Time Varying Independent Component Analysis

Ray-Bing Chen

Ying Chen

Wolfgang Karl Härdle

National Cheng Kung University

National University of Singapore

Humboldt-Universität zu Berlin

Source extraction and dimension reduction

High dimensional and complex financial time series are neither Gaussian distributed nor stationary.

Multivariate Data Analysis (MDA)

Let $\mathbf{X}_t \in \mathbb{R}^p$ denote the returns of financial assets.

- Principal component analysis: $\mathbf{X}_t = \Gamma \times \mathrm{PC}_t$,
- Factor analysis: $\mathbf{X}_t = \Gamma\Lambda^{1/2}F_t + U_t$,

Jolliffe (2002), Härdle and Simar (2012)

Under Gaussianity, cross-uncorrelatedness indicates independence.

Jacobian transformation for a linear transformation $X = AZ$:

$$f_Z(z) = \prod_{j=1}^{p} f_{Z_j}(z_j), \quad f_X(x) = abs(|A|^{-1}) \cdot f_Z(A^{-1}X)$$

Fact: Financial time series are heavy-tailed distributed.

Independent Component Analysis (ICA)

Let $\mathbf{X}_t \in \mathbb{R}^p$ denote the returns of financial assets:

$$
\begin{aligned}
\text{IC}_t &= B\mathbf{X}_t = (b_1, \cdots, b_p)^\top \mathbf{X}_t \\
\begin{pmatrix} \text{IC}_{1t} \\ \vdots \\ \text{IC}_{pt} \end{pmatrix} &= \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ . & \cdots & . \\ b_{p1} & \cdots & b_{pp} \end{pmatrix} \begin{pmatrix} x_{1t} \\ \vdots \\ x_{pt} \end{pmatrix} \\
\text{equivalently } \mathbf{X}_t &= A \times \text{IC}_t
\end{aligned}
$$

where B is a nonsingular filter matrix: $B^{-1} = A$.

How to find ICs?

$$\mathbf{X}_t = A \times \mathrm{IC}_t$$

Jones and Sibson (1987): projection pursuit

Hyvärinen and Oja (1997): FastICA

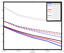
Hyvärinen, Karhunen and Oja (2001): MLE and others

Chen, Guo, Härdle and Huang (2011): COPICA

The observed series as well the ICs are assumed to be stationary.

The filter A (or B) is constant over time.

Fact: Turbulences in financial markets indicate nonstationary.



Demonstration

Log returns of HD, HPQ and IBM.

$$\mathbf{X}_t = \begin{cases} A_1 \mathrm{IC}_t & t \in [1, 300] \\ A_2 \mathrm{IC}_t & t \in [301, 600] \end{cases}$$

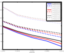
where IC_t are NIG distributed, see Barndorff-Nielson (1997).

Two ICA filters are:

$$A_1 = 10^{-3} \begin{pmatrix} 0.6 & 13.0 & 6.2 \\ 3.8 & 2.7 & 13.0 \\ 7.9 & 5.9 & 4.8 \end{pmatrix}, \qquad A_2 = 10^{-3} \begin{pmatrix} -0.1 & 0.8 & 5.3 \\ 7.0 & 1.9 & 1.6 \\ 0.1 & 4.2 & 1.1 \end{pmatrix}.$$

2008/09/03 – –2009/08/31, (a period with market turbulence)

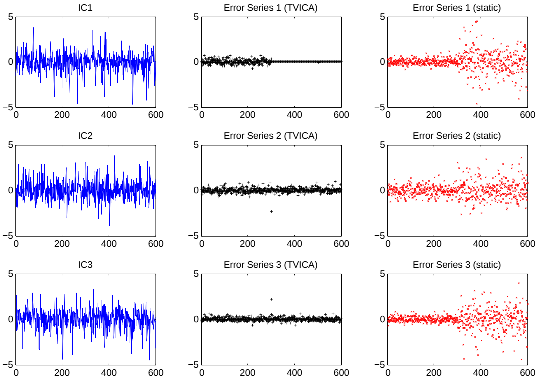
2004/07/30 – –2006/12/29 (a relatively quiet period)

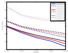


Demonstration (Continued)

Static ICA: average value of RMSEs is 0.886 (1.196 after change)

Time varying ICA: average value of RMSEs is 0.201 (0.160 after change)



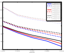


Literature review

Matteson and Tsay (2009): allow the mixing matrix B to vary over time via a smooth function of other transition variables.

