Metrika (1998) 48: 83–97

Metrika

Asymptotic properties of the least squares estimators of a two dimensional model

Debasis Kundu[1,*], Rameshwar D. Gupta[2,]**

[1] Department of Mathematics, Indian Institute of Technology Kanpur, Pin 208016, India (e-mail: kundu@iitk.ac.in)
[2] Department of Mathematics, Statistics and Computer Science, The University of New Brunswick, Saint John, P.O. Box 5050, Canada, E2L 4L5 (e-mail: gupta@unbasj.ca)

Received March 1997

Abstract. We consider a particular two dimensional model, which has a wide applications in statistical signal processing and texture classifications. We prove the consistency of the least squares estimators of the model parameters and also obtain the asymptotic distribution of the least squares estimators. We observe the strong consistency of the least squares estimators when the errors are independent and identically distributed double array random variables. We show that the asymptotic distribution of the least squares estimators are multivariate normal. It is observed that the asymptotic dispersion matrix coincides with the Cramer-Rao lower bound. This paper generalizes some of the existing one dimensional results to the two dimensional case. Some numerical experiments are performed to see how the asymptotic results work for finite samples.

Key words: Strong consistency, texture classification, statistical signal processing

1. Introduction

We consider the following two dimensional model:

$$y(m,n) = \sum_{k=1}^{p} A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) + X(m,n);$$
$$\text{for } m = 1,\ldots,M, \quad n = 1,\ldots,N \tag{1}$$

---

* Part of the work has been supported by a grant of the Department of Science and Technology, Government of India, Grant No. $SR/OY/M$–06/93.
** Part of the work has been supported by the Natural Sciences and Engineering Research Council of Canada, Grant No. OGP-0004850.

where A_k^0's are unknown real numbers, λ_k^0's, μ_k^0's are unknown numbers, where $\lambda_k^0 \in (-\pi, \pi)$ and $\mu_k^0 \in (0, \pi)$. $X(m,n)$ is a two dimensional (2-D) sequence of independent and identically distributed random variables. 'p' is assumed to be a known integer. Given a sample $y(m,n)$; $m = 1, \ldots, M$, $n = 1, \ldots, N$, the problem is to estimate A_k^0's, λ_k^0's, μ_k^0's for $k = 1, \ldots p$ and the error variance σ^2.

This is also a well discussed model in Multidimensional Signal Processing, when $X(m,n)$'s are independently and identically distributed (i.i.d.) random variables on a 2-D plane. See for example the works of Barbieri and Barone (1992), Cabrera and Bose (1993), Chun and Bose (1995), Hua (1995), Lang and McClellan (1982) and see the references there for the different estimation procedures. It is interesting to observe that the model (1) is the 2-D extension of the one dimensional frequency model, which was originally discussed by Hannan (1971) and Walker (1971) in the time series analysis.

It is also observed that the model (1) can be used to model textures. To see how this model represents different textures the readers are referred to the work of Manderekar and Zhang (1995). They provided nice 2-D image plots of $y(m,n)$, whose grey level at (m,n) is proportional to the value of $y(m,n)$ and when it is corrupted by independent Gaussian noise field. Manderakar and Zhang (1995) considered this problem and obtained the consistency properties of the estimators of λ_k's and μ_k's when the estimators are obtained through the periodogram analysis and when $X(m,n)$'s are from a stationary random field. Their results based on the work of Lai and Wei (1982) and they did not provide the asymptotic distribution of the estimators.

But no where, at least not known to the authors, the properties of the least squares estimators have been discussed of the model (1). It is important to observe that the model (1) is a nonlinear regression model and unfortunately it does not satisfy the sufficient conditions stated by Jennrich (1969) or Wu (1981) for the least squares estimators (LSE) to be consistent. It may noted that when $p = 1$, $M = 1$ and $\lambda_k^0 = 0$, this model coincides with the one dimensional model discussed in Hannan (1971), Walker (1971), Kundu (1993) and Kundu and Mitra (1996). It was shown in Kundu (1993) that even the one dimensional model does not satisfy the sufficient conditions of Jennrich (1969) or Wu (1981). Therefore the consistency or the asymptotic normality of the LSE's is not immediate in this case. It is also worth mentioning at this stage that Bunke and Bunke (1989) also considered several nonlinear regression models in their book, but did not consider this particular model. One of the major difference of this particular model with the other usual nonlinear models is in the rate of convergence. Usually in the nonlinear model we observe the rate of convergence to be $\sqrt{n}$ but for this model, at least for the frequencies, the rate of convergence is $n^{3/2}$.

The main idea of this paper is to study the properties of the least squares estimators of the parameters of the model (1) and see how the asymptotic results behave for finite sample. We prove the strong consistency of the least squares estimators (LSE's) of A_k's, λ_k's and μ_k's when the errors are i.i.d. double array random variables. We obtain that the asymptotic distribution of the least squares estimators are multivariate normal. The explicit expression of the asymptotic dispersion matrix of the LSE's are obtained, which may be useful to obtain the confidence bounds. It is observed that the asymptotic dispersion matrix coincides with the Cramer Rao bound, which may not be very surprising. We prove that the LSE of σ^2 is strongly consistent when the error variance is finite and the asymptotic distribution of the LSE of σ^2 can be obtained when the fourth order moment of error random variables are finite. Our approach is different from that of Mandrekar and Zhang (1995). Our results extend some of the existing one dimensional results of Walker (1971), Hannan (1971), Kundu (1993), and Kundu and Mitra (1996) to the two dimensional case.

