On a Matrix Factor Model

Feiyu Jiang[1], JJiayuan Zhou[2], Ke Zhu[2], Wai Keung Li[2,3]

[1]Centre for Statistical Science, Tsinghau University
[2]Department of Statistics and Actuarial Science
The University of Hong Kong
[3]Department of Mathematics and Information Technology
Education University of Hong Kong

The authors thank the HK RGC GRF 17303315 for partial support

Outline

Realized Covariance Matrices

- Suppose we have intraday price data of n assets, $X_{s,t}$
  - Here t means the t^{th} day, and s is the time within the day. $0 \leq s \leq m$ where m represents time in each day.
- Specify a sampling frequency τ, we can sample several time points $(s_k)_{k=0,1,\ldots,\tau}$
- Then the RCOV matrix for time t, denoted by Y_t can be calculated as

$$Y_t = \sum_{k=1}^{\tau} r_{k,t} r'_{k,t}; \quad \text{where } r_{k,t} = X_{s_k,t} - X_{s_{k-1},t}.$$

Realized Covariance Matrices

- Estimation of RCOVs:
  - Wang et al. (2010) introduced the ARVM estimation.
  - in Tao et al. (2011) and Tao et al. (2013): thresholded estimator
- RCOV matrices can approximate the underlying daily return covariance structure.

Model Setting

- We focus on modeling and forecasting of the RCOV matrices.
- The conditonal volatility model that has been widely used is the traditional GARCH and ARCH models. Here let Y_t denote the realized covariance matrix at day t
- $$Y_t = \Sigma_t^{1/2} \Delta_t \Sigma_t^{1/2}$$
- Here Δ_t is a sequence of i.i.d. positive definite random matrices. Σ_t is a sequence of positive definite matrices containnig the parameters.
- Let $\mathcal{G}_t$ represents the sigma algebra generated by $(Y_t, Y_{t-1}, Y_{t-2}, \ldots)$. We assume $\Sigma_t \in \mathcal{G}_{t-1}$ with $E(Y_t|\mathcal{G}_{t-1}) = \Sigma_t$

BEKK

- BEKK:
  - $$\Sigma_t = \Omega + \sum_{i=1}^{P}\sum_{k=1}^{K} A_{ki} Y_{t-i} A'_{ki} + \sum_{j=1}^{Q}\sum_{k=1}^{K} B_{kj}\Sigma_{t-j}B'_{kj}$$
- Ω and the initial states of Y's and Σ's are restricted to be positive definite. A and B are $n \times n$ matrices. P, Q, K are non-negative integers.
- Stationary condition and its proof are motivated by results in Boussama et al. (2011).

Stationary Condition

- Denote $M = max(P, Q)$, $A_i^* = \sum_{k=1}^{K} A_{ik}^{\otimes 2}$, $B_i^* = \sum_{k=1}^{K} B_{ik}^{\otimes 2}$

Theorem

Suppose that Δ_t is an i.i.d. $n \times n$ positive definite random matrix with $E\|\Delta_t\| < \infty$, and

(H1) *the distribution of Δ_1, denoted by Γ, is absolute continuous with respect to the Lebesgue measure;*

(H2) *the point I_n is in the interior of the support of Γ;*

(H3) $\rho\left(\sum_{i=1}^{M}(A_i^* + B_i^*)\right) < 1.$

Then, Y_t in model is strictly stationary with $E\|Y_t\| < \infty$. Moreover, Y_t is positive Harris recurrent and geometrically ergodic.

Model with Matrix-F Innovation

- If we assume the distribution of these random vectors, the short term return vectors, follow from multivariate normal distribution. Then we end up with Wishart type of distribution of Δ_t. See WAR model (2006) and CAW model (2009).
- To deal with heavy tailed cases exist in real application, we consider the matrix-F distribution.
- See Opschoor et al. (2017) for a somewhat different approach.

Model with Matrix-F Innovation

- The matrix-F innovation: assume Δ_t in previous setting follows the $F(\nu_1, \nu_2, \frac{\nu_2 - n - 1}{\nu_1} I_n)$ distribution, where the density of $F(\nu, \Sigma)$ is

$$f(x; \nu, \Sigma) = \Lambda(\nu_1, \nu_2) \times \frac{|\Sigma|^{-\nu_1/2} |x|^{(\nu_1 - n - 1)/2}}{|I_n + \Sigma^{-1} x|^{(\nu_1 + \nu_2)/2}}, \quad \text{for } x \in \mathbb{R}^+_{n \times n},$$

where $\Lambda(\nu_1, \nu_2) = \frac{\Gamma_n((\nu_1+\nu_2)/2)}{\Gamma_n(\nu_1/2)\Gamma(\nu_2/2)}$ with $\Gamma_n(x) = \pi^{n(n-1)/4} \prod_{i=1}^{n} \Gamma(x + (1 - i)/2)$; and $\mathbb{R}^+_{n \times n}$ represents the space of positive definite $n \times n$ real matrices.

