TEDAS - Tail Event Driven ASset allocation

Wolfgang Karl Härdle

Sergey Nasekin

Alla Petukhina

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

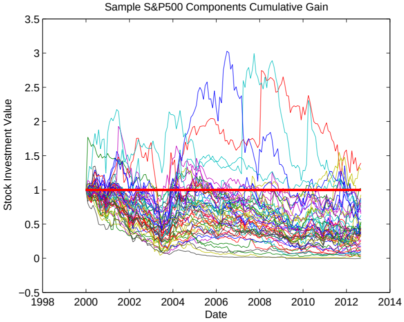
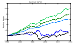


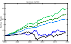
Figure 1: 50 random S&P 500 Sample Components' Cumulative Return: **94%** of stocks lost the value of the initial investment (thick red line)



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

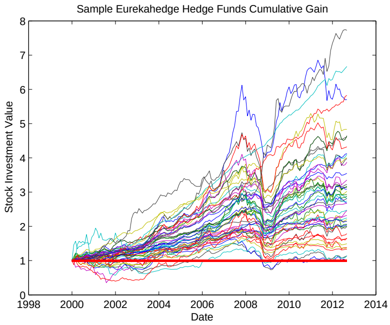


Figure 2: 50 Eurekahedge Hedge Funds Indices' Cumulative Return: **0%** of funds lost the value of the initial investment (thick red line)

Diversification

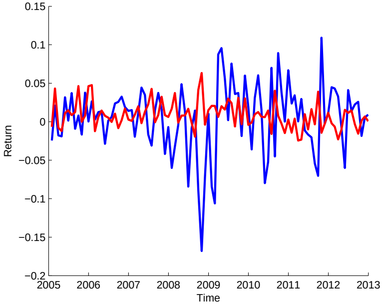


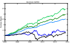
Figure 3: S&P 500 and Eurekahedge North America Macro Hedge Fund Index monthly returns in 20050131-20121231

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 1: 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



Tail Risk

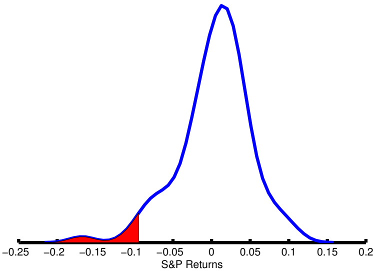


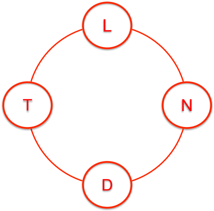
Figure 4: Estimated density of S&P 500 returns

The TLND challenge

- **T**ail dependence of assets
- **L**arge universe of financial assets: $p > n$
- **N**on normality
- **D**ynamic nature of distributions

TEDAS Objectives

- Hedge tail events
  - Quantile regression
  - Variable selection in high dimensions
- Improve Asset Allocation
  - Higher-order moments' optimization
  - Modelling of moments' dynamics



Outline

1. Motivation ✓
2. TEDAS framework
3. Empirical Application
4. Conclusions
5. Technical Details

Tail Events

- $Y \in \mathbb{R}^n$ core returns; $X \in \mathbb{R}^{n \times p}$ satellite assets' returns, $p > n$

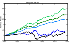
-

$$q_\tau(x) \stackrel{\text{def}}{=} F^{-1}_{y|x}(\tau) = x^\top \beta(\tau) = \arg \min_{\beta \in \mathbb{R}^p} \mathsf{E}_{Y|X=x}\, \rho_\tau\{Y - X\beta\},$$

$$\rho_\tau(u) = u\{\tau - \mathbf{I}(u < 0)\}$$

- L_1 penalty $\lambda_n \|\hat{\omega}^\top \beta\|_1$ to nullify "excessive" coefficients; λ_n and $\hat{\omega}$ controlling penalization; constraining $\beta \leq 0$ yields ALQR ▸ Details

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

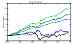


TEDAS Step 1

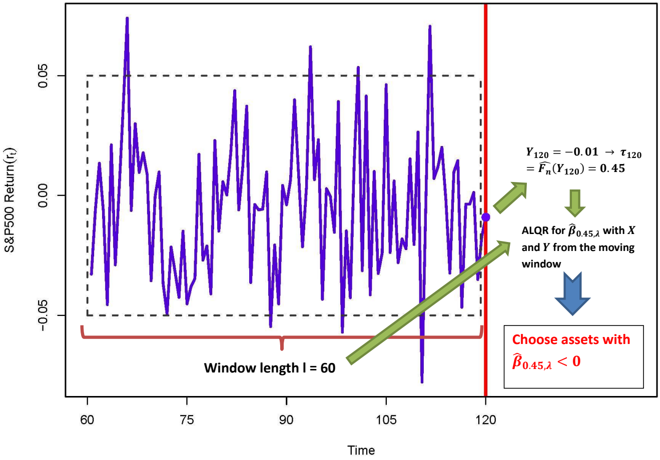
Initial wealth $W_0 = \$1$, at $t = l, \dots, n$; $l = 60$ length of the moving window

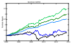
- Portfolio constituents' selection

  1. determine core asset return Y_t, set $\tau_t = \widehat{F}_n(Y_t)$ ▸ Notation
  2. ALQR for $\hat{\beta}_{\tau_t,\lambda_n}$ using the observations $X \in \mathbb{R}^{t-l+1,\dots,t \times p}$, $Y \in \mathbb{R}^{t-l+1,\dots,t}$
  3. if $Y_t < 0$, choose X_j, $j = 1, \dots, p$ with $\hat{\beta}_{\tau_t,\lambda_n} < 0$; if $Y_t > 0$, choose X_j, $j = 1, \dots, p$ with $\hat{\beta}_{\tau_t,\lambda_n} > 0$



TEDAS Step 1

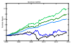

S&P500 Return(r_t)
0.05
0.00
−0.05
60
75
90
105
120
Time
Window length l = 60
$Y_{120} = -0.01 \rightarrow \tau_{120}$
$= \widehat{F_n}(Y_{120}) = 0.45$
ALQR for $\widehat{\beta}_{0.45,\lambda}$ with X and Y from the moving window
Choose assets with $\widehat{\beta}_{0.45,\lambda} < 0$



TEDAS Step 2

- Portfolio allocation

    1. apply a one of TEDAS Gestalt to satellite assets X_j to determine $\widehat{w}_t$
    2. determine the realized portfolio wealth for $t+1$:
       $W_{t+1} = W_t(1 + \widehat{w}_t^\top X_{t+1})$