The rest of the papers is organized as follows. In Section 2, we prove the strong consistency of the LSE's of A_k's, λ_k's and μ_k's when the errors are i.i.d. random variables. In Section 3, the asymptotic normality results of those estimators are established under the same set of assumptions. The consistency and the asymptotic normality results of the estimator of σ^2 are obtained in section 4. A summary of numerical experiments is given in Section 5 and finally we draw conclusions from our work in Section 6.

2. Consistency of the LSE's of A_k, λ_k and μ_k

We need the following lemma to prove the necessary results. We denote the set of positive integers by Z.

Lemma 1. *Let $\{X(m,n);\ m,n \in Z\}$ be a i.i.d. sequence of double array random variables with mean zero and finite variance then;*

$$\sup_{\alpha,\beta}\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N}X(m,n)\cos(m\alpha)\cos(n\beta)\right| \to 0 \quad a.s. \quad when \min\{M,N\} \to \infty \tag{2}$$

where a.s. means almost surely.

Proof of Lemma 1: Consider the following random variables;

$$\begin{aligned} Z(m,n) &= X(m,n) \quad \text{if } |X(m,n)| < (mn)^{3/4} \\ &= 0 \qquad\qquad \text{otherwise} \end{aligned}$$

First we will show that $Z(m,n)$ and $X(m,n)$ are equivalent sequences. Consider

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}P\{X(m,n) \neq Z(m,n)\} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}P\{|X(m,n)| > (mn)^{3/4}\}$$

Now observe that there are at most $2^k k$ combinations of (m,n)'s such that $mn < 2^k$, therefore we have

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} P\{|X(m,n)| \geq (mn)^{3/4}\}$$
$$\leq \sum_{k=1}^{\infty}\sum_{2^{k-1}\leq r<2^k} P\{|X(m,n)| \geq r^{3/4}\} \quad [\text{here } r = mn]$$
$$\leq \sum_{k=1}^{\infty} k2^k P\{|X(1,1)| \geq 2^{(k-1)3/4}\}$$
$$\leq C\sum_{k=1}^{\infty} k2^k \frac{E|X(1,1)|^2}{2^{(k-1)3/2}} \leq C\sum_{k=1}^{\infty} \frac{k}{2^{k/2}} < \infty \tag{3}$$

Here C is a constant and note that it may represent different constant at different places. Therefore, $X(m,n)$ and $Z(m,n)$ are equivalent sequences. So

$$P\{X(m,n) \neq Z(m,n) \quad i.o.\} = 0 \tag{4}$$

Here i.o. means infinitely often. Let $U(m,n) = Z(m,n) - E(Z(m,n))$, then

$$\sup_{\alpha,\beta}\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N} E(Z(m,n))\cos(m\alpha)\cos(n\beta)\right| \leq \frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N}|E(Z(m,n))|$$

Since $E(Z(m,n)) \to 0$ as $M,N \to \infty$, therefore as $M,N \to \infty$

$$\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N}|E(Z(m,n))| \to 0 \tag{5}$$

Therefore, it is enough to prove that

$$\sup_{\alpha,\beta}\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n)\cos(m\alpha)\cos(n\beta)\right| \to 0 \quad \text{a.s.} \tag{6}$$

Now for any fixed $\varepsilon > 0$, $-\pi < \alpha, \beta < \pi$ and $0 < h \leq \frac{1}{2(MN)^{3/4}}$, we have

$$P\left\{\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n)\cos(m\alpha)\cos(n\beta)\right| \geq \varepsilon\right\}$$
$$\leq 2e^{-hMN\varepsilon}\prod_{m=1}^{M}\prod_{n=1}^{N} Ee^{hU(m,n)\cos(m\alpha)\cos(n\beta)} \tag{7}$$

Since $|hU(m,n)\cos(m\alpha)\cos(n\beta)| \leq 1/2$, using $e^x < 1 + x + x^2$ for $|x| < 1/2$, we have

$$2e^{-hMN\varepsilon}\prod_{m=1}^{M}\prod_{n=1}^{N} Ee^{hU(m,n)\cos(m\alpha)\cos(n\beta)} \leq 2e^{-hMN\varepsilon}(1 + h^2\sigma^2)^{MN}. \tag{8}$$

Now choose $h = \frac{1}{2(MN)^{3/4}}$, therefore for large M and N

$$P\left\{\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n)\cos(m\alpha)\cos(n\beta)\right| \geq \varepsilon\right\}$$
$$\leq Ce^{-(MN)^{1/4}\varepsilon/2}e^{\sigma^2/4} \quad (C \text{ is a constant}).$$