Model with Matrix-F Innovation

- In a constructive manner, Δ_t can be generated by products of Wishart matrices.

$$\Delta_t = \frac{\nu_2 - n - 1}{\nu_1} L_t^{1/2} R_t^{-1} L_t^{1/2}$$

  where $L_t \sim Wishart(\nu_1, I_n)$ and $R_t \sim Wishart(\nu_2, I_n)$ are two independent Wishart processes.
- We can see when $\nu_2 \to \infty$, $\Delta_t \xrightarrow{d} Wishart(\nu_1, \frac{1}{\nu_1} I_n)$. Which is the CAW model.
- We call the proposed model with matrix-F likelihood as CBF model, short for conditional BEKK matrix-F model.

MLE

Theorem

Suppose that Y_t is strictly stationary, the model is identifiable, and $E\|Y_t\| < \infty$. Then, $\hat{\theta} \xrightarrow{a.s.} \theta_0$ as $T \to \infty$.

Theorem

Let l_t represent the negative of the log likelihood. Suppose that Y_t is strictly stationary, the model is identifiable, and $E\|Y_t\|^3 < \infty$. And suppose

$$\mathcal{O} = -E\left(\frac{\partial^2 l_t(\theta_0)}{\partial\theta\partial\theta'}\right) \text{ is invertible.}$$

Then $\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, \mathcal{O}^{-1})$ as $T \to \infty$.

VTMLE

- We assume that Y_t is strictly stationary with a finite mean $S = E(Y_t)$. By taking expectation on both sides of the BEKK regression we end up with

$$\Omega = S - \sum_{i=1}^{P}\sum_{k=1}^{K} A_{ki} S A'_{ki} - \sum_{j=1}^{Q}\sum_{k=1}^{K} B_{kj} S B'_{kj}$$

-

$$\Sigma_t = S - \sum_{i=1}^{P}\sum_{k=1}^{K} A_{ki} S A'_{ki} - \sum_{j=1}^{Q}\sum_{k=1}^{K} B_{kj} S B'_{kj} + \sum_{i=1}^{P}\sum_{k=1}^{K} A_{ki} Y_{t-i} A'_{ki} + \sum_{j=1}^{Q}\sum_{k=1}^{K} B_{kj} \Sigma_{t-j} B'_{kj}.$$

VTMLE

- Variance target, as discussed in Pedersen and Rahbek (2014), is essentially a two-step estimation
  - Replace S with the sample mean of observed Y_t
  - Optimize likelihood with respect to A's and B's

Theorem

Suppose that Y_t is strictly stationary, the model is identifiable, and $E\|Y_t\| < \infty$. Then, $\hat{\theta}_v \xrightarrow{a.s.} \theta_{v0}$ as $T \to \infty$.

- To study the asymptotics of VTMLE, we seperate θ_v as $\theta_v = (s', \nu', u')'$ where $\nu = (\nu_1, \nu_2)'$, $s = vec(S)$, and u represents the stacking of A's and B's, and we write $\zeta = (\nu', u')'$ as the part of parameters needed in optimization.

VTMLE

Theorem

Suppose that Y_t is strictly stationary, the model is identifiable, $E\|Y_t\|^3 < \infty$, and

$$J_1 = -E\left[\frac{\partial^2 l_{vt}(\theta_{v0})}{\partial\zeta\partial\zeta'}\right] \text{ is invertible.}$$

Then $\sqrt{T}(\hat{\theta}_v - \theta_{v0}) \xrightarrow{d} N(0, \mathcal{O}_v)$ *as* $T \to \infty$
where $\mathcal{O}_v = \begin{pmatrix} I_{n^2} & 0 \\ -J_1^{-1}J_2 & J_1^{-1} \end{pmatrix} E\left[w_t w_t'\right] \begin{pmatrix} I_{n^2} & 0 \\ -J_1^{-1}J_2 & J_1^{-1} \end{pmatrix}'$ *with*
$J_2 = -E\left[\frac{\partial^2 l_{vt}(\theta_{v0})}{\partial\zeta\partial s'}\right]$, $w_t = \begin{pmatrix} \Psi(u_0) vec\left(Y_t - \Sigma_t\right) \\ \partial l_{vt}(\theta_{v0})/\partial\zeta \end{pmatrix}$ *and*
$\Psi(u) = \left(I_{n^2} - \sum_{i=1}^M A_i^* - \sum_{i=1}^M B_i^*\right)^{-1}\left(I_{n^2} - \sum_{i=1}^M B_i^*\right)$

Matrix Factor

- Motivated by Tao et al. (2011); Shen, Yao and Li (2015)
- The matrix factor model accounts for the series of big $n \times n$ matrices by a lower rank dynamic, a series of $r \times r$ matrices and $n \times r$ factor loading.
- More specifically, $Y_t^* = FY_{ft}^* F' + Y_0^*$ where Y_{ft}^* is an $r \times r$ positive definite factor. F is an $n \times r$, factor loading matrix and is orthonormal with $F'F = I_r$, Y_0^* is an $n \times n$ constant matrix

