1.00				
UK (UK)	0.69	1.00			
SW (Switzerland)	0.51	0.58	1.00		
GER (Germany)	0.60	0.59	0.50	1.00	
JAP (Japan)	0.47	0.45	0.38	0.24	1.00

Table 1: Correlation statistics for traditional asset class indices; based on monthly data Jan. 1994 - Aug. 2001; table from Lhabitant (2002, p.158)

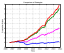


Traditional Assets/Hedge Funds

Hedge Funds	US	UK	SW	GER	JAP
Conv. arb.	0.10	0.08	0.07	0.10	−0.02
Dedic. sh. bias	−0.77	−0.53	−0.33	−0.46	−0.48
Fix. inc. arb.	0.10	0.13	0.01	0.08	−0.10
Glob. macro	0.30	0.19	0.10	0.27	−0.11
Man. fut.	−0.10	0.02	−0.09	−0.03	0.03

Table 2: Correlation statistics for traditional asset class and hedge funds' indices; based on monthly data Jan. 1994 - Aug. 2001; table from Lhabitant (2002, p.164)

▸ Details for Hedge Funds Strategies ▸ More



Asset Allocation and Hedge Funds

Combine several assets to maximize risk-adjusted performance consistently with the investor's preferences.

Hedge funds:

- may in reality be a conservative investment
- offer superior risk-adjusted returns
- better diversification
- dynamic.

Risk-Adjusted Return

Hedge funds offer superior risk-adjusted returns

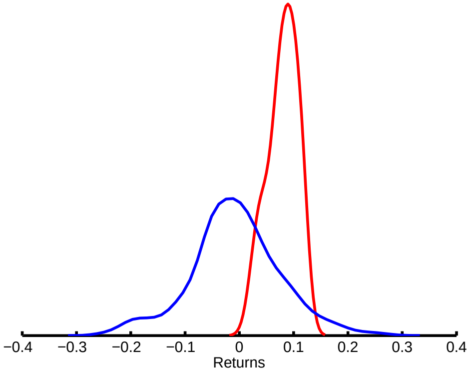


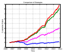
Figure 4: Estimated kernel densities for S&P 500 components (in blue) and for Eurekahedge hedge funds indices (in red) returns

Efficient Frontier

Efficient frontier is the set of optimal portfolio risk and return values $(\sigma_p(w), \mu_p)$ such that $\mu_p \geq \mu_{min.var.}$, i.e., such portfolios which have their expected return at least as large as the minimum variance portfolio.

Portfolios that lie below the efficient frontier are *sub-optimal*: they do not provide enough return for the level of risk.

Portfolios made solely of stocks are sub-optimal to those which include hedge funds.



Efficient Frontier Example

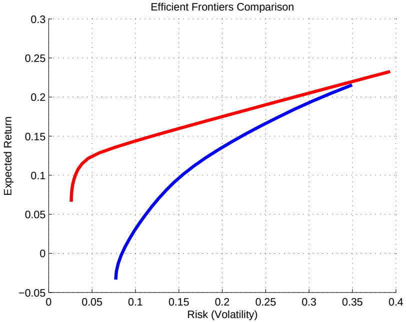
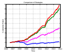
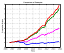


Figure 5: Efficient frontiers built by using all S&P 500 components only (in blue) and by randomly mixing them with Eurekahedge hedge funds indices (in red)



Alternative Allocation Approach

- hedge funds' returns: negative skewness and/or positive excess kurtosis - mean-variance approach underestimates portfolio risk
- returns' covariance structure is time-changing
- VaR as the objective optimization function
- adjust VaR via Cornish-Fisher (CF) expansion
- multivariate dynamic volatility/higher moments to model covariances



ACF diagnostics

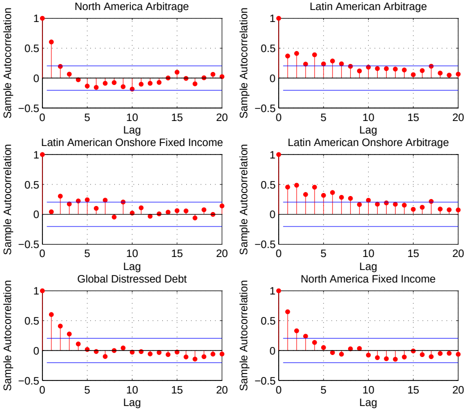


Figure 6: Sample ACFs of the squared returns for selected hedge funds

Test for ARCH Effects

H0: no conditional heteroscedasticity in zero-mean returns e_t
H1: the ARCH of e_t:

$$e_t^2 = \alpha_0 + \sum_{k=1}^{p} \alpha_k e_{t-k}^2 + u_t, \quad u_t \sim \mathrm{N}(0, \sigma^2), \text{i.i.d.} \tag{1}$$

at least one $\alpha_k \neq 0$, $k = 1, \ldots, p$; test statistic $n \cdot R^2 \sim \chi_p^2$, n sample size, R^2 coefficient of determination in (1)

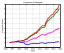
Test for ARCH Effects - Results

Table 3: Test for ARCH effects in selected hedge funds' returns residuals

Hedge Fund Name	Test Results			
	Test Stat.	Crit. Value	P-value	Conclusion
North America Arbitrage	33.66	5.99	0.00	H0 rej.
Latin American Arbitrage	6.73	5.99	0.03	H0 rej.
Latin American Onshore Fixed Income	2.75	5.99	0.25	H0 not rej.
Latin American Onshore Arbitrage	9.53	5.99	0.01	H0 rej.
Global Distressed Debt FoF	60.06	5.99	0.00	H0 rej.
North America Fixed Income	47.20	5.99	0.00	H0 rej.

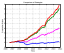
The significance level for the test is 0.05; H0: no ARCH effect

$p = 2$ lags assumed



Modelling Variance-Covariance Structure

1. Exponentially Weighted Moving Average estimator
2. Orthogonal GARCH framework: modelling $\widehat{\Sigma}_t = B_t \Delta_t B_t^\top + \Omega_t$ ▸ Details
3. Dynamic Conditional Correlation (DCC) framework: modelling $\widehat{\Sigma}_t = D_t R_t D_t$ ▸ Details
4. Multivariate time-varying distribution approach ▸ Details



Cornish-Fisher VaR Optimization

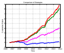
The alternative asset allocation (Favre, Galeano, 2002)

$$\begin{aligned} \underset{w \in \mathbb{R}^d}{\text{minimize}} \quad & W_t\{-q_\alpha(w_t) \cdot \sigma_p(w_t)\} \\ \text{subject to} \quad & w_t^\top \mu = \mu_p, \;\; w_t^\top 1 = 1, \;\; w_{t,i} \geq 0 \end{aligned}$$

here $W_t \stackrel{\text{def}}{=} W_0 \cdot \prod_{j=1}^{t-1} w_{t-j}^\top (1 + r_{t-j})$, $\tilde{w}$, W_0 initial wealth,
$\sigma_p^2(w) \stackrel{\text{def}}{=} w_t^\top \Sigma_t w_t$,

$$q_\alpha(w) \stackrel{\text{def}}{=} z_\alpha + (z_\alpha^2 - 1)\frac{S_p(w)}{6} + (z_\alpha^3 - 3z_\alpha)\frac{K_p(w)}{24} - (2z_\alpha^3 - 5z_\alpha)\frac{S_p(w)^2}{36},$$

here $S_p(w)$ skewness, $K_p(w)$ kurtosis, z_α is $N(0,1)$ α-quantile
▸ Details. If $S_p(w)$, $K_p(w)$ zero, then obtain Markowitz allocation



TEDAS Idea

How to select hedge funds which are related to S&P 500 tail events (TE)?