Let $K = M^2N^2$, choose K points, $\theta_1 = (\alpha_1, \beta_1), \ldots, \theta_K = (\alpha_K, \beta_K)$, such that for each point $\theta = (\alpha, \beta) \in (-\pi, \pi)$, we have a point θ_j satisfying

$$|\alpha_j - \alpha| + |\beta_j - \beta| \leq \frac{2\pi}{M^2N^2} \tag{9}$$

Note that

$$\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n)\{\cos(m\alpha)\cos(n\beta) - \cos(m\alpha_j)\cos(n\beta_j)\}\right|$$
$$\leq C\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\frac{MN}{M^2N^2}[m+n] \to 0 \quad \text{as } M, N \to \infty.$$

Therefore for large M and N, we have

$$P\left\{\sup_{\alpha,\beta}\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n)\cos(m\alpha)\cos(n\beta)\right| \geq 2\varepsilon\right\}$$
$$\leq P\left\{\max_{j \leq M^2N^2}\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n)\cos(m\alpha_j)\cos(n\beta_j)\right| \geq \varepsilon\right\}$$
$$\leq CM^2N^2e^{-(MN)^{1/4}/2}. \tag{10}$$

Since $\sum_{t=1}^{\infty} t^2 e^{-t^{1/4}} < \infty$,

$$\sup_{\alpha,\beta}\left|\frac{1}{N}\frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n)\cos(m\alpha)\cos(n\beta)\right| \to 0 \quad \text{a.s.} \tag{11}$$

Where (11) follows from the Borel Cantelli lemma.

Consider the following assumptions:

Assumption 1: Let $A_1^0, \ldots A_p^0$ be arbitrary real numbers not anyone of them are identically equal to zero, $\lambda_1^0, \ldots \lambda_p^0 \in (-\pi, \pi)$ and they are distinct, similarly $\mu_1^0, \ldots \mu_p^0 \in (0, \pi)$ and they are distinct.

Assumption 2: Let $\{X(m,n); m,n \in Z\}$ be i.i.d. sequence of double array random variables and $E(X(m,n)) = 0$, $E(X(m,n)^2) = \sigma^2$.

Now we would like to prove the following results:

Theorem 1: *Under the assumptions 1 and 2, the least squares estimators of the parameters of the model (1), i.e., which are obtained by minimizing*

$$\sum_{m=1}^{M}\sum_{n=1}^{N}\left(y(m,n) - \sum_{k=1}^{p} A_k \cos(m\lambda_k + n\mu_k)\right)^2 \tag{12}$$

with respect to the unknown parameters, are strongly consistent estimators of the corresponding parameters.

Proof of Theorem 1: We take $p = 2$ for notational convenience. Let's use the following notations: $\mathbf{A} = (A_1, A_2)$, $\lambda = (\lambda_1, \lambda_2)$ and $\mu = (\mu_1, \mu_2)$. Consider the following expression;

$$\begin{aligned}
&R_{M,N}(\mathbf{A}, \lambda, \mu) \\
&\quad = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(y(m,n) - \sum_{k=1}^{2} A_k \cos(m\lambda_k + n\mu_k)\right)^2 \\
&\quad = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(\sum_{k=1}^{2} A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) - \sum_{k=1}^{2} A_k \cos(m\lambda_k + n\mu_k)\right)^2 \\
&\qquad + \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n)^2 + 2\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n) \\
&\qquad \times \left(\sum_{k=1}^{2} A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) - \sum_{k=1}^{2} A_k \cos(m\lambda_k + n\mu_k)\right).
\end{aligned}$$

Consider the following set;

$$\begin{aligned}
S_{\delta,T} = \{(A_1, A_2, \lambda_1, \lambda_2, \mu_1, \mu_2);\ & |A_1 - A_1^0| \ge \delta,\ |A_1| \le T,\ |A_2| \le T \quad \text{or} \\
& |A_2 - A_2^0| \ge \delta,\ |A_1| \le T,\ |A_2| \le T \quad \text{or} \\
& |\lambda_1 - \lambda_1^0| \ge \delta,\ |A_1| \le T,\ |A_2| \le T \quad \text{or} \\
& |\lambda_2 - \lambda_2^0| \ge \delta,\ |A_1| \le T,\ |A_2| \le T \quad \text{or} \\
& |\mu_1 - \mu_1^0| \ge \delta,\ |A_1| \le T,\ |A_2| \le T \quad \text{or} \\
& |\mu_2 - \mu_2^0| \ge \delta,\ |A_1| \le T,\ |A_2| \le T\}
\end{aligned}$$

First we will prove the following;

$$\underline{\lim} \inf_{(A,\lambda,\mu)\in S_{\delta,T}} R_{MN}(\mathbf{A},\lambda,\mu) > \sigma^2 \quad \text{for all } \delta > 0 \tag{13}$$

here $\underline{\lim}$ denotes $\liminf$. Now suppose $(\hat{A}_{MN}, \hat{\lambda}_{MN}, \hat{\mu}_{MN})$ be the least squares estimators of (A^0, λ^0, μ^0) and they are not consistent, therefore either;

Case 1: For all subsequences (M_K, N_K) of (M, N), $|\hat{A}_{M_K N_K}| \to \infty$ or

Case 2: There exists a $\delta > 0$ and a $T < \infty$, and a subsequence (M_K, N_K) of (M, N) such that;

$$(\hat{A}_{M_K N_K}, \hat{\lambda}_{M_K N_K}, \hat{\mu}_{M_K N_K}) \in S_{\delta,T} \tag{14}$$

for all $K = 1, 2, \ldots$. Now

$$R_{M_K,N_K}(\hat{A}_{M_K N_K}, \hat{\lambda}_{M_K N_K}, \hat{\mu}_{M_K N_K}) \le R_{M_K,N_K}(A^0, \lambda^0, \mu^0) \tag{15}$$

as $(\hat{A}_{M_K N_K}, \hat{\lambda}_{M_K N_K}, \hat{\mu}_{M_K N_K})$ is the least squares estimator of (A^0, λ^0, μ^0), when $M = M_K$ and $N = N_K$. As $K \to \infty$, the left hand side of (15) converges to a number which is strictly greater than σ^2 due to (13), whereas the right hand side of (15) converges to σ^2. That gives a contradiction. Hence the consistency of the least squares estimators can be established once we prove (13).