Matrix Factor

- Let Y_t be the threshold MSRVM estimator of Y_t^*
- Define $\bar{Y} = \frac{1}{T}\sum_{t=1}^{T} Y_t$, $\bar{S} = \frac{1}{T}\sum_{t=1}^{T}\left\{Y_t - \bar{Y}\right\}^2$.
- Factor loading can be found by $\tilde{Y}_{ft} = \tilde{F}' Y_t \tilde{F}$, $\tilde{Y}_0^* = \bar{Y} - \tilde{F}\tilde{F}'\bar{Y}\tilde{F}\tilde{F}'$ and $\tilde{F} = (\tilde{f}_1, \cdots, \tilde{f}_r)$, where $\tilde{f}_1, \cdots, \tilde{f}_r$ are the eigenvectors of $\bar{S}$ corresponding to its r largest eigenvalues.
- How do we select r?
    - Eigen value sequential ratios
    - Eigen value over summation

Matrix Factor

Assumption-1 All row vectors of F' and Y_0^* satisfy the sparsity condition below. For an n-dimensional vector $(x_1, \cdots, x_n)$, we say it is sparse if it satisfies

$$\sum_{i=1}^{n} |x_i|^{\delta_*} \leq U\pi(n),$$

where $\delta_* \in [0, 1)$, U is a positive constant, and $\pi(n)$ is a deterministic function of n that grows slowly in n with typical examples $\pi(n) = 1$ or $\log(n)$.

Matrix Factor

Assumption-2 The factor model has r fixed factors, and matrices Y_0^* and Y_{ft}^* satisfy $\|Y_0^*\| < \infty$ and $\max_{1 \leq t \leq T} \|Y_{ft,jj}^*\| = O_p(B(T))$ for $j = 1, 2, \cdots, r$, where $Y_{ft,jj}^*$ is the j-th diagonal entry of Y_{ft}^*, and $1 \leq B(T) = o(T)$.

Matrix Factor

Assumption-3 $\max_{1\leq t\leq T} \|Y_t^* - Y_t\| = O_p(A(n, m, T))$ for some rate function $A(n, m, T)$ such that $A(n, m, T)B^5(T) = o(1)$. where m is the average intra-day sample size across all assets and all days.

Tao et al. 2011 suggested that we can select $A(n, m, T) = \pi(n)[e_m(n^2 T)^{1/\beta}]^{1-\delta_*} \log T$, with $e_m \sim m^{-1/4}$ and $B(T) = \log T$, so that $A(n, m, T)B^5(T) = o(1)$ for large β.

Matrix Factor

Theorem

Suppose that Assumptions 1-3 hold. Then, as n, m, T go to infinity,

$$\text{(i)}\quad F'\tilde{F} - I_r = O_p(A(n,m,T)B(T)),$$

$$\text{(ii)}\quad \tilde{Y}_{ft} - Y_{ft} = O_p(A^{1/2}(n,m,T)B^{3/2}(T)),$$

where $Y_{ft} = Y^*_{ft} + F'Y^*_0F$, and $F = (f_1, \cdots, f_r)$ with $f_1, \cdots, f_r$ being the eigenvectors of $\bar{S}^*$ corresponding to its r largest eigenvalues.

Matrix Factor

- As Y_{ft} is not observable, and what we should estimate are based on $\tilde{Y}_{ft}$. Consider the MLE case denoted by:

$$\hat{\theta}_{1f} = (\hat{\gamma}'_{1f}, \hat{\nu}'_{1f})' = \arg\min_{\theta \in \Theta} \tilde{L}_f(\theta).$$

And the ideal MLE denoted by: $\hat{\theta}_{2f} = (\hat{\gamma}'_{2f}, \hat{\nu}'_{2f})' = \arg\min_{\theta \in \Theta} L_f(\theta)$,

Theorem

Suppose that the conditions 1-3 and conditions for asymptotics of MLE hold. Then, as n, m, T go to infinity,
$\hat{\theta}_{1f} - \hat{\theta}_{2f} = O_p(B(T)/T) + O_p(A^{1/2}(n, m, T)B^{5/2}(T))$.

