SUPPLEMENT TO "ON THE ESTIMATION OF INTEGRATED COVARIANCE MATRICES OF HIGH DIMENSIONAL DIFFUSION PROCESSES"

By Xinghua Zheng * and Yingying Li *

Hong Kong University of Science and Technology

The proposition/lemma/equation numbers below refer to the main article Zheng and Li (2010).

Proof of Proposition 4. By (2.5) and (2.6), for any $t \in (0,1]$,

$$\frac{\int_0^t \sigma_s^{(j)}\sigma_s^{(k)} d\langle W^{(j)}, W^{(k)}\rangle_s}{\sqrt{\int_0^t (\sigma_s^{(j)})^2\,ds \cdot \int_0^t (\sigma_s^{(k)})^2\,ds}} = r^{(jk)} \frac{\int_0^t \sigma_s^{(j)}\sigma_s^{(k)}\,ds}{\sqrt{\int_0^t (\sigma_s^{(j)})^2\,ds \cdot \int_0^t (\sigma_s^{(k)})^2\,ds}} = \rho^{(jk)},$$

i.e., for any $t \in (0,1]$,

$$\frac{\rho^{(jk)}}{r^{(jk)}} = \frac{\int_0^t \sigma_s^{(j)}\sigma_s^{(k)}\,ds/t}{\sqrt{\int_0^t (\sigma_s^{(j)})^2\,ds/t \cdot \int_0^t (\sigma_s^{(k)})^2\,ds/t}}.$$

Letting $t \downarrow 0$, using the l'Hospital's rule and noting that $\sigma_t^{(j)}$ are càdlàg, we see that

$$\frac{\rho^{(jk)}}{r^{(jk)}} = \frac{\sigma_0^{(j)}\sigma_0^{(k)}}{\sqrt{(\sigma_0^{(j)})^2 \cdot (\sigma_0^{(k)})^2}} = 1 \quad \text{or} \quad -1.$$

Therefore for all $t \in (0,1]$,

$$\left|\int_0^t \sigma_s^{(j)}\sigma_s^{(k)}\,ds\right| = \sqrt{\int_0^t (\sigma_s^{(j)})^2\,ds \cdot \int_0^t (\sigma_s^{(k)})^2\,ds}.$$

By Cauchy-Schwartz inequality, this holds only if $\sigma_s^{(j)}$ and $\sigma_s^{(k)}$ are proportional to each other. Therefore, almost surely, there exists a scalar process $\gamma_t \in D([0,1);\mathbb{R})$ and a p-dimensional vector $(\sigma^{(1)},\ldots,\sigma^{(p)})^T$ such that

$$(\sigma_t^{(1)},\ldots,\sigma_t^{(p)})^T = \gamma_t \cdot (\sigma^{(1)},\ldots,\sigma^{(p)})^T.$$

*Research partially supported by DAG (HKUST) and GRF 606811 of the HKSAR.

Now we show that $\mathbf{X}_t := (X_t^{(j)})$ belongs to Class $\mathcal{C}$. In fact one can always find a p-dimensional standard Brownian motion $\widetilde{\mathbf{W}}_t$ such that

$$\mathbf{W}_t = R^{1/2}\widetilde{\mathbf{W}}_t,$$

where $\mathbf{W}_t = (W_t^{(1)}, \ldots, W_t^{(p)})^T$, R is the (constant) correlation matrix of $\mathbf{W}_t$, and $R^{1/2}$ stands for its square root matrix. Hence, writing $\mu_t = (\mu_t^{(1)}, \ldots, \mu_t^{(p)})^T$, we have

$$d\mathbf{X}_t = \mu_t\, dt + \mathrm{diag}(\sigma_t^{(1)}, \ldots, \sigma_t^{(p)})\, d\mathbf{W}_t = \mu_t\, dt + \gamma_t \cdot \mathrm{diag}(\sigma^{(1)}, \ldots, \sigma^{(p)}) R^{1/2}\, d\widetilde{\mathbf{W}}_t.$$