- Volatility and co-volatility literature, see e.g. Baillie and Morana (2009), Scharth and Medeiros (2009),
- Incorporate changes via Markov-Switching or mixture of multiplicative error specifications,
- Need a globally given mechanism for this time variation.

Mercurio and Spokoiny (2004) use a local change point (LCP) approach: completely data driven approach.

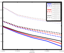


TVICA

Let $\mathbf{X}_t \in \mathbb{R}^p$ denote the returns of financial assets, TVICA model:

$$\mathbf{X}_t = A_t \mathrm{IC}_t$$

- Time varying independent source extraction,
- For each time point t, LCP identifies a "trust interval" $I_t = [t - m_t, t]$, over which the filter A_t $\approx$const.,
- Neither prior information (on say states of the market) nor distributional assumption is required. Data-driven and applicable for various kinds of breaks (macroeconomic or political changes).



Outline

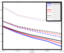
1. Motivation ✓
2. TVICA and estimation
3. Simulation study
4. Real data analysis
5. Conclusion

TVICA

Let $\mathbf{X}_t \in \mathbb{R}^p$ denote the returns of financial assets,
$\mathbf{Z}_t = \{z_1(t), \cdots, z_p(t)\}^\top$ are cross independent.

$$\text{TVICA model:} \quad \mathbf{X}_t = A_t \mathbf{Z}_t, \quad \mathbf{Z}_t = B_t^{-1} \mathbf{X}_t$$

Local Homogeneity: for any particular time point t there exists a past time interval $I_t = [t - m_t, t]$, over which the linear filter A_t is approximately constant, i.e. $A_s \approx A, \forall s \in I_t$.



Estimation: under homogeneity

Suppose that at time point t, an interval of **homogeneity** $I_t = [t - m_t, t)$ is given with m_t indicating the length of the interval.

The log-likelihood function on the interval I_t is:

$$L(I_t, B_t) = \sum_{s=t-m_t}^{t} \sum_{j=1}^{r} \log\{f_j(b_{jt}^{\top} \mathbf{X}_s)\} + (m_t + 1) \log |\det B_t|, \quad (1)$$

where $f_j(z_j)$ is the pdf of IC z_j, $j = 1, \cdots, p$. MLE is $\tilde{B}_t$.

Estimation: under local homogeneity

Small modeling bias: divergence of a time varying model (local homogeneity) to a static model (homogeneity) is small, Spokoiny (2011).
For $r, \rho > 0$, the fitted log likelihood with $B_t = B^*$ satisfies:

$$\mathsf{E}_{B^*}|L(I_k, \tilde{B}_t^{(k)}, B^*)|^r = \mathsf{E}_{B^*}|L(I_k, \tilde{B}_t^{(k)}) - L(I_k, B^*)|^r \le R_r(B^*), \tag{2}$$

where $R_r(B^*) = \max_{k \le K} \mathsf{E}_{B^*}|L_{I_k}(\tilde{B}_k, B^*)|^r$.

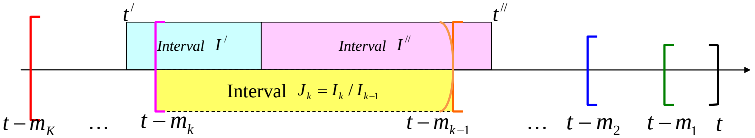
Goal: For any time point t and nested intervals, $I_0 \subset I_1 \subset \cdots \subset I_{K-1} \subset I_K$, LCP method finds the longest interval of local homogeneity.
The identification of the trust interval is done via a sequential testing algorithm.

LCP algorithms

H_0 : I_k is a local homogeneous interval *given that I_{k-1} was not rejected.*

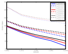
Initialization: I_0 is accepted $\hat{B}_t^{(0)} = \tilde{B}_t^{(0)}$.

Next for $k = 1, \cdots, K$, screen $J_k = I_k \setminus I_{k-1} = [t - m_k, t - m_{k-1})$ and check for a change point.



$$T_{I,t} = \max_{B'', B'} \{L_{I''}(B'') + L_{I'}(B')\} - \max_{B} L_I(B), \quad (3)$$

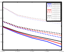
$$T_k = \max_{t \in J_k} T_{I,t} \begin{cases} \leq \eta_k & H_0 \text{ is not rejected: } \hat{B}_t^{(k)} = \tilde{B}_t^{(k)} \\ > \eta_k & H_0 \text{ is rejected, terminate} \end{cases} \quad (4)$$



LCP parameters

Set of intervals: $I_k = [t - m_k, t]$ with $m_k = m_0 a^k$.