Matrix Factor

- For the VTMLE, consider the estimators

$$\hat{s}_{1fv} = \frac{1}{T}\sum_{t=1}^{T} \tilde{Y}_{ft},$$
$$\hat{\zeta}_{1fv} = (\hat{u}'_{1fv}, \hat{\nu}'_{1fv})' = \operatorname*{arg\,min}_{\zeta \in \Theta_u \times \Theta_\nu} \tilde{L}_{fv}(\hat{s}_{1fv}, \zeta),$$

While the ideal estimators

$$\hat{s}_{2fv} = \frac{1}{T}\sum_{t=1}^{T} Y_{ft},$$
$$\hat{\zeta}_{2fv} = (\hat{u}'_{2fv}, \hat{\nu}'_{2fv})' = \operatorname*{arg\,min}_{\zeta \in \Theta_u \times \Theta_\nu} L_{fv}(\hat{s}_{2fv}, \zeta),$$

Matrix Factor

Theorem

Suppose that the conditions 1-3 and conditions for VTMLE asymptotics hold. Then, as n, m, T go to infinity,

$$\text{(i)}\quad \hat{s}_{1fv} - \hat{s}_{2fv} = O_p(A^{1/2}(n,m,T)B^{3/2}(T)),$$

$$\text{(ii)}\quad \hat{\zeta}_{1fv} - \hat{\zeta}_{2fv} = O_p(B(T)/T) + O_p(A^{1/2}(n,m,T)B^{5/2}(T)).$$

Diagonal Coefficients

- We can further propose special structures for the coefficients A's and B's.
- Most used assumptions are A's and B's are diagonal matrices, which propose the account that series dependency are more generally in than across pannel. See Engle et al. (1995) for details.

Portmanteau Tests for Model Diagnostic Checks

- Define $\mathfrak{z}_t(\gamma) = vec\left(\Sigma_t^{-1/2}(\gamma) Y_t \Sigma_t^{-1/2}(\gamma) - I_n\right)$ as the residual of the CBF model.
- Define $\mathrm{b}_{t,j}(\gamma) = \mathfrak{z}_t'(\gamma)\mathfrak{z}_{t-j}(\gamma)$ as the process of inner product of residuals.
- Concatenate the inner product as

$$\mathcal{V}_l(\gamma) = \frac{1}{T}\sum_{t=l+1}^{T} \begin{pmatrix} \mathrm{b}_{t,1}(\gamma) \\ \mathrm{b}_{t,2}(\gamma) \\ \vdots \\ \mathrm{b}_{t,l}(\gamma) \end{pmatrix}.$$

- By the following theorem we construct the asymptotic normality of $\mathcal{V}_l(\lambda)$

Portmanteau Tests for Model Diagnostic Checks

Theorem

Under assumptions of the MLE asymptotics and with $E\|Y_t\|^4 < \infty$,
$\sqrt{T}\mathcal{V}_l(\hat{\gamma}) \xrightarrow{d} N\left(\ 0, \quad \mathrm{V}\ \right)$ *where* $\mathrm{V} = \left[\ I_l \quad \mathfrak{R}\ \right] E\left(\mathrm{e}_t \mathrm{e}_t'\right) \left[\ I_l \quad \mathfrak{R}\ \right]'$,

$$\mathfrak{R} = E\begin{pmatrix} \mathfrak{z}_{t-1}'(\gamma_0)\left(\partial \mathfrak{z}_t(\gamma_0)/\partial\theta'\right) \\ \mathfrak{z}_{t-2}'(\gamma_0)\left(\partial \mathfrak{z}_t(\gamma_0)/\partial\theta'\right) \\ \vdots \\ \mathfrak{z}_{t-l}'(\gamma_0)\left(\partial \mathfrak{z}_t(\gamma_0)/\partial\theta'\right) \end{pmatrix} \times \mathcal{O}^{-1} \text{ and } \mathrm{e}_t = \begin{pmatrix} \mathrm{b}_{t,j}(\gamma_0) \\ \mathrm{b}_{t,j}(\gamma_0) \\ \vdots \\ \mathrm{b}_{t,j}(\gamma_0) \\ \partial l_t(\theta_0)/\partial\theta \end{pmatrix}.$$

Therefore we have $\prod = T\mathcal{V}_l(\hat{\gamma})'\mathrm{V}^{-1}\mathcal{V}_l(\hat{\gamma}) \xrightarrow{d} \chi_l^2$ and a large $\prod$ indicates rejection of adquate fit.

Portmanteau Tests for Model Diagnostic Checks

Similarly, for VTCBF we have:

Theorem

Under assumptions of the VTMLE asymptotics and with $E\|Y_t\|^4 < \infty$,
$\sqrt{T}\mathcal{V}_l(\hat{\delta}_v) \xrightarrow{d} N\left(\ 0,\ \ \mathrm{V}_v\ \right)$ *where* $\mathrm{V}_v = \left[\ I_l \quad \mathfrak{R}_v\ \right] E\left(\mathrm{u}_t\mathrm{u}_t'\right)\left[\ I_l \quad \mathfrak{R}_v\ \right]'$

$$\mathfrak{R}_v = E\begin{pmatrix} \mathfrak{z}_{t-1}'\left(\delta_{v0}\right)\left(\partial\mathfrak{z}_t\left(\delta_{v0}\right)/\partial\theta'\right) \\ \mathfrak{z}_{t-2}'\left(\delta_{v0}\right)\left(\partial\mathfrak{z}_t\left(\delta_{v0}\right)/\partial\theta'\right) \\ \vdots \\ \mathfrak{z}_{t-l}'\left(\delta_{v0}\right)\left(\partial\mathfrak{z}_t\left(\delta_{v0}\right)/\partial\theta'\right) \end{pmatrix} \times \begin{pmatrix} I_{n^2} & 0 \\ -J_1^{-1}J_2 & J_1^{-1} \end{pmatrix}$$