□

Next, we give a detailed explanation of the second statement in Remark 3, namely, if $w_s \not\equiv \int_0^1 w_t\, dt$ on $[0,1)$, then except in the trivial case when H is a delta measure at 0, the LSD $F^w \neq F$, where F is the LSD in Proposition 1 determined by $H(\cdot/\int_0^1 w_t\, dt)$. The reason is as follows: firstly, for any $w = (w_s)$, the family of distributions $\{F^{cw} : c > 0\}$ is a scale family, i.e., for any $x \geq 0$, $F^{cw}(x) = F^w(x/c)$. Hence it suffices to show that for any w such that $0 < w_s \not\equiv 1$ on $[0,1)$ and $\int_0^1 w_s\, ds = 1$, $F^w \neq F$. In fact, $F^w = F$ if and only if $m_{F^w}(z) = m_F(z)$ for all $z \in \mathbb{C}_+$, and by the uniqueness of the solution to the equation (1.3), this can be true only if $M(z) = -1/z \cdot (1 - y(1 + zm(z)))$. Suppose this is true, then by (3.28), $(1 + zm(z))/\widetilde{m}(z) = 1 - y(1 + zm(z))$, or, $1/(1 + y\widetilde{m}(z)) = (1 + zm(z))/\widetilde{m}(z) = 1 - y(1 + zm(z))$. By (3.23), the last equation implies that

$$\int_0^1 \frac{1}{1 + y\widetilde{m}(z) w_s}\, ds = \frac{1}{1 + y\widetilde{m}(z)}.$$

Since both sides are analytical in $\widetilde{m}(z)$, we get

$$\int_0^1 \frac{1}{1 + u w_s}\, ds = \frac{1}{1 + u}, \quad \text{for all } u \in \mathbb{C} \text{ such that } \mathrm{Re}(u) > 0.$$

In particular, focusing on $u > 0$, multiplying both sides by u, letting $u \to \infty$ and applying the monotone convergence theorem to the left hand side yield

$$\int_0^1 \frac{1}{w_s}\, ds = 1. \tag{5.1}$$

However, by the Cauchy-Schwarz inequality, since $\int_0^1 w_s\, ds = 1$, (5.1) holds only if $w_s \equiv 1$ on $[0,1)$. □

Now we give the proofs of the various lemmas in Section 3.1

PROOF OF LEMMA 1. For notational ease we shall omit the superscript p, and write $\mathbf{v}_\ell = \mathbf{v}_\ell^{(p)}$, $\mathbf{w}_\ell = \mathbf{w}_\ell^{(p)}$, and $v_\ell^{(p,j)} = v_\ell^{(j)}$ etc. For each p, set $A = A_p = (\mathbf{v}_1 + \mathbf{w}_1, \ldots, \mathbf{v}_n + \mathbf{w}_n)$, and $B = B_p = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$. Then $\widetilde{S}_n = AA^T$, and $S_n = BB^T$. By Lemma 2.7 of Bai (1999),

$$(L(F^{\widetilde{S}_n}, F^{S_n}))^4 \le \frac{2}{p^2} \operatorname{tr}((\mathbf{v}_1, \ldots, \mathbf{v}_n)(\mathbf{v}_1, \ldots, \mathbf{v}_n)^T) \cdot \operatorname{tr}(\widetilde{S}_n + S_n).$$

By condition (ii), $\operatorname{tr}((\mathbf{v}_1, \ldots, \mathbf{v}_n)(\mathbf{v}_1, \ldots, \mathbf{v}_n)^T) \le np \cdot \varepsilon_p^2$. Moreover,

$$\begin{aligned}
\operatorname{tr}(\widetilde{S}_n) &= \sum_{\ell=1}^n \sum_{j=1}^p (v_\ell^{(j)} + w_\ell^{(j)})^2 \\
&\le 2\sum_{\ell=1}^n \sum_{j=1}^p (v_\ell^{(j)})^2 + 2\sum_{\ell=1}^n \sum_{j=1}^p (w_\ell^{(j)})^2 \\
&\le 2np \cdot \varepsilon_p^2 + 2\operatorname{tr}(S_n).
\end{aligned}$$