To prove (13), consider the following sets:

$$A_{1\delta} = \{(A_1, A_2, \mu_1, \mu_2, \lambda_1, \lambda_2);\ |A_1| \le T,\ |A_2| \le T,\ |A_1 - A_1^0| \ge \delta\}$$
$$A_{2\delta} = \{(A_1, A_2, \mu_1, \mu_2, \lambda_1, \lambda_2);\ |A_1| \le T,\ |A_2| \le T,\ |A_2 - A_2^0| \ge \delta\}$$
$$M_{1\delta} = \{(A_1, A_2, \mu_1, \mu_2, \lambda_1, \lambda_2);\ |A_1| \le T,\ |A_2| \le T,\ |\mu_1 - \mu_1^0| \ge \delta\}$$
$$M_{2\delta} = \{(A_1, A_2, \mu_1, \mu_2, \lambda_1, \lambda_2);\ |A_1| \le T,\ |A_2| \le T,\ |\mu_2 - \mu_2^0| \ge \delta\}$$
$$\Lambda_{1\delta} = \{(A_1, A_2, \mu_1, \mu_2, \lambda_1, \lambda_2);\ |A_1| \le T,\ |A_2| \le T,\ |\lambda_1 - \lambda_1^0| \ge \delta\}$$
$$\Lambda_{2\delta} = \{(A_1, A_2, \mu_1, \mu_2, \lambda_1, \lambda_2);\ |A_1| \le T,\ |A_2| \le T,\ |\lambda_2 - \lambda_2^0| \ge \delta\}$$

Since we have

$$S_{\delta T} = A_{1\delta} \cup A_{2\delta} \cup M_{1\delta} \cup M_{2\delta} \cup \Lambda_{1\delta} \cup \Lambda_{2\delta} \tag{16}$$

therefore (13) can be established if we can prove that

$$\underline{\lim} \inf_{(A,\lambda,\mu)\in V} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(\sum_{k=1}^{2} A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) - \sum_{k=1}^{2} A_k \cos(m\lambda_k + n\mu_k) \right)^2$$
$$> 0 \quad \text{for } \delta > 0 \tag{17}$$

where V is any one of the sets $A_{1\delta}, A_{2\delta}, M_{1\delta}, M_{2\delta}, \Lambda_{1\delta}, \Lambda_{1\delta}$ or $\Lambda_{2\delta}$. We will prove (17) when $V = A_{1\delta}$. The other cases can be proved similarly. Observe that

$$\underline{\lim}\ \inf_{A_{1\delta}} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(\sum_{k=1}^{2} A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) - \sum_{k=1}^{2} A_k \cos(m\lambda_k + n\mu_k) \right)^2$$

$$= (A_1^0 - A_1)^2 \lim_{M,N\to\infty} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \cos^2(m\lambda_k^0 + n\mu_k^0)^2$$

$$\geq \frac{1}{2}\delta^2 > 0 \quad \text{a.s.} \tag{18}$$

Hence (13) has been established, so Theorem 1 is proved.

3. Asymptotic normality of $\hat{A}$, $\hat{\lambda}$, $\hat{\mu}$

In this section we obtain the asymptotic distribution of the least squares estimators of the parameters of the model (1).

Theorem 2: *Under the assumptions 1 and 2, the limiting distribution of* $\{M^{1/2}N^{1/2}(\hat{A}-A^0),\ M^{3/2}N^{1/2}(\hat{\lambda}-\lambda^0),\ M^{1/2}N^{3/2}(\hat{\mu}-\mu^0)\}$ *as* $\mathrm{Min}(M,N) \to \infty$, *is a* $3p$ *variate normal with mean vector zero and covariance matrix* $2\sigma^2\Sigma^{-1}$, *when* Σ^{-1} *has the following structure:*

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{11}^{-1} & 0 & 0 \\ 0 & \Sigma_{22}^{-1} & \Sigma_{23}^{-1} \\ 0 & \Sigma_{32}^{-1} & \Sigma_{33}^{-1} \end{bmatrix} \tag{19}$$

where $\Sigma_{11}^{-1} = \mathbf{I_p}$, $\Sigma_{22}^{-1} = \Sigma_{33}^{-1} = \frac{48}{7}\, diag\{A_1^{0^{-2}}, \ldots, A_p^{0^{-2}}\}$, *and* $\Sigma_{23}^{-1} = \Sigma_{32}^{-1} = -\frac{36}{7}\, diag\{A_1^{0^{-2}}, \ldots, A_p^{0^{-2}}\}$,