$$\text{and}\ \ \mathrm{u}_t = \begin{pmatrix} \mathrm{b}_{t,j}\left(\delta_{v0}\right) \\ \mathrm{b}_{t,j}\left(\delta_{v0}\right) \\ \vdots \\ \mathrm{b}_{t,j}\left(\delta_{v0}\right) \\ \omega_t(\delta_{v0}) \end{pmatrix}$$

Therefore we have $\prod_v = T\mathcal{V}_l(\hat{\delta}_v)'\mathrm{V}_v^{-1}\mathcal{V}_l(\hat{\delta}_v) \xrightarrow{d} \chi_l^2$ and a large $\prod_v$ indicates rejection of adquate fit.

Simulation 1

- $P = 1$, $Q = 1$, $K = 1$.
- $Y_t = \Sigma_t^{1/2} \Delta_t \Sigma_t^{1/2}$ with $\Sigma_t = \Omega + A Y_{t-1} A' + B \Sigma_{t-1} B'$,
-

$$A = \begin{bmatrix} 0.4 & 0 & 0 \\ 0 & 0.55 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}, \; B = \begin{bmatrix} 0.4 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.5 \end{bmatrix},$$

$$\Omega = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.2 & 0.5 & 0.25 \\ 0.3 & 0.25 & 0.5 \end{bmatrix},$$

MLE results

		ν_1	ν_2	A_{ii}			B_{ii}			$vech(\Omega)'$					
True		10	8	0.40	0.55	0.50	0.40	0.30	0.50	0.50	0.20	0.30	0.50	0.25	0.50
T=1000	Bias	0.0478	0.0188	-0.0019	-0.0022	-0.0019	-0.0116	-0.0109	-0.0087	-0.0041	0.0028	0.0041	-0.0013	0.0041	0.0042
	SD	0.4316	0.2684	0.0267	0.0274	0.0248	0.1243	0.0986	0.0760	0.0649	0.0190	0.0352	0.0431	0.0252	0.0634
	AD	0.4174	0.2590	0.0263	0.0265	0.0242	0.1135	0.0916	0.0552	0.0646	0.0186	0.0328	0.0441	0.0244	0.0507
T=2000	Bias	0.0202	0.0112	-0.0011	-0.0009	-0.0008	-0.0055	-0.0036	-0.0031	-0.0016	0.0011	0.0020	-0.0014	0.0013	0.0013
	SD	0.2911	0.1843	0.0187	0.0189	0.0173	0.0830	0.0643	0.0485	0.0447	0.0129	0.0238	0.0295	0.0169	0.0420
	AD	0.2872	0.1782	0.0181	0.0183	0.0167	0.0781	0.0631	0.0380	0.0447	0.0128	0.0226	0.0304	0.0168	0.0348
True		15	10	0.40	0.55	0.50	0.40	0.30	0.50	0.50	0.20	0.30	0.50	0.25	0.50
T=1000	Bias	0.1266	0.0220	-0.0018	-0.0010	-0.0015	-0.0186	-0.0134	-0.0098	-0.0017	0.0036	0.0063	-0.0015	0.0046	0.0050
	SD	0.8858	0.3839	0.0260	0.0249	0.0241	0.1324	0.1040	0.0756	0.0689	0.0191	0.0368	0.0446	0.0252	0.0628
	AD	0.8301	0.3113	0.0253	0.0251	0.0231	0.1172	0.0936	0.0667	0.0661	0.0177	0.0339	0.0434	0.0235	0.0599
T=2000	Bias	0.0659	0.0114	-0.0009	-0.0007	-0.0010	-0.0076	-0.0051	-0.0038	-0.0013	-0.0001	0.0026	-0.0012	0.0017	0.0020
	SD	0.5911	0.2538	0.0179	0.0169	0.0168	0.0897	0.0698	0.0491	0.0470	0.0129	0.0252	0.0301	0.0170	0.0427
	AD	0.5870	0.2216	0.0175	0.0173	0.0159	0.0806	0.0644	0.0457	0.0452	0.0122	0.0233	0.0298	0.0162	0.0412
True		20	10	0.40	0.55	0.50	0.40	0.30	0.50	0.50	0.20	0.30	0.50	0.25	0.50
T=1000	Bias	0.2134	0.0228	-0.0014	-0.0013	-0.0005	-0.0131	-0.0122	-0.0102	-0.004	0.0027	0.0046	-0.0019	0.0037	0.0039
	SD	1.5644	0.3677	0.0254	0.0245	0.0233	0.1276	0.1001	0.0782	0.0664	0.0179	0.0355	0.0422	0.0241	0.0634
	AD	1.5297	0.3664	0.0258	0.0247	0.0234	0.1118	0.0944	0.0669	0.0630	0.0172	0.0324	0.0440	0.0236	0.0595
T=2000	Bias	0.0671	0.0143	-0.0004	-0.0014	-0.0005	-0.0066	-0.0053	-0.0062	-0.0013	0.0016	0.0025	-0.0002	0.0022	0.0037
	SD	1.0333	0.2442	0.0166	0.0173	0.0161	0.0836	0.0683	0.0485	0.0455	0.0122	0.0237	0.0299	0.0166	0.0421
	AD	1.0525	0.2522	0.0177	0.0170	0.0161	0.0769	0.0650	0.0461	0.0433	0.0118	0.0223	0.0303	0.0162	0.0410