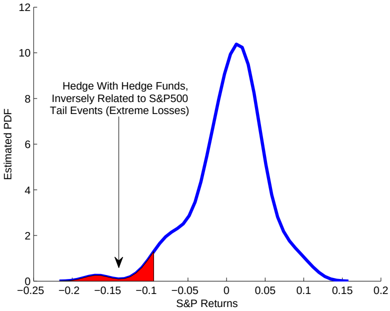
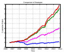


Figure 7: Hedge funds can provide insurance for S&P 500 for tail events

TEDAS Technique

- consider a data vector $Y \in \mathbb{R}^n$ of S& P 500 returns and a matrix $X \in \mathbb{R}^{n\times p}$ of hedge funds' returns, $p > n$;
- negative S& P 500 TE and hedge funds TE: $q_\tau(x) \stackrel{\text{def}}{=} F^{-1}_{y|x}(\tau)$ $= x^\top \beta(\tau) = \arg\min_{\beta\in\mathbb{R}^p} \mathsf{E}_{Y|X=x}\, \rho_\tau\{Y - X\beta\}$,
  $\rho_\tau(u) = u\{\tau - \mathbf{I}(u<0)\}$, $\tau \in (0,1)$, β tail dependence;
- deal with $p > n$ by introducing L_1 penalty $\lambda_n \|\hat{\omega}^\top \beta\|_1$ to nullify "excessive" coefficients; λ_n and $\hat{\omega}$ controlling penalization; constraining $\beta \le 0$ gives the final Adaptive Lasso Quantile Regression (ALQR) estimator ▸ Details

$$\hat{\beta}^{\text{adapt}}_{\tau,\lambda_n} = \arg\min_{\beta\in\mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top\beta) + \lambda_n \|\hat{\omega}^\top\beta\|_1, \quad \beta \le 0$$

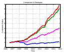


TEDAS Monte-Carlo Analysis

- $\lambda_n = 0.25\sqrt{\|\hat{\beta}^{\text{init}}\|_0} \log(n \vee p)(\log n)^{0.1/2}$, $\hat{\omega}_j = 1/|\hat{\beta}_j^{\text{init}}| \wedge \sqrt{n}$; $\hat{\beta}^{\text{init}}$ are taken as $\hat{\beta}_{\tau,\lambda}$ from (4) (for oracle ▸ Details results to hold);
- $X_i \sim \text{N}(0, \Omega)$, $n = 50$, $p = 300$, $\beta = (-5, -2, -1, 3, 1, 0.5, 0, ..., 0)$, $q = 3$, $\varepsilon_i \sim \text{N}(0, \sigma^2)$;
- $\Omega_{i,j} = 0.5^{|i-j|}$, $\sigma = 0.1, 0.5, 1$ (three levels of noise);
- λ for $\hat{\beta}_{\tau,\lambda}$ is chosen to minimize the BIC criterion:

$$\text{BIC}_{\tau,\lambda} \stackrel{\text{def}}{=} \log\left\{ n^{-1} \cdot \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \hat{\beta}_{\tau,\lambda}) \right\} + \frac{\log(n)}{2n} \cdot \widehat{\text{df}}(\lambda)$$

where $\widehat{\text{df}}(\lambda) \stackrel{\text{def}}{=} \|\hat{\beta}_{\tau,\lambda}\|_0 = \hat{q}$. ▸ Results



ALQR Accuracy Criteria

Model selection is assessed according to a number of accuracy criteria:

1. Standardized L_2-norm

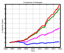
$$\text{Dev} \stackrel{\text{def}}{=} \frac{\|\beta - \hat{\beta}\|_2}{\|\beta\|_2}$$

2. Sign consistency

$$\text{Acc} \stackrel{\text{def}}{=} \sum_{j=1}^{p} |\operatorname{sign}(\beta_j) - \operatorname{sign}(\hat{\beta}_j)|$$

3. Least angle

$$\text{Angle} \stackrel{\text{def}}{=} \frac{< \beta, \hat{\beta} >}{\|\beta\|_2 \|\hat{\beta}\|_2}$$



Accuracy Criteria

4. Estimate of q

$$\text{Est} \stackrel{\text{def}}{=} \hat{q}$$

5. Empirical risk

$$\text{Risk} \stackrel{\text{def}}{=} \sqrt{n^{-1} \sum_{i=1}^{n} \left\{ X_i^\top (\beta - \hat{\beta}) \right\}^2}$$

Monte-Carlo Analysis Results

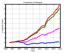
Table 4: Criteria Results under Different Models and Quantiles

Accuracy Crit. and Model		Noise Levels and Quantile Indices					
		σ=0.1		σ=0.5		σ=1	
		τ=0.1	τ=0.9	τ=0.1	τ=0.9	τ=0.1	τ=0.9
Dev	**ALQR**	*0.61(0.05)*	*0.61(0.06)*	*0.61(0.06)*	*0.60(0.05)*	*0.62(0.06)*	*0.60(0.05)*
	LQR	0.65(0.07)	0.68(0.08)	0.65(0.06)	0.64(0.06)	0.66(0.07)	0.64(0.07)
Acc	**ALQR**	*5.93(1.34)*	*6.14(1.41)*	*6.04(1.45)*	*6.04(1.62)*	*6.08(1.52)*	*5.88(1.14)*
	LQR	13.55(15.38)	21.06(19.66)	14.28(15.66)	10.93(12.35)	16.42(16.98)	10.77(13.46)
Angle	**ALQR**	*0.43(0.07)*	*0.45(0.09)*	*0.46(0.11)*	*0.45(0.09)*	*0.48(0.12)*	*0.44(0.05)*
	LQR	0.61(0.18)	0.66(0.20)	0.61(0.15)	0.62(0.17)	0.62(0.15)	0.64(0.20)
Est	**ALQR**	*3.13(1.44)*	*3.26(1.40)*	*3.12(1.25)*	*3.36(1.19)*	*3.15(1.54)*	*3.04(1.15)*
	LQR	10.95(15.80)	18.48(20.02)	11.56(16.01)	8.31(12.55)	13.65(17.47)	8.15(13.70)
Risk	**ALQR**	*0.58(0.25)*	*0.58(0.24)*	*0.89(0.26)*	*0.84(0.24)*	*0.89(0.27)*	*0.88(0.24)*
	LQR	0.67(0.27)	0.72(0.30)	0.95(0.26)	0.93(0.23)	0.94(0.30)	0.98(0.27)

Model notation: ALQR - Adaptive Lasso quantile regression; LQR - simple Lasso quantile regression

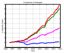
Standard deviations are given in brackets

Number of replications is 100



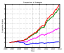
Data for Analysis

- data on 166 monthly log-returns of 164 Eurekahedge hedge funds indices in the period of 31.01.2000 - 31.10.2013 (source: *Bloomberg*)
- data on 166 monthly log-returns of S&P 500 in the period of 31.01.2000 - 31.10.2013
- the idea is to hedge the benchmark asset (e.g.,S&P 500) with a security (e.g., hedge fund) moving in opposite direction at different quantiles



TEDAS Build-Up

- moving window, width 80 months: the covariate matrix $X_l \in \mathbb{R}^{80\times164}$, $(n = 80,\ p = 164)$; the response vector $Y_l \in \mathbb{R}^{80}$
- $\tau_{1,2,3,4,5} = (0.05, 0.15, 0.25, 0.35, 0.50)$
- $\hat{q}_\tau \stackrel{\text{def}}{=} F_n^{-1}(\tau)$ is the S&P 500 log-returns sample τ-quantiles obtained from the S&P 500 log-returns edf F_n
- $\hat{\beta}_{\tau,\lambda_n}$ are the estimated non-zero ALQR coefficients
- the GARCH process for individual assets is assumed to be GARCH (1,1) with mean equation specified as ARMA (1,1) model



TEDAS Explained

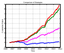
Initial wealth $W_0 = \$1$, at $t = l, \ldots, n$:

1. determine S& P 500 return r_t, select $\tau_{j,t}, j = 1, \ldots, 5$ according to the right $\hat{q}_{\tau_t}$ in: $r_t \leq \hat{q}_{\tau_{1,t}}$, $\hat{q}_{\tau_{1,t}} < r_t \leq \hat{q}_{\tau_{2,t}}$, $\hat{q}_{\tau_{2,t}} < r_t \leq \hat{q}_{\tau_{3,t}}$, $\hat{q}_{\tau_{3,t}} < r_t \leq \hat{q}_{\tau_{4,t}}$, $\hat{q}_{\tau_{4,t}} < r_t \leq \hat{q}_{\tau_{5,t}}$

2. ALQR for $\hat{\beta}_{\tau_{j,t},\lambda_n}$ using the observations $X \in \mathbb{R}^{t-l+1,\ldots,t \times p}$, $Y \in \mathbb{R}^{t-l+1,\ldots,t}$, apply CF-VaR to hedge funds with $\hat{\beta}_{\tau_{j,t},\lambda_n} \neq 0$ to determine optimal weights

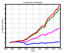
3. buy hedge funds with $\hat{\beta}_{\tau_{j,t},\lambda_n} \neq 0$ taken with estimated weights

4. if none of the inequalities from Step 2 holds, invest into the benchmark asset (S&P 500) at r_t.



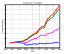
TEDAS Example

1. Suppose $t = 58$ (Oct. 2009), accumulated wealth $W_{58} = \$1.429$, $r_{58} = -1.85\%$.
2. It occurs that $\hat{q}_{0.15,58} = -4.18\%$, $\hat{q}_{0.25,58} = -1.85\%$ and so $\hat{q}_{0.15,58} < r_{58} \leq \hat{q}_{0.25,58}$.
3. ALQR using $X \in \mathbb{R}^{9,\ldots,58\times 170}$, $Y \in \mathbb{R}^{9,\ldots,58}$ yields $\hat{\beta}_{0.25,58} = (-0.77, -1.12, -0.41)$, i.e. three hedge funds, *Latin American Arbitrage*, *North America Macro*, *Emerging Markets CTA/Managed Futures*.
4. CF-VaR optimization problem yields $w = (0.22, 0.16, 0.62)$ as solution, this portfolio yields a return of 0.62% ($W_{59} = \$1.438$), while the benchmark asset return has been -1.85%.



All Strategies

- "Strategy 1" - ALQR-based TEDAS with CF-VaR optimization and DCC-modeled covariance structure
- "Strategy 2" - base case S&P 500 "buy-and-hold"
- "Strategy 3" - ALQR-based TEDAS with "naive" diversification (always equal weights, no optimization)
- "Strategy 4" - based on Orthogonal GARCH model and simple variance-covariance VaR optimization



Strategies' Returns Comparison

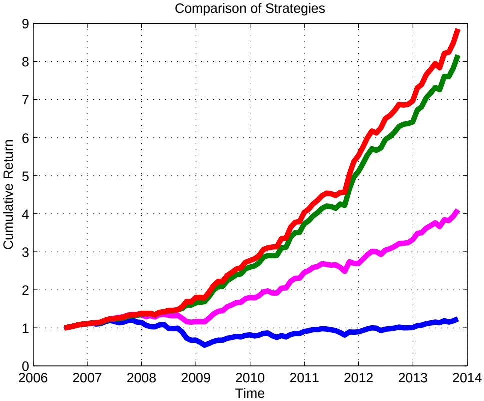


Figure 8: Strategies' cumulative returns' comparison: Strategy 1 (in red), Strategy 2 (in blue), Strategy 3 (in green), Strategy 4 (in magenta)

Strategies' Utilities Comparison

- logarithmic Bernoulli utility function $u(x) = \log(x)$

- expansion around $\overline{W} \stackrel{\text{def}}{=} W_0 \cdot w^\top(1+\overline{r})$ up to the 4th order and taking expectations yields expected utility

$$
\begin{aligned}
\mathsf{E}\{u(W_t)\} &\simeq u(m_1) + u'(m_1)\,\mathsf{E}(Y - m_1) + \frac{1}{2}u''(m_1)\,\mathsf{E}\{(Y - m_1)^2\} \\
&+ \frac{1}{3!}u^{(3)}(m_1)\,\mathsf{E}\{(Y - m_1)^3\} + \frac{1}{4!}u^{(4)}(m_1)\,\mathsf{E}\{(Y - m_1)^4\} \\
&= u(\overline{W}) + u'(\overline{W})\,\mathsf{E}(W_t - \overline{W}) + \frac{1}{2}u''(\overline{W})\,\mathsf{E}\{(W_t - \overline{W})^2\} \\
&+ \frac{1}{3!}u^{(3)}(\overline{W})\,\mathsf{E}\{(W_t - \overline{W}^3\} + \frac{1}{4!}u^{(4)}(\overline{W})\,\mathsf{E}\{(W_t - \overline{W})^4\} \\
&\simeq \log(\overline{W}) - \frac{1}{2\overline{W}^2}\sigma_P^2 + \frac{1}{3!\overline{W}^3}S_P - \frac{1}{4!\overline{W}^4}K_P
\end{aligned}
$$

Strategies' Utilities Comparison

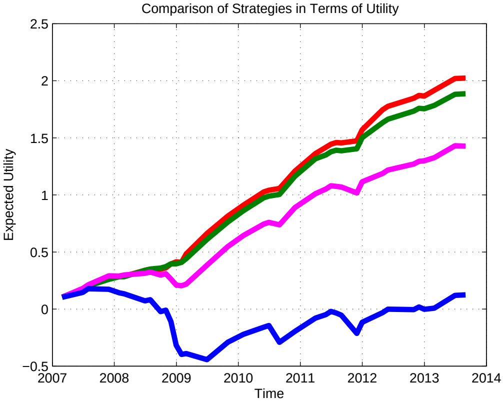


Figure 9: Comparison of expected utilities for 4 strategies: Strategy 1 (in red), Strategy 2 (in blue), Strategy 3 (in green), Strategy 4 (in magenta)

Strategies' Returns Comparison

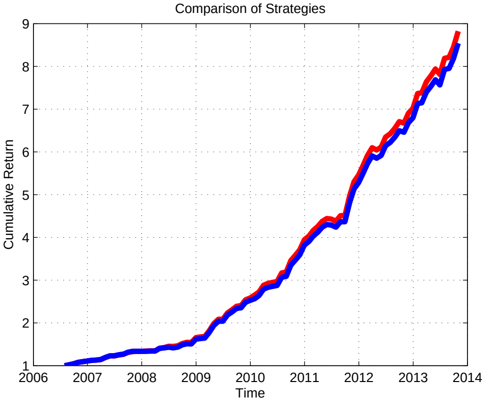
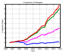


Figure 10: Strategies' cumulative returns' comparison: Strategy using dynamic distribution (in red), Strategy 1 (in blue)



Dynamics of Conditional Moments

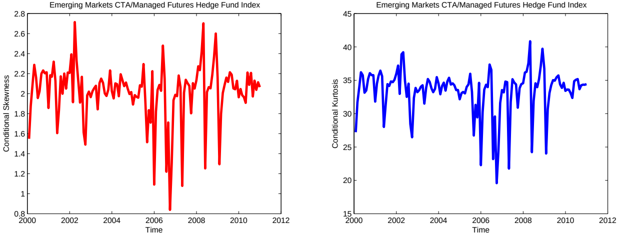


Figure 11: Dynamics of conditional skewness and kurtosis for Emerging Markets CTA/Managed Futures Hedge Fund Index

