Flexible copula-based vector MEM

Nikolaus Hautsch
Ostap Okhrin
Alexander Ristig

Chair of Econometrics
Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
http://www2.hu-berlin.de/oekonometrie
http://lvb.wiwi.hu-berlin.de
http://case.hu-berlin.de

Liquidity risk

Example 1

Problem: Let Q be an huge order influencing the market price; thus Q cannot be executed at the actual market price.

Solution: Split Q into smaller amounts $\delta_i \widehat{V}_i$, such that $Q = \sum_{i=1}^{n} \delta_i \widehat{V}_i$, with $\delta_i \in (0,1)$, where $\widehat{V}_i$ is the forecasted trading volumes, $i = 1, \ldots, n$.

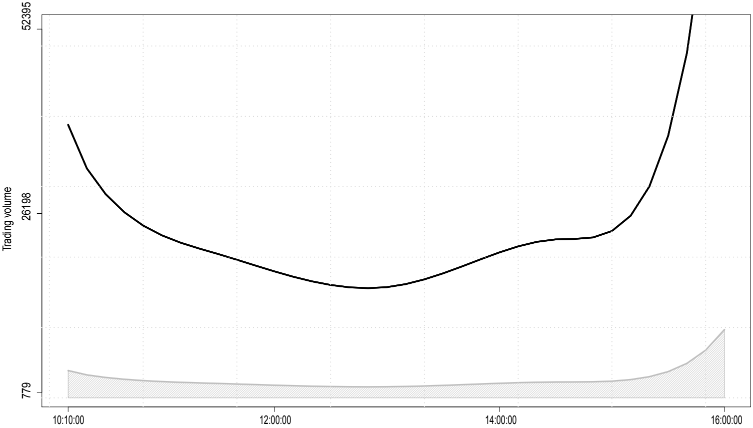


Figure 1: Trading volume of Skyworks Solutions Inc. (SWKS) at the 2009-08-08. Forecasted market volume; limited trade sizes with $\delta_i = 0.05$.

Multiplicative error model (MEM)

Engle (2002) proposed the MEM, i.e.

$$x_i = \mu_i \, \varepsilon_i ,$$
$$\mu_i \stackrel{\text{def}}{=} \mathrm{E}\left(x_i | \mathcal{F}_{i-1}; \xi\right),$$

with

- $\{x_i\}_{i=1}^n$ is a positive valued stochastic process,
- ε_i is iid with $\mathrm{E}\left(\varepsilon_i\right) = 1$ and density $f : \mathbb{R}^+ \to \mathbb{R}^+$,
- Vector of parameters ξ,
- μ_i measurable w.r.t. $\mathcal{F}_{i-1}$.

Research questions

- Joint modeling of volumes, price variations and trading intensities
- Basic ingredients are positive: *no* natural distribution on $\mathbb{R}_+^d$
- More accurate forecasts due to large information set
- Challenges:
  - Non-Gaussian dependencies
  - Two step estimation instead of computationally demanding full-ML estimation

Vector MEM (VMEM)

Cipollini and Gallo (2010) formalized the VMEM as

$$x_i = \mu_i \odot \varepsilon_i,$$

$$\mu_i \stackrel{\text{def}}{=} \mathsf{E}\left(x_i|\mathcal{F}_{i-1};\xi\right) = \underbrace{\omega}_{(d\times 1)} + \underbrace{A}_{(d\times d)} x_{i-1} + \underbrace{B}_{(d\times d)} \mu_{i-1},$$

where "$\odot$" is the componentwise Hadamard product, $\varepsilon_{ij} \geq 0 \ \forall i$, $j = 1, \ldots, d$, $\xi = (\omega, A, B)$ and $\varepsilon_i|\mathcal{F}_{i-1} \sim D\left(1_d, \Sigma\right)$.