VTMLE results

		ν_1	ν_2	A_{ii}			B_{ii}			$vech(\Omega)'$					
True		10	8	0.40	0.55	0.50	0.40	0.30	0.50	0.50	0.20	0.30	0.50	0.25	0.50
T=1000	Bias	0.0051	0.0457	-0.0011	-0.0019	-0.0014	-0.0092	-0.0082	-0.0077	-0.0065	0.0018	0.0028	-0.0038	0.0027	0.0017
	SD	0.4474	0.2767	0.0274	0.0300	0.0262	0.1247	0.0988	0.0764	0.0658	0.0237	0.0375	0.0447	0.0293	0.0647
	AD	0.4152	0.2921	0.0266	0.0285	0.0255	0.1143	0.1002	0.0930						
T=2000	Bias	-0.0016	0.0235	-0.0006	-0.0005	-0.0006	-0.0044	-0.0024	-0.0025	-0.0029	0.0004	0.0010	-0.0027	0.0001	-0.0002
	SD	0.2912	0.1963	0.0194	0.0212	0.0183	0.0833	0.0643	0.0486	0.0454	0.0162	0.0258	0.0303	0.0199	0.0431
	AD	0.2943	0.2080	0.0208	0.0193	0.0172	0.0788	0.0652	0.0593						
True		15	10	0.40	0.55	0.50	0.40	0.30	0.50	0.50	0.20	0.30	0.50	0.25	0.50
T=1000	Bias	0.0656	0.0407	-0.0012	0.0008	-0.0011	-0.0169	-0.0120	-0.0092	-0.0027	0.0031	0.0058	-0.0025	0.0042	0.0041
	SD	0.8861	0.3855	0.0262	0.0258	0.0244	0.1326	0.1037	0.0758	0.0693	0.0210	0.0377	0.0451	0.0269	0.0631
	AD	0.9330	0.4018	0.0275	0.0263	0.0237	0.1271	0.1082	0.0682						
T=2000	Bias	0.0358	0.0214	-0.0007	-0.0005	-0.0009	-0.0070	-0.0045	-0.0036	-0.0018	0.0009	0.0023	-0.0018	0.0014	0.0015
	SD	0.5965	0.2556	0.0182	0.0175	0.0171	0.0899	0.0698	0.0492	0.0472	0.0143	0.0258	0.0307	0.0183	0.0432
	AD	0.6192	0.2429	0.0194	0.0186	0.0166	0.0825	0.0765	0.0607						
True		20	10	0.40	0.55	0.50	0.40	0.30	0.50	0.50	0.20	0.30	0.50	0.25	0.50
T =1000	Bias	0.0882	0.0483	-0.0009	-0.0009	-0.0006	-0.0120	-0.0109	-0.0098	-0.0049	0.0024	0.0040	-0.0026	0.0037	0.0027
	SD	1.5377	0.3668	0.0277	0.0252	0.0236	0.1277	0.1002	0.0781	0.0670	0.0196	0.0365	0.0428	0.0256	0.0638
	AD	1.6297	0.3917	0.0269	0.0265	0.0248	0.1216	0.0993	0.0921						
T =2000	Bias	0.0690	0.0282	-0.0001	-0.0013	-0.0003	-0.0060	-0.0047	-0.0060	-0.0016	0.0015	0.0023	-0.0007	0.0019	0.0033
	SD	1.0280	0.2451	0.0168	0.0177	0.0165	0.0835	0.0686	0.0484	0.0457	0.0135	0.0242	0.0302	0.0177	0.0424
	AD	1.0733	0.2690	0.0188	0.0181	0.0177	0.0809	0.0711	0.0650						

Simulation 2

- To check the size of the proposed portmanteau test, $P = 1$, $Q = 1$, $K = 1$.
- $Y_t = \Sigma_t^{1/2} \Delta_t \Sigma_t^{1/2}$ with $\Sigma_t = \Omega + A_1 Y_{t-1} A_1' + B_1 \Sigma_{t-1} B_1'$,
-

$$A_1 = \begin{bmatrix} 0.42 & 0 & 0 \\ 0 & 0.45 & 0 \\ 0 & 0 & 0.44 \end{bmatrix}, B_1 = \begin{bmatrix} 0.72 & 0 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0 & 0.65 \end{bmatrix},$$

$$\Omega = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.2 & 0.5 & 0.25 \\ 0.3 & 0.25 & 0.5 \end{bmatrix}.$$

