Statistical methods for high-dimensional volatility

Mark Podolskij (University of Luxembourg)

join work with Grégoire Szymanski

Plan

- Brief review of high frequency methods
- Literature review on high dimensional volatility
- Random matrix theory
- Estimation of the spectral distribution

High frequency setting

- In this talk we consider a d-dimensional diffusion model of the type

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s$$

where a is a d-dimensional **drift** process, σ is a $\mathbb{R}^{d\times d}$-valued **volatility** process and W is a d-dimensional Brownian motion.

- We observe the data

$$X_0,\ X_{1/n},\ X_{2/n}, \ldots, X_{(n-1)/n},\ X_1 \qquad \text{with} \quad n \to \infty.$$

For a fixed $d \in \mathbb{N}$ define the **realised variance** estimator via

$$\widehat{\Sigma}_n = \sum_{i=1}^{n} (X_{i/n} - X_{(i-1)/n})^{\otimes 2}$$

with $x^{\otimes 2} = xx^\star$.

Theorem (Barndorff-Nielsen, Graversen, Jacod, P. and Shephard (06))

Denote by

$$\Sigma := \int_0^1 \sigma_s^{\otimes 2} ds$$

the ***quadratic variation*** *of X. It holds that*

$$\Delta_n^{-1/2} \left(\widehat{\Sigma}_n - \Sigma \right) \xrightarrow{d} \mathcal{MN} \left(0, \int_0^1 A_s ds \right)$$

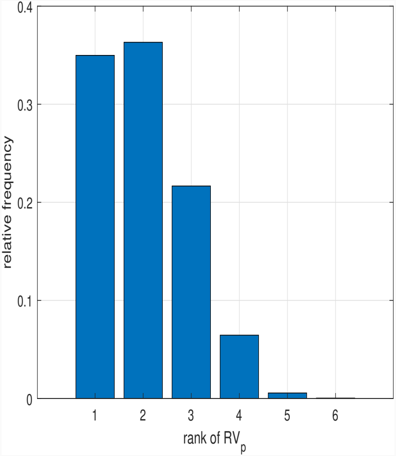
where $A_s^{jk,j'k'} = c_s^{jj'} c_s^{kk'} + c_s^{jk'} c_s^{kj'}$ and $c_s = \sigma_s^{\otimes 2}$.

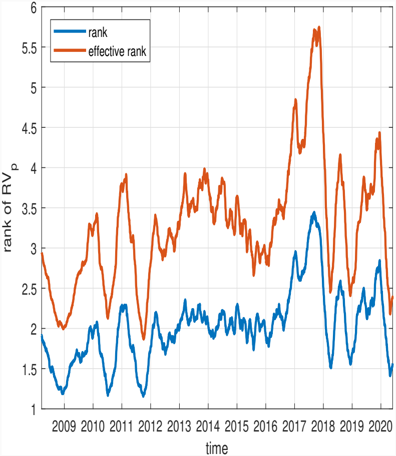
Central question of this talk

Can we infer the volatility in the high dimensional setting $d \sim n$?

- *Volatility estimation under sparsity constraints:* Wang and Zou (10), Fan and Kim (18), Koike (20), Christensen, Nielsen and P. (22).
- *Principal component analysis and factor modelling:* Aït-Sahalia and Xiu (17,20), Li, Linton and Zhang (25), Li, Ding and Zheng (21).
- *Estimation of spectral distribution:* Zheng and Li (11), Heinrich and P. (14).

Average ranks over 90-days window from $d = 30$ members of Dow Jones Industrial Average. We use 5-minutes data over the period January 2007-May 2020.





Question

Can we infer the volatility in the high dimensional setting without sparsity constraints?

Spectral distribution

- Let $\Sigma \in \mathbb{R}^{d \times d}$ be symmetric positive definite. In the following discussion we focus on the **spectral distribution** of Σ:

$$F_\Sigma(x) := \frac{1}{d} \sum_{j=1}^{d} 1_{\{\lambda_j(\Sigma) \leq x\}}, \qquad x \in \mathbb{R},$$

  where $\lambda_j(\Sigma)$'s denote the eigenvalues of Σ.

- Consider iid observations

$$Y_1, \ldots, Y_n \sim \mathcal{N}_d(0, \Sigma)$$

  Let Σ_n denote the empirical covariance matrix.

