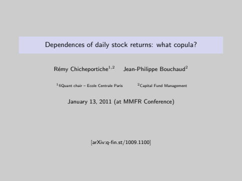
Dependences of daily stock returns: what copula?

Rémy Chicheportiche[1,2] Jean-Philippe Bouchaud[2]

[1] fiQuant chair – Ecole Centrale Paris [2] Capital Fund Management

January 13, 2011 (at MMFR Conference)

[arXiv:q-fin.st/1009.1100]



Good afternoon to everyone.

The talk I'm going to give bases on the results of a paper we wrote together with Jean-Philippe Bouchaud some months ago and that is currently being reviewed.

I will be presenting some empirical observations on daily stock returns and their dependences, and argue that their distribution is not compatible with a well-known class of multivariate distribution, namely the elliptical distributions.

I will then procede and expose our attempts in modeling the stock dependences with a structural model, the intuition of which relies on hierarchical classification.

Understand — and eventually model — the origin and structure of dependences across financial data, like changes in traded stock prices, companies default probabilities, etc.

Motivations:

- Risk Management (Optimal diversification, VaR, contagion and cascades, Expected Shortfall, etc.)
- Optimal portfolio

Understand — and eventually model — the origin and structure of dependences across financial data, like changes in traded stock prices, companies default probabilities, etc.

Motivations:

- Risk Management (Optimal diversification, VaR, contagion and cascades, Expected Shortfall, etc.)
- Optimal portfolio

The aim of this work is to isolate stylized facts regarding assets dependences, come up with a description of the structure of dependences, and to make sense out of it.

Potential application ranges are in Risk Management and Optimal Portfolio design

Linear covariances $\langle r_i r_j \rangle$ are often not sufficient

- Crises and extreme market conditions: large events
- Gamma-risk of Δ-hedged option portfolios: $\langle r_i^2 r_j^2 \rangle - \langle r_i^2 \rangle \langle r_j^2 \rangle$
- Correlated companies defaults: tail probabilities

Higher-order correlations needed ... but empirically noisier

The copula fashion: how relevant is this description of multivariate dependences

In what extent could alternative dependence measures be described through the standard linear correlation coefficient? What "new" information do they provide?

Linear covariances $\langle r_i r_j \rangle$ are often not sufficient

- Crises and extreme market conditions: large events
- Gamma-risk of Δ-hedged option portfolios: $\langle r_i^2 r_j^2 \rangle - \langle r_i^2 \rangle \langle r_j^2 \rangle$
- Correlated companies defaults: tail probabilities

Higher-order correlations needed . . . but empirically noisier

The copula fashion: how relevant is this description of multivariate dependences

In what extent could alternative dependence measures be described through the standard linear correlation coefficient? What "new" information do they provide?

In terms of dependences, several preliminary observations and motivating questions can be formulated.

- extreme market conditions: non-linear regime of correlations
- portfolios of derivatives with non-linear payoff: higher-order correlations of underlying asset
- defaults: rare events

Higher order correlations: noisy, in particular for tail dependences (by definition)

The copula, and alternative measures

Measures of dependence (pairwise)

Pearson's correlation coefficients

$$\rho^{(d)} = \frac{Cov(R_i^d, R_j^d)}{\sqrt{\mathbb{V}[R_i^d]\mathbb{V}[R_j^d]}}, \qquad d \in \mathbb{N} \qquad \rho^{(\mathsf{a})} = \frac{Cov(|R_i|, |R_j|)}{\sqrt{\mathbb{V}[|R_i|]\mathbb{V}[|R_j|]}}$$

Coefficient of tail dependence

$$\tau_{UU}(p) = \mathbb{P}\left[R_i > \mathcal{P}_{i<}^{-1}(p) \mid R_j > \mathcal{P}_{j<}^{-1}(p)\right]$$

and similarly for $\tau_{LL}, \tau_{UL}, \tau_{LU}$

Measures of dependence (pairwise)