Strategies' Transaction Cost

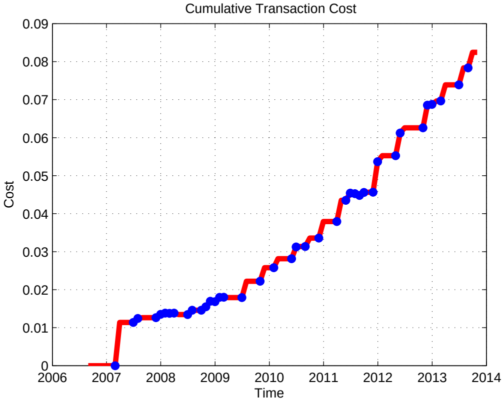


Figure 12: Cumulative transaction cost for Strategies 1,3,4 (each time 1% of trade value); blue dots denote portfolio rebalancing

Strategies' Risk Comparison

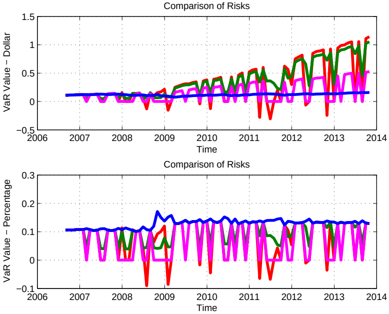


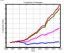
Figure 13: Absolute (upper) and relative(lower) CF-VaR for Strategies: 1 (in red), 2 (in blue), 3 (in green), 4(in magenta)

-$\widehat{\beta}$ in each window, $\tau = 0.05$

Figure 14: Different -$\widehat{\beta}$ in application; $\tau = 0.05$

-$\widehat{\beta}$ in each window, $\tau = 0.15$

Figure 15: Different $-\widehat{\beta}$ in application; $\tau = 0.15$ ▸ Selected Funds



Histograms of $\hat{q}$

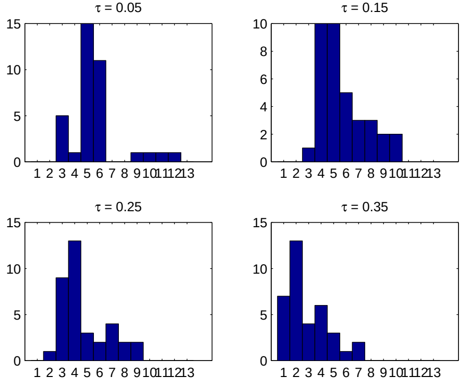
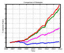


Figure 16: Frequency of the number of selected variables for 4 different τ



The estimated value of λ

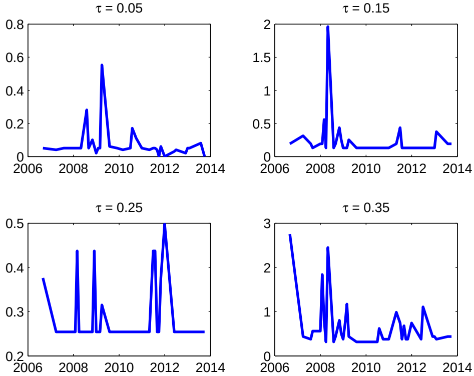


Figure 17: The estimated λ for 4 different τ

The Influential Variables

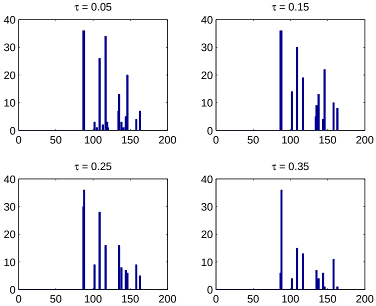


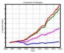
Figure 18: The frequency of the hedge funds

Selected Hedge Funds

Table 5: The selected hedge funds for $\tau = 0.05$

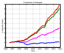
Top 9 influential hedge funds	Frequency
Emerging Markets Arbitrage Hedge Fund Index	36
Emerging Markets CTA/Managed Futures Hedge Fund Index	36
CTA/Managed Futures Hedge Fund Index	34
Asia inc. Japan Macro Hedge Fund Index	26
Europe CTA Managed Futures Hedge Fund Index	20
Asia CTA Hedge Fund Index	13
Asia Convertible Arbitrage Hedge Fund Index	7
Europe Relative Value Hedge Fund Index	7
Europe Arbitrage Hedge Fund Index	5

▸ Return to betas



Conclusions

- Improving Asset Allocation
  - Hedge funds successfully replace conventional assets in portfolios
  - Diversification with hedge funds is superior to traditional financial instruments
- Hedging tail events
  - TEDAS strategy performs better than benchmark models
  - Modelling portfolio dynamics produces gains in terms of risk-adjusted return



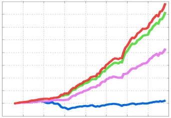
TEDAS - Tail Event Driven Asset Allocation

Wolfgang Karl Härdle

Sergey Nasekin

David Lee

Phoon Kok Fai



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://www.case.hu-berlin.de

Lasso Shrinkage

Linear model: $Y = X\beta + \varepsilon$; $Y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$, $\beta \in \mathbb{R}^p$, $\{\varepsilon_i\}_{i=1}^n$ i.i.d., independent of $\{X_i; i = 1, ..., n\}$

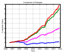
The optimization problem for the lasso estimator:

$$\begin{aligned} \hat{\beta}^{\text{lasso}} = & \ \arg \min_{\beta \in \mathbb{R}^p} f(\beta) \\ & \text{subject to} \quad g(\beta) \geq 0 \end{aligned} \tag{2}$$

where

$$f(\beta) = \frac{1}{2}(y - X\beta)^\top (y - X\beta)$$
$$g(\beta) = t - \|\beta\|_1$$

where t is the size constraint on $\|\beta\|_1$ Back to "ALQR Estimator"



Lasso Duality

If (1) is convex programming problem, then the Lagrangian is

$$L(\beta, \lambda) = f(\beta) - \lambda g(\beta).$$

and the primal-dual relationship is

$$\underbrace{\underset{\beta}{\text{minimize}} \ \sup_{\lambda \geq 0} \ L(\beta, \lambda)}_{\text{primal}} \geq \underbrace{\underset{\lambda \geq 0}{\text{maximize}} \ \inf_{\beta} \ L(\beta, \lambda)}_{\text{dual}}$$

Then the dual function $L^*(\lambda) = \inf_{\beta} \ L(\beta, \lambda)$ is

$$L^*(\lambda) = \frac{1}{2} y^\top y - \frac{1}{2} \hat{\beta}^\top X^\top X \hat{\beta} - t \frac{(y - X\hat{\beta})^\top X \hat{\beta}}{\|\hat{\beta}\|_1}$$

with $(y - X\hat{\beta})^\top X\hat{\beta} / \|\hat{\beta}\|_1 = \lambda$ Back to "ALQR Estimator"

Paths of Lasso Coefficients

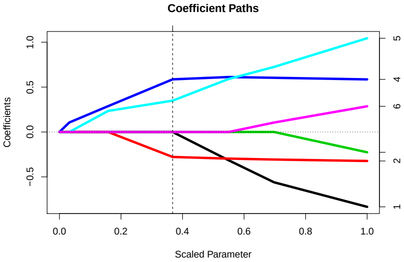
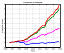


Figure 19: Lasso shrinkage of coefficients in the hedge funds dataset example (6 covariates were chosen for illustration); each curve represents a coefficient as a function of the scaled parameter $\hat{s} = t/\|\beta\|_1$; the dashed line represents the model selected by the BIC information criterion ($\hat{s} = 3.7$)

Back to "ALQR Estimator"



Example of Lasso Geometry

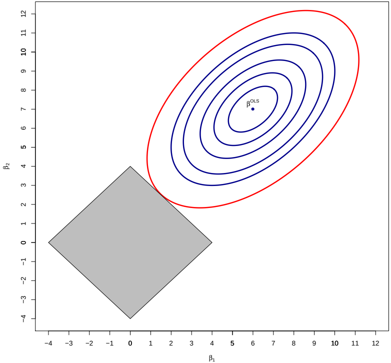


Figure 20: Contour plot of the residual sum of squares objective function centered at the OLS estimate $\widehat{\beta}^{ols} = (6, 7)$ and the constraint region $\sum |\beta_j| \leq t$ MVAlassocontour

Quantile Regression

The loss $\rho_\tau(u) = u\{\tau - \mathbf{I}(u < 0)\}$ gives the (conditional) quantiles $F^{-1}_{y|x}(\tau) \stackrel{\text{def}}{=} q_\tau(x)$.

