

Online Supplementary Material to
“Exploratory Analysis and Modeling of Stock Returns”

Kimihiro Noguchi Alexander Aue Prabir Burman*

Abstract

In this document, a detailed proof of Theorem 5.1 in Noguchi, Aue and Burman (2014) is given. It concerns the rate of convergence of the spline-based estimators as developed in Section 5 of that paper.

1 Proof of Theorem 5.1

The proof of Theorem 5.1 requires the use of the following auxiliary results. The first lemma can be established by repeatedly applying integration by parts.

Lemma 1. *Let r be an integrable function on $[0, 1]^2$ and let*

$$R(x_1, x_2) = \int_0^{x_1} \int_0^{x_2} r(u_1, u_2) du_1 du_2.$$

Let h_1 and h_2 be functions on $[0, 1]$ whose derivatives \dot{h}_1 and \dot{h}_2 are integrable. If $R(1, 1) = 0$, it follows that

$$\int_0^1 \int_0^1 h_1(u_1) h_2(u_2) r(u_1, u_2) du_1 du_2$$

*K. Noguchi (E-mail: Kimihiro.Noguchi@wwu.edu) is Assistant Professor, Department of Mathematics, Western Washington University, Bellingham, WA 98225. Alexander Aue (E-mail: aaue@ucdavis.edu) is Associate Professor and Prabir Burman (E-mail: pburman@ucdavis.edu) is Professor, Department of Statistics, University of California, One Shields Avenue, Davis, CA 95616. This research was partially supported by NSF grants DMS 09-05400, DMS 09-07622, DMS 12-09226, DMS 13-05858, and DMS 14-07530.

$$\begin{aligned}
&= -h_1(1) \int_0^1 \dot{h}_2(u_2) R(1, u_2) du_2 - h_2(1) \int_0^1 \dot{h}_1(u_1) R(u_1, 1) du_1 \\
&\quad + \int_0^1 \int_0^1 \dot{h}_1(u_1) \dot{h}_2(u_2) R(u_1, u_2) du_1 du_2.
\end{aligned}$$

Lemma 2. *If Assumptions 5.1 and 5.2 are satisfied, then with probability approaching one all the eigenvalues of kD , where $D = T^{-1} \sum_{t=1}^T B_t B_t'$, are between two positive constants.*

Proof. The proof is given in two parts. In step (a), it will be shown that the matrix $k\bar{D}$, with $\bar{D} = E[D]$, has all eigenvalues between two positive constants. This requires the use of part (a) of Assumption 5.2. In step (b), it will then be verified that $k\|D - \bar{D}\| \rightarrow 0$ in probability. Under the α -mixing condition of part (b) in Assumption 5.2, this requires that $k^3/T \rightarrow 0$ as assumed in Theorem 5.1. The relationship between k and T can be improved further: Under the simplified ρ -mixing condition

$$\sup_t \sup_{h_1, h_2} |\text{Corr}(h_1(X_t), h_2(X_{t+h}))| \leq \rho(h),$$

only $k^2/T \rightarrow 0$ and $\sum_{h=0}^{\infty} \rho(h) < \infty$ are needed. Under the simplified Ψ -mixing condition

$$\sup_t \sup_{x_1, x_2} |f_{t,t+h}(x_1, x_2) - f_t(x_1)f_{t+h}(x_2)| \leq \Psi(h),$$

a further simplification to $k/T \rightarrow 0$ and $\sum_{h=0}^{\infty} \Psi(h) < \infty$ is possible. A good summary of properties of various strong mixing conditions may be found in Bradley (2005).

(a) Note that for any vector $\beta = (\beta'_0, \beta'_1)'$,

$$\beta' \bar{D} \beta = \frac{1}{T} \sum_{t=1}^T \int [\beta'_0 B_0(t/T) + \beta'_1 \Phi_1(x)]^2 f_t(x) dx.$$

Part (a) of Assumption 5.2 implies that

$$\begin{aligned}
\beta' \bar{D} \beta &\geq \frac{\underline{f}}{T} \sum_{t=1}^T \int [\beta'_0 B_0(t/T) + \beta'_1 \Phi_1(x)]^2 dx, \\
\beta' \bar{D} \beta &\leq \frac{\bar{f}}{T} \sum_{t=1}^T \int [\beta'_0 B_0(t/T) + \beta'_1 \Phi_1(x)]^2 dx.
\end{aligned}$$