Pearson's correlation coefficients

$$\rho^{(d)} = \frac{Cov(R_i^d, R_j^d)}{\sqrt{\mathbb{V}[R_i^d]\mathbb{V}[R_j^d]}}, \qquad d \in \mathbb{N} \qquad \rho^{(a)} = \frac{Cov(|R_i|, |R_j|)}{\sqrt{\mathbb{V}[|R_i|]\mathbb{V}[|R_j|]}}$$

Coefficient of tail dependence

$$\tau_{UU}(p) = \mathbb{P}\left[R_i > \mathcal{P}_{i<}^{-1}(p) \mid R_j > \mathcal{P}_{j<}^{-1}(p)\right]$$

and similarly for $\tau_{LL}, \tau_{UL}, \tau_{LU}$

Let me introduce some tools and define the observables that I will be talking about later

Similarly to the moments of a univariate distibution, define the covariance of the powered variables of a pair

Tail dependence: the joint — or rather conditional — probability that one variable of a pair is beyond its p-th quantile *knowing* that the other one is, too.

Symmetric, provided the marginal cdf's are identical $\mathcal{P}_{i<}(p) = \mathcal{P}_{j<}(p)$ (Bayes formula)

Mostly interesting when $p \approx 1$

Copulas: a global picture of dependences

Sklar's theorem: any N-variate distribution can be decomposed into

- its univariate marginals $\mathcal{P}_i(r_i)$
- a "copula" (or copula density) function describing the dependence structure between N $\mathcal{U}[0,1]$ r.v.: $c(u_1, u_2, \ldots, u_N)$

$$\boxed{c(u,v) = d\mathbb{P}[R_1 \leq \mathcal{P}_{1<}^{-1}(u) \text{ and } R_2 \leq \mathcal{P}_{2<}^{-1}(v)]}$$

Construction:

- from its very definition and properties: e.g. Archimedean copulas
  Almost any $c(u_1, u_2, \ldots, u_N)$ with required properties is allowed
- Sklar's theorem applied to mathematically convenient multivariate distributions: e.g. Gaussian, or Elliptically Contoured copulas
  Brute force parametric estimation for calibration
- structural model
  Statistical tests are not enough: intuition and interpretation are needed (underlying structure, mechanisms of dependence).

Copulas: a global picture of dependences

Sklar's theorem: any N-variate distribution can be decomposed into

- its univariate marginals $\mathcal{P}_i(r_i)$
- a "copula" (or copula density) function describing the dependence structure between N $\mathcal{U}[0,1]$ r.v.: $c(u_1, u_2, \dots, u_N)$

$$\boxed{c(u,v) = d\mathbb{P}[R_1 \leq \mathcal{P}_{1<}^{-1}(u) \text{ and } R_2 \leq \mathcal{P}_{2<}^{-1}(v)]}$$

Construction:

- from its very definition and properties: e.g. Archimedean copulas
  Almost any $c(u_1, u_2, \dots, u_N)$ with required properties is allowed
- Sklar's theorem applied to mathematically convenient multivariate distributions: e.g. Gaussian, or Elliptically Contoured copulas
  Brute force parametric estimation for calibration
- structural model
  Statistical tests are not enough: intuition and interpretation are needed (underlying structure, mechanisms of dependence).

Copulas provide a global picture of dependences because, according to Sklar's theorem, blabla (s'arreter sur la definition)

Example for a pair ($N = 2$).