- How does F_{Σ_n} relate to F_Σ when $d \sim n$?

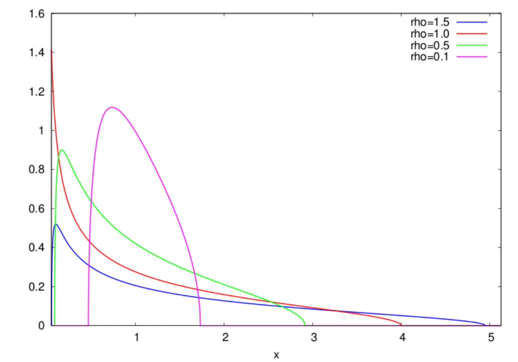
Marčenko-Pastur distribution

Theorem (Marčenko and Pastur (67))

Assume that $\Sigma = I_d$ *and* $d/n \to \rho \in (0,1)$. *Then* $F_{\widehat{\Sigma}_n} \to G$ *weakly* $\mathbb{P}$*-almost surely, where the measure* G *has a Lebesgue density* g*:*

$$g(x) = \frac{1}{2\pi} \frac{\sqrt{(\rho_+ - x)(x - \rho_-)}}{\rho x} 1_{[\rho_-, \rho_+]}(x)$$

where $\rho_+ = (1 + \sqrt{\rho})^2$ *and* $\rho_- = (1 - \sqrt{\rho})^2$.



Marčenko-Pastur distribution: General version

Theorem

Assume that $d/n \to \rho \in (0,\infty)$ and $F_\Sigma \to F$ weakly $\mathbb{P}$-almost surely as $d \to \infty$. Introduce the ***Stieltjes transform*** *of a distribution μ:*

$$m_\mu(z) := \int_{\mathbb{R}} \frac{1}{z-x}\mu(dx) \qquad z \in \mathbb{C}_+$$

Define the function

$$v_{\Sigma_n}(z) := -z^{-1}\left(1 - \frac{d}{n}\right) + \frac{d}{n} m_{F_{\Sigma_n}}(z)$$

Then it holds

$$v_{\Sigma_n}(z) \overset{a.s.}{\to} v(z)$$

and v is related to F via the ***Marčenko-Pastur equation****:*

$$-\frac{1}{v(z)} = z - \rho \int_0^\infty \frac{x}{1 + xv(z)} F(dx) \qquad \forall z \in \mathbb{C}_+$$

Statistical application

- El Karoui (08) has proposed to use this probabilistic result to recover the theoretical distribution F. First, he assumed that

$$F_\Sigma = F$$

- In the second step the function $v_{\Sigma_n}(z)$ is used as an approximation of $v(z)$.

- In the last step a numerical procedure is proposed to recover the spectral distribution function F from the Marčenko-Pastur equation

$$-\frac{1}{v(z)} = z - \rho \int_0^\infty \frac{x}{1 + xv(z)} F(dx)$$

when the function v is known.

Question

Does this procedure work for **time-varying** volatility process?

Univariate time variation

- Zheng and Li (11) consider the semimartingale model

$$X_t = X_0 + \int_0^t f_s \sigma dW_s$$

where $\sigma \in \mathbb{R}^{d\times d}$ is a constant matrix and $f : [0,1] \to \mathbb{R}$.

- In this setting the spectral distribution $F_{\widehat{\Sigma}_n}$ does not converge in general!

- However, the statistic

$$\widetilde{\Sigma}_n := \frac{\operatorname{tr}(\widehat{\Sigma}_n)}{n} \sum_{i=1}^{n} \frac{(X_{i/n} - X_{(i-1)/n})^{\otimes 2}}{\|X_{i/n} - X_{(i-1)/n}\|^2}$$

does satisfy the Marčenko-Pastur theorem.

Time varying volatility

- Heinrich and P. (14) consider time varying volatility model

$$X_t = X_0 + \int_0^t \sigma_s dW_s$$

with

$$\sigma_s := A\ 1_{[0,1/2]}(s) + B\ 1_{(1/2,1]}(s).$$

where $\sigma \in \mathbb{R}^{d\times d}$ is a constant matrix and $f : [0,1] \to \mathbb{R}$.

- They proved that $F_{\widehat{\Sigma}_n}$ converges via **method of moments**. However, very strong conditions on the matrices A and B are required ($\to$ *free probability*).