Since $\int \beta' \Phi(x) dx = 0$, it follows that

$$\frac{1}{T} \sum_{t=1}^T \int [\beta'_0 B_0(t/T) + \beta'_1 \Phi_1(x)]^2 dx$$

$$\begin{aligned}
&= \frac{1}{T} \sum_{t=1}^T [\beta_0' B_0(t/T)]^2 + \frac{1}{T} \sum_{t=1}^T \int [\beta_1' \Phi_1(x)]^2 dx \\
&= \beta_0' \left[\frac{1}{T} \sum_{t=1}^T B_0(t/T) B_0(t/T)' \right] \beta_0 + \beta_1' \Psi_1 \left(\int B_1(x) B_1(x)' dx \right) \Psi_1' \beta_1.
\end{aligned}$$

Following the arguments of Burman (1991), all eigenvalues of $kT^{-1} \sum_t B_0(t/T) B_0(t/T)'$ and $\int B_1 B_1'$ are bounded between two positive constants. Since $\Psi_1 \Psi_1' = I$, it follows that all eigenvalues of

$$\frac{k}{T} \sum_{t=1}^T \int [\beta_0' B_0(t/T) + \beta_1' \Phi_1(x)]^2 dx$$

and hence those of $k\bar{D}$ are bounded between two positive constants as well.

(b) Note that it is enough to show that $k\|V_{01}\| \rightarrow 0$ in probability and $k\|V_{11}\| \rightarrow 0$ in probability, where

$$\begin{aligned}
V_{01} &= \frac{1}{T} \sum_{t=1}^T B_0(t/T) \{B_1(X_t) - E[B_1(X_t)]\}', \\
V_{11} &= \frac{1}{T} \sum_{t=1}^T \{B_1(X_t) B_1(X_t)' - E[B_1(X_t) B_1(X_t)']\}.
\end{aligned}$$

The result is proved if it can be shown that $k^2 E[\text{tr}(V_{01}' V_{01})] \rightarrow 0$ and $k^2 E[\text{tr}(V_{11}' V_{11})] \rightarrow 0$.

0. Observe that

$$\begin{aligned}
k^2 E[\text{tr}(V_{01}' V_{01})] &= \left(\frac{k}{T}\right)^2 \sum_{i,j} \sum_{s,t=1}^T B_{0i}(s/T) B_{0i}(t/T) \text{Cov}(B_{1j}(X_s), B_{1j}(X_t)) \\
&= \left(\frac{k}{T}\right)^2 \sum_{i,j} \sum_{s,t \in I_i} B_{0i}(s/T) B_{0i}(t/T) \text{Cov}(B_{1j}(X_s), B_{1j}(X_t)),
\end{aligned}$$

where $I_i \subset \{1, \dots, T\}$ and $B_{0i}(t/T) \neq 0$ for $t \in I_i$. Then, $T_i = \text{card}(I_i) \sim T/k$.

Furthermore,

$$\text{Cov}(B_{1j}(X_s), B_{1j}(X_t)) = \int_0^1 B_{1j}(x_1) B_{1j}(x_2) [f_{s,t}(x_1, x_2) - f_s(x_1) f_t(x_2)] dx_1 dx_2,$$

where $f_{s,t}$ is the joint probability density function of X_s and X_t . Now let $r = f_{s,t} - f_s f_t$ and $h_1 = h_2 = B_{1j}$. Since $|R(x_1, x_2)| = |F_{s,t}(x_1, x_2) - F_s(x_1) F_t(x_2)| \leq \alpha(|s-t|)$ by part (b) of Assumption 5.2 and since $\int |\dot{B}_{1j}| \leq c$ for all j and some constant $c > 0$,

Lemma 1 yields that $|\text{Cov}(B_{1j}(X_s), B_{1j}(X_t))| = \mathcal{O}(1)\alpha(|s - t|)$. Hence,

$$\begin{aligned} k^2 E[\text{tr}(V'_{01} V_{01})] &= \mathcal{O}(1) \left(\frac{k}{T}\right)^2 \sum_{i,j} \sum_{s,t \in I_i} B_{0i}(s/T) B_{0i}(t/T) \alpha(|s - t|) \\ &= \mathcal{O}(1) \left(\frac{k}{T}\right)^2 \sum_{i,j} T_i \sum_{h=0}^{\infty} \alpha(h) = \mathcal{O}(1) \left(\frac{k}{T}\right)^2 \frac{k^2 T}{k} \sum_{h=0}^{\infty} \alpha(h) \\ &= \mathcal{O}(1) \frac{k^3}{T} \sum_{h=0}^{\infty} \alpha(h) = o(1). \end{aligned}$$

