How to Estimate Eigenvalues and Eigenvectors of Covariance Function when Parameters Are Estimated

Kohtaro Hitomi
Kyoto Institute of Technology*

6th March 2002

Abstract

This paper introduces an estimation method of the eigenvalues and eigenvectors from the sample covariance function that involves estimated parameters. We prove that the estimated eigenvalues and eigenvectors are consistent.

It also includes the approximation method of the critical values of ICM test using estimated eigenvalues.

*Department of Architecture and Design, Kyoto Institute of Technology, Sakyo, Kyoto 606-8585 Japan.

1 Introduction

This article proposes an estimation method for eigenvalues and eigenvectors of covariance functions. This is an extension of Principal Components Analysis (PCA) of finite dimensional random vectors to random "functions."

Over the past few decades, a considerable number of studies have been conducted on functional data analysis in statistics. Some surveys of the field are given in Ramsey and Silverman (1997) and Ramsey and Daizill (1991).

Such methods might be useful in econometrics. For example, suppose the ICM test (Bierens and Ploberger 1997) for functional form $E[y|x] = Q(\theta, x)$. This test uses the following random function on $\Xi \subset R^k$,

$$z_n = \frac{1}{\sqrt{n}} \sum (y_t - Q(\hat{\theta}, x_t)) \exp(\xi' x_t) \quad \xi \in \Xi \subset R^k,$$

and the null distribution of the test statistic depends on the eigenvalues of the covariance function of z_n. Thus if we can estimate the eigenvalues, critical values of the test statistic are easily calculated. See Hitomi (2000) for detail.

In many econometric models, sample covariance functions include estimated parameters and are defined on $R^k \times R^k$ instead of $R^1 \times R^1$. Ramsey and Silverman (1997) have used a discrete approximation method for estimating eigenvalues and eigenvector of covariance functions on a subset of $R^1 \times R^1$. It is difficult, however, to extend their method to higher dimensions. Dauxois, Pousse and Romain (1982) have investigated the convergence of estimated eigenvalues and eigenvectors of sample covariance functions on separable Hilbert space. Their sample covariance function has not included estimated parameters and they have proposed no estimation method, however.

This article solves the above problems for applying the functional data analysis to econometric models. It proposes an estimation method of eigenvalues and eigenvectors from a sample covariance function on a subset of $R^k \times R^k$, which involves estimated parameters, and proves consistency of estimated eigenvalues and eigenvectors.

The plan of the paper is the following. Section 2 explains the model and the estimation method. The consistency of the estimated eigenvalues and eigenvectors is proved under high-level assumptions in section 3. As the example, low-level assumptions for the ICM test are derived in section 3. The last section is concluding remarks. Some mathematical proofs are included in Appendix.

2 Model and Estimation Method

Suppose that we are interested in eigenvalues and eigenvectors of a continuous covariance function $\Gamma_0(\xi_1, \xi_2)$ on $\Xi \times \Xi$, where Ξ is a compact subset of R^k, $k \geq 1$. We assume that Γ_0 satisfies

$$\iint |\Gamma_0(\xi_1, \xi_2)| \, d\mu(\xi_1) d\mu(\xi_2) < \infty,$$

where $\mu(\xi)$ is a known probability measure on Ξ.

An eigenvalue λ and an eigenvector $\psi(\xi)$ of $\Gamma_0(\xi_1, \xi_2)$ are the solution of characteristic equation

$$\int \Gamma_0(\xi_1, \xi)\psi(\xi_1) d\mu(\xi_1) = \lambda\psi(\xi). \tag{1}$$

Assume that there is a consistent estimator Γ_n of Γ_0 , which involves an estimate of unknown parameters $\theta_0 \in \Theta \subset R^q$,

$$\Gamma_n(\hat{\theta}, \xi_1, \xi_2) = \frac{1}{n} \sum_{t=1}^{n} a_n(\hat{\theta}, w_t, \xi_1) a_n(\hat{\theta}, w_t, \xi_2), \tag{2}$$

where $a_n(.) : \Theta \times R^d \times \Xi \to R^1$ is a function that satisfies $\|a_n(\theta, w_t, \xi)\|_2 < \infty$ for all $(\theta, w_t) \in \Theta \times R^d$, $w_t \in R^d$ is an i.i.d. random variable and $\hat{\theta}$ is a consistent estimator of θ_0.

We begin by introducing some notation. $< f, g >$ is the inner product in $L_2(\mu(\xi))$, i.e.

$$< f, g >= \int f(\xi) g(\xi) d\mu(\xi),$$

and $\|\cdot\|_2$ is $L_2(\mu(\xi))$ norm. Let $\Phi : L_2(\mu(\xi)) \to L_2(\mu(\xi))$ be a bounded linear operator, we write Φf when we apply operator Φ to $f \in L_2(\mu(\xi))$. Thus $\Gamma_n f$ implies

$$\Gamma_n f = \int \Gamma_n(\hat{\theta}, \xi_1, \xi) f(\xi_1) d\mu(\xi_1).$$

$\|\cdot\|_F$ is the uniform operator norm, i.e.

$$\|\Phi\|_F = \sup_{\|f\|_2 = 1} \|\Phi f\|_2 .$$

For notational simplicity, sometime we abbreviate $a_n(\hat{\theta}, w_i, \xi)$ by $a_i(\xi)$ or a_i.

We estimate eigenvalues and eigenvectors of $\Gamma_n(\hat{\theta}, \xi_1, \xi_2)$. The operator Γ_n maps arbitrary function $f(\xi) \in L_2(\mu(x))$ on a finite dimensional space, which is spanned by $\{a_1(\xi), a_2(\xi), \ldots, a_n(\xi)\}$, since

$$\begin{aligned}
\Gamma_n f &= \int \Gamma_n(\hat{\theta}, \xi_1, \xi) f(\xi_1) d\mu(\xi_1) \\
&= \int \frac{1}{n} \sum_{t=1}^{n} a_t(\xi_1) a_t(\xi) f(\xi_1) d\mu(\xi_1) \\
&= \frac{1}{n} \sum_{t=1}^{n} \langle a_t, f \rangle a_t(\xi) \\
&= \sum_{t=1}^{n} b_t a_t(\xi),
\end{aligned}$$

where $b_t = \frac{1}{n} \langle a_t, f \rangle$. Let $\mathcal{H}_n$ be the space that is spanned by $\{a_1(\xi), a_2(\xi), \ldots, a_n(\xi)\}$.