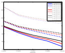
- The starting value m_0 should be sufficiently small to provide a reasonable local homogeneity.
- The coefficient $a > 1$ controls the increasing speed of the candidate intervals.



LCP parameters

Critical values $\{\eta_k\}$ are calculated under H_0.

- MC: generate homogeneous series $\mathbf{X}_t = (B^*)^{-1}\mathrm{IC}_t$.
- The final estimate $\hat{B} = \hat{B}_K$ depends on the critical values $\{\eta_k\}_{k=1}^{K}$.
- Small modeling bias: $\mathrm{E}_{B^*}|L(I_k, \tilde{B}_t^{(k)}, \hat{B})|^r \le \rho R_r(B^*)$,
  - B^* is the MLE over I_0.
  - The hyperparameter r specifies the loss function that measures the divergence of a time varying model to a static model.
  - The hyperparameter ρ is similar to the test level parameter.
  - Given the values of r and ρ, $R_r(B^*)$ can be computed straightforwardly.



Find ICs

Pre-whitening: use the Mahalanobis transformation $\tilde{\Sigma}_x^{-1/2}\mathbf{X}_t$.

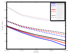
Quasi maximum likelihood estimation: for leptokurtic sources

$$\log f_j(x_j) = \alpha_1 - 2\log\cosh(x_j) = \alpha_1 - 2\log\{\frac{1}{2}(e^{x_j} + e^{-x_j})\}.$$

The first derivative of $\log f_j$:

$$g_j(x_j) = -2\tanh(x_j) = -\frac{2\{\exp(2x_j) - 1\}}{\exp(2x_j) + 1}, \quad \forall\, j = 1, \ldots, p,$$

A small misidentification in the density doesn't affect the consistency of the QMLE, Hyvärinen and Oja (1999).



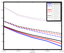
Data

$\mathbf{X}_t \in \mathbb{R}^{10}$: log returns of HD, HPQ, IBM, INTC, JNJ, JPM, KFT, KO, MCD and MMM over a stationary time period: 2010/01/14–2010/10/28. Fit IC_t under NIG assumption. Generate 10 independent univariate series, with 610 sample points for each series and with 1000 replications.

Homogeneity scenario (HOMO): $\mathbf{X}_t = A_t \mathrm{IC}_t$ with $A_t = I_{10}$,
Jump scenario (JPLM and JPEM): a sudden change after $t = 250$.
Smooth change scenario (SLEM): interval with changes: $[220, 380]$

Investigate detection power and location of the change point.

Analyze impact of the hyperparameters (r, ρ) on the LCP algorithm.

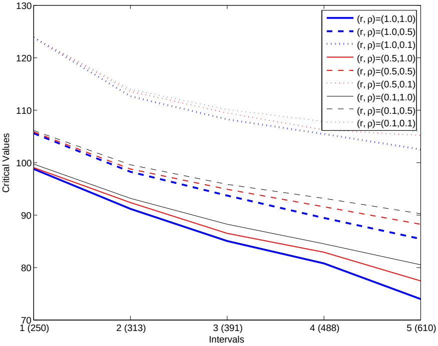


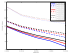
Critical values

Set of intervals: $m_k = m_0 a^k$ with $m_0 = 200$, $a = 1.25$ and $K = 5$

$$I_0 = 200,\ I_1 = 250,\ I_2 = 313,\ I_3 = 391\ I_4 = 488,\ I_5 = 610,$$

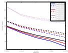
r and ρ are assigned to be $1, 0.5$ and 0.1





Result: rejection ratio and location

ρ	@	$r = 0.1$ l_1	l_2	l_3	l_4	$r = 0.5$ l_1	l_2	l_3	l_4	$r = 1.0$ l_1	l_2	l_3	l_4
0.1	HOMO	— 0.6 —				— 0.6 —				— 0.7 —			
	JPLF	–	–	100	–	–	–	100	–	–	–	100	–
	JPEM	–	–	99.2	0.8	–	–	99.4	0.6	–	–	99.4	0.6
	SLEM	–	5.9	93.1	1.0	–	6.8	92.4	0.8	–	7.9	91.3	0.8
0.5	HOMO	— 4.9 —				— 5.9 —				— 8.3 —			
	JPLF	0.1	–	99.9	–	0.1	0.1	99.8	–	0.1	0.1	99.8	–
	JPEM	–	0.1	99.5	0.4	–	0.2	99.5	0.3	–	0.2	99.6	0.2
	SLEM	0.2	32.4	67.4	–	0.2	34.4	65.4	–	0.2	36.1	63.7	–
1.0	HOMO	— 15.3 —				— 20.3 —				— 26.8 —			
	JPLF	0.2	0.4	99.4	–	0.2	0.4	99.4	–	0.2	0.7	99.1	–
	JPEM	–	0.4	99.5	0.1	–	0.6	99.4	–	–	0.8	99.2	–
	SLEM	0.2	49.5	50.3	–	0.2	52.6	47.2	–	0.4	56.4	43.2	–