In the above arguments, it was used that $B_0(u) \leq 1$ for all u and that $\sum_{h=0}^{\infty} \alpha(h) < \infty$ by 5.2. Next, consider

$$k^2 E[\text{tr}(V'_{11} V_{11})] = \left(\frac{k}{T}\right)^2 \sum_{i,j} \sum_{s,t=1}^T \text{Cov}(B_{1i}(X_s) B_{1j}(X_s), B_{1i}(X_s) B_{1j}(X_t)).$$

Using Lemma 1 again, this time with $r = f_{s,t} - f_s f_t$ and $h_1 = h_2 = B_{1i} B_{1j}$, it follows that $|\text{Cov}(B_{1i}(X_s) B_{1j}(X_s), B_{1i}(X_s) B_{1j}(X_t))| \leq \alpha(|s - t|)$ for all i and j for some constant $c > 0$. Thus,

$$k^2 E[\text{tr}(V'_{11} V_{11})] = \mathcal{O}(1) \left(\frac{k}{T}\right)^2 \sum_{|i-j| \leq 2d} \sum_{s,t=1}^T \alpha(|s - t|),$$

where the double over i and j has been reduced to a sum over $|i - j| \leq 2d$, since $B_{1i}(x) B_{1j}(x) = 0$ whenever $|i - j| > 2d$. Consequently,

$$k^2 E[\text{tr}(V'_{11} V_{11})] = \mathcal{O}(1) \frac{k^3}{T} \sum_{h=0}^{\infty} \alpha(h) = o(1).$$

This completes the proof. □

Lemma 3. *If Assumptions 5.1 and 5.2 are satisfied, then*

$$\begin{aligned} (a) \quad S_{T,1} &= \frac{1}{T} \sum_{t=1}^T (\hat{m}_t - \bar{m}_t)^2 = \mathcal{O}_P\left(\frac{k}{T}\right), \\ (b) \quad S_{T,2} &= \frac{1}{T} \sum_{t=1}^T (\bar{m}_t - m_t)^2 = \mathcal{O}_P(k^{-2\nu}), \end{aligned}$$

as $T \rightarrow \infty$.

Proof. (a) Note that

$$S_{T,1} = (\hat{\beta} - \bar{\beta})' D (\hat{\beta} - \bar{\beta}) \leq \|D^{-1}\| \|\hat{\beta} - \bar{\beta}\|^2 = \mathcal{O}_P(k) \left\| \frac{1}{T} \sum_{t=1}^T B_t Z_t \right\|^2, \quad (1.1)$$

where the $\mathcal{O}_P(k)$ rate follows from Lemma 2. Now

$$\left\| \frac{1}{T} \sum_{t=1}^T B_t Z_t \right\|^2 = \left\| \frac{1}{T} \sum_{t=1}^T B_0(t/T) Z_t \right\|^2 + \left\| \frac{1}{T} \sum_{t=1}^T \Phi_1(X_t) Z_t \right\|^2.$$

Observe that for the first squared norm on the right-hand side, it holds that

$$E \left[\left\| \frac{1}{T} \sum_{t=1}^T B_t Z_t \right\|^2 \right] = \frac{1}{T^2} \sum_j \sum_{t=1}^T B_{0j}^2(t/T) \sigma_Z^2 \leq \frac{\sigma_Z^2}{T}, \quad (1.2)$$

on account of $\sum_j B_{0j}^2(u) \leq 1$. Next, using $\Phi_1(x) = \Psi_1 B_1(x)$, it follows that

$$\left\| \frac{1}{T} \sum_{t=1}^T \Phi_1(X_t) Z_t \right\|^2 \leq \|\Psi_1\|^2 \left\| \frac{1}{T} \sum_{t=1}^T B_1(X_t) Z_t \right\|^2 \leq \left\| \frac{1}{T} \sum_{t=1}^T B_1(X_t) Z_t \right\|^2,$$