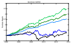


TEDAS Gestalten

TEDAS gestalt	Dynamics modelling	Weights optimization
TEDAS Naïve	NO	Equal weights
TEDAS Hybrid	NO	Mean-variance optimization of weights Details
TEDAS Basic	DCC volatility Details	CF-VaR optimization Details
TEDAS Advanced	Time-Varying Details	Cornish-Fisher-CVaR minimization Details
TEDAS Expert	Conditional Distributions	Expected utility optimization Details

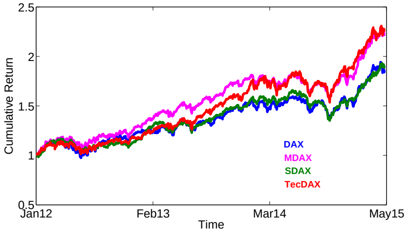
Hedge funds' data

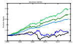
- Monthly data
  - Core asset (Y): S&P 500, Nikkei225, DAX 30, FTSE 100
  - Satellite assets (X): 164 Eurekahedge hedge funds indices
- Span: 20000131-20131031 (166 months)
- Source: Bloomberg



German stocks' data

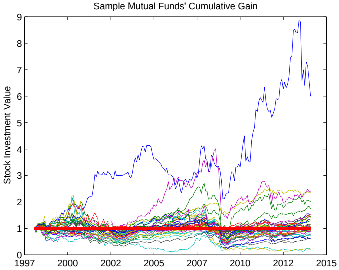
- Frankfurt Stock Exchange (Xetra), weekly data
  - Core asset (Y): DAX index
  - Satellites assets (X): 125 stocks - SDAX (48), MDAX (47) and TecDAX (50) as on 20140801
- Span: 20121221 - 20141127
- Source: Datastream

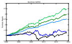




Mutual Funds' Data

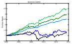
- Monthly data
  - Core asset (Y): S&P500
  - Satellite assets (X): 583 Mutual funds
- Span: 19980101 - 20131201
- Source: Datastream





Benchmark Strategies

1. RR: dynamic risk-return optimization ▸ Details
2. PESS: tail risk optimization ▸ Details
3. Risk-parity portfolio (equal risk contribution) ▸ Details
4. 60/40 portfolio ▸ Details



TEDAS with Y= S&P 500

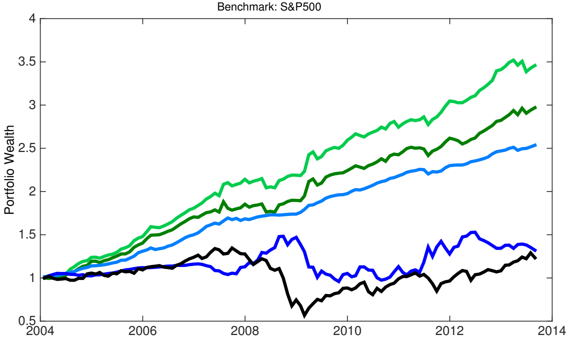
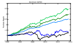


Figure 5: Cumulative portfolio wealth comparison: TEDAS Expert, TEDAS Advanced, RR, PESS, S&P 500 buy & hold; X = hedge funds' indices' returns matrix



TEDAS with Y= Nikkei 225

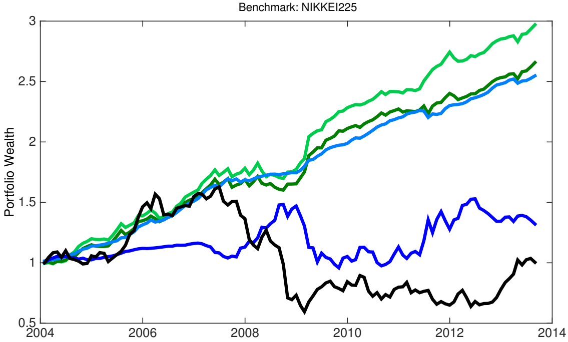
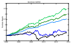


Figure 6: Cumulative portfolio wealth comparison: TEDAS Expert, TEDAS Advanced, RR, PESS, Nikkei 225 buy & hold; X = hedge funds' indices' returns matrix



TEDAS with Y= FTSE 100

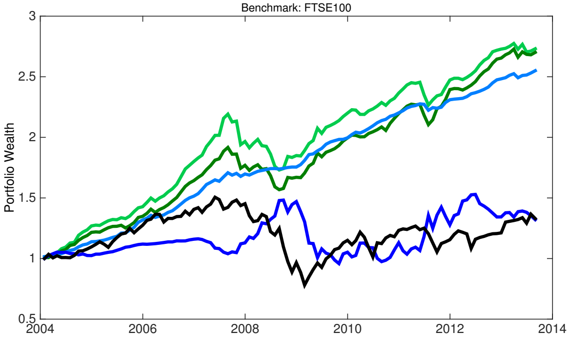


Figure 7: Cumulative portfolio wealth comparison: TEDAS Expert, TEDAS Advanced, RR, PESS, FTSE100 buy & hold; X = hedge funds' indices' returns matrix

TEDAS with Y= DAX30

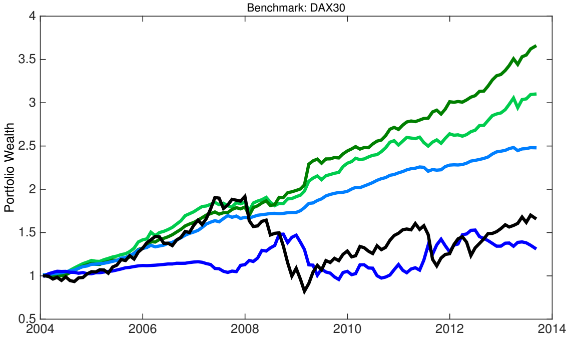


Figure 8: Cumulative portfolio wealth comparison: TEDAS Expert, TEDAS Advanced, RR, PESS, DAX30 buy & hold; X = hedge funds' indices' returns matrix

Histograms of $\hat{q}$

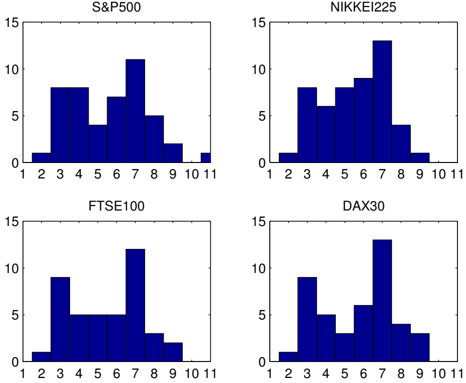
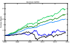


Figure 9: Frequency of the number of selected variables for 4 different Y

Selected Hedge Funds: S&P 500

Table 2: The selected hedge funds for S&P 500 benchmark

Top 5 influential hedge funds	Frequency
Latin American Onshore Fixed Income Hedge Fund Index	19
Emerging Markets Dual Approach Absolute Return Fund Index	18
Large North American Hedge Fund Index	15
Europe Macro Hedge Fund Index	15
Asia Macro Hedge Fund Index	14