Proof of Theorem 2: Let's denote $\theta = (A, \lambda, \mu)$, $\hat{\theta} = (\hat{A}, \hat{\lambda}, \hat{\mu})$ and

$$Q(\theta) = \sum_{m=1}^{M} \sum_{n=1}^{N} \left[y(m,n) - \sum_{k=1}^{p} A_k \cos(m\lambda_k + n\mu_k) \right]^2. \tag{20}$$

Then

$$Q'(\hat{\theta}) - Q'(\theta^0) = (\hat{\theta} - \theta^0) Q''(\bar{\theta}), \tag{21}$$

here $Q'(\hat{\theta})$ is a $1 \times 3p$ vector defined as follows:

$$Q'(\hat{\theta}) = \left[\frac{\delta Q(\theta)}{\delta A}, \frac{\delta Q(\theta)}{\delta \lambda}, \frac{\delta Q(\theta)}{\delta \mu} \right] \tag{22}$$

and $Q''(\theta)$ is a $3p \times 3p$, matrix as follows;

$$Q''(\theta) = \begin{bmatrix} \frac{\delta Q(\theta)}{\delta A \delta A^T} & \frac{\delta Q(\theta)}{\delta A \delta \lambda^T} & \frac{\delta Q(\theta)}{\delta A \delta \mu^T} \\ \frac{\delta Q(\theta)}{\delta \lambda \delta A^T} & \frac{\delta Q(\theta)}{\delta \lambda \delta \lambda^T} & \frac{\delta Q(\theta)}{\delta \lambda \delta \mu^T} \\ \frac{\delta Q(\theta)}{\delta \mu \delta A^T} & \frac{\delta Q(\theta)}{\delta \mu \delta \lambda^T} & \frac{\delta Q(\theta)}{\delta \mu \delta \mu^T} \end{bmatrix}. \tag{23}$$

$\hat{\theta}$ denotes the LSE of θ and $\bar{\theta}$ is a point lying between the line joining the point $\hat{\theta}$ and θ^0. Consider the following $3p \times 3p$ diagonal matrix $\mathbf{D}$ as follows;

$$\mathbf{D} = \begin{bmatrix} M^{-1/2}N^{-1/2}\mathbf{I}_p & 0 & 0 \\ 0 & M^{-3/2}N^{-1/2}\mathbf{I}_p & 0 \\ 0 & 0 & M^{-1/2}N^{-3/2}\mathbf{I}_p \end{bmatrix}. \tag{24}$$

Since $\hat{\theta}$ is the LSE of θ, therefore $Q'(\hat{\theta}) = 0$, and hence (21) can be written as;

$$(\hat{\theta} - \theta^0) = -Q'(\theta^0)[Q''(\bar{\theta})]^{-1} \tag{25}$$

and

$$(\hat{\theta} - \theta^0)\mathbf{D}^{-1} = -[Q'(\theta^0)\mathbf{D}][\mathbf{D}Q''(\bar{\theta})\mathbf{D}]^{-1}. \tag{26}$$

First we will compute $Q'(\theta^0)\mathbf{D}$. Observe that for $i = 1, \ldots, p$,

$$\begin{aligned} \frac{\delta Q(\theta)}{\delta A_i} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N}\left(y(m,n) - \sum_{k=1}^{p} A_k \cos(m\lambda_k + n\mu_k)\right)\cos(m\lambda_i + n\mu_i), \\ \frac{\delta Q(\theta)}{\delta \lambda_i} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N}\left(y(m,n) - \sum_{k=1}^{p} A_k \cos(m\lambda_k + n\mu_k)\right)A_i m \sin(m\lambda_i + n\mu_i), \\ \frac{\delta Q(\theta)}{\delta \mu_i} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N}\left(y(m,n) - \sum_{k=1}^{p} A_k \cos(m\lambda_k + n\mu_k)\right)A_i n \sin(m\lambda_i + n\mu_i). \end{aligned} \tag{27}$$

Therefore, we have

$$\begin{aligned} \frac{\delta Q(\theta^0)}{\delta A_i} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n)\cos(m\lambda_i^0 + n\mu_i^0), \\ \frac{\delta Q(\theta^0)}{\delta \lambda_i} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} A_i^0 m X(m,n)\sin(m\lambda_i^0 + n\mu_i^0), \\ \frac{\delta Q(\theta^0)}{\delta \mu_i} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} A_i^0 n X(m,n)\sin(m\lambda_i^0 + n\mu_i^0). \end{aligned} \tag{28}$$

Observe that (28) satisfy the Lindeberg-Feller condition (see Chung; 1978). Therefore, $Q'(\theta^0)$ with proper normalization will be asymptotically normal. We now look at the asymptotic covariance matrix of the vector $Q'(\theta^0)$. We need the following results for $\beta \neq 0$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} \sin^2(\beta t) = \frac{1}{2}, \quad \text{and} \quad \lim_{N\to\infty} \frac{1}{N^2}\sum_{t=1}^{N} t\sin^2(\beta t) = \frac{1}{6}. \tag{29}$$