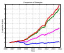
Minimize

$$\hat{\beta}_\tau = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta).$$

Re-write:

$$\underset{(\xi,\zeta)\in\mathbb{R}^{2n}_{+}}{\text{minimize}} \quad \left\{ \tau 1_n^\top \xi + (1-\tau) 1_n^\top \zeta | X\beta + \xi - \zeta = Y \right\}$$

with ξ, ζ are vectors of "slack" variables Back to "ALQR Estimator"

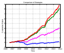


Non-Positive (NP) Lasso-Penalized QR

The lasso-penalized QR problem with an additional non-positivity constraint takes the following form:

$$
\begin{aligned}
\underset{(\xi,\zeta,\eta,\tilde{\beta}) \in \mathbb{R}_{+}^{2n+p} \times \mathbb{R}^{p}}{\text{minimize}} \quad & \tau 1_n^\top \xi + (1-\tau) 1_n^\top \zeta + \lambda 1_n^\top \eta \\
\text{subject to} \quad & \xi - \zeta = Y + X\tilde{\beta}, \\
& \xi \geq 0, \\
& \zeta \geq 0, \\
& \eta \geq \tilde{\beta}, \\
& \eta \geq -\tilde{\beta}, \\
& \tilde{\beta} \geq 0, \qquad \tilde{\beta} \stackrel{\text{def}}{=} -\beta
\end{aligned} \tag{3}
$$

Back to "ALQR Estimator"



Solution

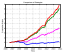
Transform into matrix (I_p is $p \times p$ identity matrix; $E_{p\times n} = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$):

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax = b, \;\; Bx \leq 0 \end{aligned}$$

where $A = \begin{pmatrix} I_n & -I_n & 0 & X \end{pmatrix}$, $b = Y$, $x = \begin{pmatrix} \xi & \zeta & \eta & \beta \end{pmatrix}^\top$,

$$c = \begin{pmatrix} \tau 1_n \\ (1-\tau)1_n \\ \lambda 1_p \\ 0 1_p \end{pmatrix}, \quad B = \begin{pmatrix} -E_{p\times n} & 0 & 0 & 0 \\ 0 & -E_{p\times n} & 0 & 0 \\ 0 & 0 & -I_p & I_p \\ 0 & 0 & -I_p & -I_p \\ 0 & 0 & 0 & I_p \end{pmatrix}$$

Back to "ALQR Estimator"



Solution - Continued

The previous problem may be reformulated into *standard form*

$$
\begin{aligned}
\text{minimize} \quad & c^\top x \\
\text{subject to} \quad & Cx = d, \\
& x + s = u, \;\; x \geq 0, s \geq 0
\end{aligned}
$$

and the dual problem is:

$$
\begin{aligned}
\text{maximize} \quad & d^\top y - u^\top w \\
\text{subject to} \quad & C^\top y - w + z = c, \;\; z \geq 0, w \geq 0
\end{aligned}
$$

Back to "ALQR Estimator"

Solution - Continued

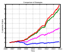
The KKT conditions for this linear program are

$$F(x, y, z, s, w) = \left\{ \begin{array}{c} Cx - d \\ x + s - u \\ C^\top y - w + z - c \\ x \circ z \\ s \circ w \end{array} \right\} = 0,$$

with $y \geq 0$, $z \geq 0$ dual slacks, $s \geq 0$ primal slacks, $w \geq 0$ dual variables.

This can be solved by a primal-dual path following algorithm based on the *Newton method*

▸ Back to "ALQR Estimator"



Adaptive Lasso Procedure

Lasso estimates $\hat{\beta}$ can be inconsistent (Zou, 2006) in some scenarios.

Lasso soft-threshold function gives biased results

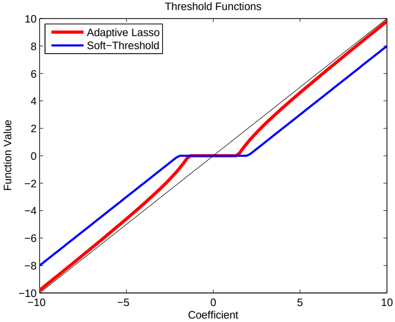


Figure 21: Threshold functions for simple and adaptive Lasso

Adaptive Lasso Procedure

The adaptive Lasso (Zou, 2006) yields a sparser solution and is less biased.

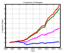
L_1 - penalty replaced by a re-weighted version; $\hat{\omega}_j = 1/|\hat{\beta}_j^{\text{init}}|^{\gamma}$, $\gamma = 1$, $\hat{\beta}^{\text{init}}$ is from (2)

The adaptive lasso estimates are given by:

$$\hat{\beta}_{\lambda}^{\text{adapt}} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} (Y_i - X_i^{\top}\beta)^2 + \lambda \|\hat{\omega}^{\top}\beta\|_1$$

(Bühlmann, van de Geer, 2011): $\hat{\beta}_j^{\text{init}} = 0$, then $\hat{\beta}_j^{\text{adapt}} = 0$

Back to "ALQR Estimator"



Simple and Adaptive Lasso Penalized QR

Simple lasso-penalized QR optimization problem is:

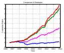
$$\hat{\beta}_{\tau,\lambda} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta) + \lambda \|\beta\|_1 \qquad (4)$$

Adaptive lasso-penalized QR model uses the re-weighted penalty:

$$\hat{\beta}_{\tau,\lambda}^{\text{adapt}} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta) + \lambda \|\hat{\omega}^\top \beta\|_1 \qquad (5)$$

Adaptive lasso-penalized QR procedure can ensure oracle properties for the estimator ▸ Details

▸ Back to "ALQR Estimator"



Algorithm for Adaptive Lasso Penalized QR

The optimization for the adaptive lasso quantile regression can be re-formulated as a lasso problem:

- the covariates are rescaled: $\tilde{X} = (X_1 \circ \hat{\beta}_1^{\text{init}}, \ldots, X_p \circ \hat{\beta}_p^{\text{init}})$;
- the lasso problem (4) is solved:

$$\hat{\hat{\beta}}_{\tau,\lambda} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau (Y_i - \tilde{X}_i^\top \beta) + \lambda \|\beta\|_1$$

- the coefficients are re-weighted as $\hat{\beta}_{\tau,\lambda}^{\text{adapt}} = \hat{\hat{\beta}}_{\tau,\lambda} \circ \hat{\beta}^{\text{init}}$

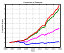
Back to "ALQR Estimator"

Oracle Properties of an Estimator

An estimator has oracle properties if (Zheng et al., 2013):

- it selects the correct model with probability converging to 1;
- the model estimates are consistent with an appropriate convergence rate (He, Shao, 2000);
- estimates are asymptotically normal with the same asymptotic variance as that knowing the true model

Back to "Simple and Adaptive Lasso Penalized QR"



Oracle Properties for Adaptive Lasso QR

In the linear model, let $Y = X\beta + \varepsilon = X^1\beta^1 + X^2\beta^2 + \varepsilon$, where $X = (X^1, X^2)$, $X^1 \in \mathbb{R}^{n\times q}$, $X^2 \in \mathbb{R}^{n\times(p-q)}$; β^1_q are true nonzero coefficients, $\beta^2_{p-q} = 0$ are noise coefficients; $q = \|\beta\|_0$.