Selected Hedge Funds: Nikkei 225

Table 3: The selected hedge funds for Nikkei 225 benchmark

Top 5 influential hedge funds	Frequency
Small Latin American Hedge Fund Index	22
Taiwan Hedge Fund Index	20
North America Top-Down Absolute Return Fund Index	16
Large North American Hedge Fund Index	15
Latin American Onshore Fixed Income Hedge Fund Index	15

Selected Hedge Funds: FTSE 100

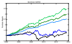
Table 4: The selected hedge funds for FTSE 100 benchmark

Top 5 influential hedge funds	Frequency
Small Latin American Hedge Fund Index	15
Europe Macro Hedge Fund Index	15
Latin American Onshore Fixed Income Hedge Fund Index	14
Asia Pacific Top-Down Absolute Return Fund Index	14
Latin American Fixed Income Hedge Fund Index	13

Selected Hedge Funds

Table 5: The selected hedge funds for DAX 30 benchmark

Top 5 influential hedge funds	Frequency
Emerging Markets Dual Approach Absolute Return Fund Index	21
North America Macro Hedge Fund Index	19
Taiwan Hedge Fund Index	16
Asia CTA Hedge Fund Index	14
Europe Macro Hedge Fund Index	14



Evolution of ξ_t and ν_t

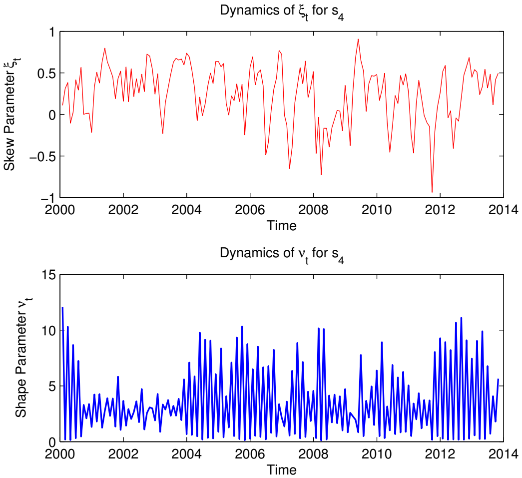
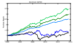


Figure 10: Evolution of the skew and shape parameters for s_4 in Table **??**



Conditional Skewness and Kurtosis

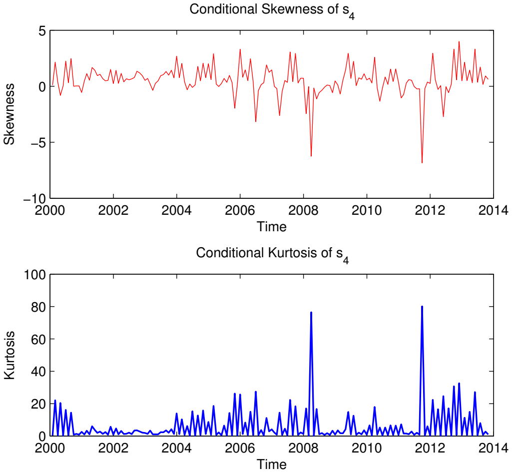


Figure 11: Evolution of the conditional skewness and kurtosis for s_4

TEDAS: DAX results

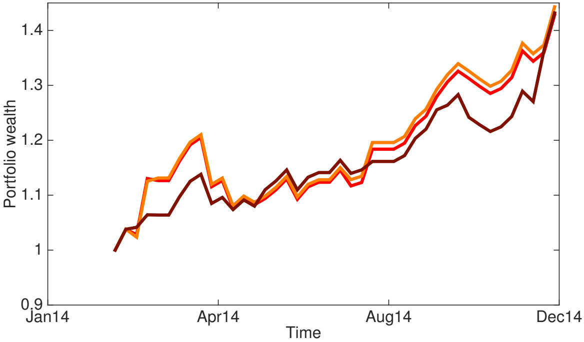


Figure 12: Strategies' cumulative returns' comparison (with transaction costs 1% of portfolio value): TEDAS Naïve, TEDAS Hybrid, TEDAS basic

TEDAS: DAX results

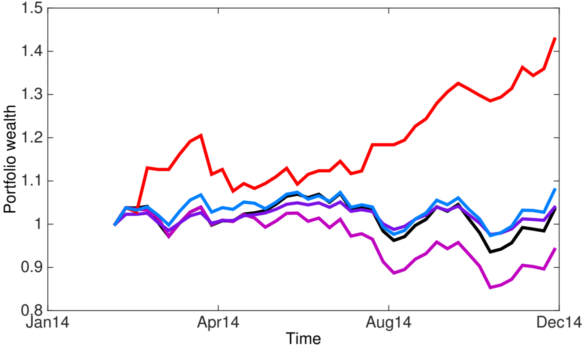
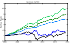


Figure 13: Strategies' cumulative returns' comparison (with transaction costs 1% of portfolio value): TEDAS Hybrid, DAX Buy-and-hold, 60/40, Risk-parity, RR



TEDAS: Mutual Funds' results

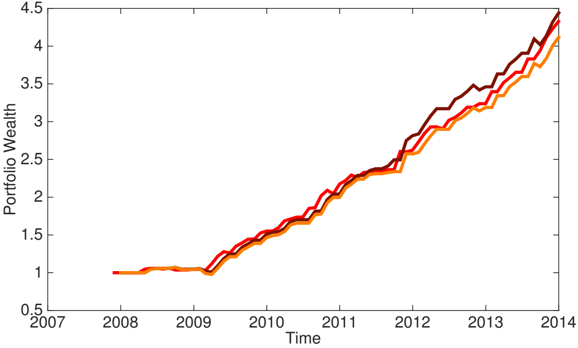
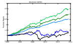


Figure 14: Strategies' cumulative returns' comparison (with transaction costs 1% of portfolio value): TEDAS Naïve, TEDAS Hybrid, TEDAS basic

TEDAS_strategie



TEDAS: Mutual Funds' results

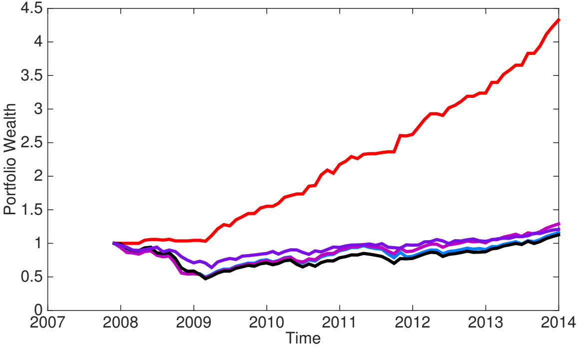
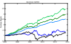