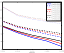
Data and experiments

$\mathbf{X}_t \in \mathbb{R}^{10}$: log returns of HD, HPQ, IBM, INTC, JNJ, JPM, KFT, KO, MCD and MMM.

The set of intervals: $m_k = m_0 a^k$ with $m_0 = 200$, $a = 1.25$ and $K = 5$.

The parameters $(r, \rho) = (0.5, 0.5)$ and $(r, \rho) = (0.1, 0.1)$ are considered respectively.

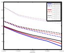
B^*: MLE over I_0 or identity matrix.



Data and experiments

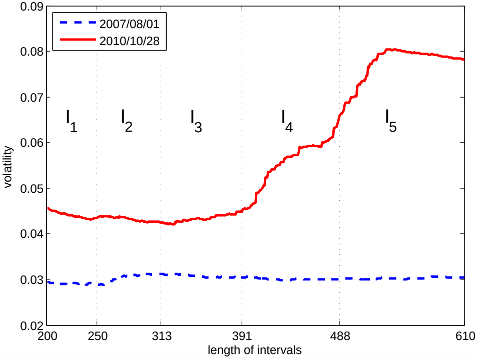
The first experiment considers the time interval 2005/03/01–2007/08/01, during which no influential economic or financial events occurred.
The second experiment considers the time interval 2008/05/30–2010/10/28, during which the stock market crash occurred in 2008.

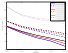
Does the proposed method detect intervals of local homogeneity? Can we identify an interval in a post-financial crisis world that indicates a relatively stationary period?



Empirical evidence

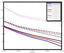
Realized volatility recursively computed for the 1st August 2007 and the 28th October 2010. The set of intervals with $m_0 = 200$, $a = 1.25$ and $K = 5$ is marked in the plot to highlight the underlying pattern across the intervals.





Results: CVs and test statistics

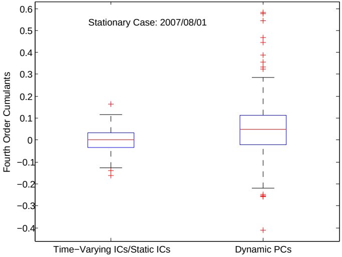
	2005/03/01-2007/08/01					2008/05/30-2010/10/28				
	CV				T_I	CV				T_I
(r, ρ)	$(0.5, 0.5)$		$(0.1, 0.1)$			$(0.5, 0.5)$		$(0.1, 0.1)$		
B^*	MLE	Identity	MLE	Identity		MLE	Identity	MLE	Identity	
I_1	107.23	102.84	122.37	120.89	74.36	108.87	105.85	126.51	123.74	69.81
I_2	98.40	98.45	117.43	113.21	76.62	101.71	98.67	116.86	113.95	81.97
I_3	93.15	92.35	112.30	108.44	66.86	96.32	94.92	113.91	110.05	**265.35**
I_4	89.64	88.81	109.53	105.57	77.52	92.59	91.57	111.18	107.80	**469.99**
I_5	86.28	85.74	106.82	103.01	72.79	88.72	88.21	108.99	105.85	**205.60**



Results: Independence under homogeneity

Fourth order cross-cumulant is used as a measure of statistical independence:

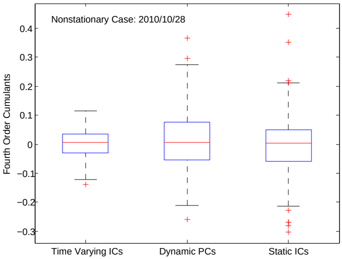
$$\mathrm{cum}(z_i, z_j, z_k, z_l) = \mathrm{E}(z_i z_j z_k z_l) - \mathrm{E}(z_i z_j)\,\mathrm{E}(z_k z_l) - \mathrm{E}(z_i z_k)\,\mathrm{E}(z_j z_l) - \mathrm{E}(z_i z_l)\,\mathrm{E}(z_j z_k),$$

