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

National Central University

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.

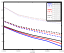
- Principal component analysis: $\mathbf{X}_t = \Gamma \times \mathrm{PC}_t$,
- Factor analysis: $\mathbf{X}_t = \Gamma \Lambda^{1/2} F_t + U_t$, where Γ is the eigenvector corresponding to the nonzero eigenvalues Λ of the covariance of $\mathbf{X}$, see Jolliffe (2002), Härdle and Simar (2011)

Under Gaussianity, cross-uncorrelated 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.

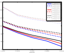


Independent Component Analysis (ICA)

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

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

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



How to find ICs?

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

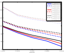
Jones and Sibson (1987): projection pursuit

Hyvärinen and Oja (1997): FastICA

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

The observed series and as well the ICs are assumed to be stationary. The filter B is constant over time.

Fact: The ever occurring turbulences in financial markets.



Demonstration

Log returns of HD, HPQ and IBM.

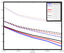
$$\mathbf{X}_t = \begin{cases} A_1 \mathrm{IC}_t & t \in [1, 300] \\ A_2 \mathrm{IC}_t & t \in [301, 600] \end{cases}$$

where IC_t are NIG distributed, see Barndorff-Nielson (1997).
The theoretical ICA filters are:

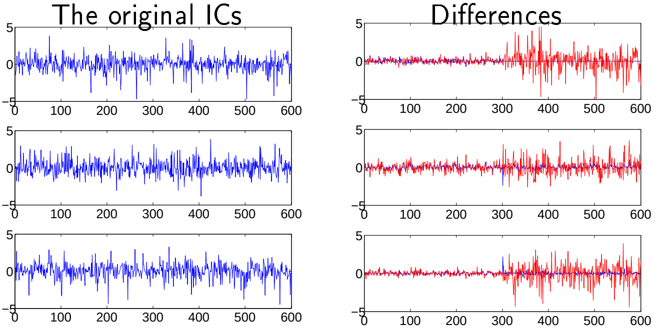
$$A_1 = \begin{pmatrix} 0.0006 & 0.0130 & 0.0062 \\ 0.0038 & 0.0027 & 0.0130 \\ 0.0079 & 0.0059 & 0.0048 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} -0.0001 & 0.0008 & 0.0053 \\ 0.0070 & 0.0019 & 0.0016 \\ 0.0001 & 0.0042 & 0.0011 \end{pmatrix},$$

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

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

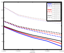


Demonstration (Continued)



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 multiplicative error specifications,
- Need a globally given mechanism for this time variation.

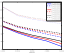
Mercurio and Spokoiny (2004) use a local change point (LCP) approach: completely data driven approach.

TVICA

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

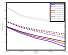
$$\mathbf{X}_t = A_t \mathrm{IC}_t$$

- 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 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) with different magnitudes and abrupt or smooth types.



Outline

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

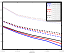


TVICA

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. $A_s \approx A, \; \forall \; s \in I_t$.



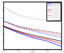
Estimation: under homogeneity

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:

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

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



Estimation: under local homogeneity

Local homog.: B_t does not deviate much from a const. filter B^*.

Small modeling bias: divergence of a time varying model (local homogeneity) to a static model (homogeneity) is small, see Spokoiny (2009).

For $r, \rho > 0$, the fitted log likelihood with $B_t = B^*$ satisfies:

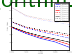
$$\mathsf{E}_{B^*}|L_{I_k}(\tilde{B}_k, B^*)|^r \leq \rho R_r(B^*), \quad t \in I_K \tag{2}$$

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

Goal: For any time point t, and given a family of nested intervals, $I_0 \subset I_1 \subset \cdots \subset I_{K-1} \subset I_K$, LCP method attempts to find the longest interval of local homogeneity among them.

The longer the length of intervals, the smaller the variance of the estimator (under homogeneity) but the higher the bias.

The identification of the trust interval is done via a sequential testing algorithm.



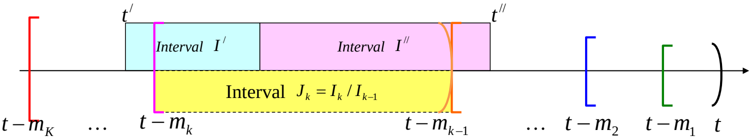
LCP algorithms

H_0 : I_k is a local homogeneous interval *given that the test was not rejected at* $k-1$

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

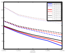
Next for $k = 1, \cdots, K$, the procedure is to sequentially screen

$J_k = I_k \setminus I_{k-1} = [t - m_k, t - m_{k-1})$ and check it for any possible change point.



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

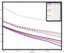
$$T_k = \max_{t \in J_k} T_{I,t} \begin{cases} \leq \eta_k & H_0 \text{ is not rejected: } \hat{B}_t^{(k)} = \tilde{B}_t^{(k)} \\ \geq \eta_k & H_0 \text{ is rejected, terminate} \end{cases} \quad (4)$$



LCP parameters

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

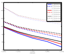
- The starting value m_0 should be sufficiently small to provide a reasonable local homogeneity.
- The coefficient $a > 1$ controls the increasing speed of the candidate intervals.