Let λ_n and ψ_n be a solution of the sample characteristic equation

$$\int \Gamma_n(\hat{\theta}, \xi_1, \xi) \psi_n(\xi_1) d\mu(\xi_1) = \lambda_n \psi_n(\xi) \tag{3}$$

$$\Leftrightarrow \quad \Gamma_n \psi_n = \lambda_n \psi_n. \tag{4}$$

ψ_n is a linear combination of $\{a_1, a_2, \ldots, a_n\}$ since $\Gamma_n \psi_n \in \mathcal{H}_n$. Therefore we can express ψ_n as

$$\psi_n = \sum_{t=1}^{n} \alpha_t a_t(\xi)$$

and put it into (3), we get

$$\begin{aligned}
& \Gamma_n \psi_n &&= \lambda_n \sum_{t=1}^{n} \alpha_t a_t \\
\Leftrightarrow \quad & \tfrac{1}{n} \textstyle\sum_{t=1}^{n} \langle a_t, \psi_n \rangle a_t &&= \lambda_n \sum_{t=1}^{n} \alpha_t a_t \\
\Leftrightarrow \quad & \textstyle\sum_{t=1}^{n} \langle a_t, \sum_{s=1}^{n} \alpha_s a_s \rangle a_t &&= \lambda_n \sum_{t=1}^{n} \alpha_t a_t \\
\Leftrightarrow \quad & \textstyle\sum_{t=1}^{n} \sum_{s=1}^{n} \alpha_s \langle a_t, a_s \rangle a_t &&= \lambda_n \sum_{t=1}^{n} \alpha_t a_t.
\end{aligned}$$

Comparing the coefficients of a_t, we get

$$\sum_{s=1}^{n} \alpha_s \langle a_t, a_s \rangle = \lambda_n \alpha_t \tag{5}$$

Now we define the $n \times n$ matrix A such that the (i, j) element of A is $\langle a_i, a_j \rangle$,

$$A = \{\langle a_i, a_j \rangle\}$$

and the $m \times 1$ vector α as $\alpha = (\alpha_1, \ldots, \alpha_n)'$. Then the matrix expression of (5) is

$$A\alpha = \lambda_n \alpha. \tag{6}$$

This implies that an eigenvalue of (3) is an eigenvalue of matrix A and an eigenvector of (3) is $\psi_n = \sum_{t=1}^{n} \alpha_t a_t$, where α_t is the t-th element of the eigenvector of matrix A corresponding λ_n.

We got the following lemma,

Lemma 1. *Suppose* $\|a_n(\theta, w_t, \xi)\|_2 < \infty$ *for all* $(\theta, w_t) \in \Theta \times R^d$.

The following statements are equivalent,

1. λ_n *and* ψ_n *is a solution of the characteristic equation*

$$\int \Gamma_n(\hat{\theta}, \xi_1, \xi)\psi_n(\xi_1)d\mu(\xi_1) = \lambda_n \psi_n(\xi).$$

2. λ_n *is a eigenvalue of* A, $\psi_n = \sum_{t=1}^{n} \alpha_t a_t(\xi)$, *where* α_t *is the* t*-th element of the corresponding eigenvector of* λ_n *and*

$$A = \{\langle a_i, a_j \rangle\}.$$

3 Consistency

We assume the following two sets of high-level assumptions. An example of low-level assumptions is discussed in section 4.

Assumption a.s.

1. (uniform convergence) Θ and Ξ are compact subset of R^q and R^k respectively. Let $\Gamma_n(\theta, \xi_1, \xi_2)$ be

$$\Gamma_n(\theta, \xi_1, \xi_2) = \frac{1}{n}\sum_{t=1}^{n} a_n(\theta, w_t, \xi_1)a_n(\theta, w_t, \xi_2).$$

   $\Gamma_n(\theta, \xi_1, \xi_2)$ converges to a nonrandom continuous function $\Gamma(\theta, \xi_1, \xi_2)$ a.s. uniformly on $\Theta \times \Xi \times \Xi$. And $\Gamma(\theta_0, \xi_1, \xi_2) = \Gamma_0(\xi_1, \xi_2)$ for all $(\xi_1, \xi_2) \in \Xi \times \Xi$.

2. (consistency of $\hat{\theta}$) $\hat{\theta}$ converges to $\theta_0 \in \Theta$ a.s.

Assumption pr

1. (uniform convergence) Θ and Ξ are compact subset of R^q and R^k respectively. Let $\Gamma_n(\theta, \xi_1, \xi_2)$ be

$$\Gamma_n(\theta, \xi_1, \xi_2) = \frac{1}{n} \sum_{t=1}^{n} a_n(\theta, w_t, \xi_1) a_n(\theta, w_t, \xi_2).$$

   $\Gamma_n(\theta, \xi_1, \xi_2)$ converges to a nonrandom continuous function $\Gamma(\theta, \xi_1, \xi_2)$ in probability uniformly on $\Theta \times \Xi \times \Xi$. And $\Gamma(\theta_0, \xi_1, \xi_2) = \Gamma_0(\xi_1, \xi_2)$ for all $(\xi_1, \xi_2) \in \Xi \times \Xi$.

2. (consistency of $\hat{\theta}$) $\hat{\theta}$ converges to $\theta_0 \in \Theta$ in probability.

The first set of assumptions corresponds to almost sure convergence of the eigenvalues and the eigenvectors and the second set of assumptions corresponds to convergence in probability.

Let $\{\lambda_i\}$ be the decreasing sequence of the non-null eigenvalues of $\Gamma_0(\xi_1, \xi_2)$ and $\{\psi_i\}$ be the corresponding sequence of the eigenvectors of $\Gamma_0(\xi_1, \xi_2)$, $\{\lambda_{ni}\}$ be the decreasing sequence of the non-null eigenvalues of $\Gamma_n(\hat{\theta}, \xi_1, \xi_2)$ and $\{\psi_{ni}\}$be the corresponding sequence of the eigenvectors of $\Gamma_n(\hat{\theta}, \xi_1, \xi_2)$ and define the set I_i as the following,

$$I_i = \{j | \lambda_i = \lambda_j\}, \quad |I_i| = k_i.$$

Proposition 2. *Suppose Assumption a.s. is satisfied. When λ_i is of order k_i, there are k_i sequences $\{\lambda_{nj} | j \in I_i\}$ converging to λ_i a.s. If Assumption pr is satisfied instead of Assumption a.s., $\{\lambda_{nj} | j \in I_i\}$ converge to λ_i in probability. Moreover the convergence is uniform in j.*

Proof. First we think the almost sure convergence case.