- 95%

Size

Table: Size of Test

	MLE					VT-MLE				
Lags	1	2	3	4	5	1	2	3	4	5
T=1000	0.068	0.064	0.079	0.071	0.075	0.071	0.069	0.076	0.070	0.072
T=2000	0.062	0.052	0.064	0.055	0.055	0.055	0.054	0.058	0.059	0.060

Simulation 2

- To check the power of the test, we specify a model with $P = 2$, $Q = 1$, $K = 1$.
- Specify $A_2 = \lambda \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ for different values of λ.
- But we estimate the simulated data with $P = 1$, $Q = 1$, $K = 1$.

Power

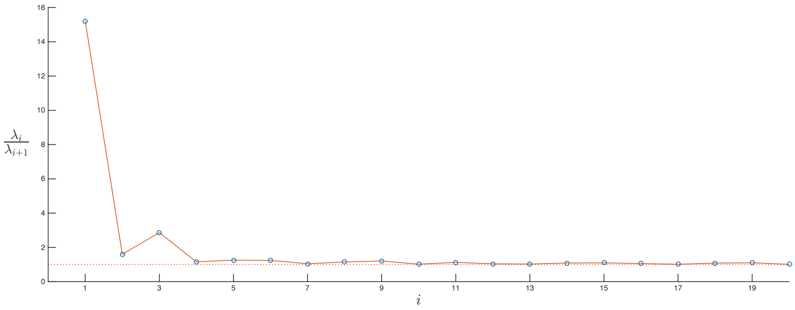
Table: The propotion of rejection for MLE ($\Pi(l)$) and VTMLE ($\Pi_v(l)$)

		$l=1$		$l=2$		$l=3$		$l=4$		$l=5$	
λ	T	$\Pi(l)$	$\Pi_v(l)$	$\Pi(l)$	$\Pi_v(l)$	$\Pi(l)$	$\Pi_v(l)$	$\Pi(l)$	$\Pi_v(l)$	$\Pi(l)$	$\Pi_v(l)$
0.1	1000	0.373	0.379	0.531	0.522	0.520	0.524	0.518	0.507	0.511	0.516
	2000	0.421	0.440	0.614	0.620	0.627	0.648	0.655	0.660	0.648	0.653
0.2	1000	0.420	0.433	0.520	0.551	0.518	0.507	0.524	0.520	0.515	0.520
	2000	0.501	0.503	0.626	0.642	0.611	0.629	0.630	0.624	0.627	0.632
0.3	1000	0.619	0.622	0.782	0.779	0.749	0.762	0.744	0.720	0.730	0.719
	2000	0.690	0.651	0.824	0.809	0.801	0.817	0.767	0.799	0.753	0.782
0.4	1000	0.768	0.749	0.801	0.792	0.799	0.810	0.797	0.803	0.811	0.799
	2000	0.844	0.838	0.881	0.895	0.901	0.897	0.884	0.895	0.897	0.903
0.5	1000	0.901	0.905	0.916	0.904	0.914	0.901	0.916	0.940	0.912	0.952
	2000	0.980	0.958	0.986	0.956	0.991	0.969	0.990	0.985	0.991	0.988

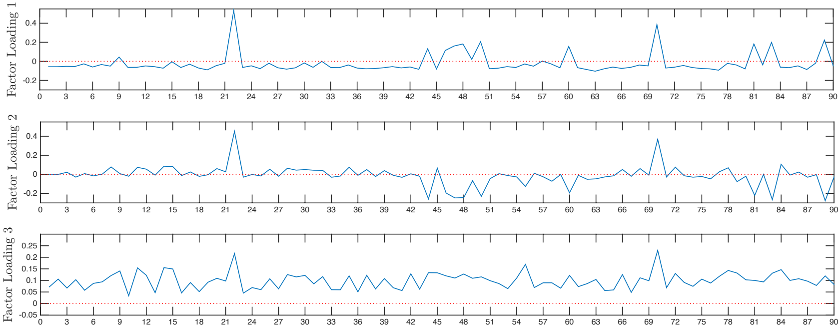
Application

- We collected intraday data for 90 stocks from S&P500, the period is from January 3rd 2005 to June 2nd 2008. In total 858 days, we use the first 758 days to fit the model and last 100 days as forecast period.
- We impose the factor model first, then consider model the lower ranked part by the CBF model, versus the CAW model.
- Based on eigen value ratio plot, we consider number of factors $r = 3$.