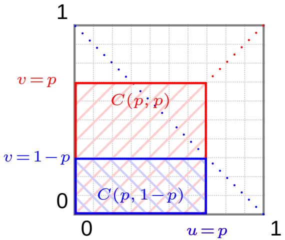
The copula is the joint probability that the variables are all below given quantiles of their marginal distributions

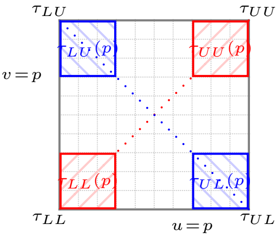
A copula can be buid essentially in three ways:

- from definition: independence copula, Archimedean copulas (Clayton, Frank, Joe, Gumbel, ...) . $C(\boldsymbol{u}) = \Phi^{-1}\left(\Phi(u_1) + \ldots + \Phi(u_N)\right)$,
  where Φ (the generator function) is $N-2$ times continuously differentiable, and such that $\Phi(1) = 0$, $\lim_{u \to 0} \Phi(u) = \infty$ and $\Phi^{(N-2)}$ is decreasing convexe.
- Sklar's theorem: usual classes of multivariate distributions, max. likelihood
- implicit in a structural model for underlying variables

Copulas (bivariate)

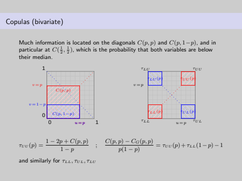
Much information is located on the diagonals $C(p,p)$ and $C(p,1-p)$, and in particular at $C(\frac{1}{2},\frac{1}{2})$, which is the probability that both variables are below their median.





$$\tau_{UU}(p) = \frac{1-2p+C(p,p)}{1-p} \quad ; \quad \frac{C(p,p)-C_G(p,p)}{p(1-p)} = \tau_{UU}(p) + \tau_{LL}(1-p) - 1$$

and similarly for $\tau_{LL}, \tau_{UL}, \tau_{LU}$



We're not looking at the copula over the whole space, but instead reduce the dimension by considering only two directions: diagonals do a good job, and are naturally linked to the coefficients of tail dependence

normalized difference with respect to gaussian case better catches fine differences, and has interesting limits

Elliptical distribution: definition and properties

Let $R = \sigma\epsilon$, with $\epsilon \sim \mathcal{N}(0, \Sigma)$ and $\sigma \sim \mathcal{P}(\sigma; \nu)$

Denote $r_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$ and $f^{(d)} = \langle \sigma^{2d} \rangle / \langle \sigma^d \rangle^2$

Then

- $\rho_{ij}^{(1)} = r_{ij}$
- $\rho_{ij}^{(2)} = \frac{f^{(2)}(1+2r_{ij}^2)-1}{3f^{(2)}-1}$
- $\rho_{ij}^{(\mathrm{a})} = \frac{f^{(1)}(\sqrt{1-r_{ij}^2}+r_{ij}\arcsin r_{ij})-1}{\frac{\pi}{2}f^{(1)}-1}$
- $C_{ij}(\frac{1}{2}, \frac{1}{2}) = \frac{1}{4} + \frac{1}{2\pi}\arcsin r_{ij}$
- the tail dependences $\tau(p)$ have a finite limit when $p \to 1$: simultaneous extreme events are made possible because of the common volatility.

The linear correlation coefficient and $C_{ij}(\frac{1}{2}, \frac{1}{2})$ are invariant in the class of elliptical r.v. The tail dependence coefficients, are invariant in the subclass of (aproximate) power-law decaying distributions.

Elliptical distribution: definition and properties

Let $R = \sigma\epsilon$, with $\epsilon \sim \mathcal{N}(0, \Sigma)$ and $\sigma \sim \mathcal{P}(\sigma; \nu)$

Denote $r_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$ and $f^{(d)} = \langle \sigma^{2d} \rangle / \langle \sigma^{d} \rangle^2$

Then

- $\rho^{(1)}_{ij} = r_{ij}$
- $\rho^{(2)}_{ij} = \frac{f^{(2)}(1+2r^2_{ij})-1}{3f^{(2)}-1}$
- $\rho^{(a)}_{ij} = \frac{f^{(1)}(\sqrt{1-r^2_{ij}}+r_{ij}\arcsin r_{ij})-1}{\frac{\pi}{2}f^{(1)}-1}$
- $C_{ij}(\frac{1}{2}, \frac{1}{2}) = \frac{1}{4} + \frac{1}{2\pi}\arcsin r_{ij}$
- the tail dependences $\tau(p)$ have a finite limit when $p \to 1$: simultaneous extreme events are made possible because of the common volatility.