The conclusion follows since $\varepsilon_p = o(1/\sqrt{p})$ and hence $o(1/\sqrt{n})$. □

REMARK 4. *When applying this lemma to conclude that in Proposition 5 the drift terms are negligible, the $\mathbf{v}_\ell$'s are $\int_{\tau_{n,\ell-1}}^{\tau_{n,\ell}} \mu_t \, dt$, whose entries are $O(1/n) = o(1/\sqrt{p})$. Similar remark applies to Proposition 8.*

PROOF OF LEMMA 3. Let $A = UDU^*$ be the eigen-decomposition of A, where $D = \operatorname{diag}(d_j)$ with $d_j \ge 0$. Then $(wA + I)^{-1} = U(wD + I)^{-1}U^*$, and

$$||(wA + I)^{-1}|| = \frac{1}{\min_j |wd_j + 1|} \le 1.$$

□

Lemma 5 can be proved similarly.

PROOF OF LEMMA 4. The proof relies on the following three facts: for any $p \times p$ matrices C and D,
(1) $|\operatorname{tr}(CD)| \le \sqrt{\operatorname{tr}(CC^*) \cdot \operatorname{tr}(DD^*)} \le p \cdot ||C|| \cdot ||D||$;
(2) if C and D are invertible, then $C^{-1} - D^{-1} = C^{-1}(D - C)D^{-1}$;
(3) $||CD|| \le ||C|| \cdot ||D||$.
Therefore

$$\begin{aligned}
&|\operatorname{tr}\left(B\left((w_1A + I)^{-1} - (w_2A + I)^{-1}\right)\right)| \\
\le& p \cdot ||B|| \cdot ||(w_1A + I)^{-1}|| \cdot ||(w_1 - w_2)A|| \cdot ||(w_2A + I)^{-1}||.
\end{aligned}$$

By Lemma 3 we see that (i) holds.

As to (ii),

$$\begin{aligned}
&|\mathbf{q}^*B(w_1A+I)^{-1}\mathbf{q}-\mathbf{q}^*B(w_2A+I)^{-1}\mathbf{q}|\\
=&|\mathbf{q}^*B(w_1A+I)^{-1}((w_1-w_2)A)(w_2A+I)^{-1}\mathbf{q}|\\
\leq&||B||\cdot||(w_1A+I)^{-1}((w_1-w_2)A)(w_2A+I)^{-1}||\cdot|\mathbf{q}|^2\\
\leq&||B||\cdot|w_1-w_2|\cdot||A||\cdot|\mathbf{q}|^2,
\end{aligned}$$

where in the last inequality we used Lemma 3 again. □

Proof of Lemma 6. Since $-1/z = i/v$, it suffices to show that $1 + \tau\mathbf{q}^*(A-zI)^{-1}\mathbf{q} \in Q_1$. In fact, let $A = UDU^*$ be the eigen-decomposition of A with $D = \operatorname{diag}(d_j)$ and $d_j \geq 0$. Then $\mathbf{q}^*(A-zI)^{-1}\mathbf{q} = \mathbf{q}^*U\operatorname{diag}(1/(d_j - iv))U^*\mathbf{q}$. Let $\mathbf{q}^*U = (w_1,\ldots,w_p)$. Then

$$\operatorname{Re}(\mathbf{q}^*(A-zI)^{-1}\mathbf{q}) = \sum_j |w_j|^2 d_j/(d_j^2+v^2) \geq 0, \tag{5.2}$$

and

$$\operatorname{Im}(\mathbf{q}^*(A-zI)^{-1}\mathbf{q}) = \sum_j |w_j|^2 v/(d_j^2+v^2) \geq 0.$$

□

Proof of Lemma 7. This is by direct calculation. Using the fact that for any two invertible matrices C and D, $C^{-1}+D^{-1} = C^{-1}(C+D)D^{-1}$, and plugging in $z = iv$ we get

$$\begin{aligned}
\operatorname{Re}(\operatorname{tr}(A(B-zI)^{-1}A^*)) &= \frac{\operatorname{tr}(A(B-ivI)^{-1}A^* + A(B+ivI)^{-1}A^*)}{2}\\
&= \frac{\operatorname{tr}(A(B-ivI)^{-1}(2B)(B+ivI)^{-1}A^*)}{2}.
\end{aligned}$$