$$\begin{aligned}
& \sup_{\xi_1, \xi_2} \left| \Gamma_n(\hat{\theta}, \xi_1, \xi_2) - \Gamma_0(\xi_1, \xi_2) \right| \\
\leq \; & \sup_{\xi_1, \xi_2} \left| \Gamma_n(\hat{\theta}, \xi_1, \xi_2) - \Gamma(\hat{\theta}, \xi_1, \xi_2) \right| + \sup_{\xi_1, \xi_2} \left| \Gamma(\hat{\theta}, \xi_1, \xi_2) - \Gamma(\theta_0, \xi_1, \xi_2) \right| \\
\leq \; & \sup_{\theta, \xi_1, \xi_2} \left| \Gamma_n(\theta, \xi_1, \xi_2) - \Gamma(\theta, \xi_1, \xi_2) \right| + \sup_{\xi_1, \xi_2} \left| \Gamma(\hat{\theta}, \xi_1, \xi_2) - \Gamma(\theta_0, \xi_1, \xi_2) \right|.
\end{aligned}$$

The first term of the last inequality converges to zero a.s. since $\Gamma_n(\theta, \xi_1, \xi_2)$ converges to $\Gamma(\theta, \xi_1, \xi_2)$ a.s. uniformly on $\Theta \times \Xi \times \Xi$. $\Gamma(\theta, \xi_1, \xi_2)$ is uniformly continuous because $\Gamma(\theta, \xi_1, \xi_2)$ is continuous on $\Theta \times \Xi \times \Xi$ and $\Theta \times \Xi \times \Xi$ is a compact set. Thus the second term converges to zero a.s. since $\hat{\theta} \to \theta_0$ a.s. Therefore $\Gamma_n(\hat{\theta}, \xi_1, \xi_2) \to \Gamma_0(\xi_1, \xi_2)$ a.s. uniformly on $\Xi \times \Xi$.

Let think the distance between Γ_n and Γ_0 in the uniform operator norm on

$L_2(\mu(\xi)) \to L_2(\mu(\xi))$.

$$\begin{aligned}
& \|\Gamma_n - \Gamma_0\|_F \\
= & \sup_{\|f\|_2=1} \|\Gamma_n f - \Gamma_0 f\|_2 \\
\leq & \sup_{\|f\|_2=1} \left\| \int \left(\Gamma_n(\hat{\theta}, \xi_1, \xi) - \Gamma_0(\xi_1, \xi) \right) f(\xi_1) d\mu(\xi_1) \right\|_2 \\
\leq & \sup_{\|f\|_2=1} \left(\int \left(\int \left(\Gamma_n(\hat{\theta}, \xi_1, \xi) - \Gamma_0(\xi_1, \xi) \right) f(\xi_1) d\mu(\xi_1) \right)^2 d\mu(\xi) \right)^{1/2} \\
\leq & \sup_{\|f\|_2=1} \left(\int \left(\left(\int \left| \Gamma_n(\hat{\theta}, \xi_1, \xi) - \Gamma_0(\xi_1, \xi) \right|^2 d\mu(\xi_1) \right)^{1/2} \left(\int f(\xi_1)^2 d\mu(\xi_1) \right)^{1/2} \right)^2 d\mu(\xi) \right)^{1/2} \\
= & \sup_{\|f\|_2=1} \left(\iint \left| \Gamma_n(\hat{\theta}, \xi_1, \xi) - \Gamma_0(\xi_1, \xi) \right|^2 d\mu(\xi_1) d\mu(\xi) \right)^{1/2} \|f(\xi)\|_2 \\
\leq & \sup_{\xi_1, \xi_2} \left| \Gamma_n(\hat{\theta}, \xi_1, \xi) - \Gamma_0(\xi_1, \xi) \right| .
\end{aligned}$$

Thus Γ_n converges to Γ_0 a.s. in the uniform operator norm. The conclusion follows directly from lemma 5 of Donford and Schwartz (1988, p1091).

The proof of the convergence in probability case is similar. Since $\Gamma_n(\theta, \xi_1, \xi_2) \to \Gamma(\theta, \xi_1, \xi_2)$ uniformly on $\Theta \times \Xi \times \Xi$ in probability and $\hat{\theta} \to \theta_0$ in probability, then $\Gamma_n(\hat{\theta}, \xi_1, \xi_2) \to \Gamma_0(\xi_1, \xi_2)$ uniformly on $\Xi \times \Xi$ in probability. Thus any subsequence n_k in sequence n contains a further subsequence n_{k_i} such that $\Gamma_{n_{k_i}} \to \Gamma_0$ a.s. uniformly on $\Theta \times \Xi \times \Xi$. Therefore $\Gamma_{n_{k_i}} \to \Gamma_0$ a.s. in the uniform operator norm. Then apply lemma 5 of Donford and Schwartz (1988, p1091) to the sequence of operator $\Gamma_{n_{k_i}}$. Therefore for $j \in I_i$, $\lambda_{n_{k_i} j} \to \lambda_i$ a.s. This implies for $j \in I_i$, $\lambda_{nj} \to \lambda_i$ in probability. □

The following proposition is worth mentioning in passing.

Proposition 3. *Soppose Assumption a.s. (Assumption pr)is satisfied. Then*

$$\begin{aligned}
\sum_{j=1}^{n} |\lambda_{nj}| & \to \sum |\lambda_j| \quad a.s. \, (in \, pr) \\
\sum_{j=1}^{n} |\lambda_{nj}|^2 & \to \sum |\lambda_j|^2 \quad a.s. \, (in \, pr).
\end{aligned}$$

Proof. Since

$$\sum_{j=1} |\lambda_j| = \int \Gamma_0(\xi, \xi) d\mu(\xi)$$

and

$$\sum_{j=1} |\lambda_j|^2 = \iint \Gamma_0(\xi_1, \xi_2)^2 d\mu(\xi_1) d\mu(\xi_2),$$

$$\begin{aligned}
\left| \sum_{j=1}^{n} |\lambda_{nj}| - \sum_{j=1}^{\infty} |\lambda_j| \right| &= \left| \int \Gamma_n(\xi, \xi) d\mu(\xi) - \int \Gamma_0(\xi, \xi) d\mu(\xi) \right| \\
&\leq \int |\Gamma_n(\xi, \xi) - \Gamma_0(\xi, \xi)| \, d\mu(\xi) \\
&\leq \sup_{\xi_1, \xi_2} \|\Gamma_n(\xi_1, \xi_2) - \Gamma_0(\xi_1, \xi_2)\| \\
&\to 0 \quad a.s. \, (in \, pr).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\left| \sum_{j=1}^{n} |\lambda_{nj}|^2 - \sum_{j=1}^{\infty} |\lambda_j|^2 \right| &= \left| \iint \Gamma_n(\xi_1, \xi_2)^2 d\mu(\xi_1) d\mu(\xi_2) - \iint \Gamma_n(\xi_1, \xi_2)^2 d\mu(\xi_1) d\mu(\xi_2) \right| \\
&= \left| \iint \left((\Gamma_0 - (\Gamma_0 - \Gamma_n))^2 - \Gamma_0^2 \right) d\mu(\xi_1) d\mu(\xi_2) \right| \\
&\leq 2 \left| \iint \Gamma_0(\Gamma_0 - \Gamma_n) d\mu(\xi_1) d\mu(\xi_2) \right| + \left| \iint (\Gamma_0 - \Gamma_n)^2 d\mu(\xi_1) d\mu(\xi_2) \right| \\
&\leq 2 \sup_{\xi_1, \xi_2} |\Gamma_0(\xi_1, \xi_2) - \Gamma_n(\xi_1, \xi_2)| \iint \Gamma_0 d\mu(\xi_1) d\mu(\xi_2) \\
&\quad + \sup_{\xi_1, \xi_2} |\Gamma_0(\xi_1, \xi_2) - \Gamma_n(\xi_1, \xi_2)|^2 \\
&\to 0 \quad a.s. \, (in \, pr).
\end{aligned}$$