The linear correlation coefficient and $C_{ij}(\frac{1}{2}, \frac{1}{2})$ are invariant in the class of elliptical r.v. The tail dependence coefficients, are invariant in the subclass of (aproximate) power-law decaying distributions.

First, what *are* elliptically distributed variables in the language of finance?

I'm not going to be too technical, it will be sufficient here to see elliptical distributions as generated by a stochastic volatility model of multivariate data.

ϵ is a random vector, σ is a random variable, COMMON TO ALL INDIVIDUALS, with its own distribution $\mathcal{P}$ possibly indexed by a parameter ν (ex: Student, when $\mathcal{P}$ is inverse-Gamma).

Pearson's correlations are just ... the normalized elements of the dispersion matrix Σ

the higher order correlations are related to the moments of the stochastic volatility

$f^{(2)} = \frac{\kappa_\sigma + 1}{3}$

Data

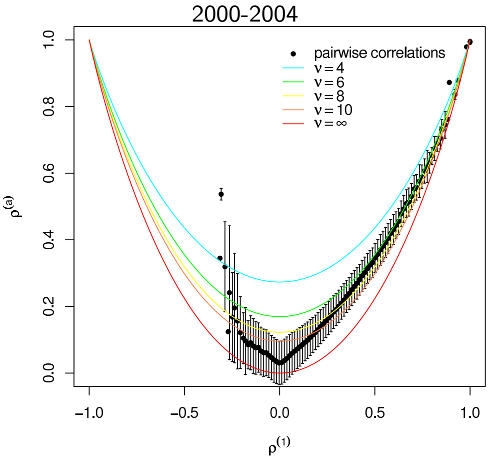
- US daily stock returns since 1990
- rescaled time series
- a pair is only relevant if 200 common trading dates
- First: 6 pools (LMS Caps $\times$ 2 subperiods)
- Then: 4 pools (all caps over 4 subperiods)
- Same thing for Japan

Data

- US daily stock returns since 1990
- rescaled time series
- a pair is only relevant if 200 common trading dates
- First: 6 pools (LMS Caps $\times$ 2 subperiods)
- Then: 4 pools (all caps over 4 subperiods)
- Same thing for Japan

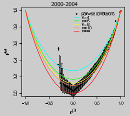
rescaled time series: not a study/model of stock *prices* but *dependences*

absolute vs linear correlation: data and Student



"more Gaussian" for small ρ

absolute vs linear correlation: data and Student



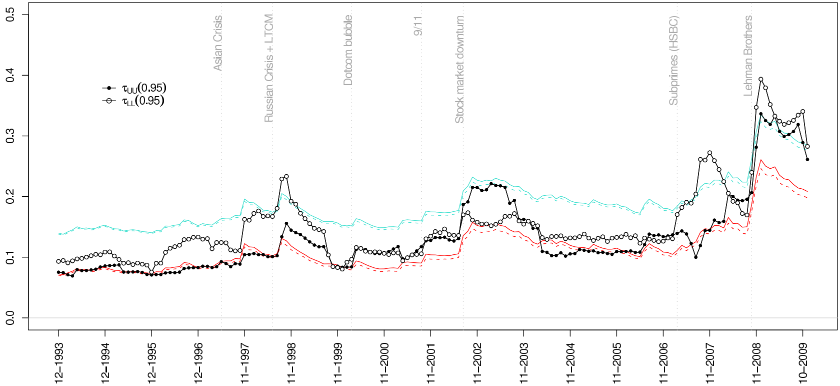
"more Gaussian" for small ρ

$$\rho_{ij}^{(\mathsf{a})} = \frac{f^{(1)}(\sqrt{1-r_{ij}^2}+r_{ij}\arcsin r_{ij})-1}{\frac{\pi}{2}f^{(1)}-1},\ f^{(1)} = \frac{2}{\nu-2}\left(\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu-1}{2})}\right)^2$$