Figure 15: Strategies' cumulative returns' comparison (with transaction costs 1% of portfolio value): TEDAS Hybrid, DAX Buy-and-hold, 60/40, Risk-parity, RR

TEDAS_strategie



Conclusions

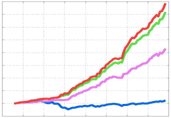
- Hedge tail events
  - Quantile regression for tail dependence estimation
  - ALQR for asstes' univerce's dimensionality reduction
- Improve Asset Allocation
  - Dynamic distribution structure of the portfolio
  - VAR, CVAR and utility higher-order moments' optimization
  - Out-of-sample performance gain

TEDAS - Tail Event Driven Asset Allocation

Wolfgang Karl Härdle

Sergey Nasekin

Alla Petukhina



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

Notation

- $\hat{q}_\tau \stackrel{\text{def}}{=} \widehat{F}_n^{-1}(\tau)$, where the log-returns edf is

$$\widehat{F}_n(r) \stackrel{\text{def}}{=} \int_{-\infty}^{r} \hat{f}_n(u)\, du = \frac{1}{n}\sum_{i=1}^{n} H\left(\frac{r-r_i}{h}\right), \tag{2}$$

  where $\hat{f}_n(r) \stackrel{\text{def}}{=} (1/nh)\sum_{i=1}^{n} K\{(r-r_i)/h\}$, $H(x) = \int_{-\infty}^{x} K(u)du$, $K(\cdot)$ is a normal kernel function, h bandwidth

- $\hat{\beta}_{\tau,\lambda_n}$ are the estimated non-zero ALQR coefficients

Back to "Back to framework"

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 $\{X_i; i = 1, ..., n\}$

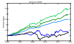
The optimization problem for the lasso estimator:

$$\begin{aligned} \hat{\beta}^{\text{lasso}} = \quad & \arg\min_{\beta \in \mathbb{R}^p} f(\beta) \\ & \text{subject to} \quad g(\beta) \geq 0 \end{aligned} \tag{3}$$

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 "Tail Events"



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 "Tail Events"

Paths of Lasso Coefficients

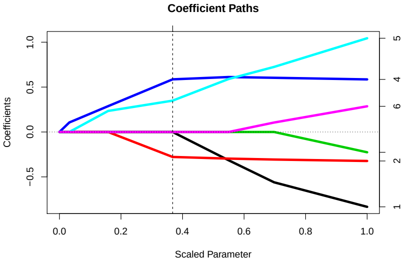
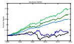


Figure 16: Lasso shrinkage of coefficients in the hedge funds dataset example (6 covariates were chosen for illustration); each curve represents a coefficient as a function of the scaled parameter $\hat{s} = t/\|\beta\|_1$; the dashed line represents the model selected by the BIC information criterion ($\hat{s} = 3.7$)

Back to "Tail Events"



Example of Lasso Geometry

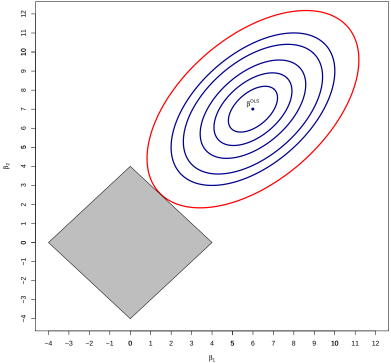


Figure 17: Contour plot of the residual sum of squares objective function centered at the OLS estimate $\widehat{\beta}^{ols} = (6, 7)$ and the constraint region $\sum |\beta_j| \leq t$ MVAlassocontour

Quantile Regression

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

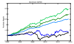
Minimize

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

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 "Tail Events"

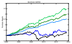


Non-Positive (NP) Lasso-Penalized QR

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

$$
\begin{aligned}
\underset{(\xi,\zeta,\eta,\tilde{\beta})\in\mathbb{R}_+^{2n+p}\times\mathbb{R}^p}{\text{minimize}} \quad & \tau 1_n^\top \xi + (1-\tau) 1_n^\top \zeta + \lambda 1_n^\top \eta \\
\text{subject to} \quad & \xi - \zeta = Y + X\tilde{\beta}, \\
& \xi \geq 0, \\
& \zeta \geq 0, \\
& \eta \geq \tilde{\beta}, \\
& \eta \geq -\tilde{\beta}, \\
& \tilde{\beta} \geq 0, \qquad \tilde{\beta} \overset{\text{def}}{=} -\beta
\end{aligned}
\tag{4}
$$

Back to "Tail Events"



Solution

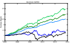
Transform into matrix (I_p is $p \times p$ identity matrix; $E_{p\times n} = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$):

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

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 "Tail Events"



Solution - Continued

The previous problem may be reformulated into *standard form*

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

and the dual problem is:

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

Back to "Tail Events"

Solution - Continued

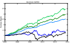
The KKT conditions for this linear program are

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

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 "Tail Events"



Adaptive Lasso Procedure

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

Lasso soft-threshold function gives biased results

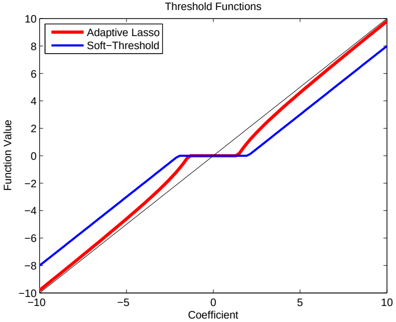
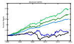


Figure 18: Threshold functions for simple and adaptive Lasso



Adaptive Lasso Procedure

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

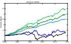
L_1 - penalty replaced by a re-weighted version; $\hat{\omega}_j = 1/|\hat{\beta}_j^{\text{init}}|^{\gamma}$, $\gamma = 1$, $\hat{\beta}^{\text{init}}$ is from (3)

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 "Tail Events"



Simple and Adaptive Lasso Penalized QR

Simple lasso-penalized QR optimization problem is:

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

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

$$\hat{\beta}_{\tau,\lambda}^{\text{adapt}} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta) + \lambda \|\hat{\omega}^\top \beta\|_1 \tag{6}$$

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

Back to "Tail Events"

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: $\tilde{X} = (X_1 \circ \hat{\beta}_1^{\text{init}}, \dots, X_p \circ \hat{\beta}_p^{\text{init}})$;
- the lasso problem (5) 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 "Tail Events"

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

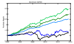
Back to "Simple and Adaptive Lasso Penalized QR"

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:

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

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

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

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

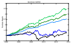
where γ^* is the τth quantile and f is the pdf of ε Back

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

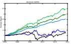


Selected Hedge Funds' Strategies

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.