□

Let $\phi_i(\xi)$ and $\phi_{ni}(\xi)$ be normalized eigenvectors corresponding to λ_i and λ_{ni} respectively, thus

$$\phi_i(\xi) = \frac{\psi_i(\xi)}{\|\psi_i(\xi)\|_2}, \ \phi_{ni}(\xi) = \frac{\psi_{ni}(\xi)}{\|\psi_{ni}(\xi)\|_2} \text{ and } \|\phi_i(\xi)\| = \|\phi_{ni}(\xi)\| = 1 \text{ for all } i \text{ and } n.$$

When the multiplicity of the eigenvalue λ_i is larger than one, corresponding normalized eigenvectors are not uniquely determined. Therefore we could not have a convergence property for each corresponding eigenvector sequence. However the corresponding eigenspace is unique. Therefore next we think about the convergence property of the projection operators that maps functions on the eigenspace corresponding to the eigenvalue λ_i.

Let $P_j(\xi_1, \xi_2) = \sum_{k \in I_j} \phi_k(\xi_1) \phi_k(\xi_2)$ be an orthogonal projection operator that maps $L_2(\mu(\xi))$ on the eigenspace corresponding to the eigenvalue λ_i, and

$P_{nj}(\xi_1,\xi_2)=\sum_{k\in I_j}\phi_{nk}(\xi_1)\phi_{nk}(\xi_2)$ be an estimator of P_j. Note that the covariance function $\Gamma(\xi_1,\xi_2)$ and $\Gamma_n(\hat{\theta},\xi_1,\xi_2)$ could be decomposed as the followings,

$$\begin{aligned}\Gamma(\xi_1,\xi_2) &= \sum_{j=1}^{\infty}\lambda_j P_j(\xi_1,\xi_2)\\ \Gamma_n(\hat{\theta},\xi_1,\xi_2) &= \sum_{j=1}^{m}\lambda_{nj} P_{nj}(\xi_1,\xi_2).\end{aligned}$$

Proposition 4. *Suppose Assumption a.s. (Assumption pr) is satisfied. Then for each j, P_{nj} converges P_j a.s. (in probability) in the uniform operator norm.*

Proof. From the proof of Proposition 2, Γ_n converges to Γ a.s. (in probability) in the uniform operator norm. In the proof of Proposition 3 in Dauxois, Pousse and Romain (1982), they have shown that $\|P_{nj}-P_j\|_F\to 0$ if $\|\Gamma_n-\Gamma\|_F\to 0$. We have proved the proposition. □

When the multiplicity of the eigenvalue λ_i is equal to one, we could talk about the convergence of eigenvectors. However if $\phi_i(\xi)$ is the normalized eigenvector corresponding to λ_i, $-\phi_i(\xi)$ also satisfies the characteristic equation and $\|-\phi_i(\xi)\|_2=1$. So we need further specification. Choose one of normalized eigenvector, named $\phi_i^1(\xi)$, corresponding to λ_i and choose a sequence of eigenvectors $\phi_{ni}^1(\xi)$ corresponding to λ_{nj} such that $\left\langle\phi_i^1(\xi),\phi_{ni}^1(\xi)\right\rangle\geq 0$.

Corollary 5. *Suppose Assumption a.s. (or Assumption pr) is satisfied and the multiplicity of the eigenvalue λ_i is one. Then $\phi_{ni}^1(\xi)$ converges to $\phi_i^1(\xi)$ a.s. (in probability) on $L_2(\mu(\xi))$.*

Proof. Since the multiplicity of the eigenvalue λ_i is 1, $P_i(\xi_1,\xi_2)=\phi_i^1(\xi_1)\phi_i^1(\xi_2)$ and $P_{ni}(\xi_1,\xi_2)=\phi_{ni}^1(\xi_1)\phi_{ni}^1(\xi_2)$. Then

$$\begin{aligned}\|P_i-P_{nj}\|_F &= \sup_{\|f\|_2=1}\left\|\int\left(\phi_i^1(\xi_1)\phi_i^1(\xi)-\phi_{ni}^1(\xi_1)\phi_{ni}^1(\xi)\right)f(\xi_1)d\mu(\xi_1)\right\|_2\\ &= \sup_{\|f\|_2=1}\left\|\left\langle\phi_{ni}^1,f\right\rangle\phi_{ni}^1-\left\langle\phi_i^1,f\right\rangle\phi_i^1\right\|_2\\ &\geq \left\|\left\langle\phi_{ni}^1,\phi_i^1\right\rangle\phi_{ni}^1-\left\langle\phi_i^1,\phi_i^1\right\rangle\phi_i^1\right\|_2\\ &= \left(1-\left\langle\phi_{ni}^1,\phi_i^1\right\rangle^2\right)^{1/2}\\ &= \left(\frac{1}{2}\left\|\phi_{ni}^1-\phi_i^1\right\|_2^2\left(1+\left\langle\phi_{ni}^1,\phi_i^1\right\rangle\right)\right)^{1/2}\geq 0.\end{aligned}$$

Now $1+\left\langle\phi_{ni}^1,\phi_i^1\right\rangle\geq 1$ by the construction of ϕ_{ni}^1 and $\|P_i-P_{nj}\|_F$ converges to zero a.s. (in probability). Therefore $\left\|\phi_{ni}^1-\phi_i^1\right\|_2$ converges to zero a.s. (in probability). □