Similar results hold for cosine function also (see Walker; 1971 or Mangulis; 1965). So we have;

$$M^{-1}N^{-1}Cov\left(\frac{\delta Q(\theta^0)}{\delta A_i}, \frac{\delta Q(\theta^0)}{\delta A_j}\right) \to \begin{Bmatrix} 0 & \text{if } i \neq j \\ 2\sigma^2 & \text{if } i = j \end{Bmatrix} \tag{30}$$

$$M^{-3}N^{-1}Cov\left(\frac{\delta Q(\theta^0)}{\delta \lambda_i}, \frac{\delta Q(\theta^0)}{\delta \lambda_j}\right) \to \begin{Bmatrix} 0 & \text{if } i \neq j \\ \frac{2}{3}\sigma^2 A_i^{0^2} & \text{if } i = j \end{Bmatrix} \tag{31}$$

$$M^{-1}N^{-3}Cov\left(\frac{\delta Q(\theta^0)}{\delta \mu_i}, \frac{\delta Q(\theta^0)}{\delta \mu_j}\right) \to \begin{Bmatrix} 0 & \text{if } i \neq j \\ \frac{2}{3}\sigma^2 A_i^{0^2} & \text{if } i = j \end{Bmatrix} \tag{32}$$

$$M^{-2}N^{-2}Cov\left(\frac{\delta Q(\theta^0)}{\delta \lambda_i}, \frac{\delta Q(\theta^0)}{\delta \mu_j}\right) \to \begin{Bmatrix} 0 & \text{if } i \neq j \\ \frac{1}{2}\sigma^2 A_i^{0^2} & \text{if } i = j \end{Bmatrix} \tag{33}$$

$Q'(\theta^0)\mathbf{D}$ converges to a $3p$ Multivariate Normal with mean vector $\mathbf{0}$ and the dispersion matrix $2\sigma^2\Sigma$, where

$$\Sigma = \begin{bmatrix} \mathbf{I}_p & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_1 & \mathbf{D}_2 \\ 0 & \mathbf{D}_2 & \mathbf{D}_1 \end{bmatrix}, \tag{34}$$

here both $\mathbf{D_1}$ and $\mathbf{D_2}$ are $p \times p$ matrices and they are as follows;

$$\mathbf{D}_1 = diag\{\tfrac{1}{3}A_1^{0^2}, \ldots, \tfrac{1}{3}A_p^{0^2}\},$$

$$\mathbf{D}_2 = diag\{\tfrac{1}{4}A_1^{0^2}, \ldots, \tfrac{1}{4}A_p^{0^2}\}.$$

Observe that $\bar{\theta}$ converges to θ^0 a.s., and

$$\lim_{M,N\to\infty}(\mathbf{D}Q''(\bar{\theta})\mathbf{D}) = \lim_{M,N\to\infty}(\mathbf{D}Q''(\theta^0)\mathbf{D}) = \Sigma \tag{35}$$

Therefore

$$(\hat{\theta} - \theta^0)\mathbf{D^{-1}} \to N_{3p}(\mathbf{0}, 2\sigma^2\Sigma^{-1}) \tag{36}$$

where Σ^{-1} is as defined before.

4. Consistency and asymptotic normality of $\hat{\sigma}^2$

In Sections 2 and 3 we considered the LSE's of A_k's, λ_k's, μ_k's and obtained the consistency and asymptotic normality properties. In this section we consider the estimation procedures of σ^2 and obtain its asymptotic properties. If $\hat{A}_k$'s, $\hat{\lambda}_k$'s and $\hat{\mu}$'s are the LSE's of A_k, λ_k, μ_k respectively, then the LSE of σ^2 will be

$$\hat{\sigma}^2 = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(y(m,n) - \sum_{k=1}^{p} \hat{A}_k \cos(m\hat{\lambda}_k + n\hat{\mu}_k) \right)^2. \tag{37}$$

First we will show that $\hat{\sigma}^2$ is a consistent estimator of σ^2. To prove that we need the following lemmas;

Lemma 2: *Let $\{X(m,n); m,n \in Z\}$ be a i.i.d. sequence of double array random variables and $E(X(m,n)) = 0$, $E(X(m,n)^2) = \sigma^2$, then*

$$\sup_{\alpha,\beta} \frac{1}{M^2N} \sum_{m=1}^{M} \sum_{n=1}^{N} mX(m,n) \cos(m\alpha) \cos(n\beta) \to 0 \quad a.s.$$

$$\sup_{\alpha,\beta} \frac{1}{MN^2} \sum_{m=1}^{M} \sum_{n=1}^{N} nX(m,n) \cos(m\alpha) \cos(n\beta) \to 0 \quad a.s.$$

Proof of Lemma 2: It follows similarly as that of Lemma 1.

Lemma 3: *If $\hat{\lambda}_k$ and $\hat{\mu}_k$ are the LSE's of λ_k and μ_k respectively of the model (1), then*

$$M(\hat{\lambda}_k - \lambda_k^0) \to 0 \quad a.s.$$
$$N(\hat{\mu}_k - \mu_k^0) \to 0 \quad a.s.$$

as $\{M,N\} \to \infty$.