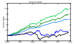
5. *Market neutral* hedge funds seek to neutralize certain market risks by taking offsetting long and short positions in instruments with actual or theoretical relationships. Most of them are in fact long/short equity hedge funds.

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

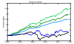
Hedge Fund Indices	MSCI Indices								
	WRD	EUR	US	UK	FR	SW	GER	JAP	PAC
Asia CTA	−0.01	0.02	−0.02	−0.06	0.01	−0.09	0.04	−0.03	0.02
Asia Distressed Debt	0.30	0.30	0.24	0.31	0.31	0.26	0.27	0.26	0.34
Asia Macro	−0.01	−0.01	−0.04	0.01	−0.02	0.07	−0.03	0.06	0.06
Global CTA FoF	0.02	0.08	−0.08	0.09	0.10	0.09	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

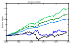


Risk-Return Asset Allocation

For random log-returns $X_t \in \mathbb{R}^p$:

$$\begin{aligned} \min_{w_t \in \mathbb{R}^p} \quad & \sigma^2_{P,t}(w_t) \stackrel{\text{def}}{=} w_t^\top \Sigma_t w_t \\ \text{s.t.} \quad & \mu_{P,t}(w_t) = r_T, \\ & w_t^\top 1_p = 1, \\ & w_{i,t} \geq 0 \end{aligned} \tag{7}$$

where $\Sigma_t \stackrel{\text{def}}{=} \mathrm{E}_{t-1}\{(X_t - \mu)(X_t - \mu)^\top\}$, Σ_t is modeled with a GARCH model ▸ Details ▸ Back to "Benchmark Strategies" ▸ Back to "Framework"

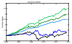


Tail Risk Asset Allocation

Given portfolio returns $X \in \mathbb{R}^{n \times p}$, Bassett et al. (2004)

$$\begin{aligned} \min_{(\beta,\alpha)^\top \in \mathbb{R}^p} \quad & \sum_{t=1}^{n} \rho_\tau \left\{ X_{t1} - \sum_{j=2}^{p} (X_{t1} - X_{tj}) \beta_j - \alpha \right\} \\ \text{s.t.} \quad & w^\top \hat{\mu} = r_T, \\ & w^\top 1_p = 1, \end{aligned} \tag{8}$$

where $w = w(\beta) = (1 - \sum_{j=2}^{p} \beta_j, \beta^\top)^\top$, $\rho_\tau(u) = u\{\tau - \mathbf{I}(u < 0)\}$, $\tau \in (0,1)$, $\hat{\mu} \stackrel{\text{def}}{=} \overline{X}$ sample returns' mean Back to "Benchmark Strategies"



Multi-Moment Utility Optimization

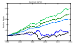
The (dynamic) investment decision: $U(\cdot)$ utility function; $X_t \in \mathbb{R}^p$ log-returns, w_t weights, $\mu_{P,t}(w_t) \stackrel{\text{def}}{=} w_t^\top \mu$, $\mu \stackrel{\text{def}}{=} \mathsf{E}_{t-1}(X_t)$, r_T "target" return:

$$\max_{w_t \in \mathbb{R}^p} \mathsf{E}_{t-1}\left\{U(W_t)\right\}, \quad \text{s.t. } \mu_{P,t}(w_t) = r_T, \; w^\top 1_p = 1, w_{i,t} \geq 0, \tag{9}$$

$$\mathsf{E}_{t-1}\left\{U(W_t)\right\} \approx U\{\overline{W}_t\} + \frac{1}{2}U^{(2)}\{\overline{W}_t\}\sigma^2_{W_t} + \\ + \frac{1}{3!}U^{(3)}\{\overline{W}_t\}S_{W_t} + \frac{1}{4!}U^{(4)}\{\overline{W}_t\}K_{W_t},$$

where $W_t \stackrel{\text{def}}{=} 1 + w_t^\top X_t$ is the end-of-period t wealth, $\overline{W}_t \stackrel{\text{def}}{=} \mathsf{E}_{t-1}(W_t)$, $\sigma^2_{W_t} \stackrel{\text{def}}{=} \mathsf{E}_{t-1}\left\{(W_t - \overline{W}_t)^2\right\}$, $S_{W_t} \stackrel{\text{def}}{=} \mathsf{E}_{t-1}\left\{(W_t - \overline{W}_t)^3\right\}$, $K_{W_t} \stackrel{\text{def}}{=} \mathsf{E}_{t-1}\left\{(W_t - \overline{W}_t)^4\right\}$; $U^{(n)}(\cdot)$ is the nth derivative of $U(\cdot)$

Back to "Strategies"



Utility Function Example

- CARA utility:

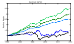
$$U(W) = -\exp(-\eta W),$$

where η coefficient of risk aversion

- then:

$$\begin{aligned} \mathsf{E}_{t-1}\left\{U(W_t)\right\} &= \mathsf{E}_{t-1}\left\{-\exp(-\eta W_t)\right\} \\ &\approx -\exp(-\eta \overline{W}_t)\left(1 + \frac{\eta^2}{2}\sigma^2_{W_t} - \frac{\eta^3}{3!}S_{W_t} + \frac{\eta^4}{4!}K_{W_t}\right) \end{aligned}$$

Back to "Strategies"



Portfolio Moments

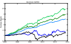
The porfolio moments:

$$\sigma^2_{W_t} = w_t^\top M_t^2 w_t$$
$$S_{W_t} = w_t^\top M_t^3 (w_t \otimes w_t)$$
$$K_{W_t} = w_t^\top M_t^4 (w_t \otimes w_t \otimes w_t),$$

where $\otimes$ Kronecker product,

$$M_t^2 \stackrel{\text{def}}{=} \mathsf{E}_{t-1}(r_t - \mu)^2 \tag{10}$$
$$M_t^3 \stackrel{\text{def}}{=} \mathsf{E}_{t-1}\{(r_t - \mu)(r_t - \mu)^\top \otimes (r_t - \mu)^\top\} \tag{11}$$
$$M_t^4 \stackrel{\text{def}}{=} \mathsf{E}_{t-1}\{(r_t - \mu)(r_t - \mu)^\top \otimes (r_t - \mu)^\top \otimes (r_t - \mu)^\top\}, \tag{12}$$

A dynamic distribution model is used to obtain M_t^2, M_t^3, M_t^4 in (10), (11), (12)



Dynamic Distribution Model

- joint normality is questionable
- possible persistence in the dynamics of moments
- reaction of distribution parameters to past shocks
- computational feasibility

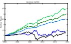
Descriptive Statistics

Table 7: Monthly returns of 3 Eurekahedge hedge funds' indices

	Japan Multi-Strategy		North America Fixed Income		Europe Arbitrage	
Univariate statistics						
Normality tests						
JB	533.775	(0.000)	294.089	(0.000)	610.407	(0.000)
KS	0.503	(0.000)	0.473	(0.000)	0.485	(0.000)
Omnibus	82.773	(0.000)	43.761	(0.000)	171.079	(0.000)
Dynamic conditional moments' tests						
ARCH	11.227	(0.000)	34.966	(0.000)	26.592	(0.000)
Bera-Lee	48.469	(0.000)	36.475	(0.000)	40.783	(0.000)
Bera-Zuo	203.723	(0.000)	20.149	(0.166)	421.847	(0.000)
Multivariate statistics						
Test						
Omnibus	326.226	(0.000)				
Mardia	301.199	(0.000)				
Henze-Zirkler	9.862	(0.000)				

Standard errors and p-values are given in parentheses.