Example 2

$\varepsilon_{ij} \sim Exp\left(1\right) = F$ and $D = C\left\{F\left(x_{i1}\right), \ldots, F\left(x_{id}\right); \theta\right\}$

Outline

1. Motivation ✓
2. Estimation
3. HAC
4. Data
5. Preliminary estimation results
6. Summary

Full ML estimation

From VMEM has x_i a conditional density

$$
\begin{aligned}
f_x\left(x_{i1},\ldots,x_{id}|\mathcal{F}_{i-1};\theta,\xi,\alpha\right) &= c\left\{F_1(\varepsilon_{i1};\alpha_1),\ldots,F_d(\varepsilon_{id};\alpha_d);\theta\right\} \\
&\quad \times \prod_{j=1}^{d} \frac{f_j\left(\varepsilon_{ij};\alpha_j\right)}{\mu_{ij}\left(\xi_j\right)},
\end{aligned}
$$

for $\varepsilon_{ij} \stackrel{\text{def}}{=} \varepsilon_{ij}(\xi_j)$, $i = 1,\ldots,n$ and copula density c. Then, the log-likelihood is given through

$$
\begin{aligned}
\ell\left(\theta,\xi,\alpha|\mathcal{F}_{i-1}\right) &= \sum_{i=1}^{n}\sum_{j=1}^{d}\left[\log\left\{\varepsilon_{ij} f_j\left(\varepsilon_{ij};\alpha_j\right)\right\} - \log x_{ij}\right] \\
&\quad + \sum_{i=1}^{n} \log c\left\{F_1(\varepsilon_{i1};\alpha_1),\ldots,F_d(\varepsilon_{id};\alpha_d);\theta\right\}.
\end{aligned}
$$

Advantage:

- Efficient and asymptotically unbiased estimates

Disadvantages:

- If C is flexible structured copula, ℓ has to be optimized for each possible structure
- Curse of dimensionality
  - $2d^2 + 4d - 1$ parameters to estimate (assuming 2 parameters for each margin and $d - 1$ for the copula)
- Parameter restrictions to ensure stationarity and positivity
- Computationally intensive
- ...

Multistage procedure

Similar to Chen and Fan (2006)

1. Estimate $\mu_i(\xi) = \{\mu_{i1}(\xi_1), \dots, \mu_{id}(\xi_d)\}^\top$, $i = 1, \dots, n$
2. Extract the residuals $\hat{\varepsilon}_{i1}, \dots, \hat{\varepsilon}_{id}$ and determine the edfs $\widehat{F}_j$
3. $\hat{\theta} = \arg\max_{\theta} \sum_{i=1}^{n} \log c \left\{ \widehat{F}_1(\hat{\varepsilon}_{i1}), \dots, \widehat{F}_d(\hat{\varepsilon}_{id}); \theta \right\}$

Advantages:

- Separate estimation of $\mu_i(\xi)$ and $C(\cdot;\theta)$
- Applicable for high dimensional time series, conditional on the copula

Disadvantages:

- Matrices A and B should be diagonal
- Estimation equation by equation is inefficient otherwise

Consequence: no real VMEM and no superior forecasting properties

Maximization by parts

Set $\delta = (\xi, \alpha)^\top$ and rewrite the log-likelihood

$$\ell\left(\theta, \delta | \mathcal{F}_{i-1}\right) = \overbrace{\sum_{i=1}^{n}\sum_{j=1}^{d}\left[\log\left\{\varepsilon_{ij} f_j\left(\varepsilon_{ij};\alpha_j\right)\right\} - \log x_{ij}\right]}^{working}$$
$$\underbrace{+\sum_{i=1}^{n}\log c\left\{F_1(\varepsilon_{i1};\alpha_1),\ldots,F_d(\varepsilon_{id};\alpha_d);\theta\right\}}_{error}$$

in two parts

$$\ell\left(\delta, \theta\right) = \ell_w\left(\delta\right) + \ell_e\left(\delta, \theta\right).$$

Shortcuts: $\ell_w^\delta = \partial \ell_w / \partial \delta$, $\ell_e^\delta = \partial \ell_e / \partial \delta$ and $\ell_e^\theta = \partial \ell_e / \partial \theta$

Algorithm (Song et al. (2005))

For $k = 1$:

1. *Solve* $\ell_w^\delta(\delta) = \mathbf{0}$ *for* $\hat{\delta}^1$
2. *Solve* $\ell_e^\theta(\hat{\delta}^1, \theta) = \mathbf{0}$ *for* $\hat{\theta}^1$

For $k = 2, 3, \ldots$:

1. *Solve* $\ell_w^\delta(\delta) = -\ell_e^\delta(\hat{\delta}^{k-1}, \hat{\theta}^{k-1})$ *for* $\hat{\delta}^k$
2. *Solve* $\ell_e^\theta(\hat{\delta}^{k-1}, \theta) = \mathbf{0}$ *for* $\hat{\theta}^k$

HAC

Example 3

3-dimensional fully nested HAC with generators ϕ and ψ:

$$C(u_1, u_2, u_3) = \phi\left\{\phi^{-1} \circ \psi\left\{\psi^{-1}(u_1) + \psi^{-1}(u_2)\right\} + \phi^{-1}(u_3)\right\}$$

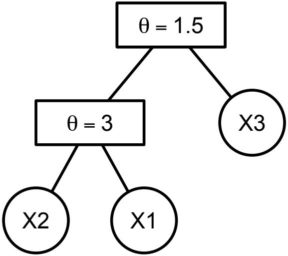


Figure 2: Structure of 3-dim HAC

Estimation

Okhrin et al. (2011) proposed a multi-stage QML-estimation, i.e.

$$\left(\frac{\partial \ell_1}{\partial \theta_1}, \dots, \frac{\partial \ell_{d-1}}{\partial \theta_{d-1}} \right)^\top = \mathbf{0}_{d-1}$$

with quasi log-likelihood

$$\ell(\theta|\mathbf{X}) = n^{-1} \sum_{i=1}^{n} \log \left[c \left\{ \widehat{F}_1(x_{i1}), \dots, \widehat{F}_d(x_{id}), \theta \right\} \right].$$

Data

- Datasource: NASDAQ (LOBSTER)
- Stock: SWKS
- 2nd of November 2009 to 31th of December 2009
- Series is cleaned and adjusted for deterministic intraday patterns
- Aggregation level:
    - 10 min
    - 1331 observations

Variables

High-frequency data sets provide information about

$$\{(p_j, s_j, t_j)\}_{j=1}^{n^*},$$

with

- Price series p_j,
- Amount of traded shares s_j,
- Time stamps t_j,
- Number of observation within a day n^*.

Construct the lower frequent equi-distant series

$$\{(\text{HL}_i, \text{Vol}_i, \text{NT}_i, t_i)\}_{i=1}^{n},$$

with

- $n < n^*$
- $\text{HL}_i = \max\{p_j | t_j \in (t_{i-1}, t_i]\} - \min\{p_j | t_j \in (t_{i-1}, t_i]\}$,
- $\text{NT}_i = \#\{t_j | t_j \in (t_{i-1}, t_i]\}$ and
- $\text{Vol}_i = \text{NT}_i^{-1} \sum_{i_j \in (t_{i-1}, t_i]} s_i$.

Multistage estimation

$$\mu_i = c + a x_{i-1} + b\mu_{i-1}$$

	HL	NT	Vol
c	0.02	0.02	0.18
a	0.11	0.20	0.28
b	0.87	0.77	0.56
α_1	-	0.94	0.50
α_2	-	5.00	11.74
p_{LB10}	0.03	0.01	0.10
p_{LB20}	0.09	0.17	0.40
p_{LB30}	0.15	0.06	0.55

Table 1: Estimated parameters and p-values of the Ljung-Box (LB) test; sample size: 1331. Assumptions: $\varepsilon_{\text{HL}} \sim Exp(1)$, $\varepsilon_{\text{NT}} \sim Ga(\alpha_1, \alpha_2)$ and $\varepsilon_{\text{Vol}} \sim Ga(\alpha_1, \alpha_2)$.