4 Example: ICM test

Let $w_t = (y_t, x_t)$ be a sequence of i.i.d. random variables on $R^1 \times R^d$. The ICM test statistics for testing $H_0 : y_t = Q(x_t, \theta_0) + u_t$ uses the random function

$$z_n(\xi) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (y - Q(x_t, \hat{\theta})) \exp(\xi' \Phi(x_t)),$$

where $\hat{\theta}$ is the nonlinear least squares estimator of θ_0 and $\Phi(x)$ is a bounded one-to-one function on $R^d \to R^d$. The test statistics is

$$T_n = \int z_n(\xi)^2 d\mu(\xi).$$

Under Assumption A in Bierens (1990), which is also included in Appendix A for readers' convenience, $z_n(\xi)$ converges a Gaussian process $z(\xi)$ with covariance function

$$\begin{aligned}\Gamma_0(\xi_1, \xi_2) \quad &= \quad E\left[u_t^2 \left(\exp(\xi_1' \Phi(x_t)) - b(\theta_0, \xi_1) A^{-1} \frac{\partial Q(x_t, \theta_0)}{\partial \theta}\right)\right. \\ &\quad \left. \times \left(\exp(\xi_2' \Phi(x_t)) - b(\theta_0, \xi_2) A^{-1} \frac{\partial Q(x_t, \theta_0)}{\partial \theta}\right)\right],\end{aligned}$$

where $b(\theta_0, \xi) = E[(\partial/\partial\theta) Q(x_t, \theta_0) \exp(\xi' \Phi(x_t))]$ and

$$A = E\left[\frac{\partial Q(x_t, \theta_0)}{\partial \theta} \frac{\partial Q(x_t, \theta_0)}{\partial \theta'}\right].$$

A natural estimator of $\Gamma_0(\xi_1, \xi_2)$ is

$$\begin{aligned}\Gamma_n(\hat{\theta}, \xi_1, \xi_2) \quad &= \quad \frac{1}{n} \sum_{t=1}^{n} \left(y_t - Q(x_t, \hat{\theta})\right)^2 \left(\exp(\xi_1' \Phi(x_t)) - b_n(\hat{\theta}, \xi_1) A_n(\hat{\theta})^{-1} \frac{\partial Q(x_t, \hat{\theta})}{\partial \theta}\right) \\ &\quad \times \left(\exp(\xi_2' \Phi(x_t)) - b_n(\hat{\theta}, \xi_2) A_n(\hat{\theta})^{-1} \frac{\partial Q(x_t, \hat{\theta})}{\partial \theta}\right),\end{aligned}$$

where

$$\begin{aligned}b_n(\theta, \xi) \quad &= \quad \frac{1}{n} \sum_{t=1}^{n} \frac{\partial Q(x_t, \theta)}{\partial \theta} \exp(\xi' \Phi(x_t)), \\ A_n(\theta) \quad &= \quad \frac{1}{n} \sum_{t=1}^{n} \frac{\partial Q(x_t, \theta)}{\partial \theta} \frac{\partial Q(x_t, \theta)}{\partial \theta'}.\end{aligned}$$

The low level assumptions for the consistency of the estimated eigenvalues and the eigenvectors are not so restrictive. The following assumptions, which were used in Bierens (1990) and Bierens and Ploberger (1997) to derive the asymptotic distribution of the ICM test, ensure the consistency of the estimated eigenvalues, the projection operators and the eigenvectors.

Assumption ICM

1. Assumption A is satisfied
2. Ξ is a compact subset of R^d. The probability measure $\mu(\xi)$ is chosen absolutely continuous with respect to Lebesgue measure.

Proposition 6. *Assumption ICM implies Assumption a.s.*

Proof. See Appendix B. □

One application of estimated eigenvalues is to estimate critical values of the ICM test. Bierens and Ploberger (1997) could not get critical values of their test since it depends on the distribution of independent variables. They reported only case independent upper bounds of the critical values. It might be too conservative. It might be possible to apply Hansen's bootstrapping method (Hansen 1996), however it is very time consuming. With the estimated eigenvalues we could estimate the critical values.

As shown in Bierens and Ploberger (1997), under the null hypothesis the test statistics T_n converges to

$$T_n \xrightarrow{d} q = \sum_{i=1}^{\infty} \lambda_i Z_i^2,$$

where λ_is are the eigenvalues of Γ_0 and Z_i's are independent standard normal random variables.

Construct new random variable q_n based on the estimated eigenvalues as the following,

$$q_n = \sum_{i=1}^{n} \lambda_{ni} Z_i^2,$$

where λ_{ni}'s are the eigenvalues of Γ_n. The critical values of T_n could be estimated by the critical values of q_n. The following theorem justifies the above method.

Theorem 7. *Suppose Assumption ICM is satisfied. Let* $\{\lambda_j : j = 1, 2, \ldots, \infty\}$ *and* $\{\lambda_{nj} : j = 1, 2, \ldots, n\}$ *be the decreasing sequences of the eigenvalues of* $\Gamma_0(\xi_1, \xi_2)$ *and* $\Gamma_n(\hat{\theta}, \xi_1, \xi_2)$ *respectively and* $\phi_q(t)$ *and* $\phi_{q_n}(t)$ *be the characteristic function of* q *and the conditional characteristic function of* q_n *on* $\Gamma_n(\hat{\theta}, \xi_1, \xi_2)$ *respectively. Then*

$$\phi_{q_n}(t) \to \phi_q(t) \quad a.s.$$

Proof. See Appendix B. $\square$

Consider local alternatives to the following form

$$H_n: \quad y_t = Q(x_t, \theta_0) + \frac{g(x_t)}{\sqrt{n}} + u_t, \tag{7}$$

where $g(x)$ satisfies $0 < E[g(x)^2] < \infty$.

Under the local alternative (7), $z_n(\xi)$ converges to a Gaussian process with mean function

$$\eta(\xi) = E\left[g(x_t)\left(\exp(\xi_1'\Phi(x_t)) - b_n(\hat{\theta}, \xi_1)A_n(\hat{\theta})^{-1}\frac{\partial Q(x_t, \hat{\theta})}{\partial \theta}\right)\right]$$

and the covariance function $\Gamma_0(\xi_1, \xi_2)$ as shown in Theorem 2 of Bierens and Ploberger (1997). Under the Assumption A and the local alternative assumpton, It could be shown that

$$\begin{aligned}
\sqrt{n}(\hat{\theta} - \theta_0) \quad &\xrightarrow{d} \quad N(E[g(x_t)\frac{\partial Q}{\partial \theta}], A^{-1}E\left[u_t^2 \frac{\partial Q}{\partial \theta}\frac{\partial Q}{\partial \theta'}\right]A^{-1}) \\
\hat{\theta} \quad &\to \quad \theta_0 \quad \text{a.s.}
\end{aligned}$$