ARCH, Bera-Lee and Bera-Zuo stand for the test statistics of the ARCH test by Engle (1982) and information matrix tests for testing variation in second, third and fourth conditional moments



Generalized Hyperbolic (GH) Distribution

A vector X has a multivariate GH distribution if

$$X = \mu + W\delta + \sqrt{W}AZ, \tag{13}$$

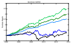
where

(i) $Z \sim \mathrm{N}(0, I_k)$

(ii) $A \in \mathbb{R}^{d \times k}$

(iii) μ, $\delta \in \mathbb{R}^d$

(iv) $W \geq 0$, scalar-valued random variable, independent of Z, $W \sim GIG(\lambda, \alpha, \beta)$; GIG is the generalized inverse Gaussian distribution



Multivariate Affine GH Distribution

- margins of the (MGH) distribution not mutually independent for some choice of $\Sigma = AA^\top$
- MAGH distribution, Schmidt et al. (2006), models margins and dependency independently

$Y \sim MAGH(\lambda, \alpha, \beta, \mu, \Sigma)$ if

(i) $X = (X_1, \ldots, X_d)^\top$, $X_i \sim GH(0, 1, \alpha_i, \beta_i)$, $i = 1, \ldots, d$

(ii) $Y = AX + \mu$, $AA^\top = \Sigma$ positive definite

Normal Inverse Gaussian (NIG) Distribution

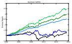
- obtained from the GH distribution with $\lambda = -0.5$
- "semi-heavy tails" property: fits financial data well

The density is written as:

$$f_{NIG}(x) = \frac{\alpha\delta}{\pi} \exp\left\{\delta\sqrt{\alpha^2 - \beta^2} + \beta(x-\mu)\right\} \frac{K_1\left\{\alpha\sqrt{\delta^2 + (x-\mu)^2}\right\}}{\sqrt{\delta^2 + (x-\mu)^2}},$$

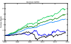
where $0 \leq |\beta| \leq \alpha$, $\delta > 0$, K_1 is the modified Bessel function of the third kind and order 1

Location-Scale Property: let $\overline{\alpha} \stackrel{\text{def}}{=} \delta\alpha$ and $\overline{\beta} \stackrel{\text{def}}{=} \delta\beta$, then
$X \sim NIG(\overline{\alpha}, \overline{\beta}, \mu, \delta) \Leftrightarrow (X-\mu)/\delta \sim NIG(\overline{\alpha}, \overline{\beta}, 0, 1)$



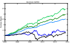
Choice of the Matrix A

- assume $X = As$, X random signal generated by another random vector $s = (s_1, \dots, s_d)$, s_i statistically independent, $i = 1, \dots, d$ and a mixing matrix A, both unknown
- the *independent component analysis* (ICA) technique separates source signals s from a set of mixed signals X without or with very little aid of information about f or the mixing process A
- ICA estimates A and s by maximizing the nongaussianity of linear combinations of X



The Model for Portfolio Returns

- assume $\varepsilon_t = As_t$, $\mathsf{E}\,\varepsilon_t = 0$, $\mathsf{E}\,\varepsilon_t\varepsilon_t^\top = I_d$, $\mathsf{E}\,s_t = 0$, $\mathsf{E}\,s_ts_t^\top = I_d$
- define $\mathsf{E}(s_t|\mathcal{F}_t) = 0$, $D_t \stackrel{\text{def}}{=} \mathsf{E}(s_ts_t^\top|\mathcal{F}_t) \stackrel{\text{def}}{=} \text{diag}(d_{1t}, \ldots, d_{dt})$
- let $z_{it} \sim NIG(\overline{\alpha}_{it}, \overline{\beta}_{it}, 0, 1)$, then $s_{it} \sim NIG(\overline{\alpha}_{it}/\sqrt{d_{it}}, \overline{\beta}_{it}/\sqrt{d_{it}}, 0, \sqrt{d_{it}})$
- *MANIG*: multivariate affine normal inverse Gaussian distribution
- model for portfolio returns $r_t = m_t + \varepsilon_t$, $r_t|\mathcal{F}_t \sim MANIG(m_t, \Sigma_t, \omega_t)$, where $\omega_t = (\omega_{1t}, \ldots, \omega_{dt})^\top$ and $\omega_{it} = (\alpha_{it}, \beta_{it})^\top$, $i = 1, \ldots, d$, $\Sigma_t = M_t^2 = AD_tA^\top$, d_{it} can be modeled as GARCH-type processes



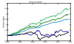
Moment Dynamics

- reparametrize the model to have asymmetry and shape parameters $\xi_{it} = \beta_{it}/\alpha_{it}$, $\nu_{it} = \sqrt{\alpha_{it}^2 - \beta_{it}^2}$
- introduce asymmetric GARCH-like dynamics:

$$\nu_{i,t} = a_{i,0} + a_{i,1}^{-}|s_{i,t-1}|N_{i,t-1} + a_{i,1}^{+}|s_{i,t-1}|P_{i,t-1} + a_{i,2}\nu_{i,t-1} \tag{14}$$

$$\xi_{i,t} = b_{i,0} + b_{i,1}^{-}s_{i,t-1}N_{i,t-1} + b_{i,1}^{+}s_{i,t-1}P_{i,t-1} + b_{i,2}\xi_{i,t-1}, \tag{15}$$

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



Portfolio Moments

$$M_t^3 = A M_{s_t}^3 (A \otimes A)^\top, \quad M_t^4 = A M_{s_t}^4 (A \otimes A \otimes A)^\top,$$

where

$$M_{s_t}^3 = \mathsf{E}_{t-1}(s_{i,t} s_{j,t} s_{k,t}) = \sum_{r=1}^{p} d_{ir,t} d_{jr,t} d_{kr,t} sk_{rt}^s$$

$$M_{s_t}^4 = \mathsf{E}_{t-1}(s_{i,t} s_{j,t} s_{k,t} s_{l,t})$$

$$= \sum_{r=1}^{p} d_{ir,t} d_{jr,t} d_{kr,t} d_{lr,t} kurt_{rt}^s + \sum_{r=1}^{p} \sum_{s \neq r} \psi_{rs,t},$$