Application



Application



Application

Table: Top 10 magnitude of Factor loadings for each factor

Factor 1	Number	22	44	47	48	50	60	70	81	83	89
	Stock	LEN	MRO	NBL	NE	NFX	OXY	PHM	PXD	RDC	RRC
	Factor loading	0.5318	0.1317	0.1606	0.1822	0.2062	0.1566	0.3854	0.1825	0.1988	0.2228
Factor 2	Number	22	44	46	47	48	50	70	81	83	89
	Stock	LEN	MRO	MUR	NBL	NE	NFX	PHM	PXD	RDC	RRC
	Factor loading	0.4518	-0.2580	-0.1962	-0.2476	-0.2442	-0.2304	0.3684	-0.2215	-0.2655	-0.2767
Factor 3	Number	9	11	14	15	22	44	55	70	78	84
	Stock	JEC	JPM	KEY	KIM	LEN	MRO	NUE	PHM	PSA	RF
	Factor loading	0.1410	0.1544	0.1553	0.1496	0.2159	0.1336	0.1693	0.2302	0.1432	0.1468

Application

Table: Results of diagnostic test in Application 2 with $P = 1, Q = 3$

Lags	1	2	3	4	5
P Value	0.1529	0.2433	0.1215	0.1805	0.1045

Application

Table: Summary of predictive results in Application 2

		(P, Q)								
		$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$	$(3,1)$	$(3,2)$	$(3,3)$
Forecast error†	F-CBF model	77.868	77.314	76.846	77.924	77.313	76.998	77.982	77.312	76.974
	F-CAW model	78.244	77.871	77.328	78.429	77.870	77.554	78.465	77.836	77.419
DM test‡		0.0033	0.0144	0.0015	0.0089	0.0157	0.0114	0.0009	0.0126	0.1083

† One-step forecast error in Frobenius norm
‡ The p-value of the DM test

Comment

- We have developed under the matrix-F a framework of conditional volatility model for Realized Covariance Matrices, and explored the asymptotics and stationarity conditions. Several diagonostic checking methodsd have been studied in this work
- In application, the CBF model can achieve significant better forecast performance than the CAW model.
- Real data suggest that some long memory model or change point may occur in the matrix process, which are of interest in future research.

Reference

- At-Sahalia, Y., J. Fan, and D. Xiu (2010). High-frequency covariance estimates with noisy and asynchronous nancial data. Journal of the American Statistical Association 105 (492), 1504-1517.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. Econometrica 71 (2), 579-625.
- Baillie, R. T. and T. Bollerslev (1989). Common stochastic trends in a system of exchange rates. the Journal of Finance 44 (1), 167-181.
- Barndor-Nielsen, O. E. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64 (2), 253-280.
- Boussama, F., Fuchs, F., and Stelzer, R. (2011). Stationarity and geometric ergodicity of bekk multivariate garch models. Stochastic Processes and their Applications, 121(10):23312360.
- Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. Econometric theory, 11(01):122150.
- Li, W. and McLeod, A. (1981). Distribution of the residual autocorrelations in multivariate arma time series models. Journal of the Royal Statistical Society. Series B (Methodological), pages 231239.
- Ling, S. and Li, W.-K. (1997). Diagnostic checking of nonlinear multivariate time series with multivariate arch errors. Journal of Time Series Analysis, 18(5):447464.
- Pedersen, R. S. and Rahbek, A. (2014). Multivariate variance targeting in the bekkgarch model. The Econometrics Journal, 17(1):2455.

Reference

- Konno, Y. (1991). A note on estimating eigenvalues of scale matrix of the multivariate f-distribution. Annals of the Institute of Statistical Mathematics 43 (1), 157-165.
- Leung, P. L. and M. Lo (1996). An identity for the noncentral multivariate f distribution with application. Statistica Sinica, 419-431.
- Shen, K., Yao, J., and Li, W. K. (2015). Forecasting high-dimensional realized volatility matrices using a factor model. arXiv preprint arXiv:1504.03454.
- Tao, M., Wang, Y., Yao, Q., and Zou, J. (2011). Large volatility matrix inference via combining low-frequency and high-frequency approaches. Journal of the American Statistical Association, 106(495):10251040.
- Tao, M., Wang, Y., Zhou, H. H., et al. (2013). Optimal sparse volatility matrix estimation for high-dimensional itô processes with measurement errors. The Annals of Statistics, 41(4):18161864.
- Wang, Y., Zou, J., et al. (2010). Vast volatility matrix estimation for highfrequency financial data. The Annals of Statistics, 38(2):943978.
- Tse, Y. K. (2002). Residual-based diagnostics for conditional heteroscedasticity models. The Econometrics Journal 5 (2), 358-374.
- Yu, P. L., Li, W., and Ng, F. (2017). The generalized conditional autoregressive wishart model for multivariate realized volatility. Journal of Business & Economic Statistics, pages 115.
- Zhu, K. and W. K. Li (2015). A new pearson-type qmle for conditionally heteroscedastic models. Journal of Business & Economic Statistics 33 (4), 552-565.

Thank You