Proof of Lemma 3: Observe that from (21) we have

$$(\hat{\theta} - \theta^0) = -Q'(\theta^0)[Q''(\bar{\theta})]^{-1}. \tag{38}$$

Now suppose $\mathbf{R}$ is a $3p \times 3p$ diagonal matrix as given below;

$$\mathbf{R} = \begin{bmatrix} \mathbf{I}_p & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{M}\mathbf{I}_p & \mathbf{0} \\ 0 & \mathbf{0} & \frac{1}{N}\mathbf{I}_p \end{bmatrix}. \tag{39}$$

From (38),

$$(\hat{\theta}-\theta_0)\mathbf{R}^{-1} = Q'(\theta_0)\mathbf{R}[\mathbf{R}Q''(\bar{\theta})\mathbf{R}]^{-1}$$

$$= \left[\frac{1}{MN}Q'(\theta_0)\mathbf{R}\right]\left[\frac{1}{\sqrt{MN}}\mathbf{R}Q''(\bar{\theta})\frac{1}{\sqrt{MN}}\mathbf{R}\right]^{-1}. \quad (40)$$

Observe that $\frac{1}{\sqrt{MN}}\mathbf{R} = \mathbf{D}$ (as defined in Section 3). Therefore

$$\left[\frac{1}{\sqrt{MN}}\mathbf{R}Q''(\bar{\theta})\frac{1}{\sqrt{MN}}\mathbf{R}\right]^{-1} \to \Sigma^{-1} \quad (41)$$

and because of Lemma 2, we have

$$\frac{1}{MN}Q'(\theta_0)\mathbf{R} \to 0 \quad \text{a.s.} \quad (42)$$

which proves the result.

Theorem 3: *Under the assumptions 1 and 2 of the model (1), $\hat{\sigma}^2$ is a strongly consistent estimator of σ^2.*

Proof of Theorem 3: Note that

$$\hat{\sigma}^2 = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(y(m,n) - \sum_{k=1}^{p}\hat{A}_k\cos(m\hat{\lambda}_k + n\hat{\mu}_k)\right)^2$$

$$= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}X(m,n)^2 + 2\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}X(m,n)$$

$$\times\left[\sum_{k=1}^{p}A_k^0\cos(m\lambda_k^0 + n\mu_k^0) - \sum_{k=1}^{p}\hat{A}_k^0\cos(m\hat{\lambda}_k + n\hat{\mu}_k)\right]$$

$$+\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[\sum_{k=1}^{p}A_k^0\cos(m\lambda_k^0 + n\mu_k^0) - \sum_{k=1}^{p}\hat{A}_k^0\cos(m\hat{\lambda}_k + n\hat{\mu}_k)\right]^2$$

$$= T_1 + T_2 + T_3 \quad \text{(say)}. \quad (43)$$

Observe that T_2 converges to zero almost surely because of Lemma 1 and T_3 converges to zero almost surely because of Lemma 3. Since T_1 converges to σ^2 almost surely because of Strong Law of Large Numbers, it proves the theorem.

Now to prove the asymptotic normality of $\hat{\sigma}^2$, the second order moment conditions of $X(m,n)$'s are not sufficient. We need a stronger assumption than the second order moment conditions. We need the following assumption:

Assumption 3: Let $\{X(m,n);\ m,n \in Z\}$ be a i.i.d. sequence of double array random variables and $E(X(m,n)) = 0$, $E(X(m,n)^2) = \sigma^2$ and $E(X(m,n)^4) = a$ (say).

The asymptotic normality results of $\hat{\sigma}^2$ is established by the following result:

Theorem 4: *Under the Assumptions 1 and 3 of the model (1), $\sqrt{MN}(\hat{\sigma}^2 - \sigma^2)$ is asymptotically normal with mean 0 and variance $(a - \sigma^4)$.*

To prove Theorem 4, we need the following lemmas.

Lemma 4: *Let $\{X(m,n); m,n \in Z\}$ be a i.i.d. sequence of double array random variables satisfying Assumption 2 and $\hat{\lambda}_k$ and $\hat{\mu}_k$ are the LSE's of λ_k and μ_k respectively of the model (1). Then*

$$\frac{1}{\sqrt{MN}}\sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n)[\cos(m\lambda_k^0 + n\mu_k^0) - \cos(m\hat{\lambda}_k + n\hat{\mu}_k)] \rightarrow 0 \quad \textit{in Prob.} \tag{44}$$

Proof of Lemma 4: Using Lemma 2 and Theorem 2, the result is immediate.

Lemma 5: *Let $\{X(m,n); m,n \in Z\}$ be a i.i.d. sequence of double array random variables satisfying Assumption 2 and $\hat{A}_k$, $\hat{\lambda}_k$ and $\hat{\mu}_k$ are the LSE's of A_k, λ_k and μ_k respectively of the model (1) then*

$$\frac{1}{\sqrt{MN}}\sum_{m=1}^{M}\sum_{n=1}^{N}[\cos(m\lambda_k^0 + n\mu_k^0) - \cos(m\hat{\lambda}_k + n\hat{\mu}_k)]^2 \rightarrow 0 \quad \textit{in Prob.} \tag{45}$$

Proof of Lemma 5: Using Lemma 3 and Theorem 2, the result can be proved.