Also assume that $\lambda q/\sqrt{n} \to 0$ and $\lambda/\{\sqrt{q}\log(n \vee p)\} \to \infty$ and certain regularity conditions are satisfied ▸ Details

▸ Back

Oracle Properties for Adaptive Lasso QR

Then the adaptive L_1 QR estimator has the oracle properties (Zheng et al., 2013):

1. Variable selection consistency:

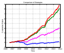
$$\mathrm{P}(\beta^2 = 0) \geq 1 - 6\exp\left\{-\frac{\log(n \vee p)}{4}\right\}.$$

2. Estimation consistency: $\|\beta - \hat{\beta}\| = \mathcal{O}_p(\sqrt{q/n})$

3. Asymptotic normality: $u_q^2 \stackrel{\text{def}}{=} \alpha^\top \Sigma_{11}\alpha$, $\forall \alpha \in \mathbb{R}^q$, $\|\alpha\| < \infty$,

$$n^{1/2} u_q^{-1} \alpha^\top (\beta^1 - \hat{\beta}^1) \xrightarrow{\mathcal{L}} \mathrm{N}\left\{0, \frac{(1-\tau)\tau}{f^2(\gamma^*)}\right\}$$

where γ^* is the τth quantile and f is the pdf of ε Back

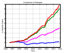


Selected Hedge Funds' Strategies

1. *Convertible arbitrage* hedge funds focus on the mispricing of convertible bonds. A typical position involves a long position in the convertible bond and a short position in the underlying asset.

2. *Fixed income arbitrage* hedge funds tend to profit from price anomalies between related securities and/or bet on the evolution of interest rates spreads. Typical trading strategies are butterfly-like structures, cash/futures basis trading strategies or relative swap spread trades.

3. *Event-driven* hedge funds focus on price movements generated by an anticipated corporate event, such as a merger, an acquisition, a bankruptcy, etc. Return

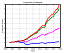


Selected Hedge Funds' Strategies

4. *Long/short equity* hedge funds represent the original hedge fund model. They invest in equities both on the long and the short sides, and generally have a small net long exposure. They are genuinely opportunistic strategies and could be classified as "double alpha, low beta" funds.

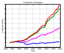
5. *Market neutral* hedge funds seek to neutralize certain market risks by taking offsetting long and short positions in instruments with actual or theoretical relationships. Most of them are in fact long/short equity hedge funds.

6. *Dedicated short bias* hedge funds are essentially long/short equity hedge funds, that maintain a consistent net short exposure, therefore attempting to capture profits when the market declines. Return



Selected Hedge Funds' Strategies

7. *Emerging market* hedge funds invest in equities and fixed-income securities of emerging markets around the world.
8. *Global macro* hedge funds take very large directional bets on overall market directions that reflect their forecasts of major economic trends and/or events.
9. *Managed futures* hedge funds implement discretionary or systematic trading in listed financial, commodity and currency futures around the world. The managers of these funds are known as commodity trading advisors (CTAs).
10. *Multi-strategy* hedge funds regroup managers acting in several of the above-mentioned strategies. Return



Traditional Assets/Hedge Fund Indices

Table 6: Correlation statistics for MSCI and hedge funds' indices returns

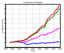
Hedge Fund Indices	MSCI Indices								
	WRD	EUR	US	UK	FR	SW	GER	JAP	PAC
Asia CTA	−0.01	0.02	−0.02	−0.06	0.01	−0.09	0.04	−0.03	0.02
Asia Distressed Debt	0.30	0.30	0.24	0.31	0.31	0.26	0.27	0.26	0.34
Asia Macro	−0.01	−0.01	−0.04	0.01	−0.02	0.07	−0.03	0.06	0.06
Global CTA FoF	0.02	0.08	−0.08	0.09	0.10	0.09	0.07	0.06	0.10
Global Event Driven FoF	0.65	0.59	0.58	0.66	0.59	0.50	0.57	0.47	0.67
Global Macro FoF	0.19	0.22	0.07	0.24	0.22	0.18	0.20	0.23	0.31
CTA/Managed Futures	−0.04	0.02	−0.13	0.03	0.03	0.07	−0.01	0.04	0.05
Event Driven	0.82	0.75	0.75	0.78	0.75	0.64	0.75	0.62	0.83
Fixed Income	0.70	0.65	0.63	0.70	0.65	0.56	0.62	0.51	0.78
Long Short Equities	0.82	0.78	0.74	0.76	0.77	0.64	0.77	0.64	0.82
Asia inc Japan Distr. Debt	0.30	0.30	0.24	0.31	0.31	0.26	0.27	0.26	0.34
Asia inc Japan Macro	0.34	0.33	0.31	0.27	0.33	0.24	0.35	0.31	0.40

Calculations based on monthly data Jan. 2000 - Jul. 2012

WRD - World, EUR - Eurozone, FR - France, SW - Switzerland, PAC - Pacific ex. Japan

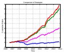
FoF means "fund of funds"

Back



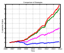
The Orthogonal GARCH Model

- Y_t is a time-dependent matrix of asset returns, $\Gamma_t = B_t \in \mathbb{R}^{p \times p}$ is the matrix of standardized eigenvectors of $\frac{1}{n} Y_t^\top Y_t$ ordered according to decreasing magnitude of eigenvalues
- $F_t = P_t \stackrel{\text{def}}{=} Y_t \Gamma_t$ is the matrix of principal components of Y_t
- retaining only the first k most important factors f and introducing noise terms u_i gives
  $y_j = b_{j1} f_1 + b_{j2} f_2 + \ldots + b_{jk} f_k + u_i$ or $Y_t = F_t B_t^\top + U_t$
- then $\Sigma_t = \mathrm{Var}(Y_t) = \mathrm{Var}(F_t B_t^\top) + \mathrm{Var}(U_t) = B_t \Delta_t B_t^\top + \Omega_t$, where $\Delta_t = \mathrm{Var}(F_t)$ is a diagonal matrix of principal component variances at t: can be separately modeled by univariate GARCH processes Back



The Orthogonal GARCH Model - continued

- B_t does not change much from day to day and can be approximated by B_{t-1} without introducing large errors in the calculation of the covariance matrix;
- Ω_t assumed to be constant and diagonal: it may be calculated from residuals $E_t = Y_t - F_t B_t^\top$, where each ω_j^2 on the diagonal is equal to $\omega_j^2 = \frac{1}{n}\sum_{i=1}^{n}(y_{ij} - f_i^\top \tilde{b}_j)^2$ with $\tilde{B} = B^\top$;
- the rule how to choose k can be based on the "proportion of total variation" explained by the first k principal components, which is calculated as the ratio of the sum of the first k eigenvalues of the matrix $\frac{1}{n} Y_t^\top Y_t$ to the sum of all p eigenvalues of this matrix Back

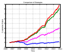


The Dynamic Conditional Correlations Model

The DCC (1,1) model separately estimates a series of univariate GARCH models and their correlation: $r_t|\mathcal{F}_{t-1} \sim \mathrm{N}(0, D_t R_t D_t)$, where

$$
\begin{aligned}
D_t^2 &= \mathrm{diag}(\omega_i) + \mathrm{diag}(\alpha_i) \odot r_{t-1} r_{t-1}^\top + \mathrm{diag}(\beta_i) \odot D_{t-1}^2, \\
\varepsilon_t &= D_t^{-1} r_t, \\
Q_t &= S \odot (\imath\imath^\top - A - B) + A \odot \{P_{t-1}\varepsilon_{t-1}\varepsilon_{t-1}^\top P_{t-1}\} + B \odot Q_{t-1}, \\
R_t &= \{\mathrm{diag}(Q_t)\}^{-1} Q_t \{\mathrm{diag}(Q_t)\}^{-1}
\end{aligned}
$$

where r_t is an $d \times 1$ vector of returns t, D_t is an $d \times d$ diagonal matrix of standard deviations σ_{it}, $i = 1, \ldots, d$, modeled by univariate GARCH, ε_t is an $d \times 1$ vector of standardized returns with $\varepsilon_{it} \stackrel{\mathrm{def}}{=} r_{it}\sigma_{it}^{-1}$, $\imath$ is a vector of ones; $P_{t-1} \stackrel{\mathrm{def}}{=} \{\mathrm{diag}(Q_t)\}^{1/2}$