LCP parameters

Critical values $\{\eta_k\}$ are calculated under the null of homog.

- 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: $\mathrm{E}_{B^*}|L_{I_K}(\tilde{B}_K, \hat{B})|^r \le \rho R_r(B^*)$,
  - 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.



Finding ICs

Pre-whitening: the Mahalanobis transformation $\tilde{\Sigma}_x^{-1/2}\mathbf{X}_t$, where $\tilde{\Sigma}_x$ is the sample covariance based on the available data.

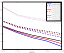
Quasi maximum likelihood estimation: for leptokurtic sources

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

where α_1 is a normalizing constant to make this function a pdf. 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 ML estimator, see Hyvärinen et al. (2001).



Data

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

Homogeneity scenario: $\mathbf{X}_t = A_t \mathrm{IC}_t$ with $A_t = A$ is an identity matrix,
Change point scenario: $\mathbf{X}_t = A_t \mathrm{IC}_t$ with a_{21} changes from 0 to 3 at $t = 251$.

Investigate detection power and locate the position of the change point properly,

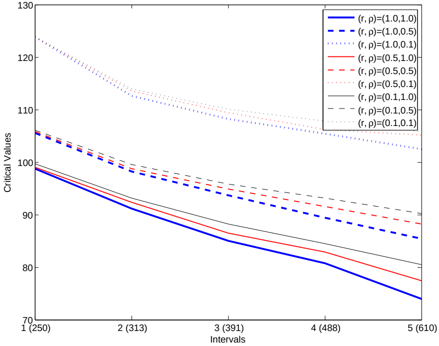
Analyze impact of the hyperparameters (r, ρ) on the LCP algorithm.

Critical values

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

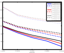
$$I_0 = 200,\ I_1 = 250,\ I_2 = 313,\ I_3 = 391\ I_4 = 488,\ I_5 = 610,$$

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



Result: rejection ratio and location

Parameter (r, ρ)	Homogeneity	Change Point at $t = 251$ $I_3/I_2 = [219, 298]$	
$(1.0, 1.0)$	0.268	$@I_1 = 0.001$ $@I_2 = 0.008$, $@I_3 = 0.991$	sum $= 1$
$(1.0, 0.5)$	0.083	$@I_1 = 0.001$ $@I_2 = 0.002$, $@I_3 = 0.997$	sum $= 1$
$(1.0, 0.1)$	0.007	$@I_1 = 0.000$ $@I_2 = 0.000$, $@I_3 = 1.000$	sum $= 1$
$(0.5, 1.0)$	0.203	$@I_1 = 0.001$ $@I_2 = 0.005$, $@I_3 = 0.994$	sum $= 1$
$(0.5, 0.5)$	0.059	$@I_1 = 0.001$ $@I_2 = 0.001$, $@I_3 = 0.998$	sum $= 1$
$(0.5, 0.1)$	0.006	$@I_1 = 0.000$ $@I_2 = 0.000$, $@I_3 = 1.000$	sum $= 1$
$(0.1, 1.0)$	0.153	$@I_1 = 0.001$ $@I_2 = 0.004$, $@I_3 = 0.995$	sum $= 1$
$(0.1, 0.5)$	0.049	$@I_1 = 0.001$ $@I_2 = 0.001$, $@I_3 = 0.998$	sum $= 1$
$(0.1, 1.0)$	0.006	$@I_1 = 0.000$ $@I_2 = 0.000$, $@I_3 = 1.000$	sum $= 1$



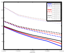
Data and experiments

$\mathbf{X}_t \in \mathbb{R}^6$: log returns of HD, HPQ, IBM, JNJ and JPM.
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.

The first experiment considers the time interval 2007/03/30–2009/08/31, during which the stock market crash occurred in 2008.
The second experiment considers the time interval 2004/07/30–2006/12/29, during which no influential economic or financial events occurred.

Does the proposed method detect intervals of local homogeneity?

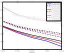
Can we identify an interval in a post-financial crisis world that indicates a relatively stationary period?



Results: CVs and test statistics

	2007/03/30-2009/08/31			2004/07/30-2006/12/29		
	CV		T_I	CV		T_I
(r, ρ)	$(0.5, 0.5)$	$(0.1, 0.1)$		$(0.5, 0.5)$	$(0.1, 0.1)$	
I_1	70.10(69.45)	83.38(85.35)	43.77	59.24(60.03)	71.71(74.22)	20.94
I_2	63.55(62.76)	80.00(79.16)	**83.86**	53.81(54.74)	66.99(70.29)	30.79
I_3	59.49(58.73)	75.52(76.73)	**76.14**	50.93(50.94)	64.95(66.30)	41.60
I_4	56.62(56.02)	73.85(74.47)	**128.24**	48.44(48.48)	63.59(64.48)	34.71
I_5	53.35(53.22)	73.17(72.42)	**188.71**	46.25(46.27)	61.48(62.91)	38.94

For experiment 1: the interval $I_1 = [2008/09/03, 2009/08/31]$.



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.