$\Gamma_n(\theta, \xi_1, \xi_2)$ could be decomposed as the followings,

$$\begin{aligned}
\Gamma_n(\theta, \xi_1, \xi_2) \quad &= \quad \frac{1}{n}\sum_{t=1}^{n}\left(y_t - Q(x_t, \theta)\right)^2 w(x_t, \xi_1, \theta)w(x_t, \xi_2, \theta) \\
&= \quad \frac{1}{n}\sum_{t=1}^{n}\left(u_t + Q(x_t, \theta_0) - Q(x_t, \theta)\right)^2 w(x_t, \xi_1, \theta)w(x_t, \xi_2, \theta) \\
&\quad + \frac{2}{n}\sum_{t=1}^{n}\left(u_t + Q(x_t, \theta_0) - Q(x_t, \theta)\right)\frac{g(x_t)}{\sqrt{n}}w(x_t, \xi_1, \theta)w(x_t, \xi_2, \theta) \\
&\quad + \frac{1}{n}\sum_{t=1}^{n}\frac{g(x_t)^2}{n}w(x_t, \xi_1, \theta)w(x_t, \xi_2, \theta),
\end{aligned}$$

where

$$w(x_t, \xi_1, \theta) = \exp(\xi_1'\Phi(x_t)) - b_n(\theta, \xi_1)A_n(\theta)^{-1}\frac{\partial Q(x_t, \theta)}{\partial \theta}.$$

Using the same argument in the proof of Theorem 7, the first term converges to $\Gamma(\theta, \xi_1, \xi_2)$ a.s. uniformly on Ξ. The second and the third term converges to zero a.s. uniformly on Ξ.

Therefore $\Gamma_n(\hat{\theta}, \xi_1, \xi_2)$ converges to $\Gamma_0(\xi_1, \xi_2)$ a.s. uniformly on Ξ under the local alternative. This implies that we can consistently estimate λ_is and Theorem 7 also holds under the local alternative.

5 Concluding Remarks

This paper introduces an estimation method of the eigenvalues and eigenvectors from the sample covariance function, which involves estimated parameters. Then prove the consistency of the estimated eigenvalues and the eigenvectors.

One drawback of the method is that it needs considerable computation time. The estimated eigenvalues are the eigenvalues of the $n \times n$ matrix A where the i-j element of A is

$$\int a_n(\hat{\theta}, w_i, \xi) a_n(\hat{\theta}, w_j, \xi) d\mu(\xi).$$

In general, it might be difficult to integrate analytically. We need about order $n \times n$ times numerical integration to get A. Therefore some effective numerical integration method or approximation method are required.

Unsolved problem is the asymptotic distribution of the estimated eigenvalues and eigenvectors. Some central limit theorems in Hilbert space might be applicable. However, this further elaboration is beyond the scope of the present paper.

References

[1] Philippe Besse and J. O. Ramsay. Principal components analysis of sampled functions. *Psychometrika*, 51(2):285–311, 1986.

[2] Herman J. Bierens. A consistent conditional moment test of functional form. *Econometrica*, 58(6):1443–1458, 1990.

[3] Herman J. Bierens. *Topics in Advanced Econometrics*. Cambridge University Press, 1994.

[4] Herman J. Bierens and Werner Ploberger. Asymptotic theory of integrated conditional moment tests. *Econometrica*, 65(5):1129–1151, 1997.

[5] Nelson Dunford and Jacob T. Schwartz. *Linear Operators Part II Spectral Theory*. Wiley, 1988.

[6] B. E. Hansen. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 64(2):413–430, 1996.

[7] Kohtaro Hitomi. Common structure of consistent misspecification tests and a new test. *mimeo*, 2000.

[8] A. Pousse J. Dauxois and Y. Romain. Asymptotic theory for the principal component analysis of vector random function: Some applications to statistical inference. *Journal of Multivariate Analysis*, 12:136–154, 1982.

[9] James O. Ramsay and Bernard W. Silverman. *Functional Data Analysis*. Springer-Verlog, 1997.

[10] J. O. Ramsey and C J. Dalzell. Some tools for functional data analysis (with discussion). *Journal of Royal Statistical Society Series B*, 53(3):539–572, 1991.

[11] Alexander Weinstein and William Stenger. *Methods of Intermediate Problems for Eigenvalues*, volume 89 of *Mathematics in Science and Engineering*. Academic Press inc., 1972.

A Maintained Assumption for ICM test

The following is Assumption A in Bierens (1990).

Assumption A

1. Let $\{w_t = (y_t, x_t)|i = 1, 2, \ldots, n\}$ be a sequence of i.i.d. random variable on $R \times R^d$. Moreover, $E[y_t^2] < \infty$.

2. The parameter space Θ is a compact and convex subset of R^q and $Q(x_t, \theta)$ is for each $\theta \in \Theta$ a Borel measurable real function on R^d and for each d-vector x a twice continuously differentiable real function on Θ. Moreover, $E[\sup_{\theta\in\Theta} Q(x_t, \theta)] < \infty$ and for $i_1, i_2 = 1, 2, \ldots, q$,

$$
\begin{aligned}
E\left[\sup_{\theta\in\Theta} \frac{\partial Q(x_t,\theta)}{\partial\theta_{i_1}} \frac{\partial Q(x_t,\theta)}{\partial\theta_{i_1}}\right] &< \infty, \\
E\left[\sup_{\theta\in\Theta} (y_t - Q(x_t,\theta))^2 \frac{\partial Q(x_t,\theta)}{\partial\theta_{i_1}} \frac{\partial Q(x_t,\theta)}{\partial\theta_{i_2}}\right] &< \infty, \\
E\left[\sup_{\theta\in\Theta} (y_t - Q(x_t,\theta)) \frac{\partial^2 Q(x_t,\theta)}{\partial\theta_{i_1}\partial\theta_{i_2}}\right] &< \infty.
\end{aligned}
$$

3. $E[(y_t - Q(x_t, \theta))^2]$ takes a unique minimum on Θ at θ_0. Under H_0 the parameter vector θ_0 is an interior point of Θ.

4. $A = E[(\partial/\partial\theta)Q(x_t, \theta_0)(\partial/\partial\theta')Q(x_t, \theta_0)]$ is nonsingular.