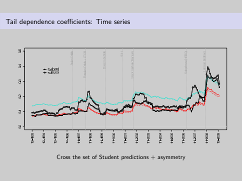
As a first observable, we looked at the relationship between the absolute correlation (Y axis) and linear correlation coefficients (X axis).
The black point and dispersion bars correspond to the empirical data averaged in bins of rho.
The coloured lines are the Student predictions (equivalent for any elliptical distribution) for different values of the degree of freedom parameter ν
Link to tail index (lower ν, fatter tails, more probability of *jointly* large events)

The larger ρ, the more "Student-like" with low degree of freedom !

Tail dependence coefficients: Time series



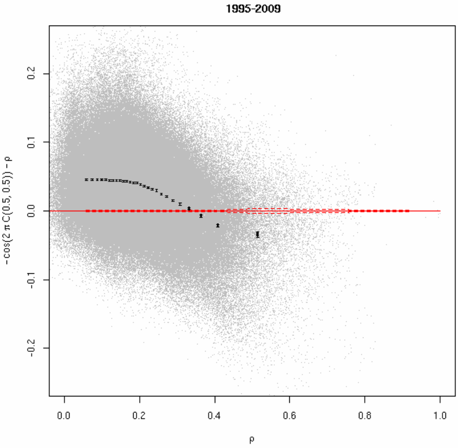
Cross the set of Student predictions + asymmetry



I present here the time evolution of the tail dependence at a quantile level $p = 0.95$ (95-th centile): data in black for U and L tails. Given the time-evolving ρ, Student descriptions predict a tail dependence , drawn here for $\nu = 4$ and $\nu = \infty$.

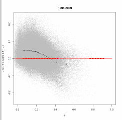
The larger ρ, the more "Student-like" with low degree of freedom !

Center point of the copula



$$C_{\mathsf{emp}}(\tfrac{1}{2},\tfrac{1}{2} \mid \rho) \neq \frac{1}{4} + \frac{1}{2\pi}\arcsin\rho \quad \text{i.e.} \quad -\cos\left(2\pi C_{\mathsf{emp}}(\tfrac{1}{2},\tfrac{1}{2} \mid \rho)\right) - \rho \neq 0$$

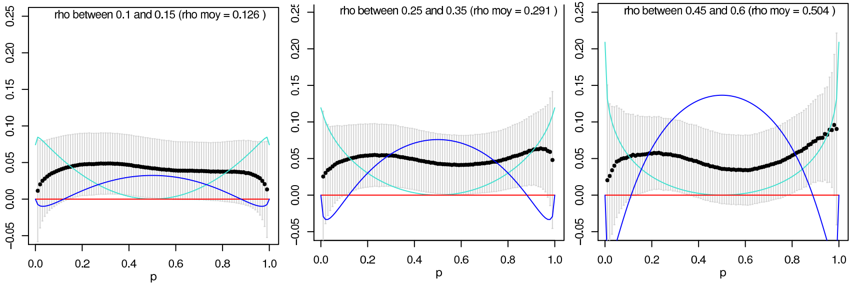
Center point of the copula



$$C_{\mathsf{emp}}(\tfrac{1}{2}, \tfrac{1}{2} \mid \rho) \neq \frac{1}{4} + \frac{1}{2\pi} \arcsin \rho \quad \text{i.e.} \quad -\cos\left(2\pi C_{\mathsf{emp}}(\tfrac{1}{2}, \tfrac{1}{2} \mid \rho)\right) - \rho \neq 0$$

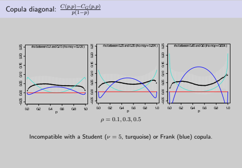
NOT a trivial consequence of the asymmetry in the tails !