- In particular, these results can't be used to infer the spectral distribution F_Σ where $\Sigma = (A^{\otimes 2} + B^{\otimes 2})/2$.

Subsampling approach

- To overcome this problem we apply a simple **subsampling method**. Define random variables

$$Y_i := \left(X_{i/n} - X_{(i-1)/n} + X_{(i+\lfloor n/2 \rfloor)/n} - X_{(i+\lfloor n/2 \rfloor - 1)/n}\right)$$

for $i = 1, \ldots, \lfloor n/2 \rfloor$.

- When the matrices A and B are deterministic:

$$(Y_i) \text{ are iid} \sim \mathcal{N}_d\,(0, (A+B)/n)$$

- Consequently, the matrix

$$\Sigma_n := \sum_{i=1}^{\lfloor n/2 \rfloor} Y_i^{\otimes 2}$$

can be used to infer F_Σ as proposed in El Karoui (08).

Piecewise constant volatility model

- We consider a time varying volatility model of the form

$$X_t = X_0 + \int_0^t \sigma_s dW_s$$

with

$$\sigma_s := \sum_{j=1}^{m} \sigma_j 1_{[t_j, t_{j+1})}(s), \qquad t_1 := 0, \ t_{m+1} := 1$$

where $\sigma_j \in \mathbb{R}^{d \times d}$ are independent of the Brownian motion W.

- We define $k_j^\star$ such that $t_j \in [k_j^\star/n, (k_j^\star + 1)/n)$.

- Our first goal is to perform a multiple change point detection to infer $k_1^\star, \dots, k_m^\star$.

Wild binary segmentation method

- We employ the **wild binary segmentation** method proposed by Fryzlewicz (14). We introduce the CUSUM statistic

$$U^{(a,b)}(s,t,k) = \sqrt{\frac{t-k}{(t-s)(k-s)}} \sum_{i=s+1}^{k} z_i^a z_i^b - \sqrt{\frac{k-s}{(t-s)(t-k)}} \sum_{i=k+1}^{t} z_i^a z_i^b,$$

  where $z_i := \sqrt{n}(X_{i/n} - X_{(i-1)/n})$.

- We introduce the maximizer

$$\widehat{k}^{(a,b)}(s,t) := \mathrm{argmax}_k |U^{(a,b)}(s,t,k)|$$

  and define the set

$$\mathcal{D}_\tau(s,t) = \left\{(a,b) : \max_{s<k<t} |U^{(a,b)}(s,t,k)| > \tau \log(n),\ 1 \le a \le b \le d\right\}.$$

Algorithm

For $\epsilon = \epsilon_n \sim n^{1/2}$ run initiated at $s = 0$ and $t = n$:

Algorithm 1 Binary Segmentation$(s,\ t)$

if $\mathcal{D}_\tau(s,t)$ is empty **then**

  return $\emptyset$

else

  Choose $(a,b) \in \mathcal{D}_\tau(s,t)$

  return

$$\left\{\widehat{k}^{(a,b)}(s,t)\right\} \cup \text{Binary Segmentation}\left(s, \widehat{k}^{(a,b)}(s,t) - \lfloor \epsilon \rfloor\right) \cup \text{Binary Segmentation}\left(\widehat{k}^{(a,b)}(s,t) + \lfloor \epsilon \rfloor, t\right)$$

end if

The algorithm produces quantities $\left(\widehat{k}_j\right)_{j=1,\ldots,\widehat{m}}$.

Concentration bound

Theorem (P. and Szymanski (25))

Let $\epsilon_n \sim n^{1/2}$. Then it holds

$$\mathbb{P}\left(\widehat{m} = m, |\widehat{k}_1 - k_1^\star| \leq \epsilon_n, \ldots, |\widehat{k}_m - k_m^\star| \leq \epsilon_n\right) \to 1.$$

Putting things together

- *Multiple change point detection:* Use the wild binary segmentation method to infer m, $k_1^\star, \ldots, k_m^\star$.
- *Subsamlping:* Construct weighted subsampling increments $Y_1, \ldots, Y_{n_\star}$ and define

$$\Sigma_n := \sum_{i=1}^{n_\star} Y_i^{\otimes 2}$$

- *Inversion:* Use the Marčenko-Pastur law to infer F_Σ from F_{Σ_n}.

Thank you very much for your attention!