Lemma 6: *Under the same Assumptions 1 and 2 of the model (1), $\hat{\sigma}^2$ can be represented as;*

$$\hat{\sigma}^2 = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n)^2 + o_p(M^{-1/2}N^{-1/2}) \tag{46}$$

where $Z = o_p((MN)^{-1/2})$ means that $\sqrt{MN}Z$ converges to zero in probability.

Proof of Lemma 6: From the expression (43), using Lemma 4 and Lemma 5, the result can be obtained.

Proof of Theorem 4: From Lemma 6, Theorem 4 follows using the Central Limit Theorem.

5. Numerical experiments and discussions

In this section we give summary of some numerical experiments performed to see how the asymptotic results behave for finite sample sizes. We performed all the experiments in PC-486, using the random deviate generator of Press et al. (1993). We considered the following model:

$$y(m,n) = 4.0\cos(2.0m + 1.0n) + 5.0\cos(2.5m + 1.5n) + X(m,n), \qquad (47)$$

$\{X(m,n);\ m = 1,\ldots,M,\ n = 1,\ldots,N\}$ are i.i.d. Gaussian random variables with mean zero and finite variance σ^2. We considered $M = N = 10, 20, 30, 40, 50$ and $\sigma = .25, .50, .75, 1.0$. For each sample size and for each σ we computed the LSE's of A_1, A_2, λ_1, λ_2, μ_1 and μ_2 and observed the average estimates and the average mean squared errors (MSE's) over five hundred replications. However we do not provide the results due to space, but provide the central findings of our simulation works below.

From the simulations it became very clear that as sample size increases or the variance decreases, the average MSE's and biases of all the estimators decrease. It shows the consistency properties of all the estimators. It is clear that the MSE's of the estimators of the non-linear parameters are smaller than that of the linear parameters even for small sample sizes. Some of the asymptotic behaviors are present even at small sample sizes. For example if $A_1 < A_2$, then it is observed that the MSE's of $\hat{\mu}_2$ and $\hat{\lambda}_2$ are smaller than that of $\hat{\mu}_1$ and $\hat{\lambda}_1$ respectively. It is also observed that as sample size increases the MSE's become closer to the asymptotic variances. Therefore looking at the behavior of the MSE's we can say that the asymptotic results can be used to draw the small sample inferences for the different model parameters.

6. Conclusions

In this paper we consider the estimation of the parameters of a two dimensional model, which has wide applicability in Statistical Signal Processing and Texture classifications. We study the asymptotic properties of the LSE's of the model parameters and show that the LSE's estimators are strongly consistent. We also obtain the asymptotic distribution of the LSE's, which may be useful to obtain the confidence region or testing of hypothesis problem. This paper generalizes some of the existing one dimensional results to the 2-D case. Numerical experiments suggest that the asymptotic results can be used to draw the small sample inferences for the linear and non-linear parameters.

Acknowledgments: The authors thank the referees for corrections and suggestions for improvements. The editorial comments are also appreciated which led to better presentation of the paper.

References

[1] Barbieri MM, Barone P (1992) A two dimensional Prony's method for Spectral estimation. IEEE Trans. Signal Processing 40, 11:2747–2756

[2] Bunke H, Bunke O (1989) Nonlinear regression, functional relations and robust methods. Vol. 2, Wiley and Sons, Berlin

[3] Cabrera SD, Bose NK (1993) Prony's methods for two dimensional complex exponential signals modelling. In: Tzafestas SG (ed.) Applied control, Chap. 15, Dekker, New York, pp. 401–411
[4] Chun J, Bose NK (1995) Parameter estimation via signal selectivity of signal subspaces (PESS) and its applications to 2-D wave number estimation. Digital Signal Processing 5:58–76
[5] Chung KL (1974) A course in probability theory. Academic Press, New York
[6] Hannan EJ (1971) Nonlinear time series regression. Jour. Applied Probability 8:767–780
[7] Hua Y (1992) Estimating two dimensional frequencies by matrix enhancement and matrix Pencil. IEEE Trans. Signal Processing 40, 9:2267–2280
[8] Jennrich RI (1969) Asymptotic properties of the non-linear least squares estimation. Annals of Mathematical Statistics 40:633–643
[9] Kundu D (1993) Asymptotic theory of the least squares estimator of a particular non-linear regression model. Statistics and Probability Letters 18:13–17
[10] Kundu D, Mitra A (1996) Asymptotic theory of the least squares estimator of a non-linear regression model. Communications in Statistics, Theory and Methods 25, 1:133–141
[11] Lai TL, Wai ZC (1982) A law of iterated logarithm for double array of independent random variables with applications to regression and time series models. Ann Probability 10, 2:320–335
[12] Lang SW, McClellan JH (1982) The extension of Pisarenko's method to multiple dimensions. Proc. ICASSP-82:125–128
[13] Mandrekar V, Zhang H (1995) Determination of discrete spectrum of random fields with application to texture classification. Technical Report, Michigan State University, USA
[14] Mangulis V (1965) Handbook of series for scientists and engineers. Academic Press, New York
[15] Press et al. (1993) Numerical recipes in Fortran, Cambridge University Press, Cambridge
[16] Walker AM (1971) On the estimation of a harmonic component in a time series with stationary residuals. Biometrika 58, 1:21–36
[17] Wu CFJ (1981) Asymptotic theory of non-linear least squares estimation. Annals of Statistics 9:510–513