$\psi_{rs,t} = d_{ir,t} d_{jr,t} d_{ks,t} d_{ls,t} + d_{ir,t} d_{js,t} d_{kr,t} d_{ls,t} + d_{is,t} d_{jr,t} d_{kr,t} d_{ls,t}$,
$D_t^{1/2} = (d_{ij,t})_{i,j=1,\ldots,p}$, sk_{it}^s, $kurt_{it}^s$ are obtained with α_{it}, β_{it}

Back to "Example"

Conditional VaR (CVaR) Optimization

Given $\alpha > 0.5$ confidence level,

$$\min_{w_t \in \mathbb{R}^p} \text{CVaR}_\alpha(w_t), \quad \text{s.t. } \mu_{P,t}(w_t) = r_T, w_t^\top 1_p = 1, w_{i,t} \geq 0, \tag{16}$$

$$\text{CVaR}_\alpha(w_t) = -\frac{1}{1-\alpha} q_\alpha^*(w_t) \sigma_{P,t}(w_t), \tag{17}$$

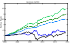
Proof

where (via Cornish-Fisher (CF) expansion):

$$q_{\alpha^*}^*(w_t) = \left\{ 1 + \frac{S_{P,t}(w_t)}{6} z_{\alpha^*} + \frac{K_{P,t}(w_t)}{24}(z_{\alpha^*}^2 - 1) - \frac{S_{P,t}^2(w_t)}{36}(2z_{\alpha^*}^2 - 1) \right\} \varphi(z_{\alpha^*}), \tag{18}$$

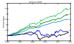
where $\alpha^* \overset{\text{def}}{=} 1 - \alpha$

Back to "Strategies"



The Orthogonal GARCH Model

- X_t matrix of asset returns, $\Gamma_t = B_t \in \mathbb{R}^{p \times p}$ matrix of standardized eigenvectors of $n^{-1} X_t^\top X_t$ ordered according to decreasing magnitude of eigenvalues
- $F_t = P_t \stackrel{\text{def}}{=} X_t \Gamma_t$ matrix of principal components of X_t
- keep the first k most important factors f, introduce noise u_i, then $y_j = b_{j1} f_1 + b_{j2} f_2 + \ldots + b_{jk} f_k + u_i$ or $Y_t = F_t B_t^\top + U_t$
- then $\Sigma_t = \text{Var}(X_t) = \text{Var}(F_t B_t^\top) + \text{Var}(U_t) = B_t \Delta_t B_t^\top + \Omega_t$, $\Delta_t = \text{Var}(F_t)$ diagonal matrix of principal component variances at t: can be separately modeled by univariate GARCH processes Return to "Risk-Return Asset Allocation"



Dynamic Conditional Correlations Model

Assume: $r_t | \mathcal{F}_{t-1} \sim \mathrm{N}(0, D_t R_t D_t)$, 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, $p \times 1$ vector of returns t, D_t is an $p \times p$ diagonal matrix of standard deviations σ_{it}, $i = 1, \ldots, p$, ε_t $p \times 1$ vector of standardized returns with $\varepsilon_{it} \stackrel{\text{def}}{=} r_{it}\sigma_{it}^{-1}$, $\imath$ vector of ones; $P_{t-1} \stackrel{\text{def}}{=} \{\mathrm{diag}(Q_t)\}^{1/2}$, ω_i, α_i, β_i, A, B coefficients, $\odot$ Hadamard product

The DCC Model - Continued

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

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

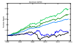
where the volatility part is the sum of individual GARCH likelihoods jointly maximized by separately maximizing each term

$$\begin{aligned} L_V(\theta) &= -\frac{1}{2}\sum_{t=1}^{T}\left\{n\log(2\pi) + \log|D_t|^2 + r_t^\top D_t^{-2} r_t\right\} \\ &= -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{d}\left\{\log(2\pi) + \log(\sigma_{it}^2) + \frac{r_{it}^2}{\sigma_{it}^2}\right\}, \end{aligned}$$

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



Cornish-Fisher VaR Optimization

The alternative asset allocation (Favre, Galeano, 2002)

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

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 If $S_p(w)$, $K_p(w)$ zero, then obtain Markowitz allocation

Back to "Framework"

Risk Parity (Equal risk contribution)

Let $\sigma(w) = \sqrt{w^\top \Sigma w}$. Euler decomposition:

$$\sigma(w) = \sum_{i=1}^{n} \sigma_i(w) = \sum_{i=1}^{n} w_i \frac{\sigma(w)}{\partial w_i}$$

where $\frac{\sigma(w)}{\partial w_i}$ is the marginal risk contribution and $\sigma_i(w) = w_i \frac{\sigma(w)}{\partial w_i}$ the risk contribution of i-th asset. The idea of ERC strategy is to find risk balanced porfolio, such that:

$$\sigma_i(w) = \sigma_j(w)$$

i.e. the risk contribution is the same for all assets of the portfolio

Back to "Empirical results"

60/40 allocation strategy

60/40 portfolio allocation strategy implies the investing of 60% of the portfolio value in stocks (often via a broad index such as S&P500) and 40% in government or other high-quality bonds, with regular rebalancing to keep proportions steady.

Back to "Empirical results"

Portfolio Skewness and Kurtosis

Skewness S_P and excess kurtosis K_P are given by moment expressions

$$S_P(w) = \frac{1}{\sigma_P^3(w)}(m_3 - 3m_2m_1 + 2m_1^3)$$

$$K_P(w) = \frac{1}{\sigma_P^4(w)}(m_4 - 4m_3m_1 + 6m_2m_1^2 + 3m_1^4) - 3$$

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

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

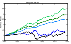
where $\sigma_P^2(w) = w^\top \Sigma w$ and $S_{ijk} = \mathrm{E}(r_i \times r_j \times r_k)$, $K_{ijkl} = \mathrm{E}(r_i \times r_j \times r_k \times r_l)$ can be computed via sample averages from returns data.