B Mathematical Proofs

Lemma 8. *Suppose $A_n(\theta)$ and $B_n(\theta)$ be random functions on a compact subset $\Theta \subset R^k$, and $A_n(\theta)$ and $B_n(\theta)$ converges to nonrandom continuous functions $A(\theta)$ and $B(\theta)$ a.s. uniformly on Θ respectively. Then $A_n(\theta)B_n(\theta)$ converges to $A(\theta)B(\theta)$ a.s. uniformly on Θ.*

Proof. There are null set N_1 and N_2 such that for every $\varepsilon > 0$ and every $\omega \in \Omega \setminus (N_1 \cup N_2)$,

$$
\sup_{\theta\in\Theta} |A_n(\omega,\theta) - A(\theta)| \le \varepsilon \text{ and } \sup_{\theta\in\Theta} |B_n(\omega,\theta) - B(\theta)| \le \varepsilon \text{ if } n > n_0(\omega, \varepsilon).
$$

Since $A(\theta)$ and $B(\theta)$ are continuous function on a compact set Θ, there is M such that $\sup_{\theta\in\Theta} |A(\theta)| < M$ and $\sup_{\theta\in\Theta} |B(\theta)| < M$. And for every $\omega \in$

$\Omega \setminus (N_1 \cup N_2)$,

$$\begin{aligned}
& \sup_{\theta\in\Theta} |A_n(\theta)B_n(\theta) - A(\theta)B(\theta)| \\
\leq\ & \sup_{\theta\in\Theta} |A_n(\theta) - A(\theta)| \sup_{\theta\in\Theta} |B(\theta)| + \sup_{\theta\in\Theta} |B_n(\theta) - B(\theta)| \sup_{\theta\in\Theta} |A(\theta)| \\
& + \sup_{\theta\in\Theta} |A_n(\theta) - A(\theta)| \sup_{\theta\in\Theta} |B_n(\theta) - B(\theta)| \\
\leq\ & 2M\varepsilon + \varepsilon^2
\end{aligned}$$

if $n > n_0(\omega, \varepsilon)$. Thus $\sup_{\theta\in\Theta} |A_n(\theta)B_n(\theta) - A(\theta)B(\theta)| \to 0$ a.s. □

Theorem 9. (theorem 2.7.5 Bierens 1994) *Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variable in R^d. Let $f(x, \theta)$ be a Borel measurable function continuous on $R^d \times \Theta$, where Θ is a compact Borel set in R^k, which is continuous in θ for each $x \in R^d$. If $E[\sup_{\theta\in\Theta} |f(X_j, \theta)|] < \infty$, then $\frac{1}{n}\sum_{i=1}^n f(X_i, \theta) \to E[f(x, \theta)]$ a.s. uniformly on Θ.*

Proof of Proposition 6

Proof. Almost sure convergence of $\hat\theta$ and $\Gamma_n(\hat\theta, \xi_1, \xi_2) \to \Gamma_0(\xi_1, \xi_2)$ a.s. point wisely satisfied under Assumption A using standard argument. So we concentrate a.s. uniform convergence of $\Gamma_n(\theta, \xi_1, \xi_2)$.

We could decompose $\Gamma_n(\theta, \xi_1, \xi_2)$ as the followings,

$$\begin{aligned}
\Gamma_n(\theta, \xi_1, \xi_2) \quad = \quad & \frac{1}{n}\sum_{t=1}^n (y_t - Q(x_t, \theta)^2) \exp(\xi_1'\Phi(x_t)) \exp(\xi_2'\Phi(x_t)) \qquad : \Gamma_n^1 \\
& - b_n(\theta, \xi_1) A_n(\theta)^{-1} \frac{1}{n}\sum_{t=1}^n \frac{\partial Q(x_t, \theta)}{\partial \theta'} (y_t - Q(x_t, \theta) \exp(\xi_2'\Phi(x_t)) \qquad : b_n A_n^{-1} \Gamma_n^2 \\
& - b_n(\theta, \xi_2) A_n(\theta)^{-1} \frac{1}{n}\sum_{t=1}^n \frac{\partial Q(x_t, \theta)}{\partial \theta'} (y_t - Q(x_t, \theta) \exp(\xi_1'\Phi(x_t)) \qquad : b_n A_n^{-1} \Gamma_n^2 \\
& + b_n(\theta, \xi_1) A_n(\theta)^{-1} b_n(\theta, \xi_2).
\end{aligned}$$

Because of lemma 8, it is enough to show that

$$\Gamma_n^1(\theta, \xi_1, \xi_2) \quad \to \quad E[(y_t - Q(x_t, \theta))^2 \exp(\xi_1'\Phi(x_t)) \exp(\xi_2'\Phi(x_t))] \quad \text{a.s. uniformly} \tag{8}$$

$$\Gamma_n^1(\theta, \xi_1, \xi_2) \quad \to \quad E\left[\frac{\partial Q(x_t, \theta)}{\partial \theta'} (y_t - Q(x_t, \theta) \exp(\xi'\Phi(x_t))\right] \quad \text{a.s. uniformly} \tag{9}$$

$$b_n(\theta, \xi) \quad \to \quad E\left[\frac{\partial Q(x_t, \theta)}{\partial \theta} \exp(\xi'\Phi(x_t))\right] \text{a.s. uniformly} \tag{10}$$

$$A_n(\theta) \quad \to \quad E\left[\frac{\partial Q(x_t, \theta)}{\partial \theta} \frac{\partial Q(x_t, \theta)}{\partial \theta'}\right] \quad \text{a.s. uniformly.} \tag{11}$$

Since Ξ is compact and $\Phi(x)$ is a bounded function, there exists M such that for all x, $|\exp(\xi'\Phi(x))| < M$. Thus

$$\begin{aligned}
& E\left[\sup_{\theta,\xi_1,\xi_2} \left|(y_t - Q(x_t,\theta))^2 \exp(\xi_1'\Phi(x_t)) \exp(\xi_2'\Phi(x_t))\right|\right] \\
\leq \quad & M^2 E\left[\sup_{\theta} (y_t - Q(x_t,\theta))^2\right] \\
\leq \quad & 2M^2 E[y_t^2] + 2ME\left[\sup_{\theta} Q(x_t,\theta)^2\right] \\
< \quad & \infty,
\end{aligned}$$

by Assumption A.1 and A.2. Therefore (8) is satisfied by Theorem 9.

Similarly,

$$\begin{aligned}
& E\left[\sup_{\theta,\xi} \left|\frac{\partial Q(x_t,\theta)}{\partial \theta} (y_t - Q(x_t,\theta)) \exp(\xi'\Phi(x_t))\right|\right] \\
\leq \quad & ME\left[\sup_{\theta} \left|\frac{\partial Q(x_t,\theta)}{\partial \theta} (y_t - Q(x_t,\theta))\right|\right] \\
\leq \quad & M\left(E\left[\sup_{\theta} \left|\frac{\partial Q(x_t,\theta)}{\partial \theta}\right|^2\right]\right)^{1/2} \left(E\left[\sup_{\theta} \left|(y_t - Q(x_t,\theta))^2\right|\right]\right)^{1/2} \\
< \quad & \infty,
\end{aligned}$$

by Assumption A.1 and A.2. (9) is satisfied.