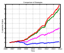


The DCC Model - Continued

- the correlation targeting gives $S = (1/T)\sum_{t=1}^{T} \varepsilon_t \varepsilon_t^\top$
- then provided that $Q_0 = \varepsilon_0 \varepsilon_0^\top$ is positive definite, each subsequent Q_t will also be positive definite
- the procedure will yield consistent but inefficient estimates of the parameters: the log-likelihood function

$$L(\theta, \phi) = -\frac{1}{2} \sum_{t=1}^{T} \left(n \log(2\pi) + 2 \log |D_t| + \log |R_t| + \varepsilon_t^\top R_t^{-1} \varepsilon_t \right),$$

where θ denotes the parameters in D and ϕ denotes additional correlation parameters in R, is maximized by parts



The DCC Model - Continued

The log-likelihood is rewritten:

$$L(\theta, \phi) = L_V(\theta) + L_C(\theta, \phi),$$

where the volatility part is the sum of individual GARCH likelihoods jointly maximized by separately maximizing each term

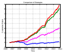
$$L_V(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left(n\log(2\pi) + \log|D_t|^2 + r_t^\top D_t^{-2} r_t\right)$$

$$= -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{d}\left(\log(2\pi) + \log(\sigma_{it}^2) + \frac{r_{it}^2}{\sigma_{it}^2}\right),$$

and the correlation part is

$$L_C(\theta, \phi) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log|R_t| + \varepsilon_t^\top R_t^{-1}\varepsilon_t - \varepsilon_t^\top \varepsilon_t\right).$$

Back



Multivariate Time-Changing Distribution Model

- motivation: $r_t|\mathcal{F}_{t-1} \sim \mathrm{N}(0, D_t R_t D_t)$ is not realistic
- fit a multivariate distribution capturing tail dependence and skewness
- model:

$$
\begin{aligned}
r_t &= \mu_t(\theta) + \varepsilon_t, \\
\varepsilon_t &= \Sigma_t^{1/2}(\theta) z_t, \\
z_t &\sim SK(\eta),
\end{aligned}
$$

where $SK(\eta)$ is the multivariate skewed-t distribution, θ vector of conditional mean and variance parameters, η vector of distribution shape parameters

Multivariate Time-Changing Distribution

The density of the innovations z_t is given by

$$g(z_t|\eta) = \prod_{i=1}^{p} \frac{2b_i}{\xi_i + \frac{1}{\xi_i}} \frac{\Gamma(\frac{\nu_i+1}{2})}{\sqrt{\pi(\nu_i - 2)}\Gamma(\frac{\nu_i}{2})} \left(1 + \frac{\kappa_{i,t}^2}{\nu_i - 2}\right)^{-\frac{\nu_i+1}{2}}$$

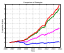
where

$$\begin{cases} (b_i z_{i,t} + a_i)\xi_i & \text{if } z_{i,t} \leq -a_i/b_i, \\ (b_i z_{i,t} + a_i)/\xi_i & \text{if } z_{i,t} > -a_i/b_i, \end{cases}$$

with

$$a_i = \frac{\Gamma(\frac{\nu_i - 1}{2})\sqrt{\nu_i - 2}}{\sqrt{\pi}\Gamma(\frac{\nu_i}{2})}\left(\xi_i - \frac{1}{\xi_i}\right), \quad b_i^2 = \xi_i^2 + \frac{1}{\xi_i^2} - 1 - a_i^2.$$

and shape parameters $\eta = (\nu_1, \ldots, \nu_p; \xi_1, \ldots, \xi_p)$, $2 < \nu_i < \infty$, $\xi_i > 0$



Multivariate Time-Changing Distribution

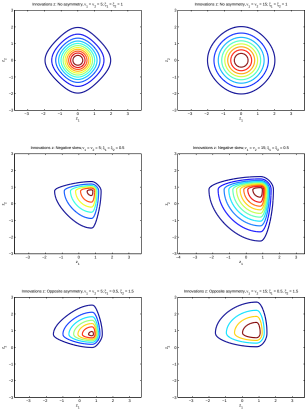
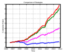


Figure 22: Contour plots of the skewed t distribution in the bivariate case



Multivariate Time-Changing Distribution

The rth moment of $z_{i,t} \sim SK(\nu_i, \xi_i)$ is defined as

$$M_{i,r} = m_{i,r} \frac{\xi_i^{r+1} + \frac{(-1)^r}{\xi_i^{r+1}}}{\xi_i + \frac{1}{\xi_i}}$$

with

$$m_{i,r} = \frac{\Gamma(\frac{\nu_i - r}{2})\Gamma(\frac{r+1}{2})(\nu_i - 2)^{\frac{r+1}{2}}}{\sqrt{\pi(\nu_i - 2)}\Gamma(\frac{\nu_i}{2})}$$

The skewness and kurtosis of $z_{i,t}$ are given by

$$sk_i^z = M_{i,3} - 3M_{i,1}M_{i,2} + 2M_{i,1}^3,$$
$$kurt_i^z = M_{i,4} - 4M_{i,1}M_{i,3} + 6M_{i,2}M_{i,1}^2 - 3M_{i,1}^4$$

Multivariate Time-Changing Distribution

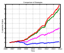
Given the properties of z_t,

$$S^{\varepsilon}_{ijk,t} = \sum_{r=1}^{p} \delta_{ir,t}\delta_{jr,t}\delta_{kr,t} sk^z_r$$

and

$$K^{\varepsilon}_{ijkl,t} = \sum_{r=1}^{p} \delta_{ir,t}\delta_{jr,t}\delta_{kr,t}\delta_{lr,t} kurt^z_r + \sum_{r=1}^{p}\sum_{s\neq r} \psi_{rs,t},$$

where $\psi_{rs,t} = \delta_{ir,t}\delta_{jr,t}\delta_{ks,t}\delta_{ls,t} + \delta_{ir,t}\delta_{js,t}\delta_{kr,t}\delta_{ls,t} + \delta_{is,t}\delta_{jr,t}\delta_{kr,t}\delta_{ls,t}$, $\Sigma_t^{1/2} = (\delta_{ij,t})_{i,j=1,\dots,p}$