Copula diagonal: $\frac{C(p,p)-C_G(p,p)}{p(1-p)}$



$$\rho = 0.1, 0.3, 0.5$$

Incompatible with a Student ($\nu = 5$, turquoise) or Frank (blue) copula.



- curvature
- tail limits
- copula center point

Frank: $\Phi(p) = -\log \frac{e^{-\theta p}-1}{e^{-\theta}-1}$

The larger ρ, the more "Student-like" with low degree of freedom !

Conclusions of the empirical study

- Elliptical copulas fail in describing the multivariate dependence structure of stocks
- Obvious intuitive reason: one expects more than one volatility factor to affect stocks !
- How to describe entanglement of volatility factors with idiosyncratic returns?

2011-01-20

Conclusions of the empirical study

- Elliptical copulas fail in describing the multivariate dependence structure of stocks
- Obvious intuitive reason: one expects more than one volatility factor to affect stocks !
- How to describe entanglement of volatility factors with idiosyncratic returns?

Bla bla

We have to find something inbetween: add structure inside the vol, but depart from "pseudo-elliptical" constructions.

Intuitions

1. Events occurence/amplitudes rely on factors
2. Factors are hierarchical. In what sense?

Take a one-factor model,

$$R_i = \beta_i I_0 + \eta_i$$

Empirical fact:
the dispersion of residuals η_i increases with the vol of the market index I_0.

Or, expliciting the stochastic volatilities,

$$R_i = \beta_i \sigma_0 \epsilon_0 + \sigma_0^{\alpha} \sigma_i \epsilon_i, \qquad \alpha > 0$$

Intuitions

1. Events occurence/amplitudes rely on factors
2. Factors are hierarchical. In what sense?

Take a one-factor model,

$$R_i = \beta_i I_0 + \eta_i$$

Empirical fact:
the dispersion of residuals η_i increases with the vol of the market index I_0.

Or, expliciting the stochastic volatilities,

$$R_i = \beta_i \sigma_0 \epsilon_0 + \sigma_0^{\alpha} \sigma_i \epsilon_i, \qquad \alpha > 0$$

Now, how to procede from here? The conclusions of the empirical study suggest a rich structure of the volatilities. The following intuitions shall guide us toward a model of stock dependences.

Obviously, the stock return is driven by the factor it is exposed to. (see literature on hierarchical clustering). Here we go farther in considering hierarchy also for the volatilities.

$\alpha > 0$ (increase), but discuss $\alpha < 1$ or $\alpha > 1$.
$\alpha = 1 \Rightarrow$ factorize σ, so back to roughly elliptical case.

A hierarchical model

More levels generalization: the vol of the market influences that of sectors, which in turn influence that of more indisyncratic factors.

Example: 2 levels (market + sectors) with log-normal volatilities:

$$I_0 = \epsilon_0 \, \mathrm{e}^{s_0 \xi_0}$$

$$F_f = \Gamma_f \, I_0 + \epsilon_f \, \mathrm{e}^{\alpha_f^I s_0 \xi_0 + s_f \xi_f}$$

$$R_i = \beta_i \, F_f + \epsilon_i \, \mathrm{e}^{\alpha_i^I s_0 \xi_0 + \alpha_i^F s_f \xi_f + s_i \xi_i}$$

with ϵ, ξ Gaussian.

A hierarchical model

More levels generalization: the vol of the market influences that of sectors, which in turn influence that of more indisyncratic factors.

Example: 2 levels (market + sectors) with log-normal volatilities:

$$
\begin{aligned}
I_0 &= \epsilon_0 \, e^{s_0 \xi_0} \\
F_f &= \Gamma_f \, I_0 + \epsilon_f \, e^{\alpha_f^I s_0 \xi_0 + s_f \xi_f} \\
R_i &= \beta_i \, F_f + \epsilon_i \, e^{\alpha_i^I s_0 \xi_0 + \alpha_i^F s_f \xi_f + s_i \xi_i}
\end{aligned}
$$

with ϵ, ξ Gaussian.