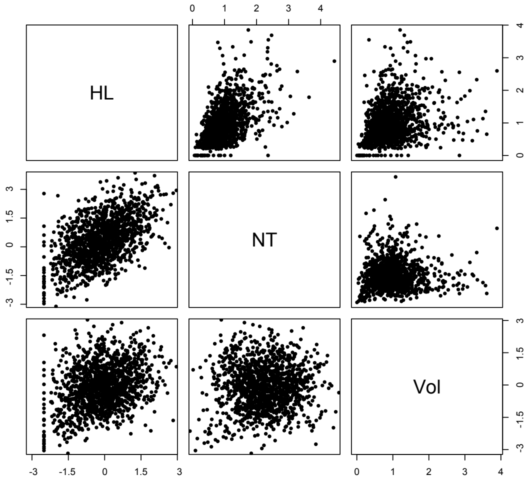


Figure 3: Scatterplot of the residuals.

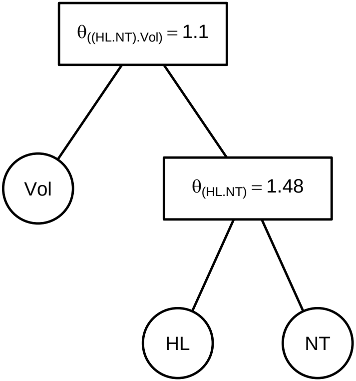


Figure 4: Estimated HAC for the residuals of the univariate processes.

Full ML

$$\mu_i = \begin{pmatrix} 0.22 \\ 0.22 \\ 0.24 \end{pmatrix} + \begin{pmatrix} 0.29 & 0.08 & 0.07 \\ 0.19 & 0.08 & 0.07 \\ 0.09 & 0.24 & 0.22 \end{pmatrix} x_{i-1} + \begin{pmatrix} 0.33 & 0.16 & 0.07 \\ 0.01 & 0.04 & 0.01 \\ 0.00 & 0.00 & 0.15 \end{pmatrix} \mu_{i-1}$$

Stucture	log-lik
$((\mathrm{HL.NT})_{1.85}\mathrm{Vol})_{1.27}$	-2740
$((\mathrm{HL.Vol})_{1.49}\mathrm{NT})_{1.43}$	-2828
$((\mathrm{NT.Vol})_{1.44}\mathrm{HL})_{1.44}$	-2838

With $\alpha_1 = (1.50, 2.73, 3.68)^\top$, $\alpha_2 = (2.98, 2.07, 1.04)^\top$ for $(\mathrm{HL}, \mathrm{NT}, \mathrm{Vol})^\top$.

Conclusion

- Aim at forecasting economic variables
- Two step estimation for VMEM
- Benefit from well developed univariate models and modern copulas

To do:

- Maximization by parts
- Time varying relationships
- Penalization of the parameter space's dimensionality

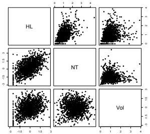
Flexible copula-based vector MEM

Nikolaus Hautsch

Ostap Okhrin

Alexander Ristig

Chair of Econometrics
Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
http://www2.hu-berlin.de/oekonometrie
http://lvb.wiwi.hu-berlin.de
http://case.hu-berlin.de



References

Chen, X. and Y. Fan,

Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models under Copula Misspecification

Journal of Econometrics, 135(1-2), 125–154, 2006

Cipollini, F. and G. M. Gallo

Automated Variable Selection in Vector Multiplicative Error Models

Computational Statistics and Data Analysis 54, 2470–2486, 2010

Engle, R. F.

New Frontiers for ARCH Models

Journal of Applied Econometrics, 17(5), 425–446, 2002

Härdle, W. K., O. Okhrin, and Y. Okhrin
Time Varying Hierarchical Archimedean Copulae
SFB 649 discussion paper 2010, 018, SFB 649, Economic Risk, Berlin, 2010

Harry Joe
Multivariate Models and Dependence Concepts
Chapman & Hall, 1997

Okhrin O., Y. Okhrin and W. Schmid
Determining the Structure and Estimation of Hierarchical Archimedean Copulas
Journal of Econometrics, under revision

Song P., Y. Fan and J. D. Kalbfleisch
Maximization by Parts in Likelihood Inference
Journal of the American Statistical Association, Vol. 100, No. 472, 1145–1158, 2005