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

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

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

▸ Back



Regularity Conditions for Adaptive Lasso QR

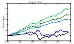
A1 Sampling and smoothness: $\forall x$ in the support of X_i, $\forall y \in \mathbb{R}$, $f_{Y_i|X_i}(y|x)$, $f \in \mathcal{C}^k(\mathbb{R})$, $|f_{Y_i|X_i}(y|x)| < \overline{f}$, $|f'_{Y_i|X_i}(y|x)| < \overline{f'}$; $\exists \underline{f}$, such that $f_{Y_i|X_i}(x^\top \beta_\tau|x) > \underline{f}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

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 $j \notin T$; $T = \{0, 1, ..., s\}$, $\overline{T}(\delta, m) \subset \{0, 1, ..., p\} \backslash T$, then $\exists m \geq 0$, $c \geq 0$ such that

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

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:

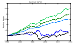
$$\frac{q^3\{\log(n \vee p)\}^{2+\eta}}{n} \to 0, \eta > 0$$

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 $\forall j \leq q$, $|\hat{\beta}_j| > b_0$



Proof of the CF-CVaR Expansion 1

Define the Cornish-Fisher expansion:

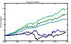
$$q_{1-\alpha} \stackrel{\text{def}}{=} z_{1-\alpha} + (z_{1-\alpha}^2 - 1)s + (z_{1-\alpha}^3 - 3z_{1-\alpha})k - (2z_{1-\alpha}^3 - 5z_{1-\alpha})s^2,$$

where $s \stackrel{\text{def}}{=} S/6$, $k \stackrel{\text{def}}{=} K/24$, S and K are skewness and excess kurtosis, respectively; $z_{1-\alpha} \stackrel{\text{def}}{=} \Phi^{-1}(1-\alpha)$.

Re-write:

$$q_{1-\alpha} = a_0 + a_1 z_{1-\alpha} + a_2 z_{1-\alpha}^2 + a_3 z_{1-\alpha}^3, \tag{19}$$

where $a_0 = -s$, $a_1 = 1 - 3k + 5s^2$, $a_2 = s$, $a_3 = k - 2s^2$



Proof of the CF-CVaR Expansion 2

Define the *conditional Value-at-Risk* (CVaR) or *expected shortfall* (ES):

$$\mathrm{CVaR}_\alpha \stackrel{\mathrm{def}}{=} \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_q \mathrm{d}\, q,$$

where $\mathrm{VaR}_q \stackrel{\mathrm{def}}{=} -\Phi^{-1}(\alpha)\sigma\sqrt{T}$

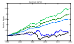
Observe:

$$\mathrm{CVaR}_\alpha = -\frac{1}{1-\alpha}\int_\alpha^1 \Phi^{-1}(\alpha)\sigma\sqrt{T}\mathrm{d}\, q \quad (20)$$

$$= -\frac{\sigma\sqrt{T}}{1-\alpha}\int_\alpha^1 \Phi^{-1}(\alpha)\mathrm{d}\, q \quad (21)$$

$$= -\frac{\sigma\sqrt{T}}{1-\alpha}\int_{-\infty}^{-z_\alpha} u\varphi(u)\mathrm{d}\, u, \quad (22)$$

where (22) follows from the change of variable: $u = z_q = \Phi^{-1}(q)$



Proof of the CF-CVaR Expansion 3

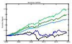
Substitute (19) into (22):

$$\begin{aligned}\mathrm{CVaR}_\alpha &= -\frac{\sigma\sqrt{T}}{1-\alpha}\int_{-\infty}^{-z_\alpha}\left(a_0 + a_1 z + a_2 z^2 + a_3 z^3\right)\varphi(z)\mathrm{d}\,z\\ &= a_0A_0 + a_1A_1 + a_2A_2 + a_3A_3,\end{aligned}$$

$$\begin{aligned}A_0 &= -\frac{\sigma\sqrt{T}}{1-\alpha}\int_{-\infty}^{-z_\alpha}\varphi(z)\mathrm{d}\,z = -\sigma\sqrt{T},\\ A_1 &= -\frac{\sigma\sqrt{T}}{1-\alpha}\int_{-\infty}^{-z_\alpha} z\varphi(z)\mathrm{d}\,z = \frac{\sigma\sqrt{T}}{1-\alpha}\varphi(z_\alpha),\\ A_2 &= -\frac{\sigma\sqrt{T}}{1-\alpha}\int_{-\infty}^{-z_\alpha} z^2\varphi(z)\mathrm{d}\,z = -\sigma\sqrt{T}\left(\frac{\varphi(z_\alpha)z_\alpha}{1-\alpha}+1\right),\\ A_3 &= -\frac{\sigma\sqrt{T}}{1-\alpha}\int_{-\infty}^{-z_\alpha} z^3\varphi(z)\mathrm{d}\,z = \frac{\sigma\sqrt{T}}{1-\alpha}(z_\alpha^2+2)\varphi(z_\alpha).\end{aligned}$$

Collecting terms and simplifying gives the desired result. □

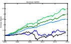
▸ Return to "Conditional VaR Optimization"



References I

Bassett, G. W., Koenker, R., Kordas, G.
Pessimistic Portfolio Allocation and Choquet Expected Utility
Journal of Financial Econometrics, Vol. 2, No. 4 (2004)

Buehlmann, P., van de Geer, S.
Statistics for High-Dimensional Data: Methods, Theory and Applications
Springer, 2011



References II

Chang, C. and Tsay, R.
Estimation of Covariance Matrix via the Sparse Cholesky Factor with Lasso
Journal of Statistical Planning and Inference, Vol. 140, No. 12, 3858–3873 (2004)

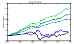
Dowd, K.
Measuring Market Risk. 2nd Edition
Wiley, 2005

References III

Engle, R.
Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation
Econometrica, Vol. 50, No. 4 (1982)

Fan, J., Y. Zhang and Ke Yu.
Asset Allocation and Risk Assessment with Gross Exposure Constraints for Vast Portfolios
Working paper, Princeton University

Ghalanos, A., Rossi, E. and Urga, G.
Independent Factor Autoregressive Conditional Density Model
Econometric Reviews, Vol. 34, No. 5, 594–616 (2015)



References IV

Jondeau, E. and Rockinger, M.

The Impact of Shocks on Higher Moments

Journal of Financial Econometrics, Vol. 7, No.2, 77–105 (2009)

Koenker, R., Bassett, G. W.
Regression Quantiles
Econometrica, Vol. 46, No. 1 (January, 1978)

Lhabitant, F.-S.
Hedge Funds. Myths and Limits
Wiley, 2002

References V

Li, Y., Zhu, J.

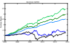
L_1-norm Quantile Regression

Journal of Computational and Graphical Statistics, Vol. 17, No. 1, pp. 1-23

Markowitz, H.,

Portfolio Selection

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



References VI

Schmidt, R, Hrycej, T. and Stützle, E.

Multivariate Distribution Models with Generalized Hyperbolic Margins

Computational Statistics and Data Analysis (2006), 50, pp. 2065–2096

Tibshirani, R.

Regression Shrinkage and Selection via the Lasso

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

References VII

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