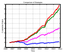


Multivariate Time-Changing Distribution

Generalize the model above assuming $z_t \sim SK(\eta_t)$: $\nu_{i,t}$, ξ_{it} follow asymmetric GARCH-like dynamics

$$\begin{aligned}
\ln(\nu_{i,t} - \underline{\nu}) &= c_{i,0} + c_{i,1}^{-}|z_{i,t-1}|N_{i,t-1} + c_{i,1}^{+}|z_{i,t-1}|P_{i,t-1} \\
&\quad + c_{i,2}\ln(\nu_{i,t-1} - \underline{\nu}) \\
\ln(\xi_{i,t}) &= d_{i,0} + d_{i,1}^{-} z_{i,t-1} N_{i,t-1} + d_{i,1}^{+} z_{i,t-1} P_{i,t-1} + d_{i,2}\ln(\xi_{i,t-1}),
\end{aligned}$$

where $N_{i,t} = \mathbf{I}(z_{i,t} \leq 0)$, $P_{i,t} = 1 - N_{i,t}$ Back

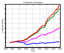


Multivariate Time-Changing Distribution

The skewed t distribution for the multivariate DCC model yields maximum likelihood in one step

$$\ln(r_1, \ldots, r_T | \theta, \eta) = \sum_{t=1}^{T} \left[\ln \left\{ g(\Sigma_t(\theta)^{-1/2}(r_t - \mu_t(\theta)) | \eta) \right\} - 0.5 \ln |\Sigma_t(\theta)| \right]$$

where $\theta = (\theta_1, \ldots, \theta_p, a, b)^\top$ with
$\theta_i = (\mu_i, \phi_{i,1}, \ldots, \phi_{i,p}, \omega_i, \alpha_i, \beta_i)^\top$, $\eta = (\eta_1, \ldots, \eta_p)^\top$, with
$\eta_i = (c_{i,0}, c_{i,1}^-, c_{i,1}^+, c_{i,2}, d_{i,0}, d_{i,1}^-, d_{i,1}^+, d_{i,2})^\top$



Portfolio Skewness and Kurtosis

Skewness S_p and excess kurtosis K_p are given by moment expressions

$$S_p(w) = \frac{1}{\sigma_p^3(w)}(m_3 - 3m_2 m_1 + 2m_1^3)$$

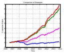
$$K_p(w) = \frac{1}{\sigma_p^4(w)}(m_4 - 4m_3 m_1 + 6m_2 m_1^2 + 3m_1^4) - 3$$

where portfolio non-central moments also depend on w:

$$m_1 = \mu_p(w) \stackrel{\text{def}}{=} w^\top \mu$$

$$m_2 = \sigma_p^2 + m_1^2$$

$$m_3 = \sum_{i=1}^{d} \sum_{j=1}^{d} \sum_{k=1}^{d} w_i w_j w_k S_{ijk}$$



Portfolio Skewness and Kurtosis - Continued

$$m_4 = \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d} w_i w_j w_k w_l K_{ijkl},$$

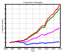
where $\sigma_p^2(w) = w^\top \Sigma w$ and $S_{ijk} = \mathrm{E}(r_i \times r_j \times r_k)$, $K_{ijkl} = \mathrm{E}(r_i \times r_j \times r_k \times r_l)$ can be computed via sample averages from returns data.

S_{ijk}, K_{ijkl} determine the d-dimensional portfolio co-skewness and co-kurtosis tensors

$$S \overset{\text{def}}{=} \{S_{ijk}\}_{i,j,k=1,\dots,d} \in \mathbb{R}^{d\times d\times d}$$

$$K \overset{\text{def}}{=} \{K_{ijkl}\}_{i,j,k,l=1,\dots,d} \in \mathbb{R}^{d\times d\times d\times d}.$$

Back



Regularity Conditions for Adaptive Lasso QR

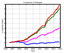
A1 Sampling and smoothness: $\forall x$ in the support of X_i, $\forall y \in \mathbb{R}$, $f_{Y_i|X_i}(y|x)$, $f \in \mathcal{C}^k(\mathbb{R})$, $|f_{Y_i|X_i}(y|x)| < \overline{f}$, $|f'_{Y_i|X_i}(y|x)| < \overline{f'}$; $\exists \underline{f}$, such that $f_{Y_i|X_i}(x^\top \beta_\tau | x) > \underline{f} > 0$

A2 Restricted identifiability and nonlinearity: let $\delta \in \mathbb{R}^p$, $T \subset \{0, 1, ..., p\}$, δ_T such that $\delta_{Tj} = \delta_j$ if $j \in T$, $\delta_{Tj} = 0$ if $j \notin T$; $T = \{0, 1, ..., s\}$, $\overline{T}(\delta, m) \subset \{0, 1, ..., p\} \setminus T$, then $\exists m \geq 0$, $c \geq 0$ such that

$$\inf_{\delta \in A, \delta \neq 0} \frac{\delta^\mathrm{T} \mathrm{E}(X_i X_i^\top)\delta}{\|\delta_{T \cup \overline{T}(\delta, m)}\|^2} > 0, \qquad \frac{3\underline{f}^{3/2}}{8\overline{f}'} \inf_{\delta \in A, \delta \neq 0} \frac{\mathrm{E}[|X_i^\mathrm{T}\delta|^2]^{3/2}}{\mathrm{E}[|X_i^\mathrm{T}\delta|^3]} > 0,$$

where $A \stackrel{\text{def}}{=} \{\delta \in \mathbb{R}^p : \|\delta_{T^c}\|_1 \leq c\|\delta_T\|_1, \|\delta_{T^c}\|_0 \leq n\}$

▸ Back



Regularity Conditions - Continued

A3 Growth rate of covariates:

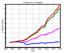
$$\frac{q^3\{\log(n \vee p)\}^{2+\eta}}{n} \to 0, \eta > 0$$

A4 Moments of covariates: Cramér condition

$$\mathsf{E}[|x_{ij}|^k] \leq 0.5 C_m M^{k-2} k!$$

for some constants C_m, M, $\forall k \geq 2$, $j = 1, ..., p$

A5 Well-separated regression coefficients: $\exists b_0 > 0$, such that $\forall j \leq q$, $|\hat{\beta}_j| > b_0$



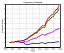
The GJR (1,1) model

The Glosten-Jagannathan-Runkle (GJR-GARCH) (1,1) model with mean equation specified as autoregressive moving average (ARMA) (1,1), may be represented as follows

$$
\begin{aligned}
r_t &= c + \phi_1 r_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t, \\
\varepsilon_t &= \sigma_t z_t, \qquad z_t \sim \mathrm{N}(0,1), \text{i.i.d.}, \\
\sigma_t^2 &= \omega + \beta \sigma_{t-1}^2 + \gamma \varepsilon_{t-1}^2 + \xi \mathbf{I}(\varepsilon_{t-1} < 0) \varepsilon_{t-1}^2
\end{aligned}
$$

where c, ϕ_1, θ_1, ω, β, γ, ξ are model parameters, $\omega > 0$, $\beta \geq 0$, $\gamma \geq 0$, $\gamma + \xi \geq 0$, $\beta + \gamma + \xi < 1$, $\mathbf{I}(\cdot)$ is the indicator function.

Back



References

Peter Buehlmann, Sara van de Geer
Statistics for High-Dimensional Data: Methods, Theory and Applications
Springer, 2011

Dowd, K.
Measuring Market Risk. 2nd Edition
Wiley, 2005

Favre, L., Galeano, J.-A.
Mean-Modified Value-at-Risk Optimization with Hedge Funds
The Journal of Alternative Investments, Vol. 5, No. 2 (2002)

References

Lhabitant, François-Serge
Hedge Funds. Myths and Limits
Wiley, 2002

Roger Koenker, Gilbert Bassett, Jr
Regression Quantiles
Econometrica, Vol. 46, No. 1 (January, 1978)

Youjuan Li, Ji Zhu
L_1-norm Quantile Regression
Journal of Computational and Graphical Statistics, Vol. 17, No. 1, pp. 1-23

References

Markowitz, Harry

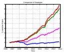
Portfolio Selection

Journal of Finance, Vol. 7, No. 1 (Mar., 1952), pp. 77-91

Robert Tibshirani

Regression Shrinkage and Selection via the Lasso

J. R. Statist. Soc (1996), No.1 pp. 267-288



References

Qi Zheng, Colin Gallagher, K.B. Kulasekera

Adaptive Penalized Quantile Regression for High-Dimensional Data

Journal of Statistical Planning and Inference, 143 (2013) 1029-1038

Hui Zou

The Adaptive Lasso and Its Oracle Properties

Journal of the American Statistical Association, Dec., 2006, Vol. 101, No. 476