The last term is nonnegative because the matrix inside the trace function is Hermitian nonnegative definite. Similarly,

$$\begin{aligned}
\operatorname{Im}(\operatorname{tr}(A(B-zI)^{-1}A^*)) &= \frac{\operatorname{tr}(A(B-ivI)^{-1}A^* - A(B+ivI)^{-1}A^*)}{2i}\\
&= \frac{\operatorname{tr}(A(B-ivI)^{-1}(2ivI)(B+ivI)^{-1}A^*)}{2i} \geq 0.
\end{aligned}$$

□

PROOF OF LEMMA 8. Suppose otherwise that there exist $z_1 = a + ib \neq z_2 = c + id$ with $a, b, c, d \geq 0$ such that

$$\int_0^1 \frac{1}{1+z_1 w_s}\,ds - \int_0^1 \frac{1}{1+z_2 w_s}\,ds = (z_2 - z_1)\int_0^1 \frac{w_s}{(1+z_1 w_s)(1+z_2 w_s)}\,ds = 0,$$

i.e.,

$$\int_0^1 \frac{w_s(1+\overline{z_1}w_s)(1+\overline{z_2}w_s)}{|1+z_1 w_s|^2 |1+z_2 w_s|^2}\,ds = 0. \tag{5.3}$$

Let us first study the imaginary part. We have that

$$\mathrm{Im}((1+\overline{z_1}w_s)(1+\overline{z_2}w_s)) = -(b+d)w_s - (ad+bc)w_s^2 \leq 0,$$

hence in order for (5.3) to hold, it has be that $b = d = 0$. It follows that

$$\mathrm{Re}((1+\overline{z_1}w_s)(1+\overline{z_2}w_s)) = 1 + (a+c)w_s + ac w_s^2 \geq 1,$$

contradicting (5.3). □

Finally we prove Proposition 7. The proof uses the following Lemma which is of some independent interest.

LEMMA 10. *Suppose that X_t is a one dimensional diffusion process satisfying*

$$dX_t = \mu_t\,dt + \sigma_t\,dW_t, \quad t \in [0,1],$$

and there exist constants $C_\mu, c_\sigma, C_\sigma$ such that

$$|\mu_t| \leq C_\mu, \quad \text{and} \quad 0 < c_\sigma \leq |\sigma_t| \leq C_\sigma < \infty, \quad \text{for all } t \in [0,1] \quad \text{almost surely.}$$

Suppose further that the observation times (not necessarily independent of X_t!) $0 = \tau_0 < \tau_1 < \ldots < \tau_n = 1$ satisfy

$$\max_{1 \leq i \leq n} n|\tau_i - \tau_{i-1}| \leq C_\Delta < \infty.$$

Then the realized volatility $[X,X]_1 := \sum_{i=1}^n |X_{\tau_i} - X_{\tau_{i-1}}|^2$ satisfies

$$P\left(\sqrt{n}\left|[X,X]_1 - \int_0^1 \sigma_t^2\,dt\right| \geq x\right) \leq \sqrt{2}\exp\left(3C_\mu^2/(2c_\sigma^2)\right)\cdot\exp\left(-x^2/(16C_\sigma^4 C_\Delta^2)\right)$$

for all $0 \leq x \leq C_\sigma^2 C_\Delta \sqrt{n}$.

PROOF. The proof makes use of Lemma 3 in Fan, Li and Yu (2010) and the Girsanov Theorem. Denote $A(x) = A_n(x) = \{\sqrt{n}|[X,X]_1 - \int_0^1 \sigma_t^2 \, dt| \ \geq \ x\}$. Define

$$Z_t = \exp\left(-\int_0^t \frac{\mu_s}{\sigma_s} \, dW_s - \frac{1}{2}\int_0^t \frac{\mu_s^2}{\sigma_s^2} \, ds\right), \quad t \in [0,1],$$