Bla bla

Spend time defining and explaining ("financial meaning") each parameter.

Exposure parameters

Usual way: regress R_i on F_f to get β_i, and F_f on I_0 to get Γ_f. But

- what is F_f and I_0?
- endogeneity

Instead: use the predicted correlations

$$\langle R_i R_j \rangle = \begin{cases} 1 & , i = j \\ \beta_i \beta_j & , f_i = f_j, i \neq j \quad \text{(same sector)} \\ \beta_i \Gamma_{f_i} \beta_j \Gamma_{f_j} & , f_i \neq f_j \quad \text{(diff. sector)} \end{cases}$$

(Almost) no knowledge needed of what the factors are, how they are built.
Still relies on predefined hierarchical classification: for us, 9 Bloomberg sectors!

(back)

Exposure parameters

Usual way: regress R_i on F_f to get β_i, and F_f on I_0 to get Γ_f. But

- what is F_f and I_0?
- endogeneity

Instead: use the predicted correlations

$$\langle R_i R_j \rangle = \begin{cases} 1 & , i = j \\ \beta_i \beta_j & , f_i = f_j, i \neq j \quad \text{(same sector)} \\ \beta_i \Gamma_{f_i} \beta_j \Gamma_{f_j} & , f_i \neq f_j \quad \text{(diff. sector)} \end{cases}$$

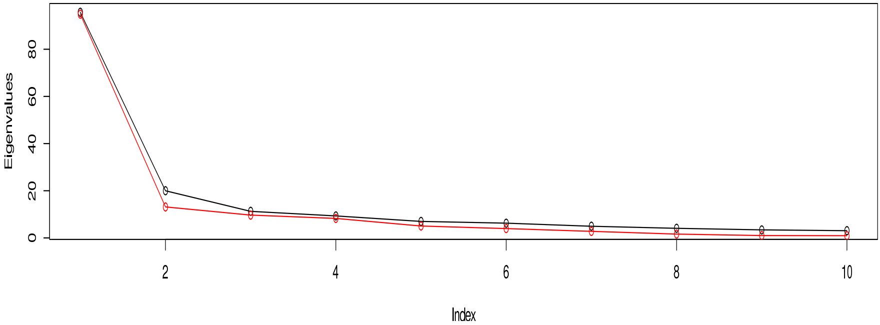
(Almost) no knowledge needed of what the factors are, how they are built.
Still relies on predefined hierarchical classification: for us, 9 Bloomberg sectors!

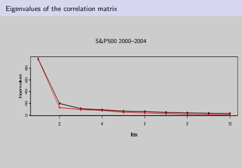
(back)