where the last step is justified since $\|\Psi_1\| \leq 1$. Now

$$\begin{aligned} E \left[\left\| \frac{1}{T} \sum_{t=1}^T B_1(X_t) Z_t \right\|^2 \right] &= \sum_j E \left[\left(\frac{1}{T} \sum_{t=1}^T B_{1j}(X_t) Z_t \right)^2 \right] \\ &= \sum_j \frac{\sigma_Z^2}{T^2} \sum_{t=1}^T E[B_{1j}^2(X_t)] \leq \frac{\sigma_Z^2}{T}, \end{aligned} \quad (1.3)$$

where the inequality is implied by

$$E[B_{1j}(X_s) B_{1j}(X_t) Z_s Z_t] = \begin{cases} 0, & s \neq t, \\ \sigma_Z^2 E[B_{1j}^2(X_t)], & s = t, \end{cases}$$

and $\sum_j B_{1j}^2(x) \leq 1$ for any x . Equations (1.1)–(1.3) consequently show that $S_{T,1} = \mathcal{O}_P(k/T)$ and part (a) is verified.

(b) There are splines s_0^* and s_1^* with k equispaced knots (of degree d) such that

$$\|g_0 - s_0^*\|_\infty = \mathcal{O}_P(k^{-\nu}) \quad \text{and} \quad \|g_1 - s_1^*\|_\infty = \mathcal{O}_P(k^{-\nu}).$$

Let $\tilde{s}_1 = s_1^* - \int s_1^*(x) dx$, then $\int \tilde{s}_1(x) dx = 0$. Since $\bar{g}_0(t/T) = \bar{\beta}'_0 B_0(t/T)$ and $\bar{g}_1(X_t) = \bar{\beta}'_1 \Phi_1(X_t)$ are obtained by minimizing $T^{-1} \sum_t [m_t - \bar{\beta}'_0 B_0(t/T) - \bar{\beta}'_1 \Phi_1(X_t)]^2$, it follows that

$$S_{T,2} = \frac{1}{T} \sum_{t=1}^T [m_t - \bar{\beta}'_0 B_0(t/T) - \bar{\beta}'_1 \Phi_1(X_t)]^2$$

$$\begin{aligned}
&\leq \frac{1}{T} \sum_{t=1}^T [m_t - s_0^*(t/T) - \tilde{s}_1^*(X_t)]^2 \\
&= \frac{1}{T} \sum_{t=1}^T \left[\{g_0(t/T) - s_0^*(t/T)\} + \{g_1(X_t) - \tilde{s}_1^*(X_t)\} \right]^2 \\
&\leq \frac{2}{T} \sum_{t=1}^T \{g_0(t/T) - s_0^*(t/T)\}^2 + \frac{2}{T} \sum_{t=1}^T \{g_1(X_t) - \tilde{s}_1^*(X_t)\}^2 \\
&\leq 2\|g_0 - s_0^*\|_\infty^2 + 2\|g_1 - \tilde{s}_1^*\|_\infty^2.
\end{aligned}$$

Now, $\|g_0 - s_0^*\|_\infty = \mathcal{O}_P(k^{-\nu})$ by construction and $\|g_1 - \tilde{s}_1^*\|_\infty \leq \|g_1 - s_1^*\|_\infty + |\int s_1^*| \leq 2\|g_1 - s_1^*\|_\infty = \mathcal{O}_P(k^{-\nu})$ because $|\int s_1^*| = |\int g_1 - \int s_1^*| \leq \|g_1 - s_1^*\|_\infty$. This shows that $S_{T,2} = \mathcal{O}_P(k^{-2\nu})$, as required for part (b). \square

Proof of Theorem 5.1. The assertion follows from Lemma 3, noticing that the two rates are balanced if k is of the order $T^{1/(2\nu+1)}$. In this case, $k/T = k^{-2\nu} = T^{-2\nu/(2\nu+1)}$. This completes the proof. \square

References

- [1] Noguchi, K., Aue, A., and Burman, P. (2014), “Exploratory Analysis and Modeling of Stock Returns,” *Journal of Computational and Graphical Statistics*.
- [2] Bradley, R.C. (2005), “Basic Properties of Strong Mixing Conditions: a Survey and Some Open Questions,” *Probability Surveys*, 2, 107–144.
- [3] Burman, P. (1991), “Regression Function Estimation from Dependent Observations,” *Journal of Multivariate Analysis*, 36, 263–279.