which is a martingale (under P), and define a measure $\widetilde{P}$ via Radon - Nikodym derivative $d\widetilde{P}/dP|\mathcal{F}_t = Z_t$. Then under measure $\widetilde{P}$, X_t satisfies

$$dX_t = \sigma_t \, d\widetilde{W}_t,$$

where $\widetilde{W}_t$ is a standard Brownian-motion under $\widetilde{P}$. Hence by Lemma 3 in Fan, Li and Yu (2010),

$$\widetilde{P}\left(A(x)\right) \leq 2\exp\left(-x^2/(8C_\sigma^4 C_\Delta^2)\right), \quad \text{for all } 0 \leq x \leq C_\sigma^2 C_\Delta \sqrt{n}.$$

Now by the Cauchy-Schwartz inequality,

$$P(A(x)) = \widetilde{E}\left(\mathbf{1}_{A(x)} \cdot 1/Z_1\right) \leq \sqrt{\widetilde{P}(A(x)) \cdot \widetilde{E}\left(1/Z_1^2\right)},$$

where $\widetilde{E}$ stands for the expectation taken under measure $\widetilde{P}$. However, by a similar argument as in the proof of Lemma 3 of Fan, Li and Yu (2010) (namely, firstly using a time-change for martingales and then applying the optional sampling theorem) one gets that

$$\begin{aligned}\widetilde{E}\left(1/Z_1^2\right) &= \widetilde{E}\exp\left(2\int_0^1 \frac{\mu_s}{\sigma_s} \, d\widetilde{W}_s + \int_0^1 \frac{\mu_s^2}{\sigma_s^2} \, ds\right) \\ &\leq \widetilde{E}\exp\left(2\widetilde{B}_{C_\mu^2/c_\sigma^2} + C_\mu^2/c_\sigma^2\right) = \exp(3C_\mu^2/c_\sigma^2),\end{aligned}$$

where $\widetilde{B}$ is a Brownian-motion under $\widetilde{P}$ resulted from the time change of the martingale $\int_0^t \mu_s/\sigma_s \, d\widetilde{W}_s$. The conclusion follows. □

PROOF OF PROPOSITION 7. It suffices to show that $(\mathrm{tr}(\Sigma_p^{RCV}) - \mathrm{tr}(\Sigma_p))/p \to 0$ almost surely. By Lemma 10, there exist constants $c, C_1, C_2 > 0$ such that for all $1 \leq j \leq p$ and $0 \leq x \leq c\sqrt{n}$,

$$P\left(\sqrt{n}|\Sigma_{p;jj}^{RCV} - \Sigma_{p;jj}| > x\right) \leq C_1 \exp\{-C_2 x^2\}.$$

Hence for any $0 < \varepsilon < c$ and for all p,

$$\begin{aligned}P\left(\left|\frac{\mathrm{tr}(\Sigma_p^{RCV}) - \mathrm{tr}(\Sigma_p)}{p}\right| \geq \varepsilon\right) \leq &\sum_{j=1}^p P\left(|\Sigma_{p,jj}^{RCV} - \Sigma_{p;jj}| \geq \varepsilon\right) \\ \leq &p \cdot C_1 \exp(-C_2\varepsilon^2 \cdot n).\end{aligned}$$

The almost sure convergence then follows from the Borel-Cantelli Lemma. □

REFERENCES

BAI, Z. D. (1999). Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica* **9** 611–677. With comments by G. J. Rodgers and Jack W. Silverstein; and a rejoinder by the author. MR1711663 (2000e:60044)

FAN, J., LI, Y. and YU, K. (2010). Vast Volatility Matrix Estimation using High Frequency Data for Portfolio Selection. *Manuscript*.

ZHENG, X. and LI, Y. (2010). "On the Estimation of Integrated Covariance Matrices of High Dimensional Diffusion Processes".

DEPARTMENT OF INFORMATION SYSTEMS,
BUSINESS STATISTICS AND OPERATIONS MANAGEMENT
HONG KONG UNIVERSITY OF SCIENCE
AND TECHNOLOGY
CLEAR WATER BAY, KOWLOON, HONG KONG.
E-MAIL: xhzheng@ust.hk, yyli@ust.hk