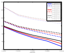


Results: Independence under inhomogeneity

Fourth order cross-cumulant is used as a measure of statistical independence:

$$\mathrm{cum}(z_i, z_j, z_k, z_l) = \mathrm{E}(z_i z_j z_k z_l) - \mathrm{E}(z_i z_j)\,\mathrm{E}(z_k z_l) - \mathrm{E}(z_i z_k)\,\mathrm{E}(z_j z_l) - \mathrm{E}(z_i z_l)\,\mathrm{E}(z_j z_k),$$





Conclusion

- Develop a time varying modeling for independent source extraction, ✓
- For each time point t, LCP approach helps to identify a "trust interval" $I_t = [t - m_t, t)$, over which the linear filter A_t (or B_t) is approximately const., ✓
- Simulation study and real data analysis show that the TVICA method is data driven. It provides a stable performance for different parameter selection and works well, ✓
- A universal statistical MDA method that is applicable for non-Gaussian and non-stationary financial time series.

Appendix

HD: The Home Depot

HPQ: Hewlett-Packard

IBM: International Business Machines

INTC: Intel

JNJ: Johnson & Johnson

JPM: JPMorgan Chase

KFT: Kraft Foods

KO: Coca-Cola

MCD: McDonald's

MMM: 3M

References

Back, A. and Weigend, A.
A first application of independent component analysis to extracting structure from stock returns
International Journal of Neural Systems, 8, 473-484, 1998.

Baillie, R. T. and Morana, C.
Modelling long memory and structural breaks in conditional variances: An adaptive FIGARCH approach
Journal of Economics Dynamics and Control, 33, 1577-1592, 2009.

References

Barndorff-Nielsen, O.
Normal inverse gaussian distributions and stochastic volatility modelling
Scandinavian Journal of Statistics, 24, 1-13, 1997.

Ray-Bing Chen, Mei-Hui Guo, Wolfgang K. Hardle, and Shih-Feng Huang
COPICA - Independent Component Analysis via Copula Techniques
submitted, 2011.

Chen, Y. and Härdle, W. and Spokoiny, V.
GHICA risk analysis with GH distributions and independent components
Journal of Empirical Finance, 17, 255-269, 2010.

References

Hamilton, J. D. and Susmel, R.
Autoregressive conditional heteroskedasticity and changes in regime
Journal of Econometrics, 64, 307-333, 1994.

Härdle W. and Panov, V. and Spokoiny, V. and Wang, W.
Modern Mathematical Statistics, Exercises and Solutions
Springer-Verlag Berlin Heidelberg New York, 2012.

Härdle, W. and Simar, L.
Applied Multivariate Statistical Analysis, 4th edn
Springer-Verlag Berlin Heidelberg New York, 2012.

References

Hyvärinen, A. and Karhunen, J. and Oja, E.
Independent Component Analysis
John Wiley & Sons, Inc., 2001.

Hyvärinen, A. and Oja, E.
A fast fixed-point algorithm for independent component analysis
Neural Computation, 9, 1483-1492, 1997.

Hyvärinen, A. and Oja, E.
Independent component analysis: Algorithm and applications
Neural Networks, 13, 411-430, 1999.

References

Jolliffe, I. T.
Principal component analysis
Springer-Verlag Berlin Heidelberg New York, 2002.

Kouontchou, P. and Maillet, B.
ICA-based high frequency VaR for risk management
ESANN'2007 proceedings - European Symposium on Artificial Neural Networks, Bruges, Belgium, 2007.

Lanne, M.
A mixture multiplicative error model for realized volatility
Journal of Financial Econometrics, 4, 594-616, 2006.

References

Matteson, D. S. and Tsay, R. S.
Independent component analysis for multivariate financial time series
submitted, 2009.

Mercurio, D. and Spokoiny, V.
Statistical inference for time-inhomogeneous volatility models
Ann. Statist., 12, 577-602, 2004.

Scharth, M. and Medeiros, M. C.
Asymmetric effects and long memory in the volatility of dow jones stocks
International Journal of Forecasting, 25, 304-327, 2009.

References

So, M. K. P. and Lam, K. and Li, W. K.
A stochastic volatility model with markov switching
Journal of Business & Economic Statistics, 16, 244-253, 1998.

Spokoiny, V.
Mathematical Statistics
Springer-Verlag Berlin Heidelberg New York, 2011.

Wu, E. and Yu, P. and Li, W.
Value at risk estimation using independent component analysis-generalized autoregressive conditional heteroscedasticity (ICA-GARCH) models
International Journal of Neural Systems, 16, 371-382, 2006.