For $b_n(\theta,\xi)$,

$$\begin{aligned}
& E\left[\sup_{\theta,\xi} \left|\frac{\partial Q(x_t,\theta)}{\partial \theta} \exp(\xi'\Phi(x_t))\right|\right] \\
\leq \quad & ME\left[\sup_{\theta} \left|\frac{\partial Q(x_t,\theta)}{\partial \theta}\right|\right] \\
\leq \quad & ME\left[\sup_{\theta} \left|\frac{\partial Q(x_t,\theta)}{\partial \theta}\right|^2\right]^{1/2} E\left[\sup_{\theta} \left|\frac{\partial Q(x_t,\theta)}{\partial \theta}\right|^2\right]^{1/2} \\
< \quad & \infty,
\end{aligned}$$

by Assumption A.2. Therefore (10) is satisfied by Theorem 9. And (11) also satisfied by Assumption A.2 and Theorem 9.

Then (8)-(11) were all satisfied. Thus by lemma 8, $\Gamma_n(\theta,\xi_1,\xi_2) \to \Gamma(\theta,\xi_1,\xi_2)$ a.s. uniformly on $\Theta \times \Xi \times \Xi$. We have proved the proposition. $\square$

Proof of Theorem 7

Proof. The characteristic function of q is

$$\phi_q(t) = \prod_{j=1}^{\infty}(1 - 2i\lambda_j t)^{-1/2},$$

and the conditional characteristic function of q_n on Γ_n is

$$\phi_{q_n}(t) = \prod_{j=1}^{\infty}(1 - 2i\lambda_{nj} t)^{-1/2},$$

where $i = \sqrt{-1}$ and $\lambda_{nj} = 0$ if $j > n$. It is enough to show that

$$\log \phi_{q_n}(t) \to \log \phi_q(t) \quad a.s. \tag{12}$$

First, we prove that

$$-\frac{1}{2}\sum_{j=1}^{\infty}\{\log(1 - 2i\lambda_{nj}t) + 2i\lambda_{nt}t\} \to -\frac{1}{2}\sum_{j=1}^{\infty}\{\log(1 - 2i\lambda_j t) + 2i\lambda_t t\} \quad a.s. \tag{13}$$

as n increases. Let decompose the difference between the both side of the above equation as the following, take a large enough integer r then

$$\begin{aligned}
& -\frac{1}{2}\sum_{j=1}^{\infty}\{\log(1 - 2i\lambda_{nj}t) + 2i\lambda_{nt}t\} + \frac{1}{2}\sum_{j=1}^{\infty}\{\log(1 - 2i\lambda_j t) + 2i\lambda_t t\} \\
= \quad & -\frac{1}{2}\sum_{j=1}^{r}\{\log(1 - 2i\lambda_{nj}t) + 2i\lambda_{nt}t\} + \frac{1}{2}\sum_{j=1}^{r}\{\log(1 - 2i\lambda_j t) + 2i\lambda_t t\} \quad : S_r \\
& -\frac{1}{2}\sum_{j=r+1}^{\infty}\{\log(1 - 2i\lambda_{nj}t) + 2i\lambda_{nt}t\} + \frac{1}{2}\sum_{j=r+1}^{\infty}\{\log(1 - 2i\lambda_j t) + 2i\lambda_t t\} \quad : R_r \\
= \quad & S_r + R_r.
\end{aligned}$$

Since λ_{nj} converges to λ_j by Proposition 2, choosing large enough n we get

$$|S_r| < \varepsilon.$$

On the other hand there is a constant C_1 that satisfies

$$|\log(1 - 2i\lambda t) + 2i\lambda t| < C_1\lambda^2 t^2.$$

Therefore

$$
\begin{aligned}
|R_r| &\leq C_1 t^2 \sum_{j=r+1}^{\infty} \lambda_{nj}^2 + C_1 t^2 \sum_{j=r+1}^{\infty} \lambda_j^2 \\
&\leq C_1 t^2 \lambda_{n(r+1)} \sum_{j=1}^{\infty} \lambda_{nj} + C_1 t^2 \lambda_{r+1} \sum_{j=1}^{\infty} \lambda_j
\end{aligned}
$$

since λ_{nj} and λ_j are decreasing sequence. By Proposition 2, λ_{nj} converges to λ_j uniformly in j almost surely and $\sum_{j=1}^{\infty} \lambda_{nj}$ converges to $\sum_{j=1}^{\infty} \lambda_j$ almost surely by Proposition 2,

$$
\lim_{r\to\infty} |R_r| = 0 \quad a.s.
$$

Thus (13) is satisfied.

Next, we will show that

$$
-\frac{1}{2}\sum_{j=1}^{n} \log(1-2i\lambda_{nj}t) \to -\frac{1}{2}\sum_{j=1}^{\infty} \log(1-2i\lambda_j t) \quad a.s.
$$

The left hand side converges to

$$
\begin{aligned}
-\frac{1}{2}\sum_{j=1}^{n} \log(1-2i\lambda_{nj}t) &= -\frac{1}{2}\sum_{j=1}^{n} \{\log(1-2i\lambda_{nj}t) + 2i\lambda_{nj}t\} + \frac{1}{2}2it\sum_{j=1}^{n}\lambda_{nj} \\
&\to -\frac{1}{2}\sum_{j=1}^{\infty} \{\log(1-2i\lambda_j t) + 2i\lambda_j t\} + \frac{1}{2}2it\sum_{j=1}^{\infty}\lambda_j \quad a.s.
\end{aligned}
$$

by (13) and Proposition 3. On the other hand,

$$
\begin{aligned}
\lim_{n\to\infty} -\frac{1}{2}\sum_{j=1}^{n} \log(1-2i\lambda_j t) &= \lim_{n\to\infty}\left\{-\frac{1}{2}\sum_{j=1}^{n} \{\log(1-2i\lambda_j t) + 2i\lambda_j t\} + \frac{1}{2}2it\sum_{j=1}^{n}\lambda_j\right\} \\
&= \lim_{n\to\infty}\left\{-\frac{1}{2}\sum_{j=1}^{n} \{\log(1-2i\lambda_j t) + 2i\lambda_j t\}\right\} + \lim_{n\to\infty}\frac{1}{2}2it\sum_{j=1}^{n}\lambda_j
\end{aligned}
$$

since the both term in the last equation converge. This implies that

$$
-\frac{1}{2}\sum_{j=1}^{n} \log(1-2i\lambda_{nj}t) \to -\frac{1}{2}\sum_{j=1}^{\infty} \log(1-2i\lambda_j t) \quad a.s.
$$

We have proved. $\square$