Bla bla

Eigenvalues of the correlation matrix

S&P500 2000–2004



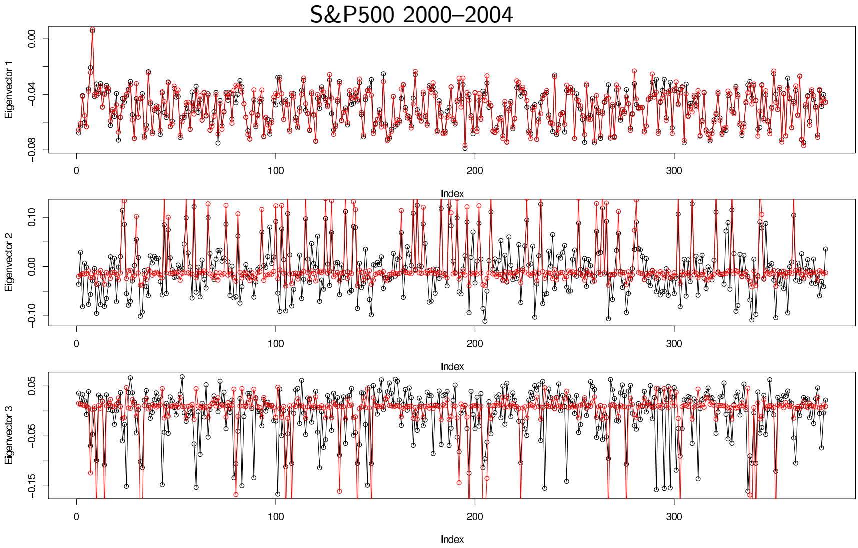


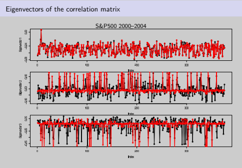
9 sectors

First 10 eigenvalue sorted by decreasing size

black: raw correlation matrix
red: calibration matrix of model with calibrated parameters

Eigenvectors of the correlation matrix





Three first eigenvectors, according to the ranking of their corresponding eigenvalues.

Evoquer le role des stocks "bi" et l'implication en terme de cleaning des vecteurs propres

black: raw correlation matrix
red: calibration matrix of model with calibrated parameters

CLEANING. We couldn't have done it by naked eye (what is noise and what is structure)

Volatility amplifiers and volatility dispersion

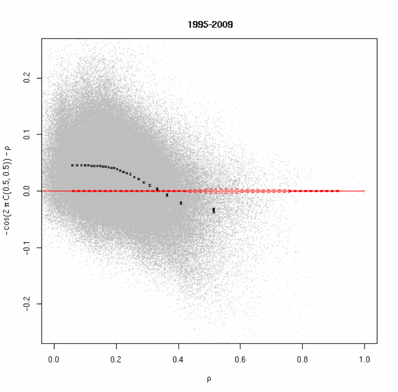
$$\alpha_f^I \approx 0.9 - 1.5 \gtrsim 1$$
$$\alpha_i^F \approx 0 - 1 < 1 \quad \text{(very noisy, though)}$$
$$\alpha_i^I \approx 0.6 - 1 \lesssim 1$$

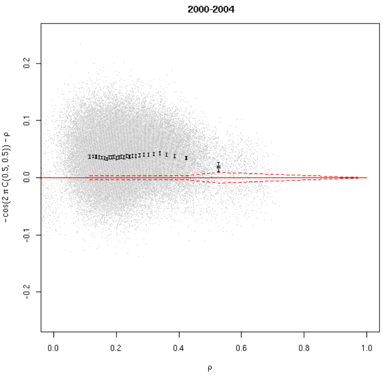
$$s_0 \approx 0.4 \qquad (\nu_{eff} \approx 5)$$
$$s_f \approx 0.15 - 0.25$$
$$s_i \quad \text{distributed around } 0.37 \pm 0.13$$

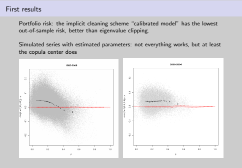
First results

Portfolio risk: the implicit cleaning scheme “calibrated model” has the lowest out-of-sample risk, better than eigenvalue clipping.

Simulated series with estimated parameters: not everything works, but at least the copula center does







linear correl: meilleur out-of-sample risque (cleaning scheme), rappeler le role des stocks "bi"

non-linear correl: pas très bon pour l'instant. (Theoretical) Noise reduction?

étude des résidus: $\alpha > 1$ ou $\alpha < 1$? Discussion

- Classification algo: either rely on outside clustering techniques, or generate moves in existing classification (e.g. Bloomberg) and evaluate a cost function. What are the optimal moves? Cost function? (go)
- Revealed factor structure ⇒ noise undressing of correlation matrix (values AND vectors !!!)
- Extension to account for U/L asymmetry and leverage effects
- Higher frequency: intra-day data
- Another field of application: default probabilities of companies — spreads of CDS's